

**Міжнародний науково-технічний  
журнал**

**ВИМІРЮВАЛЬНА ТА  
ОБЧИСЛЮВАЛЬНА ТЕХНІКА  
В ТЕХНОЛОГІЧНИХ  
ПРОЦЕСАХ**

---

**2024, № 3**

---

**International scientific-technical  
journal**

**MEASURING AND COMPUTING  
DEVICES IN TECHNOLOGICAL  
PROCESSES**

---

**2024, Issue 3**

**Хмельницький 2024  
Khmelnyskyi 2024**

МІЖНАРОДНИЙ НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ  
ВИМІРЮВАЛЬНА ТА ОБЧИСЛЮВАЛЬНА ТЕХНІКА В ТЕХНОЛОГІЧНИХ ПРОЦЕСАХ

Затверджений як фахове видання (перереєстрація), група «Б»  
Наказ МОН 28.12.2019 №1643

*Засновано в травні 1997 р.*

*Виходить 4 рази на рік*

**Хмельницький, 2024, № 3 (79)**

**Засновник і видавець:** Хмельницький національний університет  
(до 2005 р. — Технологічний університет Поділля, м. Хмельницький)

Наукова бібліотека України ім. В.І. Вернадського <http://nbuv.gov.ua/j-tit/vott>

Журнал включено до наукометричних баз:

Index Copernicus <http://jml2012.indexcopernicus.com/p24781565.3.html>  
Google Scholar [http://scholar.google.com.ua/citations?user=nwN\\_nusAAAAJ&hl=uk](http://scholar.google.com.ua/citations?user=nwN_nusAAAAJ&hl=uk)  
CrossRef <http://doi.org/10.31891/2219-9365>

**Головний редактор** Мартинюк В. В., д. т. н., професор, завідувач кафедри автоматизації, комп'ютерно-інтегрованих технологій і телекомунікацій Хмельницького національного університету

**Заступник головного редактора** Бойко Ю. М., д. т. н., професор кафедри телекомунікацій та радіотехніки, начальник науково-дослідної частини Хмельницького національного університету

**Відповідальний секретар** Кравчик Ю. В., к. е. н., старший викладач кафедри економіки, менеджменту та адміністрування Хмельницького національного університету

**Ч л е н и р е д к о л е г і ї**

Бармак О. В., д.т.н., Бедратюк Л. П., д.фіз.-мат.н., Бубулис Алгимантас, д.т.н. (Литва), Васілевський О. М., д.т.н., Горященко К. Л., к.т.н., Здоренко В. Г., д.т.н., Калачинський Томаш, PhD (Польща), Косенков В. Д., к.т.н., Кулаков П. І., д.т.н., Кухарчук В. В., д.т.н., Кучерук В. Ю., д.т.н., Лампасі Алессандро, PhD, (Італія), Лукасевич Марцін, PhD, (Польща), Мрозинський Адам, PhD, (Польща), Мусяль Януш, PhD, (Польща), Ортігейра Мануель Дуарте, PhD, (Португалія), Походило Є. В., д.т.н., Психалінос Костас, PhD, (Греція), Савенко О. С., д.т.н., Семенко А. І., д.т.н., Сурду М. М., д.т.н., Шарпан О. Б., д.т.н.

*Технічний редактор* Кравчик Ю. В., к. е. н., доцент.

Рекомендовано до друку рішенням Вченої ради Хмельницького національного університету,  
протокол № 1 від 29.08.2024

Адреса редакції: Україна, 29016,  
м. Хмельницький, вул. Інститутська, 11,  
Хмельницький національний університет,  
Редакція журналу "Вимірвальна та обчислювальна техніка в технологічних процесах"  
☎ 067-347-74-57  
e-mail: [vottp@khmnu.edu.ua](mailto:vottp@khmnu.edu.ua)  
web: <http://vottp.khmnu.edu.ua>

Зареєстровано Міністерством України у справах преси та інформації.  
Свідоцтво про державну реєстрацію друкованого засобу масової інформації  
Серія КВ № 24923-14863 ПР від 12 липня 2021 року (перереєстрація)

© Хмельницький національний університет, 2024  
© Редакція журналу «Вимірвальна та обчислювальна техніка в технологічних процесах», 2024

## ЗМІСТ

<b>РУДНЄВ МИКИТА</b> АНАЛІЗ КОМПЛЕКСІВ БЕЗПЛОТНИХ ЛІТАЛЬНИХ ТА НАЗЕМНИХ МОБІЛЬНИХ СИСТЕМ ДЛЯ ЗАДАЧ МОНІТОРИНГУ ЕКОЛОГІЧНО-КРИТИЧНИХ ОБ'ЄКТІВ .....	7
<b>ВОЛОКИТА АРТЕМ, МОРОЗОВ-ЛЕОНОВ ОЛЕКСАНДР</b> ПІДХОДИ ТА МЕТОДИ ПРИЙНЯТТЯ РІШЕНЬ В СЕРЕДОВИЩАХ ІЗ НЕПОВНОЮ ТА НЕВИЗНАЧЕНОЮ ІНФОРМАЦІЄЮ .....	19
<b>ЗАЩЕПКИНА НАТАЛІЯ, МЕЛЬНИЧЕНКО ДМИТРО</b> АВТОНОМНА СПЕКТРОФОТОМЕТРИЧНА СИСТЕМА ДЛЯ АНАЛІЗУ НІТРАТІВ І НІТРИТІВ У ВОДІ .....	29
<b>ДАЩЕНКО ВОЛОДИМИР, ВОЛКОВИЙ ПАВЛО</b> ВИМІРЮВАННЯ ПОВЕРХНЕВОЇ ЩІЛЬНОСТІ МАТЕРІАЛУ З ОДНОРІДНОЮ СТРУКТУРОЮ УЛЬТРАЗВУКОВИМ МЕТОДОМ .....	34
<b>КРУПА ДМИТРО</b> ОЦІНКА СУМІСНОСТІ ОКРЕМИХ КОМПОНЕНТ ВЕБ-СЕРВЕРІВ З ВИКОРИСТАННЯМ ОНТОЛОГІЧНОГО ПІДХОДУ ТА НЕЧІТКОГО ЛОГІЧНОГО ВИВЕДЕННЯ .....	40
<b>ПАНЧИШИН БОГДАН</b> АНАЛІЗ ПРИНЦИПІВ РОБОТИ КОМПОНЕНТІВ ФРЕЙМВОРКА REACT JS .....	46
<b>ТИМОЩУК ДМИТРО, ЯЦКІВ ВАСИЛЬ</b> ВИКОРИСТАННЯ ГІПЕРВІЗОРІВ ДЛЯ СТВОРЕННЯ КІБЕРПОЛІГОНУ .....	52
<b>БЕРНАТОВИЧ АНАТОЛІЙ</b> АНАЛІЗ ВИМОГ ДО РОЗРОБКИ УНІВЕРСАЛЬНОГО ФІЗИЧНОГО РУШЯ ІЗ ПІДТРИМКОЮ ПЛАГІНІВ, КОНФІГУРАЦІЄЮ ПАЙПЛАЙНУ ТА АПАРАТНИМ ПРИСКОРЕННЯМ .....	57
<b>МЕЩЕРЯКОВ ОЛЕКСАНДР</b> ОГЛЯД МЕТОДІВ ТА ЗАСОБІВ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МОДЕЛЕЙ КЛАСИФІКАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ .....	64
<b>ТІТОВА ВІРА, КЛЬОЦ ЮРІЙ, ВОЛИНЕЦЬ ВІТАЛІЙ, ПЕТЛЯК НАТАЛІЯ, ОГОРОДНИК МАКСИМ</b> РОЗРОБЛЕННЯ ПОЛІТИКИ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ПРИВАТНОГО ПІДПРИЄМСТВА .....	79
<b>НАГІРНИЙ ВАСИЛЬ, КУЦЬ ВІКТОР</b> РОЗВИТОК ЗАСОБІВ ДІАГНОСТИКИ СЕНСОРНИХ МЕРЕЖ ДВОКОНТУРНИХ СИСТЕМ ОПАЛЕННЯ .....	84
<b>ПИРИГ ЯРОСЛАВ, ПИРИГ ЮЛІЯ</b> БАГАТОКРИТЕРІАЛЬНИЙ ПІДХІД НА ОСНОВІ ГЕНЕТИЧНОЇ ЕВОЛЮЦІЇ ДЛЯ ПОШУКУ ОПТИМАЛЬНОГО МАРШРУТУ ПЕРЕДАЧІ ДАНИХ У БЕЗПРОВІДНИХ СЕНСОРНИХ МЕРЕЖАХ .....	88
<b>ЛАКТІОНОВ ОЛЕКСАНДР, ПЕДЧЕНКО НАЗАР, ЯНКО АЛІНА, БОРЯК БОГДАН</b> МОДЕЛЮВАННЯ БАЗОВОЇ КОНСТРУКЦІЇ РОБОТОТЕХНІЧНОЇ ПЛАТФОРМИ .....	95
<b>БОЖКО КОСТЯНТИН, МУШКЕТ КОСТЯНТИН</b> АКТИВНА АНТЕНА У ВИМІРЮВАННЯХ ЕЛЕКТРИЧНИХ ПАРАМЕТРІВ СОНЯЧНИХ ЕЛЕМЕНТІВ .....	100

---

<b>RUDNIEV MYKYTA</b> ANALYSIS OF UNMANNED AERIAL AND GROUND MOBILE SYSTEMS FOR MONITORING ENVIRONMENTALLY CRITICAL OBJECTS .....	7
<b>VOLOKYTA ARTEM, MOROZOV-LEONOV OLEKSANDR</b> APPROACHES AND METHODS OF DECISION-MAKING IN ENVIRONMENTS WITH INCOMPLETE AND UNCERTAIN INFORMATION .....	19
<b>ZASHCHEPKINA NATALIYA, MELNYCHENKO DMYTRO</b> AUTONOMOUS SPECTROPHOTOMETRIC SYSTEM FOR NITRATE AND NITRITE ANALYSIS IN WATER .....	29
<b>DASHCHENKO VOLODYMYR, VOLKOVYI PAVLO</b> MEASUREMENT OF THE AREA DENSITY OF A MATERIAL WITH A UNIFORM STRUCTURE USING THE ULTRASONIC METHOD .....	34
<b>KRUPA DMYTRO</b> EVALUATION OF THE COMPATIBILITY OF INDIVIDUAL COMPONENTS OF WEB SERVERS USING THE ONTOLOGICAL APPROACH AND FUZZY LOGICAL INFERENCE .....	40
<b>PANCHYSHYN BOHDAN</b> ANALYSIS OF THE WORKING PRINCIPLES OF REACT JS FRAMEWORK COMPONENTS .....	46
<b>TYMOSHCHUK DMYTRO, YATSKIV VASYL</b> USING HYPERVISORS TO CREATE A CYBER POLYGON .....	52
<b>BERNATOVYCH ANATOLII</b> ANALYSIS OF REQUIREMENTS FOR THE DEVELOPMENT OF A UNIVERSAL PHYSICS ENGINE WITH PLUG-IN SUPPORT, PIPELINE CONFIGURATION, AND HARDWARE ACCELERATION .....	57
<b>MESHCHERIAKOV OLEKSANDR</b> REVIEW OF METHODS AND TOOLS FOR IMPROVING THE EFFICIENCY OF MEDICAL IMAGE CLASSIFICATION MODELS .....	64
<b>TITOVA VIRA, KLOTS YURIY, VOLYNETS VITALII, PETLIAK NATALIYA, OHORODNYK MAKSYM</b> DEVELOPING THE INFORMATION SECURITY POLICY OF A PRIVATE ENTERPRISE .....	79
<b>NAHIRNYI VASYL, KUTS VIKTOR</b> DEVELOPMENT OF DIAGNOSTIC TOOLS OF SENSOR NETWORKS OF DOUBLE-CIRCUIT HEATING SYSTEMS .....	84
<b>PYRIH YAROSLAV, PYRIH YULIIA</b> A MULTI-CRITERIA APPROACH BASED ON GENETIC EVOLUTION FOR FINDING THE OPTIMAL DATA TRANSMISSION ROUTE IN WIRELESS SENSOR NETWORKS .....	88
<b>LAKTIONOV OLEKSANDR, PEDCHENKO NAZAR, YANKO ALINA, BORIAK BOHDAN</b> MODELING THE BASIC STRUCTURE OF A ROBOTIC PLATFORM .....	95
<b>BOGKO KONSTANTIN, MUSHKET KOSTIANTYN</b> ACTIVE ANTENNA IN MEASUREMENTS OF ELECTRICAL PARAMETERS OF SOLAR ELEMENTS ....	100





<https://doi.org/10.31891/2219-9365-2024-79-1>

УДК

РУДНІЄВ Микита

Національний аерокосмічний університет ім. М. С. Жуковського «Харківський авіаційний інститут»

e-mail: [m.rudniev@student.csn.khai.edu](mailto:m.rudniev@student.csn.khai.edu)

## АНАЛІЗ КОМПЛЕКСІВ БЕЗПЛОТНИХ ЛІТАЛЬНИХ ТА НАЗЕМНИХ МОБІЛЬНИХ СИСТЕМ ДЛЯ ЗАДАЧ МОНІТОРИНГУ ЕКОЛОГІЧНОКРИТИЧНИХ ОБ'ЄКТІВ

У статті досліджено особливості сучасних безпілотних літальних та наземних мобільних систем (БЛМНС) і їх застосування для моніторингу екологічно критичних об'єктів. Розглянуто класифікацію безпілотних апаратів, методи їх зв'язку та архітектуру Internet of Drones (IoD). Визначено переваги та виклики використання БЛМНС, обґрунтовано необхідність інтеграції різних технологій для підвищення ефективності моніторингу. Запропоновано оптимальні конфігурації безпілотних систем для виконання моніторингу та реагування на надзвичайні ситуації.

Ключові слова: безпілотні літальні апарати, безпілотні наземні апарати, рої дронів, Internet of Drones, системи моніторингу, екологічна безпека.

RUDNIEV Mykyta

National Aerospace University "Kharkiv Aviation Institute", NAU "KhAI"

## ANALYSIS OF UNMANNED AERIAL AND GROUND MOBILE SYSTEMS FOR MONITORING ENVIRONMENTALLY CRITICAL OBJECTS

The article examines the features of modern unmanned aerial and ground mobile systems (UAGMS) and their application for monitoring environmentally critical objects. The classification of unmanned vehicles, communication methods, and the architecture of the Internet of Drones (IoD) are considered. The advantages and challenges of using UAGMS are identified, and the necessity of integrating various technologies to enhance monitoring efficiency is substantiated. Optimal configurations of unmanned systems are proposed for monitoring and responding to emergency situations.

Keywords: unmanned aerial vehicles, unmanned ground vehicles, drone swarms, Internet of Drones, monitoring systems, environmental safety.

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Згідно з дослідженнями ДСНС 2011-2020-го року [1] Україна є найбільш критичним регіоном Європи з техногенного навантаження та в п'ять разів перевищує середньоєвропейський рівень. Ризики виникнення НС техногенного характеру є досить високим, що насамперед пов'язано з високим рівнем техногенного навантаження на регіони, наявністю комплексу енергетичних, хімічних, гірничодобувних об'єктів, значною кількістю промислово-міських агломерацій і високою щільністю населення у промислово розвинутих регіонах країни.

Агресія росії на території України призвела до ще більшого погіршення екологічного стану. На сьогодні більшість екологічних ризиків в Україні пов'язані з пошкодженням промислових об'єктів, будинків, енергетичної інфраструктури та екосистем, спричинених конфліктом, що триває [2]. Станом лише на кінець 2022-го року було пошкоджено або зруйновано понад 400 підприємств, включаючи хімічні заводи, площі лісових пожеж та інших насаджень сягнули 75 тис. га. Крім того, було розбомблено приблизно 40 складів нафти та палива, що спричинило викид 680 000 тонн нафтопродуктів. Війна створила серйозну загрозу для ядерної та радіаційної безпеки, поставивши під небезпеку 15 діючих ядерних реакторів, що може мати транскордонні наслідки.

Високі рівні ризику виникнення НС потребують якісного та своєчасного моніторингу екологічно-критичних об'єктів. В Україні діють відповідні контрольно-спостережні служби, що виконують завдання екологічного моніторингу, але сучасне завдання стоїть не лише у виконанні моніторингу, але й у підвищенні його наукового і технологічного рівня, ступеня репрезентивності та інформативності [3, 4].

Для виконання цього завдання на допомогу приходять безпілотні літальні та наземні апарати. Синергія використання безпілотних літальних апаратів (БПЛА) та безпілотних наземних апаратів (БПНА) відкриває нові можливості для моніторингу та забезпечення екологічної безпеки [5,6,7]. БПЛА можуть забезпечувати швидке та ефективне обстеження великих територій, зокрема, завдяки можливості отримання аерофотознімків та збирання даних з різних сенсорів. З іншого боку, БПНА можуть виконувати завдання на місцевості, такі як збір проб, вимірювання радіаційного фону чи аналіз стану ґрунтів. Комбінація цих двох типів безпілотників дозволяє забезпечити комплексний підхід до моніторингу, що включає збір, аналіз та інтеграцію даних з різних джерел, підвищуючи ефективність та точність оцінки екологічної ситуації. Такий

підхід дозволяє швидко реагувати на НС, знижуючи ризики для довкілля та населення. Тому, завданням цієї роботи буде аналіз можливих рішень для моніторингу екологічно-критичних об'єктів.

### АНАЛІЗ ОСТАННІХ РІШЕНЬ І ПУБЛІКАЦІЙ

БЛНМС можна описати як з'єднання двох або більше робототехнічних систем для виконання місії як одна команда. Цю концепцію можна порівняти з біологічним симбіозом, який визначається як «будь-який тип тісної і довготривалої біологічної взаємодії між двома різними біологічними організмами, будь то мутуалістичний, коменсалістський або паразитичний» [8]. Хоча ролі кожного робота в системі різняться за складністю і важливістю, основна мета використання БЛНМС полягає в тому, щоб використовувати спеціалізацію багатьох робототехнічних конструкцій, замість того, щоб створювати одного універсального робота, який би виконував усі задачі. Таким чином, кожен робот призначений для виконання конкретної місії, яка є частиною більшої мети. Термін БЛНМС охоплює не лише конструкцію і функції окремих роботів-агентів, але й правила взаємодії, зв'язку і співпраці між роботами, а також одночасну роботу цих численних спеціалізованих роботів з метою створення цілісної ефективної системи.

Для початку варто ознайомитися з фундаментальною класифікацією різних роботів та індивідуальними функціональними можливостями кожної класифікації. Для цього було виконано класифікацію та синтез інформації з різних наукових статей.

#### Класифікація конструкцій БПНА

За особливостями конструкції БПНА можна поділити на:

1. Колісні апарати: є одним з найпоширеніших типів наземних роботів, що застосовуються для дослідження навколишнього середовища, незалежно від того, чи це добре облаштоване середовище, чи несприятливе середовище [9]. Зазвичай, такі апарати призначені для виконання спеціалізованої місії, що базується на результатах роботи його модуля розвідки і картографування. Колісні конструкції мають широкий спектр конфігурацій, від широко використовуваних чотириколісних конструкцій до складних багатоколісних роверів, призначених для пересування нерівними ландшафтами [10]. Розмір, конструкція і тип коліс можуть змінюватися залежно від робочого простору і спеціалізованих завдань, покладених на апарат. Колісні роботи використовуються в багатьох наукових галузях, промислових та військових цілях [11]. Міжнародні виробники та дистриб'ютори виготовляють різні комерційні апарати, які доступні на ринку. Ці апарати можуть бути оснащені широким спектром датчиків, механічних маніпуляторів і пристроїв управління, а також ШІ для полегшення процесів майже у всіх галузях промисловості та сферах застосування.

2. Гусеничні апарати є надійною та популярною альтернативою колісним [12]. Розподіляючи вагу апарату на більшу площу поверхні, створюється більше зчеплення з ґрунтом. Це дозволяє цим типам транспортних засобів працювати на нерівній місцевості з м'яким ґрунтом. Хоча вони повільніші за швидкістю, переваги підвищеного зчеплення і стійкості поєднуються зі стабільним і простим механізмом приводу, який дозволяє гусеничним транспортним засобам обертатися навколо себе. Більшість гусеничних платформ складається з двох гусеничних систем з обох боків апарату, кожна з яких має різноспрямоване керування [13].

3. Роботи на ногах: ножні конструкції можуть використовуватися для ходьби, бігу або навіть стрибків [14,15,16,17,18]. Конструкції на основі ніг можуть містити будь-яку кількість ніг на основі тварин, які імітуються, наприклад, двоногі, чотириногі, шестиногі і т.д. Різні конструкції можуть допомогти для пересування на різних типах поверхонь, залежно від кількості ніг та їхнього розташування.

4. Стрибаючі роботи: намагаються імітувати стрибки багатьох видів тварин, таких як жаби та кенгуру. Такий спосіб пересування може бути корисним для робота, коли він пересувається місцевістю з перешкодами, які неможливо подолати за допомогою традиційного колісного методу пересування. Однією з труднощів стрибкового способу пересування є проектування робота, який залишався б ідеально збалансованим під час кожного наступного стрибка, а для визначення центру тяжіння робота необхідно провести обширні випробування. Багато стрибучих роботів, таких як робот-кенгуру, використовують рухомі противаги для імітації хвоста. Це утримує робота від надмірного або недостатнього обертання під час кожного стрибка і налаштовує його на наступний стрибок [19].

#### Класифікація конструкцій БПЛА

Зростаючий інтерес до БПЛА в останні роки призвів до активної появи різних типів літальних апаратів з різною конфігурацією та компонентами, що відрізняються за формою і розмірами. БПЛА поділяються на чотири типи: однороторні, багатороторні, літальні та гібридні.

1. Однороторні (або вертольоти): ця категорія БПЛА злітає і сідає вертикально. Як правило, вони використовують основний гвинт для керування положенням (крен, тангаж і ривання) і хвостовий гвинт для керування напрямком. Основною перевагою є здатність перевозити важкі вантажі протягом тривалого часу польоту. Найбільшою складністю при розробці дрона з одним ротором є керування горизонтальним напрямком польоту. Використання одного ротора дозволяє дуже легко і ефективно здійснювати



вертикальний політ, але при використанні лише одного ротора важко досягти бокового повороту в певному напрямку [20].

2. Багатороторні [21] (або мультикоптери): це БПЛА, що мають більше двох роторів. Ця категорія БПЛА поділяється на п'ять підкатегорій: біроторні, трироторні, квадатороторні, гексатороторні та восьмироторні (октокоптери). Як і однороторні БПЛА, багатороторні БПЛА також забезпечують вертикальний зліт і посадку. Вони швидкі та маневрені в польоті, що дозволяє їм виконувати складні маневри та польоти в обмеженому просторі. Однак короткий час польоту є основним недоліком цих типів літальних апаратів.

3. Фіксоване крило [22]: принцип польоту цієї категорії БПЛА ґрунтується на простій конструкції фіксованого жорсткого крила. Класифікація цих дронів базується не лише на типі крила, але й на корпусі та системі живлення (літій-іонні, літій-полімерні акумулятори або газові). Вони поділяються на чотири підкатегорії: звичайні, стрілоподібні, зворотні стрілоподібні і дельтаплани. Крім того, вони можуть нести важче корисне навантаження, ніж багатороторні. Недоліком цих БПЛА є обмежена маневреність у польоті, що не дозволяє їм виконувати складні маневри і літати над обмеженим простором, а також необхідність злітно-посадкової смуги для зльоту і посадки.

4. Гібридні [23]: це вдосконалена версія, яка використовує переваги як багатороторних, так і фіксованих БПЛА. Вони пропонують хорошу маневреність і швидкість на великих відстанях. Вони можуть нести велике корисне навантаження і не потребують злітно-посадкової смуги. Основними недоліками є висока ціна, складна механіка, а також нижчі показники стабільності польоту і обмежений діапазон швидкостей.

### Методи зв'язку БА

У сфері безпілотних систем, що постійно розвивається, ефективний зв'язок між БПЛА та БПНА має першочергове значення. Така синергія забезпечує безперебійну координацію і розширює можливості автономних систем у різних сферах застосування - від моніторингу довкілля до реагування на катастрофи. У цьому розділі розглядаються можливі методи зв'язку для БА, які забезпечують стабільний, надійний і ефективний обмін даними.

1. Радіочастотний зв'язок [24]: мільярди пристроїв по всьому світу спілкуються за допомогою бездротових радіочастотних хвиль. Передавачі та приймачі пристроїв зв'язуються між собою на певній радіочастоті. Цей метод зв'язку часто використовується БА.

Переваги	Недоліки
Специфічні радіочастотні сигнатури означають, що навіть у жвавому міському середовищі сигнальні піки безпілотників досить чіткі, що робить цю технологію надійною. Технологія перевірена, масштабована і в цілому безпечна.	Сигнали можуть бути навмисно або ненавмисно перехоплені за допомогою засобів глушіння сигналів. Сигнал має обмежений радіус дії і повинен працювати в межах прямої видимості.

2. 4G зв'язок з використанням однієї або декількох SIM-карт: ліцензовані та регульовані стільникові мережі пропонують модулі та мережеві послуги 4G. Вони забезпечують широке покриття та рівномірну якість.

Переваги	Недоліки
З'єднання 4G дозволяє передавати дані швидше і в більших обсягах, ніж попередні покоління мережевих послуг. Ці послуги пропонують хороший радіус дії на великих відстанях, за умови, що територія покрита достатньою кількістю веж стільникового зв'язку.	Віддалені території можуть бути недостатньо або зовсім не покриті стільниковим зв'язком, використання БА з цим методом зв'язку обмежене або неможливе.

### 3. Супутниковий зв'язок

Переваги	Недоліки
Супутникові технології забезпечують глобальний постійний зв'язок. Покриття дуже широке, а час безвідмовної роботи стабільно високий.	Наразі технологія супутникового зв'язку в основному використовується у великих БПЛА військового типу, які літають на великі відстані і висоти. Комерційні дрони поки що не мають життєздатного рішення, яке б включало супутниковий зв'язок.

4. 5G зв'язок: технологія, що забезпечує швидкість передачі даних у сотні разів вищу, ніж її попередник 4G.

Переваги	Недоліки
Ідеально підходить для завдань з високими ресурсними запитами, таких як передача відео високої чіткості в реальному часі. Він має кращі показники затримки та швидкості передачі даних, ніж попередні покоління.	Будь-яка залежність від однієї лінії зв'язку 5G є небезпечною, особливо при низькому покритті. Це означає, що операції в сільській місцевості будуть під загрозою. Щоб дрони могли ефективно використовувати мережу 5G, покриття має бути всеосяжним, що на даний момент далеко від реальності.

### Компоненти архітектури IoD

Коли виконання завдання потребує створення системи з роїв БА, симбіотично взаємодіючих для виконання спільної мети, на допомогу приходить Internet of Drones (IoD). IoD - це об'єднана мережа дронів, яка використовує існуючу інфраструктуру Інтернету речей (IoT) для полегшення виконання місій за допомогою передачі даних і навігаційних послуг в режимі реального часу.

IoD може відрізнитися в залежності від архітектури, проміжного ПЗ, методів інтеграції та спільного використання даних та безпеки. Основою архітектури IoD є два основні елементи: компоненти архітектури та протоколи зв'язку. Архітектура IoD відіграє життєво важливу роль у контролі та адмініструванні БА для більш ефективного виконання їх роботи [24].

1. Архітектурні елементи (компоненти): архітектура IoD містить низку компонентів IoT такі як сенсори, хмарні платформи, зв'язок тощо. Роль та операції цих компонентів залежать від їхньої взаємодії з архітектурою IoD. Стабільність, збір даних і методи комунікації в IoD визначаються компонентами його архітектури [25]. Компоненти архітектури IoD призначені для виконання операцій, головним чином пов'язаних із прийняттям рішень та управлінням дронами, а також забезпеченням того, щоб дані досягали правильного місця призначення від вузлів-джерел [26].

2. Протоколи зв'язку: IoD дозволяє використовувати різні протоколів зв'язку для підтримки та передачі даних між вузлами. Наприклад, одними із застосовуваних протоколів зв'язку є протоколи MAVLink і ROSLink. MAVLink - це полегшений протокол маршрутизації повідомлень, а ROSLink - протокол інтегрований з операційною системою робота (Robot Operating System), яка забезпечує підключення роботів до IoT [27]. Ефективність протоколу зв'язку обумовлюється переміщенням носіїв та їх активністю в мережі [28]. Протокол зв'язку для IoD повинен бути швидкодіючим в питаннях маршрутизації, де передача даних між вузлами-джерелами та вузлами-цільми є критично важливою [29].

У середовищі підключених апаратів IoD рівень проміжного програмного забезпечення відіграє важливу роль, функціонуючи як посередник між різними вузлами та додатками [30]. Рівень проміжного ПЗ забезпечує взаємодію між різними інтерфейсами IoD, а саме мовою програмування, операційною системою, мережами та архітектурою [31]. Існує розглядаються два підходи до проміжного програмних забезпечення - сервісне та хмарне.

1. Проміжне ПЗ на основі сервісів: полегшує архітектурі IoD доступ до мережі, локальну доставку повідомлень, кешування та вирішення імен. Сервісне проміжне ПЗ може забезпечити надійний зв'язок для всієї архітектури IoD [32]. Крім того, воно може добре інтегруватися з іншими мережевими рівнями для забезпечення ефективної співпраці, продуктивності та адаптивності до архітектури. Проміжне ПЗ на основі сервісів може допомогти забезпечити ефективну передачу даних і забезпечити тривалий час з'єднання між мережевими комунікаціями [33].

2. Хмарне проміжне ПЗ: дозволяє різноманітним додаткам інтегрувати свої операції в мережеву архітектуру. Найбільш поширеним проміжним ПЗ в робототехніці є ROS (Robot Operating System), яка вивантажує обчислення і обробку в хмару для ефективного використання ресурсів. Цей процес отримав назву хмарної робототехніки [34]. Хмарне проміжне ПЗ забезпечує ефективний зв'язок між наземною мережею та БА. Хмарне проміжне ПЗ забезпечує дуже швидко реакцію на запитуваний сервіс. Крім того, це дає змогу IoD активно взаємодіяти з іншими мережами.

Інтеграція та спільне використання даних дозволяє обробляти і об'єднувати кілька джерел даних для отримання правильної інформації для прийняття рішень. Крім того, алгоритми на основі злиття даних для роїв додатково підтримують більш ефективну роботу в інтелектуальному середовищі. У цьому підрозділі буде розглянуто три різні типи об'єднання і спільного використання даних для IoD, а саме: розподілений, централізований і хмарний.

1. Розподілене використання даних: децентралізує дані в локальну взаємодію, і жодні компоненти не є важливими для будь-якої іншої операції, що дає кілька переваг для роботи IoD, а саме: масштабованість, відмовостійкість, сумісність і легкість перепроєктування [35]. Ці переваги допомагають

переналаштувати або видалити конкретний недоступний або відключений дрон з мережі. Крім того, розподілене середовище зручніше для самостійного прийняття рішень, а в разі надзвичайної ситуації може працювати в режимі спільного використання.

2. Централізоване використання даних: потребує потужного комунікаційного обладнання для безперервного виконання операцій. В такому механізмі вся інформація розподіляється через центр злиття з іншими пристроями в мережі, що забезпечує надання більш точної інформації про операцію [36]. На ранній стадії розгортання і роботи на коротких відстанях централізований обмін даними є більш ефективним. Однак, коли мова йде про більші відстані і пізніші етапи операцій, передислокувати пристрій або керувати ним буде набагато складніше [37].

3. Хмарне використання даних: повністю контролюється хмарними інтерфейсами, які об'єднують різні сервіси, щоб допомогти клієнту прийняти ефективне рішення. Крім того, в архітектурі можуть бути реалізовані такі послуги, як аналітика та інтелектуальне управління. Розроблені фреймворки для інтеграції різних джерел для управління дорожнім рухом в додатках розумних міст використовують хмарне злиття даних. Крім того, ці фреймворки пропонують послуги з безпеки, уникнення зіткнень та навігації з урахуванням ризиків [38].

На основі проведеного аналізу БПЛА та БПНА можна зробити висновок про те, що БА можуть виконувати широкий спектр задач включаючи моніторинг екологічно-критичних об'єктів. Користуючись досягненнями в області інженерії БА можна обрати конструкцію БА відповідно до середовища та специфіки виконуваного завдання. Розвиток сенсорів, машинного навчання та алгоритмів розпізнавання надає можливості ефективного моніторингу та ідентифікації об'єктів. В свою чергу, технологія ІoD дозволяє об'єднувати різні БА в рої дронів, створюючи корисні симбіози для виконання спільних місій.

Відповідно рішень проаналізованих в цьому пункті в наступному пункті будуть запропоновані можливі варіанти БА та БЛМНС, які б могли виконувати задачі моніторингу екологічно-критичної інфраструктури.

#### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Для того, щоб обрати оптимальні варіанти побудови БЛМНС необхідно визначити конкретну задачу для якої буде використовуватися система, та провести детальний аналіз вимог [39, 40]. Це включає:

1. Оцінку середовища експлуатації: Визначення умов, у яких буде працювати система, таких як тип місцевості, кліматичні умови, можливі перешкоди та ризики.
2. Технічні вимоги: Визначення необхідних технічних характеристик БА.
3. Вимоги до комунікацій: Визначення методів передачі даних між БА.
4. Інтеграція та сумісність: Огляд можливостей інтеграції системи з існуючими технологіями та платформами.

Як потенційну задачу для БЛМНС будемо вважати моніторинг уявного торфовища. Відштовхуючись від задачі та можливого місця її виконання можна продовжити аналіз вимог.

Оцінка середовища експлуатації: торфовище площею 30га, пересічена рівна місцевість, помірний клімат.

Технічні вимоги: камера високої роздільної здатності, тепловізор, тривалість роботи від 2-х годин, вантажопідйомність для БПНА від 90кг.

Вимоги до комунікацій: радіочастотний зв'язок між БПЛА та БПНА.

Інтеграція та сумісність: інтеграція технології розпізнавання .

#### Огляд побудови БПЛА

Для вибору конструкції БПЛА необхідно проаналізувати специфіку території та задачі моніторингу. Так як місцевість пересічена, то можна припустити про відсутність взлітної смуги, отже БПЛА повинен робити вертикальний взліт та посадку. БПЛА має бути оснащений камерою та тепловізором та мати змогу фокусуватися над містом пожежі для її контролю.

Таблиця 1

Аналіз властивостей БПЛА

Властивість/Тип БПЛА	Однороторний	Квадрокоптер	Жорстке крило	Гібридний
Достатня підйомна сила	✓	✓	✓	✓
Вертикальний зліт/посадка	✓	✓		✓
Маневреність		✓		✓
Швидкість польоту	✓	✓	✓	✓
Тривалість польоту	✓	✓	✓	✓

Властивість/Тип БПЛА	Однороторний	Квадрокоптер	Жорстке крило	Гібридний
Нескладна механіка		✓	✓	
Результат	4	6	4	5

Згідно з виконаним аналізом приходимо до висновку, що оптимальним вибором БПЛА для виконання моніторингу в цій місцевості буде квадрокоптер, адже він відповідає усім необхідним критеріям.

#### Огляд побудови БПНА

Враховуючи специфіку місцевості та виконуваного завдання можна визначити тип використовуваного БПНА аналогічно тому, як це було зроблено в попередньому підрозділі. Необхідно враховувати, що БПНА повинен мати високу прохідну спроможність та високу підйомну силу для транспортування засобів гасіння пожежі.

Таблиця 2

#### Аналіз властивостей БПНА

Властивість/Тип БПНА	Колісний	Гусеничний	На ногах	Стрибаючий
Достатня вантажопідйомність	✓	✓		
Висока прохідна можливість		✓	✓	
Маневреність	✓	✓	✓	
Швидкість	✓	✓	✓	✓
Час автономної роботи	✓	✓		✓
Нескладна механіка	✓	✓		
Результат	5	6	3	2

Зважаючи на потребу у підвищеній прохідності, гусеничний БПЛА має перевагу над колісним, тому він найкраще підходить для виконання такого завдання.

#### Комбінування БПЛА та БПНА для виконання сумісних завдань

Комбінування безпілотних літальних апаратів (БПЛА) та безпілотних наземних апаратів (БПНА) дозволяє створити ефективну систему для комплексного моніторингу та реагування на надзвичайні ситуації в екологічно-критичних об'єктах. Такий підхід забезпечує інтеграцію переваг кожного типу апаратів, підвищуючи загальну ефективність та надійність системи [41]. Використовуючи таблиці 1 та 2 можна можна співставити БА, об'єднавши апарати, які можуть доповнювати один одного в складі БЛМНС. В результаті ми отримуємо різні конфігурації здатні для виконання різних завдань.

Таблиця 3

#### Варіанти можливих конфігурацій БЛМНС

№ Конфігурації	Тип БПЛА	Типа БПНА	Переваги	Недоліки	Можливе застосування
1	Квадрокоптер	Гусеничний	Висока маневреність, проходимість, вантажопідйомність	Обмежена тривалість польоту	Моніторинг та гасіння пожеж
2	Жорстке крило	Колісний	Висока швидкість, велика тривалість польоту	Потреба у злітно-посадній смузі, рівна поверхня	Інспекція інфраструктури
3	Гібридний апарат	Робот на ногах	Проходимість та універсальність	Вартість, вантажопідйомність, швидкість, складність	Моніторинг важкодоступних територій

Конфігурація 1 підходить для заданої задачі моніторингу та керування пожежами на торфовищах. Ця конфігурація з квадрокоптера та гусеничного апарату буде розглянута більш детально. БА в цій БЛМНС можуть виконувати наступні функції:

1. Квадрокоптер з камерою та тепловізором:
  - Моніторинг з повітря: Використання квадрокоптера для регулярного обльоту території, виявлення теплових аномалій, диму або інших ознак небезпеки.
  - Оперативне реагування: Виявлення аномалій у реальному часі та передача даних на станцію управління та БПНА для подальших дій.
2. Гусеничний наземний апарат з системою гасіння пожеж:
  - Реагування на землі: Після отримання координат від квадрокоптера, гусеничний апарат рухається до проблемної ділянки для детального огляду та виконання необхідних дій.
  - Гасіння пожеж: Використання системи гасіння пожеж для ліквідації вогню.

#### Використання ІоД для створення БЛМНС

Для взаємодії багатьох БА при виконанні сумісної місії може використовуватися ІоД з хмарним розподіленням даних та використанням протоколу зв'язку ROSLink. Дані рішення можуть обраними так як вони відповідають специфіці виконуваної задачі та мають низку переваг:

1. Впровадження ROSLink:
  - ROSLink дозволяє ефективно інтегрувати роботів, обладнаних ROS, з ІоТ [42].
  - ROSLink забезпечує доступ до будь-якого робота через Інтернет за допомогою легкого асинхронного комунікаційного протоколу.
2. Переваги хмарного розподілення даних:
  - Віртуалізація роботів та їх управління через Інтернет відкриває нові можливості для хмарних робототехнічних додатків.
  - Перенесення інтенсивних обчислень до хмарних ресурсів дозволяє подолати обмеження щодо обчислювальних потужностей, зберігання та енергоспоживання роботів.
3. Ефективність та надійність:
  - Тести ROSLink у хмарі показали його ефективність та надійність для керування роботами через Інтернет.
  - Використання високошвидкісних з'єднань та достатньої пропускної здатності сучасного Інтернету дозволяє забезпечити високу якість обслуговування для додатків, розгорнутих за допомогою ROSLink.
4. Масштабованість та гнучкість:
  - ROSLink дозволяє легко масштабувати систему, забезпечуючи можливість додавання нових роботів та користувачів без значних змін у її архітектурі.



Рис. 1 Схема взаємодії БЛМНС

#### Алгоритм роботи БЛМНС

Процес спільної роботи наведених вище безпілотних апаратів може виглядати так:

1. Виявлення проблеми:
  - Квадрокоптер здійснює обліт території та за допомогою камер і тепловізорів виявляє потенційні проблеми (наприклад, підвищення температури, задимлення).
  - Дані з квадрокоптера передаються в реальному часі на центральну станцію управління.
2. Аналіз та планування:
  - Оператори на станції управління аналізують отримані дані та визначають необхідні дії.
  - Визначаються координати проблемної ділянки та план дій для наземного апарату.

3. Реагування на землі:

- Гусеничний апарат отримує координати проблемної ділянки та вирушає на місце.
- Виконується детальний огляд території за допомогою камер та сенсорів на гусеничному апараті.
- Використовується система гасіння пожеж для ліквідації вогню або роботизована рука для виконання інших необхідних дій.

4. Оцінка результатів:

- Після завершення операції квадрокоптер знову облітає територію для оцінки результатів виконаної роботи.
- Дані передаються на центральну станцію для аналізу та підтвердження успішності операції.

**Алгоритм розпізнавання вогню та диму Fire-YOLO**

Для успішного виявлення пожежі окрім камери та тепловізору необхідний і алгоритм, який міг би проаналізувати отримане зображення та зробити висновки щодо наявності або відсутності пожежі в зоні моніторингу. Для вирішення цієї задачі може бути використаний алгоритм YOLO [43].

YOLO (You Look Only Once) - це алгоритм виявлення об'єктів у реальному часі, розроблений Джозефом Редмонном та Алі Фархаді у 2015 році. Це одноетапний детектор об'єктів, який використовує згорткову нейронну мережу (CNN) для прогнозування обмежувальних рамок та ймовірностей класів об'єктів на вхідних зображеннях.

Fire-YOLO є полегшеною версією алгоритму YOLO. В той час, як YOLO може працювати з 9000 тисячами класів, Fire-YOLO зосереджений на роботі з двома класами - дим та вогонь, і відповідно пропонує покращену швидкодію.

Виявлення класів Fire-YOLO алгоритмом виконується в декілька кроків.

По-перше, зображення буде поділено на сітку з клітинками розміру  $S \times S$  утворюючи вихідний тензор. Кожна клітинка може містити декілька об'єктів.



Рис. 2 Поділ вхідного зображення на  $S \times S$  клітин

Після цього визначається, чи є дим чи полум'я всередині кожної окремої клітини. Потім кожна клітина прогнозує три межові рамки (прямокутне утворення з декількох клітин в якому міститься об'єкт) і дає оцінку достовірностей цих межових рамок. Для цього кожній клітині вихідного тензору повинен бути співставлений вектор розміру  $B \cdot (5 + C)$ , де  $B$  - це кількість межових рамок,  $5$  - чотири координати межових рамок та один показник впевненості,  $C$  - кількість класів.

Загальна формула розміру вихідного тензору буде виглядати так:  $S \times S \times (B \cdot 5 + C)$ . Після цього застосовуються функції втрат, які допомагають отримати межові рамки зі значеннями вирогідності наявності об'єкта та певного класу [43]:

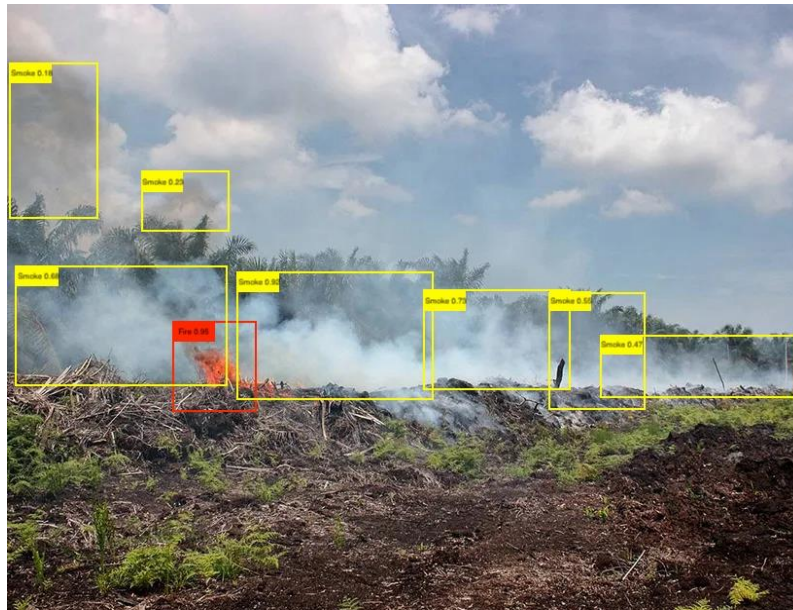


Рис. 3 Вихідне зображення з межовими рамками та визначеними класами

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

В процесі роботи були досліджені можливості БПЛА та БПНА для задач моніторингу екологічно-критичних об'єктів в складі БЛМНС. В ході роботи були розглянуті різні типи БА та проведено аналіз їх можливостей для виконання задач моніторингу. Квадрокоптер на гученичний апарат виявились найбільш придатними для задач повітряного та наземного моніторингу.

В ході аналізу дотичних робіт було виявлено, що використання БЛМНС може бути ефективнішим, ніж використання поодинокого БА [9]. Інтеграція різних типів БА дозволяє створити ефективні, адаптивні системи, що можуть оперативно реагувати на змінні умови та виявлені загрози. Виявлено, що комбінація БПЛА та БПНА забезпечує підвищену ефективність моніторингу та реагування завдяки інтеграції переваг кожного типу апаратів. Це включає використання квадрокоптерів для повітряного огляду та гусеничних апаратів для детального наземного аналізу та гасіння пожеж.

В ході роботи були розглянуті можливі методи взаємодії БА в БЛМНС. Було розглянуто можливості IoD з хмарним розподіленням даних та протокол ROSLink для комунікації приладів в системі.

Був досліджений алгоритм та математичний апарат розпізнавання об'єктів YOLO. Фреймворк YOLO-Fire вважається одним з найефективніших алгоритмів для ідентифікації пожеж та може бути використаний для задач моніторингу.

Загалом, проведене дослідження підтвердило високий потенціал використання БЛМНС для задач моніторингу екологічно-критичних об'єктів та запропонувало один із можливих варіантів такої системи.

Майбутні дослідження та розробки в сфері БЛМНС можуть зосередитися на:

1. Розвитку автономності: Підвищення рівня автономності безпілотних апаратів для зменшення залежності від людського фактора.
2. Забезпеченню безпеки: Розробка і впровадження нових методів захисту даних та систем від кіберзагроз.
3. Ефективному управлінні ресурсами: Оптимізація використання енергетичних і обчислювальних ресурсів для підвищення ефективності системи.
4. Інтеграція з іншими системами: Забезпечення сумісності та інтеграції БЛМНС з іншими існуючими системами моніторингу та управління.
5. Покращення показників ефективності: Час реакції, точність моніторингу, стійкість системи тощо.

Це сприятиме підвищенню якості моніторингу та швидкості реагування на надзвичайні ситуації, що в свою чергу покращить захист навколишнього середовища та безпеку населення.

#### Література

1. Бондар Д.В., Гурник А.В., Литовченко А.О., Хижняк В.В., Шевченко В.Л., Ядченко Д.М. Застосування безпілотних авіаційних систем у сфері цивільного захисту. Київ : ГО «Європейська наукова платформа», 2022. С. 16-22.

2. Горішний Є., Міняйло А., Дмитренко Л. Екологічні ризики та битки доквіллю України внаслідок війни. Київ : Український науково-дослідний інститут спеціальної техніки та судових експертиз СБУ, 2023. DOI 10.31733/2078-3566-2023-4-76-83
3. Банчук М.В. Система моніторингу санітарно-епідеміологічної ситуації в Україні як один з найважливіших елементів стратегії державного управління з забезпечення здоров'я населення. Державне управління: удосконалення та розвиток № 10, 2011.
4. Клименко М.О., Прищепя А.М., Вознюк Н.М. Моніторинг доквілля. Київ : Академія, 2006. 360 с.
5. Skorobohatko, S., Fesenko, H., Kharchenko, V. et al. Architecture and Reliability Models of Hybrid Sensor Networks for Environmental and Emergency Monitoring Systems. *Cybern Syst Anal* 60, 293–304 (2024). DOI 10.1007/s10559-024-00670-x.
6. Leichenko, K. et al. Deployment of a UAV swarm-based LiFi network in the obstacle-ridden environment: algorithms of finding the path for UAV placement. *Radioelectronic and Computer Systems*, [S.l.], v. 2024, n. 1, p. 176-195, feb. 2024. DOI 10.32620/reks.2024.1.14.
7. Fesenko, H.; Illiashenko, O.; Kharchenko, V.; Kliushnikov, I.; Morozova, O.; Sachenko, A.; Skorobohatko, S. Flying Sensor and Edge Network-Based Advanced Air Mobility Systems: Reliability Analysis and Applications for Urban Monitoring. *Drones* 2023, 7, 409. DOI 10.3390/drones7070409.
8. Louie, A.H. Relational biology of symbiosis. 2010. P. 1. DOI 10.1007/s10516-010-9117-9.
9. Borges, P.; Peynot, T.; Liang, S.; Arain, B.; Wildie, M.; Minareci, M.; Lichman, S.; Samvedi, G.; Sa, I.; Hudson, N.; et al. A Survey on Terrain Traversability Analysis for Autonomous Ground Vehicles: Methods, Sensors, and Challenges. *Field Robot.* 2022. pp. 1567-1627.
10. Lanctot, S.; Herkenhoff, B.; Hassanalіan, M. Unmanned Launching & Landing Rover (ULLR) for Moon and Martian Missions on Ice-Caps. In *Proceedings of the 2021 ASCEND Conference*, Las Vegas, NV, USA, 15–17 November 2021. DOI 10.2514/6.2021-4201.
11. John, T.S. Advancements in robotics and its future uses. *Int. J. Sci. Eng. Res.* 2011. pp. 1–6.
12. Gonzalez-De-Santos, P.; Fernández, R.; Sepúlveda, D.; Navas, E.; Armada, M. Unmanned ground vehicles for smart farms. In *Agronomy-Climate Change and Food Security*; Intech UK: Horwich, UK, 2020. P. 73.
13. Michael R. Blackburn, Richard Bailey, Brent Lytle, "Improved mobility in a multi-degree-of-freedom unmanned ground vehicle," *Proc. SPIE 5422, Unmanned Ground Vehicle Technology VI*, (2 September 2004). DOI 10.1117/12.544715.
14. Herkenhoff, B.; Lanctot, S.; Bjorkman, T.; Serda, N.; Hassanalіan, M. Preliminary Design Concept of Locust Inspired Jumping Moon Robot Swarm. In *Proceedings of the AIAA Propulsion and Energy 2021 Forum*, Virtual Event, 9–11 August 2021. P. 3270. DOI 10.2514/6.2021-3270.
15. Western, A.; Cervantes, R.; Dunning, C.; Haghshenas Jaryani, M.; Hassanalіan, M. Bioinspired Robot with Walking, Rolling, and Jumping Capabilities for Planetary Exploration. In *Proceedings of the AIAA Aviation 2021 Forum*, Virtual Event, 2–6 August 2021. P. 2784. DOI 10.2514/6.2021-2784.
16. Herkenhoff, B.K.; Lanctot, S.I.; Fisher, J.M.; Serda, N.; Bjorkman, T.S.; Martinez, V.; Johnsonand, T.; Davis, C.; Yazzie, T.; Vadiee, N.; et al. Preliminary Design Concept of Locust Inspired Jumping Moon Robot Swarm. In *Proceedings of the Lunar and Planetary Science Conference*, Online, 15–19 March 2021. P. 2754. DOI 10.2514/6.2021-3270.
17. Western, A.; Haghshenas-Jaryani, M.; Hassanalіan, M. Golden wheel spider-inspired rolling robots for planetary exploration. *Acta Astronaut.* 2022. pp. 34–48. DOI 10.1016/j.actaastro.2022.12.030.
18. Carlo, M.; Metin, S. A biomimetic climbing robot based on the gecko. *J. Bionic Eng.* 2006. pp. 115–125. DOI 10.1016/S1672-6529(06)60015-2.
19. Liu, GH., Lin, HY., Lin, HY. et al. A Bio-Inspired Hopping Kangaroo Robot with an Active Tail. *J Bionic Eng* 11. pp. 541–555 (2014). DOI 10.1016/S1672-6529(14)60066-4.
20. Darvishpoor, S.; Roshanian, J.; Raissi, A.; Hassanalіan, M. Classifications, configurations, and flight mechanisms of unmanned aerial systems: A review. *Prog. Aerosp. Sci.* 2021, 121, 100694. DOI 10.1016/j.paerosci.2020.100694.
21. Segui-Gasco, P.; Al-Rihani, Y.; Shin, H.S.; Savvaris, A. A Novel Actuation Concept for a Multi Rotor UAV. *J. Intell. Robot. Syst.* 2014. pp. 173–191.
22. Cai, G.; Lum, K.Y.; Chen, B.M.; Lee, T.H. A brief overview on miniature fixed-wing unmanned aerial vehicles. In *Proceedings of the IEEE ICCA 2010*, Xiamen, China, 9–11 June 2010. pp. 285–290.
23. Saeed, A.S.; Younes, A.B.; Cai, C.; Cai, G. A survey of hybrid Unmanned Aerial Vehicles. *Prog. Aerosp. Sci.* 2018. pp. 91–105.
24. Al-Ghafri, Y.; Asif, H.M.; Tarhuni, N.; Nadir, Z. Advancing Non-Line-of-Sight Communication: A Comprehensive Review of State-of-the-Art Technologies and the Role of Energy Harvesting. *Sensors* 2024, 24, 4671. DOI 10.3390/s24144671
25. Mohanta, K.; Al-Rubaye, S. Towards 6G Satellite–Terrestrial Networks: Analysis of Air Mobility Operations. *Electronics* 2024, 13, 2855. DOI 10.3390/electronics13142855.
26. Choudhary, G.; Sharma, V.; You, I. Sustainable and secure trajectories for the military Internet of Drones (IoD) through an efficient Medium Access Control (MAC) protocol. *Comput. Electr. Eng.* 2019. pp. 59–73. DOI 10.1016/j.compeleceng.2019.01.007.



27. Lagkas, T.; Argyriou, V.; Bibi, S.; Sarigiannidis, P. UAV IoT Framework Views and Challenges: Towards Protecting Drones as “Things”. *Sensors* 2018, 18, 4015. DOI 10.3390/s18114015.
28. M. Gharibi, R. Boutaba and S. L. Waslander, "Internet of Drones," in *IEEE Access*, vol. 4. pp. 1148-1162, 2016, doi: 10.1109/ACCESS.2016.2537208.
29. Bousbaa, F.Z.; Kerrache, C.A.; Mahi, Z.; Tahari, A.E.K.; Lagraa, N.; Yagoubi, M.B. GeoUAVs: A new geocast routing protocol for fleet of UAVs. *Comput. Commun.* 2020. pp. 259–269. DOI 10.1016/j.comcom.2019.10.026.
30. A. I. A. Ahmed et al., "Service Management for IoT: Requirements, Taxonomy, Recent Advances and Open Research Challenges," in *IEEE Access*, vol. 7. pp. 155472-155488, 2019. DOI 10.1109/ACCESS.2019.2948027.
31. Vallejo, D.; Castro-Schez, J.; Glez-Morcillo, C.; Albusac, J. Multi-agent architecture for information retrieval and intelligent monitoring by UAVs in known environments affected by catastrophes. *Eng. Appl. Artif. Intell.* 2020, 87, 103243. DOI 10.1016/j.engappai.2019.103243.
32. Juan López, Pablo Royo, Enric Pastor, Cristina Barrado, and Eduard Santamaria. 2007. A middleware architecture for unmanned aircraft avionics. In *Proceedings of the 2007 ACM/IFIP/USENIX international conference on Middleware companion (MC '07)*. Association for Computing Machinery, New York, NY, USA, Article 24, 1–6. DOI 10.1145/1377943.1377962.
33. Sharma, V.; Kumar, R.; Patiala, P. Service-oriented middleware for multi-UAV guided ad hoc networks. *IT Converg. Pract.* 2014. pp. 24–33.
34. J. P. Ribeiro et al., "UAV cooperative perception based on DDS communications network," *OCEANS 2017 - Anchorage, Anchorage, AK, USA, 2017*. pp. 1-8.
35. S. Aggarwal, M. Shojafar, N. Kumar and M. Conti, "A New Secure Data Dissemination Model in Internet of Drones," *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019. pp. 1-6. DOI 10.1109/ICC.2019.8761372.
36. Xu, S.; Doğançay, K.; Hmam, H. Distributed pseudolinear estimation and UAV path optimization for 3D AOA target tracking. *Signal. Process.* 2017. pp. 64–78. DOI 10.1016/j.sigpro.2016.10.012.
37. Patra, A.N.; Regis, P.A.; Sengupta, S. Distributed allocation and dynamic reassignment of channels in UAV networks for wireless coverage. *Pervasive Mob. Comput.* 2019. pp. 58–70. DOI 10.1016/j.pmcj.2019.02.005.
38. Mohamed, N.; Al-Jaroodi, J.; Jawhar, I.; Idries, A.; Mohammed, F. Unmanned aerial vehicles applications in future smart cities. *Technol. Forecast. Soc. Chang.* 2018, 153, 119293. DOI 10.1016/j.techfore.2018.05.004.
39. Кобріна Н.В. Застосування безпілотних авіаційних комплексів для вирішення екологічних завдань. Харків : ДП “УКРНТЦ ЕНЕРГОСТАЛЬ”, 2014. С. 88-90.
40. Пітак І.В., Негадайлов А.А., Масікевич Ю.Г. Геоінформаційні технології в екології. Чернівці, 2012. С. 273.
41. Dinelli, C.; Racette, J.; Escarcega, M.; Lotero, S.; Gordon, J.; Montoya, J.; Dunaway, C.; Androulakis, V.; Khaniani, H.; Shao, S.; et al. Configurations and Applications of Multi-Agent Hybrid Drone/Unmanned Ground Vehicle for Underground Environments: A Review. *Drones* 2023, 7, 136. DOI 10.3390/drones7020136.
42. Koubaa, A.; Alajlan, M.; Qureshi, B. ROSLnk: Bridging ROS with the Internet-of-Things for cloud robotics. *Studies in Computational Intelligence*, 2017. DOI 10.1007/978-3-319-54927-9\_8.
43. Luan, T.; Zhou, S.; Zhang, G.; Song, Z.; Wu, J.; Pan, W. Enhanced Lightweight YOLOX for Small Object Wildfire Detection in UAV Imagery. *Sensors* 2024, 24, 2710. DOI 10.3390/s24092710.

## References

1. Bondar, D.V., Gurnik, A.V., Lytovchenko, A.O., Khizhnyak, V.V., Shevchenko, V.L., Yadchenko, D.M. Application of Unmanned Aerial Systems in Civil Protection. Kyiv: NGO "European Scientific Platform", 2022. pp. 16-22.
2. Horishniy, Y., Minyaylo, A., Dmytrenko, L. Environmental Risks and Damage to Ukraine's Environment as a Result of the War. Kyiv: Ukrainian Scientific Research Institute of Special Equipment and Forensic Science of the Security Service of Ukraine, 2023. DOI: 10.31733/2078-3566-2023-4-76-83
3. Banchuk, M.V. The Monitoring System of the Sanitary and Epidemiological Situation in Ukraine as One of the Most Important Elements of the State Management Strategy for Ensuring Public Health. *Public Administration: Improvement and Development*, No. 10, 2011.
4. Klymenko, M.O., Pryshchepa, A.M., Vozniuk, N.M. Environmental Monitoring. Kyiv: Akademia, 2006. 360 pages. Skorobohatko, S., Fesenko, H., Kharchenko, V. et al. Architecture and Reliability Models of Hybrid Sensor Networks for Environmental and Emergency Monitoring Systems, 293–304 (2024). DOI 10.1007/s10559-024-00670-x.
5. Skorobohatko, S., Fesenko, H., Kharchenko, V. et al. Architecture and Reliability Models of Hybrid Sensor Networks for Environmental and Emergency Monitoring Systems. *Cybern Syst Anal* 60, 293–304 (2024). DOI 10.1007/s10559-024-00670-x.
6. Leichenko, K. et al. Deployment of a UAV swarm-based LiFi network in the obstacle-ridden environment: algorithms of finding the path for UAV placement. *Radioelectronic and Computer Systems, [S.I.]*, v. 2024, n. 1, p. 176-195, feb. 2024. DOI 10.32620/reks.2024.1.14.
7. Fesenko, H.; Iliashenko, O.; Kharchenko, V.; Kliushnikov, I.; Morozova, O.; Sachenko, A.; Skorobohatko, S. Flying Sensor and Edge Network-Based Advanced Air Mobility Systems: Reliability Analysis and Applications for Urban Monitoring. *Drones* 2023, 7, 409. DOI 10.3390/drones7070409.
8. Louie, A.H. Relational biology of symbiosis. 2010. P. 1. DOI 10.1007/s10516-010-9117-9.
9. Borges, P.; Peynot, T.; Liang, S.; Arain, B.; Wildie, M.; Minareci, M.; Lichman, S.; Samvedi, G.; Sa, I.; Hudson, N.; et al. A Survey on Terrain Traversability Analysis for Autonomous Ground Vehicles: Methods, Sensors, and Challenges. *Field Robot.* 2022. pp. 1567-1627.
10. Lancot, S.; Herkenhoff, B.; Hassanalain, M. Unmanned Launching & Landing Rover (ULLR) for Moon and Martian Missions on Ice-Caps. In *Proceedings of the 2021 ASCEND Conference, Las Vegas, NV, USA, 15–17 November 2021*. DOI 10.2514/6.2021-4201.

11. John, T.S. Advancements in robotics and its future uses. *Int. J. Sci. Eng. Res.* 2011. pp. 1–6.
12. Gonzalez-De-Santos, P.; Fernández, R.; Sepúlveda, D.; Navas, E.; Armada, M. Unmanned ground vehicles for smart farms. In *Agronomy–Climate Change and Food Security*; Intech UK: Horwich, UK, 2020. P. 73.
13. Michael R. Blackburn, Richard Bailey, Brent Lytle, "Improved mobility in a multi-degree-of-freedom unmanned ground vehicle," *Proc. SPIE 5422, Unmanned Ground Vehicle Technology VI*, (2 September 2004). DOI 10.1117/12.544715.
14. Herkenhoff, B.; Lanctot, S.; Bjorkman, T.; Serda, N.; Hassanalian, M. Preliminary Design Concept of Locust Inspired Jumping Moon Robot Swarm. In *Proceedings of the AIAA Propulsion and Energy 2021 Forum, Virtual Event*, 9–11 August 2021. P. 3270. DOI 10.2514/6.2021-3270.
15. Western, A.; Cervantes, R.; Dunning, C.; Haghshenas Jaryani, M.; Hassanalian, M. Bioinspired Robot with Walking, Rolling, and Jumping Capabilities for Planetary Exploration. In *Proceedings of the AIAA Aviation 2021 Forum, Virtual Event*, 2–6 August 2021. P. 2784. DOI 10.2514/6.2021-2784.
16. Herkenhoff, B.K.; Lanctot, S.I.; Fisher, J.M.; Serda, N.; Bjorkman, T.S.; Martinez, V.; Johnsonand, T.; Davis, C.; Yazzie, T.; Vadiie, N.; et al. Preliminary Design Concept of Locust Inspired Jumping Moon Robot Swarm. In *Proceedings of the Lunar and Planetary Science Conference, Online*, 15–19 March 2021. P. 2754. DOI 10.2514/6.2021-3270.
17. Western, A.; Haghshenas-Jaryani, M.; Hassanalian, M. Golden wheel spider-inspired rolling robots for planetary exploration. *Acta Astronaut.* 2022. pp. 34–48. DOI 10.1016/j.actaastro.2022.12.030.
18. Carlo, M.; Metin, S. A biomimetic climbing robot based on the gecko. *J. Bionic Eng.* 2006. pp. 115–125. DOI 10.1016/S1672-6529(06)60015-2.
19. Liu, GH., Lin, HY., Lin, HY. et al. A Bio-Inspired Hopping Kangaroo Robot with an Active Tail. *J Bionic Eng* 11. pp. 541–555 (2014). DOI 10.1016/S1672-6529(14)60066-4.
20. Darvishpoor, S.; Roshanian, J.; Raissi, A.; Hassanalian, M. Classifications, configurations, and flight mechanisms of unmanned aerial systems: A review. *Prog. Aerosp. Sci.* 2021, 121, 100694. DOI 10.1016/j.paerosci.2020.100694.
21. Segui-Gasco, P.; Al-Rihani, Y.; Shin, H.S.; Savvaris, A. A Novel Actuation Concept for a Multi Rotor UAV. *J. Intell. Robot. Syst.* 2014. pp. 173–191.
22. Cai, G.; Lum, K.Y.; Chen, B.M.; Lee, T.H. A brief overview on miniature fixed-wing unmanned aerial vehicles. In *Proceedings of the IEEE ICCA 2010, Xiamen, China*, 9–11 June 2010. pp. 285–290.
23. Saeed, A.S.; Younes, A.B.; Cai, C.; Cai, G. A survey of hybrid Unmanned Aerial Vehicles. *Prog. Aerosp. Sci.* 2018. pp. 91–105.
24. Al-Ghafri, Y.; Asif, H.M.; Tarhuni, N.; Nadir, Z. Advancing Non-Line-of-Sight Communication: A Comprehensive Review of State-of-the-Art Technologies and the Role of Energy Harvesting. *Sensors* 2024, 24, 4671. DOI 0.3390/s24144671
25. Mohanta, K.; Al-Rubaye, S. Towards 6G Satellite–Terrestrial Networks: Analysis of Air Mobility Operations. *Electronics* 2024, 13, 2855. DOI 10.3390/electronics13142855.
26. Choudhary, G.; Sharma, V.; You, I. Sustainable and secure trajectories for the military Internet of Drones (IoD) through an efficient Medium Access Control (MAC) protocol. *Comput. Electr. Eng.* 2019. pp. 59–73. DOI 10.1016/j.compeleceng.2019.01.007.
27. Lagkas, T.; Argyriou, V.; Bibi, S.; Sarigiannidis, P. UAV IoT Framework Views and Challenges: Towards Protecting Drones as “Things”. *Sensors* 2018, 18, 4015. DOI 10.3390/s18114015.
28. M. Gharibi, R. Boutaba and S. L. Waslander, "Internet of Drones," in *IEEE Access*, vol. 4. pp. 1148-1162, 2016, doi: 10.1109/ACCESS.2016.2537208.
29. Bousbaa, F.Z.; Kerrache, C.A.; Mahi, Z.; Tahari, A.E.K.; Lagraa, N.; Yagoubi, M.B. GeoUAVs: A new geocast routing protocol for fleet of UAVs. *Comput. Commun.* 2020. pp. 259–269. DOI 10.1016/j.comcom.2019.10.026.
30. A. I. A. Ahmed et al., "Service Management for IoT: Requirements, Taxonomy, Recent Advances and Open Research Challenges," in *IEEE Access*, vol. 7. pp. 155472-155488, 2019. DOI 10.1109/ACCESS.2019.2948027.
31. Vallejo, D.; Castro-Schez, J.; Glez-Morcillo, C.; Albusac, J. Multi-agent architecture for information retrieval and intelligent monitoring by UAVs in known environments affected by catastrophes. *Eng. Appl. Artif. Intell.* 2020, 87, 103243. DOI 10.1016/j.engappai.2019.103243.
32. Juan López, Pablo Royo, Enric Pastor, Cristina Barrado, and Eduard Santamaria. 2007. A middleware architecture for unmanned aircraft avionics. In *Proceedings of the 2007 ACM/IFIP/USENIX international conference on Middleware companion (MC '07)*. Association for Computing Machinery, New York, NY, USA, Article 24, 1–6. DOI 10.1145/1377943.1377962.
33. Sharma, V.; Kumar, R.; Patiala, P. Service-oriented middleware for multi-UAV guided ad hoc networks. *IT Converg. Pract.* 2014. pp. 24–33.
34. J. P. Ribeiro et al., "UAV cooperative perception based on DDS communications network," *OCEANS 2017 - Anchorage, Anchorage, AK, USA*, 2017. pp. 1-8.
35. S. Aggarwal, M. Shojafar, N. Kumar and M. Conti, "A New Secure Data Dissemination Model in Internet of Drones," *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*, Shanghai, China, 2019. pp. 1-6. DOI 10.1109/ICC.2019.8761372.
36. Xu, S.; Doğançay, K.; Hmam, H. Distributed pseudolinear estimation and UAV path optimization for 3D AOA target tracking. *Signal. Process.* 2017. pp. 64–78. DOI 10.1016/j.sigpro.2016.10.012.
37. Patra, A.N.; Regis, P.A.; Sengupta, S. Distributed allocation and dynamic reassignment of channels in UAV networks for wireless coverage. *Pervasive Mob. Comput.* 2019. pp. 58–70. DOI 10.1016/j.pmcj.2019.02.005.
38. Mohamed, N.; Al-Jaroodi, J.; Jawhar, I.; Idries, A.; Mohammed, F. Unmanned aerial vehicles applications in future smart cities. *Technol. Forecast. Soc. Chang.* 2018, 153, 119293. DOI 10.1016/j.techfore.2018.05.004.
39. Kobrina, N.V. Application of Unmanned Aerial Complexes for Solving Environmental Tasks. Kharkiv: State Enterprise "Ukrainian Scientific and Technical Center ENERGOSTAL", 2014. pp. 88-90.
40. Pitak, I.V., Negadaylov, A.A., Masikevich, Y.G. Geoinformation Technologies in Ecology. Chernivtsi, 2012. p. 273. Dinelli, C.; Racette, J.; Escarcega, M.; Lotero, S.; Gordon, J.; Montoya, J.; Dunaway, C.; Androulakis, V.; Khaniani, H.; Shao, S.; et al. Configurations and Applications of Multi-Agent Hybrid Drone/Unmanned Ground Vehicle for Underground Environments: A Review. *Drones* 2023, 7, 136. DOI 10.3390/drones7020136.
41. Dinelli, C.; Racette, J.; Escarcega, M.; Lotero, S.; Gordon, J.; Montoya, J.; Dunaway, C.; Androulakis, V.; Khaniani, H.; Shao, S.; et al. Configurations and Applications of Multi-Agent Hybrid Drone/Unmanned Ground Vehicle for Underground Environments: A Review. *Drones* 2023, 7, 136. DOI 10.3390/drones7020136.
42. Koubaa, A.; Alajlan, M.; Qureshi, B. ROSLnk: Bridging ROS with the Internet-of-Things for cloud robotics. *Studies in Computational Intelligence*, 2017. DOI 10.1007/978-3-319-54927-9\_8.
43. Luan, T.; Zhou, S.; Zhang, G.; Song, Z.; Wu, J.; Pan, W. Enhanced Lightweight YOLOX for Small Object Wildfire Detection in UAV Imagery. *Sensors* 2024, 24, 2710. DOI 10.3390/s24092710.

<https://doi.org/10.31891/2219-9365-2024-79-2>

УДК 004.8

ВОЛОКИТА Артем

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

<https://orcid.org/0000-0001-9069-5544>

e-mail: [artem.volokita@kpi.ua](mailto:artem.volokita@kpi.ua)

МОРОЗОВ-ЛЕОНОВ Олександр

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

e-mail: [olmorleon@yahoo.com](mailto:olmorleon@yahoo.com)

<https://orcid.org/0009-0001-8283-0248>

## ПІДХОДИ ТА МЕТОДИ ПРИЙНЯТТЯ РІШЕНЬ В СЕРЕДОВИЩАХ ІЗ НЕПОВНОЮ ТА НЕВИЗНАЧЕНОЮ ІНФОРМАЦІЄЮ

*В сучасному світі часто виникають задачі, в яких необхідно приймати рішення на основі неповної або невизначеної інформації. У цій статті зроблено огляд деяких сучасних методів та підходів прийняття рішень, наведені їхні сильні та слабкі сторони, особливості застосування та можливості до інтеграції в інші галузі.*

*Розгляд методів та підходів показав що саме методи, що базуються на навчанні з підкріпленням, є найбільш універсальними, ефективними та мають потенціал для подальшого покращення. Показані також можливі рішення для прийняття рішень із плануванням.*

*Ключові слова: прийняття рішень, неповна інформація, невизначена інформація, навчання з підкріпленням.*

VOLOKYTA Artem, MOROZOV-LEONOV Oleksandr  
National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute"

## APPROACHES AND METHODS OF DECISION-MAKING IN ENVIRONMENTS WITH INCOMPLETE AND UNCERTAIN INFORMATION

*Artificial intelligence (AI) plays an increasingly important role in decision-making in the modern world, providing effective solutions in various spheres of activity. One of the most important areas of decision-making for the functioning of modern society is the navigation and safety of autonomous vehicles and cyber security.*

*With the growing complexity and scale of the environments in which these artificial intelligence systems work, new challenges arise: yes, in many cases, autonomous agents must interact with dynamic and multifactorial environments, where an accurate description of all the factors necessary for the agent's work is often impossible. This requires the development and application of special approaches and algorithms capable of adapting to unpredictable conditions and making decisions based on incomplete or inaccurate information.*

*In the future, the scope and role of using artificial intelligence for decision-making will continue to grow, which will lead to even greater complexity of the environments in which these systems operate. This indicates the great potential of scientific research in this direction.*

*In the modern world, tasks often arise in which it is necessary to make decisions based on incomplete or uncertain information. This article provides an overview of some modern decision-making methods and approaches, their strengths and weaknesses, application features, and possibilities for integration into other industries.*

*The review of methods and approaches has shown that methods based on reinforcement learning are the most versatile, effective and have the potential for further improvement. Possible solutions for decision-making with planning are also shown.*

*The purpose of this paper is to review approaches and methods for solving decision-making problems with incomplete or uncertain information in various fields, to determine the feasibility of using them based on the requirements of the application area, and to analyze their flexibility and versatility.*

*Keywords: decision making; incomplete information; uncertain information; reinforcement learning.*

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Штучний інтелект (ШІ) відіграє все більшу роль у прийнятті рішень у сучасному світі, забезпечує ефективні рішення в різних сферах діяльності. Одними із найбільш важливих сфер прийняття рішень для функціонування сучасного суспільства є навігація та безпека автономних транспортних засобів та кібербезпека.

Зі зростанням складності та масштабів середовищ, у яких працюють ці системи штучного інтелекту, виникають нові виклики: так, в багатьох випадках автономні агенти повинні взаємодіяти із динамічними та багатофакторними середовищами, де точний опис всіх необхідних для роботи агента факторів часто є неможливим. Це вимагає розробки та застосування спеціальних підходів та алгоритмів, здатних адаптуватися до непередбачуваних умов та приймати рішення на основі неповної або неточної інформації.

У майбутньому масштаби та роль використання штучного інтелекту для прийняття рішень будуть зростати і надалі, що призводитиме до ще більшої складності середовищ, у яких діють ці системи. Це свідчить про великий потенціал наукових досліджень у цьому напрямку.

## ПОСТАНОВКА ПРОБЛЕМИ

У багатьох реальних ситуаціях, у яких агенти мають приймати рішення, інформація часто буває неповною чи неточною через випадкові збурення у динаміці системи чи через обмежені можливості спостереження за поточним станом або поведінкою інших агентів. Розвиток методів та підходів прийняття рішень в середовищах із таким характером інформації є критично важливим для досягнення надійності та безпеки роботи систем в умовах ризику та невпевненості та мінімізації можливих втрат. У зв'язку із різноманіттям галузей та конкретних рішень існує необхідність в їх огляді, аналізі їх слабких та сильних сторін та особливостей застосування.

## АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Методам прийняття рішень із неповною чи невизначеною інформацією присвячені наукові праці [1]-[13]. Стаття [1] приводить класифікацію підходів до вирішення задач прийняття рішень для автономних транспортних засобів, а саме підходи теорії ігор, вірогіднісні підходи, частково спостережувані марківські процеси вирішування (англ. partially observable Markov decision process, POMDP) та підходи, що базуються на навчанні. В роботі [2] розглядається проблематика невизначеності у кібербезпеці. Автори статті [3] розглядають методи наближеного розв'язання задач, сформальованих у POMDP. У статті [4] розглядається адаптація MCTS на неперервні простори дій (англ. continuous action spaces) із станом, що не спостерігається напряму, та поєднання із навчанням з підкріпленням (reinforcement learning, RL) для тактичного прийняття рішень автономними транспортними засобами. У статті [5] розглянуто використання ансамблю нейронних мереж для прийняття рішень в ситуаціях із високою невизначеністю, навіть якщо вони суттєво відрізняються від навчальних даних, та із дотриманням вимог безпеки через перехід до заздалегідь безпечної дії при великій мірі невизначеності.

## ВИДІЛЕННЯ НЕДОСЛІДЖЕНИХ ЧАСТИН ЗАГАЛЬНОЇ ПРОБЛЕМИ

В результаті аналізу публікацій показано, що існує потреба у об'єднанні та абстрагуванні методів вирішення задач прийняття рішень від конкретних сфер застосування, виходячи із загальних вимог та обмежень, що притаманні багатьом із них. Існує необхідність в розгляді доцільності використання різних методів та підходів, виходячи із вимог швидкості навчання, швидкодії та точності.

## ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є огляд підходів та методів розв'язання задач прийняття рішень при неповній чи невизначеній інформації у різних сферах, визначення доцільності використання виходячи із вимог сфери застосування, та аналіз їх гнучкості і універсальності. Особливу увагу приділено аналізу конкретних покращень та розробок, впроваджених у різних сферах застосування, з метою виявлення універсальних підходів, що можуть бути ефективно адаптовані та використані в інших галузях для підвищення ефективності прийняття рішень в умовах невизначеності.

## ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Неповнота та невизначеність інформації в середовищах стосується ситуацій, коли агент або система не має доступу до всіх необхідних для прийняття рішення даних, або коли ці дані є неоднозначними або містять суттєві похибки. Наприклад, в системах автопілоту транспортних засобів невідомими є наміри водіїв інших транспортних засобів. Автопілот може адекватно вимірювати положення та швидкість оточуючих транспортних засобів, проте передбачити зміни смуг або різке гальмування інших водіїв він не може. Ще одним прикладом невизначеної інформації можуть бути GPS-дані про положення транспортних засобів. В сфері кібербезпеки прикладами такої інформації можуть бути невідомі вектори атаки та неповна мережева інформація про стан системи, що зазнає кібератаки.

Основним сучасним підходом до формулювання таких задач є застосування математичного апарату POMDP. Існують традиційні методи точного вирішення задач, сформульованих за його допомогою, однак їх застосування не є доцільним для сучасних застосувань штучного інтелекту через обчислювальну складність та конкретні вимоги та обмеження, що визначаються сферою застосування: так, для задач прийняття рішень для автономних транспортних засобів та для захисту від кібератак важливим фактором є швидкість отримання рішення, хоча б наближеного.

Основною групою методів для розв'язання цих задач є навчання з підкріпленням, зокрема глибоке (deep reinforcement learning, DRL). В ньому агент вчиться приймати рішення шляхом взаємодії із середовищем та отримання зворотнього зв'язку у вигляді винагород. В умовах невизначеності чи неповноти інформації агент поступово набуває досвіду та оптимізує свою стратегію дій, навіть не маючи повної інформації про всі можливі стани середовища або наслідки своїх дій.

За допомогою деяких засобів роботи із неповною та невизначеною інформацією можливе поступове уточнення чи апроксимація початково невідомих факторів середовища. Це дозволяє також використовувати

модифіковані версії більш традиційних методів прийняття рішень у середовищах із повною інформацією у поєднанні із навчанням з підкріпленням.

Приклад такого поєднання наводять автори [4]. В статті розглядається задача тактичного прийняття рішень автономним транспортним засобом на багатополосній трасі серед трафіку. Класична форма алгоритму пошуку по дереву Монте-Карло (англ. Monte Carlo tree search, MCTS) є одним із найбільш часто застосовуваних засобів теорії ігор, проте не підходить для прийняття рішень у середовищі із неповною чи невизначеною інформацією. Автори адаптували метод на неперервні простори дій (англ. continuous action spaces) зі станом, що не спостерігається напряму, та поєднали його із навчанням з підкріпленням, одержаний алгоритм автори назвали MCTS/NN. Пошук по дереву надає можливості для планування, що може бути перерване у будь-який час із приблизним рішенням, а за наявності більшого часу для обчислень результат покращується. Для опрацювання невизначеної інформації щодо станів водіїв оточуючих транспортних засобів використовується частковий фільтр (англ. particle filter), що надає апроксимуючу інформацію про найімовірніші рішення оточуючих водіїв, виходячи із історії їх дій у минулому. Порівняння із класичними методами навігації транспортних засобів на багатополосних трасах та із застосуванням лише компоненту навчання чи компоненту планування ясно вказує на переваги у швидкості переміщення та уникнення аварій при застосуванні розробленого методу. Автори вказують на перевагу MCTS/NN над використанням глибоких Q-мереж (Deep Q-Network, DQN) в ефективності використання окремих зразків даних при навчанні завдяки застосуванню пошуку по дереву: для навчання їх потрібно менше, що може бути особливо важливо для застосувань метода у сферах, де неможливо чи недоцільно використовувати комп'ютерну симуляцію, а необхідний саме збір реальних даних для навчання. Також автори наголошують на гнучкості та загальності розробленого методу.

Застосування покращеної версії Double DQN (DDQN) описується в [6]. Автори пропонують новий метод навчання модифікації Double DQN - Rainbow DQN - із більш ефективним використанням навчальних даних, що поєднує прийняття тактичних рішень у динамічному середовищі, дані в якому невизначені та зашумлені, із дотриманням обмежень, що накладаються вимогами безпеки. Отримана реалізація Rainbow DQN перевершує звичайний Double DQN у ефективності, а новітній спосіб використання шару захисту (safety layer) для винагород, названий авторами safety feedback, значно покращує ефективність використання даних при навчанні та продуктивність роботи (рис. 1). Експерименти в симуляції демонструють здатність отриманої реалізації знаходити рішення в ситуаціях, коли необхідне довгострокове планування, та добре справляється із невизначеністю та зашумленістю інформації.

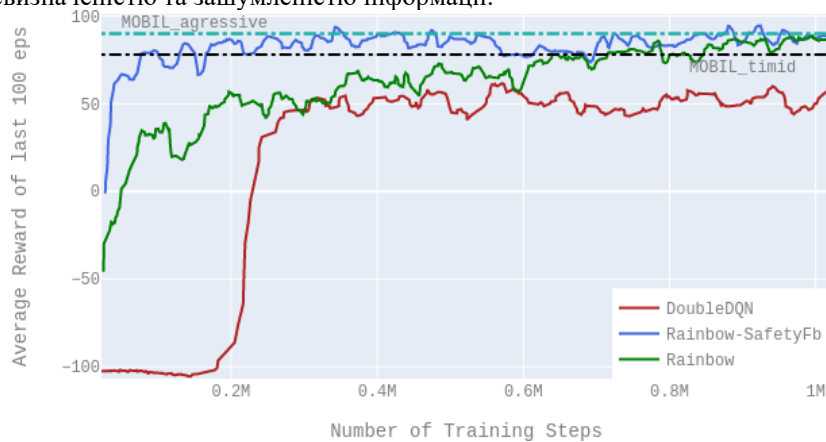


Рис. 1. Ефективність навчання стандартного DDQN та двох версій Rainbow DQN: із safety feedback та без нього  
Джерело: [6]

Покращення DQN із використанням методу байєсівського навчання з підкріпленням (Bayesian reinforcement learning) описується в [5]. Пропонується використання ансамблю нейронних мереж із рандомізованими апіорними функціями (randomized prior functions, RPF). Для приблизної оцінки невизначеності використовується коефіцієнт варіації у значеннях Q-функцій, що оцінюються мережами ансамблю. На основі цього введено критерій, що дозволяє визначити ступінь впевненості агента для прийняття певного рішення. В результаті експериментальних порівнянь зі стандартним DQN виявлено, що розроблений метод дозволяє приймати рішення в ситуаціях із високою невизначеністю, навіть якщо вони суттєво відрізняються від навчальних даних. Показано, що розроблений метод добре підходить для вибору безпечних дій в ситуаціях із невизначеністю при проходженні транспортним засобом перехрестя (рис. 2). Уникнення зіткнень було одним із важливих завдань при розробці цього методу, і алгоритм обирає апіорно безпечні дії у випадку надто великої невизначеності, що виділяє його серед інших методів і свідчить про перспективність його використання у інших критичних середовищах.

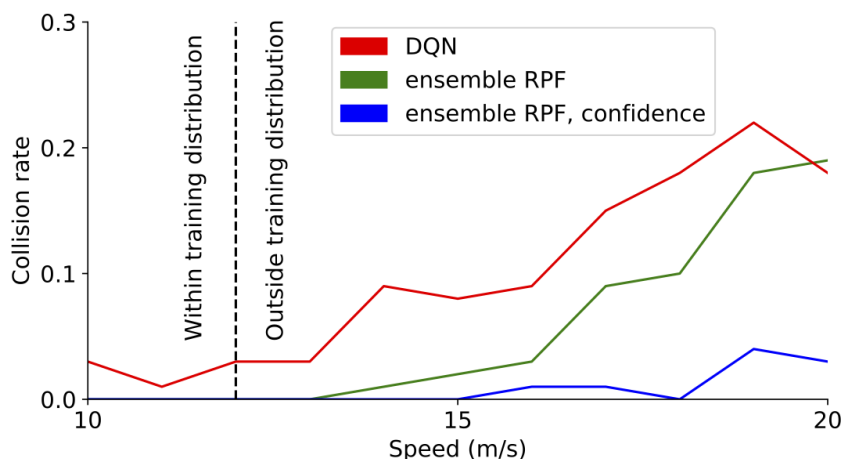


Рис. 2. Частота зіткнень при використанні стандартного DQN, ансамблю мереж із PRF без критерію впевненості та із ним  
Джерело: [5]

Однією із проблем навчання з підкріпленням є дилема exploration-exploitation, автори [7] пропонують новий алгоритм навчання heuristic decaying state entropy (HDSE), що пришвидшує навчання агента RL. Для вирішення проблем невизначеності середовища використано модель future integrated risk assessment model. З поєднанням цієї моделі та long short-term memory (LSTM) для передбачення траєкторій руху транспортних засобів, отримана модель показує кращі результати щодо уникнення зіткнень та покращення ефективності трафіку в модельованих середовищах із великою та малою щільністю трафіку.

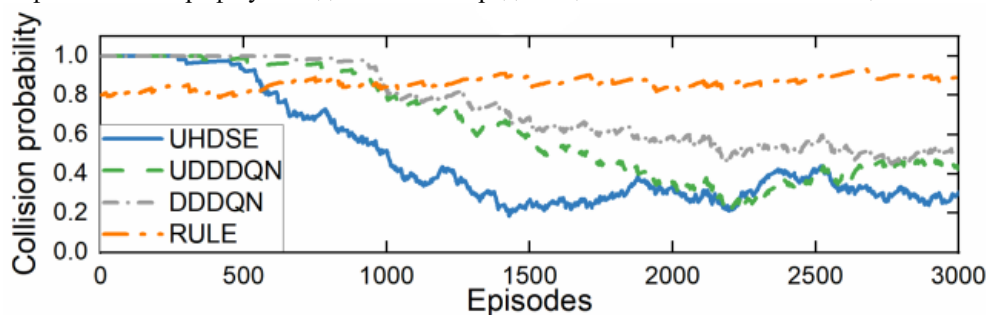


Рис. 3. Частота зіткнень впродовж навчання при використанні агентів RL із новим алгоритмом навчання та future integrated risk assessment model (UHDSE) та зі звичайним навчанням (UDDDQN, DDDQN)  
Джерело: [7]

В статті [8] розглядаються задачі прийняття рішень при проходженні транспортним засобом перехрестя. Автори роблять акцент на перевагах POMDP для задач в реальному світі із притаманними йому факторами невизначеності та непередбачуваності, однак також вони звертають увагу на те, що вирішення задач, сформульованих в POMDP, за допомогою RL часто вимагає зберігання великої кількості спостережень, та на те, що така система є обчислювально неефективною при використанні в неперервному просторі дій. Для вирішення цих проблем, пропонується моделювання задачі за допомогою MDP та застосування ієрархічних варіантів (hierarchical options) для навчання з підкріпленням. Такий підхід (Hierarchical Options MDP, HOMDP) дозволяє зберігати за допомогою LSTM лише поточне спостереження. В результаті порівняння із звичайним POMDP виявлено, що розроблений метод показує кращі результати в навігації та уникненні зіткнень (таб. 1).

В сфері безпеки комп'ютерних систем на сьогоднішній день розрізняють одноступеневі та багаступеневі кібератаки. До одноступеневих атак належать, наприклад, SQL-ін'єкції (SQL injection) чи міжсайтовий скриптинг (Cross-Site Scripting, XSS). Одноступеневі атаки на сьогодні мають вже досить довгу історію супротиву та протидії, та архітектура комп'ютерних систем часто створюється з урахуванням їх загрози: у багатьох випадках одноступенева атака може загрожувати одному окремому компоненту у системі, часто периферійному, і пошкодження такої загрози можуть бути виправлені досить швидко. На відміну від них, багаступеневі атаки поєднують у собі одноступеневі у послідовність, що дозволяє зловмисникам заволодіти окремим компонентом у системі і одразу ж продовжити дії, вже виходячи із можливостей скомпрометованого елемента. Такий підхід використовується щоб проникнути до ключових компонентів системи. Для окремих ступенів зловмисники використовують окремі вразливості поточних версій програмного забезпечення, протоколів безпеки тощо. Традиційний підхід - виявлення вразливостей та внесення відповідних виправлень у систему є процесом тривалим та виконується він загалом вручну. Автори [9] розглядають засоби адаптивної кібербезпеки (adaptive cyber defense, ACD) для вирішення цієї

проблеми. Важливим засобом дослідження взаємозалежностей вразливостей та шляху зараження системи є графічні моделі, проте вони носять детерміністичний характер представлення прогресу зловмисника і не підходять для вироблення ефективних шляхів протидії з боку ACD. В розробленому рішенні вводяться байєсівські графи нападу (Bayesian attack graphs, BAGs), що представляють вірогіднісну інформацію про можливе використання тих чи інших шляхів кібератаки. Завдяки цьому, стає можливим використання MDP для формулювання оптимальних стратегій як кібератаки, так і протидії їй, а також використання POMDP у випадках, коли система захисту має доступ лише до підмножини компонентів системи через вади засобів виявлення атаки чи відмову компонентів. Для загальної оцінки захищеності мережі та вибору оптимальних дій для протидії нападу використовується навчання з підкріпленням, а саме Q-навчання. Для вирішення проблеми невідомих вірогідностей переходів між станами у BAG використовується виборка Томпсона (Thompson sampling), що дозволяє оновлювати початкові довільно задані вірогідності із часом та балансує exploration-exploitation навчання з підкріпленням. Було проведено тестування розробленої системи для захисту мережі із 10 комп'ютерів, показано ефективність розробленого методу протидії у симуляціях, що базуються на реальних випадках багатоступеневих кібератак.

Таблиця 1

**Порівняння POMDP та HOMDP за метриками частоти успіху, зіткнень, незавершення дії та сукупної винагорода**

Задача	Метрика	POMDP	HOMDP
Пріоритетно	% успіху	97.1	98.3
	% зіткнень	1.7	1.7
	% незавершених	1.2	0.0
	Винагорода	621	873
Правий поворот	% успіху	99.5	99.8
	% зіткнень	0.5	0.2
	% незавершених	0.0	0.0
	Винагорода	892	903
Лівий поворот	% успіху	95.6	97.3
	% зіткнень	2.4	2.6
	% незавершених	2.0	0.1
	Винагорода	213	632

Джерело: таблиця складена на основі даних зі статті [8]

У статті [10] розглядається проблематика прийняття рішень у техніках захисту Moving Target Defense (MTD), що змінюють характеристики програмного забезпечення системи в реальному часі для того, щоб зловмиснику було складніше скористатися можливими вразливостями або для передбачення поведінки зловмисника. Зокрема, зазначається що розробка та впровадження систем із MTD представляють труднощі, що виходять із невизначеності під час роботи систем, а також що існуючі реалізації не враховують факторів невизначеності у параметрах самої моделі та стану системи та обмежені в можливостях до адаптації. Пропонується вирішення проблематики застосування MTD завдяки використанню POMDP та байєсівського навчання (Bayesian learning). Впроваджено байєс-адаптивний POMDP (Bayes-Adaptive POMDP, BA-POMDP), що дозволяє вирішувати проблему невизначеності параметрів моделі. Розроблена реалізація здатна адаптуватись до невизначеності, що впливає із ходу кібератаки та стану компонентів системи, що захищається, та краще координує необхідні засоби та стратегії захисту. Початкові експерименти із застосування розробленої системи показують перспективність цього напрямку, автори вважають, що система може добре масштабуватись та здатна протистояти багатьом векторам атак одночасно, але також зазначається що для достовірних висновків щодо ефективності роботи системи необхідне подальше, більш глибоке експериментальне тестування.

Однією із задач кібербезпеки є введення зловмисника в оману (Cyber Deception, CD) стосовно стану системи, що атакується, для вивчення стратегій, тактик, можливостей та намірів нападника. Часто рішення про першочергові цілі та методи кібератак приймаються зловмисником виходячи із попередніх знань про систему та її інфраструктуру. При застосуванні засобів CD захисник може використати цю особливість для того, щоб змінити вектори атаки на неефективні та затратні за часом, «підкидаючи» нападнику хибні шляхи для атаки. Успішність застосування CD визначається тим, наскільки багато таких шляхів нападник буде змушений виокремлювати та відрізняти від справжніх вразливостей системи. Автори [11] приводять метод такого кібер-захисту для інтернету речей (Internet of Things, IoT). Через динамічну природу мережевих зв'язків саме цей клас комп'ютерних систем може використати CD повною мірою. Розроблена реалізація використовує алгоритм частково спостережуваного планування Монте-Карло (partially observable Monte Carlo planning, POMCP) для представлення невизначеності векторів атаки. Це один із евристичних методів пошуку по дереву для приблизного розв'язання задач, сформульованих в POMDP, що надає задовільні рішення при великому масштабі мережевої структури системи, що захищається. Дії, які виконує система у відповідь на атаки - це створення хибних вузлів системи із заздалегідь заданими характеристиками затримки

та пропускної здатності з'єднання. Розроблена реалізація дозволяє перешкоджати кібератакам, значно зменшуючи їх вплив на систему, та разом із тим надає можливості для вивчення стратегій та засобів зловмисника.

В статті [12] досліджується перспективність застосування підвиду навчання з підкріпленням, табличного Q-навчання (Tabular Q-learning), для створення повністю автономних засобів захисту від кібератак. Цей підвид використовує Q-матрицю, таблицю відповістей всіх очікуваних винагород та пар «стан-дія», що дозволяє швидко оновлювати значення Q-функції без необхідності в додаткових обчисленнях, проте вимагає додаткового об'єму пам'яті для зберігання. Розглядаючи дилему exploration-exploitation, в даному випадку алгоритм віддає перевагу діям, що надають пріоритет пошуку в глибину (exploitation). На відміну від наведених авторами існуючих засобів, що використовують POMDP та враховують невизначеність інформації, цей підхід використовує звичайний MDP. За результатами експериментального порівняння із існуючими реалізаціями систем кібер-захисту, було зроблено висновок про те що розроблене рішення не є оптимальним через відсутність гнучкості: так, окремі конфігурації створеної системи показували результати кращі за деякі аналоги, проте жоден із варіантів конфігурацій не перевершував всі аналоги. Для малих задач POMDP можна отримати приблизне рішення із використанням Tabular Q-Learning, але ефективність цих рішень швидко падає при зростанні складності. Тим не менше, деякі варіанти Tabular Q-Learning можливо використовувати для деяких задач кібер-захисту та специфічних задач прийняття рішень в середовищах із неповною чи невизначеною інформацією в інших галузях.

Автори [13] описують повністю автономний механізм кібер-захисту від динамічних I-DDoS атак великої складності під назвою Horde. Він має розподілену багатоагентну архітектуру для захисту критичних ланок мережі без втручання людини. Цей підхід надає можливості для створення комбінацій обмеження трафіку та перенаправлення в мережі, виходячи із динамічної поведінки зловмисника та спостереження стану мережі. Horde може обчислювати оптимальні послідовності дій для критичного зниження впливу атаки на систему, виявлення стратегій та засобів нападника. Для покращення роботи методів навчання з підкріпленням Horde відкидає неважливі в даний момент секції мережі, тим самим покращуючи ефективність пошуку в ширину (exploration). Для обробки невизначеної інформації використовуються моделі POMDP, пропонується новітній підхід BRITE Loop, що керує діяльністю захисних агентів так, що вони оновлюють стан власної моделі виходячи зі своїх поточних спостережень та отримують оцінку своїх дій щодо захисту системи. Таким чином, в мережі одночасно знаходиться множина автономних захисних агентів, кожен із яких вивчає власні уявлення про стан системи та вектори атаки та вносить свій вклад в формування загальної стратегії протидії. Розроблена система захисту найкраще придатна до використання у комп'ютерних системах великого масштабу. Недоліком даної системи є слабкість захисту проти атак що використовують досі невідомі вразливості.

Окрім автономних транспортних засобів та кібербезпеки, задачі прийняття рішень у середовищах із неповною та невизначеною інформацією розглядаються та вирішуються і в інших областях застосування (таб. 2).

Таблиця 2

Дослідження прийняття рішень в інших сферах застосування

Сфера застосування	Методи прийняття рішень
Автопілот морських дронів	Deep Deterministic Policy Gradients (DDPG)/DRL ([14]), Proximal Policy Optimization (PPO)/DRL ([15]), POMDP/PPO ([16])
Автопілот літальних апаратів	DDQN ([17]), DRL/MDP ([18]), POMDP/RL ([19])
Організація обслуговування та підтримки систем	POMDP/HMM ([20]), POMDP ([21])
Біржові торги	DRL/Gated Recurrent Unit (GRU) ([22]), Hidden Markov Model (HMM) ([23])

Джерело: розроблено авторами

Отже, POMDP являє собою досить гнучкий та універсальний засіб для опису задач прийняття рішень при неповній чи невизначеній інформації у різноманітних сферах.

Автори [24] пропонують новий швидкий та точний підхід до розв'язання задач POMDP для систем підтримки прийняття рішень, названий авторами Permutable POMDP. Цей підхід вимагає значно менше обчислень для досягнення тієї ж точності та якості вирішення. Тестування показало прискорення обчислень на декілька порядків (рис. 4), що потенційно дозволить розглядати, моделювати та вирішувати задачі із просторами дій значно більшими, ніж ті, що є вирішуваними зі звичайним POMDP. Цей метод вимагає деяких особливих властивостей у структурі POMDP, проте багато задач підтримки прийняття рішень, як вважають автори, можливо апроксимувати як такі, що мають саме таку структуру.



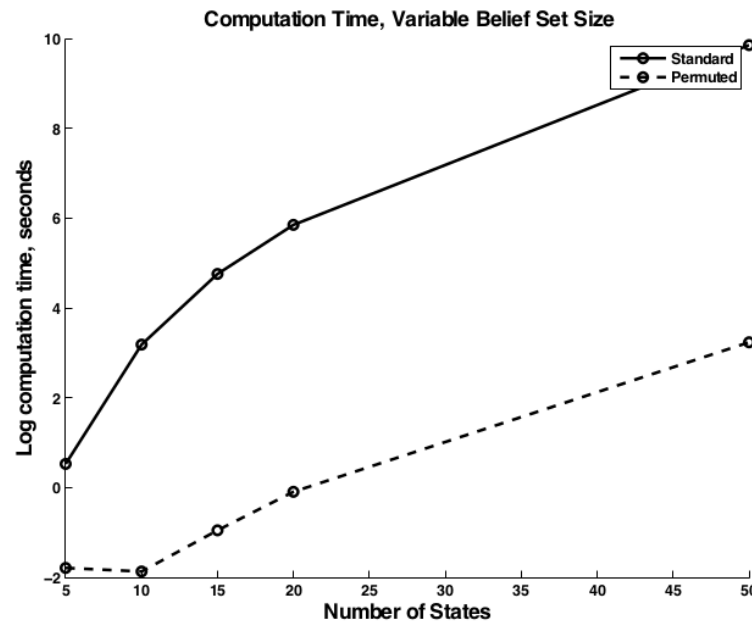


Рис. 4. Залежність логарифму часу обчислень від розміру простору станів для стандартного POMDP та для Permuted POMDP  
Джерело: [24]

У статті [25] пропонується новий метод навчання з підкріпленням для розв'язання задач, сформульованих у POMDP, що використовує лише потік винагород та неповні або зашумлені спостереження. Цей підхід, названий авторами deep variational reinforcement learning (DVRL), відрізняється від існуючих методів глибокого навчання з підкріпленням, що використовують рекуррентну нейронну мережу (recurrent neural network, RNN), наприклад, LSTM, для зберігання історії спостережень - таких як deep recurrent Q-network (DRQN). Результати порівняння із ними на деяких комп'ютерних іграх (Mountain Nike, що визначається як неперервна задача контролю із зашумленими спостереженнями, та модифікація ігор Atari із неповною інформацією) показують, що розроблений метод перевершує існуючі методи, що базуються на RNN.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Огляд досліджень із розробки, розвитку та застосування методів та підходів до прийняття рішень в середовищах із неповною та невизначеною інформацією показав, які методи є ефективними та гнучкими, в якому напрямку варто вести дослідження, та яка проблематика цієї теми.

Навчання з підкріпленням є найпоширенішим підходом до знаходження наближеного розв'язання задачі, сформульованої у POMDP. Розповсюджені методи глибокого навчання, проте звичайний DQN та DDQN поступаються більш новим модифікаціям, що швидше навчаються та показують кращі результати за метриками ефективності та точності. В деяких випадках, коли окрім прийняття рішень у реальному часі доцільне також і планування дій у майбутньому, поєднання навчання з підкріпленням із модифікаціями методів пошуку по дереву демонструє кращі результати. Планування із використанням модифікації MCTS є досить гнучким та добре відповідає вимогам швидкодії. Модифікації POMDP відкривають можливості для розв'язання деяких складних задач, що неможливо зі стандартним POMDP, чи розв'язання тих самих задач із набагато меншою кількістю обчислень, та приймати рішення без зберігання довгої історії минулих спостережень.

Обраний напрямок є перспективним для подальших досліджень, що впливає із активного розвитку автономних транспортних засобів, дронів, розвитку комп'ютерних систем та необхідності у їх захисті. Покращення та нові розробки різних сфер можуть бути застосовані у інших, наприклад, використання планування за допомогою MCTS може застосовуватись для кращого передбачення дій зловмисника при протидії кібератаці. Також варто розглянути, які задачі, що формуються у POMDP, можливо апроксимувати для застосування Permuted POMDP, що має набагато кращу ефективність.

#### Література

1. Schwarting W., Alonso-Mora J., Rus D. Planning and Decision-Making for Autonomous Vehicles, Annual Review of Control, Robotics, and Autonomous Systems, Vol. 1 (2018), P. 187-210, 2018. [DOI:10.1146/annurev-control-060117-105157](https://doi.org/10.1146/annurev-control-060117-105157)

2. Jajodia S., Cybenko G., Liu P., Wang C., Wellman M. Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Survey. Switzerland: Springer Nature, 2019. 262 с. URL: <https://link.springer.com/book/10.1007/978-3-030-30719-6> (дата звернення: 18.08.2024).
3. Bowyer C. M. Approximation Methods for Partially Observed Markov Decision Processes (POMDPs), arXiv:2108.13965, 2021, 59 p. URL: <https://arxiv.org/abs/2108.13965> DOI:10.48550/arXiv.2108.13965
4. Hoel C.-J., Driggs-Campbell K., Wolff L., Kochenderfer M. J. Combining Planning and Deep Reinforcement Learning in Tactical Decision Making for Autonomous Driving, IEEE Transactions on Intelligent Vehicles, Vol. 5 no. 2, P. 294-305, 2020. DOI:10.1109/TIV.2019.2955905
5. Hoel C.-J., Tram T., Sjöberg J. Reinforcement Learning with Uncertainty Estimation for Tactical Decision-Making in Intersections, arXiv:2006.09786, 2020, 7 p. URL: <https://arxiv.org/abs/2006.09786> DOI:10.48550/arXiv.2006.09786
6. Ugur Yavas M., Kemal Ure N., Kumbasar T. A New Approach for Tactical Decision Making in Lane Changing: Sample Efficient Deep Q Learning with a Safety Feedback Reward, arXiv:2009.11905, 2020, 7 p. URL: <https://arxiv.org/abs/2009.11905> DOI:10.48550/arXiv.2009.11905
7. Deng H., Zhao Y., Wang Q., Nguyen A.-T. Deep Reinforcement Learning Based Decision-Making Strategy of Autonomous Vehicle in Highway Uncertain Driving Environments, Automotive Innovation, Vol. 6, P. 438-452, 2023. DOI:10.1007/s42154-023-00231-6
8. Qiao Zh., Mülling K., Dolan J. M., Palanisamy P. POMDP and Hierarchical Options MDP with Continuous Actions for Autonomous Driving at Intersections. In: 2018 IEEE International Conference on Intelligent Transportation Systems (ITSC). IEEE, Maui, USA, p. 2377-2382. DOI:10.1109/ITSC.2018.8569400
9. Hu Zh., Zhu M., Liu P. Adaptive Cyber Defense Against Multi-Stage Attacks Using Learning-Based POMDP, ACM Transactions on Privacy and Security (TOPS), Vol. 24 no. 1, P. 1-25, 2020. DOI:10.1145/3418897
10. Liu R., Tahvildari L. Using POMDP-based Approach to Address Uncertainty-Aware Adaptation for Self-Protecting Software, arXiv:2308.02134, 2023, 7 p. URL: <https://arxiv.org/abs/2308.02134> DOI:10.48550/arXiv.2308.02134
11. Al Amin M. A. R., Shetty S., Njilla L. L., Tosh D. K., Kamhoua C. A. Dynamic Cyber Deception Using Partially Observable Monte-Carlo Planning Framework. Modeling and Design of Secure Internet of Things / Edited by Charles A. Kamhoua, Laurent L. Njilla, Alexander Kott and Sachin Shetty. John Wiley & Sons, Inc, 2020. p. 331-355. DOI:10.1002/9781119593386.ch14
12. Applebaum A., Dennler C., Dwyer P., Moskowicz M., Nguyen H., Nichols N., Park N., Rachwalski P., Rau F., Webster A., Wolk M. Bridging Automated to Autonomous Cyber Defense: Foundational Analysis of Tabular Q-Learning. In: AISec'22: Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security. ACM, New York, p. 149-159. DOI:10.1145/3560830.3563732
13. Dutta A. Autonomous Cyber Defense: Formal Models and Applications: дисертація. University of North Carolina, Charlotte, 2021. 178 с.
14. Cui Zh., Guan W., Zhang X. Collision avoidance decision-making strategy for multiple USVs based on Deep Reinforcement Learning algorithm, Ocean Engineering, Vol. 308, 2024. DOI:10.1016/j.oceaneng.2024.118323
15. Guan W., Luo W., Cui Zh. Intelligent decision-making system for multiple marine autonomous surface ships based on deep reinforcement learning, Robotics and Autonomous Systems, Vol. 172, 2024. DOI: 10.1016/j.robot.2023.104587
16. Zheng K., Zhang X., Wang Ch., Zhang M., Cui H. A partially observable multi-ship collision avoidance decision-making model based on deep reinforcement learning, Ocean & Coastal Management, Vol. 242, 2023. DOI:10.1016/j.ocecoaman.2023.106689
17. Han H., Cheng J., Lv M. Interpretable DRL-based Maneuver Decision of UCAV Dogfight, arXiv:2407.01571, 2024, 6 p. URL: <https://arxiv.org/abs/2407.01571> DOI:10.48550/arXiv.2407.01571
18. Alvarez L. E., Brittain M. W., Young S. D. Tradeoffs When Considering Deep Reinforcement Learning for Contingency Management in Advanced Air Mobility, arXiv:2407.00197, 2024, 16 p. URL: <https://arxiv.org/abs/2407.00197> DOI:10.48550/arXiv.2407.00197

19. Chen Y., Dong Q., Shang X., Wu Zh., Wang J. Multi-UAV Autonomous Path Planning in Reconnaissance Missions Considering Incomplete Information: A Reinforcement Learning Method, *Drones*, Vol. 7(1) no. 10, 2023. DOI:10.3390/drones7010010
20. Arcieri G., Hoelzl C., Schwery O., Straub D., Papakonstantinou K. G., Chatzi E. Bridging POMDPs and Bayesian decision making for robust maintenance planning under model uncertainty: An application to railway systems, *Reliability Engineering & System Safety*, Vol. 239, 2023. DOI:10.1016/j.ress.2023.109496
21. Deep A., Zhou Sh., Veeramani D., Chen Y. Partially observable Markov decision process-based optimal maintenance planning with time-dependent observations, *European Journal of Operational Research*, Vol. 311 no. 2, P. 533-544, 2023. DOI:10.1016/j.ejor.2023.05.022
22. Ansari Y., Yasmin S., Naz Sh., Zaffar H., Ali Z., Moon J., Rho S. A Deep Reinforcement Learning-Based Decision Support System for Automated Stock Market Trading, *IEEE Access*, Vol. 10, P. 127469-127501, 2022. DOI:10.1109/ACCESS.2022.3226629
23. Zhang L., Li Zh., Xu Y., Li Y. Multi-period mean variance portfolio selection under incomplete information, *Applied Stochastic Models in Business and Industry*, Vol. 32 no. 6, P. 753-774, 2016. DOI:10.1002/asmb.2191
24. Doshi F., Roy N. The Permutable POMDP: Fast Solutions to POMDPs for Preference Elicitation. In: 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008). Padgham, Parkes, Müller and Parsons (eds.), Estoril, Portugal, p. 493-500. URL: <https://groups.csail.mit.edu/rrg/papers/aamas08-fd.pdf> (дата звернення: 18.08.2024)
25. Igl M., Zintgraf L., Le T. A., Wood F., Whiteson Sh. Deep Variational Reinforcement Learning for POMDPs. In: Proceedings of the 35th International Conference on Machine Learning. PMLR 80, Stockholm, p. 2117-2126. URL: <http://proceedings.mlr.press/v80/igl18a/igl18a.pdf> (дата звернення: 18.08.2024)

#### References

1. W. Schwarting, J. Alonso-Mora, D. Rus, "Planning and Decision-Making for Autonomous Vehicles". *Annual Review of Control, Robotics, and Autonomous Systems*, Vol. 1 (2018), P. 187-210, 2018. DOI:10.1146/annurev-control-060117-105157
2. S. Jajodia, G. Cybenko, P. Liu, C. Wang, M. Wellman, *Adversarial and Uncertain Reasoning for Adaptive Cyber Defense: Survey*. Switzerland: Springer Nature, 2019, 262 p. URL: <https://link.springer.com/book/10.1007/978-3-030-30719-6> (Last accessed: 18.08.2024).
3. C. M. Boweyer, "Approximation Methods for Partially Observed Markov Decision Processes (POMDPs)", arXiv:2108.13965, 2021, 59 p. URL: <https://arxiv.org/abs/2108.13965> DOI:10.48550/arXiv.2108.13965
4. C.-J. Hoel, K. Driggs-Campbell, L. Wolff, M. J. Kochenderfer, "Combining Planning and Deep Reinforcement Learning in Tactical Decision Making for Autonomous Driving". *IEEE Transactions on Intelligent Vehicles*, Vol. 5 no. 2, P. 294-305, 2020. DOI:10.1109/TIV.2019.2955905
5. C.-J. Hoel, T. Tram, J. Sjöberg, "Reinforcement Learning with Uncertainty Estimation for Tactical Decision-Making in Intersections", arXiv:2006.09786, 2020, 7 p. URL: <https://arxiv.org/abs/2006.09786> DOI:10.48550/arXiv.2006.09786
6. M. Ugur Yavas, N. Kemal Ure, T. Kumbasar, "A New Approach for Tactical Decision Making in Lane Changing: Sample Efficient Deep Q Learning with a Safety Feedback Reward", arXiv:2009.11905, 2020, 7 p. URL: <https://arxiv.org/abs/2009.11905> DOI:10.48550/arXiv.2009.11905
7. H. Deng, Y. Zhao, Q. Wang, A.-T. Nguyen, "Deep Reinforcement Learning Based Decision-Making Strategy of Autonomous Vehicle in Highway Uncertain Driving Environments". *Automotive Innovation*, Vol. 6, P. 438-452, 2023. DOI:10.1007/s42154-023-00231-6
8. Zh. Qiao, K. Mülling, M. Dolan, P. Palanisamy, "POMDP and Hierarchical Options MDP with Continuous Actions for Autonomous Driving at Intersections", in 2018 *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, Maui, 2018, pp. 2377-2382. DOI:10.1109/ITSC.2018.8569400
9. Zh. Hu, M. Zhu, P. Liu, "Adaptive Cyber Defense Against Multi-Stage Attacks Using Learning-Based POMDP", *ACM Transactions on Privacy and Security (TOPS)*, Vol. 24 no. 1, P. 1-25, 2020. DOI:10.1145/3418897
10. R. Liu, L. Tahvildari, "Using POMDP-based Approach to Address Uncertainty-Aware Adaptation for Self-Protecting Software", arXiv:2308.02134, 2023, 7 p. URL: <https://arxiv.org/abs/2308.02134> DOI:10.48550/arXiv.2308.02134
11. M. A. R. Al Amin, S. Shetty, L. L. Njilla, D. K. Tosh, C. A. Kamhoua, "Dynamic Cyber Deception Using Partially Observable Monte-Carlo Planning Framework", in *Modeling and Design of Secure Internet of Things*. Charles A. Kamhoua, Laurent L. Njilla, Alexander Kott and Sachin Shetty, Eds. John Wiley & Sons, Inc, 2020. pp. 331-355. DOI:10.1002/9781119593386.ch14
12. A. Applebaum, C. Dennler, P. Dwyer, M. Moskowitiz, H. Nguyen, N. Nichols, N. Park, P. Rachwalski, F. Rau, A. Webster, M. Wolk, "Bridging Automated to Autonomous Cyber Defense: Foundational Analysis of Tabular Q-Learning", in *AISeC'22: Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security*, New York, pp. 149-149. DOI:10.1145/3560830.3563732
13. A. Dutta, "Autonomous Cyber Defense: Formal Models and Applications", PhD thesis, University of North Carolina, Charlotte, 2021.
14. Zh. Cui, W. Guan, X. Zhang, "Collision avoidance decision-making strategy for multiple USVs based on Deep Reinforcement Learning algorithm". *Ocean Engineering*, Vol. 308, 2024. DOI:10.1016/j.oceaneng.2024.118323
15. W. Guan, W. Luo, Zh. Cui, "Intelligent decision-making system for multiple marine autonomous surface ships based on deep reinforcement learning". *Robotics and Autonomous Systems*, Vol. 172, 2024. DOI: 10.1016/j.robot.2023.104587
16. K. Zheng, X. Zhang, Ch. Wang, M. Zhang, H. Cui, "A partially observable multi-ship collision avoidance decision-making model based on deep reinforcement learning". *Ocean & Coastal Management*, Vol. 242, 2023. DOI:10.1016/j.ocecoaman.2023.106689

17. H. Han, J. Cheng, M. Lv, "Interpretable DRL-based Maneuver Decision of UCAV Dogfight", arXiv:2407.01571, 2024, 6 p. URL: <https://arxiv.org/abs/2407.01571> DOI:10.48550/arXiv.2407.01571
18. L. E. Alvarez, M. W. Brittain, S. D. Young, "Tradeoffs When Considering Deep Reinforcement Learning for Contingency Management in Advanced Air Mobility", arXiv:2407.00197, 2024, 16 p. URL: <https://arxiv.org/abs/2407.00197> DOI:10.48550/arXiv.2407.00197
19. Y. Chen, Q. Dong, X. Shang, Zh. Wu, J. Wang, "Multi-UAV Autonomous Path Planning in Reconnaissance Missions Considering Incomplete Information: A Reinforcement Learning Method". *Drones*, Vol. 7(1) no. 10, 2023. DOI:10.3390/drones7010010
20. G. Arcieri, C. Hoelzl, O. Schwery, D. Straub, K. G. Papakonstantinou, E. Chatzi, "Bridging POMDPs and Bayesian decision making for robust maintenance planning under model uncertainty: An application to railway systems". *Reliability Engineering & System Safety*, Vol. 239, 2023. DOI:10.1016/j.ress.2023.109496
21. A. Deep, Sh. Zhou, D. Veeramani, Y. Chen, "Partially observable Markov decision process-based optimal maintenance planning with time-dependent observations". *European Journal of Operational Research*, Vol. 311 no. 2, P. 533-544, 2023. DOI:10.1016/j.ejor.2023.05.022
22. Y. Ansari, S. Yasmin, Sh. Naz, H. Zaffar, Z. Ali, J. Moon, S. Rho, "A Deep Reinforcement Learning-Based Decision Support System for Automated Stock Market Trading". *IEEE Access*, Vol. 10, P. 127469-127501, 2022. DOI:10.1109/ACCESS.2022.3226629
23. L. Zhang, Zh. Li, Y. Xu, Y. Li, "Multi-period mean variance portfolio selection under incomplete information". *Applied Stochastic Models in Business and Industry*, Vol. 32 no. 6, P. 753-774, 2016. DOI:10.1002/asmb.2191
24. F. Doshi, N. Roy, "The Permutable POMDP: Fast Solutions to POMDPs for Preference Elicitation", in *7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Estoril, Portugal, 2008, pp. 493-500. URL: <https://groups.csail.mit.edu/rg/papers/aamas08-fd.pdf> (Last accessed: 18.08.2024)
25. M. Igl, L. Zintgraf, T. A. Le, F. Wood, Sh. Whiteson, "Deep Variational Reinforcement Learning for POMDPs", in *Proceedings of the 35th International Conference on Machine Learning*, Stockholm, 2018, pp. 2117-2126. URL: <http://proceedings.mlr.press/v80/igl18a/igl18a.pdf> (Last accessed: 18.08.2024)

<https://doi.org/10.31891/2219-9365-2024-79-3>

УДК 543.421

ЗАЩЕПКИНА Наталія

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

<https://orcid.org/0000-0001-9397-6632>

e-mail: [nanic1604@gmail.com](mailto:nanic1604@gmail.com)

МЕЛЬНИЧЕНКО Дмитро

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

<https://orcid.org/0009-0008-7056-123X>

e-mail: [melnik.dimoss@gmail.com](mailto:melnik.dimoss@gmail.com)

## АВТОНОМНА СПЕКТРОФОТОМЕТРИЧНА СИСТЕМА ДЛЯ АНАЛІЗУ НІТРАТІВ І НІТРИТІВ У ВОДІ

У статті представлено новий підхід до визначення концентрацій нітратів і нітритів у воді на основі спектрофотометричного аналізу без використання хімічних реагентів. Запропонована методика включає використання фотометричної системи з одним фотодіодом та чотирма світлодіодами на довжинах хвиль 260 нм, 310 нм, 365 нм і 405 нм. Для підвищення точності та зменшення впливу зовнішніх факторів застосовується двокаскадний підсилювач, а фотодіод працює у фотопровідному режимі, підключений до трансимпедансного підсилювача. Додатково, для усунення впливу домішок на результати аналізу використовується вугільний фільтр. Методика базується на законі Бугера-Ламберта-Бера, який описує ослаблення інтенсивності світла при проходженні через поглинаюче середовище. Для підвищення чутливості вимірювань використовується кювета з довжиною оптичного шляху 100 мм.

Проведення подальших тестувань необхідне для визначення точності та селективності системи. У перспективі, розроблена установка має потенціал для автономного моніторингу якості води в умовах обмежених ресурсів. Як напрямок подальших досліджень пропонується вивчення можливості використання фотодіода на довжині хвилі 210 нм для покращення визначення нітратів, що може підвищити чутливість і точність методу.

Ключові слова: спектрофотометрія, нітрати, нітритами, водні зразки, безреагентний аналіз, автономна система, фотометрія, поглинання світла, закон Бугера-Ламберта-Бера, моніторинг якості води, вугільний фільтр, оптичний шлях, точність вимірювань, довжина хвилі, аналітична хімія.

ZASHCHEPKINA Nataliia, MELNYCHENKO Dmytro

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

## AUTONOMOUS SPECTROPHOTOMETRIC SYSTEM FOR NITRATE AND NITRITE ANALYSIS IN WATER

This article presents a novel approach to determining the concentrations of nitrates and nitrites in water based on spectrophotometric analysis without the use of chemical reagents. The proposed method involves using a photometric system with a single photodiode and four LEDs at wavelengths of 260 nm, 310 nm, 365 nm, and 405 nm. To enhance accuracy and reduce the influence of external factors, a dual-stage amplifier is used, with the photodiode operating in photoconductive mode and connected to a transimpedance amplifier. Additionally, a carbon filter is employed to eliminate the effects of impurities on the analysis results. The methodology is based on the Beer-Lambert law, which describes the attenuation of light intensity as it passes through an absorbing medium. To improve measurement sensitivity, a cuvette with a 100 mm optical path length is used.

Further testing is necessary to determine the accuracy and selectivity of the system. The developed system has the potential for autonomous water quality monitoring in resource-limited environments. As a future research direction, the possibility of using a photodiode at a wavelength of 210 nm to improve nitrate detection, which could enhance the sensitivity and accuracy of the method, is proposed.

Keywords: Spectrophotometry, Nitrates, Nitrites, Water samples, Reagent-free analysis, Autonomous system, Photometry, Light absorption, Beer-Lambert law, Water quality monitoring, Carbon filter, Optical path, Measurement accuracy, Wavelength, Analytical chemistry.

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У сучасному світі проблема забруднення водою нітратами та нітритами викликає значну тривогу через їхні потенційні шкідливі впливи на здоров'я людини та навколишнє середовище. З огляду на це, розробка та удосконалення методик аналізу забруднювачів води стає надзвичайно важливою для забезпечення безпеки питної води та охорони водних ресурсів [1,2].

У статті запропоновано інноваційний підхід до визначення концентрацій нітратів і нітритів у воді, заснований на вдосконаленому спектрофотометричному аналізі без використання хімічних реагентів.

Методика передбачає застосування фотометричної системи, що складається з одного фотодіода та чотирьох світлодіодів на довжинах хвиль 260 нм, 310 нм, 365 нм і 405 нм.

Використання єдиного фотодіода з двокаскадним підсиленням забезпечує високу точність вимірювань, зменшуючи вплив зовнішніх факторів. Фотодіод працює в фотопровідному режимі, а

трансмпедансний підсилювач перетворює фотострум у напругу, яка потім додатково підсилюється перед передачею на АЦП.(аналогово-цифровий перетворювач)

Закон Бугера-Ламберта-Бера використовується для опису ослаблення інтенсивності світла під час проходження через поглинаюче середовище, а довгий оптичний шлях (100 мм) у кюветі дозволяє значно підвищити чутливість системи до концентрацій нітратів і нітритів. Для додаткового очищення проби від домішок використовується вугільний фільтр.

Запропонована методика має значний потенціал для автономного використання в системах моніторингу якості води, зокрема в умовах обмежених ресурсів. Пропонується як перспективний напрямок подальших досліджень вивчення можливостей використання світлодіода на довжині хвилі 210 нм для підвищення точності визначення нітратів у воді.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Спектрофотометрія є одним із найбільш поширених і ефективних методів аналізу хімічного складу водних проб, включаючи визначення концентрацій нітратів і нітритів. Цей метод базується на вимірюванні інтенсивності світла, яке поглинається або відбивається зразком при проходженні через нього пучка світла певної довжини хвилі. Поглинання світла певної довжини хвилі залежить від концентрації специфічних хімічних сполук, що дозволяє здійснювати кількісний аналіз.[4]

Застосування спектрофотометрії для визначення нітратів і нітритів зазвичай передбачає використання хімічних реагентів, які взаємодіють з цими сполуками і утворюють забарвлені комплекси.

Інтенсивність забарвлення пропорційна концентрації нітратів або нітритів у зразку. Одним з найбільш відомих прикладів є реакція нітритів з сульфаніламидами та нафтилетилендіаміном, яка утворює азосполуку, що поглинає світло при довжині хвилі близько 540 нм. Такий підхід забезпечує високу точність і чутливість, але він також має низку обмежень.[5]

Попри високу ефективність, традиційні спектрофотометричні методи аналізу мають певні недоліки. По-перше, використання хімічних реагентів додає додатковий етап підготовки проби, що збільшує час аналізу і вимагає наявності реагентів, які можуть бути дорогими або небезпечними. Це робить метод менш ефективним для автономного використання або застосування в польових умовах, де доступ до хімікатів обмежений.

По-друге, необхідність ретельного калібрування і підготовки проби підвищує ймовірність помилок, що може вплинути на точність аналізу. Наприклад, присутність інших хімічних речовин у воді може вплинути на реакцію з реагентами, що призведе до хибних результатів.

Крім того, традиційні методи часто вимагають складного і дорогого обладнання, що обмежує їх доступність для невеликих або віддалених лабораторій.

У зв'язку з цим виникає потреба у розробці нових підходів, які б забезпечували високу точність і чутливість, але не залежали від використання хімічних реагентів і складного обладнання.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою даного дослідження є розробка і впровадження методу для визначення нітратів і нітритів у воді на основі вдосконаленого спектрофотометричного аналізу, який не вимагає використання дорогіснних реагентів і може працювати автономно.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Запропонована методика передбачає використання вугільного фільтра для попереднього очищення води від домішок, а також фотометричної системи, що складається з одного приймача, підсилювача, АЦП і чотирьох фотодіодів на довжинах хвиль 260 нм, 310 нм, 365 нм і 405 нм. [3] Завданням є розробка алгоритму, який на основі поглинання світла на зазначених довжинах хвиль дозволить з високою точністю і селективністю визначити наявність і концентрацію нітратів і нітритів у водних пробах.

Запропонований метод має забезпечити високу швидкість, точність і економічність аналізу, що дозволить застосовувати його для моніторингу якості води у віддалених регіонах або в умовах обмеженого доступу до лабораторного обладнання.

Останні дослідження спрямовані на вдосконалення спектрофотометричних методів шляхом використання альтернативних підходів, які зменшують або повністю виключають необхідність використання хімічних реагентів. Одним із перспективних напрямів є застосування фотометричних систем, які використовують кілька довжин хвиль для аналізу поглинання світла у зразку, що дозволяє проводити більш точний і селективний аналіз без необхідності у реагентах.

Такі підходи передбачають використання специфічних довжин хвиль, які забезпечують максимальну чутливість до нітратів і нітритів, при цьому зменшуючи вплив інших компонентів води.

Наприклад, використання фотодіодів, налаштованих на 310 нм для нітритів і 365 нм для нітратів, може значно підвищити точність аналізу.

Таким чином, вдосконалення методів спектрофотометрії, спрямоване на зменшення залежності від реагентів і підвищення автономності систем, є актуальним і перспективним напрямом досліджень, що дозволяє розширити можливості аналізу якості води у різних умовах.

Запропоновано вдосконалену фотометричну систему, яка використовує кілька довжин хвиль для визначення концентрацій нітратів і нітритів у водних пробах без застосування хімічних реагентів.

Основні компоненти системи включають джерела випромінювання, фотодіод для прийому світлового сигналу, підсилювач, аналогово-цифровий перетворювач та мікроконтролер для обробки даних.

Фотометрична система використовує чотири світлодіоди, кожен з яких налаштований на певну довжину хвилі: 260 нм, 310 нм, 365 нм і 405 нм. Світлодіоди на 310 нм і 365 нм вибрані для специфічного виявлення нітритів та нітратів відповідно, тоді як світлодіоди на 260 нм та 405 нм виконують допоміжні функції, забезпечуючи додаткову інформацію для підвищення селективності і точності вимірювань.

Для забезпечення максимальної точності та зменшення впливу зовнішніх факторів усі чотири світлодіоди використовують один фотодіод, який працює в прямому режимі і підключений до двох каскадної схеми підсилення.

Перший каскад являє собою трансимпедансний підсилювач, який перетворює фотострум фотодіода в напругу. Другий каскад додатково підсилює цю напругу перед подачею сигналу на АЦП. Це дозволяє отримати більш точні вимірювання і мінімізувати можливі похибки, пов'язані з варіаціями в характеристиках окремих фотодіодів (рис.1).

Для зменшення впливу можливих домішок, які можуть спотворити результати спектрофотометричного аналізу, перед початком вимірювань вода проходить через вугільний фільтр. Цей фільтр ефективно видаляє великі частки, залишки хлору, метали та інші потенційно заважаючі домішки, що можуть вплинути на точність вимірювань.

Обробка результатів відбувається на мікроконтролері, який обробляє отримані цифрові сигнали, порівнюючи їх з калібрувальними значеннями, отриманими на етапі попередньої калібровки системи. Калібрування проводилося з використанням стандартних розчинів без нітратів і нітритів, що дозволяє врахувати базову поглинальну здатність води.

Після цього аналізується рівень поглинання на кожній довжині хвилі, що дозволяє визначити концентрації нітратів і нітритів у пробі.

У дослідженні застосовується закон Бугера-Ламберта-Бера, який є фундаментальним у спектрофотометричному аналізі. Цей закон описує ослаблення інтенсивності монохроматичного світла при його проходженні через поглинаюче середовище і формулюється наступним чином:

$$A = \epsilon \times c \times L \times A$$

де:  $A$  — оптична густина (абсорбція) середовища;  $\epsilon$  — молярний коефіцієнт поглинання, який залежить від довжини хвилі світла та природи речовини;  $c$  — концентрація поглинаючої речовини у розчині;  $L$  — довжина оптичного шляху (товщина шару поглинаючого середовища).

Згідно з цим законом, величина абсорбції прямо пропорційна концентрації поглинаючої речовини та довжині оптичного шляху, через який проходить світло. Тобто, чим більша концентрація нітратів або нітритів у зразку та чим більший оптичний шлях, тим вища буде оптична густина, що реєструється фотодіодом.

Для забезпечення більшої точності та чутливості вимірювань у цьому дослідженні використовується кювета з довжиною оптичного шляху 100 мм.

Використання довшої кювети дозволяє збільшити оптичний шлях, що підвищує чутливість системи до навіть невеликих концентрацій нітратів і нітритів у воді. Це важливо для точного визначення цих сполук у пробах з низькими концентраціями, що може бути критичним для контролю якості води.[6]

Таким чином, застосування закону Бугера-Ламберта-Бера разом із використанням довгої кювети дозволяє досягти високої точності та надійності аналізу, що є ключовим у задачах моніторингу якості води. Запропонований алгоритм аналізу даних включає кілька етапів:

1. **Калібрування світлодіодів:** Забезпечення однакової інтенсивності випромінювання всіх світлодіодів для отримання достовірних результатів.
2. **Зняття базової лінії:** Вимірювання поглинання чистої води для врахування природної поглинальної здатності без нітратів і нітритів.
3. **Аналіз поглинання:** Визначення концентрацій нітратів і нітритів на основі вимірюваного поглинання на кожній з чотирьох довжин хвиль.
4. **Корекція результатів:** Використання допоміжних довжин хвиль (260 нм і 405 нм) для корекції результатів і зменшення впливу можливих перешкод.

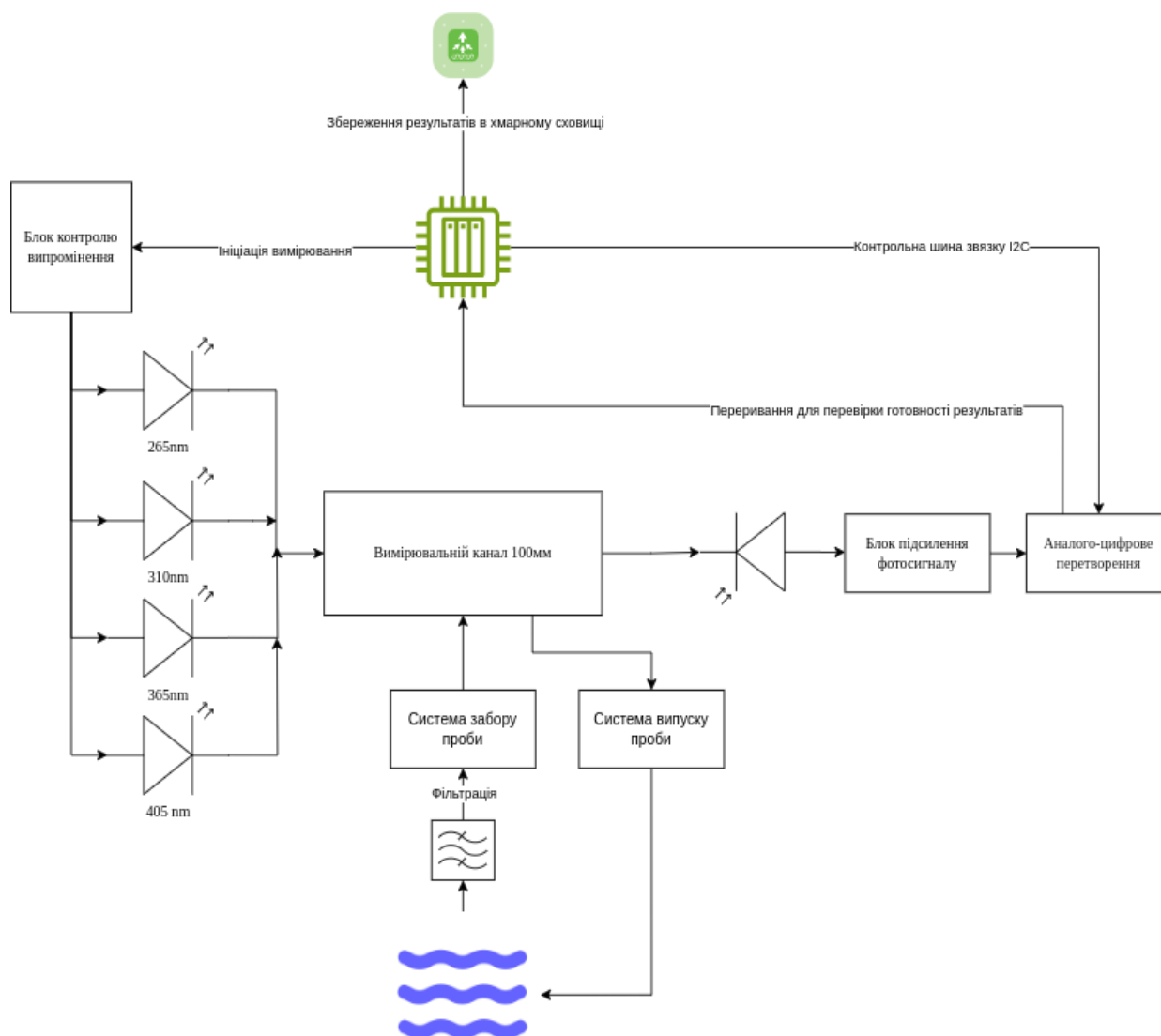


Рис. 1. Блок схема вимірювального приладу

Система розроблена для автономної роботи, що дозволяє проводити аналіз проб води без участі оператора. Це досягається завдяки використанню попередньо налаштованих алгоритмів, які автоматично визначають наявність і концентрацію нітратів та нітритів. Завдяки цьому система може бути використана у віддалених регіонах або в умовах, де відсутні лабораторії та кваліфікований персонал.

Ця методологія забезпечує високу точність, швидкість і надійність аналізу, дозволяючи проводити моніторинг якості води з мінімальними витратами і максимальною ефективністю.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Розроблена установка має потенціал для широкого використання у системах моніторингу якості води, особливо в умовах, де доступ до лабораторного обладнання обмежений.

Завдяки можливості автономної роботи та використанню недорогих компонентів, ця система може бути впроваджена для регулярного моніторингу нітратів і нітритів у колодязній, річковій та іншій воді.

Особливо актуальним є застосування системи в сільських та віддалених районах, де можливості для швидкого аналізу якості води обмежені. На даному етапі розробки запропонованої установки отримано проміжні результати, однак попередні теоретичні розрахунки та модельні експерименти свідчать про потенціал запропонованого підходу для використання у спектрофотометричному аналізі нітратів і нітритів у воді.

Для повної оцінки ефективності розробленої системи необхідно провести подальші тестування, спрямовані на визначення її точності та селективності при аналізі реальних водних зразків. Це включає експериментальне порівняння отриманих результатів із традиційними методами аналізу, а також перевірку



стійкості системи до можливих перешкод, таких як присутність інших хімічних речовин у воді. Крім того, важливо провести калібрування системи на різних концентраціях нітратів і нітритів для створення надійної бази даних для подальших вимірювань.

Як перспективний напрямок подальших досліджень пропонується вивчення можливості використання фотодіода на довжині хвилі 210 нм для визначення нітратів у воді.

Перший абсорбційний пік нітратів знаходиться саме на цій довжині хвилі, що може забезпечити вищу чутливість до їх концентрації.

Дослідження цього питання дозволить оцінити можливості розширення спектрального діапазону аналізу та потенційно підвищити точність і надійність вимірювань.

Проведення таких досліджень сприятиме подальшому вдосконаленню запропонованої системи та розширенню її функціональних можливостей.

Ці кроки дозволять остаточно оцінити ефективність запропонованої методології та визначити оптимальні умови її застосування для моніторингу якості води.

### Література

1. EPA. National primary drinking water regulations [Електронний ресурс]. – 2023. – Режим доступу: <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations>
2. World Health Organization. Nitrate and nitrite in drinking-water: Background document for development of WHO Guidelines for Drinking-water Quality [Електронний ресурс]. – 2011. – Режим доступу: <https://www.who.int/docs/default-source/wash-documents/wash-chemicals/nitrate-nitrite-background-document.pdf>
3. Защепкіна, Н., Мельниченко, Д. Розробка прототипу інформаційно-вимірювальної системи для експрес-вимірювання нітратів в рідинах // Вісник Хмельницького національного університету. – 2023. – Вип. 168. – DOI: <https://doi.org/10.31891/2219-9365-2023-75-6>
4. Hollas, J. M. Modern Spectroscopy [Електронний ресурс]. – 4th ed. – Chichester: John Wiley & Sons, 2004. – Розділ 2: "Introduction to Spectrophotometry".
5. Cogan, D. The application of spectrophotometry in the detection of nitrates and nitrites in environmental samples [Електронний ресурс]. – 2013. – Режим доступу: [https://doras.dcu.ie/19530/1/cest13\\_0400\\_DCogan.pdf](https://doras.dcu.ie/19530/1/cest13_0400_DCogan.pdf)
6. Harris, D. C. Quantitative Chemical Analysis. 9th ed. – New York: W. H. Freeman and Company, 2015. – P. 379-383.

### References

1. EPA. National primary drinking water regulations [Electronic resource]. – 2023. – Access mode: <https://www.epa.gov/ground-water-and-drinking-water/national-primary-drinking-water-regulations> World Health Organization. Nitrate and nitrite in drinking-water: Background document for development of WHO Guidelines for Drinking-water Quality [Electronic resource]. – 2011. – Access mode: <https://www.who.int/docs/default-source/wash-documents/wash-chemicals/nitrate-nitrite-background-document.pdf>
2. Zashchepkina, N., Melnychenko, D. Development of a prototype information-measuring system for express measurement of nitrates in liquids // Bulletin of the Khmelnytskyi National University. – 2023. – Issue 168. – DOI: <https://doi.org/10.31891/2219-9365-2023-75-6>
3. Hollas, J. M. Modern Spectroscopy [Electronic resource]. – 4th ed. – Chichester: John Wiley & Sons, 2004. – Chapter 2: "Introduction to Spectrophotometry".
4. Cogan, D. The application of spectrophotometry in the detection of nitrates and nitrites in environmental samples [Electronic resource]. – 2013. – Access mode: [https://doras.dcu.ie/19530/1/cest13\\_0400\\_DCogan.pdf](https://doras.dcu.ie/19530/1/cest13_0400_DCogan.pdf)
5. Harris, D.C. Quantitative Chemical Analysis. 9th ed. - New York: W. H. Freeman and Company, 2015. - P. 379-383

<https://doi.org/10.31891/2219-9365-2024-79-4>

УДК 676.22.017

ДАЩЕНКО Володимир

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

<https://orcid.org/0009-0000-9298-2993>

[apph.dvd@gmail.com](mailto:apph.dvd@gmail.com)

ВОЛКОВИЙ Павло

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

<https://orcid.org/0009-0000-6032-9084>

[pavel98volk@gmail.com](mailto:pavel98volk@gmail.com)

## ВИМІРЮВАННЯ ПОВЕРХНЕВОЇ ЩІЛЬНОСТІ МАТЕРІАЛУ З ОДНОРІДНОЮ СТРУКТУРОЮ УЛЬТРАЗВУКОВИМ МЕТОДОМ

*В результаті виготовлення, переробки паперової сировини та отримання на виході готової продукції, яка має однорідну структуру, виробник прагне отримати високу якість готового виробу. Для досягнення вказаної мети актуальною є задача вимірювання та контролю поверхневої щільності однорідного паперового матеріалу в процесі його виробництва, яка безпосередньо впливає на його якість. На сьогодні, підприємства паперової промисловості використовують контактний руйнівний метод визначення поверхневої щільності матеріалу, що не дає можливості реалізувати постійний моніторинг параметру. Для вирішення даної проблеми необхідно розробити нові комп'ютеризовані системи і методи вимірювання поверхневої щільності матеріалу із використанням саме безконтактних ультразвукових перетворювачів, оскільки тільки вони дають можливість забезпечити моніторинг параметру в режимі реального часу при невеликих грошових витратах на їх установку та експлуатацію у порівнянні з аналогами.*

*Для розробки безконтактного методу вимірювання поверхневої щільності однорідного матеріалу необхідно вирішити наступні завдання: дослідити теорію розповсюдження та ослаблення ультразвукових хвиль в контрольованому з однорідною структурою матеріалі; навести основні співвідношення амплітуд ультразвукових хвиль, які взаємодіють з матеріалом; розробити алгоритм вимірювальних та розрахункових процедур для визначення параметру поверхневої щільності матеріалу, що будуть застосовані в комп'ютеризованій системі.*

*Для контролю матеріалу необхідно враховувати особливості складу матеріалу і його вплив на інформативні параметри самого зондувального ультразвукового сигналу під час вимірювання. При безконтактному ультразвуковому методі матеріал опромінюється хвилями, а значення його поверхневої щільності визначається співвідношенням амплітуд ультразвукової хвилі, що пройшла крізь матеріал з однорідною структурою, і хвилі, яка падає на нього. Для налаштування комп'ютеризованої інформаційно-вимірювальної системи під параметри навколишнього середовища перед початком вимірювання поверхневої щільності матеріалу використовується опорний канал.*

*Ключові слова: поверхнева щільність, паперовий матеріал, однорідна структура, ультразвуковий метод, безконтактні перетворювачі, вимірювання, комп'ютеризована інформаційно-вимірювальна система.*

DASHCHENKO Volodymyr, VOLKOVYI Pavlo

National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

## MEASUREMENT OF THE AREA DENSITY OF A MATERIAL WITH A UNIFORM STRUCTURE USING THE ULTRASONIC METHOD

*As a result of the production, processing of paper raw materials and obtaining at the output of the finished product, which is a homogeneous structure, the manufacturer strives to obtain a high quality of the finished product. To achieve this goal, the task of measuring and controlling the area density of homogeneous paper material during its production, which directly affects its quality, is relevant. Today, paper industry enterprises use the contact destructive method of determining the area density of the material, which does not provide the opportunity to implement constant monitoring of the parameter. In order to solve this problem, it is necessary to develop new computerized systems and methods of measuring the area density of the material using non-contact ultrasonic transducers, since only they provide an opportunity to monitor the parameter in real time at low financial costs for their installation and operation in comparison with analogues.*

*To develop a non-contact method of measuring the area density of a homogeneous material, the following tasks must be solved: to investigate the theory of propagation and attenuation of ultrasonic waves in a controlled material with a homogeneous structure; state the main amplitude ratios of ultrasonic waves that interact with the material; to develop an algorithm of measurement and calculation procedures to determine the area density parameter of the material, which will be applied in the computerized system.*

*To control the material, it is necessary to take into account the features of the material composition and its influence on the informative parameters of the probing ultrasonic signal itself during the measurement. With the non-contact ultrasonic method, the material is irradiated with waves, and the value of its area density is determined by the ratio of the amplitudes of the ultrasonic wave that passed through the material with a uniform structure and the wave that falls on it. A reference channel is used to adjust the computerized information and measurement system to the environmental parameters before starting the measurement of the area density of the material.*

*Key words: area density, paper material, homogeneous structure, ultrasonic method, non-contact transducers, measurement, computerized information and measurement system.*

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

До матеріалів з однорідною суцільною структурою можна віднести полімерну плівку, папір, фольгу та інші матеріали, що використовуються сьогодні у широкому вжитку у всіх галузях господарства людини. Розглянемо, як приклад, більш детально паперовий матеріал. Відповідно до класифікації, яка застосовується в целюлозно-паперовій промисловості, папір традиційно поділяють за призначенням на 9 груп: папір для друку; пакувальний та обгортковий; для писання, креслення та малювання; оздоблювальний; електротехнічний; світлочутливий; для виготовлення тютюнових виробів; вбирний; технічно-промисловий папір різного призначення. Кожна група об'єднує види паперу, які різняться між собою волокнистим складом, масою  $1 \text{ м}^2$  (поверхневою щільністю), кількістю наповнювача, ступенем проклеювання та іншими властивостями. Оскільки сьогодні є необхідність в забезпеченні високої якості паперу особливо для друку, тому актуальним є також і завдання застосування безконтактних перетворювачів для моніторингу поверхневої щільності матеріалу на виробництві.

Папір для друку визначається як папір для друкування видавничої продукції, що включає такі види – газетний, типографський, тонкий типографський, офсетний, для глибокого друку, картографічний, документний, обкладинковий, етикетковий, титульний тощо. Слабким місцем цієї класифікації є те, що види паперу визначені як за сферою застосування, так і за способом друку. Критеріями визначення більшості видів паперу є ступінь проклеювання і зольність, а в деяких випадках склад та інші показники.

Класифікації паперу для друку можуть відрізнятися в різних країнах, але основним параметром паперу є маса  $1 \text{ м}^2$  або по іншому поверхнева щільність матеріалу. За державними стандартами, папір має масу до  $250 \text{ г/м}^2$ , а картон – понад  $250 \text{ г/м}^2$ . Однак, за міжнародними стандартами більшості країн, папером вважають матеріал з масою до  $225 \text{ г/м}^2$ , а картоном – понад  $225 \text{ г/м}^2$ .

Таким чином, виникає важлива наукова та практична проблема, яка полягає в забезпеченні оперативного вимірювання поверхневої щільності та контролю паперового матеріалу за цим параметром в процесі його виробництва та подальшої переробки.

## ПОСТАНОВКА ПРОБЛЕМИ І ОГЛЯД ОСТАННІХ ДОСЛІДЖЕНЬ

Існують різні методи контролю поверхневої щільності матеріалів, такі як ємнісний, радіоізотопний, метод відбиття радіохвиль, ультразвуковий та руйнівний метод вирізання і зважування зразків матеріалу на вагах.

Найпростішим є метод вирізання пробних зразків матеріалу та подальше їх зважування [1]. Для визначення поверхневої щільності, беруть зразок розміром приблизно  $10 \text{ см}^2$  та зважують його. Маса  $10 \text{ см}^2$  перераховують за формулою ділення маси зразка на його виміряне значення площі. Проте такий метод вимірювання не дозволяє проводити безперервний оперативний моніторинг параметру, а тільки забезпечує вимірювання в лабораторних умовах.

Ємнісний метод, який використовує електричні перетворювачі, є перспективним для контролю щільності, але має обмежену чутливість до зміни параметрів матеріалу. Відомі системи сканування матеріалу з ємнісними перетворювачами [2], які можна використовувати для контролю поверхневої щільності, за правильним їх налаштуванням.

Радіоізотопний метод [2], що визначає щільність через розсіювання чи поглинання випромінювання, ефективний для суцільних структур матеріалів, але потребує суворого дотримання заходів безпеки. Метод має переваги: незалежність похибки від умов навколишнього середовища і можливість контролювати паперову масу значної товщини.

Методи відбиття радіохвиль дозволяють визначати щільність матеріалу за опромінюючим сигналом [3], який взаємодіє з об'єктом вимірювання, після його відбиття і формування реєстраційного інформаційного сигналу.

Ультразвукові безконтактні методи вимірювання та контролю є найбільш ефективними, оскільки забезпечують високу точність і зниження вартості приладів порівняно з аналогічними вимірювальними засобами із використанням інших видів датчиків. Зазвичай в ультразвукових вимірювальних системах випромінювання спрямовується перпендикулярно до поверхні матеріалу, а для їх налаштування використовують опорний канал з повітряним середовищем [4]. Ці методи застосовуються для швидкого отримання вимірювальної інформації про параметри матеріалів у виробничому процесі, що забезпечує високу ефективність виробництва.

Сьогодні не існує у світі стандартного безконтактного методу вимірювання поверхневої щільності матеріалів з однорідною структурою, який дозволив би оперативно контролювати цей параметр, та був би простим в реалізації та експлуатації. Тому для контролю матеріалу з однорідною структурою необхідно застосовувати метод, який буде враховувати особливості складу матеріалу і його вплив на інформаційні параметри самого зондувального ультразвукового сигналу під час вимірювання. Для безконтактного оперативного вимірювання поверхневої щільності різних матеріалів з однорідною суцільною структурою

необхідно провести дослідження розповсюдження ультразвукових хвиль в цих середовищах. Проведений аналіз показав, що найбільш перспективними методами вимірювання та контролю поверхневої щільності різних матеріалів є безконтактні ультразвукові методи [5-12]. При цьому матеріал опромінюється ультразвуковими хвилями, а значення його параметру визначається співвідношенням амплітуд ультразвукової хвилі, що пройшла крізь текстильний матеріал, і хвилі, яка падає на нього.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є дослідження методу та системи для вимірювання і контролю поверхневої щільності матеріалів з однорідною структурою, задля подальшої реалізації безперервного моніторингу цього параметру в процесі виробництва.

Для досягнення мети необхідно вирішити наступні завдання:

- дослідити теорію розповсюдження та ослаблення ультразвукових хвиль в контрольованому суцільному матеріалі малої товщини;
- навести основні співвідношення амплітуд ультразвукових хвиль, які описують принципи реалізації безконтактного метода вимірювання;
- розробити алгоритм вимірювальних та розрахункових процедур для визначення параметру поверхневої щільності матеріалу з однорідною структурою.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Вираз для модуля  $|W(\omega)|$  комплексного коефіцієнта проходження хвиль або, іншими словами, для співвідношення амплітуд ультразвукових хвиль, однієї, що пройшла контрольоване середовище з поверхневою щільністю  $m_s$ , до другої, яка тільки падає на нього, можна подати як [6]:

$$|W(\omega)| = \left( \left( \frac{4Z_1 Z_2}{(Z_1 + Z_2)^2} \right)^2 \left( \sum_{N=0}^{\infty} \left( \frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^{2N} \cdot \cos \left( (2N + 1) \frac{m_s \omega}{\pi Z_2} \right) \times \right. \right. \\ \left. \left. \times e^{-\alpha_s (2N+1) \frac{m_s}{\pi \rho_2}} \right)^2 + \left( \sum_{N=0}^{\infty} \left( \frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^{2N} \cdot \sin \left( (2N + 1) \frac{m_s \omega}{\pi Z_2} \right) \times \right. \right. \\ \left. \left. \times e^{-\alpha_s (2N+1) \frac{m_s}{\pi \rho_2}} \right)^2 \right)^{\frac{1}{2}}, \quad (1)$$

де  $Z_1$  – акустичний опір середовища, з якого падає хвиля;  $Z_2, \rho_2$  – акустичний опір та об'ємна щільність середовища матеріалу, в яке потрапляє ультразвукова хвиля;  $\omega$  – колова частота;  $\alpha_s$  – коефіцієнт згасання ультразвукових хвиль в контрольованому матеріалі;  $N$  – кількість перевідбивань ультразвукових хвиль від меж середовищ з різними акустичними опорами.

Вираз для модуля  $|V(\omega)|$  комплексного коефіцієнта відбиття хвиль або, іншими словами, для співвідношення амплітуд ультразвукових хвиль, однієї, що відбилася від контрольованого середовища з поверхневою щільністю  $m_s$ , до другої, яка тільки падає на нього, можна подати так [6]:

$$|V(\omega)| = \left( \left( \frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^2 \cdot \left( \sum_{N=0}^{\infty} \left( \frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^{2N} \cdot \left( \cos \left( 2N \frac{m_s \omega}{\pi Z_2} \right) \times \right. \right. \right. \\ \left. \left. \times e^{-\alpha_s 2N \frac{m_s}{\pi \rho_2}} - \cos \left( (2N + 2) \frac{m_s \omega}{\pi Z_2} \right) \times e^{-\alpha_s (2N+2) \frac{m_s}{\pi \rho_2}} \right) \right)^2 + \\ \left. \left( \sum_{N=0}^{\infty} \left( \frac{Z_2 - Z_1}{Z_1 + Z_2} \right)^{2N} \cdot \left( \sin \left( 2N \frac{m_s \omega}{\pi Z_2} \right) \cdot e^{-\alpha_s 2N \frac{m_s}{\pi \rho_2}} - \right. \right. \right. \\ \left. \left. \left. - \sin \left( (2N + 2) \frac{m_s \omega}{\pi Z_2} \right) \cdot e^{-\alpha_s (2N+2) \frac{m_s}{\pi \rho_2}} \right) \right) \right)^2 \right)^{\frac{1}{2}}. \quad (2)$$

Отримані залежності (1) та (2) є базовими для створення вимірювальних систем на основі

амплітудного методу опрацювання сигналу, що проходить крізь контрольований матеріал, та сигналу, який відбивається від нього.

Відносні зміни амплітуд хвиль, що спричиняються згасанням сигналів, для різних матеріалів, враховуючи залежності (1) та (2), після спрощення складових виразів із переходом від колової частоти  $\omega$  до  $f$  та зважаючи на те, що величини  $\frac{Z_1}{Z_2} \ll 1$ ,  $\frac{m_s \omega}{\pi Z_2} \ll 1$  можна подати так:

$$|W| = \frac{1}{\sqrt{1 + \left(\frac{m_s f}{Z_1}\right)^2}}, \quad (3)$$

$$\delta_{|W|} = \left( |W| \sqrt{1 + \left(\frac{m_s f}{Z_1}\right)^2} - 1 \right) \cdot 100\%,$$

$$|V| = \frac{1}{\sqrt{1 + \left(\frac{Z_1}{m_s f}\right)^2}}, \quad (4)$$

$$\delta_{|V|} = \left( |V| \sqrt{1 + \left(\frac{Z_1}{m_s f}\right)^2} - 1 \right) \cdot 100\% .$$

Використаємо формулу (3) для реалізації амплітудного ультразвукового метода вимірювання поверхневої щільності паперу на проходження хвиль для розробки структури комп'ютеризованої інформаційно-вимірювальної системи (КІВС). На рисунку 1 показано структурну схему КІВС, за допомогою якої можна контролювати поверхневу щільність паперових матеріалів з використанням амплітудних залежностей ультразвукових сигналів.

Запропонована КІВС містить генераторну частину електричних імпульсів на мікроконтролері 1, блок живлення формувача зондувальних імпульсів 2, формувач радіоімпульсів на ключах 3, підсилювач потужності 4, випромінюючі перетворювачі 5 та 7, приймаючі перетворювачі 6 і 8, підсилювач з автоматичним регульованим підсиленням (АРП) 9, регулюючий атенюатор 10, детектори пікової напруги 11 і 12, мікроконтролер 13, перетворювач рівнів сигналів (ПР) 14 і персональний комп'ютер (ПК) 15, контрольований паперовий матеріал 16.

КІВС працює наступним чином. Електричні коливання з генераторної частини на мікроконтролері 1 потрапляють до формувача радіоімпульсів на ключах 3, який використовує напругу з блоку живлення 2 для підсилення імпульсів, з виходу якого радіоімпульси додатково підсилюються підсилювачем 4 за потужністю і поступають на випромінюючий перетворювач 5. Ультразвукові коливання, що випромінюються приймаються перетворювачем 6, після проходження ними контрольованого паперового матеріалу 16 (вимірювальний канал). Радіоімпульси подаються також на випромінюючий перетворювач 7, який випромінює ультразвукові коливання в повітряне середовище (опорний канал), що потрапляють в приймаючий перетворювач 8. Електричні коливання з другого приймаючого перетворювача 8 після проходження атенюатора 10 та пікового детектора 11 подаються на один вхід мікроконтролера 13, на інший його вхід подається вихідна напруга приймаючого перетворювача 6 після підсилення її підсилювачем з АРП 9 та виділення максимальної амплітуди напруги піковим детектором 12. Вихідний код інформаційного параметру мікроконтролера 13 треба подати у вигляді добутку вхідної напруги опорного каналу та модулю коефіцієнту проходження  $|W|$  ультразвукового сигналу, який можна визначити з виразу (3). Вхідні напруги перетворюються за допомогою внутрішнього АЦП мікроконтролера 13 в цифрові коди і записуються в пам'ять ПК 15 після передачі інформації з ПР 14.

Приймаючи, що  $Z_1/f = \text{const}$  в ПК 15 визначається поверхнева щільність паперового матеріалу за формулою:

$$m_s = \frac{\sqrt{\left(\frac{1}{|W|^2} - 1\right)} Z_1}{f}. \quad (5)$$

Результат відображається на моніторі ПК 15. В пам'ять ПК 15 у вигляді констант введені значення частоти  $f$ , швидкості розповсюдження ультразвукових коливань  $c_1$  в навколишньому середовищі (повітрі), щільність цього середовища  $\rho_1$  за нормальних умов. При відсутності контрольованого паперового матеріалу 16 у вимірювальному каналі перед початком вимірювань за допомогою регулюючого атенюатора 10 проводиться зрівнювання амплітуд напруг на входах мікроконтролера 13 за їх кодованими значеннями (при цьому  $|W| = 1$ ), отже із виразу (5) отримаємо  $m_s = 0$ .

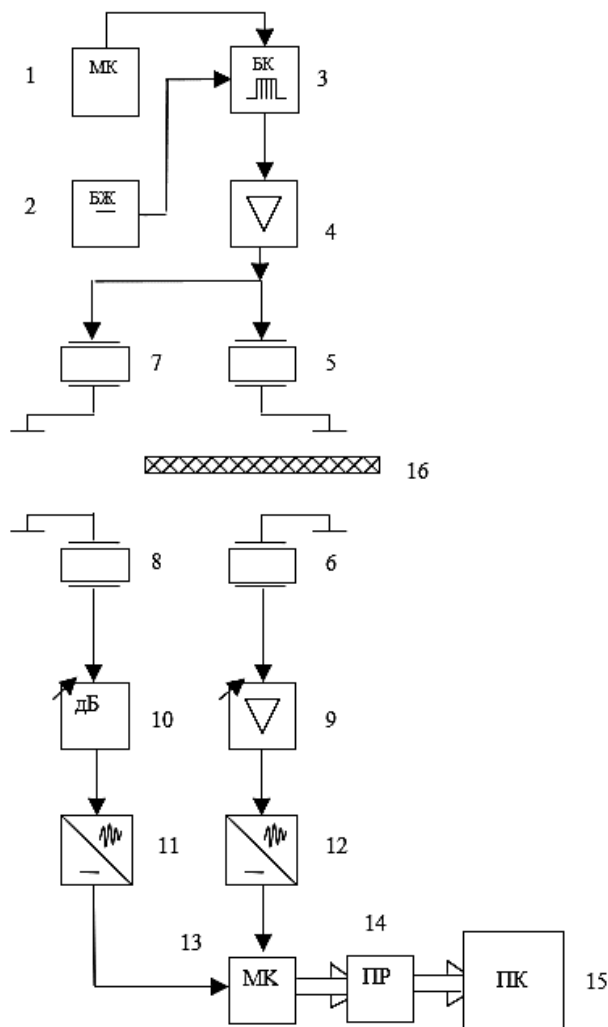


Рис. 1. Структурна схема КІВС для вимірювання та контролю поверхневої щільності паперових матеріалів

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Проведені дослідження показали, що за допомогою безконтактного ультразвукового методу та інформаційно-вимірювальної системи можна визначити поверхневу щільність паперового матеріалу з врахуванням зміни параметрів навколишнього середовища. Тому подальший розвиток безконтактних ультразвукових методів вимірювання поверхневої щільності матеріалів з однорідною структурою дасть можливість у майбутньому створювати високоточні засоби вимірювання різних технологічних параметрів матеріалів вже з неоднорідною структурою. Це дозволить проводити безперервний оперативний контроль як поверхневої щільності, так і інших технологічних параметрів різних матеріалів, а також підвищити якість готової продукції.

#### Література

1. Здоренко В.Г., Барилко С.В. Дослідження похибки розрахункового методу визначення поверхневої щільності тканин. Вісник КНУТД. 2013. № 2 (70). С. 16–21.
2. Gloy Y.-S., Gries T., Spies G. (2013). Non Destructive Testing of Fabric Weight in the Weaving Process. 13th International Symposium on Nondestructive Characterization of Materials (NDCM-XIII), 20-24 May 2013, Le Mans, France. NDT.net Issue: 2014-04).

3. ТОЕ. Електромагнітне поле: підручник / Ю. О. Карпов, Ю. Г. Ведміцький, В. В. Кухарчук. – Херсон: ОЛДІ-ПЛЮС, 2014. – 392 с.
4. Здоренко В.Г., Барилко С.В. Дослідження проходження імпульсного ультразвукового сигналу крізь тканину при контролі поверхневої щільності. Вісник ХНУ. 2013. № 3. С. 90–96.
5. Barylko S, Zdorenko V, Kuzymchuk O, et al (2019). Adaptive ultrasonic method for controlling the basis weight of knitted fabrics, *Journal of Engineered Fibers and Fabrics*, 14, P. 1-7.
6. Zdorenko V. Development of a Computerized Control System Structure and Study of Ultrasonic Wave Propagation in Various Textile Materials / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 33 – 82.
7. Zdorenko V. Research on the Interaction of Ultrasonic Waves with Various Textile Materials in the Process of Non-contact Control / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 83 – 124.
8. Zdorenko V. Application of Non-contact Methods to Control the Technological Parameters of Textile Materials in the Production Process / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 125 – 165.
9. Zdorenko V. Design of the Models and Methods of Constructing Computerized Control Systems of Technological Parameters of Textile Materials / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 167 – 214.
10. Zdorenko V. Development of Experimental Samples of Computerized Systems and Non-contact Control Over Technological Parameters of Textile / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 215 – 289.
11. Здоренко В.Г. Застосування ультразвукового пристрою для визначення поверхневої густини текстильної волоконної маси / Здоренко В.Г., Барилко С.В., Лісовець С.М., Шипко Д.О. // Вісник Київського національного університету технологій та дизайну. №5 (150). 2020. – С. 67 –73.
12. Здоренко В.Г., Барилко С.В., Чорноморченко В.К., Защепкіна Н.М. Дослідження згасання амплітуд ультразвукових хвиль при контролі тканин з різною поверхневою щільністю. Вісник Київського національного університету технологій та дизайну. 2013. № 3 (71). С. 24–29.

#### References

1. Zdorenko V.G., Barylko S.V. Doslidzhennia pokhybky rozrakhunkovoho metodu vyznachennia poverkhnevoi shchilnosti tkanyn. *Visnyk KNUVD*. 2013. № 2 (70). P. 16–21.
2. Gloy Y.-S., Gries T., Spies G. (2013). Non Destructive Testing of Fabric Weight in the Weaving Process. 13th International Symposium on Nondestructive Characterization of Materials (NDCM-XIII), 20-24 May 2013, Le Mans, France. *NDT.net Issue: 2014-04*.
3. ТОЕ. Elektromagnitne pole: pidruchnyk / Yu. O. Karpov, Yu. H. Vedmitskiy, V. V. Kukharchuk. – Kherson: OLDI-PLIUS, 2014. – 392 p.
4. Zdorenko V.G., Barylko S.V. Doslidzhennia prokhodzhennia impulsnoho ultrazvukovoho syhnalu kriz tkanynu pry kontroli poverkhnevoi shchilnosti. *Visnyk KhNU*. 2013. № 3. P. 90–96.
5. Barylko S, Zdorenko V, Kuzymchuk O, et al (2019). Adaptive ultrasonic method for controlling the basis weight of knitted fabrics, *Journal of Engineered Fibers and Fabrics*, 14, P. 1-7.
6. Zdorenko V. Development of a Computerized Control System Structure and Study of Ultrasonic Wave Propagation in Various Textile Materials / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 33 – 82.
7. Zdorenko V. Research on the Interaction of Ultrasonic Waves with Various Textile Materials in the Process of Non-contact Control / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 83 – 124.
8. Zdorenko V. Application of Non-contact Methods to Control the Technological Parameters of Textile Materials in the Production Process / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 125 – 165.
9. Zdorenko V. Design of the Models and Methods of Constructing Computerized Control Systems of Technological Parameters of Textile Materials / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 167 – 214.
10. Zdorenko V. Development of Experimental Samples of Computerized Systems and Non-contact Control Over Technological Parameters of Textile / Zdorenko V., Barylko S., Zashchepkina N., Zaporozhets A., Lisovets S., Kiva I. // *Studies in Systems, Decision and Control*. 460. 2023. – P. 215 – 289.
11. Zdorenko V.G. Zastosuvannia ultrazvukovoho prystroiu dlia vyznachennia poverkhnevoi hustyny tekstylnoi volokonnoi masy / Zdorenko V.H., Barylko S.V., Lisovets S.M., Shypko D.O. // *Visnyk Kyivskoho natsionalnoho universytetu tekhnologii ta dizainu*. №5 (150). 2020. – P. 67 –73.
12. Zdorenko V.G., Barylko S.V., Chornomorchenko V.K., Zashchepkina N.M. Doslidzhennia zghasannia amplitud ultrazvukovykh khvyl pry kontroli tkanyn z riznoiu poverkhnevoiu shchilnistiu. *Visnyk Kyivskoho natsionalnoho universytetu tekhnologii ta dizainu*. 2013. № 3 (71). P. 24–29.

<https://doi.org/10.31891/2219-9365-2024-79-5>

УДК 004

КРУПА Дмитро

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0002-6087-9988>

e-mail: [dima-krupa@ukr.net](mailto:dima-krupa@ukr.net)

## ОЦІНКА СУМІСНОСТІ ОКРЕМИХ КОМПОНЕНТ ВЕБ-СЕРВЕРІВ З ВИКОРИСТАННЯМ ОНТОЛОГІЧНОГО ПІДХОДУ ТА НЕЧІТКОГО ЛОГІЧНОГО ВИВЕДЕННЯ

*У статті описаний метод перевірки сумісності окремих компонент зі використанням онтологічного підходу та нечіткого логічного виведення. Запропонована структура методу оцінки сумісності об'єктів між собою. Наведений приклад використання онтологічного підходу до оцінки сумісності компонент. Запропоновані нечіткі змінні та модифіковано метод нечіткого логічного виведення. Наведений приклад роботи методу.*

*Ключові слова: сумісність компонент, онтологічний підхід, нечітке логічне виведення*

KRUPA Dmytro

Lviv Polytechnic National University

## EVALUATION OF THE COMPATIBILITY OF INDIVIDUAL COMPONENTS OF WEB SERVERS USING THE ONTOLOGICAL APPROACH AND FUZZY LOGICAL INFERENCE

*System compatibility is significant problem in software development, especially when systems can be changed independently. To check their compatibility, their application programming interfaces (APIs) should be compatible. But not only APIs ensure compatibility, software developers also should be sure that their apps fit logically. This process can be done using manual process of checking systems and their interfaces or by covering both apps with integration and API tests. Additionally, expert should proof that systems are compatible and their OOP structure fits to both business models.*

*In this article, system compatibility problem was raised and described. Was suggested method to check object compatibility. Object compatibility is initial step to check interface compatibility. Also, ontology approach is described and suggested how it can be used in method. In suggested method ontologies can be used by experts or ontology reasoning checks can be executed. Ontology assessment was divided into structural and logical parts. Where structural part can be done by automated check to ensure that requested object is compatible with ontology concept. Logical part of assessment can be done by using ontology reasoning or expert interaction.*

*Was suggested structure of base object for representation each instance that should be checked. This is initial step for creating compatibility checking process and/or framework.*

*Beside this, fuzzy logical inference was modified to fits suggested methods. Some base rules were given to explicitly show how logical deduction works in this case. Also, expert's intervention in the process has more impact than automatic check to decrease possibility of mistake.*

*Also, simple example was processed using suggested method to presents how it works. In this example was checked compatibility between objects "Room" and "Cabinet". Was suggested that they are for different systems and checks if these objects are compatible.*

*Keywords: component compatibility, ontology approach, fuzzy logical inference.*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

При розробці комплексних рішень для бізнесу часто виникає потреба інтеграції із іншими системами для спрощення виконання завдань, які обумовлюються бізнес вимогами. Наприклад, відправка емейлів, збереження файлів, редагування відео, тощо. Дані сервіси повністю вирішують поставлену задачу, а системі, яка з ними інтегрується, достатньо лише надати дані для опрацювання.

Окремо можемо виділити вимоги, що потребують інтеграції з іншими системами, які оперують тими ж самими або схожими поняттями. Для спрощення розуміння проблеми у статті буде розглянуто серверні технології, зокрема інтеграція двох систем на основі серверів, бізнес яких працює в одній області та оперують тотожними або схожими поняттями.

Із наведених прикладів такі поняття як «емейл», «файл», «відео» є загальновідомими, тому при їх використанні не потрібні спеціальні знання у даній області. Але виникає проблема, коли дві системи інтегруються між собою та оперують складними, нетривіальними або неоднозначними поняттями як, наприклад, «замовлення товару» чи «доставка». В одній системі поняття «доставка» може бути визначено як адресна доставка до покупця, а в іншій – завантаження товару у фургон перевізника на складі.

Якщо говорити про сумісність компонент для серверних технологій, то проблема зводиться до сумісності між запитами (а саме об'єктами), що передаються між двома системами. Вони повинні не лише мати необхідні дані, але і логічно (з точки зору бізнесу) підходити для виконання завдання.



## АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Зазвичай перевірка сумісності веб-серверів є складним та затратним процесом. Для перевірки сумісності компонент можна використовувати інтеграційні тести [8], але вони будуть ефективні лише при незмінності однієї із компонент.

Проте виникає проблема, коли системи, що були інтегровані одна в іншу, починаються змінюватися і їхня інтеграція перестає працювати коректно. Важливим завданням є не лише перевірка факту сумісності двох веб-серверів, але і надання максимальної інформації про їхню взаємодію та що саме потрібно змінити, щоб система в цілому почала функціонувати як було заплановано.

При представленні знань про домену область кожної із компонент у вигляді онтологій варто зазначити що онтологія, як і сама система, теж може та має еволюціонувати щоб завжди чітко відповідати своїй сфері застосування [3].

Проблема сумісності компонент та пропозиції її вирішення була розглянута у роботах [4–7]. Ці методи дозволяють оцінити та забезпечити сумісність серверів між собою. Але не завжди є можливість створити програмне рішення яке повністю покриє усі можливі проблеми із сумісністю компонентів зважаючи на обмеженість ресурсів (проблема «залізного трикутника») [9]. Тому потрібно розробити просте та гнучке рішення, яке буде допомагати розробникам забезпечувати сумісність систем між собою.

## ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

**Метою роботи** є розробити метод визначення ступеня сумісності між класами, використовуючи нечітке виведення та онтологічний підхід.

## ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

У даній статті будемо користуватися загальновідомими термінами ООП: клас, об'єкт, поле [10]. Також система, з якою планується інтеграція, матиме об'єкти, які ми вважатимемо еталонними, бо саме з ними будемо порівнювати власні об'єкти. Важливим є наявність онтології, яка формалізовано відображає структуру системи з якою відбувається інтеграція.

Створимо структуру методу, що буде надавати оцінку сумісності класів, зважаючи на їхню роль в бізнес-домени.

Структура методу

- Потрібно представити два об'єкти як множину полів з метаданими
- Визначити поля, що є обов'язковими для порівняння
- Додати обмеження по типу та значеннях
- Використання онтологічного підходу для оцінки сумісності класів
- Використати нечітке виведення для перевірки відповідності об'єкту еталонному
- Надати оцінку сумісності класів

### Представлення об'єктів як множину полів з метаданими

Для початку представимо поле об'єкту як новий об'єкт із метаданими.

Розглянемо клас як множину характеристик, а об'єкт – як набір понять «ключ-значення». Також створимо універсальний об'єкт (тут і далі будемо його називати MetaObject), який буде інкапсулювати метадані поля для подальшого порівняння. Такий об'єкт може містити дані: назва поля, тип поля, список допустимих значень, список обмежень, шлях у ієрархії до найабстрактнішого класу (наприклад, у Java – це клас Object).

Для порівняння сумісності(відповідності) двох полів потрібно представити їх у форматі MetaObject та використати функцію, яка даватиме числове представлення сумісності двох полів.

*Function*  $checkCompatibility(x,y) = [0;1]$ ,

де  $x$  та  $y$  – екземпляри класу MetaObject,  $[0;1]$  – межі результату обчислення.

Реалізація функції *checkCompatibility* може змінюватися в залежності від вимог проекту.

Структура MetaObject у форматі псевдокоду:

```
Об'єкт MetaObject {
  Поле: {
    Ім'я: "Name",
    Тип: "string"
  },
  Поле: {
    Ім'я: "Type",
    Тип: "string"
  }
}
```

```

    },
    Поле: {
        Ім'я: "allowedValues",
        Тип: "Array"
    },
    },
    Поле: {
        Ім'я: "restrictions",
        Тип: "Array"
    },
    },
    Поле: {
        Ім'я: "hierarchy",
        Тип: "String"
    }
}
    }
}
    
```

#### 1. Використання онтологічного підходу

Для того, щоб визначити чи відповідає порівнюваний клас предметній області, необхідно перевірити його відповідність онтології, зокрема як він відповідає еталонному класу. У цьому кроці необхідна оцінка експерта, що порівняє запитований клас із відповідним йому в онтології [2].

На цьому етапі нас цікавлять дві оцінки: оцінка відповідності структурі та оцінка логічної відповідності.

Оцінку відповідності структурі можна отримати програмним шляхом: чи клас відповідає концепту, має усі його властивості та дотримується його аксіом. Наприклад, *Кімната*  $\sqsubseteq$  *Змає.Вікно*.

Оцінку логічної відповідності може надати експерт, або спробувати співставити назву та властивості об'єкта із концептом, використовуючи тезаурус з конкретної предметної області. Наприклад, "surname" можна замінити словом "lastname".

#### 2. Оцінка сумісності класів

Для роботи із заданими параметрами використаємо визначення нечіткої змінної, що складається із трійки параметрів  $\langle a, X, A \rangle$  [1], де

$a$  – ім'я змінної

$X$  – універсальна множина

$A$  – нечітка підмножина множини  $X$ , що описує обмеження на значення змінної  $a$  ( $\mu_A(x)$ )

Таким чином у нас сформовано три нечіткі змінні:

$\langle$ оцінка основних властивостей,  $[0;1]$ ,  $(\mu_{A1}(X)) \rangle$

$\langle$ оцінка додаткових властивостей,  $[0;1]$ ,  $(\mu_{A2}(X)) \rangle$

$\langle$ оцінка відповідності онтології,  $[0;1]$ ,  $(\mu_{A3}(X)) \rangle$

Наступним кроком створимо правило, яке на основі оцінок визначає ступінь логічної відповідності.

Оскільки ми працюємо із нечіткими даними, використаємо нечітке логічне виведення.

Для спрощення розглянемо еталонний об'єкт  $E$  та порівнюваний об'єкт  $A$ .

$\mu_{A1}(X) = \sum_i \text{checkRequiredCompatibility}(x_i, y_i)$

$\mu_{A2}(X) = \sum_i \text{checkOptionalCompatibility}(x_i, y_i)$

Логічна оцінка =  $\min$ (оцінка експерта, програмна оцінка)

Структурна оцінка =  $\min$ (оцінка експерта, програмна оцінка)

$\mu_{A3}(X) = 0.65 * \text{логічна оцінка} + 0.35 * \text{структурна оцінка}$

Коефіцієнти 0.65 вказує на перевагу логічній відповідності, яка надається експертом у галузі, для якої розроблена онтологія, а 0.35 – на важливість відповідності структурі онтології, що необхідно для розробки програмного рішення.

#### 1. Фазицікація

А) оцінка основних властивостей

Функція приналежності:

- НеВідповідає ( $x$ ) =  $\max(0, \min(1, 1 - x))$ ;
- ЧастковоВідповідає ( $x$ ) =  $1 - |2x - 1|$
- Відповідає ( $x$ ) =  $\max(0, \min(x, 1))$ ;

Б) оцінка додаткових властивостей

Функція приналежності:

- НеВідповідає ( $x$ ) =  $\max(0, \min(1, 1 - x))$ ;
- ЧастковоВідповідає ( $x$ ) =  $1 - |2x - 1|$
- Відповідає ( $x$ ) =  $\max(0, \min(x, 1))$ ;

В) оцінка відповідності онтології

Оцінка відповідності онтології може як надаватися експертом у нечіткому вигляді і не потребувати

фазифікації, так і може бути у вигляді чіткого числа, що необхідно фазифікувати.

Функція приналежності:

- НеВідповідає  $(x) = \max(0, \min(1, 1 - x))$ ;
- ЧастковоВідповідає  $(x) = 1 - |2x - 1|$
- Відповідає  $(x) = \max(0, \min(x, 1))$ ;

## 2. Пошук логічного висновку

Сформуємо правила формату «ЯКЩО ... ТО ...» для пошуку логічного висновку [2].

- Якщо оцінка основних властивостей відповідає І оцінка додаткових властивостей відповідає І оцінка відповідності онтології відповідає ТО системи повністю сумісні
- Якщо оцінка основних властивостей *частково відповідає* І оцінка додаткових властивостей *відповідає* І оцінка відповідності онтології *не відповідає* ТО системи *не сумісні*
- Якщо оцінка основних властивостей *відповідає* І оцінка додаткових властивостей *не відповідає* І оцінка відповідності онтології *не відповідає* ТО системи *потребують додаткової сумісності*
- Якщо оцінка основних властивостей *частково відповідає* І оцінка додаткових властивостей *не відповідає* І оцінка відповідності онтології *не відповідає* ТО системи *потребують додаткової сумісності*
- Якщо оцінка основних властивостей *відповідає* І оцінка додаткових властивостей *не відповідає* І оцінка відповідності онтології *відповідає* ТО системи *сумісні*

Кількість та умови правил можуть змінюватися в залежності від вимог системи

## 3. Дефазифікація

Призначимо числові ваги вихідним термінам:

- повністю сумісні = 1
- не сумісні = 0
- потребують додаткової сумісності = 0.5

Якщо ж активуються декілька правил, то будемо використовувати формулу:  
 $Fazz = \min(\text{rule1}, \text{rule 2}, \text{rule 3}, \text{rule 4}, \text{rule 5})$ .

Причина полягає в тому, що якщо виникає сумнів в інтеграції систем – необхідно приділити цьому увагу та вирішити імовірну проблему.

## 3. Розглянемо роботу методу на прикладі

Як приклад розглянемо вищенаведений клас «Кімната» та чи буде він сумісним із класом «Кабінет».

Розглянемо клас «Кімната», як один із популярних прикладів, які використовуються для вивчення ООП. Характеристики (поля) «Кімнати» {«ширина», «довжина», «висота», «кількість вікон», «колір стін»}.

Перетворимо об'єкт класу «Кімната» та список об'єктів типу MetaObject.

```
{
  "Name": "ширина"
  "Type": "Number",
  "allowedValues": ["x > 0"],
  "restrictions": [],
  "hierarchy": "Number -> Object"
}
```

Аналогічні об'єкти будуть розроблені для полів «довжина» та «висота». «Кількість вікон» буде відрізнятися типом даних: цілий тип (Integer) замість Number.

«Колір стін» можна перетворити у такий об'єкт:

```
{
  "Name": "колір стін"
  "Type": "String",
  "allowedValues": ["білий", "чорний", "зелений", "синій" ... ],
  "restrictions": [],
  "hierarchy": "String -> Object"
}
```

Обов'язковими полями є: «ширина», «довжина», «висота» та «колір стін». Додатковими: «кількість вікон».

Клас «Кабінет» має поля: «ширина», «довжина», «висота», «робочий стіл», «колір стін»}. Аналогічно перетворюємо його поля.

Онтологічний підхід:

Припустимо, що експертна оцінка «так, класи сумісні». Тобто,

Логічна оцінка = 1

Перевіримо виконання аксіом з концепту “Кімната” для об’єкту “Кабінет”.

- $Кімната \sqsubseteq \exists має Колір Стін. Колір = true = 1$
- $Кімната \sqsubseteq \exists має Ширину. Ширина = true = 1$
- $Кімната \sqsubseteq \exists має Довжину. Довжина = true = 1$
- $Кімната \sqsubseteq \exists має Висоту. Висота = true = 1$
- $Кімната \sqsubseteq \exists має Вікно. Вікно = false = 0$

Для даного прикладу обчислимо структурну оцінку як середнє значення аксіом:

Структурна оцінка = 0.8

$$\mu_{A1}(X) = \sum_i \text{checkRequiredCompatibility}(x_i, y_i) = 1$$

$$\mu_{A2}(X) = \sum_i \text{checkOptionalCompatibility}(x_i, y_i) = 0$$

$$\mu_{A3}(X) = 0.65 * 1 + 0.35 * 0.8 = 0.93$$

Результат фазифікації:

оцінка основних властивостей = Відповідає

оцінка додаткових властивостей = Не відповідає

оцінка відповідності онтології = Відповідає

Для даного прикладу застосовується правило №5 -> Системи сумісні між собою.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Розроблений метод дозволяє порівнювати об’єкти з однієї системи з бізнес-логікою іншої. Він надає універсальний інструмент перевірки сумісності об’єктів з використанням онтологічного підходу та з аналізом метаданих кожної властивості об’єкту.

Також було модифіковано структуру нечіткого логічного виведення, що повинно зменшити імовірність похибки завдяки використанню онтологій.

Окрім того, було окреслено важливість залучення експертів як із сторони бізнесу, так і експертів з розробки програмних рішень для модифікації функцій та правил нечіткого виведення з метою відповідності потребам проекту.

Підхід, розглянутий у даній статті, є початковим кроком для розробки гнучкого та не ресурсозатратного рішення для перевірки сумісності систем що активно розробляються.

### Література

1. Желдак Т. А. Нечіткі множини в системах управління та прийняття рішень / Т. А. Желдак, Л. С. Коряшкіна, С. А. Ус. – Дніпро : НТУ «Дніпровська політехніка», 2020. – 387 с.
2. Cingolani P. jFuzzyLogic: a Java Library to Design Fuzzy Logic Controllers According to the Standard for Fuzzy Control Programming / P. Cingolani, J. Alcal. // International Journal of Computational Intelligence Systems. – 2013.
3. Басюк Т. М. Онтологічний інжиніринг / Т. М. Басюк, В. В. Литвин, Д. Г. Досин. – Львів, 2016. – 264 с.
4. Liang Q. Ontology-Based Compatibility Checking for Web Service Configuration Management / Q. Liang, M. N. Huhns. – P. 15.
5. Craig D. C. Verification of Component Behavioral Compatibility / D. C. Craig, W. Zuberek. – 2007. – URL : [https://www.researchgate.net/publication/4261561\\_Verification\\_of\\_Component\\_Behavioral\\_Compatibility](https://www.researchgate.net/publication/4261561_Verification_of_Component_Behavioral_Compatibility)
6. Pahl C. An ontology for software component matching / Claus Pahl // International Journal on Software Tools for Technology Transfer. – 2007.
7. Medjahed B. Context-based matching for Web service composition / B. Medjahed, Y. Atif // Distributed and Parallel Databases. – 2007.
8. Chan W. An Overview of Integration Testing Techniques for Object-Oriented Programs / W. Chan, T. Chen, T. Tse. – 2002.
9. Stojcetovic B. Project management: cost, time and quality / Bojan Stojcetovic. – 2013. – URL : [https://www.researchgate.net/publication/305462896\\_Project\\_management\\_cost\\_time\\_and\\_quality](https://www.researchgate.net/publication/305462896_Project_management_cost_time_and_quality).
10. Object-Oriented Programming in Computer Science / R. Yilmaz, A. Sezgin, S. Kurnaz, Y. Arslan // Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics. 2019. – С. 1439–1451.

---

### References

1. Zheldak T. A. Nechitki mnozhyny v systemakh upravlinnia ta pryiniattia rishen / T. A. Zheldak, L. S. Koriashkina, S. A. Us. – Dnipro: NTU «Dniprovska politehnika», 2020. – 387 p.
2. Cingolani P. jFuzzyLogic: a Java Library to Design Fuzzy Logic Controllers According to the Standard for Fuzzy Control Programming / P. Cingolani, J. Alcal. // International Journal of Computational Intelligence Systems. – 2013.
3. Basiuk T. M. Ontolohichnyi inzhynirinh / T. M. Basiuk, V. V. Lytvyn, D. H. Dosyn. – Lviv, 2016. – 264 p.
4. Liang Q. Ontology-Based Compatibility Checking for Web Service Configuration Management / Q. Liang, M. N. Huhns. – C. 15.
5. Craig D. C. Verification of Component Behavioral Compatibility [Electronic resource] / D. C. Craig, W. Zuberek. – 2007. – Resource access mode: [https://www.researchgate.net/publication/4261561\\_Verification\\_of\\_Component\\_Behavioral\\_Compatibility](https://www.researchgate.net/publication/4261561_Verification_of_Component_Behavioral_Compatibility)
6. Pahl C. An ontology for software component matching / Claus Pahl. // International Journal on Software Tools for Technology Transfer. – 2007.
7. Medjahed B. Context-based matching for Web service composition / B. Medjahed, Y. Atif. // Distributed and Parallel Databases. – 2007.
8. Chan W. An Overview of Integration Testing Techniques for Object-Oriented Programs / W. Chan, T. Chen, T. Tse. – 2002.
9. Stojcetovic B. Project managment: cost, time and quality [Electronic resource] / Bojan Stojcetovic. – 2013. – – Resource access mode: [https://www.researchgate.net/publication/305462896\\_Project\\_managment\\_cost\\_time\\_and\\_quality](https://www.researchgate.net/publication/305462896_Project_managment_cost_time_and_quality).
10. Object-Oriented Programming in Computer Science / R. Yilmaz, A. Sezgin, S. Kurnaz, Y. Arslan // Advanced Methodologies and Technologies in Network Architecture, Mobile Computing, and Data Analytics / R. Yilmaz, A. Sezgin, S. Kurnaz, Y. Arslan., 2019. – p. 1439–1451.

<https://doi.org/10.31891/2219-9365-2024-79-6>

УДК: 681.3.06

ПАНЧИШИН Богдан

Національний Університет "Львівська Політехніка"

<https://orcid.org/0009-0009-1366-1602>

e-mail: [panchyshyn33@gmail.com](mailto:panchyshyn33@gmail.com)

## АНАЛІЗ ПРИНЦИПІВ РОБОТИ КОМПОНЕНТІВ ФРЕЙМВОРКА REACT JS

У статті проаналізовані і описані принципи роботи та структура react компонентів. Також наведені приклади їх ефективного застосування на практиці верстки веб-застосунку. Представлені варіанти створення компонентів React, які є основою для його модульної та багаторазово використовуваної архітектури.

Ключові слова: react компоненти, властивості state та props, react js, virtual dom.

PANCHYSHYN Bohdan

Lviv Polytechnic National University

## ANALYSIS OF THE WORKING PRINCIPLES OF REACT JS FRAMEWORK COMPONENTS

This topic is relevant because React is the most popular JavaScript programming language framework. Knowing the principles of building and the structure of react components has the ability to create extremely complex web applications that will be fast, functional and efficient at the same time.

The main focus of React is in managing the state of components and updating the used interface with minimal resource consumption. Below you can see the basic principles in React:

1. Components
2. Virtual DOM
3. Unidirectional data flow
4. JSX

Component reuse makes it easy to create and display similar UI elements that behave the same across applications. This facilitates code reuse and supports modularity and maintainability of the codebase. The component approach in React allows you to create small and reusable blocks of code elsewhere in your web application.

Knowledge of the principles of work and the structure of react components provides opportunities to create functional, simple, fast and optimized web applications that are easily modernized during development.

One of the key advantages of using react components is the ability to break complex and large user interfaces into small blocks that can be used anywhere in the project and more than once. Each component contains its own logic, user interface, and state, which in turn makes codebases easier to maintain, especially in large projects, and operates independently, making debugging, testing, and code analysis easier.

In addition, the examples given and analyzed above will help to better understand the very work of components, their structure, options for creation and application in projects, and how to correctly and successfully manipulate data using states and properties.

Keywords: react components, state properties and props, react js, virtual dom.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Було складно та ризиковано оновлювати інтерфейс користувача шляхом безпосереднього маніпулювання об'єктною моделлю документа (DOM). Рідке використання DOM під час прямих маніпуляцій часто призводило до неефективного рендерингу та труднощів у підтримці відповідного інтерфейсу користувача. Низька продуктивність була типовою проблемою традиційних веб-додатків, особливо коли вони мали справу з великими та динамічними інтерфейсами користувача.

DOM може бути надто сильно модифікований, що призведе до повільного відтворення та затримки роботи користувача. Функціональність інтерфейсу користувача не вдалося оновити на основі змін у стані програми. Відсутність архітектури на основі компонентів перешкоджала здатності розбити інтерфейс користувача на модульні та багаторазово використовувані компоненти у звичайних підходах. Співпраця заважала, а програми було важко масштабувати.

## АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Однією з ключових переваг використання компонентів react – це можливість розбивати складні та великі інтерфейси користувача на маленькі блоки, які можна використовувати в будь-якому місці проекту і не один раз. Кожен компонент містить в собі власну логіку, інтерфейс користувача і стан, що в свій час дозволяє легше підтримувати кодові бази, особливо у великих проектах та працює незалежно, що полегшує роботу в налагодженні, тестуванні та аналізі коду.

Структуру компонента react та його комунікацію між батьківським та дочірнім компонентами, а також застосування та призначення state і props досліджено у цій роботі. Якщо state надають можливість

керувати даними в середині компоненти, то props – забезпечують передачу цих даних між компонентами. Одно-направлений потік в компоненті дозволяє краще керувати проектом. Віртуальний DOM – рендерить тільки ту частину веб-інтерфейсу, яка саме була змінена при взаємодії з користувачем, що забезпечує кращу продуктивність.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

**Метою роботи є:** Дослідити принципи роботи та структуру компонентів React JS.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Ця тема є актуально, оскільки react є найбільш популярним фреймворком мови програмування JavaScript. Знання принципів побудови та структури react компонентів дає можливість створювати надскладні веб-застосунки, які в водночас будуть швидкими, функціональними та ефективними.

Основний фокус React полягає в управлінні станом компонентів та оновленні користувацького інтерфейсу з мінімальними витратами ресурсів. Основні принципи у React:

1. Компоненти – Користувацький інтерфейс розбивається на невеликі незалежні компоненти, кожен компонент представляє собою окрему частину інтерфейсу та містить в собі логіку і представлення про нього.

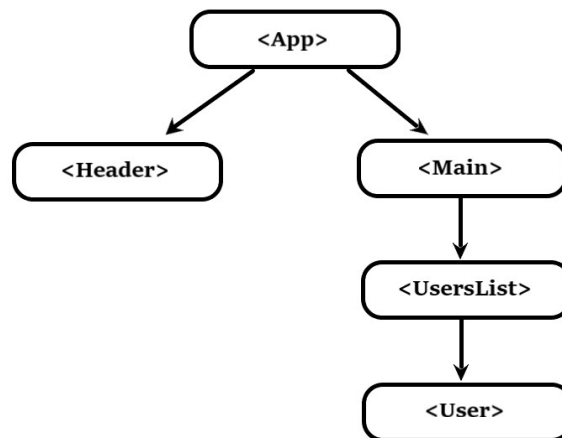


Рис. 1. Структура веб-додатку побудованого на React компонентах

2. Virtual DOM – використовується для ефективного управління оновленням інтерфейсу, замість прямого змінення реального DOM, react створює Virtual DOM та порівнює його з поточним станом для визначення мінімальних змін які потрібно внести.

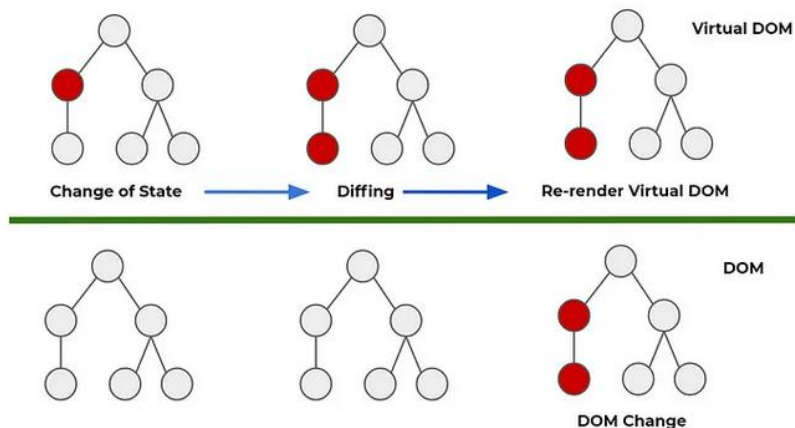


Рис. 2. Принцип роботи react над оновленням інтерфейсу з допомогою Virtual DOM

3. Одно-направлений потік даних – дані в react, як правило передаються вниз по ієрархії компонентів це означає що, зміни в батьківському компонентів також можуть впливати і на дочірні, але не навпаки, це забезпечує передбачуваність та управління при розробці додатків.

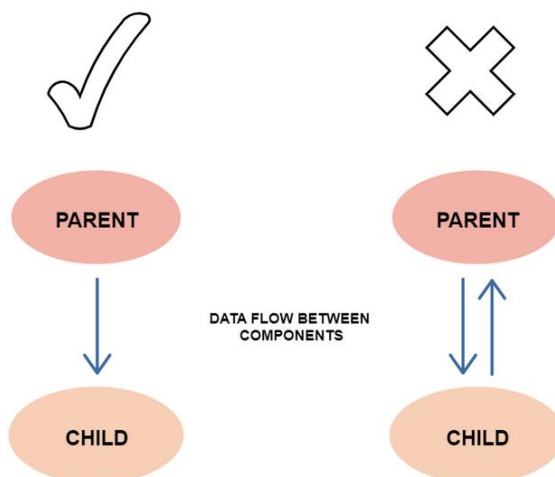


Рис.3. Односпрямований потік даних в react

4. JSX - це синтаксичне розширення JavaScript, яке дозволяє описувати структуру користувацького інтерфейсу в середині коду, це робить код більш зрозумілим та зручним для розробників.

```

CodeWithJSX.jsx X
1 import React from "react";
2
3 const Hello = () => {
4   return (
5     <div>
6       <h1>Hello, i'm using JSX</h1>
7     </div>
8   )
9 }
10
11 export default Hello;

JS CodeWithoutJSX.js X
1 import React from "react";
2
3 const Hello = () => {
4   return React.createElement("div", null,
5     React.createElement("h1", null, "Hello i'm not using JSX..."));
6 }
7
8 export default Hello;
    
```

Рис.4. Порівняння React з JSX і без JSX

Досліджено приклад коду із повторним використанням react компонентів:

**User** - це простий компонент, який приймає текстові атрибути і відображає їх в div.

```

import React from "react";
const User = ({ firstName, lastName }) => {
  return (
    <div>
      <span>{firstName}</span>
      <span>{lastName}</span>
    </div>
  );
}
export default User;
    
```

Компонент **App** імпортує та використовує User кілька разів, передаючи різні текстові значення як атрибути. Відповідно ми також можемо передати не тільки текстові значення, а й інші типи даних.

```

import './App.css';
import User from './comp/User';
function App() {
  return (
    <div className="App">
      <User firstName="John" lastName="Wick" />
      <User firstName="Harry" lastName="Potter" />
      <User firstName="Jimm" lastName="Carry" />
    </div>
  );
}
export default App;
    
```



Повторне використання компонента User полегшує створення та відображення схожих елементів інтерфейсу, які поведуться однаково в різних додатках. Це полегшує повторне використання коду та підтримує модульність і ремонтпридатність кодової бази. Компонентний підхід в React дозволяє створювати невеликі та з можливістю повторно використати блоки коду в іншому місці веб-додатку.

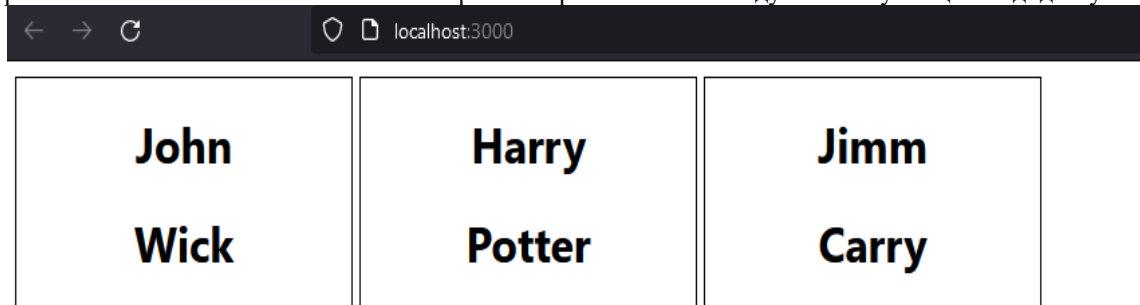


Рис.5. Результат роботи коду

Якщо ми говоримо про структуру компонентів, тут важливо виділити 4 основні критерії:

#### 1. Функціональні компоненти:

Функціональні компоненти – це функції JavaScript, які не мають стану. Вони набагато простіші та зрозуміліші ніж класові компоненти, але з появою React Hooks функціональні компоненти отримали можливість керувати локальним станом і побічними ефектами завдяки хукам useState та useEffect.

```
import React from 'react';
const MyFunctionalComponent = () => {
  return (
    <div>
      /* JSX representing the component's UI */
    </div>
  );
};
```

#### 2. Класові компоненти

Класові компоненти розширюють клас React.Component та входять до сімейства JavaScript. Вони побудовані з використанням синтаксису класу ES6 і надають більш розширені функції, такі як методи життєвого циклу та керування станом. Щоб створити компонент класу, вам потрібно успадкувати клас Component з бібліотеки React. Обов'язково потрібний метод render, він повертає вигляд компонента(JSX).

```
import React, { Component } from 'react';
class MyClassComponent extends Component {
  render() {
    return (
      <div>
        /* JSX representing the component's UI */
      </div>
    );
  }
}
```

#### 3. Props

У React props (реквізити) надають можливість комунікації від батьківського до дочірнього компонента. Іншими словами – це канал зв'язку, який дозволяє обмінюватись інформацією, при чому ця інформація може бути представлена в будь-якому типі даних. Функціональні компоненти отримують props як параметри, а компоненти класу отримують до них доступ через this.props.

```
// Functional Component
const MyFunctionalComponent = (props) => {
  return <div>{props.message}</div>;
};
```

```
// Class Component
```

```
class MyClassComponent extends Component {  
  render() {  
    return <div>{this.props.message}</div>;  
  }  
}
```

#### 4. Стани.

Стан — це вбудована функція React, яка надає можливість компонентам керувати внутрішніми даними та зберігати їх. Він представляє поточний стан компонента та може бути оновлений з часом, зазвичай у результаті взаємодії користувача або асинхронних операцій. Тобто до нього можна отримати доступ і змінити його лише в межах цього компонента або його дочірніх компонентів. Функціональні компоненти використовують хук `useState`, а класові компоненти об'єкт `this.state`.

Розберемо як працювати з властивостями `state` та `props`, на прикладі функціональних компонентів.

```
import React, { useState } from 'react';  
//Parent Component  
const Parent = () => {  
  const [count, setCount] = useState(0); // State властивість  
  
  const incrementCount = () => {  
    setCount(count + 2)  
  };  
  
  return (  
    <div>  
      <h1>Parent</h1>  
      <p>Count: {count}</p>  
      <Child count={count} />  
      <button onClick={incrementCount} >Increment</button>  
    </div>  
  );  
};  
  
//Child Component  
const Child = (props) => {  
  const { count } = props; // Props властивість  
  
  return (  
    <div>  
      <section>  
        <h2>Child</h2>  
        <p>Received Count from Parent: {count}</p>  
      </section>  
    </div>  
  );  
};  
  
export default Parent;
```

#### Батьківський компонент (Parent):

Початковий стан змінної `count` визначено завдяки хуку `useState` та дорівнює 0. У JSX компоненті `Parent` стан `count` відображається з допомогою `{count}`. Створено стрілочну функцію `incrementCount`, яка використовується, щоб оновити стан `count`, при натисканні кнопки. При рендері компонента `Child` стан `count` передається, як властивість (`count={count}`).

#### Дитячий компонент (Child):

Ми отримуємо властивість `count` від `Parent` компонента за допомогою `props`. Потім відображаємо це у `Child` компоненті за допомогою `{count}`.

У цьому прикладі показано, як батьківський компонент може керувати та оновлювати внутрішній стан (count) і передавати його як властивість дочірнім компонентам. В свою чергу дочірній компонент отримане значення використовує, щоб відобразити значення лічильника. Розуміючи різницю між state та props, ми з легкістю забезпечуємо чіткий потік даних і полегшуємо процес повторного використання компонентів. Батьківські компоненти зберігають власний стан (state), а дочірні компоненти отримують дані через властивості(props), що полегшує контроль і зв'язок між компонентами.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Знання принципів роботи та структури react компонентів надає можливості створення функціональних, простих, швидких та оптимізованих веб-додатків, які легко піддаються модернізації в ході розробки.

Окрім того наведені та проаналізовані вище приклади, допоможуть краще зрозуміти саму роботу компонентів, їх структуру, варіанти створення та застосування у проектах і як правильно та успішно маніпулювати даними за допомогою станів(state) та властивостей(props).

### Література

1. Minnick C. ReactJS Foundations. Building User Interfaces with ReactJS [Електронний ресурс] / Chris Minnick // WROX A Wiley Brand. – 2022. – Режим доступу до ресурсу: <https://dl.ebooksworld.ir/books/Beginning.ReactJS.Foundations.Chris.Minnick.Wiley.9781119685548.EBooksWorld.ir.pdf>.
2. ReactJS official documentation [Електронний ресурс] // Meta Platforms. – 2013. – Режим доступу до ресурсу: <https://react.dev/learn>.
3. Dave G. React JS Full Course for Beginners | Complete All-in-One Tutorial | 9 Hours [Електронний ресурс] / Gray Dave. – 2022. – Режим доступу до ресурсу: <https://www.youtube.com/watch?v=RVFAyFWO4go>.
4. Fedosejev A. React.js Essentials [Електронний ресурс] / Artemij Fedosejev // Packt Publishing Ltd. – 2015. – Режим доступу до ресурсу: [https://strukturfondovi.hr/wp-content/uploads/2017/03/9781783551620-REACTJS\\_ESSENTIALS.pdf](https://strukturfondovi.hr/wp-content/uploads/2017/03/9781783551620-REACTJS_ESSENTIALS.pdf).

### References

1. Minnick C. ReactJS Foundations. Building User Interfaces with ReactJS [Electronic resource] / Chris Minnick // WROX A Wiley Brand. – 2022. – Resource access mode: <https://dl.ebooksworld.ir/books/Beginning.ReactJS.Foundations.Chris.Minnick.Wiley.9781119685548.EBooksWorld.ir.pdf>.
2. ReactJS official documentation [Electronic resource] // Meta Platforms. – 2013. Resource access mode: <https://react.dev/learn>.
3. Dave G. React JS Full Course for Beginners | Complete All-in-One Tutorial | 9 Hours [Electronic resource] / Gray Dave. – 2022. – Resource access mode: <https://www.youtube.com/watch?v=RVFAyFWO4go>.
4. Fedosejev A. React.js Essentials [Electronic resource] / Artemij Fedosejev // Packt Publishing Ltd. – 2015. – Resource access mode: [https://strukturfondovi.hr/wp-content/uploads/2017/03/9781783551620-REACTJS\\_ESSENTIALS.pdf](https://strukturfondovi.hr/wp-content/uploads/2017/03/9781783551620-REACTJS_ESSENTIALS.pdf).

<https://doi.org/10.31891/2219-9365-2024-79-7>

UDC 004.056.5

TYMOSHCHUK Dmytro

Ternopil Ivan Puluj National Technical University

<https://orcid.org/0000-0003-0246-2236>

e-mail: [dmytro.tymoshchuk@gmail.com](mailto:dmytro.tymoshchuk@gmail.com)

YATSKIV Vasyi

West Ukrainian National University

<https://orcid.org/0000-0001-9778-6625>

e-mail: [jazkiv@ukr.net](mailto:jazkiv@ukr.net)

## USING HYPERVISORS TO CREATE A CYBER POLYGON

*Cyber polygon used to train cybersecurity professionals, test new security technologies and simulate attacks play an important role in ensuring cybersecurity. The creation of such training grounds is based on the use of hypervisors, which allow efficient management of virtual machines, isolating operating systems and resources of a physical computer from virtual machines, ensuring a high level of security and stability. The paper analyses various aspects of using hypervisors in cyber polygons, including types of hypervisors, their main functions, and the specifics of their use in modelling cyber threats. The article shows the ability of hypervisors to increase the efficiency of hardware resources, create complex virtual environments for detailed modelling of network structures and simulation of real situations in cyberspace.*

*Keywords: Hypervisor, cyber polygon, VMware ESXi, Microsoft Hyper-V, Xen, KVM, Cluster.*

ТИМОЩУК Дмитро

Тернопільський національний технічний університет імені Івана Пулюя

ЯЦКІВ Василь

Західноукраїнський національний університет

## ВИКОРИСТАННЯ ГІПЕРВІЗОРІВ ДЛЯ СТВОРЕННЯ КІБЕРПОЛІГОНУ

*Кіберполігони, які використовуються для тренування фахівців з кібербезпеки, тестування нових технологій захисту та моделювання атак, відіграють важливу роль у забезпеченні кібербезпеки. В основі створення таких полігонів лежить використання гіпервізорів, які дозволяють ефективно управляти віртуальними машинами, ізолюючи операційні системи та ресурси фізичного комп'ютера від віртуальних машин, забезпечуючи високий рівень безпеки та стабільності. В роботі проаналізовано різні аспекти використання гіпервізорів у кіберполігонах, включаючи типи гіпервізорів, їхні основні функції, а також специфіку застосування у моделюванні кіберзагроз та розгортанні кіберполігонів. Показано здатність гіпервізорів підвищувати ефективність використання апаратних ресурсів, створювати складні віртуальні середовища для детального моделювання мережевих структур та імітації реальних ситуацій в кіберпросторі.*

*Ключові слова: Гіпервізор, кіберполігон, VMware ESXi, Microsoft Hyper-V, Xen, KVM, Cluster.*

### INTRODUCTION

Information technology is constantly evolving, creating new challenges for cybersecurity. Cyber polygon, or environments for modelling cyber threats, are becoming a necessary tool for training specialists, testing system security, and researching new methods of countering attacks. The basis for creating such training grounds is the use of hypervisors.

A hypervisor is software that manages the operation of virtual machines. It allows you to create and effectively manage virtual machines, each with its own operating system and applications.

A cyber polygon is a virtual environment used to train cybersecurity professionals, test new security technologies, simulate attacks, and practice incident response scenarios. It can include virtual machines, network equipment, storage systems and other components that simulate real-world infrastructure.

### MAIN PART

The hypervisor distributes physical resources (such as CPU, memory, disk space, etc.) among several virtual machines, ensuring their independence and stability. This makes it possible to run several different operating systems on the same physical server, which increases the flexibility and efficiency of using hardware resources. The hypervisor ensures that virtual machines cannot interact with each other in an undesirable way or affect the host system, which provides a high level of security and stability. It can include various functions, such as managing virtual networks, data storage, and I/O processing. Thanks to its architecture, a hypervisor allows you to create and manage virtualised environments that can be used for a variety of purposes.

There are two types of hypervisors: type 1 (bare-metal) and type 2 (hosted).

Type 2 runs on top of an operating system that is already installed on a physical server (Figure 1). Examples: VMware Workstation, Oracle VirtualBox, Parallels Desktop, VMware Fusion, UTM, QEMU

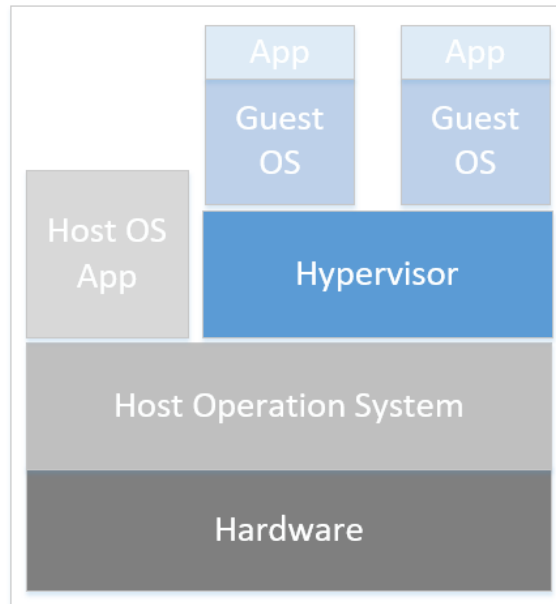


Fig. 1 Type 2 hypervisor architecture

Type 1 runs directly on the hardware without the need for an underlying operating system (Figure 2). Examples: VMware ESXi, Microsoft Hyper-V, Xen, and KVM.

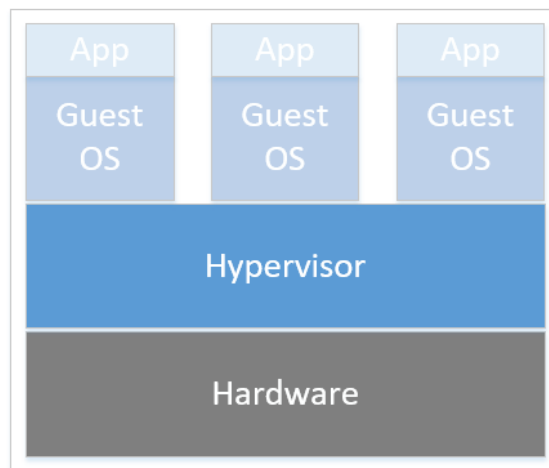


Fig. 2 Type 1 hypervisor architecture

VMware ESXi is a high-performance Type 1 hypervisor that is small and fast. The hypervisor delivers high performance while minimising virtualisation overheads thanks to its lightweight architecture [1]. High level of scalability. ESXi allows you to efficiently use server resources by running many virtual machines on a single server and automatically managing them through vCenter Server, allowing you to create large virtualised environments.

One of the main advantages of VMware ESXi is its high reliability. The hypervisor supports high availability features such as High Availability and Fault Tolerance. The vMotion function allows you to move virtual machines from one server to another without interrupting their operation, which is important for ensuring the smooth operation of services. The DRS (Distributed Resource Scheduler) tool automatically balances the load between servers by moving virtual machines according to available resources. Storage vMotion technology allows you to move virtual machines between different datastores without interrupting their operation. VMware vSAN software-defined storage system allows you to create virtualised storage based on local server disks. vSAN combines disc resources from multiple servers into a single cluster, creating scalable and high-performance storage for virtual machines. All of these are part of the VMware vSphere platform functionality and are critical technologies for ensuring the continuous operation of virtualised environments. VMware ESXi supports both Intel VT (Intel Virtualisation Technology) and AMD-V (AMD Virtualisation).

Intel VT and AMD-V are hardware-based virtualisation technologies developed by Intel and AMD, respectively. Intel VT and AMD-V are critical components for virtualisation performance in modern hypervisors. They leverage these hardware capabilities to improve virtualisation performance, security, and efficiency.

VMware ESXi can provide a solid foundation for a cyber polygon with its high-performance virtualisation capabilities, scalability, and reliability. It allows you to create complex and resource-intensive virtual environments that can mimic real-world network structures with different types of servers, operating systems, and applications. This allows you to configure various cyber threat and attack scenarios that can be tested without risking real systems. Using features such as vSphere vMotion and VMware vSAN, you can ensure that the cyber range is always running, even when performing updates or moving virtual machines between servers, which is important for long training and education sessions. VMware ESXi allows you to create complex network topologies with different network segments and access policies. This allows you to recreate real-world enterprise network conditions, including data distribution and protection. VMware ESXi also supports snapshots of virtual machines, which allows you to quickly return to previous system states after tests or attacks.

Hyper-V is a Type 1 hypervisor developed by Microsoft that allows you to create and manage virtual machines [2]. It is part of the Windows Server and Windows 10/11 (Pro, Enterprise, and Education) operating systems, providing users with the ability to efficiently virtualise servers and workstations. Hyper-V directly controls resources such as CPU, memory, and network interfaces, distributing them among virtual machines. Live Migration allows you to move running virtual machines between physical servers without interrupting their operation. Hyper-V Replica provides replication of virtual machines between two different sites, allowing for backup and disaster recovery. Despite its many benefits, Hyper-V does have some limitations. Although Hyper-V supports a variety of operating systems, some versions of Linux or other operating systems may have limited support or require additional configuration.

The Hyper-V hypervisor can serve as a solid foundation for a cyber polygon with its powerful virtualisation capabilities, integration with other Microsoft products, and flexibility in managing virtual environments. With Hyper-V, you can easily create isolated virtual environments in which to deploy a variety of operating systems and applications to simulate and defend against attacks. Integration with Active Directory allows you to control access to resources and manage users. Hyper-V also supports replication and backup functions, which allows you to preserve the state of your environment, quickly restore it after test attacks, or create backups for analysis. Hyper-V virtual switches allow you to configure complex network topologies by simulating various types of network interactions. With Hyper-V, you can implement a scalable cyber polygon that can be used for both individual training sessions and complex team exercises.

Xen is an open-source Type 1 hypervisor developed at the University of Cambridge and is one of the most popular virtualisation solutions in the world [3]. It allows you to run multiple operating systems on a single physical server, efficiently sharing hardware resources between them. Xen runs directly on the server hardware, providing a high level of performance and security by isolating virtual machines from each other.

One of the key features of Xen is its ability to support different virtualisation models, including paravirtualisation and hardware virtualisation. Paravirtualisation allows operating systems running on Xen to be optimised for virtualisation, which improves performance. Hardware virtualisation, on the other hand, allows you to run unmodified operating systems such as Windows through the use of special hardware extensions of the Intel VT or AMD-V processor. Xen supports high availability and live migration of virtual machines, allowing them to be moved between physical servers without interruption. Xen provides a command-line interface and tools for managing virtual machines, allowing administrators to effectively monitor and configure the virtualised environment.

Xen can be a good foundation for creating a cyber polygon due to its flexibility, open source, and ability to support different virtualisation models. Since Xen supports both paravirtualisation and hardware virtualisation, it allows you to create cyber polygons with different operating systems, including virtualisation-optimised versions of Linux and unmodified versions of Windows. Xen's flexibility allows you to create complex network topologies and scenarios to simulate real-world attacks. With support for live migration of virtual machines and high availability, Xen ensures the continuity of the cyber range, even during updates or the transfer of virtual machines to other physical servers. This allows you to conduct long-term training and testing without the risk of stopping the environment.

KVM (Kernel-based Virtual Machine) is a Type 1 hypervisor that is part of the Linux kernel and allows you to use hardware virtualisation features to create and manage virtual machines [4]. As a virtualisation solution, KVM delivers high performance and efficiency through tight integration with the Linux kernel, allowing it to use existing system components and resources. The main component of KVM is the kernel module, which turns the Linux kernel into a hypervisor that manages virtual machines.

One of the key advantages of KVM is that it uses standard Linux tools and mechanisms to manage virtual machines. This includes tools for monitoring, network configuration, and resource management. KVM supports a variety of operating systems, including different versions of Linux, Windows, and other OSes, making it a versatile virtualisation solution.

KVM integrates with other projects and technologies, such as libvirt, which provides interfaces for managing virtual machines via the command line or graphical interfaces.

Another important part of KVM is QEMU. QEMU (Quick Emulator) is a versatile open source emulator that allows you to create virtual machines by emulating various hardware [5].

QEMU provides emulation of a wide range of hardware components, such as processors, network cards, hard drives and other peripherals. One of the main advantages of QEMU is its ability to support different processor architectures such as x86, ARM and others. QEMU can run in both full emulation mode and hardware virtualisation mode to improve performance. In full emulation mode, QEMU provides emulation of all hardware, allowing you to run virtual machines regardless of whether the physical host supports hardware virtualisation.

KVM is an excellent basis for creating a cyber polygon due to its integration with the Linux kernel and powerful virtualisation capabilities. As a specialised environment for modelling and analysing cyber threats, a cyber polygon requires a stable, scalable and reliable platform for deploying virtual machines and simulating attacks. KVM, as part of the Linux kernel, provides high performance and efficiency through the use of hardware virtualisation via Intel VT or AMD-V. This allows you to create virtual machines with minimal overhead, which is especially important for large and resource-intensive environments such as cyber polygons. QEMU, when used in conjunction with KVM, provides the ability to emulate different hardware, allowing you to create virtual machines with different characteristics and configurations. This is useful for modelling various attack and defence scenarios within a cyber polygon. QEMU also supports virtual machine snapshots, which allows you to save system states and quickly restore them if necessary. The migration capability also helps to ensure high availability and scalability of the cyber polygon, adapting it to changing needs.

Type 1 hypervisors, which work directly with the hardware, provide higher performance than type 2 hypervisors because they do not require an underlying operating system. This is critical for cyber polygons, where high data processing speeds and realistic modelling of cyber threats are required. Due to these advantages, Type 1 hypervisors are ideal for creating a cyber polygon, providing a high level of performance, security, scalability and stability that are critical for effective testing and modelling of cyber threats.

One of the most important aspects of creating a cyber polygon is the use of routers with firewall functionality in the test environment. Routers such as the Cisco CSR1000v, IPFire, OPNsense, pfSense, MikroTik CHR, Juniper Cloud-Native Router, FortiGate VM, and Palo Alto VM-Series are a solid choice for a cyber polygon due to their advanced security, configuration, and integration with other systems [7][8][9][10][11][12][13]. Using these routers with firewall capabilities allows you to create a variety of scenarios for testing and training in the cyber lab. Their capabilities help to create complex and realistic network environments that allow you to effectively model and analyse cyber threats.

Figure 3 shows an architectural diagram of a cyber polygon created on the basis of a cluster of hypervisors.

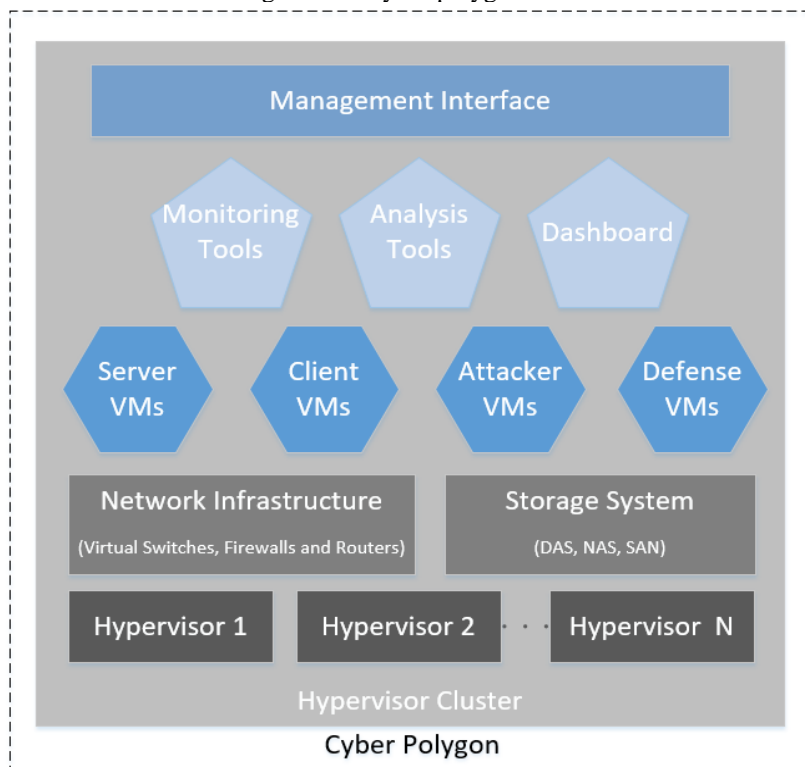


Fig. 3 The architecture of the cyber polygon

The centralised management interface allows you to manage all virtual machines, configure network configurations, and monitor the status of the entire hypervisor cluster. A hypervisor cluster consists of several physical servers with hypervisors. There are several types of virtual machines in the architecture: server machines that simulate various server roles, such as web servers, databases, file servers; client machines that simulate user endpoints, such as computers and mobile devices; attack machines configured with tools and scripts to perform cyber attacks; and machines that are equipped with security tools to detect and mitigate attacks. The network infrastructure includes virtual switches, routers and firewalls that control and monitor data flows between virtual machines and separate network segments. The storage system includes DAS (Direct-Attached Storage), NAS (Network-Attached Storage), and SAN (Storage Area Network), which provide data storage for virtual machines. Monitoring, analysis tools and a dashboard allow real-time monitoring and analysis of network activity and incident response by visualising the security status.

This architecture allows for the creation of diverse and realistic scenarios for cybersecurity training, providing the flexibility and scalability to conduct complex tests in a controlled environment.

## CONCLUSIONS

Using a Type 1 hypervisor to create a cyber polygon is an effective approach that allows you to simulate complex and realistic cyber threat scenarios. The main advantages are high performance, reliability and isolation of virtual machines, which creates a secure environment for testing and training. Type 1 hypervisors provide the ability to create scalable and adaptive cyber polygons that can be used to train cybersecurity professionals, conduct research and test new security technologies.

Thus, a Type 1 hypervisor is a key element for creating effective and secure cyber polygons that can be used in various industries, from educational institutions to corporate environments.

## References

1. VMware vSphere Documentation. URL: <https://docs.vmware.com/en/VMware-vSphere/index.html>
2. Hyper-V technology overview. URL: <https://learn.microsoft.com/en-us/windows-server/virtualization/hyper-v/hyper-v-technology-overview>
3. Xen project software overview. URL: [https://wiki.xenproject.org/wiki/Xen\\_Project\\_Software\\_Overview#PV\\_.28x86.29](https://wiki.xenproject.org/wiki/Xen_Project_Software_Overview#PV_.28x86.29)
4. Kernel Virtual Machine. URL: [https://linux-kvm.org/page/Main\\_Page](https://linux-kvm.org/page/Main_Page)
5. QEMU documentation. URL: <https://www.qemu.org/docs/master/>
6. Introduction – OPNsense documentation. URL: <https://docs.opnsense.org/intro.html>
7. IPFire documentation. URL: <https://www.ipfire.org/docs>
8. Introduction pfSense Documentation. Netgate Documentation. URL: <https://docs.netgate.com/pfsense/en/latest/general/index.html>
9. Cisco Cloud Services Router 1000v. URL: <https://www.cisco.com/c/en/us/support/routers/cloud-services-router-1000v/model.html>
10. Cloud hosted router - mikrotik documentation. URL: <https://help.mikrotik.com/docs/display/ROS/Cloud+Hosted+Router.+CHR>
11. Juniper Cloud-Native Router (JCNR). URL: <https://www.juniper.net/documentation/product/us/en/juniper-cloud-native-router/>
12. FortiGate private cloud 7.4. Fortinet Document Library. URL: <https://docs.fortinet.com/product/fortigate-private-cloud/7.4>
13. VM-Series. Palo Alto Networks Tech Docs Home. URL: <https://docs.paloaltonetworks.com/vm-series>



<https://doi.org/10.31891/2219-9365-2024-79-8>

УДК 004.78

БЕРНАТОВИЧ Анатолій

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

<https://orcid.org/0000-0002-8403-6363>

e-mail: [iracotnar1972@gmail.com](mailto:iracotnar1972@gmail.com)

## АНАЛІЗ ВИМОГ ДО РОЗРОБКИ УНІВЕРСАЛЬНОГО ФІЗИЧНОГО РУШІЯ ІЗ ПІДТРИМКОЮ ПЛАГІНІВ, КОНФІГУРАЦІЄЮ ПАЙПЛАЙНУ ТА АПАРАТНИМ ПРИСКОРЕННЯМ

Ця стаття присвячена створенню вимог до гнучкого та розширюваного фізичного рушія, здатного підтримувати модулі (плагіни), конвеєр обробки, що конфігурується, та інтуїтивно зрозумілий редактор для створення моделей, використовуючи при цьому апаратне прискорення для підвищення продуктивності. Представлено архітектуру фізичного рушія з чітким розмежуванням між основними компонентами та модулями-плагінами. Запропоновано структуру проекту, що включає чіткий розподіл завдань, з основними компонентами, модулями та графічним інтерфейсом, що зберігаються в різних директоріях. Обговорюється важливість інструментів тестування та налагодження, таких як CMake, для створення та тестування рушія на різних платформах.

Ключові слова: фізична симуляція, фізичний рушій, модульність, симуляція, C++.

BERNATOVYCH Anatolii

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

## ANALYSIS OF REQUIREMENTS FOR THE DEVELOPMENT OF A UNIVERSAL PHYSICS ENGINE WITH PLUG-IN SUPPORT, PIPELINE CONFIGURATION, AND HARDWARE ACCELERATION

The aim of the article is the explanation of the physics engine that allows for extensions through modules (plug-ins), has a pipeline of processing that can be modified, has an editor for defining the shapes of the models, and can use the hardware acceleration for computing. This will allow the researchers to apply it in different contexts like computer graphics, animation, scientific modelling, or engineering. According to the article, the extensibility of a physics engine means that the developers can include one or more new physical states or alter an existing one without altering the engine.

It is proved that the modular approach has several advantages: 1. It is easy to add new physical laws and phenomena to the engine without changing the structure of the engine. Special modules for various research or industrial requirements can be built, thus tailoring the engine to the respective domain. Pre-existing modules can therefore be reused from one project and utilized in other projects thereby enhancing the integration process and shortening the time for development. This improves scalability since the engine can expand by adding or replacing existing modules to accommodate new technologies or new requirements. The authors also explain visualization and interactive editing as factors that are essential in enhancing the effectiveness of physics engines. They also suggest a project structure with a core physics engine, modules, editor, user interface, test system, and compatibility with different physical models. Designing a configurable physics engine that can be effectively used for various games and situations is a task that can only be accomplished by a dedicated programmer with profound knowledge of the physics mechanisms. Including distributed computing features, annotated editing, hardware enhancing features, and utilizing techniques like MPI, OpenGL, Vulkan, SFML to improve this flexibility. Therefore, it is possible to list the following beneficial aspects of a physics engine that must be productive, flexible, unlike, and competitive in different researching and entertaining usages.

As for the further developments of the topic for the research, the improvements towards the integration with other systems, development of the new physical models and the usage of the various resource optimization methods like parallel computations for the better performance of the engine have to be mentioned. This makes new opportunities in the sphere of physical modelling, which is critical for science, education, entertainment, and development.

Keywords: physical simulation, physical engine, modularity, simulation, C++.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У контексті фізичного моделювання основною потребою є можливість зібрати модель і провести ряд симуляцій для дослідників або розробників. Будучи частиною додатків, фізичні рушії забезпечують галузі від комп'ютерної графіки до наукового моделювання та ігор. Багато зусиль докладається для того, щоб зробити моделі реалістичними, коли мова йде, зокрема, про фізичні процеси, і тому акцент робиться на графічних моделях. Однак фізика, яка включає в себе різні форми фізичного моделювання, повинна бути продуктивною у своїй роботі, гнучкою і розширюваною за необхідності, що є дещо складнішим завданням. На практиці, створення фізичного рушія є нетривіальною і, насправді, досить цікавою роботою для програмістів з точки зору програмної інженерії. Наприклад, сучасний світ не може обійтися без фізичного рушія для численних додатків в індустрії розваг, анімації та моделювання, симуляції в науці та техніці, а також численних обчислювальних алгоритмів в інженерії. Оскільки інструмент буде використовуватися для

роботи з різними типами вимог і специфікацій, вкрай важливо розробити структуру, що дозволяє вносити зміни «на льоту». Такий функціонал, як розширюваність або здатність використовувати апаратне забезпечення для досягнення вищої швидкості моделювання, може кардинально змінити уявлення про те, на що мають бути здатні фізичні рушії в будь-якій галузі, якщо вони є актуальними для всієї галузі.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

В роботі [1] наведено дані про важливість розробки власного фізичного рушія з огляду на недоліки наявних. Авторами вперше поставлено це питання тому інших досліджень з цього питання немає.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є: аналіз та моделювання вимог для створення фізичного рушія, який підтримує розширення через модулі (плагіни), має налаштований пайплайн обробки, інтуїтивний редактор для створення моделей і підтримує апаратне прискорення обчислень.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Часто надається великого значення розширюваність сучасного фізичного рушія. Рушій повинен мати можливість включати «модулі» (плагіни), які вводять більш реальні фізичні стани або змінюють існуючі. Це дає можливість розробникам додавати нові функції та компоненти, наприклад, модель демпфера пружинної маси, що використовується в багатьох симуляціях деформівних тіл, без необхідності модифікувати саме ядро рушія. Особи, які беруть участь у процесі розробки програмного забезпечення, можуть використовувати такі мови, як C++ [2, 3], які надають більше контролю та універсальності у написанні коду. API рушія має включати елементи для створення, реєстрації та використання плагінів. Кожен плагін реалізується як збірка, яка може включати набір компонентів, що розширюють можливості плагіна. Наприклад, розробник, який працює над моделлю автомобіля, може створити плагін, який забезпечує фізичні характеристики, такі як тертя, демпфування та інші, необхідні при моделюванні автомобіля. До речі, інший розробник, який моделює літак, може розробити інший плагін, який враховує аеродинаміку, наприклад. Це було видно в архітектурі фізичного рушія, де Pipeline є ядром. Через нього програма регулює кроки, які виконуються в процесі симуляції. Конфігурація конвеєра обробки фактично пропонує варіант, в якому будуть використані реальні фізичні моделі. Це дуже корисно, оскільки деякі моделі можуть бути працездатними, незважаючи на виключення деяких фізичних ефектів. Наприклад, візьмо створення програми, яка навчатиме пілота навичкам польоту; можна вирішити «вимкнути» тертя для програми. Як і у випадку з моделюванням, визначено місце візуалізації в повному комплексі фізичної симуляції. Редактор моделі та конвеєр процесів будуть інтерактивними, тобто дослідники зможуть виконувати роботу швидше, не звертаючи уваги на те, як працює програмування. Там, де це можливо, редактор повинен бути спроектований таким чином, щоб створити новий екземпляр фізичного об'єкта, задати його властивості та зрозуміти результати цієї дії.

Конвеєр фізичної обробки дозволить нам досліджувати різні фізичні моделі, виключаючи або включаючи окремі фізичні закони. Ось чому ця функція зручна в сценаріях, де характеристики, висвітлені вище, не є життєво важливими для оцінки змодельованого результату. Це означає, що вона допомагає дослідникам зосередитися на реальній фізиці своїх моделей, а не на аспекті симуляції, що не є критичним і не займає багато часу.

Конвеєр відносно чітко сформульований у своїх фізичних моделях і факторах, які можуть бути налаштовані відповідно до потреб користувача для декількох версій одночасно, що дає платформу для альтернативних розробок. Наприклад, якщо вчений-експериментатор, який досліджує рух небесного тіла, хоче виключити фактор тертя, то так само буде виключена статистика контакту у вакуумі або порожнечі. Фізичний рушій тепер фактично є набором елементів, і одним з таких початкових елементів є інтегрований редактор моделей і конвеєрів, який є основою для побудови конфігурації симуляції.

Редактор повинен дозволити користувачеві через веб-інтерфейс описувати тіла, їхні характеристики та взаємодії в фізичній симуляції. Користувач повинен вибрати фізичні властивості кожного об'єкта, такі як маса, пружність, тертя тощо. За допомогою інтерактивної демонстрації слід показати користувачам, як обрані ними параметри впливають на результат фізичної симуляції. Після запуску симуляції користувач отримає детальний звіт про результати, що дозволить оцінити правильність обраних параметрів. Редактор повинен відображати всі параметри, що беруть участь у симуляції, такі як сила, швидкість, напруження тощо, а також діаграми, такі як діаграми сили, вектор швидкості, тенденції деформації тощо. Редактор повинен бути більш-менш технічно орієнтованим - основною цільовою групою є науковці та інженери, які не обов'язково є гуру комп'ютерної графіки або програмування високого рівня.

Отже, для того, щоб досягти найбільшої фактичної продуктивності обчислень, необхідно інтегрувати в систему апаратне забезпечення, яке має кращу швидкість. Серед них - «командні» операції центрального процесора, який скорочено називається CPU, та графічного процесора, або скорочено GPU. Графічний процесор є кращим для розпаралелення обчислень, пов'язаних з фізичним рушієм, наприклад, при пошуку

колізій або при відображенні фізичних атрибутів для візуалізації і т.д. Це може бути записано як прискорення апаратного забезпечення для використання OpenGL, Vulkan, DirectX означає, що для підвищення швидкості роботи програм, які використовують ці графічні API, залучаються додаткові обчислювальні ресурси відеокарти. Тобто, частина навантаження з центрального процесора переноситься на відеокарту, яка спеціалізується на обробці графіки. Це дозволяє програмам працювати швидше і ефективніше. Це дозволить нам робити інтенсивні обчислення на графічних процесорах, оскільки, як було показано раніше, графічні процесори працюють краще, ніж центральні процесори в таких обчисленнях. Сучасні фізичні рушії повинні працювати на різних платформах: починаючи зі стаціонарних терміналів, таких як ПК, ноутбуки, смарт-телевізори, і переходячи до порівняно недавно класифікованих стаціонарних терміналів, включаючи гарнітури віртуальної/доповненої реальності, ігрові консолі, портативні термінали, такі як смартфони та планшети. Існує кілька підходів до вирішення цього завдання: використання спеціалізованих графічних API (OpenGL), універсальних мультимедійних бібліотек (SFML, SDL) або ж більш абстрактних, незалежних від платформи інструментів. Обробка великих і деталізованих моделей вимагає значних обчислювальних ресурсів і може потребувати спеціальних оптимізацій, подібних до тих, що використовуються при розробці дистрибутивів. Іншим інструментом є інтерфейс передачі повідомлень, часто згадуваний як MPI; він корисний при створенні програм, призначених для паралельних обчислень на різних комп'ютерах. Він корисний тим, що розбиває роботу на частини, які можна розмістити на різних вузлах, щоб підвищити ефективність роботи, і в той же час оптимізувати використання ресурсів.

Зупинимось на виборі інструментів та середовища розробки:

1. Редактор коду: Visual Studio Code - відмінний вибір завдяки широкій підтримці C++ [4], дебагуванню та великій кількості розширень. Компілятор: Microsoft Visual C++ - інтегрований у Visual Studio Code і забезпечує високу продуктивність. Система побудови: CMake - гнучка система, яка дозволяє легко керувати проектом і генерувати файли проекту для різних платформ. Система контролю версій: Git - для ефективного керування кодом і співпраці. Розподілені обчислення: стандарт MPI та бібліотеки на кшталт MPICH або OpenMPI для розподілених обчислень.

2. Структура проекту:

Спочатку ідея майбутнього проекту була така:

```
physics_engine/  
├── include/  
│   ├── physics_engine.h  
│   ├── module.h  
│   └── deformable_body_module.h  
├── src/  
│   ├── window.cpp  
│   ├── physics_engine.cpp  
│   ├── module.cpp  
│   └── deformable_body_module.cpp  
└── CMakeLists.txt
```

У Physics\_engine.h має міститися оголошення класу PhysicsEngine, який буде керувати всією симуляцією, оголошення методів для ініціалізації, оновлення та рендерингу, оголошення приватних членів для зберігання інформації про модулі та інші параметри симуляції.

physics\_engine.cpp відповідатиме за реалізацію методів класу PhysicsEngine, створення екземплярів модулів та їх додавання до вектора модулів. Логіка оновлення симуляції (викликає метод update для кожного модуля). Може містити функції для завантаження конфігурацій, збереження стану тощо.

module.h міститиме оголошення абстрактного базового класу Module. Віртуальні методи, які повинні бути реалізовані в конкретних модулях (наприклад, update, render).

module.cpp матиме деякі загальні реалізації методів для всіх модулів або допоміжні функції.

deformable\_body\_module.h міститиме оголошення класу DeformableBodyModule, який успадковується від Module, оголошення методів, специфічних для деформації тіл (наприклад, додавання сили, оновлення сітки).

deformable\_body\_module.cpp сприятиме реалізації методів класу DeformableBodyModule. Використання математичних моделей для симуляції деформації (наприклад, метод скінчених елементів).

CMakeLists.txt - файл для конфігурування проекту за допомогою CMake [5]. Визначення залежностей, компілятора, налаштування лінкера та інших параметрів.

Наступні причини пояснюють, чому було розроблено цю модель: Кожен файл належить до того сегменту системи, для якого він призначений. Базовий клас Module є основою структури модулів та різноманітних інтерфейсів між ними. Основною концепцією тут є те, що приєднання нових модулів може бути зроблено

дуже простим способом, шляхом зміни вихідних даних симуляції. CMake використовується для побудови проекту, і це робиться особливим чином, а саме в автоматичному режимі.

З цієї причини його структура не надто складна і є більш-менш базовою, а тому її легко створювати, а потім адмініструвати. Є ще одна перевага для невеликих систем, де обсяг коду досить малий: така структура проекту вигідна для переміщення системи та навігації.

Однак така структура не гарантує чіткого розподілу завдань щодо ядра рушія, модулів, графічного інтерфейсу користувача і так далі. Нелегко розширювати цю структуру далі, оскільки вона посилає неправильні сигнали групам виконавців. Наприклад, всі вихідні файли за домовленістю розміщуються в одному каталозі під назвою «src», що може стати незручним у випадку з великими об'єктами. Чим більше модулів, функцій або навіть елементів графічного інтерфейсу використовується, тим більше зростає складність системи, що призводить до ускладнення структури управління.

Тому структура була розширена та деталізована для поліпшення організації, модульності та масштабованості фізичного рушія з огляду на обмеження попередньої версії:

```
/PhysicsEngine
/src
  /core
    PhysicsEngine.cpp
    RigidBody.cpp
    SoftBody.cpp
    Force.cpp
    Collision.cpp
  /modules
    SpringMassPlugin.cpp
  /editor
    EditorGUI.cpp
    PipelineEditor.cpp
    Visualization.cpp
  /include
    PhysicsEngine.h
    RigidBody.h
    SoftBody.h
    Force.h
    Collision.h
    Module.h
  /plugins
    SpringMassPlugin.dll
  /tests
    TestPhysicsEngine.cpp
  /CMakeLists.txt
```

/PhysicsEngine - коренева директорія проекту, де зберігаються всі основні компоненти фізичного рушія, включаючи вихідні файли, заголовкові файли, модулі, плагіни, тести та конфігураційні файли.

/src - директорія містить всі вихідні коди рушія, розділені на підкатегорії для зручності.

/core містить основні компоненти фізичного рушія.

1. PhysicsEngine.cpp:

- Реалізація основного класу рушія (PhysicsEngine), який керує всіма фізичними об'єктами та обчисленнями. Включає методи для запуску симуляцій, оновлення фізики кожного кадру, керування об'єктами та взаємодії між ними.

2. RigidBody.cpp:

- Реалізація класу RigidBody, який представляє жорстке тіло в симуляції. Включає параметри, такі як маса, швидкість, імпульс, та методи для обчислення взаємодії між твердими тілами.

3. SoftBody.cpp:

- Реалізація класу SoftBody, який представляє деформівне тіло. Включає параметри, такі як еластичність, в'язкість, а також методи для обчислення внутрішніх і зовнішніх сил, які впливають на деформівне тіло.

4. Force.cpp:

- Реалізація класу Force, який представляє різні типи сил (наприклад, гравітація, пружні сили). Включає методи для додавання, видалення та обчислення дії сил на фізичні об'єкти.

5. Collision.cpp:

- Реалізація класу Collision, який керує обчисленням зіткнень між об'єктами. Включає методи для виявлення зіткнень та обробки реакцій на ці зіткнення (наприклад, відскок або злипання). /modules містить реалізації модулів (плагінів), які можуть розширювати функціональність рушія.

1. SpringMassPlugin.cpp:

- Реалізація модуля для роботи з пружинно-масовою моделлю. Включає клас, що реалізує поведінку пружин і мас, взаємодію між ними та реєструє нові компоненти у фізичному рушії. /editor містить код для редактора моделей і пайплайну обробки.

1. EditorGUI.cpp:

- Реалізація основного інтерфейсу користувача (GUI) редактора. Включає елементи для створення та налаштування моделей, запуску симуляцій, а також для керування пайплайном обробки.

2. PipelineEditor.cpp:

- Реалізація редактора пайплайну обробки. Включає інструменти для додавання, видалення і конфігурації різних етапів обробки фізичної симуляції (наприклад, обчислення сил, інтеграція руху, зіткнення).

3. Visualization.cpp:

- Реалізація модуля для візуалізації симуляційних результатів. Включає методи для рендерингу моделей, відображення результатів симуляції в реальному часі.

/include містить заголовкові файли для всіх основних класів і компонентів.

1. PhysicsEngine.h:

- Оголошення класу PhysicsEngine та його основних методів. Включає оголошення всіх необхідних змінних і функцій, пов'язаних з управлінням фізичними об'єктами та симуляціями.

2. RigidBody.h:

- Оголошення класу RigidBody, включаючи його властивості та методи. Визначає інтерфейс для роботи з жорсткими тілами.

3. SoftBody.h:

- Оголошення класу SoftBody. Визначає інтерфейс для роботи з деформівними тілами, включаючи методи для розрахунку деформацій.

4. Force.h:

- Оголошення класу Force, включаючи його типи (гравітація, пружність) та методи для їх обчислення.

5. Collision.h:

- Оголошення класу Collision, який містить методи для обробки зіткнень між фізичними об'єктами.

6. Module.h:

- Оголошення базового класу Module, який використовується для створення плагінів. Визначає інтерфейс для розширення функціональності фізичного рушія за допомогою модулів.

/plugins - директорія для збереження зібраних модулів (плагінів).

1. SpringMassPlugin.dll:

- Зібраний плагін, який можна динамічно підключати до рушія для розширення його функціональності. /tests містить тести для перевірки роботи рушія.

1. TestPhysicsEngine.cpp:

- Реалізація тестів для основних компонентів фізичного рушія. Включає юніт-тести для перевірки правильності роботи ядра рушія, фізичних моделей, модулів та взаємодії між ними.

CMakeLists.txt - конфігураційний файл для системи збірки CMake. CMakeLists.txt містить інструкції для збірки проєкту, включаючи підключення залежностей, компіляцію вихідних файлів, збірку плагінів та створення виконуваних файлів. Включає команди для створення бібліотек, підключення зовнішніх бібліотек, конфігурації тестів і налаштувань для кросплатформенності.

Ця структура забезпечує високу модульність, чітку організацію коду і полегшує розширення проєкту.

Кожен основний компонент рушія (ядро, модулі, редактор) ізольований в окремій директорії, що сприяє кращій організації коду та полегшує його розширення. Логічно пов'язані файли розташовані разом.

Наприклад, всі базові класи для фізичних об'єктів знаходяться в папці core, а файли GUI — в editor. Така структура легше масштабується, коли потрібно додавати нові модулі, компоненти або функції. Це також полегшує підтримку та тестування окремих частин проєкту. Наявність окремої папки tests дозволяє чітко організувати тести для різних частин коду, що важливо для підтримки якості проєкту.

Зважаючи на оновлену структуру проєкту, пропонуємо вимоги до реалізації основних модулів і функціональності:

1. physics\_engine.h/cpp

Основні компоненти:

- Клас PhysicsEngine для керування симуляцією:

- Основний клас рушія, що відповідає за керування симуляцією, зберігання та управління всіма фізичними об'єктами, а також інтеграцію з різними модулями.

- Функції для додавання та видалення об'єктів:

- Метод AddObject(PhysicalObject\* obj) додає фізичний об'єкт до симуляції.

- Метод `RemoveObject(PhysicalObject* obj)` видаляє фізичний об'єкт із симуляції.
- Функція для виконання кроку симуляції:
  - Метод `StepSimulation(float deltaTime)` виконує один крок симуляції, оновлюючи стан всіх фізичних об'єктів згідно з переданим часом (`deltaTime`).
- Інтерфейс для завантаження плагінів:
  - Метод `LoadPlugin(const std::string& pluginPath)` завантажує плагін (модуль) із вказаного шляху.
  - Метод `UnloadPlugin(const std::string& pluginName)` вивантажує плагін.
- 2. `module.h/cpp`  
Основні компоненти:
  - Абстрактний базовий клас для модулів:
    - Клас `Module` є абстрактним базовим класом, від якого успадковуються всі модулі. Визначає інтерфейс, який мають реалізувати всі модулі.
  - Віртуальні функції для ініціалізації, оновлення та рендерингу:
    - `virtual void Init(PhysicsEngine* engine) = 0`; ініціалізація модуля з доступом до фізичного рушія.
    - `virtual void Update(float deltaTime) = 0`; оновлення стану модуля на кожному кроці симуляції.
    - `virtual void Render() = 0`; Рендеринг візуалізації модуля (якщо це необхідно).
- 3. `deformable_body_module.h/cpp`  
Основні компоненти:
  - Клас, що успадковується від `Module`:
    - Клас `DeformableBodyModule` успадковується від `Module` і реалізує поведінку деформованих тіл.
  - Реалізує поведінку деформованих тіл (наприклад, за допомогою фізики тканин):
    - Включає параметри та методи, які дозволяють симулювати поведінку деформованих тіл, таких як тканини або еластичні матеріали, наприклад, на основі пружинно-масової моделі.
    - `void Init(PhysicsEngine* engine) override`; Ініціалізує деформоване тіло в контексті фізичного рушія.
    - `void Update(float deltaTime) override`; Оновлює стан деформованого тіла на кожному кроці симуляції.
    - `void Render() override`; Відповідає за рендеринг деформованого тіла (якщо потрібно).
- 4. Реалізація системи плагінів:
  - Механізм завантаження:
    - Варто використати динамічне завантаження бібліотек (`LoadLibrary` на Windows) для завантаження плагінів під час виконання програми.
  - Інтерфейс плагіна:
    - Визначає чіткий інтерфейс для плагінів через базовий клас `Module`, що дозволить легко інтегрувати нові модулі в рушій.
- 5. Конфігурація пайплайну:
  - Файл конфігурації:
    - Використовуйте формат JSON [6] або YAML [7] для збереження налаштувань пайплайну обробки (наприклад, черговість етапів обчислення).
  - Завантаження конфігурації:
    - Реалізуйте завантаження конфігурації на початку роботи програми, щоб автоматично налаштувати пайплайн відповідно до завантаженого файлу.
- 6. Апаратне прискорення:
  - Використання GPU:
    - Розгляньте використання CUDA або OpenCL для реалізації обчислень на графічному процесорі для прискорення симуляцій.
  - Використання SIMD:
    - Оптимізуйте обчислення на CPU за допомогою SIMD-інструкцій (наприклад, через використання бібліотеки Eigen або схожих).
- 7. Інтерфейс користувача:
  - Вибір бібліотеки GUI:
    - Використовуйте Qt [8], wxWidgets [9] або інші бібліотеки для створення графічного інтерфейсу.
  - Функціональність:
    - Реалізуйте функції для завантаження сцен, налаштування параметрів симуляції та відображення результатів у реальному часі.
- 8. Тестування та дебагування:
  - Написання тестів:
    - Створіть одиничні тести для перевірки основних функцій рушія, включаючи модулі, завантаження плагінів та оновлення фізики.
  - Використання дебагера:

- Використовуйте вбудовані інструменти для налагодження (наприклад, Visual Studio Code) для відстеження та виправлення помилок у програмі.
9. Розширення функціональності:
- Додавання нових типів об'єктів:
    - Реалізуйте нові класи, які успадковуються від Module, щоб додавати нові фізичні моделі (наприклад, рідини або газу).
  - Підтримка різних фізичних моделей:
    - Додайте підтримку різних фізичних моделей, які можуть бути інтегровані до рушія за допомогою нових модулів або плагінів.
  - Інтеграція з іншими системами:
    - Додайте можливість інтеграції з іншими програмами, наприклад, через API або інші інтерфейси, щоб використовувати фізичний рушій у різних проектах, включаючи ігри або наукові симуляції.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Коли справа доходить до симуляцій, ідея мати спільний, гнучкий і масштабований рушій обчислювальної фізики разом з можливістю використовувати апаратне прискорення має велике значення. Конвеєри допомагають користувачам працювати над дослідженнями замість того, щоб витратити час на складності моделювання, створення та візуалізацію симуляцій. Кросплатформеність рушія та можливості апаратного прискорення означають, що програма може бути більш корисною в багатьох дослідницьких контекстах. По мірі розвитку рушія, як описано тут, до нього будуть додаватися розподілені обчислення, що дозволить працювати з більшими симуляціями. Розробка фізичного рушія, який буде застосовний до всіх типів ігор і сценаріїв, - це завдання, яке повинен вирішувати програміст, що володіє достатніми знаннями про фізичну симуляцію. Можливість його подальшого розширення та зміни фреймворку, створення графічного інтерактивного редактора, перевага використання апаратного прискорення та розподілу обчислювального навантаження роблять його цінним у багатьох галузях. Таким чином, використовуючи доступні інструменти, такі як MPI, OpenGL, Vulkan або SFML, буде спроектовано та реалізовано впорядкований фізичний рушій для різних платформ, який, перш за все, буде пропонувати розумну продуктивність, але в той же час буде унікальним та конкурентоспроможним.

### Література

- Бернатович А., Стеценко І. Методи та програмні засоби фізичної симуляції / А. Бернатович, І. Стеценко // Адаптивні системи автоматичного управління. – 2023. – Т. 1. – № 42. – С. 130-140.  
<https://doi.org/10.20535/1560-8956.42.2023.279104>
1. Офіційна документація C++ : [вебсайт]. – Режим доступу : <https://en.cppreference.com/w/>  
Зеленський О.С., Лисенко В.С. Основи програмування на C++ / О.С. Зеленський, В.С. Лисенко // Кривий Ріг: Державний університет економіки і технологій. – 2023. – 269 с. – Режим доступу :  
[https://dspace.duet.edu.ua/jspui/bitstream/123456789/831/1/%D0%9D%D0%9F%20Osnovy\\_C%2B%2B.pdf](https://dspace.duet.edu.ua/jspui/bitstream/123456789/831/1/%D0%9D%D0%9F%20Osnovy_C%2B%2B.pdf)
2. Visual Studio Code : [веб-сайт]. – Режим доступу : <https://code.visualstudio.com/>  
CMake : [веб-сайт]. – Режим доступу : <https://cmake.org/cmake/help/latest/manual/cmake.1.html>
3. JSON : [веб-сайт]. – Режим доступу : <https://www.json.org/json-en.html>
4. YAML : [веб-сайт]. – Режим доступу : <https://yaml.org/>
5. Qt: офіційна документація : [вебсайт]. – Режим доступу : <https://doc.qt.io/>  
wxWidgets : [веб-сайт]. – Режим доступу : <https://docs.wxwidgets.org/3.2.5/>

### References

- Bernatovych A., Stetsenko I. Metody ta prohramni zasoby fizychnoi symuliatsii / A. Bernatovych, I. Stetsenko // Adaptivni systemy avtomatychnoho upravlinnia. – 2023. – Т. 1. – № 42. – С. 130-140. DOI:  
<https://doi.org/10.20535/1560-8956.42.2023.279104>
- Ofitsiina dokumentatsiia C++ : [veb-sait]. Rezhym dostupu : <https://en.cppreference.com/w/>  
Zelenskyi O.S., Lysenko V.S. Osnovy prohramuvannia na C++. Kryvyi Rih: Derzhavnyi universytet ekonomiky i tekhnolohii, 2023. 269 s. Rezhym dostupu :  
[https://dspace.duet.edu.ua/jspui/bitstream/123456789/831/1/%D0%9D%D0%9F%20Osnovy\\_C%2B%2B.pdf](https://dspace.duet.edu.ua/jspui/bitstream/123456789/831/1/%D0%9D%D0%9F%20Osnovy_C%2B%2B.pdf)
- Visual Studio Code : [veb-sait]. – Rezhym dostupu : <https://code.visualstudio.com/>  
CMake : [veb-sait]. – Rezhym dostupu : <https://cmake.org/cmake/help/latest/manual/cmake.1.html>
- JSON: [veb-sait]. – Rezhym dostupu : <https://www.json.org/json-en.html>
- YAML: [veb-sait]. – Rezhym dostupu : <https://yaml.org/>
- Qt: ofitsiina dokumentatsiia: [veb-sait]. – Rezhym dostupu : <https://doc.qt.io/>  
wxWidgets: [veb-sait]. – Rezhym dostupu : <https://docs.wxwidgets.org/3.2.5/>

<https://doi.org/10.31891/2219-9365-2024-79-9>

УДК 004.83

МЕЩЕРЯКОВ Олександр

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

<https://orcid.org/0009-0002-6485-9722>

e-mail: [ascellanova@gmail.com](mailto:ascellanova@gmail.com)

## ОГЛЯД МЕТОДІВ ТА ЗАСОБІВ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МОДЕЛЕЙ КЛАСИФІКАЦІЇ МЕДИЧНИХ ЗОБРАЖЕНЬ

*У статті представлені методи та засоби для покращення ефективності моделей класифікації медичних зображень. Проаналізовано різні підходи до вдосконалення точності та надійності класифікації, зокрема використання нейронних мереж, методів злиття та застосування кумулятивного ефекту від згладжування та покращення зображень на наборах даних, які використовуються для аналізу COVID-19. Розглянуто способи боротьби з аномаліями які виникають при використанні двовимірного гель-електрофорезу (2-DGE) за допомогою системи попередньої обробки. Запропоновані методи можуть бути використані для медичної діагностики, дозволяючи покращити точність та швидкість обробки медичних зображень, що в свою чергу сприяє підвищенню якості медичних послуг та скороченню часу на постановку діагнозу. Представлена в статті інформація має оглядовий характер.*

*Ключові слова: аналіз даних, машинне навчання, медичні дані, нейронна мережа, класифікація.*

MESHCHERIAKOV Oleksandr

National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute"

## REVIEW OF METHODS AND TOOLS FOR IMPROVING THE EFFICIENCY OF MEDICAL IMAGE CLASSIFICATION MODELS

*The article presents methods and tools for improving the efficiency of medical image classification models. It considers various approaches to enhancing classification accuracy and reliability, including the use of deep neural networks, fusion methods and the application of the cumulative effect of image smoothing and enhancement techniques on the datasets used for COVID-19 analysis. The datasets consist of four sets of categorised data, including COVID-19, pneumonia, viral pneumonia and bacterial pneumonia. Different combinations of CLAHE, gamma correction, HE, AMF, TVF, Gaussian and Median filtering methods are used to analyse the cumulative effect to determine the most effective combination of methods. 2-DGE images typically contain several anomalies that hinder spot detection and analysis. The article discusses methods for dealing with anomalies arising from the use of two-dimensional gel electrophoresis (2-DGE) using a pre-processing system consisting of three stages, normalization, noise reduction and background correction, are considered, which allowed to improve the image for a posteriori analysis. The fusion methods such as Uniform Voting, Distribution Summation, Dempster-Shafer, Entropy Weighting, Density-Based Weighting are analysed and compared with the newly developed BitClassification Fusion Model method on 15 benchmark data selected from the UEA and UCR time series classification repository. The proposed methods can be used for medical diagnostics, allowing to improve the accuracy and speed of medical image processing, which in turn contributes to improving the quality of medical services and reducing the time to diagnosis. Presented in the article information is an overview.*

*Keywords: data analyze, machine learning, medical data, neural network, classification*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Використання методів машинного навчання, зокрема глибокого навчання, для класифікації медичних зображень має значний вплив на сучасну медицину в питанні діагностики. Необхідність підвищення ефективності моделей класифікації медичних зображень є важливою, оскільки це дозволяє покращити діагностичні можливості та зменшити кількість помилкових діагнозів. Вдосконалення алгоритмів і засобів, що використовуються для аналізу медичних зображень, може значно поліпшити якість медичної допомоги.

Класифікація медичних зображень стикається з багатьма викликами, включаючи варіативність зображень, невеликий розмір вибірок даних, аномалії на медичних зображеннях, шум у зображеннях та необхідність високої точності. Існуючі методи не завжди можуть забезпечити необхідний рівень точності або швидкості обробки, що може призвести до помилкових діагнозів. Потрібно розглянути методи та засоби підвищення ефективності класифікації медичних зображень, зокрема за допомогою використання новітніх алгоритмів обробки зображень.

### АНАЛІЗ ОСТАННІХ ДОСЛІДЖЕНЬ І ПУБЛІКАЦІЙ

Сучасні дослідження в області класифікації медичних зображень використовують різноманітні підходи для підвищення ефективності моделей. Важливу роль відіграє попередня обробка зображень, яка включає нормалізацію, фільтрацію шуму та сегментацію зображень. В межах статті [22] реалізовано та досліджено створений алгоритм бітового злиття який застосовується до навченого класифікатора та порівняно з іншими методами злиття на 15 еталонних даних, вибраних з репозиторію класифікації часових



рядів UEA та UCR. Середня точність запропонованої моделі у відсотках є кращою, за інші методи на всіх 15 наборах даних. У статті [26] розглядається застосування накопичувального ефекту від згладжування та покращення зображень на наборах даних, які використовуються для аналізу COVID-19, що складається з 4 наборів поділених по категоріям, які включають COVID-19, пневмонія, вірусна пневмонія та бактеріальна пневмонія. Для аналізу накопичувального ефекту використовуються різні комбінації методів CLANE, гамма-корекція, HE, AMF, TVF, Gaussian та Median методи фільтрування для визначення найбільш ефективної комбінації методів. Автори статті [27] застосовують методи попередньої обробки для боротьби з аномаліями пов'язаними з процесом отримання аналізу вони виникають у вигляді дифузні плями, плями, що перекриваються, та фоновий шум при протеомному аналізі. Автор порівнює методи зниження шуму на зображеннях таких як Вейвлет-перетворення та геометричною нелінійною дифузійною фільтрацією на різних типах шуму та різною інтенсивністю шуму від 20 децибел до 8 децибел. Також нормалізацію зображень та методи корекції фону такі як порогову, багаторівневу порогову та порогову апроксимацію. На основі порівняння методів нормалізації зображень, зменшення шуму та корекції фону автор реалізує фреймворк попередньої обробки та перевіряє його з використанням реальних двовимірних гелелектрофорез зображень і у результаті кількість хибно позитивних спрацьовувань зменшилася від 71% до 93%.

### **ВИДІЛЕННЯ НЕВИРШЕНИХ РАНІШЕ ЧАСТИН ЗАГАЛЬНОЇ ПРОБЛЕМИ**

Ознайомившись з актуальними дослідженнями було знайдено невіршені питання, зокрема, ефективність моделей на малих вибірках даних. Також не достатньо досліджено вплив використання гібридних методів, які поєднують різні підходи, такі як традиційні методи обробки зображень та глибоке навчання. Ще однією проблемою є розробка методів, які б забезпечували високу точність і стабільність моделей на різноманітних медичних даних.

### **ФОРМУВАННЯ ЦІЛЕЙ СТАТТІ**

Метою статті є огляд методів та інструментів, спрямованих на підвищення ефективності моделей класифікації медичних зображень з метою поліпшення процесу діагностики. Зокрема, дослідження спрямоване на пошук найбільш ефективних технік обробки даних, а також гібридних методів, що поєднують переваги різних підходів. Це дозволить підвищити точність і швидкість класифікації медичних зображень, що, в свою чергу, сприятиме поліпшенню якості медичної діагностики.

### **ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ**

Впровадження передових методів збору даних та методів машинного навчання змінює підхід до медичної інформації. У цій статті розглядаються різні типи медичних даних а також обговорюються методології, що використовуються для покращення їх обробки. Підкреслюючи ключову роль машинного навчання, ми заглиблюємося в алгоритми і моделі, які аналізують величезні масиви даних для виявлення закономірностей, прогнозування результатів і персоналізації лікування. Вивчаючи конкретні приклади та сучасні дослідження, ми підкреслюємо глибокий вплив цих технологій на підвищення точності діагностики. Ця стаття має на меті надати огляд найбільш ефективних методів обробки даних та методів класифікації для підвищення точності машинного навчання.

З поширенням електронних медичних карток, персональних медичних пристроїв, та мобільних медичних додатків обсяг і різноманітність медичних даних, доступних на сьогодні, досягли безпрецедентного рівня. Цей стрімкий ріст даних постає як значний виклик. Проблема полягає у тому, як ефективно збирати, керувати та аналізувати цю інформацію з метою отримання релевантних висновків, для подальшого використання у діагностиці.

В основі цього виклику лежить інтеграція найефективніших методів для збору даних і технологій машинного навчання. Машинне навчання продемонструвало неабиякі можливості в обробці величезних обсягів даних для виявлення закономірностей, прогнозування майбутніх тенденцій і отримання дієвих висновків. У застосуванні до медичних даних ці методи можуть докорінно змінити наше розуміння хвороб, розробку методів лікування та надання персоналізованої медичної допомоги.

Джерела медичних великих даних, такі як записи адміністративних звернень, клінічні реєстри, електронні медичні картки, біометричні дані, інформація отримана від пацієнтів, медична візуалізація, дані біомаркерів та великі клінічні випробування.[2] Інтеграція цих джерел даних призводить до появи додаткових вимірів даних, таких як великий розмір, розрізненість джерел, різні масштаби (від секунд до років), невідповідності, неповнота і складність[1]. Не існує універсального протоколу для моделювання, порівняння або оцінки ефективності різних стратегій аналізу даних.

Необроблені медичні дані є об'ємними і неоднорідними. Медичні дані можуть бути зібрані з різних зображень, інтерв'ю з пацієнтом, лабораторних даних, а також спостережень та персональної інтерпретації лікаря.[5] Всі ці компоненти можуть вплинути на діагноз, прогноз і лікування пацієнта, і їх не можна ігнорувати. Основні сфери неоднорідності медичних даних можна згрупувати під такими заголовками:

**1.Обсяг і складність медичних даних:** Обсяг медичних даних є значним і охоплює різноманітні типи інформації: від медичних зображень і результатів лабораторних досліджень до історій хвороб і даних моніторингу в режимі реального часу. Ця складність підвищується необхідністю інтегрувати різні типи даних, які часто зберігаються в не пов'язаних одна з одною системах, щоб створити цілісне уявлення про стан здоров'я пацієнта. Передові інструменти управління даними та аналітики необхідні для вирішення цієї задачі та отримання релевантних висновків[5].

**2.Інтерпретація лікаря:** Відмінності у інтерпретації медичних даних лікарями вносять додатковий рівень варіативності. Набутий досвід, знання та суб'єктивні судження кожного лікаря можуть впливати на те, як дані розуміються і застосовуються в лікуванні пацієнтів. Ця варіативність може впливати на діагностичні рішення, плани лікування та результати лікування пацієнтів[5]. Послідовне навчання, системи підтримки прийняття рішень і стандартизовані протоколи можуть допомогти пом'якшити ці відмінності, але не можуть усунути їх повністю.

**3.Чутливість і специфічність:** Чутливість і специфічність є ключовими показниками в медичній діагностиці, що відображають точність тестів у виявленні хибно позитивних і хибно негативних результатів відповідно. Висока чутливість гарантує, що більшість пацієнтів з певним захворюванням будуть правильно ідентифіковані, в той час як висока специфічність гарантує, що більшість здорових людей не будуть помилково діагностовано як хворих [5]. Баланс цих показників має вирішальне значення, оскільки надто чутливі тести можуть призвести до хибно позитивних результатів і непотрібного лікування, тоді як надто специфічні тести можуть пропустити справжні випадки захворювання.

З клінічним текстом або записами в електронних картках пацієнтів є специфічні проблеми оскільки вони написані спеціалізованою мовою. Такі записи часто написані у телеграфному стилі, з великою кількістю інформації, для спілкування між лікарями, і не існує розроблених словників, які можна було б використовувати для перевірки орфографії чи граматики. Крім того, лікарі іноді використовують неповні речення, а також часто не згадують про такий об'єкт, як пацієнт, оскільки пацієнт вважається очевидним у тексті. Пацієнт може бути написаний або просто згаданий як «прибув з температурою 38,3 і пульсом 132» [3].

Термін «неструктуровані дані» зазвичай використовується в клінічних умовах для позначення вільного тексту, проте він також охоплює кілька інших типів даних, таких як медичні зображення. Останні охоплюють широкий спектр категорій, від статичних зображень, отриманих за допомогою діагностичних інструментів (наприклад, рентгенографії), до відеозаписів, що фіксують хід медичних процедур, таких як хірургічні втручання.

На початковому етапі медична візуалізація переважно використовувалася для термінового використання, для підтримки процесу пов'язаного з конкретною діагностичною або терапевтичною процедурою для окремого пацієнта. Таким чином, вторинне використання зображень для досліджень не розглядалося. Як наслідок, медичні зображення та відео з ходом процедури після їх первинного використання видалялися. Сьогодні продукти медичної візуалізації відіграють ширшу роль, будучи джерелом даних для клінічних досліджень, а також інструментом для підтримки догляду за пацієнтом. В результаті, клінічні зображення і відео систематично зберігаються в електронних медичних картах разом з відповідними клінічними звітами, що містять оцінки лікарів щодо інтерпретування медичних зображень [7].

Аналіз даних проводиться як на медичних зображеннях, так і на текстах клінічних звітах. Аналіз клінічних звітів відноситься до області клінічного інтелектуального аналізу текстів. Що стосується медичних зображень, їх аналіз може виконуватися з різними цілями. Деякі методи спрямовані на покращення візуальної якості зображень, щоб полегшити їх інтерпретацію людиною-спостерігачем[10]. Наприклад, вони можуть бути використані для зменшення шуму зображень. Інші підходи допомагають у прийнятті рішень щодо пацієнта. За останні кілька років були розроблені методи глибокого навчання для аналізу зображень в декількох клінічних контекстах. Наприклад, в ендоскопії такі методи автоматично виявляють та класифікують ураження[8,9]. У дерматології поширеною темою досліджень є автоматична класифікація уражень шкіри за ступенем злоскісності, а в неврології дослідження присвячені автоматичній класифікації захворювань головного мозку.

Клінічні діагностичні критерії, що використовуються для діагностики та лікування, засновані на суб'єктивному досвіді, а не на об'єктивних біомаркерах захворювань. Ці критерії, розроблені експертами, інтегровані в посібники діагностики. Однак, такий підхід має ряд обмежень. Поставлений діагноз може варіюватися в залежності від методології опитування [18], і клінічно схожі симптоми можуть бути викликані різними основними патологіями. Відтак, універсальні діагностичні критерії, що базуються лише на проявах симптомів, не завжди надійні в клінічному застосуванні. Часто вони є нестабільними в часі, неспецифічними та малоінформативними для визначення оптимальної терапії[10]. Неправильні діагнози та класифікація можуть призводити до неефективної терапії та неадекватного ведення захворювання.

Інтеграція підходів глибокого навчання в комп'ютерних системах діагностики (CAD) сприяє підвищенню точності діагностики захворювань на ранніх стадіях. Важливим етапом після отримання зразків є попередня обробка зображень, яка зменшує час навчання алгоритму, а також знижує ймовірність хибних

результатів. Для покращення зображень найбільш ефективні такі методи, як вирівнювання гістограми, CLANE та гамма-корекція, а також адаптивний медіанний фільтр (AMF), медіанний фільтр (MF), загальний варіаційний фільтр (TVF) та гауссові фільтри згладжування.

У дослідженні Шарма та інших [26] проаналізовано кумулятивний ефект від застосування методів згладжування та покращення зображень на наборах даних, які використовуються для вивчення COVID-19. Для початкового етапу авторами було обрано модель – DenseNet201. Цю модель навчали та тестували на наборі даних COVIDx, який включає зображення чотирьох категорій: COVID-19, звичайна пневмонія, вірусна пневмонія та бактеріальна пневмонія.

Враховуючи на значний обсяг набору даних та складність їх подальшої обробки, початкові тести проводилися на підмножині набору даних COVIDx. Ця підмножина включала 193 випадків COVID-19, 213 випадків пневмонії, 116 випадків вірусної пневмонії та 116 випадків бактеріальної пневмонії. Це дозволило отримати початкові результати щодо того, які покращення можуть бути досягнуті за допомогою обраної моделі та методів обробки зображень.

На Рис. 1 представлено графік залежності точності тестування від кількості епох у вигляді кривої. Графік відображає результати отримані для різних комбінацій методів покращення зображень: модель навчена на оригінальних зображеннях, вирівнювання гістограми, CLANE, гамма-корекція, HE + гамма, CLANE + гамма, AMF + гамма, TVF + гамма, TVF + гауссова + гамма, TVF + медіана + гамма, TVF + CLANE + гамма і TVF + HE + гамма.

Протягом початкового тестування на підмножині набору даних автори отримали точність для методу гамма-корекції досягла 94,21%. Точність для комбінації TVF + гамма виявилася найефективнішою та складала 95,04%. Це значно перевищувало точність моделі, навченої на оригінальних зображеннях. Усі експерименти проводилися з використання однакових параметрів: швидкості навчання  $10^{-5}$ , розмір вибірки 32, оптимізатор Adam і функція втрат у вигляді зваженої перехресної ентропії [26].

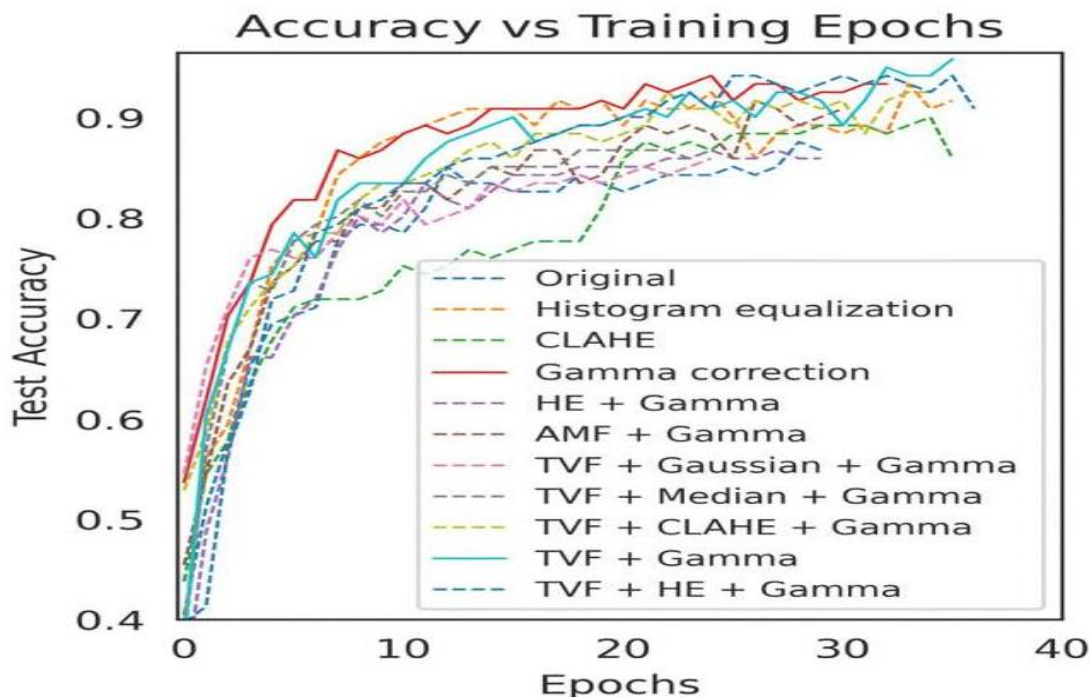


Рис. 1. Послідовність перевірених комбінацій покращення та затушовування зображень для пошуку найкращого методу покращення [26]

Відповідно до отриманих результатів у дослідженні, у верхній частині графіку деяких моделей спостерігається найвищі значення точності та чутливості до COVID-19, що вказує на переваги цих комбінацій покращення зображень у порівнянні з іншими. Застосувавши ці комбінації до повного набору даних COVIDx призводить до значного покращення в порівнянні з використанням плоских зображень. Гамма-корекція, TVF + гамма та CLANE показали найбільше підвищення чутливості у порівнянні з іншими методами покращення.

Знебарвлення зображень за допомогою TVF перед гамма-корекцією швидше досягає оптимальних результатів у порівнянні з CLANE і HE. Гамма-корекція демонструє найкращі результати, оскільки вона підвищує точність для всіх класів, а застосування повного варіаційного зашумлення перед гамма-корекцією покращує продуктивність і скорочує час збіжності на еталонних наборах даних. Крім того, гамма-корекція є однаково ефективною для всіх класів.

У дослідженні Мауріціо та інших [27] двовимірний гель-електрофорез (2-DGE) - це ключовий метод для протеомного аналізу, що дозволяє ефективно розділяти білки за їх ізоелектричною точкою та молекулярною масою. Зображення 2-DGE часто містять аномалії, пов'язані з процесом отримання аналізу, такі як: дифузні плями, перекриваючі плями, та фоновий шум.

У цьому дослідженні авторами пропонується спільна система попередньої обробки, яка поєднує можливості нелінійної фільтрації, корекції фону та методів нормалізації зображень для попередньої обробки 2-DGE зображень. Нормалізація зображень використовується для забезпечення однорідності даних, шляхом регулювання діапазону значень зображень на вході відповідно до цільових зображень.[28] Основним компонентом цієї системи є - спільна нелінійна дифузійна фільтрація, адаптивне кусково-гістограмне вирівнювання та багаторівнева порогова корекція. Ефективність методів оцінювалася як на синтетичних, так і на реальних 2-DGE-зображеннях.[27]

Основною метою нормалізації зображень є підсилення низькопоширених білків тому автори для її тестування використали синтетичне зображення з такими плямами (рис. 1а). Синтетичне зображення розміром 1024 x 1024 пікселів, з непрозорим фоном і 150 плямами, які були згенеровані відповідно до гаусівського розподілу зі стандартним відхиленням між 0,3 і 0,8. Інтенсивність плям контролювалася для відтворення низькопоширених білків з рівнем між 0,1 і 0,8. Автори здійснили порівняння методів вирівнювання гістограми, адаптивного кусково-гістограмне вирівнювання [30] та модифікацію інтенсивності фонових пікселів для нормалізації зображення, і використовували відсоток виявлених низькопоширених білків (LPD) для оцінки ефективності кожного методу.

За результатами LPD що отримали автори у ході дослідження – метод, який базується на інтенсивності фонового пікселя, виявив лише 48,7% плям з низькою щільністю. З іншого боку, гістограма та адаптивне вирівнювання за кусково-гістограмним методом виявили 82,1% та 88,9% плям відповідно. Як видно з рис. 1б і в, методи, засновані на вирівнюванні, покращують контрастність плям з низькою щільністю.

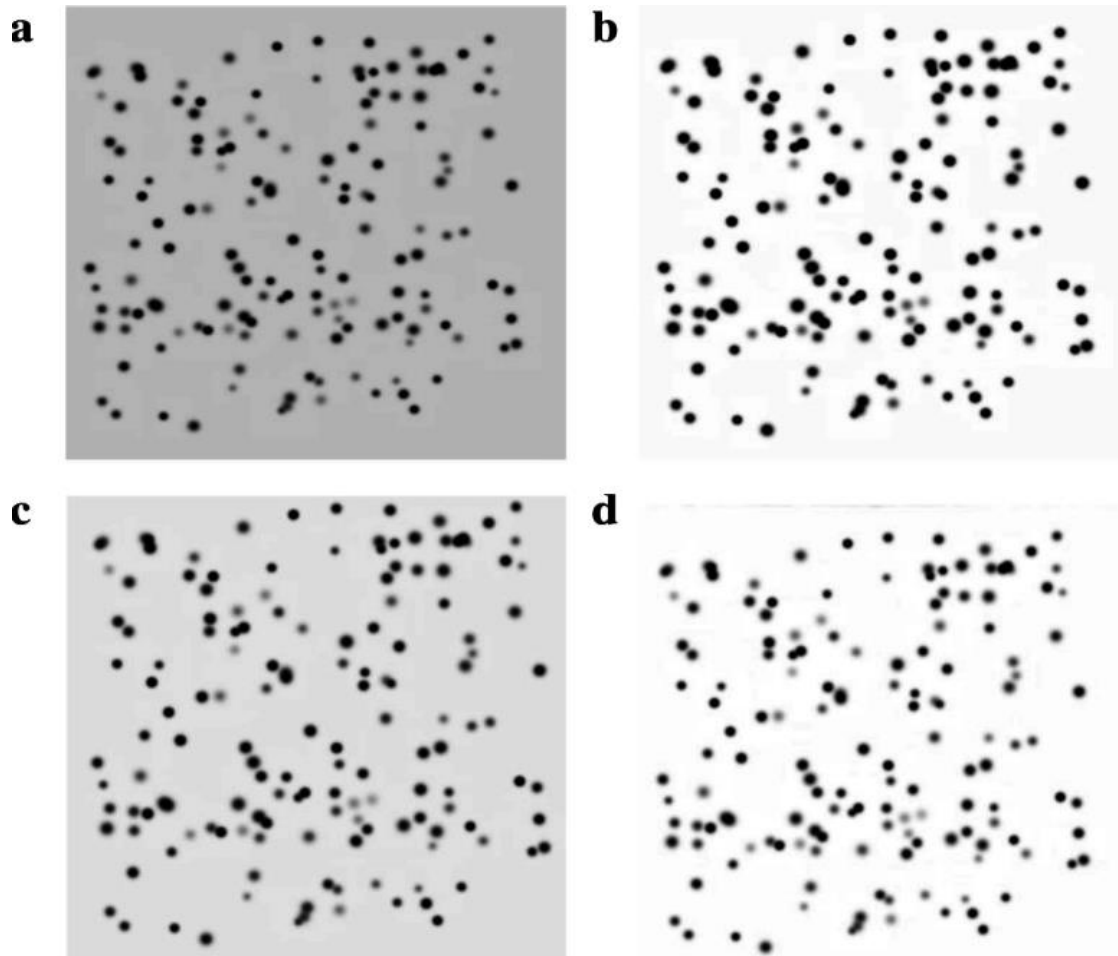


Рис. 1 Синтетичні білкові плями, змодельовані у вигляді 2-вимірного гаусівського розподілу. а - Приклад синтетичного зображення. б - Синтетичне зображення, нормалізоване за допомогою гістограмного вирівнювання. с - Синтетичне зображення, нормалізоване за допомогою адаптивного кусково-гістограмного вирівнювання і d - Синтетичне зображення, нормалізоване за допомогою модифікації інтенсивності фонового пікселя. [27]

Розглянемо порівняння методів корекції фону. У дослідженні було проведено порівняння трьох методів корекції фону: порогову, багаторівневу порогову та апроксимацію поверхню. Для оцінки цих методів було згенеровано синтетичне зображення зі зміною інтенсивності фону (див. рис. 2а). Варіації фону були отримані шляхом збільшення початкової інтенсивності до 155%. Для обох методів порогового значення було на рівні 60%. Рівняння В-сплайна [27] було використано для методів апроксимації поверхні з оптимізацією параметрів шляхом 150 ітерацій. Оцінка ефективності проводилася за допомогою індексу віднімання (SI), який порівнює кількість пікселів фону з їх оціненою кількістю.

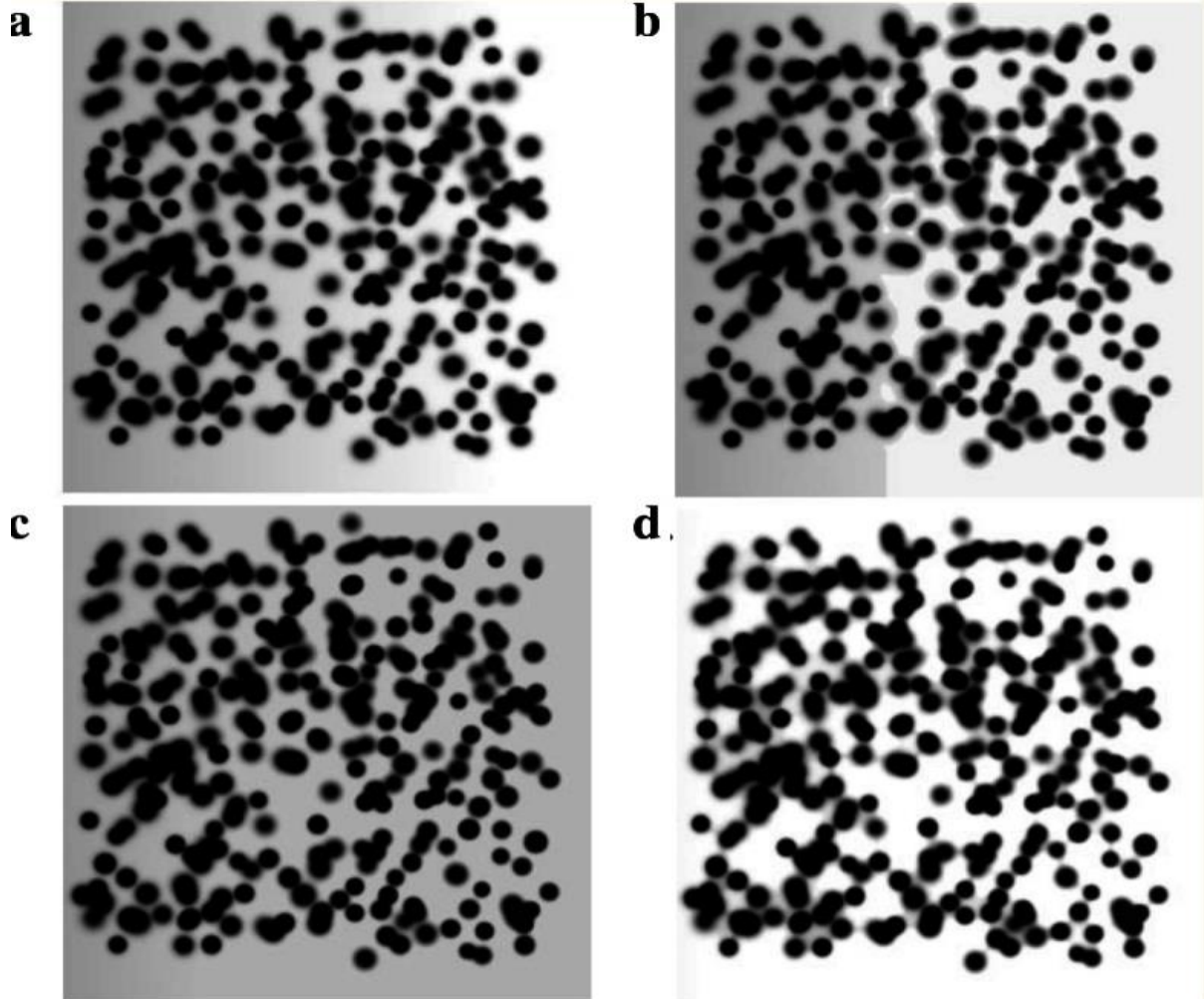


Рис. 2 Синтетичні білкові плями, змодельовані у вигляді 2-вимірного гаусівського розподілу з фоном.  
а. Приклад синтетичного зображення. б. Синтетичне зображення з корекцією фону за допомогою порогової апроксимації. с. Синтетичне зображення з корекцією фону за допомогою багаторівневої порогової апроксимації. д. Синтетичне зображення з корекцією фону за допомогою поверхневої апроксимації. [27]

Метод порогової корекції частково усунув фон, як видно з рис. 2b, фон розділено на дві різні області. У порівнянні, багаторівневе порогоування забезпечило рівномірний фон. Наближення до поверхні видалило більшу частину фону, але цей метод не зміг видалити пік селі що розтошовані близько до плям. За допомогою порогової апроксимації було виявлено 71,8% пікселів фону, тоді як за допомогою поверхневої апроксимації та багаторівневої порогової апроксимації було виявлено 97,9% та 98,5% пікселів фону для синтетичних зображень відповідно.

У статті авторами проведено аналіз методів зменшення шуму. Вейвлет-перетворення (WT) яке є одним з нелінійних фільтрів, який показує найкращі результати у зменшенні шуму на 2-DGE зображеннях [29]. Також існують інші не менш ефективні нелінійні методи, які дозволяють зменшити шум без згладжування країв плям. Автори порівняли WT з геометричною нелінійною дифузійною фільтрацією - GNDF. GNDF демонструє хороші результати на різних типах медичних зображень, але не застосовувався для 2-DGE зображень. Для фільтра WT було використано сімейство вейвлетів Даубекіса з п'ятьма рівнями декомпозиції [29]. Для GNDF було використано 35 ітерацій згладжування з коефіцієнтом дифузії 0,2 і вікнами 5x5 пікселів. Ефективність оцінювалася за точковою ефективністю [4]. WT і GNDF тестувалися з

синтезованими зображеннями з гауссівським, релеївським і експоненціальним шумом з SNR від 20 до 8 децибел. Кожне синтетичне зображення має 512x512 пікселів і містить 250 плям.

У таблиці 1 ілюструє порівняння точкової ефективності фільтрів WT і GNDF для синтетичних зображень з різними рівнями шуму. Загалом точкова ефективність обох фільтрів була дуже подібною для більшості рівнів шуму, з різницею в межах 2%. Однак для синтетичного зображення з гауссівським шумом на рівні 8 децибел GNDF показав точкову ефективність 77,86%, тоді як забезпечив ефективність на рівні WT – 67,5%.

Таблиця 1

**Оцінка ефективності методів зниження шуму за допомогою точкової ефективності (%) [27]**

Тип шуму	Метод зниження шуму	Інтенсивність шуму (dB)						
		20	18	16	14	12	10	8
Gaussian	WT	90.36	90.71	89.29	89.29	88.93	85.00	67.50
	GNDF	88.57	88.21	89.64	88.57	84.29	85.36	77.86
Rayleigh	WT	90.00	90.36	90.71	89.64	89.29	87.14	90.00
	GNDF	90.71	88.93	89.29	87.86	88.21	86.43	87.14
Exponential	WT	91.07	91.43	90.71	90.36	89.29	84.29	82.86
	GNDF	90.71	89.64	89.29	90.71	88.57	87.86	82.87

На основі аналізу методів нормалізації зображень, зменшення шуму та корекції фону автори вказують на необхідність застосування комплексного фреймворку попередньої обробки. Запропонований фреймворк включає нормалізацію зображень для підвищення контрастності низькопоширених білків, нелінійної фільтрації для зменшення шуму зі збереженням крайової інформації та корекції фону для забезпечення однорідності фонових пікселів. Згідно з попередніми результатами, ми використали кусково-гістограмне вирівнювання для нормалізації зображення, GNDF для фільтрації та багаторівневе порогове вирівнювання для корекції фону. Спільну систему попередньої обробки автори оцінювали, використовуючи як синтетичні, так і реальні 2-DGE зображення.

Оцінка інтегрованої системи попередньої обробки проводилася на синтетичному зображенні, створеному з використанням двовимірного гауссового розподілу, що містить 150 плям зі стандартним відхиленням від 0,1 до 0,8. Зображення характеризується варіацією інтенсивності фону вздовж горизонталі та містить гауссівський шум з медіаною нуль, стандартним відхиленням 1,535 і релеївським шумом з  $a = 0$  і  $b = 0,0539$ . У табл. 2 наведено результати продуктивності з використанням відсотку виявлення низькопоширених білків (LPD), точкової ефективності та індексом віднімання (SI), що порівнює кількість пік селів фону з розрахунковим значенням. Метрику SI обчислювали лише для зображень, отриманих за допомогою методів корекції фону та спільної попередньої обробки, оскільки вона вимірює фон, віднятий від зображення.

Найкращий показник LPD було отримано при використанні спільного фреймворку попередньої обробки, коли на зображенні було виявлено 60% плям з низькою щільністю. Для порівняння, цей відсоток становив 40% при застосуванні лише методу нормалізації. З точки зору ефективності виявлення плям, запропонований фреймворк виявив 63,84% плям, тоді як при використанні однієї методики було отримано нижчі відсотки: 3,57% для нормалізації, 17,69% для відфільтрованого зображення і 6,69% з використанням корекції фону. Крім того, найкращий показник віднімання також було отримано за допомогою запропонованого фреймворку – 78,62% порівняно з 11,37% при використанні лише модифікованої гістограмної техніки для корекції фону.

Таблиця 2

**Оцінка ефективності фреймворку попередньої обробки з використанням LPD, точкової ефективності ( $\Xi$ ) та відсотком пікселів ідентифікованих як фон (BSI) [27]**

Метод	LPD(%)	$\Xi$ (%)	BSI(%)
Original synthetic image	0	5.87	N.A
Piece wise equalization	40	3.57	N.A
GNDF	3.33	17.69	N.A
Modified histogram-based	0	6.69	11.37
Joint pre-processing framework	60	63.84	78.62

Інтегрована система попередньої обробки була оцінена на реальних 2-DGE зображень, отриманих з чотирьох зразків апітоксину двох зразків сечі пацієнтів з раком простати і чотирьох 2D зображень з бази даних LECB 2-D PAGE Gel Image Database. Автори зафіксували відсоткове зменшення хибних результатів, порівнюючи оригінальні та попередньо оброблені зображення. Для 2-DGE зображень апітоксину система зменшила кількість хибних спрацьовувань від 43% до 72%. У випадку зразків сечі помилкові

спрацьовування попередньо оброблених зображень знизилася на 91% і 85% відповідно. А для зображень з бази даних селевих зображень LECB 2-D PAGE кількість хибно позитивних спрацьовувань зменшилася 71% до 93%. Отримані дані вказують на ефективність системи у питанні виявлення білків, зменшуючи кількість хибних спрацьовувань, спричинених шумом і неоднорідністю фону.

Зображення, отримані методом двовимірного гель-електрофарезу (2-DGE) характеризуються наявністю кількох аномалій, що ускладнюють процес виявлення та аналізу білкових плям. У даному дослідженні авторами було проведено експериментальну оцінку і підтверджено ефективність застосування трьох етапів цифрової обробки зображень, а саме: нормалізація, зменшення шуму і корекція фону, які значно підвищили якість зображень для подальшого аналізу. Запропоновано новий інтегрований фреймворк попередньої обробки, який об'єднує найкращі методи для кожного з трьох етапів обробки зображень які допомагають у боротьбі з конкретними аномаліями.

На підставі результатів, отриманих для реальних 2-DGE зображень встановлено, що інтегрований фреймворк забезпечує значно кращі результати порівняно з окремими методами. Зокрема, для таких зображень оптимальними є методи адаптивного кусково-гістограмного вирівнювання, GNDF та багаторівневого порогового вирівнювання.

Для оцінки точності моделей застосовується метод перехресної валідації, де навчальний набір даних застосовується для побудови моделі прогнозування, а тестовий набір - для оцінки її ефективності.

K-кратна перехресна валідація є ефективним підходом для мінімізації похибок пов'язаних із випадковими особливостями вибірки навчальних даних[23].

Зазвичай використовується десятикратна перехресна перевірка, яка вважається оптимальною з точки зору обчислювальних ресурсів та часу, який потрібен для перевірки. Цей підхід передбачає розподіл даних на десять частин, де дев'ять з них використовуються як навчальний набір даних для сурогатної моделі, а одна – для її тестування.

У методі десятикратної перехресної перевірки фіксована кількість даних фіксувану кількість вибірок з набору даних розбивають на десять разів. Серед цих десяти вибірок дев'ять вибірок слугуватимуть навчальним набором даних для побудови сурогатної моделі. Для побудови моделі, тоді як решта дев'ять згорток слугуватимуть як тестовим набором даних для перевірки та підтвердження точності сурогатної моделі. Для зменшення варіабельності кожен раунд десятикратного перехресного валідації виконується на різних підмножинах даних. Підсумкову точність моделі можна виразити як середнє значення точності, отримане в результаті десяти циклів валідації.

Алгоритми інтелектуального аналізу даних розрізняються за типом навчання, існують алгоритми для контрольованого, неконтрольованого та напівконтрольованого навчання. Контрольоване навчання спрямоване на прогнозування відомого результату на основі навчальної вибірки, що містить вже класифіковані дані, з метою зробити висновки на результатах або класифікувати нові тестові дані.

У неконтрольованому навчанні відсутні попередньо визначені цільові значення для прогнозування, тому аналізатори намагаються знайти природні закономірності або кластеризацію в немаркованих даних. Напівконтрольоване навчання дозволяє досягти балансу між продуктивністю і точністю за допомогою невеликих за розміром набори маркованих або анотованих даних і використовує їх зі значно більшою колекцією немаркованих даних[8]. Потенційно неконтрольоване навчання можна використати для подолання обмежень, які виникають у підходах керованого навчання для задач візуалізації плоских хвиль (PWI). Враховуючи, що підхід неконтрольованого навчання не може бути безпосередньо використане до реальних даних, він вимагає унікальної стратегії навчання для оптимізації мереж глибокого навчання. Один із можливих варіантів – це використати глибоке когерентне навчання, яке використовує характеристики сигналу візуалізації плоских хвиль у навчальній схемі мережі. Завдяки цьому методу мережа не обмежується якістю даних і стає більш надійно за різних умов не потребуючи еталонних зображень[12].

Моделі машинного навчання ML з контрольованим навчанням застосовуються в випадках, коли цільовий результат чітко визначений, а дані мають маркування, що відповідає цьому результату[11]. Наприклад, таким результатом може бути наявність або відсутність певного захворювання, такого як діабет чи гіпертонія. Однією з ключових вимог до таких моделей є корекція маркування даних, оскільки будь-яка похибка в оцінці результатів може негативно вплинути на продуктивність моделі та зменшити її здатність узагальнювати результати на популяції поза межами навчального набору даних[17]. При використанні контрольованого навчання виникає проблема з збором значних обсягів маркованих навчальних даних. Цю проблему можна вирішити за допомогою слабо контрольованого навчання для маркування немаркованих даних[19].

Процес побудови моделей з контрольованим навчанням є добре відпрацьованим і стандартизованим, оскільки такі моделі широко використовуються. Початковим етапом для формування керованої моделі машинного навчання є постановка задачі, після чого необхідно визначити та підготувати відповідні дані для вирішення цієї задачі[13]. Попередня обробка даних є ключовим етапом, що складається з вилучення зайвих змінних, очищення даних від суперечностей і обмеження їх в рамках необхідних змінних[14]. Для навчання моделі часто обирають випадкові вибірки даних. Спосіб полягає у начинні моделі

з застосуванням відповідних алгоритмів і оцінювати продуктивність моделі за допомогою матриці плутанини та робочої характеристики приймача (ROC-крива).

Створення остаточної моделі це процес що повторюється багато разів, під час чого обирається алгоритм з оптимальною комбінацією параметрів, для забезпечення найвищої прогностичної ефективності моделі [16]. У науковій літературі можна знайти рекомендації щодо використання параметрів за замовчуванням для зменшення витрат часу на налаштування гіперпараметрів і обчислювальних зусиль. У машинному навчанні під наглядом виділяють дві основні категорії задач: класифікація та регресія.

Алгоритми класифікації використовуються для групування даних шляхом передбачення категорійної мітки або змінної на основі наданих даних. Такий підхід застосовується, коли вихідні змінні є категоричними і належать до двох або більше класів. Використання допоміжних даних у поєднанні з контекстними знаннями підвищують загальну точність класифікації [21]. Методи класифікації включають в себе: nearest neighbors, logistic regression, support vector machine (SVM), naive Bayes, linear discriminant analysis (LDA), quadratic discriminant analysis, tree-based models (decision tree, random forest, extremely randomized trees)

Прикладом використання алгоритму класифікації є спам-фільтр в електронній пошті, де модель з навчанням під наглядом використовується для передбачення, чи є повідомлення спамом. Це досягається шляхом аналізу набору даних, що містить позначені приклади спаму та звичайних електронних повідомлень [14]. Алгоритм аналізує інформацію про кожний електронний лист, включаючи дані про відправника, тему, текст та інше. Користуючись цими даними алгоритм вивчає закономірності та надає оцінку, що визначає, чи є повідомлення спамом або ні.

Алгоритми регресії використовуються для прогнозування дійсних або неперервних значень шляхом встановлення зв'язків між кількома або більше змінними. Вони поділяються на: nearest neighbors, linear regression, support vector machine regression, tree-based models (decision tree, random forest, extremely randomized trees), kernel ridge regression. Прикладом може бути прогнозування заробітної плати на основі даних пов'язаних з роботою. У цьому випадку модель навчається на основі таких вхідних даних, як тривалість роботи, галузь, місцезнаходження та інші, для її подальшого використання у прогнозуванні щодо середньої заробітної плати на основі даних щодо працівника. [15].

Метод найближчих сусідів (KNN) є одним з найбільш інтуїтивно зрозумілих підходів до машинного навчання. Він ґрунтується на наступному: для певного входу відповідний вихід, вірогідно, буде схожим на вихід при подібних входах. Наприклад, у пацієнта є симптоми характерні для іншого пацієнта, якому було діагностовано певну хворобу, то, швидше за все, що йому також може бути поставлений такий самий діагноз. Використовується для вирішення завдань неконтрольованої фільтрації, під час якої зберігається структура вихідних даних, але відсутні мітки, тому для опису міток будується матриця графу Лапласіана на основі графу побудованого методом найближчих сусідів з подальшим вивченням псевдо міток класі за допомогою спектрального аналізу [10]

Формально, методи найближчого сусіда використовують навчальні вибірки з околиці заданої точки для виконання прогнозу. У випадку задачі регресії прогноз обчислюється як середньозважене значення цільових величин. У задачах класифікації, передбачена мітка відповідає мітці яка має найбільшу суму входжень кожної мітки.

Ключовим параметром методу найближчих сусідів є метрика, що визначає несхожість. Метрика використовується для визначення околиці будь-якої точки, а також може бути використана для обчислення ваг та позначається як  $d$ .

Рисунок 3 ілюструє концепцію цього алгоритму. Алгоритм  $k$ -середніх розбиває набір з  $n$  вибірок, позначених на набір з  $k$  нез'язаних кластерів, кожен з яких пронумерований. Кожен кластер характеризується своїм центроїдом, позначеним через  $\mu$ , тобто середнім значенням вибірок у цьому кластері. Центроїди повністю визначають набір кластерів, оскільки кожна вибірка відноситься до того кластера, центроїд якого є найближчим до неї.

Мета алгоритму – знайти центроїди, які мінімізують інерцію, також відому як внутрішньокластерний критерій суми квадратів. У цьому прикладі інерційність приблизно дорівнює 184.80 і є мінімально можливою інерційністю.

Оригінальний алгоритм, який використовується для знаходження центроїдів, часто називають алгоритмом Лойда і представлено в алгоритмі. Після ініціалізації центроїдів повторюється двокроковий цикл до збіжності (коли центроїди збігаються на двох послідовних ітераціях), який складається з кроку призначення, на якому кластери оновлюються на основі поточних центроїдів та кроку оновлення, на якому центроїди оновлюються на основі поточних кластерів

Коли кластери чітко визначені, точка з певного кластера, швидше за все, залишиться в цьому кластері. Таким чином, крок призначення можна пришвидшити завдяки нерівності трикутника, відстежуючи нижню та верхню межі відстаней між точками та центрами, але за рахунок більшого використання пам'яті [16].



Некероване навчання передбачає використання алгоритмів, здатних навчатися самостійно, без попереднього маркування чи навчання даних або моделі. Модель аналізує сирі дані і самостійно виводить правила, структуруючи інформацію на основі подібностей і відмінностей без інструкцій, яким чином обробляти кожен з фрагментів даних.

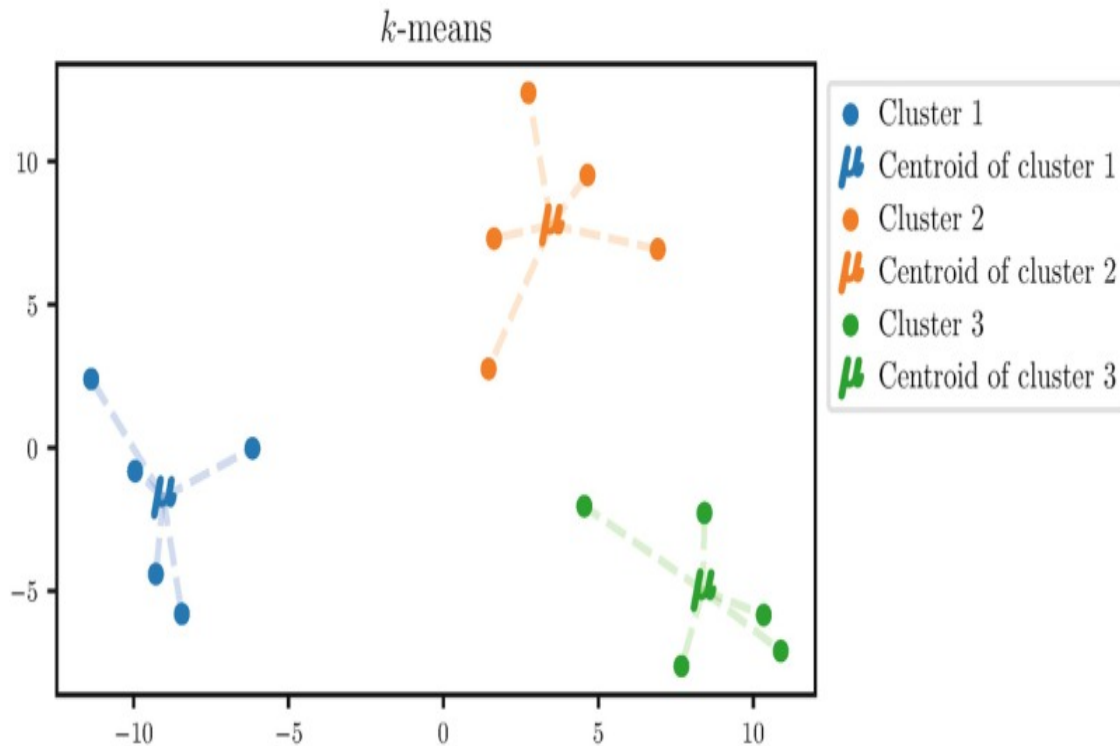


Рис. 3. Ілюстрація алгоритму k-середніх. [16]

Неконтрольоване навчання є особливо корисним для складних задач, однією з таких задач є кластеризація великих наборів даних. Алгоритм допомагає виявляти приховані ознаки, закономірності в значних об'ємах даних, що можна використовувати для подальшої категоризації цих даних.[6]

Наприклад, аналіз значного за об'ємом набору даних про погоду за допомогою такого алгоритму може призвести до групування за температурними показниками або за погодними умовами, що дослідним може інтерпретувати як класифікацію за порами року або типами погоди. До цього підходу належать методи кластеризації які поділяються на *k-means*, *Gaussian mixture model* та зменшення розмірності що в свою чергу поділяються на *principal component analysis (PCA)*, *linear discriminant analysis (LDA)*, *kernel principal component analysis*.

В межах дослідження [22] докладно аналізується алгоритм бітового злиття, застосованого до навченої моделі класифікації. Цей алгоритм злиття має на меті максимізувати точність класифікації шляхом скорочення часу виконання. Припустимо, що є множина з  $k$  класифікаторів  $\text{Classifier} = \{C1, C2, \dots, Ck\}$  та набір вхідних ознак  $X = \{x1, x2, \dots, xn\}$  у кількості  $n$  ознак, кожна з яких має  $m$  умов і набір міток класу позначається як  $\omega = \{\omega 1, \omega 2, \dots, \omega r\}$ . Кожен класифікатор  $Ci$  здійснює прогнозування мітки класу  $\omega$  для вхідного значення  $x_i$ . Для всіх класифікаторів результати формують  $r$ -вимірний вектор, який підтримує мітки  $\omega$ .

У межах набору даних ми отримаємо значення, визначене з вхідного елементу  $Ci(x)$ , яке представляє м'які мітки класів, тобто визначає, наскільки сильно класифікатор  $Ci$  вважає, що його екземпляр  $x$  належить до класу  $\omega_i$ . Він приймає значення від 0 до 1, де 0 означає що екземпляр абсолютно не відповідає класу  $\omega$ , а 1 означає повну відповідність. Методологія об'єднання класифікаторів полягає у визначенні категорії класу для входу  $x$  на основі результатів усіх класифікаторів. Наступним кроком використовується правило максимальної належності для остаточного призначення мітки класу, яке полягає у визначенні до якого класу належить об'єкт який має максимальне значення ймовірностей.

Існує дві стратегії комбінування класифікаторів: вибір класифікатора [22] та злиття класифікаторів. Ідея вибору класифікатора полягає в тому, що кожен класифікатор має досвід у деякій локальній області простору ознак. Коли вектор ознак  $x \in \mathcal{R}^n$  подається на класифікацію, класифікатор, відповідальний за околицю  $x$ , отримує найвищі повноваження для позначення  $\omega$ . Злиття класифікаторів передбачає, що всі

класифікатори мають однаковий доступ до всього простору ознак і рішення всіх  $C$  враховуються для будь-якого  $x$ . [24]

Традиційні методи злиття, до складу яких входять мажоритарне голосування, рівномірний розподіл, підсумовування розподілу, методи Демпстера-Шейфера, ентропійне зважування та зважування на основі щільності, використовують окремі вхідні дані від кожного з базових класифікаторів. Кожен з цих методів мають на меті – підвищення точності класифікації шляхом обробки і узагальнення результатів класифікації. Наприклад, при мажоритарному голосуванні, якщо більшість класифікаторів вказує на те, що екземпляр відноситься до класу 1, то алгоритм злиття автоматично присвоює клас 1 як остаточну мітку для цього екземпляра. Іноді точність може бути знижена у випадку коли реальна категорія даних відрізняється від прогнозованої. Висока часова складність мажоритарного голосування може зменшити загальну ефективність, хоча це компенсовано підвищенням точності системи [25].

Методи злиття є критично важливими для підвищення точності у задачах класифікації. Вибір ефективного методу злиття є одним із ключових чинників у вирішенні проблем розпізнавання образів. У цьому контексті запропонований ансамблевий класифікатор, заснований на бітовому злитті, пропонує значні переваги порівняно з традиційними методами злиття та вирішує проблеми пов'язані з ними, адже він не залежить від кількості класифікаторів або їх вихідних даних. Цей підхід оцінює дані на основі налаштованого власного параметра та приймає рішення відповідно до заданого порогового значення.

Запропонований авторами алгоритм бітового злиття аналізує вхідні дані у вигляді м'яких міток класів, що використовуються на стандартних наборах даних експресії генів. Цей алгоритм представляє собою ефективну та надійну структуру злиття, що перевершує традиційні підходи злиття за показниками точності, коректності та часової складності. Даний алгоритм проводить порівняння даних на основі ознак з заданим пороговим значенням і класифікує кожен ознаку як м'яку мітку класу. Основна увага алгоритму приділяється вимірюванню різноманітності в порівнянні з іншими методологіями. Після завершення класифікації традиційні алгоритми використовуються для подальшого підвищення точності.

У рамках проведеного дослідження було впроваджено і оцінено модель, яка використовує різні традиційні методи злиття. Точність, досягнута за допомогою кожного з методів відображена на Рисунках 4-5 для різних наборів даних, де вісь  $X$  вказує на відсоток даних, що використовується для тестування, а вісь  $Y$  вказує на рівень точності. З метою навчання та тестування було застосовано 10-кратну схему перехресної перевірки. Отримані результати свідчать про те, що класифікація за допомогою бітового злиття демонструє таку ж ефективність як і інші більш складні методи, у поліпшенні швидкості розпізнавання для зазначеного набору даних.

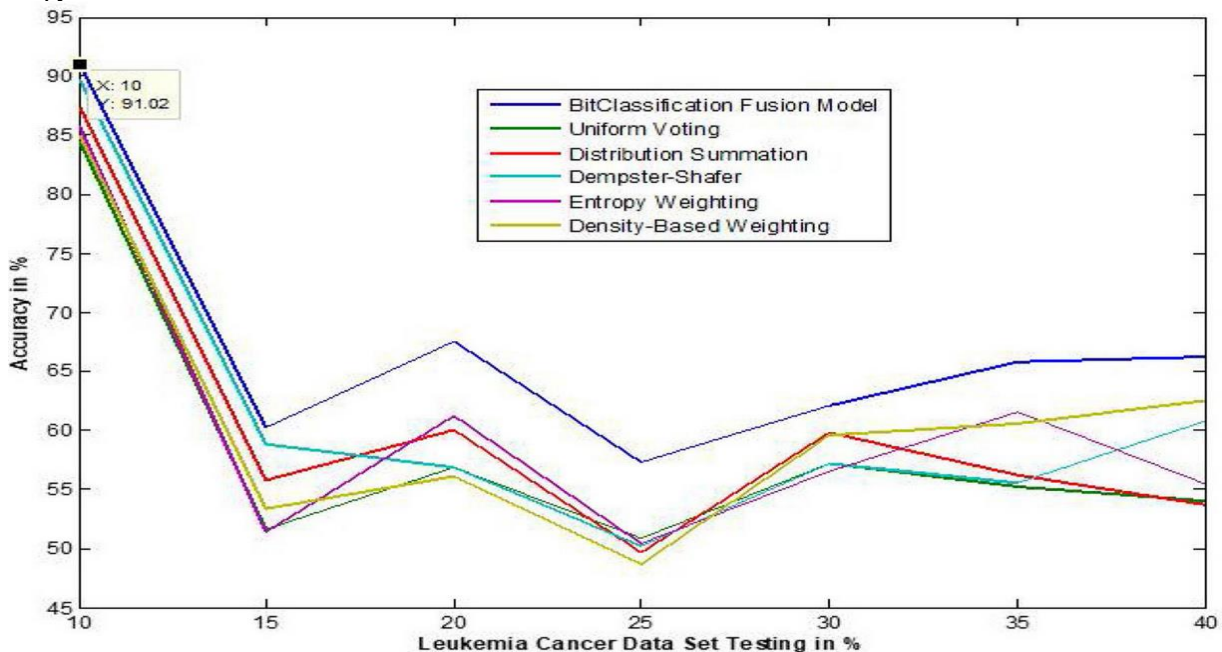


Рис. 4. Точність роботи алгоритму бітового злиття для раку лейкемії порівняно з іншими стандартними методами злиття. [22]

У ході дослідження автори провели порівняння створеного алгоритму з результатами, отриманими на 15 стандартних наборах даних, відібраних з репозиторію класифікації часових рядів UEA та UCR. Для оцінки точності злиття дослідники зіставили її з найкращими показниками інших класифікаторів.

Середня точність запропонованої моделі, виражена у відсотках, перевершила показники інших методів. При найсприятливішому сценарії, запропонована модель продемонструвала найкращі результати на всіх 15 наборах даних. Тоді, як другим найкращим методом виявився метод опорних векторів (SVM).

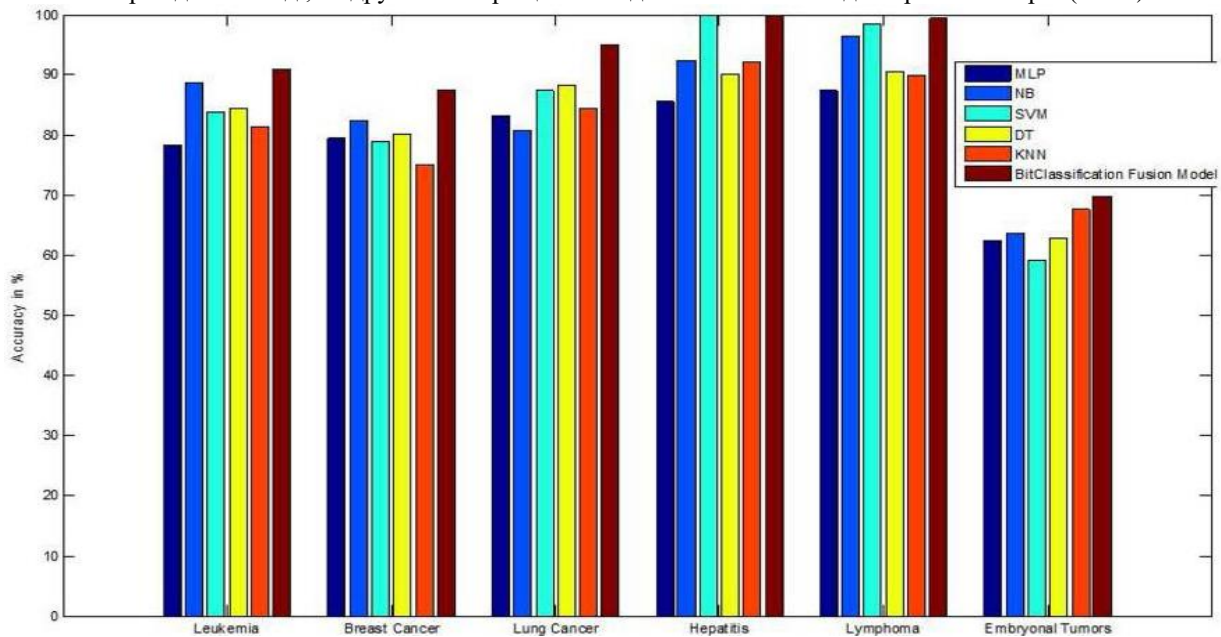


Рис. 5. Порівняльний аналіз точності всіх наборів даних. [22]

Методологія, що пропонується авторами, сприяє підвищенню точності класифікації шляхом орієнтування на категоризацію окремих значень ознак, а не на ознаки в цілому. Це дозволяє моделі досягати високих показників точності навіть на великих наборах даних, при цьому додаткові обчислювальні ресурси потрібні в незначних обсягах.

На етапі попереднього аналізу даних може бути корисним дослідити дисперсії для  $p$  ознак і виконати розрахунки  $12p(p-1)$  коваріацій або кореляцій. Проте, зі збільшенням значення  $p$  цей процес стає складнішим. Оскільки кожна окрема ознака може пояснювати лише малу частку загальної дисперсії. Тому бажаним варіантом може бути використання іншого представлення даних, яке б дозволило невеликій кількості ознак пояснювати більшу частку загальної дисперсії, тобто шляхом створення адаптованої системи координат до вхідних даних.

Метод аналізу головних компонент (PCA) полягає в побудові репрезентації даних через головні компоненти. Головні компоненти утворюють послідовність одиничних векторів, таким чином, що кожен  $i$ -й вектор забезпечує максимальне пояснення дисперсії даних, при цьому залишаючись ортогональним до перших  $i - 1$  векторів.

### ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМІ

Застосування машинного навчання в сфері охорони здоров'я передбачає інтеграцію різних типів даних із чисельних джерел для формування всебічного уявлення про стан здоров'я пацієнта. Це потребує ефективних інструментів управління даними та аналітики. Однак, існує проблема у варіативності інтерпретацій медичних даних лікарями, які ґрунтуються на персональному досвіді і судженнях лікарів, що може призвести до варіативності в наданні допомоги пацієнтам. Для ефективного вирішення цієї проблеми найкраще підходить стандартизація.

Машинне навчання відкриває нові горизонти для ідентифікації релевантних ознак, які є важливими для постановки діагнозу, допомагаючи знайти приховані закономірності. Використання моделей, які можна інтерпретувати, таких як машини опорних векторів (SVM) або дерева рішень, дозволяє дослідникам вивчати ознаки, які лежать в основі прийняття рішень алгоритмами. Наступним методом, що може бути використаний для пошуку корисних ознак – глибоке навчання. У випадках коли використовується в поєднанні з методами інтерпретації є опція використання навчання без попередньої обробки зображень. Ще один підхід визначити релевантні ознаки для класифікації – порівняти результати прогнозування різних моделей машинного навчання з різними вхідними ознаками, що надає можливість оцінити внесок кожної окремої ознаки в покращенні класифікації.

Як було зазначено у попередніх розділах, боротьба з аномаліями та артефактами у медичних зображеннях, таких як ті, що виникають при використанні 2-DGE, залишається важливим викликом. Існуючі методи попередньої обробки даних не завжди забезпечують достатню ефективність у випадках складних

аномалій, що може негативно впливати на точність діагностики. Тому що доцільним є проведення дослідження, присвяченого вдосконаленню методів попередньої обробки медичних зображень для підвищення їхньої стійкості до аномалій. Зокрема, необхідно розробити нові алгоритми нормалізації, фільтрації та корекції зображень, що будуть більш ефективними у випадках складних та нетипових даних.

Гамма-корекція, TVF + гамма та CLAHE досягають більшого покращення значення чутливості порівняно з іншими групами методів. Знебарвлення зображення за допомогою TVF перед гамма-корекцією допомагає швидше досягти найкращих результатів порівняно з методами CLAHE і HE. Гамма-корекція є найефективнішим з усіх розглянутих підходів, оскільки вона покращує оцінку точності для всіх класів, а застосування підходу повного варіаційного зашумлення перед гамма-корекцією покращує продуктивність і час збіжності на еталонних наборах даних. При цьому гамма-корекція є однаково сприятливою для всіх класів.

Серед методів корекції фону найбільш ефективними є методи поверхневої апроксимації та багаторівневої порогової апроксимації з результатами 97,9% та 98,5% пікселів фону на синтетичних зображень відповідно.

Згідно з результатами проаналізованої роботи інтегрований фреймворк для реальних 2-DGE зображень, що складається з трьох етапів обробки зображень, кожен з яких вирішує конкретну проблему, показую значно більшу ефективність порівняно з окремими методами. У цьому дослідженні методи що були визначенні як найбільш ефективні: адаптивного кусково-гістограмного вирівнювання, GNDF та багаторівневого порогового вирівнювання.

Методи злиття є критично важливими для підвищення точності у розпізнаванні образів. Класифікатор, заснований на бітовому злитті, пропонує значні переваги порівняно з традиційними методами злиття та вирішує проблеми пов'язані з ними, адже він не залежить від кількості класифікаторів або їх вихідних даних. Цей підхід оцінює дані на основі налаштованого власного параметра та приймає рішення відповідно до заданого порогового значення. У ході дослідження автори провели порівняння створеного алгоритму з результатами, отриманими на 15 стандартних наборах даних, і у кожному з наборів алгоритм показував найбільшу ефективність.

Методи злиття дозволяють значно покращити точність класифікації, особливо коли є доступ до декількох моделей або класифікаторів. Проте залишається відкритим питання оптимального налаштування параметрів таких методів, що забезпечить їх універсальне застосування.

#### Література

- 1 Tayefi M, Phuong D, Dalianis H. (2021, February). «Challenges and opportunities beyond structured data in analysis of electronic health records». doi: <https://doi.org/10.1002/wics.1549>
- 2 Gerard P, Peter V, Frede M. D. (2016, December). «Bigger data, bigger problems. J Orthop Trauma». doi: <https://doi.org/10.1097/bot.0000000000000463>
- 3 Dalianis H. (2018, May). «Evaluation Metrics and Evaluation. In Clinical Text Mining» (pp.45-53) doi: [https://link.springer.com/chapter/10.1007/978-3-319-78503-5\\_6](https://link.springer.com/chapter/10.1007/978-3-319-78503-5_6)
- 4 Tammy J, Jaimie L. Gradus, Anthony J. Rosellini. (2020, May). «Supervised Machine Learning: A Brief Primer». doi: <https://doi.org/10.1016/j.beth.2020.05.002>
- 5 Krzysztof J. Cios, Moore G. W. (2002, September) «Uniqueness of medical data mining». doi: [https://doi.org/10.1016/s0933-3657\(02\)00049-0](https://doi.org/10.1016/s0933-3657(02)00049-0).
- 6 Girish C, Ferat S. (2014, January). «A survey on feature selection methods». doi: <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- 7 Jadczyk T, Kiwic O, Raj M. Khandwalla (2019, September) «Feasibility of a voice-enabled automated platform for medical data collection: CardioCube». doi: <https://doi.org/10.1016/j.ijmedinf.2019.07.001>
- 8 Choong Ho Lee and Hyung-Jin Yoon (2017, March) «Medical big data: promise and challenges» doi: <https://doi.org/10.23876/j.krcp.2017.36.1.3>
- 9 [Maryam Tayefi, Phuong Ngo, Taridzo Chomutare](#), (2021, February) «Challenges and opportunities beyond structured data in analysis of electronic health records». doi: <https://doi.org/10.1002/wics.1549>.
- 10 Han J, Zhengya S, Hongwei H. (2015, September). «Selecting feature subset with sparsity and low redundancy for unsupervised learning». doi: <https://doi.org/10.1016/j.knosys.2015.06.008>
- 11 David A. Nardone. (1990). «Collecting and Analyzing Data: Doing and Thinking. In Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition» [H. Kenneth Walker, W. Dallas Hall, J. Willis Hurst](#).
- 12 Hyunwoo C, Seongjun P, Jinbum K, Yangmo Y, (2024, September), «Deep coherence learning: An unsupervised deep beamformer for high quality single plane wave imaging in medical ultrasound», doi:<https://doi.org/10.1016/j.ultras.2024.107408>

- 13 Faouzi J, Colliot O. (2023, July). «Classic Machine Learning Methods. In Machine Learning for Brain Disorders» (pp. 25-74). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK597496/>
- 14 Garcia S, Ramirez-gallego S, Luengo J. (2016, November). «Big data preprocessing: methods and prospect». Available at: <https://link.springer.com/article/10.1186/s41044-016-0014-0>
- 15 Kevin P. Murphy (2012, August) «Machine Learning A Probabilistic Perspective» (pp. 3-16). Available at: [https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning\\_%20A%20Probabilistic%20Perspective%20%5BMurphy%202012-08-24%5D.pdf](https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning_%20A%20Probabilistic%20Perspective%20%5BMurphy%202012-08-24%5D.pdf)
- 16 David A, Vassilvitskii S. (2007). «k-means++: The Advantages of Careful Seeding». Available at: <https://theory.stanford.edu/~sergei/papers/kMeansPP-soda.pdf>
- 17 Alanazi A. (2022) «Using machine learning for healthcare challenges and opportunities». doi: <https://doi.org/10.1016/j.imu.2022.100924>
- 18 Xie H, Ding Z, Li P, Xu X (2024, July). «An enhanced structural developmental neural network with information saturation for continual unsupervised learning». doi: <https://doi.org/10.1016/j.neucom.2024.127666>
- 19 Liyakathunisa Syed, Abdullah Alsaedi, Lina A. Alhuri, Hutaf R. Aljohani, (2023, September). «Hybrid weakly supervised learning with deep learning technique for detection of fake news from cyber propaganda». doi: <https://doi.org/10.1016/j.array.2023.100309>
- 20 Kavzoglu T. (2009, July). «Increasing the accuracy of neural network classification using refined training data». 24(7):850-858 Available at: [https://www.researchgate.net/publication/223565557\\_Increasing\\_the\\_accuracy\\_of\\_neural\\_network\\_classification\\_using\\_refined\\_training\\_data](https://www.researchgate.net/publication/223565557_Increasing_the_accuracy_of_neural_network_classification_using_refined_training_data)
- 21 Manandhar R, Inakwu O. A, Ancev T. (2009) «Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement». 1(3), 330-344. doi: <https://doi.org/10.3390/rs1030330>
- 22 Mishra S, Shaw K, Mishra D. (2022 May). « Improving the Accuracy of Ensemble Machine Learning Classification Models Using a Novel Bit-Fusion Algorithm for Healthcare AI Systems». doi: <https://doi.org/10.3389/fpubh.2022.858282>
- 23 I-Tung Yang, Husada W. (2017, March) «Improving Classification Accuracy for Single-loop Reliability-based Design Optimization». IMECS 2017. Available at: [https://www.iaeng.org/publication/IMECS2017/IMECS2017\\_pp1036-1040.pdf](https://www.iaeng.org/publication/IMECS2017/IMECS2017_pp1036-1040.pdf)
- 24 Tsiliki G, Kossida S. (2011, July) «Fusion methodologies for biomedical data». doi: <https://doi.org/10.1016/j.jprot.2011.07.001>
- 25 Kilic E, Alpaydin E. (2010, December) «Learning the areas of expertise of classifiers in an ensemble». doi: <https://doi.org/10.1016/j.procs.2010.12.014>
- 26 Sharma A, Pramod Kumar Mishra. (2022, August) «Image enhancement techniques on deep learning approaches for automated diagnosis of COVID-19 features using CXR images». doi: <https://doi.org/10.1007/s11042-022-13486-8>
- 27 Mauricio Goetz M, Maria C. Torres-Madronero, (2020, August) «Joint pre-processing framework for two-dimensional gel electrophoresis images based on nonlinear filtering, background correction and normalization techniques» . doi: <https://doi.org/10.1186/s12859-020-03713-0>
- 28 McCombe D. K, Craig G. S. (2021, Aug) «HistoClean: Open-source software for histological image pre-processing and augmentation to improve development of robust convolutional neural networks» doi: <https://doi.org/10.1016/j.csbj.2021.08.033>
- 29 Goetz M. M, Torres-Madroño M. C, Röthlisberger S, Delgado-Trejos E. (2018, February) «Preprocessing of 2-Dimensional Gel Electrophoresis Images Applied to Proteomic Analysis: A Review». doi: <https://doi.org/10.1016%2Fj.gpb.2017.10.001>
- 30 [Zhigang Ling, Yan Liang, Yaonan Wang, He Shen, Xiao Lu](#), (2015, November). «Adaptive extended piecewise histogram equalisation for dark image enhancement » doi: <https://doi.org/10.1049/iet-ipr.2014.0580>

## References

1. Tayefi M, Phuong D, Dalianis H. (2021, February). «Challenges and opportunities beyond structured data in analysis of electronic health records». doi: <https://doi.org/10.1002/wics.1549>
2. Gerard P, Peter V, Frede M. D. (2016, December). «Bigger data, bigger problems. J Orthop Trauma». doi: <https://doi.org/10.1097/bot.0000000000000463>
3. Dalianis H. (2018, May). «Evaluation Metrics and Evaluation. In Clinical Text Mining» (pp.45-53)  
a. doi: [https://link.springer.com/chapter/10.1007/978-3-319-78503-5\\_6](https://link.springer.com/chapter/10.1007/978-3-319-78503-5_6)
4. Tammy J, Jaimie L. Gradus, Anthony J. Rosellini. (2020, May). «Supervised Machine Learning: A Brief Primer». doi: <https://doi.org/10.1016/j.beth.2020.05.002>
5. Krzysztof J. Cios, Moore G. W. (2002, September) «Uniqueness of medical data mining». doi: [https://doi.org/10.1016/s0933-3657\(02\)00049-0](https://doi.org/10.1016/s0933-3657(02)00049-0)

6. Girish C, Ferat S. (2014, January). «A survey on feature selection methods». doi: <https://doi.org/10.1016/j.compeleceng.2013.11.024>
7. Jadczyk T, Kiwic O, Raj M, Khandwalla (2019, September) «Feasibility of a voice-enabled automated platform for medical data collection: CardioCube». doi: <https://doi.org/10.1016/j.ijmedinf.2019.07.001>
8. Choong Ho Lee and Hyung-Jin Yoon (2017, March) «Medical big data: promise and challenges» doi: <https://doi.org/10.23876/j.krcp.2017.36.1.3>
9. [Maryam Tayefi, Phuong Ngo, Taridzo Chomutare](#), (2021, February) «Challenges and opportunities beyond structured data in analysis of electronic health records». doi: <https://doi.org/10.1002/wics.1549>.
10. Han J, Zhengya S, Hongwei H. (2015, September). «Selecting feature subset with sparsity and low redundancy for unsupervised learning». doi: <https://doi.org/10.1016/j.knosys.2015.06.008>
11. David A. Nardone. (1990). «Collecting and Analyzing Data: Doing and Thinking. In Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition» [H. Kenneth Walker, W. Dallas Hall, J. Willis Hurst](#).
12. Hyunwoo C, Seongjun P, Jinbum K, Yangmo Y, (2024, September), «Deep coherence learning: An unsupervised deep beamformer for high quality single plane wave imaging in medical ultrasound», doi:<https://doi.org/10.1016/j.ultras.2024.107408>
13. Faouzi J, Colliot O. (2023, July). «Classic Machine Learning Methods. In Machine Learning for Brain Disorders» (pp. 25-74). Available at: <https://www.ncbi.nlm.nih.gov/books/NBK597496/>
14. Garcia S, Ramirez-gallego S, Luengo J. (2016, November). «Big data preprocessing: methods and prospect». Available at: <https://link.springer.com/article/10.1186/s41044-016-0014-0>
15. Kevin P. Murphy (2012, August) «Machine Learning A Probabilistic Perspective» (pp. 3-16). Available at: [https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning\\_%20A%20Probabilistic%20Perspective%20%5BMurphy%202012-08-24%5D.pdf](https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning_%20A%20Probabilistic%20Perspective%20%5BMurphy%202012-08-24%5D.pdf)
16. David A, Vassilvitskii S. (2007). «k-means++: The Advantages of Careful Seeding». Available at: <https://theory.stanford.edu/~sergei/papers/kMeansPP-soda.pdf>
17. Alanazi A. (2022) «Using machine learning for healthcare challenges and opportunities». doi: <https://doi.org/10.1016/j.jimu.2022.100924>
18. Xie H, Ding Z, Li P, Xu X (2024, July). «An enhanced structural developmental neural network with information saturation for continual unsupervised learning». doi: <https://doi.org/10.1016/j.neucom.2024.127666>
19. Liyakathunisa Syed, Abdullah Alsaedi, Lina A. Alhuri, Hutaf R. Aljohani, (2023, September). «Hybrid weakly supervised learning with deep learning technique for detection of fake news from cyber propaganda». doi:<https://doi.org/10.1016/j.array.2023.100309>
20. Kavzoglu T. (2009, July). «Increasing the accuracy of neural network classification using refined training data». 24(7):850-858 Available at: [https://www.researchgate.net/publication/223565557\\_Increasing\\_the\\_accuracy\\_of\\_neural\\_network\\_classification\\_using\\_refined\\_training\\_data](https://www.researchgate.net/publication/223565557_Increasing_the_accuracy_of_neural_network_classification_using_refined_training_data)
21. Manandhar R, Inakwu O. A, Ancev T. (2009) «Improving the Accuracy of Land Use and Land Cover Classification of Landsat Data Using Post-Classification Enhancement». *J(3)*, 330-344. doi: <https://doi.org/10.3390/rs1030330>
22. Mishra S, Shaw K, Mishra D. (2022 May). « Improving the Accuracy of Ensemble Machine Learning Classification Models Using a Novel Bit-Fusion Algorithm for Healthcare AI Systems». doi: <https://doi.org/10.3389/fpubh.2022.858282>
23. I-Tung Yang, Husada W. (2017, March) «Improving Classification Accuracy for Single-loop Reliability-based Design Optimization». IMECS 2017. Available at: [https://www.iaeng.org/publication/IMECS2017/IMECS2017\\_pp1036-1040.pdf](https://www.iaeng.org/publication/IMECS2017/IMECS2017_pp1036-1040.pdf)
24. Tsiliki G, Kossida S. (2011, July) «Fusion methodologies for biomedical data». doi: <https://doi.org/10.1016/j.jprot.2011.07.001>
25. Kilic E, Alpaydin E. (2010, December) «Learning the areas of expertise of classifiers in an ensemble». doi: <https://doi.org/10.1016/j.procs.2010.12.014>
26. Sharma A, Pramod Kumar Mishra. (2022, August) «Image enhancement techniques on deep learning approaches for automated diagnosis of COVID-19 features using CXR images». doi: <https://doi.org/10.1007/s11042-022-13486-8>
27. Mauricio Goetz M, Maria C. Torres-Madronero, (2020, August) «Joint pre-processing framework for two-dimensional gel electrophoresis images based on nonlinear filtering, background correction and normalization techniques» . doi: <https://doi.org/10.1186/s12859-020-03713-0>
28. McCombe D. K, Craig G. S. (2021, Aug) «HistoClean: Open-source software for histological image pre-processing and augmentation to improve development of robust convolutional neural networks» doi: <https://doi.org/10.1016/j.csbj.2021.08.033>
29. Goetz M. M, Torres-Madronero M. C, Röthlisberger S, Delgado-Trejos E. (2018, February) «Preprocessing of 2-Dimensional Gel Electrophoresis Images Applied to Proteomic Analysis: A Review». doi: <https://doi.org/10.1016%2Fj.gpb.2017.10.001>
30. [Zhigang Ling, Yan Liang, Yaonan Wang, He Shen, Xiao Lu](#), (2015, November). «Adaptive extended piecewise histogram equalisation for dark image enhancement » doi: <https://doi.org/10.1049/iet-ipr.2014.0580>

<https://doi.org/10.31891/2219-9365-2024-79-10>

УДК 004.056.53

ТИТОВА Віра

Хмельницький національний університет

<https://orcid.org/0000-0001-8668-4834>

e-mail: [titovav@khmnu.edu.ua](mailto:titovav@khmnu.edu.ua)

КЛЬОЦ Юрій

Хмельницький національний університет

<https://orcid.org/0000-0002-3914-0989>

e-mail: [klots@khmnu.edu.ua](mailto:klots@khmnu.edu.ua)

ВОЛИНЕЦЬ Віталій

<https://orcid.org/0009-0006-7999-1290>

[mailto:volynets1026@gmail.com](mailto:mailto:volynets1026@gmail.com)

ПЕТЛЯК Наталія

Хмельницький національний університет

<https://orcid.org/0000-0001-5971-4428>

e-mail: [npetlyak@khmnu.edu.ua](mailto:npetlyak@khmnu.edu.ua)

ОГОРОДНИК Максим

Хмельницький національний університет

e-mail: [maks737271@gmail.com](mailto:maks737271@gmail.com)

## РОЗРОБЛЕННЯ ПОЛІТИКИ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ПРИВАТНОГО ПІДПРИЄМСТВА

*У даній роботі представлено методику формування політики інформаційної безпеки у приватному секторі. Проведено аналіз систем даних, представлено їх техніко-економічні характеристики та визначено основні проблеми та завдання захисту інформації. Проведено порівняльний аналіз методів та засобів захисту інформації на аналогічних об'єктах інформаційної діяльності. Вибрано та продемонстровано методи захисту інформації в корпоративній мережі компанії за допомогою адміністративних заходів із запобігання загрозам інформаційної безпеки.*

*Ключові слова: політика інформаційної безпеки, методи та засоби захисту інформації.*

TITOVA Vira, KLOTS Yurii

Khmelnytskyi National University

VOLYNETS Vitalii

PETLIAK Nataliia, OHORODNYK Maksym

Khmelnytskyi National University

## DEVELOPING THE INFORMATION SECURITY POLICY OF A PRIVATE ENTERPRISE

*Based on the analysis of the main provisions of the information protection theory, it was established that in order to create an information security policy, it is necessary to develop a number of documents and instructions aimed at information protection. And in no case should stop at one method of information protection, otherwise data protection will be at risk. Data protection must be comprehensive. The comprehensive policy of information security covers the development, production and installation of technical means of protection, as well as regular inspections of the information equipment used. The development of this policy is as follows: identification of deficiencies in the company's current information protection; identification of types of threats that may arise as a result of deficiencies in the protection of information systems of the enterprise; selection of methods and ways of solving existing problems.*

*As a solution, a set of measures was developed, which consists of administrative decisions that regulate the possibility of information leakage due to the influence of the human factor. Based on the analysis of the main methods and means of information protection, it was established that organizational and legal methods and means of information protection should be aimed at countering threats to information security, reducing risks and effectively handling incidents in order to ensure a sufficient level of data protection for a long time. The evaluation of the effectiveness of the proposed measures through economic substantiation proved their feasibility of implementation in the organization.*

*Keywords: information security policy, methods and means of information protection.*

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Поява нових інформаційних технологій та розвиток потужних комп'ютерних систем для зберігання та обробки інформації підвищили вимоги до рівня захисту інформації та визначили необхідність розробки ефективних механізмів захисту інформації, сумісних із сучасними архітектурами зберігання даних.

Забезпечення захисту інформації на підприємстві – це безперервний процес, що включає контроль зовнішнього і внутрішнього середовища підприємства, організацію та проведення заходів щодо підтримки

стабільного функціонування локальної мережі та обчислювальної техніки, а також використання сучасних методів, що дозволяють мінімізувати втрати від витоку інформації. Для захисту інформації, як у мережі, так і на виробництві, компаніям необхідно сформулювати певні правила та норми, що регламентують поведінку співробітників для забезпечення безпеки, а також описати технічні та програмні засоби захисту інформації. На це й спрямовано розроблення політики безпеки.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Політика інформаційної безпеки компанії зазвичай виражається в серії документів, що відображають вимоги до захисту даних і основні напрямки діяльності компанії щодо безпеки [1]. Існує три основні рівні розробки політики безпеки: верхній, середній і нижній [2].

На верхньому рівні політики безпеки даних організації необхідно: сформулювати та продемонструвати ставлення адміністрації підприємства до системи захисту інформації та відобразити основні цілі та завдання в цій галузі; розробити індивідуальні політики безпеки, інструкції та правила, за допомогою яких регулюються окремі питання; інформувати співробітників організації про основні завдання та пріоритети в галузі інформаційної безпеки.

Політика інформаційної безпеки середнього рівня використовується для відображення корпоративних підходів і вимог, таких як: використання інформаційних систем; телекомунікаційних та інформаційних технологій, методів і підходів до обробки інформації; учасників процесів обробки інформації, від яких залежить забезпечення захисту інформації на підприємстві.

Нижній рівень політики безпеки використовується для опису конкретних процедур і документів для забезпечення інформаційної безпеки на підприємстві.

Етапи розроблення політики безпеки в організації включають: виконання оцінки особистого ставлення до загроз безпеці з боку власників і співробітників підприємства; проведення аналізу потенційно важливих інформаційних активів підприємства; виявлення існуючих загроз безпеки підприємства з подальшою оцінкою ризиків.

Розглянемо основні елементи політики інформаційної безпеки підприємства [3-5]. Захист передбачає використання організаційних засобів захисту, визначених політикою безпеки підприємства. На першому етапі необхідно визначити межі, в яких функціонуватиме політика інформаційної безпеки компанії та встановити критерії оцінки її результатів.

На етапі аналізу ризиків інформації визначають пріоритети обраних засобів захисту з розподілом їх за ступенем важливості на підприємстві, ідентифікують уразливість активів підприємства та визначають збитки. Результати аналізу ризиків інформаційної безпеки підприємства будуть застосовуватися у вигляді основи для планування роботи системи інформаційної безпеки, вибору найефективнішої стратегії та тактики. Для підвищення ефективності політики безпеки застосовуються такі прийоми, як групове визначення з використанням атрибутів та мандатне керування доступом.

Багато підприємств використовують глобальні та локальні політики безпеки, засновані на принципах управління інформаційною безпекою. Глобальні політики інформаційної безпеки спрямовані на забезпечення захисту інформації на рівні бізнес-процесів підприємства, тоді як локальні політики формуються на рівні захисту даних підприємства [4-6].

Глобальна політика підприємства являє собою правила безпеки, що описують можливі взаємодії між об'єктами, які потребують захисту інформації. Локальні політики безпеки підприємства використовуються для налаштування засобів захисту інформації, реплікації налаштувань вузлів і подальших коригувань. Зазвичай, локальні політики безпеки підприємства містять правила, які регулюють з'єднання і змінюють конфігурацію мережних пристроїв.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Отже, на основі проведеного огляду можна зробити висновок, що для розроблення політики безпеки приватного підприємства необхідно вирішити такі завдання: оцінити поточний стан інформаційної безпеки підприємства; виявити порушення в захисті інформаційної безпеки, а також виявлення найімовірніших загроз інформації; розробити пропозиції щодо реалізації адміністративних заходів із запобігання загрозам інформаційної безпеки.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Для захисту даних приватного підприємства необхідно користувачів розділити на групи з відповідними правами:

- ✓ administrator – адміністратори мережі (створення та управління політиками інформаційної безпеки, глобальні налаштування мережі тощо);
- ✓ engineer – облікові записи для повсякденного обслуговування інформаційно-обчислювальної техніки;



- ✓ worker – обліковий запис стандартного користувача (співробітника організації) з обмеженими правами;
- ✓ guest – обмежений обліковий запис (у разі необхідності доступу не співробітників організації).

Для ідентифікації користувача потрібен унікальний запис кожного користувача, який включено до відповідної групи. Тим самим здійснюється розмежування доступу (табл. 1).

Таблиця 1

Групи користувачів та їхні права

Дії	Guest	Worker	Engineer	Administrator
Створення та зміни груп користувачів	Ні	Ні	Ні	Так
Зміна налаштувань мережі	Ні	Ні	Ні	Так
Підключення до мережі нових робочих станцій	Ні	Ні	Ні	Так
Зміна налаштувань серверів	Ні	Ні	Ні	Так
Зміна прав доступу	Ні	Ні	Ні	Так
Встановлення додатків та ПЗ	Ні	Ні	Так	Так
Доступ до Інтернету	Ні	Так	Так	Так
Доступ до корпоративної електронної пошти	Ні	Так	Так	Так
Доступ до корпоративного чату	Ні	Так	Так	Так
Можливість завантажувати файли	Ні	Ні	Так	Так
Запис файлів	"Мої документи"	"Мої документи", "Робочий стіл"	Будь-яка папка на робочому ПК	Будь-який ПК мережі
Підключення флеш-дисків, зовнішніх дисків.	Ні	Ні	Так	Так

Відправною точкою для визначення економічної ефективності запропонованого підходу є очевидне припущення: з одного боку, порушення інформаційної безпеки завдає певної шкоди; з іншого боку, забезпечення інформаційної безпеки коштує дорого. Загальна очікувана вартість захисту може бути виражена як сума вартості захисту та збитків від порушення. Очевидно, що оптимальним рішенням є розподіл коштів на захист інформації таким чином, щоб мінімізувати загальну вартість захисту [7].

Також зрозуміло, що економічна ефективність заходів з інформаційної безпеки визначається розміром відверненого збитку або розміром зниження ризиків для інформаційних активів організації.

Достатньо визначити лише рівень збитків, оскільки оптимальне рішення проблеми доцільного рівня витрат на захист полягає в тому, що цей рівень дорівнює рівню збитків, які очікуються в разі порушення безпеки. Як одна з методик визначення рівня витрат можливе використання такої емпіричної залежності очікуваних втрат (ризиків)  $R$  від  $i$ -ї загрози інформації [8-9]:

$$R_i = 10^{T_i + L_i - 4}, \quad (1)$$

де  $T_i$  – коефіцієнт, що характеризує можливу частоту виникнення відповідної загрози;  $L_i$  – коефіцієнт, що характеризує значення можливого збитку в разі її виникнення.

Сумарна вартість втрат визначається формулою:

$$R = \sum_{i=1}^N R_i, \quad (2)$$

де  $N$  – кількість можливих загроз інформаційним активам.

При розрахунку сумарного показника рекомендується виходити з того, що загрози конфіденційності, цілісності та доступності здійснюються порушником незалежно. Іншими словами, припускається, що цілісність інформації порушена діями порушника, але її зміст залишається невідомим порушнику (конфіденційність не порушена), а авторизовані користувачі все ще мають доступ до активу, хоча й у спотвореному вигляді.

Для прикладу розглянемо інформаційні активи приватної компанії (табл. 2). Розрахунки показують, що ризик економічних втрат для цієї компанії становить приблизно 1 080 000 гривень. З цього можна зробити висновок, що це дуже значні втрати для підприємства. Для того, щоб зрозуміти, наскільки ефективною є розроблена політика інформаційної безпеки, необхідно розрахувати показники економічної ефективності проекту.

Таблиця 2

**Величини втрат (ризиків) для інформаційних ресурсів до впровадження розробленої політики безпеки**

Інформаційний актив	Загроза	Величина втрат (тис. грн.)
Проектна документація, розроблена організацією	конфіденційності	100
	цілісності	500
	доступності	20
Особисті дані клієнта	конфіденційності	300
	цілісності	20
	доступності	20
Особисті відомості про співробітників	конфіденційності	100
	цілісності	10
	доступності	10
Сумарна величина втрат		1 080

Для проведення розрахунків необхідно отримати дані про передбачуваний розмір втрат (ризик) ключових інформаційних ресурсів після впровадження/модернізації інформаційної безпеки. Результати базуються на висновках експертних досліджень (див. табл. 3).

Таблиця 3

**Величини втрат (ризиків) для інформаційних ресурсів після впровадження розробленої політики безпеки**

Інформаційний актив	Загроза	Величина втрат (тис. грн.)
Проектна документація, розроблена організацією	конфіденційності	10
	цілісності	50
	доступності	2
Особисті дані клієнта	конфіденційності	30
	цілісності	2
	доступності	2
Особисті відомості про співробітників	конфіденційності	10
	цілісності	1
	доступності	1
Сумарна величина втрат		108

Отже, можна зробити висновки, що впровадження розробленої політики безпеки дозволяє знизити можливі збитки в 10 разів, тобто витрати політики безпеки окупляться вже в першому кварталі. І це є зовсім невеликим навантаженням на фінансову систему організації.

**ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ  
І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ**

Аналіз основних положень теорії інформаційної безпеки показує, що для створення політики інформаційної безпеки необхідно розробити низку документів та інструкцій з метою захисту інформації. Крім того, захист інформації не повинен обмежуватися лише одним методом захисту інформації. Захист даних повинен бути комплексним. Комплексна політика захисту інформації поширюється на розробку, виготовлення та встановлення технічних засобів захисту, а також на регулярну перевірку інформаційного обладнання, що використовується.

Розробка такої політики включає в себе виявлення поточних недоліків інформаційної безпеки підприємства, визначення типів загроз, які можуть виникнути через недоліки в захисті інформаційних систем підприємства, а також вибір шляхів і засобів для вирішення існуючих проблем.

В якості рішення автори розробили комплекс заходів, що складається з адміністративних рішень, які регламентують можливість витоку інформації через вплив людського фактору. На основі аналізу основних методів та заходів захисту інформації встановлено, що організаційно-правові методи та заходи захисту інформації повинні бути спрямовані на протидію загрозам інформаційній безпеці, зниження ризиків, ефективне реагування на інциденти та забезпечення достатнього рівня захисту даних протягом тривалого часу.

Ефективність запропонованих заходів оцінено за допомогою економічного обґрунтування, що свідчить про їх доцільність в організації.

**Література**

1. Основи інформаційної безпеки: навчальний посібник/ В.А. Лужецький, А.Д. Кожухівський, О.П. Войтович. Вінниця: ВНТУ, 2013. 221 с.
2. ДСТУ ISO/IEC 27005:2019. Інформаційні технології. Методи захисту. Управління ризиками інформаційної безпеки/ Нац. стандарт України. Вид. офіц [Чинний від 2019-11-01]. Київ: ДП «УкрНДНЦ», 2019. 76 с.

3. Формування моделі політики інформаційної безпеки на основі концепцій “глибинного захисту”/ Д.В. Дячков// Підприємництво і торгівля. 2019. № 25. С. 116-121.
4. Політика інформаційної безпеки в системах інформаційно-аналітичного забезпечення підтримки прийняття організаційних рішень / С.М. Чуруброва // Проблеми програмування. 2016. № 4. С. 97-103.
5. Розробка політики інформаційної безпеки комп’ютерного контролю знань/ Н. Кухарська// Вісник Львівського державного університету безпеки життєдіяльності. 2017. №16. С.34-39.
6. Політика інформаційної безпеки об’єкта/ Ю. Хохлачова// Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. 2012 р. Вип. 2 (24). С. 23-29.
7. Оцінювання ефективності рішень в системах захисту інформації/ В. Ю. Тітова, О. С. Андрощук, В. С. Орленко, І. М. Шевчук, В. С. Даценко // Вісник Хмельницького національного університету. Технічні науки. 2020. № 5. С. 307–310.
8. Комерційна діяльність. Навч. посібник/ Л.Г. Филевич, Л.О. Попова, О.М. Прядко, Т.Л. Мітєєва, Л.А. Прибилович. Харків: ХДУХТ, 2014. 225 с.
9. Інформаційна політика в системі забезпечення фінансової безпеки держави/ А.Д. Глушко, В.В. Пантась, С.Р. Бабенко// «Ефективна економіка». 2022. №2.

### References

1. Osnovy informatsiinoi bezpeky : navchalnyi posibnyk/ V.A. Luzhetskyyi, A.D. Kozhukhivskyyi, O.P. Voitovych. Vinnytsia: VNTU, 2013. 221 s.
2. DSTU ISO/IEC 27005:2019. Informatsiini tekhnologii. Metody zakhystu. Upravlinnia ryzykamy informatsiinoi bezpeky / Nats. standart Ukrainy. Vyd. ofits [Chynnyi vid 2019-11-01]. Kyiv: DP «UkrNDNTs», 2019. 76 s.
3. Formuvannya modeli polityky informatsiinoi bezpeky na osnovi kontseptsii “hlybynnoho zakhystu/ D.V. Diachkov// Pidpriemnytstvo i torhivlia. 2019. № 25. S. 116-121.
4. Polityka informatsiinoi bezpeky v systemakh informatsiino-analitychnoho zabezpechennia pidtrymky pryiniattia orhanizatsiinykh rishen / Churubrova S. M. // Problemy prohramuvannia. 2016. № 4. S. 97-103.
5. Rozrobka polityky informatsiinoi bezpeky kompiuternoho kontroliu znan/ N. Kukharska// Visnyk Lvivskoho derzhavnoho universytetu bezpeky zhyttiediialnosti. 2017. №16. S.34-39.
6. Polityka informatsiinoi bezpeky obiekta/ Yu. Khokhlachova// Pravove, normatyvne ta metrolohichne zabezpechennia systemy zakhystu informatsii v Ukraini. 2012 r. Vyp. 2 (24). S. 23-29.
7. Otsiniuvannya efektyvnosti rishen v systemakh zakhystu informatsii/ V. Yu. Titova, O. S. Androshchuk, V. S. Orlenko, I. M. Shevchuk, V. S. Datsenko // Herald of Khmelnytskyi National University. Technical sciences. 2020. № 5. S. 307–310.
8. Komertsiiina diialnist. Navch. posibnyk/ L.H. Fylevych, L.O. Popova, O.M. Priadko, T.L. Mitiaieva, L.A. Prybylovych. Kharkiv: KhDUKhT, 2014. 225 s.
9. Informatsiina polityka v systemi zabezpechennia finansovoi bezpeky derzhavy/ A.D. Hlushko, V.V. Pantas, S.R. Babenko// «Efektyvna ekonomika». 2022. №2.

<https://doi.org/10.31891/2219-9365-2024-79-11>

УДК 620

НАГІРНИЙ Василь

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0005-7825-8858>

e-mail: [vasyl.i.nahirnyi@lpnu.ua](mailto:vasyl.i.nahirnyi@lpnu.ua)

КУЦЬ Віктор

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0003-1698-414X>

e-mail: [viktor.r.kuts@lpnu.ua](mailto:viktor.r.kuts@lpnu.ua)

## РОЗВИТОК ЗАСОБІВ ДІАГНОСТИКИ СЕНСОРНИХ МЕРЕЖ ДВОКОНТУРНИХ СИСТЕМ ОПАЛЕННЯ

*Дана стаття присвячена аналізу факторів ризику роботи усіх вузлів газових двоконтурних теплогенераторів, які вимагають постійного моніторингу, опрацювання варіантів їх побудови в поширених конструкціях, та дослідження шляхів інформаційного спраження сенсорних мереж з іншими системами кліматичного контролю (СКК).*

*Ключові слова: системи опалення, теплогенератори, системи кліматичного контролю (СКК), двофункційний теплогенератор, двоконтурні системи опалення.*

NAHIRNYI Vasyl, KUTS Viktor

Lviv Polytechnic National University

## DEVELOPMENT OF DIAGNOSTIC TOOLS OF SENSOR NETWORKS OF DOUBLE-CIRCUIT HEATING SYSTEMS

*The heating system is one of the key elements of modern housing, which provides a level of comfortable climatic environment, as a determining factor of its market competitiveness.*

*The rapid development of microprocessor technologies, personal computers, individual means of communication, the Internet, and ultimately artificial intelligence over the past few decades opens up extremely wide opportunities for improving heating systems as a component of climate control of all types of buildings and structures.*

*Together with the climate control systems, other elements of the housing infrastructure developed at a rapid pace, in particular energy. Alternative sources of energy, such as solar, wind, low-power generators, which thanks to modern means of electronic and computer technology can quickly adapt to local natural and climatic conditions, use various orientations of housing elements relative to the sides of the world, features of the terrain, insolation or shading, dominant wind directions, vegetation, etc.*

*Such building intersections can provide significant savings in energy resources, increase the level of comfort, thus capitalizing on investments in research into ways to improve climate control systems.*

*This article is devoted to the analysis of the risk factors of the operation of all nodes of gas double-circuit heat generators, which require constant monitoring, the development of options for their construction in common structures, and the study of ways of informational conjugation of sensor networks with other climate control systems (CCS).*

*Key words: heating systems, heat generators, climate control systems (CCS), dual-function heat generator, dual-circuit heating systems.*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Система опалення – один з ключових елементів сучасного житла, що забезпечує рівень комфортного кліматичного середовища, як визначального фактора його ринкової конкурентності.

Швидкий розвиток технологій мікропроцесорної техніки, персональних комп'ютерів, індивідуальних засобів комунікацій, мережі інтернету, в кінцевому - штучного інтелекту за останні кілька десятиріч відкриває надзвичайно широкі можливості для вдосконалення систем опалення, як компонента кліматичного контролю всіх типів будівель і споруд.

Разом з системами кліматичного контролю швидкими темпами розвивалися інші елементи житлової інфраструктури, зокрема енергетика. Масово впроваджуються в життя альтернативні джерела енергії, такі як сонячні, вітрові, генератори малої потужності, які завдяки сучасним засобам електронної та комп'ютерної техніки можуть швидко адаптуватися до локальних природньо-кліматичних умов, використовувати різноманітну орієнтацію елементів житла відносно сторін світу, особливості рельєфу, інсоляцію чи затінення, домінуючі вітрові напрями, рослинність, тощо.

Такі будівельні розв'язки можуть давати значну економію енергоресурсів, підвищувати рівень комфорту, капіталізуючи, таким чином, вкладення в дослідження шляхів вдосконалення систем кліматичного контролю (СКК).

Розширення функціональних можливостей СКК на основі сучасних сенсорних систем набуває актуальності не тільки з огляду на створення комфортного інтерфейсу для користувача, а й відкриває нові

напрями їх використання:

- моніторинг та діагностику режимів роботи усіх компонентів систем опалення, зокрема, газових теплогенераторів, в реальному часі;
- формування інформаційних масивів для обслуговування та ремонту обладнання;
- спостереження за параметрами систем протягом всього часу експлуатації, що є цінною інформацією для розробників та виробників.

Звичайно, такі ускладнення вимагають високої кваліфікації персоналу, що, за прогнозами аналітиків [1], в перспективі змін на ринку праці, може стати важливим фактором.

Особливої уваги потребують питання зміни стратегії використання сучасних сенсорів, які, завдяки помірній вартості і широким функціональним можливостям набувають автономного статусу, прив'язуються не до конкретних контрольно-вимірювальних функцій і пристроїв, а, радше, локалізуються у всіх важливих елементах конструкції в зонах технологічного процесу, що потребують безперервного контролю, а також є помічними для діагностики виробу у процесі обслуговування чи ремонту.

Такі сенсорні системи перетворюються на мережі, результати вимірювань передаються до пристроїв пам'яті, а їх комунікація з центральним керуванням відбувається за розробленими алгоритмами для організації керування всіма процесами і режимами роботи та діагностики.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою даної статті є аналіз будови, режимів роботи, факторів ризику основних вузлів двофункційних газових теплогенераторів та напрямів розвитку системи їх діагностики шляхом інтеграції сенсорної мережі в систему кліматичного контролю.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Позаяк, централізоване тепло- і водопостачання, яке домінувало кілька десятиліть тому, виявилось неефективним і у фінансово-економічному, і в енергетичному, і у експлуатаційному аспектах, на сьогодні в обладнанні індивідуального житла переважають системи опалення з теплогенераторами на природньому газі, що забезпечують гнучкість в технічних, експлуатаційних та економічних розв'язках. Індивідуальне опалення дозволяє самостійно регулювати температуру в помешканні, підтримувати потрібну температуру, вмикати та вимикати опалення в потрібний час, регулювати витрати на опалення [2].

Такі теплогенератори (в побуті — газові двоконтурні котли) виконують дві функції:

- теплопостачання;
- нагрівання води для санітарно-побутових потреб.

Основним елементом газового теплогенератора є система газового пальника і головного теплообмінника, які конструктивно об'єднані в камеру спалювання, що може бути "відкритою" (з природнім відведенням продуктів спалювання за рахунок різниці динамічних тисків в камері спалювання і назовні), або — "закритою" (з примусовою вентиляцією камери спалювання) [3].

Закрита камера спалювання має ряд переваг перед відкритою, бо:

- гарантує вищу безпеку, унеможливаючи потрапляння продуктів спалювання у приміщення;
- забезпечує краще спалювання, позаяк дає можливість регулювати співвідношення компонентів газоповітряної суміші;
- відкриває можливість розміщення всіх потрібних сенсорів у закритій камері спалювання для коректного моніторингу процесу горіння.

Теплообмінник забезпечує передачу тепла від газопарового потоку продуктів спалювання до теплоносія.

Природне відведення продуктів спалювання базується на різниці температур і, відповідно, густини газів у димарі, що створює тягу для відведення продуктів горіння, водночас як примусове димовидалення використовує вентилятори або аспіратори для створення необхідної тяги для видалення продуктів горіння. Обидві системи повинні включати аварійні механізми, такі як датчики диму, що автоматично відключають котел при виявленні проблем з тягою. Правильне управління тягою дозволяє підвищити загальну ефективність системи опалення, зменшити споживання пального і знижувати викиди забруднюючих речовин. Регулярне обслуговування системи відведення продуктів спалювання, включаючи чистку димаря та перевірку роботи вентиляторів, є важливим для збереження належної тяги і безпеки.

Значний вплив на ефективність теплопередачі має конденсація пари на стінках теплообмінника, наслідком чого є збільшення втрат тепла на повторне випаровування конденсату, та підвищення ризику корозії матеріалу теплообмінника.

Зменшення негативного впливу пароконденсату може досягатися як конструкцією теплообмінника, так і оптимізацією газоповітряної суміші.

Високу ефективність забезпечують т.зв. *конденсаційні* котли [4], в яких використовується подвійний теплообмінник, в одній частині якого відбувається початкове нагрівання теплоносія у зворотньому колі первинного контура за рахунок тепла, отриманого від конденсації пари зі системи відводу продуктів

спалювання, а в другій частині відбувається безпосереднє нагрівання теплоносія до потрібної температури. В такій структурі камери спалювання передбачений спеціальний відвід конденсату до каналізації.

Кількість конденсату, його хімічний склад, температура конденсату є важливими показниками якості як самого природного газу, так і якості процесу спалювання, і можуть бути використані для їх діагностики і регулювання.

Дуже ефективною конструкцією камери спалювання у конденсаційних котлах є сферична форма пальника зі сітковою поверхнею, виготовлена з каталітичного матеріалу, що забезпечує оптимальне спалювання газу з мінімізацією викидів NOx, які є найнебезпечнішими речовинами, що шкодять довкіллю і підлягають жорсткому нормуванню європейським екологічним правом [5].

У таких пальниках відбувається зміщення спектру теплового випромінювання в бік інфрачервоних хвиль, що дозволяє зменшити швидкість теплового потоку теплообмінника, стабілізує процес теплопередачі, запобігає локальному перегріву поверхонь теплообмінника, і, таким чином, продовжує його термін експлуатації.

Для діагностики процесу горіння в такому разі доцільно використати сучасні спектроскопічні сенсори, що можуть надати цінну інформацію для керування процесом горіння, а також діагностики стану пальника.

Двофункційний теплогенератор має два контури: теплопостачання та гарячого водопостачання для побутових потреб. Перемикання потоку теплоносія з одного контура в інший відбувається за допомогою двоходового клапана. Режими роботи контурів суттєво відрізняються. Так контур теплопостачання працює в режимі з малою швидкістю тапконосія зі сталою температурою, великим тепловим навантаженням та великою тепловою інерційністю. Контур гарячого водопостачання навантажений відносно малим, найчастіше пластинчастим теплообмінником проточного типу, а режим його роботи має піковий характер: швидко максимальне навантаження з різкою зміною температури теплоносія, і частими перехідними процесами нагрівання - охолодження.

Така різниця процесів потребує особливо ретельного контролю параметрів режимів роботи теплообмінників, обладнання їх відповідними сенсорами для керування системами спалювання газу та потоку теплоносія. Відповідно різними мають бути також діагностичні процедури, розробка яких передбачена наступними дослідженнями.

Тривала експлуатація теплообмінників може суттєво змінювати його технічні характеристики. Найхарактернішим видом таких змін є відкладення карбонатів на стінках, які ізолюють теплоносії від трубопровідної частини теплообмінника, збільшують його гідравлічний опір, звужують поперечний переріз, формують локальні перешкоди для теплоносія, спричиняючи точкові перегриви.

Такі відкладення є особливо дошкульними, а їх виявлення становить значну проблему.

Для дослідження стану теплообмінника нам видаються перспективними методи аналізу шуму при русі теплоносія в конкретних конструктивних елементах [6]. Вони можуть проводитися шляхом порівняння спектрів шуму в чистому теплообміннику, а також під час сервісних оглядів, що будуть інформативними для прийняття рішення про необхідність його промивання, або заміни, якщо результати свідчать про суттєве стоншення стінок трубопровідної частини і виникнення небезпеки розгерметизації гідравлічного кола.

Особливе місце в системі моніторингу мають займати питання безпеки двох частин теплогенератора: газової і електричної.

В газовій частині надзвичайно важливим є, так званий, контроль полум'я. Система контролю полум'я формує сигнал дозволу на відкриття головного газового клапана після ініціації режиму запалювання.

Поширені два варіанти будови системи контролю полум'я: зі стаціонарною свічкою або з електронним запалом. Кожна із систем має свої переваги і недоліки.

Стаціонарна свічка горить протягом усього робочого режиму теплогенератора, а її полум'я скероване на електрод гарячого з'луту термопари, термо-е.р.с якої генерує струм безпекового реле для утримання клапана свічки у відкритому стані. Таким чином, у разі аварійного зникнення газу у вхідній газовій мережі свічка гасне, термопара припиняє генерацію струму, а запобіжний клапан газового автомата закривається, не дозволяючи автоматичний запуск теплогенератора при повторній появі газу в мережі.

Така система захисту дуже надійна і використовується для теплогенераторів, що можуть працювати без електроенергії. Сервісна процедура мусить обов'язково містити діагностику цієї системи.

Принцип дії системи контролю полум'я з електронним захистом ґрунтується на вимірюванні іонізаційного струму, що виникає при успішному запалюванні контрольної допоміжної свічки після ініціації режиму запалювання пальника. Іонізаційний струм є відносно малим (одиниці, десятки мікроампер) і істотно залежить від стану поверхні електродів, між якими виникає іонізаційний струм.

Діагностика стану електродів та вимірювання іонізаційного струму є достатньо складною проблемою, має кілька поширених варіантів розв'язку, які потребують детальнішого розгляду та аналізу.

Експлуатація теплогенераторів з електронним контролем полум'я в умовах окремих районів м. Львова пов'язана з певними проблемами через особливості будови енергетичної мережі, що здійснюється за

схемою трикутника 3 х 220 В без нульового проводу. Системи контролю полум'я серійних теплогенераторів побудована для роботи в електричних мережах за схемою трифазної зірки з лінійною напругою 3х380 В з нульовим проводом, у якій напруга живлення побутових споживачів відбувається фазною напругою 220 В відносно нуля.

У будинкових мережах за схемою трикутника, на жаль, немає можливості забезпечення стабільності низького потенціалу одного з полюсів мережі. Тому, виникають збої в роботі системи захисту, відбувається блокування запуску котла. Багаторічний досвід експлуатації показує можливість розв'язку цієї проблеми шляхом електроживлення теплогенератора через спеціальний пристрій ( “Трансформатор гальванічної розв'язки ТРГ-300” ) [7]. Однак, ця проблема не має нормативно прийнятого розв'язку і потребує детального схемотехнічного аналізу і дослідження.

Необхідні також дослідження можливості використання ТРГ-300 для роботи котлів опалення від зарядних станцій, таких як EcoFlow, Bluetti, Anker та інші.

Вартим уваги є також діагностика якості електричної енергії від альтернативних джерел живлення (джерела безперебійної напруги), якими вимушено користується значна частина населення України, і які не завжди видають напругу, відповідну до стандартів якості електрики. Особливе значення має форма синусоїди, яка нормується відповідними параметрами.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

**Розвиток** сенсорних мереж для діагностики стану основних елементів газових двофункційних теплогенераторів побутових систем кліматконтролю є актуальним і перспективним завданням. Проведений нами аналіз факторів ризику показує, що зміна парадигми застосування сенсорів від їх використання як елементів певних контрольно-вимірювальних пристроїв до організації інтегральної сенсорної мережі, відкриває можливості суттєво розширити коло пристроїв чи процесів, які підлягають контролю та діагностиці .

Проаналізовані причини проблем, що виникають у процесі тривалої експлуатації в гідравлічному колі теплогенератора, зокрема, відмінності теплового навантаження у контурах системи опалення та проточного постачання теплої санітарної води, сформульована задача організації динамічного контролю температури теплообмінника гарячого водопостачання і формування на цій основі сигналів керування системою пальника.

Показана необхідність детального аналізу проблеми нестабільності системи контролю плум'я у випадку живлення теплогенератора від електромережі за схемою трикутника 3х220 В, поширену в окремих районах м. Львова, та опрацювання можливих варіантів її розв'язку.

Окреслені шляхи використання сенсорної мережі для побудови діагностичної системи основних структурних елементів теплогенератора , які дозволяють постійно моніторити стан системи, виявляти аномалії та швидко реагувати на несправності..

Враховуючи зростаючу складність та інтеграцію сенсорних мереж у будівельні технології, питання розвитку нових методів діагностики набуває особливої актуальності.

Крім того, сучасні тенденції до автоматизації та розумних будівель вимагають впровадження більш інтелектуальних діагностичних інструментів, які можуть не лише виявляти, але й передбачати можливі несправності. Це дозволяє значно знизити витрати на обслуговування та підвищити загальну ефективність систем кліматконтролю.

### References

- [1] <https://www.linkedin.com/pulse/future-work-how-technology-changing-workplace-joy-o-konyeha-loorf>
- [2] Переваги та недоліки індивідуальної системи опалення URL: <https://vn.com.ua/ua/news/preimuschestva-i-nedostatki-individualnoj-sistemyotoplenija> .
- [3] <https://www.velocityboilerworks.com/documents/106376-02-Raptor-.pdf>
- [4] <https://www.sciencedirect.com/science/article/pii/S019689041730016X>
- [5] ВИМІРЮВАННЯ ВИКИДІВ NOx ГАЗОВИХ КОТЛІВ | MEASURING AND COMPUTING DEVICES IN TECHNOLOGICAL PROCESSES (khmnu.edu.ua)
- [6][https://elib.lntu.edu.ua/sites/default/files/elib\\_upload/%D0%95%D0%9D%D0%9A%20%D0%B3%D0%BE%D1%82%D0%BE%D0%B2%D0%BE/other/tema\\_7\\_vimiryuvannya\\_tovshchini\\_detalej\\_akustichnim\\_sposobom.pdf](https://elib.lntu.edu.ua/sites/default/files/elib_upload/%D0%95%D0%9D%D0%9A%20%D0%B3%D0%BE%D1%82%D0%BE%D0%B2%D0%BE/other/tema_7_vimiryuvannya_tovshchini_detalej_akustichnim_sposobom.pdf)
- [7][https://uk.wikipedia.org/wiki/%D0%93%D0%B0%D0%BB%D1%8C%D0%B2%D0%B0%D0%BD%D1%96%D1%87%D0%BD%D0%B0\\_%D1%80%D0%BE%D0%B7%D0%B2%27%D1%8F%D0%B7%D0%BA%D0%B0](https://uk.wikipedia.org/wiki/%D0%93%D0%B0%D0%BB%D1%8C%D0%B2%D0%B0%D0%BD%D1%96%D1%87%D0%BD%D0%B0_%D1%80%D0%BE%D0%B7%D0%B2%27%D1%8F%D0%B7%D0%BA%D0%B0)

<https://doi.org/10.31891/2219-9365-2024-79-12>

УДК 004.021

ПИРИГ Ярослав

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0001-2104-8439>

e-mail: [yaroslavpyrih@gmail.com](mailto:yaroslavpyrih@gmail.com)

ПИРИГ Юлія

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-8973-4005>

e-mail: [yuliia.v.klymash@lpnu.ua](mailto:yuliia.v.klymash@lpnu.ua)

## БАГАТОКРИТЕРІАЛЬНИЙ ПІДХІД НА ОСНОВІ ГЕНЕТИЧНОЇ ЕВОЛЮЦІЇ ДЛЯ ПОШУКУ ОПТИМАЛЬНОГО МАРШРУТУ ПЕРЕДАЧІ ДАНИХ У БЕЗПРОВІДНИХ СЕНСОРНИХ МЕРЕЖАХ

В роботі наведено багатокритеріальний метод маршрутизації даних у безпроводних сенсорних мережах на основі принципів генетичної еволюції, із врахуванням динамічно змінної топології та радіусів дії вузлів. Представлено блок-схему генетичного алгоритму, який дозволяє знайти оптимальний маршрут на основі шести критеріїв, використання яких дозволяє врахувати специфіку роботи безпроводних сенсорних мереж, а саме: Евклідову відстань, рівень заряду батареї вузла, рівень сигналу вузла, рівень втрат даних, затримку передачі даних, вхідний/вихідний ступінь вузла. Для перевірки ефективності розробленого генетичного алгоритму здійснено його порівняння із жадібним, використовуючи імітаційне моделювання. На основі отриманих результатів показано ефективність роботи запропонованого рішення.

Ключові слова: генетична еволюція, безпроводні сенсорні мережі, маршрутизація, багатокритеріальність.

PYRIH Yaroslav, PYRIH Yuliia

Lviv Polytechnic National University

## A MULTI-CRITERIA APPROACH BASED ON GENETIC EVOLUTION FOR FINDING THE OPTIMAL DATA TRANSMISSION ROUTE IN WIRELESS SENSOR NETWORKS

In wireless sensor networks (WSNs), data routing is critical to ensure their efficient operation. Traditional single-criteria routing algorithms do not take into account all aspects of the dynamically changing data transmission environment, which leads to a decrease in overall network performance. The choice between single-criteria and multi-criteria routing algorithms depends on the specific conditions and requirements of the BSM. Single-criteria algorithms are suitable for simpler and more stable environments where a single parameter is critical. Multi-criteria algorithms, although more difficult to implement, provide significant advantages in adaptability and performance optimization in complex and dynamic networks.

This paper is devoted to the application of modern evolutionary computing to solve complex and multi-criteria route finding problems in modern wireless sensor networks. Genetic algorithms provide high flexibility and adaptability, which allows taking into account various criteria, which is especially important for sensor nodes with limited resources. The paper presents a multicriteria method for routing data in wireless sensor networks based on the principles of genetic evolution, taking into account the dynamically changing topology and range of nodes. A block diagram of the genetic algorithm is presented, which allows finding the optimal route based on a set of criteria, the use of which allows taking into account the specifics of wireless sensor networks. The proposed method can be applied to the implementation of both single- and multi-path routing protocols. To verify the effectiveness of the developed genetic algorithm, it is compared with the greedy one using simulation modelling. Based on the obtained results, the efficiency of the proposed solution for sensor nodes with different ranges is shown. A set of alternative routes is obtained, which is advisable to use in case of failure of the optimal route, which increases the reliability of wireless sensor networks under dynamically changing conditions.

Keywords: genetic evolution, wireless sensor networks, routing, multicriteria.

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

У безпроводних сенсорних мережах (БСМ) маршрутизація даних є критично важливою для забезпечення їх ефективного функціонування. Традиційні однокритеріальні алгоритми маршрутизації не враховують усі аспекти динамічно змінного середовища передавання даних, що призводить до зниження загальної продуктивності мережі. Вибір між однокритеріальними та багатокритеріальними алгоритмами маршрутизації залежить від конкретних умов та вимог БСМ. Однокритеріальні алгоритми підходять для простіших та більш стабільних середовищ, де один параметр є критично важливим. Багатокритеріальні алгоритми, хоча і складніші в реалізації, надають значні переваги в адаптивності та оптимізації продуктивності в умовах складних та динамічних мереж.

З розвитком обчислювальних методів виникла необхідність у впровадженні нових підходів до маршрутизації, які могли б адаптуватися до змін у мережі та забезпечувати оптимальні рішення. Обмеженість ресурсів БСМ, зокрема заряду батареї, радіусу дії вузлів, значно впливає на процес



маршрутизації даних, що зумовлює потребу в розробці енергоефективних та надійних алгоритмів. Оптимізація маршрутизації з урахуванням цих обмежень є ключовим фактором для забезпечення ефективного та стабільного функціонування БСМ. У зв'язку з цим доцільним є застосування сучасних еволюційних обчислень, що дають можливість вирішувати складні та багатокритеріальні задачі. Генетичні алгоритми забезпечують високу гнучкість та адаптивність, що дозволяє враховувати різноманітні критерії, такі як енергоспоживання, затримку передачі даних, рівень сигналу, втрат даних, а також зберігати баланс між ними. Це особливо актуально для БСМ, де обмежені ресурси та високі вимоги до якості надання послуг потребують більш складних підходів до маршрутизації даних.

Отже, дослідження генетичних алгоритмів як основи функціонування багатокритеріального підходу для пошуку оптимального маршруту у БСМ є актуальним.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Пошук робіт, який відповідає тематиці даної роботи, здійснювався у бібліотеці IEEE за ключовими словами: multi-criteria routing in wireless sensor networks.

У роботі [1] представлено кооперативний багатокритеріальний енергоощадний алгоритм для маршрутизації у безпроводних ad-hoc мережах. В якості критеріїв використано залишкову енергію вузлів та потужність передавання даних. Алгоритм здійснює вибір маршруту для кожної пари джерело-приймач у вигляді послідовності груп вузлів, що співпрацюють між собою. На основі отриманих результатів при моделюванні 20 сенсорних вузлів із різною та сталою потужністю передавання даних показано, що запропонований алгоритм дозволяє зменшити енергоспоживання та збільшити кількість доставлених пакетів.

Автори [2] здійснили дослідження практичних особливостей багатокритеріального підходу до вирішення задачі оптимальної маршрутизації в мережах зв'язку з урахуванням сукупності показників якості: час затримки пакетів, рівень втрат пакетів, вартість використання лінії зв'язку. При використанні запропонованого рішення отримано виграв у три рази за показником втрат пакетів і у півтора рази за показником вартості використання ліній зв'язку порівняно зі скалярним підходом.

Робота [3] присвячена протоколу багатошляхової QoS-маршрутизації на основі стабільності маршруту. Авторами запропонована метрика маршрутизації для маршруту з переходів, яка поєднує функцію, що базується на стабільності витрат, з такими метриками QoS як пропускна здатність та затримка. Представлені результати моделювання запропонованого рішення продемонстрували його ефективність по критерію середньої наскрізної затримки, ймовірності втрати пакетів та пропускної здатності.

У [4] представлено схему маршрутизації в БСМ із використанням багатокритеріального режиму прийняття рішень, яка динамічно знаходить наступний вузол зі списку сусідів на основі залишкового заряду батареї, частоти передачі пакетів та кількості хопів його сусідів. При цьому параметри визначаються як критерії і кожному з них присвоюється певна вага. Результати роботи запропонованої схеми демонструють рівномірне використання вузлів з ефективним використанням заряду батареї вузлів.

Робота [5] описує поєднання показників центральності та багатокритеріальні алгоритми, щоб підвищити продуктивність БСМ шляхом впровадження підходу динамічної багатокритеріальної маршрутизації. Реалізовано основну концепцію, засновану на методах багатоатрибутного прийняття рішень (MADM) для вибору найкращих альтернативи (маршрути між кожним вузлом і приймачем, а також вузли маршрутизації) для створення дерева маршрутизації відповідно до вимог користувача та зміни стану мережі. Даний процес включає кілька параметрів, які впливають на споживання енергії: вартість зв'язку від сенсорного вузла до батьківського та залишкову енергію батьківського, що допомагає сенсорним вузлам досягти збалансованого розсіювання енергії в мережі. Використовуючи середовище Matlab, здійснено імітаційне моделювання, на основі якого показано, що розроблений алгоритм дозволяє зменшити енергоспоживання та збільшити термін служби мережі.

Автори [6] розробили енергоефективний алгоритм маршрутизації, який приймає рішення про наступний крок, використовуючи метод прийняття рішень за багатьма критеріями (MCDM). При цьому враховуються три фактори, які впливають на час життя мережі та передачу пакетів: відстань до вузла-отримувача та сусідніх вузлів, залишкова енергія та довжина черги на кожному вузлі. На основі результатів моделювання показано, що розроблений алгоритм суттєво збільшує час життя мережі та зменшує швидкість втрати пакетів у порівнянні із Energy Aware Routing (EAR) та Multipath Routing (MR).

Отже, багатокритеріальна маршрутизація даних в БСМ є складною, але необхідною задачею для забезпечення ефективної роботи мережі. При цьому на основі аналізу робіт [1-6] варто відзначити, що практично не приділялась увага дослідженням щодо використання принципів еволюційних стратегій, зокрема генетичної. Таким чином, доцільним є здійснення досліджень щодо багатокритеріальної оптимізації за допомогою генетичного алгоритму для підвищення ефективності маршрутизації даних у БСМ.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є: розробка багатокритеріального методу маршрутизації даних у безпроводних сенсорних мережах на основі принципів генетичної еволюції, який здійснює пошук оптимального маршруту із врахуванням сукупності критеріїв.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Для пошуку оптимального маршруту передачі даних у БСМ використано багатокритеріальність, що дозволяє враховувати усі важливі параметри, які впливають на ефективну і надійну роботу мережі. Для визначення критеріїв, які забезпечують комплексний та збалансований пошук маршруту, було детально проаналізовано особливості функціонування БСМ та роботи, присвячені маршрутизації [7-11]. На основі цього було вибрано такі критерії як Евклідова відстань, заряд батареї вузла, рівень сигналу вузла, рівень втрат даних, затримка передачі даних та вхідний/вихідний ступінь вузла, що у сукупності дозволяє формувати оптимальні рішення в різних умовах.

Після введення вхідних даних відбувається формування матриці для кожного критерію, який використовується для пошуку маршруту. Кожен елемент такої матриці представляє з'єднання між двома вузлами і включає в себе значення певного параметру (Евклідової відстані, рівня втрат даних тощо). Далі здійснюється нормалізація утворених матриць для їх приведення до діапазону [1; 100]. Для забезпечення справедливого внеску кожного критерію в процес оцінки придатності маршруту використано ваговий коефіцієнт. Його значення є однаковим для всіх розглянутих критеріїв, щоб запобігти домінуванню одного критерію над іншими та сприяти їх коректному порівнянню і комбінуванню .

$$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \alpha_6 = 1, \alpha_i = 1/n, \quad (1)$$

де  $n$  – кількість критеріїв.

Наступний етап – це утворення однієї багатокритеріальної матриці шляхом сумування нормалізованих матриць по 6-ти критеріям. Після цього переходимо до запуску розробленого ГА, основними етапами роботи якого є застосування таких генетичних операторів: турнірний оператор відбору, впорядкований оператор схрещування, оператор перемішування для мутації, що дозволяє знайти маршрут у заданих умовах. Для покращення збіжності ГА пропонується використовувати динамічне корегування значення імовірності схрещування  $p_{cross}$  на основі оцінки значення фітнес-функції  $F(x)_{set\ r}$  для множини маршрутів та  $i$ -ого маршруту  $F(x)_{r_i}$  у певному поколінні. Таким чином, здійснюється порівняння цих двох значень, на основі чого визначається рівень якості певної особи. Динамічну зміну  $p_{cross}$  можна представити як:

$$p_{cross} = \begin{cases} \frac{a + (b - a)(F(x)_{max\_set\ r} - F(x)_{r_i})}{F(x)_{max\_set\ r} - F(x)_{aver}}, \text{ при } F(x)_{r_i} > F(x)_{aver} \\ \frac{a + b}{2}, \text{ при } F(x)_{r_i} \approx F(x)_{aver} \\ b, \text{ при } F(x)_{r_i} < F(x)_{aver} \end{cases} \quad (2)$$

де  $a, b$  – числа, які позначають нижню та верхню межу заданої імовірності схрещування відповідно;  $F(x)_{max\_set\ r}$  – максимальне значення фітнес-функції маршрутів для покоління;  $F(x)_{aver}$  – середнє значення фітнес-функції, отримане для поточного покоління;  $F(x)_{r_i}$  – значення фітнес-функції, яке отримується на основі вибору кращого значення при розгляді двох значень вибраних особин для схрещування.

Якщо мережева топологія часто змінюється через мобільність вузлів або зміну умов середовища, то доцільно далі здійснити пошук множини альтернативних маршрутів, що міститиме ті маршрути, значення фітнес-функції яких є близькими до оптимального рішення. Для цього пропонується застосувати критерій кількість транзитів (переходів)  $h$ . Відповідно визначається його значення для оптимального маршруту, після чого задається діапазон значень як  $[h - 1; h + 3]$  із кроком 1. Здійснюється повторний запуск ГА, внаслідок чого формується множина альтернативних маршрутів, що дозволить мережі здійснювати швидку адаптацію при зміні умов.

На рис. 1 представлено алгоритм для багатокритеріального пошуку оптимального маршруту та множини альтернативних маршрутів .

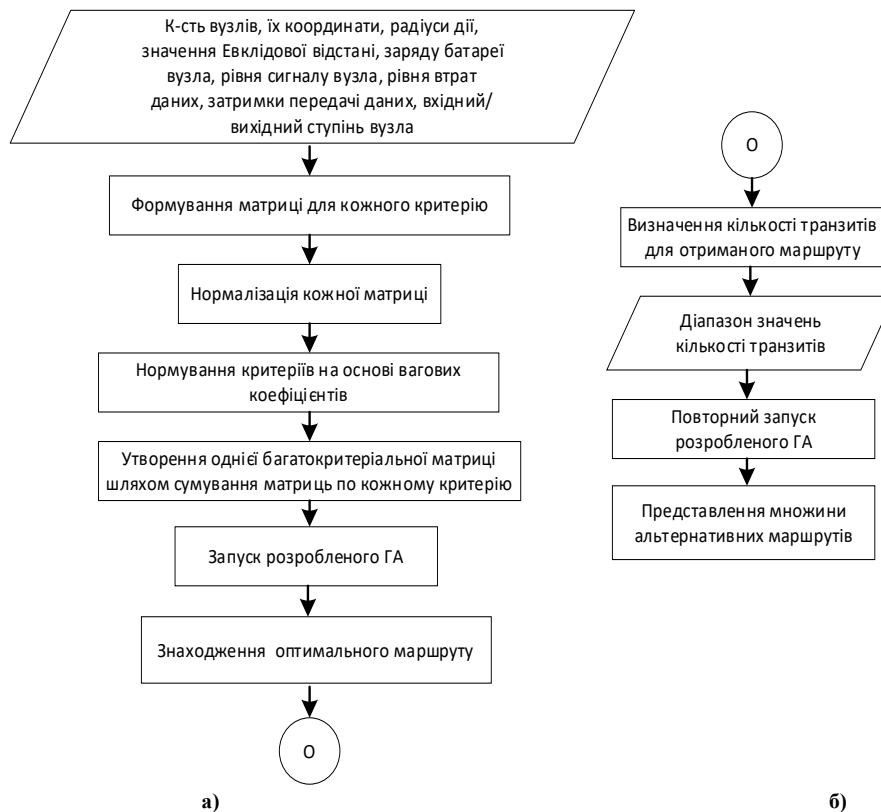


Рис. 1. Узагальнена блок-схема багатокритеріального пошуку: а) – оптимального маршруту; б) – множини резервних маршрутів по критерію кількості транзитів

Для перевірки ефективності розробленого алгоритму багатокритеріального пошуку оптимального маршруту для БСМ створено програму на мові програмування Python, основою якої є бібліотека DEAP. Для порівняння вибрано жадібний алгоритм (ЖА), який характеризується простою та ефективною обчислювальною реалізацією [12]. Крім цього цей алгоритм володіє здатністю швидко адаптуватися до змін у мережі, таких як вихід з ладу вузлів або зміна топології, що робить його доцільним для використання в динамічно змінних мережевих умовах.

Рис. 2 демонструє графову модель мережі БСМ для комплексного критерію, де круг позначає вузол мережі, у якому вказано його порядковий номер відповідно до топології, ребра – з'єднання між вузлами, на яких вказано їх вагу – нормалізоване значення. Для графової моделі використовуються вузли різних кольорів, які позначають різні радіуси їх дії: ● – 20 метрів, ● – 30 метрів, ● – 40 метрів.

На рис. 3 представлено відповідну нормалізовану матрицю.

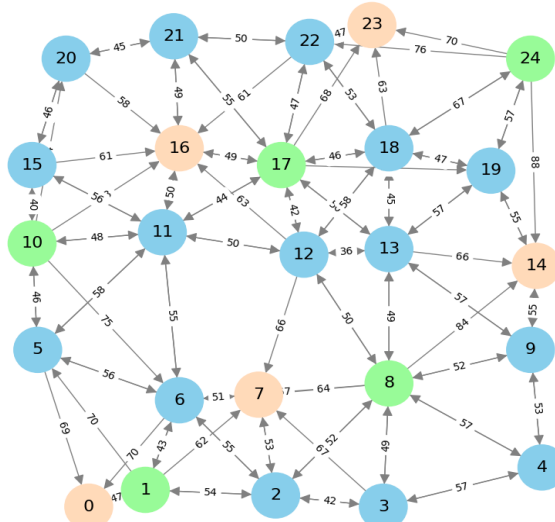


Рис. 2. Графова модель мережі БСМ для комплексного критерію

Normalized Matrix																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
1	0	50	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
2	47	0	54	1000	1000	70	43	62	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
3	1000	54	0	45	1000	1000	55	54	55	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
4	1000	1000	42	0	58	1000	1000	67	47	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
5	1000	1000	1000	57	0	1000	1000	1000	58	52	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
6	69	1000	1000	1000	1000	0	57	1000	1000	1000	47	55	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
7	70	43	55	1000	1000	56	0	52	1000	1000	1000	56	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
8	1000	1000	53	1000	1000	1000	51	0	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
9	1000	1000	52	49	57	1000	67	64	0	50	1000	1000	49	49	84	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
10	1000	1000	1000	1000	53	1000	1000	1000	52	0	1000	1000	1000	58	53	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	
11	1000	1000	1000	1000	1000	46	75	1000	1000	1000	0	55	1000	1000	1000	38	78	1000	1000	1000	74	1000	1000	1000	1000	
12	1000	1000	1000	1000	1000	58	55	1000	1000	1000	48	0	50	1000	1000	54	50	1000	1000	1000	1000	1000	1000	1000	1000	
13	1000	1000	1000	1000	1000	1000	1000	1000	66	50	1000	1000	50	0	39	1000	1000	63	40	54	1000	1000	1000	1000	1000	
14	1000	1000	1000	1000	1000	1000	1000	1000	1000	49	57	1000	1000	36	0	66	1000	1000	50	41	59	1000	1000	1000	1000	
15	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	55	1000	1000	1000	1000	0	1000	1000	1000	1000	55	1000	1000	1000	1000	
16	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	40	56	1000	1000	1000	0	61	1000	1000	1000	43	1000	1000	1000	1000	
17	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	50	1000	1000	1000	1000	0	46	1000	1000	1000	50	1000	1000	1000	
18	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	44	42	50	1000	1000	49	0	42	67	1000	54	50	68	1000	
19	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	58	45	1000	1000	1000	46	0	46	1000	1000	54	63	61	
20	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	57	55	1000	1000	1000	47	0	1000	1000	1000	58	
21	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	46	58	1000	1000	1000	0	43	1000	1000	1000	
22	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	49	55	1000	1000	1000	45	0	51	1000	1000	
23	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	61	47	53	1000	1000	50	0	50	1000	
24	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	47	0	1000	
25	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	88	1000	1000	1000	67	57	1000	1000	76	70	0

Рис. 3. Нормалізована матриця

На рис. 4 показано отриманий маршрут при використанні комплексного критерію у розробленому алгоритмі та у жадібному алгоритмі (ЖА). Варто зауважити, що для цього дослідження ЖА було модифіковано для роботи із сукупністю розглянутих критеріїв.

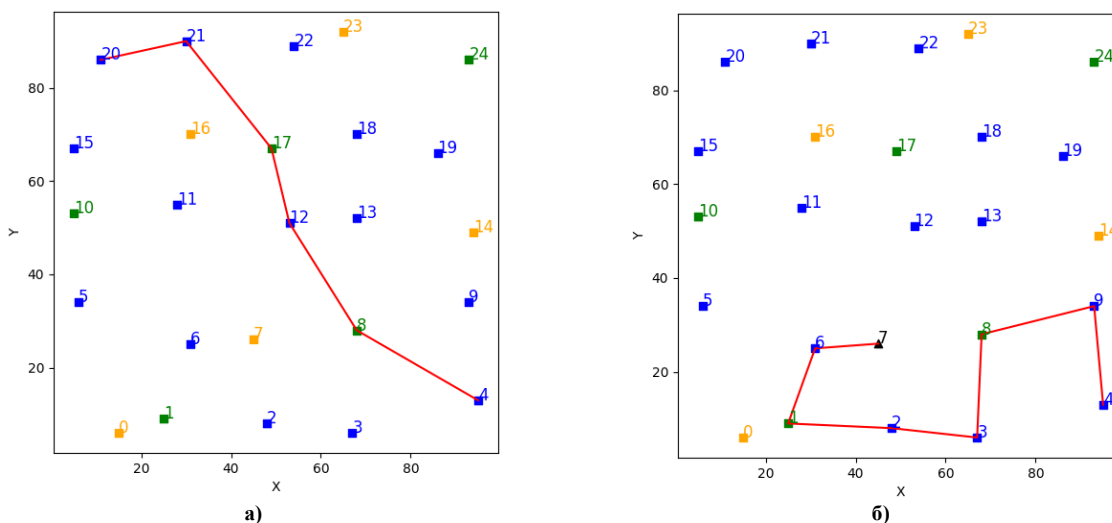


Рис. 4. Отриманий маршрут у БСМ при використанні: а) – розробленого алгоритму; б) – жадібного алгоритму

Як бачимо з рис. 4, використання ЖА не дозволило сформувати успішний маршрут між вузлами 4 і 20, у цьому випадку його робота завершена на вузлі 7.

На рис. 5 представлено результати моделювання при визначенні множини резервних маршрутів. Оскільки отриманий маршрут (рис. 4 (а)) містить 5 транзитів, то це значення було прийняте в якості номінального для формування діапазону значень кількості транзитів [4: 8].

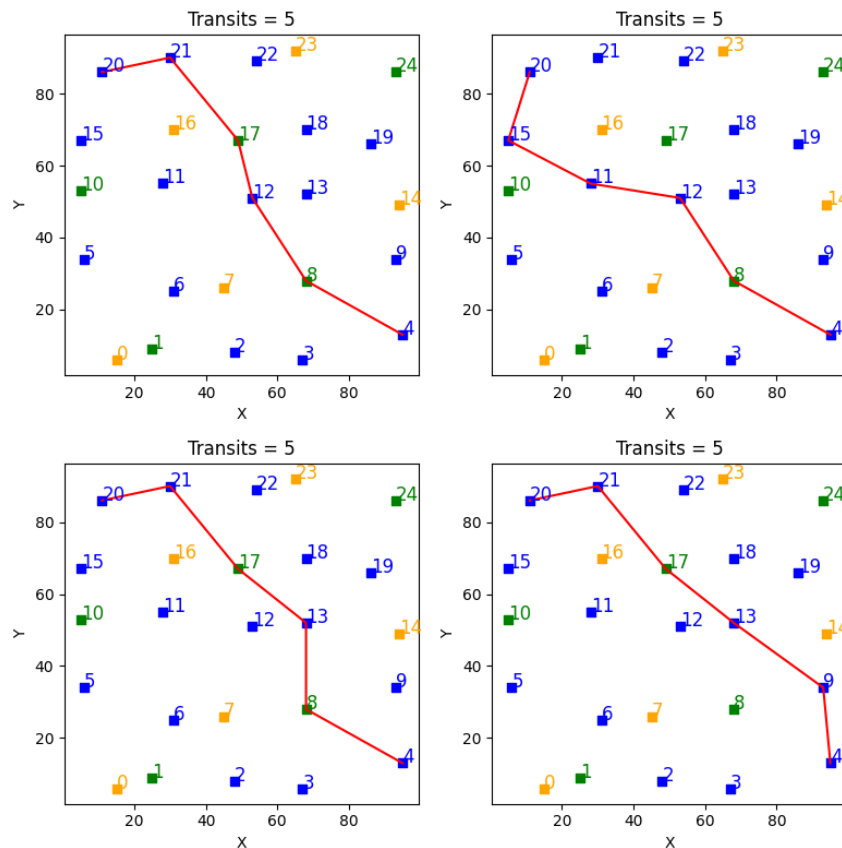


Рис. 5. Множина найкращих резервних маршрутів для  $h \in [4:8]$

Таким чином, наведені результати демонструють працездатність запропонованого рішення при використанні сукупності розглянутих критеріїв для пошуку як найкращого маршруту, так і множини резервних маршрутів у БСМ при різних радіусах дії вузлів.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМІ

Запропоновано багатокритеріальний метод маршрутизації даних у безпроводних сенсорних мережах на основі принципів генетичної еволюції, із врахуванням динамічно змінної топології та радіусів дії вузлів. Представлено блок-схему генетичного алгоритму, який дозволяє знайти оптимальний маршрут на основі шести критеріїв, використання яких дозволяє врахувати специфіку роботи БСМ, а саме: Евклідову відстань, рівень заряду батареї вузла, рівень сигналу вузла, рівень втрат даних, затримку передачі даних, вхідний/вихідний ступінь вузла. Запропонований метод маршрутизації може бути застосований для реалізації протоколів як одно-, так і багатопляхової маршрутизації. На основі отриманих результатів показано ефективність роботи запропонованого рішення при використанні комплексного критерію для пошуку маршруту передачі даних між сенсорними вузлами із різними радіусами дії. Перспективним напрямком подальших досліджень є інтеграція методів машинного навчання для динамічного налаштування параметрів генетичного алгоритму з метою покращення ефективності маршрутизації.

#### Література

1. Gravalos I., Kokkinos P., Varvarigos E. A. Multi-criteria Cooperative Energy-Aware Routing in Wireless Ad-hoc Networks // 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, Italy, 2013. – P. 387-393. – DOI: 10.1109/IWCMC.2013.6583590.
2. Guo Q., Semenets V., Bezruk V., Kaliuzhnyi M., Zheng Y. Methods for Multicriterial Selection of Optimal Routes in Communication Networks // Eastern-European Journal of Enterprise Technologies. – 2019. – P. 52–57. – DOI: 10.15587/1729-4061.2019.176783.
3. AlQahtani S., Alotaibi A. A Route Stability-Based Multipath QoS Routing Protocol in Cognitive Radio Ad Hoc Networks // Wireless Netw. – 2019. – Vol. 25. – P. 2931–2951. – DOI: 10.1007/s11276-019-02014-6.
4. Bhunia S. S., Roy S., Mukherjee N. Adaptive Learning Assisted Routing in Wireless Sensor Network Using Multi Criteria Decision Model // International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, 2014. – P. 2149-2154. – DOI: 10.1109/ICACCI.2014.6968354.

5. Fouad El H., Cherkaoui L., Khadija D. Adaptive Routing Protocol for Lifetime Maximization in Multi-Constraint Wireless Sensor Networks // Journal of Communications and Information Networks. – 2018. – Vol. 3(1). – P. 67-83. – DOI: 10.1007/s41650-018-0008-3.
6. Suh Y. H., Kim K. T., Shin D. R., Youn H. Y. Traffic-Aware Energy Efficient Routing (TEER) Using Multi-Criteria Decision Making for Wireless Sensor Network // 5th International Conference on IT Convergence and Security (ICITCS), Kuala Lumpur, Malaysia, 2015. – P. 1-5. – DOI: 10.1109/ICITCS.2015.7293029.
7. Пиріг Я., Климаш М., Пиріг Ю., Лаврів О. Генетичний алгоритм як засіб розв'язання оптимізаційних задач // Інфокомунікаційні технології та електронна інженерія. – 2023. – № 3(2). – С. 95-107.
8. Pavlenko O., Tymoshenko A., Tymoshenko O., Luntovskyy A. Searching Extreme Paths Based on Travelling Salesman's Problem for Wireless Emerging Networking // Emerging Networking in the Digital Transformation Age. Lecture Notes in Electrical Engineering, Springer, 2022. – Vol. 965. – P. 284-304. – DOI: 10.1007/978-3-031-24963-1\_16.
9. Колесніков К. В., Карапетян А. Р., Царенко Т. А. Генетичні алгоритми для задач багатокритеріальної оптимізації в мережах адаптивної маршрутизації даних // Вісник НТУ «ХПІ». Серія: Нові рішення в сучасних технологіях. – 2013. – № 56 (1029). – С. 44-50.
10. Pyrih Y., Masiuk A., Pyrih Yu., Urikova O. Investigation of a Genetic Algorithm for Solving the Travelling Salesman Problem // Digital Ecosystems: Interconnecting Advanced Networks with AI Applications. Lecture Notes in Electrical Engineering, 2024. – Vol. 1198. – P. 506-521. – DOI: 10.1007/978-3-031-61221-3\_24.
11. Rehena Z., Roy S., Mukherjee N. Efficient Data Forwarding Techniques in Wireless Sensor Networks // 3rd IEEE International Advance Computing Conference (IACC), Ghaziabad, India, 2013. – P. 449-457. – DOI: 10.1109/IAdCC.2013.6514268.
12. Klymash Y., Strykhalyuk B., Strykhalyuk I. Algorithm for Greedy Routing Based on the Thurston Algorithm in Sensor Networks // 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), Lviv, Ukraine, 2016. – P. 652-654. – DOI: 10.1109/TCSET.2016.7452142.

#### References

1. Gravalos I., Kokkinos P., Varvarigos E. A. Multi-criteria Cooperative Energy-Aware Routing in Wireless Ad-hoc Networks // 9th International Wireless Communications and Mobile Computing Conference (IWCMC), Sardinia, Italy, 2013. – P. 387-393. – DOI: 10.1109/IWCMC.2013.6583590.
2. Guo Q., Semenets V., Bezruk V., Kaliuzhnyi M., Zheng Y. Methods for Multicriterial Selection of Optimal Routes in Communication Networks // Eastern-European Journal of Enterprise Technologies. – 2019. – P. 52-57. – DOI: 10.15587/1729-4061.2019.176783.
3. AlQahtani S., Alotaibi A. A Route Stability-Based Multipath QoS Routing Protocol in Cognitive Radio Ad Hoc Networks // Wireless Netw. – 2019. – Vol. 25. – P. 2931-2951. – DOI: 10.1007/s11276-019-02014-6.
4. Bhunia S. S., Roy S., Mukherjee N. Adaptive Learning Assisted Routing in Wireless Sensor Network Using Multi Criteria Decision Model // International Conference on Advances in Computing, Communications and Informatics (ICACCI), Delhi, India, 2014. – P. 2149-2154. – DOI: 10.1109/ICACCI.2014.6968354.
5. Fouad El H., Cherkaoui L., Khadija D. Adaptive Routing Protocol for Lifetime Maximization in Multi-Constraint Wireless Sensor Networks // Journal of Communications and Information Networks. – 2018. – Vol. 3(1). – P. 67-83. – DOI: 10.1007/s41650-018-0008-3.
6. Suh Y. H., Kim K. T., Shin D. R., Youn H. Y. Traffic-Aware Energy Efficient Routing (TEER) Using Multi-Criteria Decision Making for Wireless Sensor Network // 5th International Conference on IT Convergence and Security (ICITCS), Kuala Lumpur, Malaysia, 2015. – P. 1-5. – DOI: 10.1109/ICITCS.2015.7293029.
7. Pyrih Ya., Klymash M., Pyrih Yu., Lavriv O. Henetychnyi alhorytm yak zasib rozv'iazannia optymizatsiinykh zadach // Infokomunikatsiini tekhnolohii ta elektronna inzheneriia. – 2023. – № 3(2). – С. 95-107.
8. Pavlenko O., Tymoshenko A., Tymoshenko O., Luntovskyy A. Searching Extreme Paths Based on Travelling Salesman's Problem for Wireless Emerging Networking // Emerging Networking in the Digital Transformation Age. Lecture Notes in Electrical Engineering, Springer, Cham, 2022. – Vol. 965. – P. 284-304. – DOI: 10.1007/978-3-031-24963-1\_16.
9. Kolesnikov K. V., Karapetian A. R., Tsarenko T. A. Henetychni alhorytmy dlia zadach bahatokryterialnoi optymizatsii v merezhakh adaptivnoi marshrutyzatsii danykh // Visnyk NTU «KhPI». Serii: Novi rishennia v suchasnykh tekhnolohiiakh. – 2013. – № 56 (1029). – С. 44-50.
10. Pyrih Y., Masiuk A., Pyrih Yu., Urikova O. Investigation of a Genetic Algorithm for Solving the Travelling Salesman Problem // Digital Ecosystems: Interconnecting Advanced Networks with AI Applications. Lecture Notes in Electrical Engineering, 2024. – Vol. 1198. – P. 506-521. – DOI: 10.1007/978-3-031-61221-3\_24.
11. Rehena Z., Roy S., Mukherjee N. Efficient Data Forwarding Techniques in Wireless Sensor Networks // 3rd IEEE International Advance Computing Conference (IACC), Ghaziabad, India, 2013. – P. 449-457. – DOI: 10.1109/IAdCC.2013.6514268.
12. Klymash Y., Strykhalyuk B., Strykhalyuk I. Algorithm for Greedy Routing Based on the Thurston Algorithm in Sensor Networks // 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), Lviv, Ukraine, 2016. – P. 652-654. – DOI: 10.1109/TCSET.2016.7452142.

<https://doi.org/10.31891/2219-9365-2024-79-13>

УДК 623.437-11:007.52]:621.396/.397]-047.58

ЛАКТИОНОВ Олександр

Національний університет «Полтавська політехніка імені Юрія Кондратюка»

<https://orcid.org/0000-0002-5230-524X>

e-mail: [itm.olaktionov@nupp.edu.ua](mailto:itm.olaktionov@nupp.edu.ua)

ПЕДЧЕНКО Назар

Національний університет «Полтавська політехніка імені Юрія Кондратюка»

<https://orcid.org/0000-0002-0018-4482>

e-mail: [jashsfenix@ukr.net](mailto:jashsfenix@ukr.net)

ЯНКО Аліна

Національний університет «Полтавська політехніка імені Юрія Кондратюка»

<https://orcid.org/0000-0003-2876-9316>

e-mail: [al9\\_yanko@ukr.net](mailto:al9_yanko@ukr.net)

БОРЯК Богдан

Національний університет «Полтавська політехніка імені Юрія Кондратюка»

<https://orcid.org/0000-0002-8114-7930>

e-mail: [boriakbr@nupp.edu.ua](mailto:boriakbr@nupp.edu.ua)

## МОДЕЛЮВАННЯ БАЗОВОЇ КОНСТРУКЦІЇ РОБОТОТЕХНІЧНОЇ ПЛАТФОРМИ

Стаття присвячена процесу розробки базової конструкції робототехнічної платформи із врахуванням відповідних кріплень для телекомунікаційних технологій. Проаналізовано існуючі типи конструкцій робототехнічних платформ, орієнтовані значення технічних характеристик й сформовано критерії обрання. Серед критеріїв виділено можливість кріплення спеціалізованого телекомунікаційного обладнання, швидкий доступ для обслуговування обладнання, простота у виготовленні, маневреність, прохідність. Наведено діапазони можливих значень технічних характеристик засобів радіозв'язку та сформовано відомості щодо їх кріплення на робототехнічній платформі. Створено схему базової конструкції робототехнічної платформи на основі двоколісної бази.

Ключові слова: робототехнічна платформа, конструкція, засоби зв'язку.

LAKTIONOV Oleksandr, PEDCHENKO Nazar, YANKO Alina, BORIAK Bohdan

National University «Yuri Kondratyuk Poltava Polytechnic»

## MODELING THE BASIC STRUCTURE OF A ROBOTIC PLATFORM

The hostilities in Ukraine give rise to the development of robotics, which is connected to a number of other industries, including telecommunications, electronics, programming, the Internet of Things, and mechanical engineering. Modern models of robotic platforms are multifunctional and are used for various tasks, the prerequisite for the formation of which are the technical characteristics of both the design and the means of communication.

Information about the robotic complex for extinguishing fires is highlighted in the information message. The advantages of the construction of the complex are the presence of tracks and specialized equipment for extinguishing fires. Specialized equipment is used for remote control of the complex, which provides a working distance depending on the type of terrain (open - a distance of up to two kilometers, crossed - 400 meters).

This article is dedicated to the process of developing the basic structure of a robotic platform, taking into account the necessary mounts for telecommunication technologies. The existing types of robotic platform designs have been analyzed, relevant technical specifications have been considered, and selection criteria have been established. Among the criteria highlighted are the ability to mount specialized telecommunication equipment, quick access for equipment maintenance, ease of manufacturing, maneuverability, and off-road capability. The article provides a range of possible technical specifications for radio communication equipment and offers information on how to mount them on the robotic platform. A schematic of the basic structure of a robotic platform based on a two-wheeled base has been created.

Keywords: robotic platform, design, communication equipment

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Бойові дії в Україні породжують розвиток робототехніки, котра пов'язана з низкою інших галузей, зокрема телекомунікаціями, електронікою, програмуванням, інтернетом речей, машинобудуванням. Сучасні взірці робототехнічних платформ є багатофункціональними й використовуються для різних завдань, передумовою формування котрих є технічні характеристики як конструкції, так і засобів зв'язку.

У інформаційному повідомленні [1] висвітлено інформацію про робототехнічний комплекс для гасіння пожеж. Перевагами конструкції комплексу є наявність гусениць та спеціалізованого обладнання для гасіння пожеж. Для дистанційного керування комплексом задіяне спеціалізоване обладнання, що забезпечує відстань роботи залежно від типу місцевості (відкрита – відстань до двох кілометрів, пересічена – 400 метрів).

Ідеї використання аналогічних засобів зв'язку спостерігаються у [2], де чотири колісна конструкція наземного робота має спеціалізовані кріплення не лише для органів керування, а й для вантажу.

Інформаційне повідомлення [3] містить інформацію про унікальні конструкції робототехнічних платформ, які оснащені турелями для зброї та обладнанням для розмінування. Крім засобів зв'язку вказані робототехнічні платформи повинні мати потужні електробатареї, щоб забезпечити не лише передачу сигналів керування, а й перевезення корисних вантажів [4].

Наведені конструкції робототехнічних платформ та технічні характеристики засобів зв'язку зорієнтовані для вирішення конкретних завдань, зокрема гасіння пожеж, перевезення вантажів, зброї тощо. З одного боку, конструкції є багатофункціональними, з іншого – мають чотириколісну колісну базу або гусениці. Тому цей факт актуалізує дослідження щодо розробки нових взірців конструкцій робототехнічних платформ, що передбачають наявність спеціалізованого телекомунікаційного обладнання.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Робототехнічні платформи умовно диференціюють на різновиди, зокрема роботи-сапери, спостережно-вогнєві роботи, роботи пожежники, роботи-розвідники, роботи для транспортування, роботи для запуску рою дронів тощо.

Інформаційне повідомлення [5] містить інформацію про легкий наземний робот-сапер Talon, який використовується для дистанційного розмінування. Робот має гусеничну стрічку, де швидкість переміщення робота 8,37 км/год. Унікальністю робота є багатоцільове використання. Про це свідчать різні додаткові механізми, зокрема маніпулятори, стрілецька зброя.

Український спостережно-вогнєвий робот «Мисливець» [6] довжиною 2 метри, шириною 1,25 метри, висотою 0,7 метри й вагою 120 кілограм оснащений привідною системою з шістьма моторами-редукторами по 800 Ват. Ці характеристики дозволяють роботу розвивати швидкість 10 км/год й автономно працювати до 36 год. Цей взірець також має кілька камер для керування.

Робот Recon Scout [7] довжиною 209 мм, шириною 193 мм, висотою 113 мм, вагою 0,5 кілограми на двох колесах, що розвиває швидкість 1,7 км/год. Робот оснащений різними сенсорами, зокрема інфрачервоним сенсором з кутом огляду 60 градусів. Керування, як й у напрацюваннях [5-6] реалізовано дистанційно, але на віддалі 91 й 30 метрів відповідно на різних типах місцевості.

Робот для транспортування THeMIS [8] довжиною 2,47 метри, шириною 2 метри, висотою 1,17 метри, кліренсом 0,6 метри, вагою 1630 кілограм, швидкістю до 20 км/год. Унікальністю цього робота є використання як дизельного двигуна, так і електричного. Принцип керування дистанційний на віддалі 1,5 кілометри. Робот має лідар, інфрачервоний сенсор.

Комплекс українського виробництва «Саламандра» зорієнтований для знешкодження протипіхотних мін й використовується у якості робота-камікадзе [9]. Має чотири колеса, а сама конструкція передбачає можливість прикріплення додаткових візків.

На відміну від напрацювання [9] наземна роботизована платформа TRX [10] призначена для запуску рою дронів типу Switchblade. Унікальністю платформи є наявність камер, пристроїв РЕБ, кулеметів.

Існуючі відомості щодо робототехнічних платформ містять лише основні технічні характеристики. Не зрозумілим залишається елементарна база стосовно зв'язку. Роботизована платформа наземного базування складається із рами, силової частини, ходової частини, системи радіозв'язку, системи відеозв'язку, системи обробки інформації.

Система радіозв'язку складається з передавача та приймача. Існуючі радіоприймачі RadioMaster ER8 [11], BetaFPV ELRS Micro [12], FrSky R9 Slim [13] характеризуються наступними аспектами. Робоча частота 2,4 ГГц, 2,4 ГГц, 915 МГц відповідно. Протоколи ExpressLRS, ExpressLRS, ACCESS Protocol при кількості каналів 8, 5, 6 відповідно.

Існуючі взірці передатчиків RadioMaster TX12 MKII ELRS [14], RadioMaster TX16 Mark II [15], Commando 8 ELRS [16] працюють на частотах 2,400 ГГц – 2,480 ГГц, 2,400 ГГц – 2,480 ГГц, 868/915 ГГц відповідно. Протоколи всіх передатчиків ExpressLRS забезпечують роботу по 16, 8, каналах на відстані 2 (без зовнішнього модуля), 3 кілометри відповідно.

Розглянуті існуючі взірці робототехнічних комплексів та систем радіозв'язку повинні являти собою єдину взаємопов'язану систему. Наразі не існує єдиного рішення щодо використання конкретних взірців конструкцій робототехнічних платформ у поєднанні з додатковими інструментами, зокрема системами забезпечення зв'язку тощо.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

**Мета роботи:** запропонувати конструкцію робототехнічної платформи на основі двоколісної бази з інтеграцією телекомунікаційних технологій для керування.

Завдання:

1. Провести порівняльний аналіз існуючих робототехнічних платформ й сформулювати критерії обрання оптимальної конструкції.

2. Здійснити огляд телекомунікаційного обладнання для керування робототехнічною платформою й створити орієнтовні технічні характеристики.



3. Запропонувати схему базової конструкції робототехнічної платформи.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Узагальнюючи існуючий досвід стосовно робототехнічних комплексів [5–10], наявних у вільному доступі, сформуємо діапазони основних технічних характеристик, котрими повинна володіти сучасна робототехнічна платформа.

Таблиця 1

#### Орієнтовні значення технічних характеристик робототехнічного комплексу [5–10]

Параметр	Значення
Особливості ходової частини	Колісна, гусенична стрічка
Діапазон дистанційного керування	Від 30 метрів до 1,5 кілометрів
Швидкість переміщення, км/год	1,7 – 20
Типи двигунів	Електричні, двигуни внутрішнього згорання
Додаткове обладнання	Кулемети, камери, лідари, РЕБ, інфрачервоні сенсори
Максимальне значення автономної роботи, годин	36
Орієнтована потужність двигуна	800 Вт і більше
Вага, кілограм	Від 120 до 1630

Зафіксовані у табл. 1 значення технічних характеристик є орієнтованою базою для створення конструкції робототехнічної платформи. Основний каркас конструкції зорієнтований на фіксацію органів керування та місця для перевезення вантажів. Сформуємо критерії обрання конструкції робототехнічної платформи.

Критерій № 1. Конструкція повинна передбачати кріплення спеціалізованого обладнання, зокрема відеокамер, засобів радіозв'язку, спеціалізованого обладнання для виявлення аномалій у ґрунті.

Критерій № 2. Конструкція повинна передбачати можливість використання кількох акумуляторних батарей з легким доступом для їх обслуговування.

Критерій № 3. Конструкція повинна бути простою у виготовленні й забезпечувати маневреність, прохідність, простоту виготовлення і обслуговування.

Крім критеріїв обрання конструкції важливим аспектом є розуміння технічних характеристик телекомунікаційного обладнання, що фіксуватиметься на робототехнічній платформі. Вказане обладнання робототехнічної платформи умовно можна диференціювати на групи: обладнання для керування силовою частиною робототехнічною платформою, обладнання для керування засобами додаткового зв'язку, зокрема відеокамери, турелі, інші сенсори, табл. 2.

Таблиця 2

#### Орієнтовні значення технічних характеристик радіоприймачів [11–13]

Параметр	Значення
Робоча частота	915МГц–2,4 ГГц
Протокол	ExpressLRS, ACCESS Protocol
Кількість каналів	5–8
Потужність передачі сигналу	20дБм
Вхідна напруга	3,5 В – 35 В

Як видно з табл. 2 технічні характеристики радіоприймачів впливають на кріплення відповідних кронштейнів на платформі. У табл. 3 наведено зведені відомості про передатчики.

Таблиця 3

#### Орієнтовні значення технічних характеристик передатчиків [14–16]

Параметр	Значення
Робоча частота	868/915 ГГц – 2,480 ГГц
Протокол	ExpressLRS
Кількість каналів	8, 16
Відстань передачі сигналу	2–3 кілометри

Крім робочих діапазонів частот, протоколів тощо передатчики характеризуються відстанню передачі сигналу. Для цього у конструкції також варто передбачити відповідний кронштейн кріплення та врахувати можливість використання антен.

Орієнтуючись на аналог двоколісного робота Recon Scout [17], бачимо, конструктивно відстань між антенами є максимальною. Це не заважає функціонуванню обладнання, рис. 1.

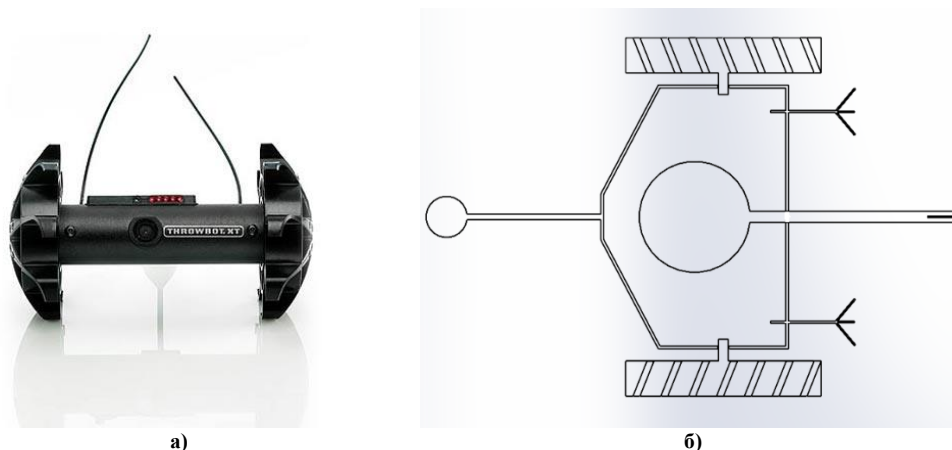


Рис. 1. Зовнішній вигляд робототехнічної платформи, (а) існуюча конструкція робототехнічної платформи, (б) запропонована конструкція робототехнічної платформи

Запропонована конструкція робототехнічної платформи матиме змогу бути оснащеною туреллю. Цей факт дає розуміння, конструктивне розміщення антен, як на Recon Scout, на етапі моделювання конструкції відповідає очікуванням. Монтаж антен на максимально допустимій відстані на корпусі робототехнічної платформи є гарним рішенням, де у майбутньому це варто дослідити відповідним інструментарієм.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

1. Завдання порівняльного аналізу існуючих робототехнічних платформ вирішується шляхом співставлення технічних характеристик існуючих платформ й створенням на основі цього критеріїв обрання платформи.
2. Завдання огляду телекомунікаційного обладнання для керування робототехнічними платформами вирішується шляхом формування орієнтовних діапазонів технічних характеристик телекомунікаційного обладнання.
3. Завдання створення схеми базової конструкції робототехнічної платформи вирішується шляхом використання двоклісної бази, що забезпечує простоту виготовлення, стійкість, маневреність, прохідність порівняно з існуючими типами конструкцій.

### Література

1. Перший в Україні пожежний наземний робот допущений до експлуатації у військах. *Міністерство оборони України*. URL: <https://www.mil.gov.ua/news/2024/08/15/pershij-v-ukraini-pozhezhnij-nazemnij-robot-dopushheni>.
2. Український наземний робот-камікадзе нищить логістику росіян на Донеччині. *АРМІЯ INFORM*. URL: <https://armyinform.com.ua/2024/04/02/ukrayinskyj-nazemnyj-robot-kamikadze-nyshhyt-logistyku-rosiyan-na-donechchyni/>
3. «Біжимо з РФ навпередки». Роботи на службі в ЗСУ: на що вони здатні і як їх вже використовують у війську – інтерв'ю з Ябчанкою. *Новини України та Світу. Головні та останні новини - НВ*. URL: <https://nv.ua/ukr/ukraine/events/nazemni-robotizovani-sistemi-yak-voni-dopomagayut-zsu-i-chi-ye-u-voroga-yabchanka-novini-ukrajini-50422466.html>
4. Для військових розробили роботизований комплекс "Мураха" (ФОТО, ВІДЕО). *Texty.org.ua - статті та журналістика даних для людей – Тексти.org.ua*. URL: <https://texty.org.ua/fragments/111564/dlya-vijskovykh-rozrobyly-robotyzovanyj-kompleks-muraha-foto-video/>.
5. TALON Tracked Military Robot - Army Technology. *Army Technology*. URL: <https://www.army-technology.com/projects/talon-tracked-military-robot/?cf-view>
6. «Мисливець» – наземний дрон, який понюхав пороху. *АрміяInform – Інформаційне агентство АрміяInform*. URL: <https://armyinform.com.ua/2021/05/27/myslyvecz-nazemnyj-dron-yakyj-ponyuhav-porohu/>
7. Recon Scout XT Reconnaissance Robot. URL: <https://www.homelandsecurity-technology.com/projects/recon-scout-xt-reconnaissance-robot/>
8. THeMIS - Milrem. *Milrem*. URL: <https://milremrobotics.com/defence/>
9. В Україні розробили новий НРК для знешкодження протипіхотних мін. | *ОПК. Новини та аналітика*. URL: <https://opk.com.ua/в-україні-розробили-новий-нрк-для-знеш/>

10. New Tracked Unmanned Vehicle Bristles With A Whopping 50 Switchblade Suicide Drones. The WARZONE. URL: <https://www.twz.com/42755/new-tracked-unmanned-vehicle-bristles-with-a-whopping-50-switchblade-suicide-drones>
11. RADiOMASTER ER8 ELRS Receiver User Manual. URL: <https://device.report/manual/8480522>
12. Приймач BetaFPV ELRS Micro Micro PWM 2.4G. FLYMOD. URL: [https://flymod.net/item/betafpv\\_elrs\\_receiver\\_pwm](https://flymod.net/item/betafpv_elrs_receiver_pwm)
13. FrSky R9 Slim+ OTA. URL: <https://rcstudio.cz/en/recievers-r9/762-frsky-r9-slim-ota.html>
14. Пульт дистанційного керування RadioMaster TX12 MKII ELRS - FPV радіоапаратура. *Phoenix drones – інтернет-магазин квадрокоптерів*. URL: <https://fx-drones.com/radioaparatura-radiomaster-tx12-mkii-elrs-pult-dystantsiinoho-keruvannia/?srsltid=AfmBOoqndNQxCn2lw6RCNcNcYaULid20Pc0IGJihUAdD1SKvJkd6e7xp>
15. Radiomaster TX16S. User Manual. URL: [https://hobbymania.com.ua/file/RadioMaster\\_TX16S.pdf](https://hobbymania.com.ua/file/RadioMaster_TX16S.pdf)
16. Апаратура керування iFlight Commando 8 ELRS 2.4G. PLANETA HOBBY. URL: <https://modelistam.com.ua/apparatura-upravleniya-iflight-commando-p-47128/>
17. Застарілі продукти. Recon Robotics. URL: <https://reconrobotics.com/products/legacy-products/>

### References

1. Pershyy v Ukraini pozhezhnyy nazemnyy robot dopushchenyy do ekspluatatsiyi u viys'kakh. Ministerstvo oborony Ukrainy. URL: <https://www.mil.gov.ua/news/2024/08/15/pershij-v-ukraini-pozhezhnij-nazemnij-robot-dopushheni>
2. Ukrainy's'kyu nazemnyy robot-kamikadze nyshchyt' lohistyku rosiyan na Donechchyni. ARMIYA INFORM. URL: <https://armyinform.com.ua/2024/04/02/ukrayinskyj-nazemnyj-robot-kamikadze-nyshchyt-logistyku-rosiyan-na-donechchyni/>
3. «Bizhymo z RF navvuperedky». Roboty na sluzhbi v ZSU: na shcho vony zdani i yak yikh v zhe vykorystovuyut' u viys'ku – interv'yu z Yabchankoyu. Novyny Ukrainy ta Svit. Holovni ta ostanni novyny - NV. URL: <https://nv.ua/ukr/ukraine/events/nazemni-robotizovani-sistemi-yak-voni-dopomagayut-zsu-i-chi-ye-u-voroga-yabchanka-novini-ukrajini-50422466.html>
4. Dlya viys'kovykh rozrobly robotyzovanyy kompleks "Murakha" (FOTO, VIDEO). Texty.org.ua - statti ta zhurnalistyka danykh dlya lyudey – Teksty.org.ua. URL: <https://texty.org.ua/fragments/111564/dlya-vijskovykh-rozrobly-robotyzovanyj-kompleks-muraha-foto-video/>
5. TALON Tracked Military Robot - Army Technology. Army Technology. URL: <https://www.army-technology.com/projects/talon-tracked-military-robot/?cf-view>
6. «Myslyvets'» – nazemnyy dron, yakyy ponyukhav porokhu. ArmiyaInform – Informatsiyne ahent'stvo ArmiyaInform. URL: <https://armyinform.com.ua/2021/05/27/myslyvecz-nazemnyj-dron-yakyy-ponyuhav-porohu/>
7. Recon Scout XT Reconnaissance Robot. URL: <https://www.homelandsecurity-technology.com/projects/recon-scout-xt-reconnaissance-robot/>
8. TheMIS - Milrem. Milrem. URL: <https://milremrobotics.com/defence/>
9. V Ukraini rozrobly novyy NRK dlya zneshkodzhennya protypikhotnykh min. | OPK. Novyny ta analytika. URL: <https://opk.com.ua/v-ukraini-rozrobly-novyj-nrk-dlya-znesh/>
10. New Tracked Unmanned Vehicle Bristles With A Whopping 50 Switchblade Suicide Drones. The WARZONE. URL: <https://www.twz.com/42755/new-tracked-unmanned-vehicle-bristles-with-a-whopping-50-switchblade-suicide-drones>
11. RADiOMASTER ER8 ELRS Receiver User Manual. URL: <https://device.report/manual/8480522>
12. Prymach BetaFPV ELRS Micro Micro PWM 2.4G. FLYMOD. URL: [https://flymod.net/item/betafpv\\_elrs\\_receiver\\_pwm](https://flymod.net/item/betafpv_elrs_receiver_pwm)
13. FrSky R9 Slim+ OTA. URL: <https://rcstudio.cz/en/recievers-r9/762-frsky-r9-slim-ota.html>
14. Pul't dystantsiynoho keruvannya RadioMaster TX12 MKII ELRS - FPV radioaaparatura. *Phoenix drones – internet-mahazyn kvadrokopteriv*. URL: <https://fx-drones.com/radioaparatura-radiomaster-tx12-mkii-elrs-pult-dystantsiinoho-keruvannia/?srsltid=AfmBOoqndNQxCn2lw6RCNcNcYaULid20Pc0IGJihUAdD1SKvJkd6e7xp>
15. Radiomaster TX16S. User Manual. URL: [https://hobbymania.com.ua/file/RadioMaster\\_TX16S.pdf](https://hobbymania.com.ua/file/RadioMaster_TX16S.pdf)
16. Апаратура керування iFlight Commando 8 ELRS 2.4G. PLANETA HOBBY. URL: <https://modelistam.com.ua/apparatura-upravleniya-iflight-commando-p-47128/>
17. Zastarili produkty. Recon Robotics. URL: <https://reconrobotics.com/products/legacy-products/>

<https://doi.org/10.31891/2219-9365-2024-79-14>

УДК 621.383.51

БОЖКО Костянтин

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»  
<https://orcid.org/0000-0002-6347-7442>

МУШКЕТ Костянтин

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»  
<https://orcid.org/0009-0005-9218-5835>

## АКТИВНА АНТЕНА У ВИМІРЮВАННЯХ ЕЛЕКТРИЧНИХ ПАРАМЕТРІВ СОНЯЧНИХ ЕЛЕМЕНТІВ

Метою роботи є обґрунтування використання активної антени в методі безконтактного вимірювання електричних параметрів окремих сонячних елементів у складі фотоелектричної сонячної панелі. Основними задачами даного наукового напрямку є розробка, моделювання та експериментальна перевірка активної антени у складі вимірювального тракту системи безконтактного вимірювання шунтового опору і електричної ємності окремих сонячних елементів у складі фотоелектричної сонячної панелі. Алгоритм вимірювання полягає у пропусканні через зовнішні контакти сонячної панелі синусоїдального сигналу від функціонального генератора. Над окремими сонячними елементами при цьому по черзі встановлюють датчик у формі плоскої антени, на якій наводиться синусоїдальний сигнал, що залежить від сумарного шунтового опору попередньо розташованих сонячних елементів. За рівномірністю спадання амплітуди сигналу від антени по мірі наближення її до кінця лінійки сонячних елементів роблять висновок про відмінність шунтового опору елементів відносно один до одного. Розрахунок шунтового опору окремих елементів виконують за рівняннями з теорії електричних кіл. До початку вимірювань окремо визначають шунтовий опір сонячної батареї в цілому. Активна антена із двома окремими секціями дозволяє отримувати підсилений сигнал, що є різницею сигналів від двох сусідніх сонячних елементів. Такий варіант методу прискорює цикл безконтактних вимірювань та підвищує їх точність. Моделювання схеми вимірювання із активною антеною було попередньо проведено в пакеті Multisim11, що прискорило розробку і допомогло визначити оптимальні параметри елементної бази, а також параметри і характеристики вимірювального сигналу.

Ключові слова: активна антена, сонячна панель, сонячний елемент, шунтовий опір, вимірювальний сигнал.

BOGKO Konstantin, MUSHKET Kostiantyn

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

## ACTIVE ANTENNA IN MEASUREMENTS OF ELECTRICAL PARAMETERS OF SOLAR ELEMENTS

The purpose of the work is to substantiate the use of an active antenna in the method of non-contact measurement of electrical parameters of individual solar cells as part of a photovoltaic solar panel. The main tasks of this scientific direction are the development, modeling and experimental verification of an active antenna as part of the measuring path of the system for non-contact measurement of shunt resistance and electric capacity of individual solar cells as part of a photovoltaic solar panel. The measurement algorithm consists in passing a sinusoidal signal from a functional generator through the external contacts of the solar panel. At the same time, a sensor in the form of a flat antenna is installed one by one above the individual solar cells, on which a sinusoidal signal is applied, which depends on the total shunt resistance of previously located solar cells. Based on the uniformity of the decrease in the signal amplitude from the antenna as it approaches the end of the line of solar cells, a conclusion is made about the difference in the shunt resistance of the cells relative to each other. The calculation of the shunt resistance of individual elements is performed according to the equations from the theory of electric circuits. Before starting the measurements, the shunt resistance of the solar battery as a whole is determined separately. An active antenna with two separate sections allows you to receive an amplified signal, which is the difference between the signals from two adjacent solar cells. This variant of the method speeds up the cycle of non-contact measurements and increases their accuracy. The simulation of the measurement scheme with an active antenna was previously carried out in the Multisim11 package, which accelerated the development and helped to determine the optimal parameters of the element base, as well as the parameters and characteristics of the measurement signal.

Keywords: active antenna, solar panel, solar cell, shunt resistance, measurement signal.

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ

#### ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Безконтактне вимірювання шунтового опору та інших електричних параметрів окремих сонячних елементів, які послідовно об'єднані між собою і мають нерозривну конструкцію у вигляді панелі, на наш погляд є актуальним у таких випадках:

- при необхідності додаткового контролю сонячної панелі перед встановленням її у важкодоступних місцях, зокрема на дахах і фасадах висотних будівель, адже тільки безконтактний метод дозволяє впевнитись у ідентичності параметрів усіх сонячних елементів і запобігти встановлення панелі, в якій може швидко початись процес деградації;
- при прогнозуванні процесу деградації сонячної панелі в умовах експлуатації, ознакою чого є поступове зменшення шунтового опору її елементів.

Безконтактний метод вимірювання шунтового опору на основі антенних датчиків був реалізований на основі осцилографічних вимірювань параметрів імпульсних сигналів [1]. Проте, цифровий осцилограф має велику похибку вимірювання напруги, яка складає до 1,5%. При типовій кількості елементів 24-вольтової сонячної панелі, яка дорівнює 60, різниця амплітуди сигналів від двох сусідніх елементів на антенних датчиках має середню величину, яка наближена до величини похибки осцилографа. За таких умов за результатами вимірювань можливо відокремити лише ті елементи, шунтовий опір яких значно (на 50% і більше) відрізняється від середнього значення. Для зменшення впливу похибки цифрового осцилографа запропоновано безпосередньо вимірювати різницевий сигнал від двох однакових антенних секцій. При цьому антена є активною, оскільки має схему підсилення диференційного сигналу.

Вимірювання електричних параметрів сонячних панелей та сонячних елементів є наразі актуальною задачею. Метою досліджень в цій сфері зокрема є створення так званих інтелектуальних сонячних модулів, тобто панелей, які мають індивідуальну систему вимірювання та діагностики, яка має використовуватись під час повного терміну їх експлуатації. Прикладом досліджень в цьому напрямку є робота українських фахівців [2], в якій теоретично і експериментально доведено, що основою діагностики сонячної панелі під час її експлуатації є вимірювання її вольт-амперної характеристики та розрахунку на її основі електричних параметрів. Сучасному стану досліджень (до 2022 року включно) з вимірювань вольт-амперних характеристик сонячних панелей при їх експлуатації присвячено огляд [3].

Вольт-амперні характеристики окремих сонячних елементів отримують на основі аналізу їх фотолюмінісцентних [4] та електролюмінісцентних [5] зображень і цей метод є наразі основним у безконтактних вимірюваннях. Кореляції між зміною різних електричних параметрів сонячних елементів присвячено роботу [6].

Інший відомий метод безконтактних вимірювань електричних параметрів сонячних елементів – метод електростатичного вольтметра [7,8]. Дослідження останніх років, які присвячено розвитку саме цього методу, наведено в роботі [9].

Для обґрунтування методу безконтактного вимірювання шунтового опору сонячних елементів розглянемо еквівалентну схему сонячного елемента (Рис. 1).

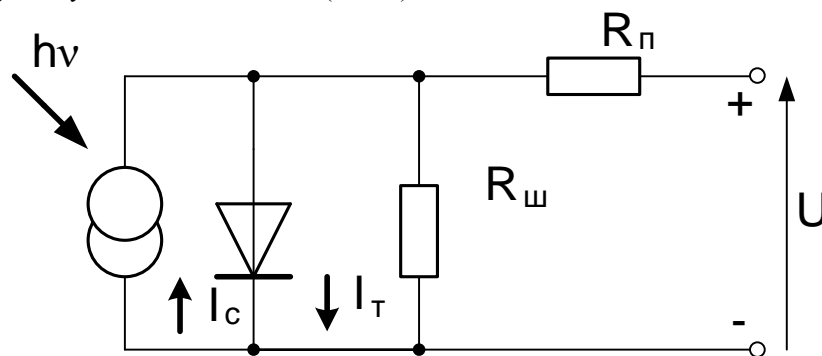


Рис. 1. Еквівалентна схема сонячного елемента:  $I_c$  – світловий струм;  $I_t$  – темновий струм;  $R_{ш}$  – шунтовий опір;  $R_p$  – послідовний опір;  $U$  – вихідна напруга сонячного елемента

Сонячна панель є конструкцією із послідовного з'єднання певної кількості сонячних елементів. Наприклад, для кремнієвої сонячної панелі з номінальною вихідною напругою 12 В ланка елементів налічує 30 або 36 штук. Із схеми (Рис. 1) впливає важлива роль шунтового опору. У ідеального сонячного елемента опір шунта є нескінченно великим. При зменшенні величини опору шунта через нього починає йти темновий струм  $I_t$  і виділяється тепло Джоуля. Внаслідок цього процесу зменшується коефіцієнт корисної дії елементів і панелі в цілому, що є ознакою процесу деградації. Фізичною причиною деградації сонячного елемента є ушкодження структури його поверхневого шару внаслідок виникнення мікро-пробовів під час експлуатації.

Слід зазначити, що еквівалентна схема (Рис.1) не враховує електричну ємність сонячного елемента, яка впливає на його динамічні характеристики, тобто на роботу в режимі змінного сигналу. Якщо на виході сонячного елемента подати змінний вимірювальний сигнал прямокутної імпульсної або синусоїдальної форми, то цей сигнал не пройде через діод, якщо його амплітуда буде меншою за напругу падіння на відкритому діоді. Для кремнію ця напруга приблизно дорівнює 600 мВ. Якщо вимірювальний сигнал подавати на сонячну панель, то умовою для обмеження сигналу є величина напруги у  $N$  раз більша, де  $N$  – кількість сонячних елементів у панелі. Висновок: сигнал від функціонального генератора із діапазоном амплітуди 0-15 В можна використовувати без додаткових обмежень при виконанні безконтактного вимірювання. Такий режим для сонячної панелі є режимом малого сигналу. Еквівалентна схема для режиму малого сигналу наведена на Рис. 2.

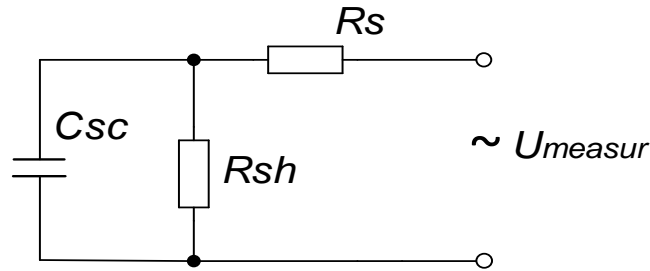


Рис. 2. Еквівалентна схема сонячного елемента в режимі малого сигналу та без освітлення:  $C_{sc}$  – ємність сонячного елемента;  $R_s$  – послідовний (serial) опір;  $R_{sh}$  – шунтовий опір;  $U_{measur}$  – напруга вимірювального сигналу від генератора

При подачі вимірювального сигналу на виводи сонячної панелі, сигнал послідовно проходить через усі елементи. По мірі проходження амплітуда сигналу буде ступінчасто спадати від максимуму на виході генератора до нуля. Якщо над сонячними елементами панелі розташувати планарні антени прямокутної форми, то на них буде наводитись такої ж амплітуди, що і на елементі. Антенний струм є малим, тому для виділення сигналу необхідно між антенною і нульовою шиною встановити резистор 1-10 МОм.

Розглянемо коректність методу з точки зору похибки, яка виникає внаслідок нехтування величиною послідовного опору  $R_s$ . Умовою такого нехтування є нерівність:

$$R_s \ll R_{sh}; \quad (1)$$

Для якісних сонячних панелей дана умова є дійсною, оскільки їх послідовний опір дорівнює 3-30 Ом, а шунтовий опір складає 3-10 кОм.

Для розрахунку шунтового опору окремих сонячних елементів необхідно спочатку виміряти шунтовий опір панелі.

Вимірювання шунтового опору панелі необхідно виконувати на достатньо низькій частоті вимірювального сигналу, щоб можна було нехтувати ємнісною складовою імпедансу панелі. При типовій ємності панелі 60 нФ на частоті  $f = 100$  Гц її реактивний опір розрахуємо за формулою:

$$X_c = 1/2\pi fC; \quad (2)$$

Розраховане значення дорівнює 2,6 Мом, що значно більше типового значення  $R_{sh}$ . Таким чином, при вимірюванні шунтового опору панелі виконується умова:

$$X_c \gg R_{sh}; \quad (3)$$

При збільшенні частоти вимірювальний сигнал буде змінювати фазу на таку величину, яку можна виміряти осцилографом. Це надає можливість виміряти імпеданс сонячної панелі і її електричну ємність.

З антенних датчиків сигнал можна подавати безпосередньо на цифровий осцилограф (Рис. 3).

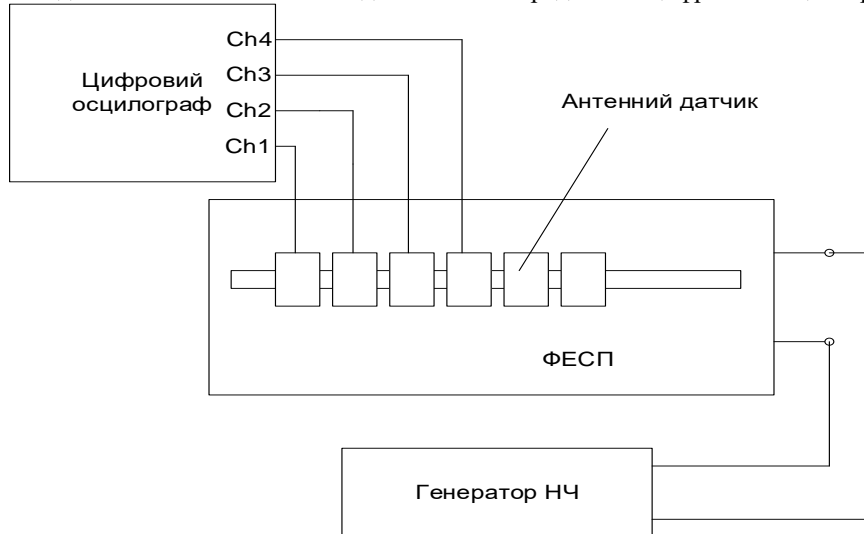


Рис. 3. Структурна схема стенду для безконтактного вимірювання: ФЕСП – фотоелектрична сонячна панель

Як зазначено вище, основним недоліком звичайного безконтактного методу вимірювання шунтового опору сонячних елементів є недостатня роздільна здатність цифрового осцилографа, причиною чого є велика систематична похибка амплітудних вимірювань. Тому розвиток методу потребує використання в якості датчиків активної антени. Для підсилення антенного сигналу авторами обрана схема диференційного підсилювача, входами якого сигнали від двох антен, які розташовують над сусідніми сонячними елементами.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

**Метою роботи** є вдосконалення методу безконтактного вимірювання електричних параметрів окремих сонячних елементів у складі фотоелектричної сонячної панелі на основі використання активної антени.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Запропонований вдосконалений варіант методу безконтактного вимірювання реалізовано в лабораторному стенді (Рис. 4).

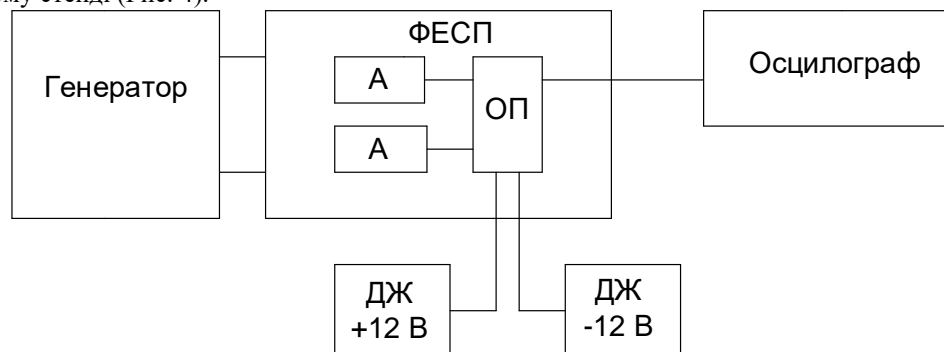


Рис.4. Лабораторний стенд вдосконаленого методу безконтактного вимірювання: А – секція антенного датчика; ДЖ – джерело живлення; ОП – схема на операційному підсилювачі; ФЕСП – фотоелектрична сонячна панель

Для лабораторного стенду були обрані такі прилади:

- цифровий осцилограф Tektronix TDS1002;
- два джерела живлення PS-1502DD;
- функціональний генератор JDS6600.

Дослідження проводили на сонячній панелі 60 Вт, 12 В із полікристалічного кремнію, яка має 36 елементів.

Антенні секції виготовлені із склотекстоліту товщиною 1 мм та з одностороннім вкриттям міддю. Розміри секцій: 30x140 мм.

В якості операційного підсилювача обрано мікросхему LM324N, яка має широке призначення.

Для визначення точних номіналів елементів електричної схеми було проведено моделювання в програмі Multisim 11. В моделі сонячні елементи представлені відповідно до еквівалентної схеми (Рис. 2). При цьому сукупність сонячних елементів до антенного датчика та після нього замінені на еквівалентні резистори та конденсатори.

Величини шунтового опору і електричної ємності сонячних елементів в моделі є типовими для кремнієвих сонячних панелей.

Результати моделювання дозволили також визначити параметри вимірювального сигналу від генератора. Форма сигналу – синусоїда. Результат роботи моделі представлено на Рис. 5.

Відповідно до моделі (Рис. 5) перший канал осцилографа виводить підсилений диференціальний сигнал від двох секцій антенного датчика, його амплітуда складає приблизно 50% від амплітуди сигналу з антени. Таким чином вирішена задача підвищення точності вимірювання шунтового опору, оскільки тут вимірюють підсилену різницю сигналів від двох антен, які розташовані над сусідніми сонячними елементами.

Самі антенні датчики в моделі представлені електричними ємностями, величина яких була розрахована за формулою для плоского конденсатора із двошаровим діелектричним заповненням: склом і склотекстолітом:

$$C = \frac{\epsilon_0 S}{\frac{d_1}{\epsilon_1} + \frac{d_2}{\epsilon_2}}; \quad (4)$$

де  $S$  – площа пластини конденсатора;  $d_1$ ,  $d_2$  – товщина першого та другого шару діелектрика;  $\epsilon_1$  та  $\epsilon_2$  – діелектрична стала першого та другого шару діелектрика;  $C$  – електрична ємність.

За формулою (4) розраховують ємність при розташуванні датчиків на лицьовій стороні сонячної панелі, яка має захисне скло. На тильній стороні скло відсутнє і ємність датчика значно зростає.

Модель побудована для сонячної панелі 270 Вт, 24 В із 60 елементами та датчиками з площею 150 кв. см кожний. Результати моделювання виявилися придатними для побудови універсальної схеми підсилювача, який і був експериментально досліджений на панелі 60 Вт, 12 В.

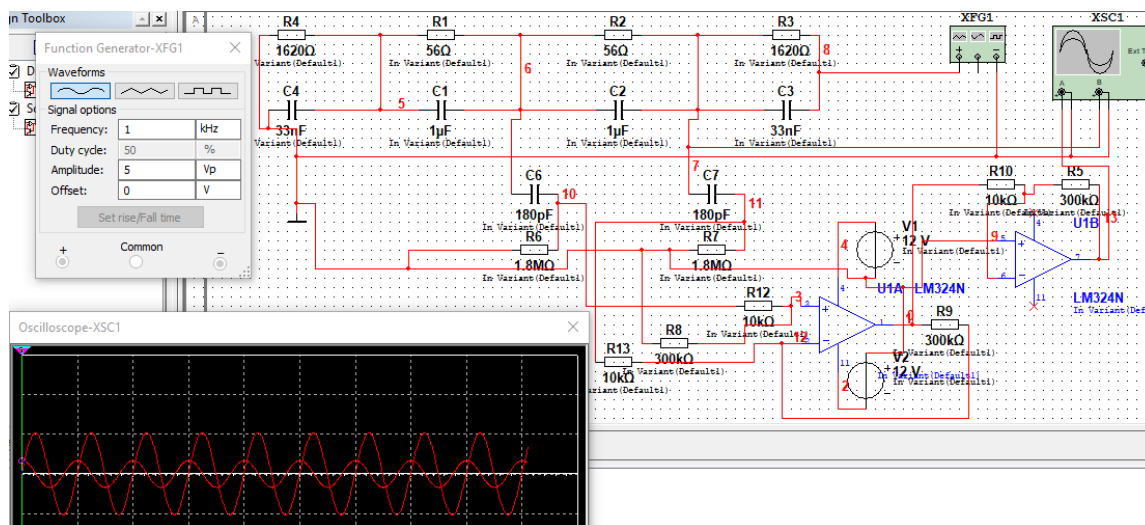


Рис. 5. Модель лабораторного стенду на основі активної антени в програмі Multisim11: розгортка осцилографа – 1мс/поділку; масштаб напруги – 5В/поділку для обох каналів; частота вимірювального сигналу 1 кГц; діюче значення сигналу на виході генератора 5В

При роботі з моделлю з'ясовано важливий факт: величину резистора навантаження для активних антенних датчиків слід значно зменшити у порівнянні із звичайною антеною, а саме з 10 МОм до 1,8 МОм, що і було здійснено практично. Необхідність зменшення величини резистора у схемі підсилювача викликана недостатнім входним опором операційного підсилювача LM324N, що і показало моделювання.

Лабораторний стенд в дії, який було зібрано відповідно до схеми (Рис. 4) наведено на Рис. 6.



Рис. 6. Лабораторний стенд з безконтактного вимірювання шунтового опору сонячних елементів в дії: на задньому плані фото видно тильну сторону сонячної панелі, самі антени датчики в кадрі відсутні



При використанні довгих дротів для з'єднання складових стенду на виході схеми підсилення з'являється імпульсна завада частотою 50 кГц, яка модульована сигналом від датчика. Джерелом завади є імпульсне джерело живлення цифрового осцилографу. При заміні осцилографа на інший завада зникає.

Процес безконтактного вимірювання шунтового опору на лицьовій стороні сонячної панелі за допомогою активної антени наведено на Рис. 7.

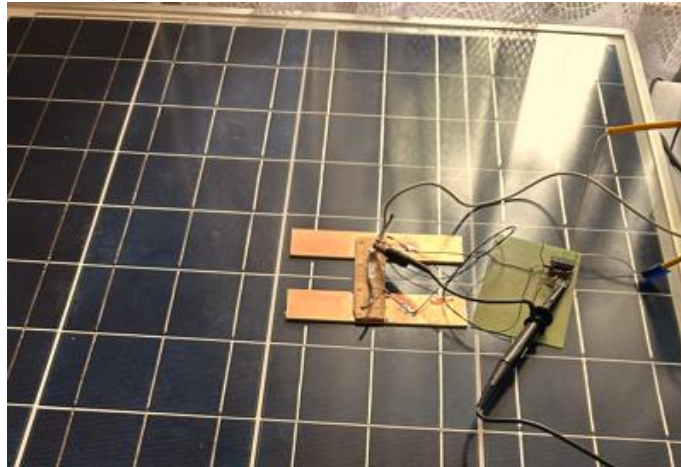


Рис. 7. Процес безконтактного вимірювання шунтового опору сонячних елементів на лицьовій стороні сонячної панелі

Щодо самої електричної схеми підсилювача, то вона має два каскади, перший з яких – диференціальний підсилювач, другий – інвертуючий підсилювач. Розрахунковий коефіцієнт підсилення для кожного із каскадів приблизно дорівнює 30.

Вихідний сигнал активної антени, розташованої на лицьовій стороні сонячної панелі, наведено на Рис. 8.

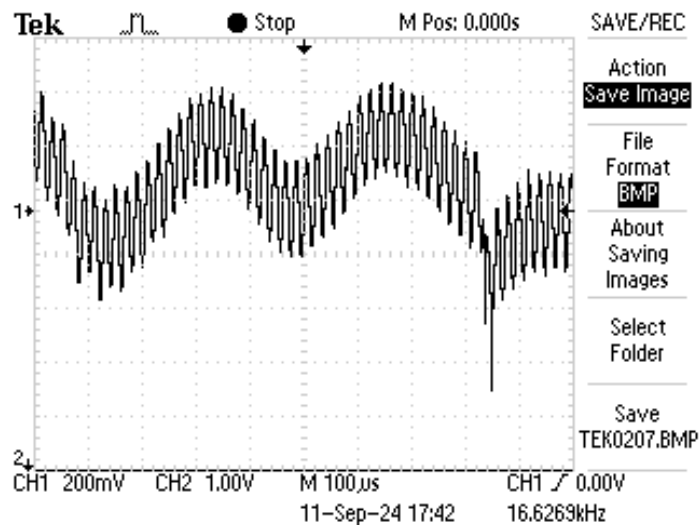


Рис. 8. Осцилограма вихідного сигналу активної антени на фоні завади: частота вимірювального сигналу 2,8 кГц; частота завади 50 кГц

Аналіз осцилограми (Рис. 8) дозволяє визначити амплітуду вихідного сигналу активної антени, яка приблизно складає 250 мВ. Це значення вдсятеро відрізняється від результату моделювання. Для збільшення амплітуди сигналу необхідно додати третій каскад у схемі підсилювача. Для придушення імпульсної високочастотної завади необхідно додати до схеми активний фільтр низьких частот, який слід побудувати на одній секції мікросхеми LM324N.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

В результаті проведених досліджень були вирішені такі наукові і практичні задачі:

- розглянуто сучасний стан основних напрямків досліджень з безконтактних методів вимірювання електричних параметрів сонячних елементів;
- розглянуто базовий метод безконтактного вимірювання шунтового опору сонячних елементів у складі панелі за допомогою антенних датчиків;

- проаналізовано обмеження базового методу і обґрунтовано його розвиток на основі використання активної антени;
- розроблено модель лабораторного стенду для безконтактних вимірювань у програмі Multisim11;
- створено діючий макет активної антени та проведено його випробування у складі лабораторного стенду;
- отримано експериментальні результати, за якими визначені основні недоліки схеми, а саме недостатній коефіцієнт підсилення та недостатня завадостійкість до високочастотних завад;
- запропоновано шляхи покращення діючої схеми активної антени на основі додавання третього каскаду підсилення і активного фільтру низьких частот.

Наступним етапом досліджень, окрім вдосконалення електричної схеми активної антени, є експериментальне визначення шунтового опору кожного із сонячних елементів на різних сонячних панелях та отримання даних щодо режимів і параметрів вимірювального сигналу для кожного із типів панелей.

### Література

1. К.М. Божко, К.Я. Мушкет, “Безконтактний метод вимірювання шунтового опору окремих сонячних елементів у складі панелі”, *Технічна інженерія*, №2(92), 2023, с.174-181. DOI: 10.26642/ten-2023-2(92)-174-181.
2. О.Ю. Гаєвський, В.Ю. Іванчук, І.О. Корнієнко, “Система вимірювання параметрів фотоелектричних модулів в реальних умовах експлуатації”, *Відновлювальна енергетика*, №2(57), 2019, с. 32-39. DOI: 10.36296/1819-8058.2019.2(57).32-39.
3. Augusto, A, Killam, S.G. Bowden, H. Wilterdink, “Measuring outdoor I–V characteristics of PV modules and systems”, *Progress in Energy*, Vol.4, 2022, 33 p. DOI: [10.1088/2516-1083/ac851c](https://doi.org/10.1088/2516-1083/ac851c).
4. J. Liu, A. Melnikov, A. Mandelis, “Contactless measurement of electrical parameters and estimation of current-voltage characteristics of Si solar cells using the illumination intensity dependence of lock-in carrierography (photoluminescence) images”, *Journal of Applied Physics*, Vol. 114, 2013, 6p. DOI: 10.1063/1.4821120.
5. K.S. Chan at al., “Application of non-contact quantum efficiency measurement for solar cell fabrication process insights”, *Solar Energy*, Vol. 233, 2022, pp. 494-503. DOI: 10.1016/j.solener.2021.12.054.
6. M. Musztyfaga-Staszuk, D. Junicki, P. Panek, “Correlation of Different Electrical Parameters of Solar Cells with Silver Front Electrodes”, *Materials*, Vol. 12, 2019, 12p. DOI: 10.3390/ma12030366.
7. S. Miyajima, K. Nishioka and Y. Hishikawa, "Non-contact Voltage Measurement of Solar Cell with Electrostatic Voltmeter," *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, Washington, DC, USA, 2017, pp. 481-483, DOI: 10.1109/PVSC.2017.8366426.
8. J. Hishikawa, K. Yamagoe, T. Onuma, “Non-contact measurement of electric potential of photovoltaic cells in a module and novel characterization technologies”, *Japan Journal of Applied Physics*, Vol.54(8S1), 2015, 08KG05. DOI: 10.7567/JJAP.54.08KG05.
9. M. Greulich at al., “Contactless measurement of current-voltage characteristics for silicon solar cells”, *Solar Energy Materials and Solar Cells*, Vol. 248, 2022, 111931. DOI: 10.1016/j.solmat.2022.111931.

### References

1. K.M. Bozhko, K.Ia. Mushket, “Bezkontaktniy metod vymiryuvannya shuntovoho oporu okremykh soniachnykh elementiv u skladi paneli”, *Tekhnichna inzheneriia*, №2(92), 2023, s.174-181. DOI: 10.26642/ten-2023-2(92)-174-181.
2. O.Iu. Haievskiyi, V.Iu. Ivanchuk, I.O. Korniienko, “Systema vymiryuvannya parametriv fotoelektrychnykh moduliv v realnykh umovakh ekspluatatsii”, *Vidnovliuvalna enerhetyka*, №2(57), 2019, s. 32-39. DOI: 10.36296/1819-8058.2019.2(57).32-39.
3. Augusto, A, Killam, S.G. Bowden, H. Wilterdink, “Measuring outdoor I–V characteristics of PV modules and systems”, *Progress in Energy*, Vol.4, 2022, 33 p. DOI:10.1088/2516-1083/ac851c.
4. J. Liu, A. Melnikov, A. Mandelis, “Contactless measurement of electrical parameters and estimation of current-voltage characteristics of Si solar cells using the illumination intensity dependence of lock-in carrierography (photoluminescence) images”, *Journal of Applied Physics*, Vol. 114, 2013, 6p. DOI: 10.1063/1.4821120.
5. K.S. Chan at al., “Application of non-contact quantum efficiency measurement for solar cell fabrication process insights”, *Solar Energy*, Vol. 233, 2022, pp. 494-503. DOI: 10.1016/j.solener.2021.12.054.
6. M. Musztyfaga-Staszuk, D. Junicki, P. Panek, “Correlation of Different Electrical Parameters of Solar Cells with Silver Front Electrodes”, *Materials*, Vol. 12, 2019, 12p. DOI: 10.3390/ma12030366.
7. S. Miyajima, K. Nishioka and Y. Hishikawa, "Non-contact Voltage Measurement of Solar Cell with Electrostatic Voltmeter," *2017 IEEE 44th Photovoltaic Specialist Conference (PVSC)*, Washington, DC, USA, 2017, pp. 481-483, DOI: 10.1109/PVSC.2017.8366426.
8. J. Hishikawa, K. Yamagoe, T. Onuma, “Non-contact measurement of electric potential of photovoltaic cells in a module and novel characterization technologies”, *Japan Journal of Applied Physics*, Vol.54(8S1), 2015, 08KG05. DOI: 10.7567/JJAP.54.08KG05.
9. M. Greulich at al., “Contactless measurement of current-voltage characteristics for silicon solar cells”, *Solar Energy Materials and Solar Cells*, Vol. 248, 2022, 111931. DOI: 10.1016/j.solmat.2022.111931.

<https://doi.org/10.31891/2219-9365-2024-79-15>

УДК 004.032.26: 004.93'1

МАСТЕНКО Ігор

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

<https://orcid.org/0000-0002-2953-4589>

[ihormas@gmail.com](mailto:ihormas@gmail.com)

СТЕЛЬМАХ Наталія

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

<https://orcid.org/0000-0003-1876-2794>

[n.stelmakh@kpi.ua](mailto:n.stelmakh@kpi.ua)

## ЗАСТОСУВАННЯ НЕЙРОМЕРЕЖ YOLO В СИСТЕМАХ АВТОМАТИЗОВАНОГО КОНТРОЛЮ ЯКОСТІ ПРИ ВИЯВЛЕННІ ДЕФЕКТІВ ВИРОБІВ ОРГАНІЧНОГО ПОХОДЖЕННЯ

*Візуальна інспекція є невід'ємною складовою контролю якості продукції виробництва. Системи машинного зору широко використовуються в різних промислових галузях, займаючи важливе місце в сучасних системах контролю та перевірки якості. Це пов'язано, перш за все, з високим рівнем конкуренції на ринку і зростаючими стандартами до якості продукції. Справжнім викликом став процес контролю якості виробів органічного походження засобами машинного зору, зокрема з використанням нейронних мереж, адже такі вироби містять багато різноманітних дефектів і кожен виріб є унікальним, не схожим на інший. Для вирішення цих задач пропонуються застосування комплексного підходу, а саме технології локалізації та класифікації дефектів виробів засобами нейронних мереж.*

*Ключові слова: дефектоскопія, автоматизована система контролю, якість продукції, машинний зір, глибоке навчання, нейронні мережі.*

MASTENKO Ihor, STELMAKH Nataliia

National Technical University of Ukraine "Ihor Sikorskyi Kyiv Polytechnic Institute"

## APPLICATION OF YOLO NEURAL NETWORKS IN DEFECT DETECTION IN AUTOMATED QUALITY CONTROL SYSTEMS FOR PRODUCTS OF ORGANIC ORIGIN

*Modern technological equipment for various purposes must fully meet the requirements of digital production and be able to quickly integrate into the structure of smart enterprises that are switching to cyber-physical technological systems. Machine vision (MV) is a key element of such systems and a promising automation tool that enables the capture and movement of various objects, including components, quality control, and safety. In addition, machine vision is increasingly being used in modular machine tools. The article analyses various applications of machine vision, in particular, its use in intelligent technological systems for product quality control. Particular attention is paid to fast and efficient quality analysis at the stage of the production process, which allows for accurate defect detection. The article investigates the feasibility of using mathematical models of artificial neural networks to create an intelligent system for monitoring the geometric condition of products. The aim of the study is to identify and classify the quality parameters of products of organic origin, namely reed tubes. For this purpose, new quality control methods based on computer vision and machine learning algorithms are proposed to identify and classify various types of defects using an integrated approach, namely, technologies for localization and classification of product defects using neural networks. As an example, we consider products of organic origin, for which the localisation and classification of defects is difficult due to their natural structure and uniqueness. To solve this problem, the use of neural networks of several YOLO architectures was first proposed. The study presented the results of training two modifications of the YOLOv10s and YOLOv10m neural networks, which were positive.*

*Keywords: flaw detection, automated quality control system, product quality, machine vision, deep learning, neural networks.*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Питання організації ефективного контролю якості залишаються ще досі актуальними, оскільки на виробництві все частіше використовують нові види матеріалів, такі наприклад як композитні або комплектуючі для виробів отриманні шляхом застосування адитивних технологій, що в свою чергу потребує особливого постоброблення, та відповідно контролю якості. Використання традиційних методів контролю якості продукції, таких як ультразвукова, магнітна, рентгенівська дефектоскопія, аналіз звукових хвиль для виявлення внутрішніх дефектів у матеріалах все ще залишаються актуальними. Натомість візуальний огляд виробів на предмет видимих дефектів, використовуючи людський фактор, таких як тріщини, подряпини, неправильна форма, втрачає популярність і це пов'язано з додатковими витратами на людську працю та значною мірою залежить від психоемоційного і фізичного стану контролерів. Важливим є також швидкий і ефективний аналіз якості продукції на етапі виробничого процесу, що забезпечує високу точність виявлення дефектів [1].

Тому впровадження сучасних комп'ютерних технологій стає дедалі актуальнішим у інтелектуальних технологічних системах для контролю якості продукції, а також для моніторингу стану технологічних

процесів і виробничого обладнання (верстатів, інструментів тощо). Зокрема, до таких належать **комп'ютерний зір** із застосуванням алгоритмів машинного навчання для автоматичного виявлення і класифікації дефектів та технології **Інтернет речей (IoT)**, а саме сенсори, що аналізують виробничий процес в реальному часі, допомагаючи виявляти аномалії. [2] Детектування та класифікація дефектів виробів є важливими етапами контролю якості на виробництві. У цій статті пропонується підхід, який досягає цих цілей, а саме детектує і класифікує дефекти на виробках органічного походження, та показує високу точність результатів.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Машинний зір є однією з найбільш динамічних областей, що розвивається в інтелектуальних технологіях. За прогнозами, загальний ринок машинного зору в промисловості в США досягне 12,29 мільярдів доларів США до 2023 року [3]. Ці дані вказують на високу потребу великих компаній у впровадженні максимальної автоматизації технологічних процесів, зокрема в операціях контролю якості продукції [4]. Зростання попиту на системи машинного зору обумовлене четвертою промисловою революцією, Індустрією 4.0, а також розвитком таких технологій, як штучний інтелект та Інтернет речей.

Сучасні науковці та розробники пропонують виробництву безліч інструментів для виявлення дефектів продукції, таких як різноманітні датчики, програмне забезпечення SPC, алгоритми штучного інтелекту і машинного навчання, рентгенівські системи, ультразвукові системи та вихрострумові системи. Визначення відповідного інструменту або об'єднання інструментів залежить від характеру дефекту та відповідної виробничої процедури. Певні аномалії можуть вимагати негайного моніторингу та аналізу за допомогою датчиків і програмного забезпечення SPC, в той час як інші можуть вимагати більш всебічного аналізу за допомогою рентгенівських або ультразвукових систем. [5]

Задача автоматизованої сегментації зображень дефектів металевих поверхонь також досить часто розглядається [6]. Метою дослідження є покращення алгоритмів сегментації шляхом використання методів глибокого навчання. Завдяки спеціальній архітектурі, нейронна мережа здатна ефективно генерувати маски сегментації. Навчальний набір даних для мережі містив зображення чотирьох типів дефектів, таких як відколи, тріщини та плями. Аналіз розподілу класів у цьому наборі показав, що класи є незбалансованими, що негативно позначається на процесі навчання.

Можливість застосування моделі глибокого навчання YOLOv5s для автоматизації виявлення поверхневих дефектів на друкованих платах досліджується досліджувалась групою науковців. [7] Модель YOLOv5s була реалізована та навчена для оцінки її ефективності в цьому завданні. Для навчання використовувався відкритий набір даних під назвою «Дефекти друкованих плат». Було проведено як якісний, так і кількісний аналіз продуктивності моделі на тестовому наборі даних. Результати показали, що модель здатна виявляти дефекти на поверхні друкованих плат з точністю 92,5% за показником mAP50. Також був проведений аналіз виявлення різних типів дефектів, після чого запропоновано рекомендації для подальшого вдосконалення системи, зокрема, використання аугментації даних та більш складної архітектури глибокого навчання.

Так у роботі [8] автори представили алгоритм комп'ютерного зору для класифікації яблук. Проведена робота по автоматичному видаленні фону зображень. Використовують сегментацію дефектів за допомогою двошарової нейронної мережі MLP для кожного пікселя на основі його значень R, G, B і H. Відокремлюють кожен об'єкт контролю до відповідної категорії якості, використовуючи 16 витягнутих ознак, включаючи 8 статистичних, 5 текстурних і 3 геометричних ознаки за трьома класифікаторами SVM, MLP і KNN. Результати показують, що після навчання класифікатора SVM на 96 (3 × 32) навчальних зображеннях яблук він досяг найкращої продуктивності 92,5% і 89,2% для двох категорій і трьох категорій якості, відповідно, протягом 24 (3 × 8) тестів. зображення яблук, враховуючи K-згортання з K = 5 для 120 (3 × 40) доступних зображень.

Фанг та інші в своїй статті [9], зосередились на плоских металевих матеріалах із подібними вимогами до контролю якості, і вона є першою, яка всебічно узагальнює методи виявлення дефектів з двовимірного та тривимірного аспектів. Введено та узагальнено методи навчання, теоретичне обговорення та розвиток застосування виявлення поверхневих дефектів промислових металевих плоских матеріалів. Ця стаття узагальнює результати досліджень автоматизованого візуального виявлення дефектів металевих пластин і смугових поверхонь за останні 30 років, більшість з яких були опубліковані за останні 10 років. Оскільки дослідники проводять експерименти з різними методами на різних наборах даних, важко порівняти ефективність виявлення різних технологій.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

**Метою даної роботи** - дослідити та представити результати розпізнавання дефектів виробів органічного походження трубчастої форми засобами нейронних мереж.

## ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Одним з найважливіших етапів створення програмного забезпечення для автоматизованої системи контролю та керування якістю виробів органічного походження трубчастої форми є вибір архітектури нейронної мережі, адже від цього залежать точність роботи системи контролю якості виробів та час обробки зображень, що в свою чергу впливає на загальний виробничий цикл виготовлення продукції. Переднавчені нейронні мережі можуть бути корисними для роботи з комплексними зображеннями під час контролю якості продукції, оскільки вони допомагають вирішити проблему нестачі анотованих даних для навчання моделей з нуля. Потрібно розглядати архітектури, які спеціалізуються на обробці зображень. Враховуючи специфіку поставленої задачі, необхідно обрати архітектуру, яка буде базуватись на класифікації зображень. Найбільш оптимальним варіантом будуть саме згорткові нейронні мережі (CNN), оскільки вони здатні виявляти локальні ознаки та патерни на зображеннях, адже вони спеціально розроблені для вирішення задач обробки зображень. Згорткові шари нейронних мереж можуть автоматично вивчати локальні ознаки на зображеннях, такі як форми, кольори та текстури, що є важливими критеріями при вирішенні задач і класифікації і детектування [10].

Вибір конкретної архітектури нейронної мережі потрібно робити на основі наступних параметрів: доступні обчислювальні ресурси, складність поставленої задачі та розмірів даних, які будуть оброблятися нейронною мережею.

Проведемо аналіз існуючих архітектур нейронних мереж, що широко використовуються для вирішення задач комп'ютерного зору та які можна використати для вирішення задач контролю та керування якістю виробів.

Однією з найбільш поширених та ефективних архітектур для вирішення багатьох завдань з обробки зображень є ResNet (Residual Neural Network). Дана архітектура завдяки додаванню скорочених з'єднань дозволяє запобігти виникненню проблеми зниклого градієнту, завдяки чому є можливість побудувати глибоку мережу, яка досягне високої продуктивності. При цьому, архітектури ResNet мають відносно складну реалізацію та потребують великого об'єму даних для навчання нейронної мережі, що можна назвати одним з недоліків даної архітектури. [11]

Варто зауважити, що все ж архітектура ResNet буде більш ефективною при вирішенні задач класифікації і не є вдалим рішенням коли необхідно детектувати та розпізнати дефекти виробу органічного походження коли на одиниці площі може бути відразу декілька класів дефектів, наприклад тріщини, пігментація та сколи після порізки матеріалу для очеретяних трубок.

YOLO (You Only Look Once) — це тип згорткової нейронної мережі, розробленої для задачі детектування об'єктів на зображеннях у реальному часі. Основними особливості YOLO є одноетапний підхід, а саме на відміну від традиційних методів, які спочатку виділяють регіони інтересу, а потім класифікують їх, YOLO обробляє зображення за один раз і це дозволяє досягти високої швидкості. Також зображення розбивається на сітку (наприклад, 7x7), де кожен осередок відповідає за виявлення об'єктів, які його центрують. Кожен осередок генерує прогнози для об'єктів, їх класів і координат. Це швидка та ефективна архітектура для великих наборів даних, яка робить прогнози обробивши фотографію лише один раз, що значно прискорює роботу. [12-13] На рис. 1 представлено типову структуру мережі YOLO [14].

Як було вже зазначено об'єктом контролю було розглянуто вироби органічного походження трубчастої форми, а саме очеретяні трубочки для напоїв рис. 2.



Рис. 1. Типова архітектура YOLO

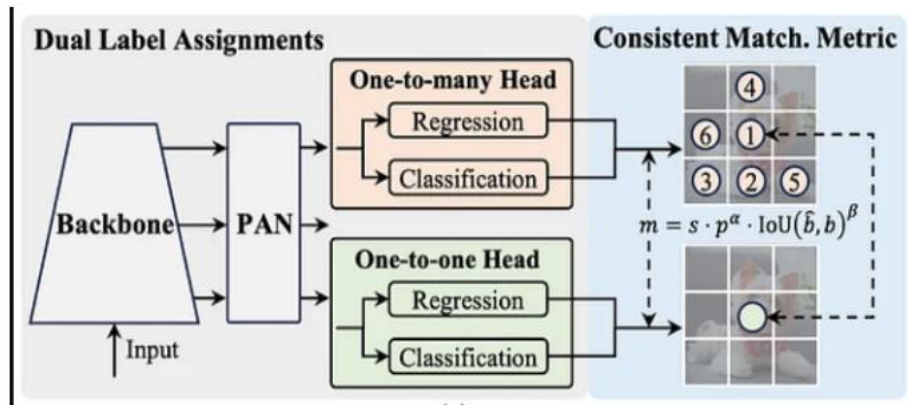


Рис. 2. Очеретяні трубочки для напоїв

Як видно з рисунка 2 очеретяні трубочки мають різний відтінок кольору, різну форму та розміри перерізу, що обумовлено їх органічним походженням і це потребує контролю та сортування виробів на одному із завершальних етапів виробництва. Окрім того на виробках є різні дефекти як внутрішні так і зовнішні, як виправні та і невивправні, які необхідно відбракувати.

Для вирішення цього завдання було запропоновано використати розроблену класифікацію дефектів.

Результат сортування об'єктів можна розділити на п'ять класів [15]:

1. Пігментація (явний, виправний; малозначний; поверхневий дефект) потребує повторної хімічної обробки.
2. Тріщини (Явний, невивправний; Значний; Поверхневий дефект) - брак.
3. Отвори в бокових стінках очеретяної трубки. (явний, невивправний, поверхневий, об'ємний дефект) - брак.
4. Наявність плівки всередині (явний, невивправний, малозначний, об'ємний дефект) потребує механічного доопрацювання.
5. Дефекти порізки (сколи торців) . Явний, невивправний; критичний; поверхневий дефект – брак.

### Характеристики набору даних

#### Dataset Split

Train Set 2838 Images (946\*3)

Valid Set 190 Images

Test Set 125 Images

Total 3153 Images (1261 unique)

#### Classes

dirty 1,397 Images

break 576 Images

hole 330 Images

plivka 166 Images

cut 140 Images

#### Augmentations

Outputs per training example: 3

Flip: Horizontal, Vertical

Shear:  $\pm 5^\circ$  Horizontal,  $\pm 15^\circ$  Vertical

Saturation: Between -30% and +30%

Brightness: Between -15% and +15%

Exposure: Between -10% and +10%

Blur: Up to 3.5px

Розглянемо результати дослідження локалізації та класифікації дефектів виробів засобами нейронних мереж для обраних архітектур YOLOv10s та YOLOv10m. Результати навчання наведені нижче.

Результати навчання YOLOv10s

45 epochs

YOLOv10s summary (fused): 383 layers, 20418862 parameters, 98.0 GFLOPs

Class	Instances	Box (P	R	mAP50	mAP50-95)
all	388	0.58	0.560	0.573	0.261
break	79	0.445	0.430	0.427	0.148
cut	18	0.749	0.833	0.874	0.500
dirty	207	0.541	0.391	0.466	0.284
hole	61	0.612	0.492	0.514	0.150
plivka	23	0.556	0.652	0.586	0.221

Speed: 0.1ms preprocess, 7.7ms inference, 0.0ms loss, 0.2ms postprocess per image

Матриця помилок та графіки навчання представлені на рис. 3,4.

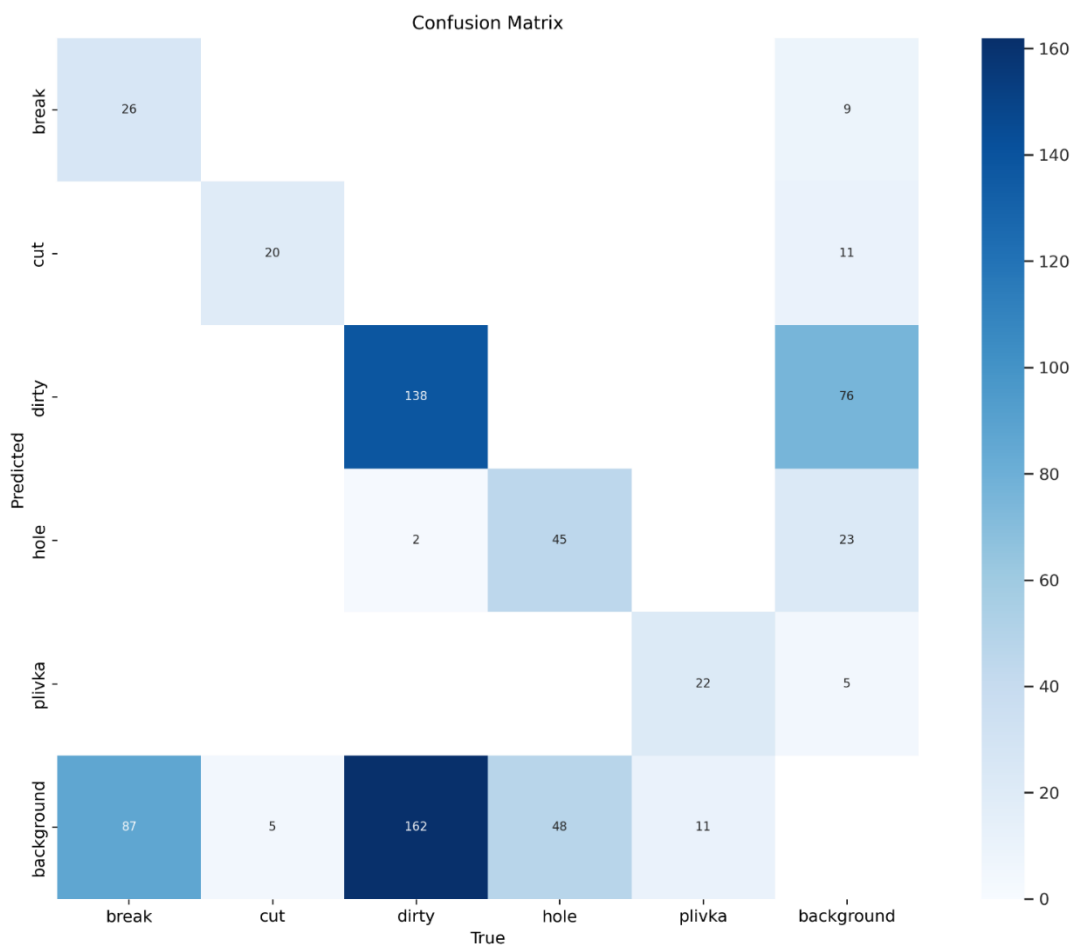


Рис. 3. Матриця помилок та графіки навчання YOLOv10s

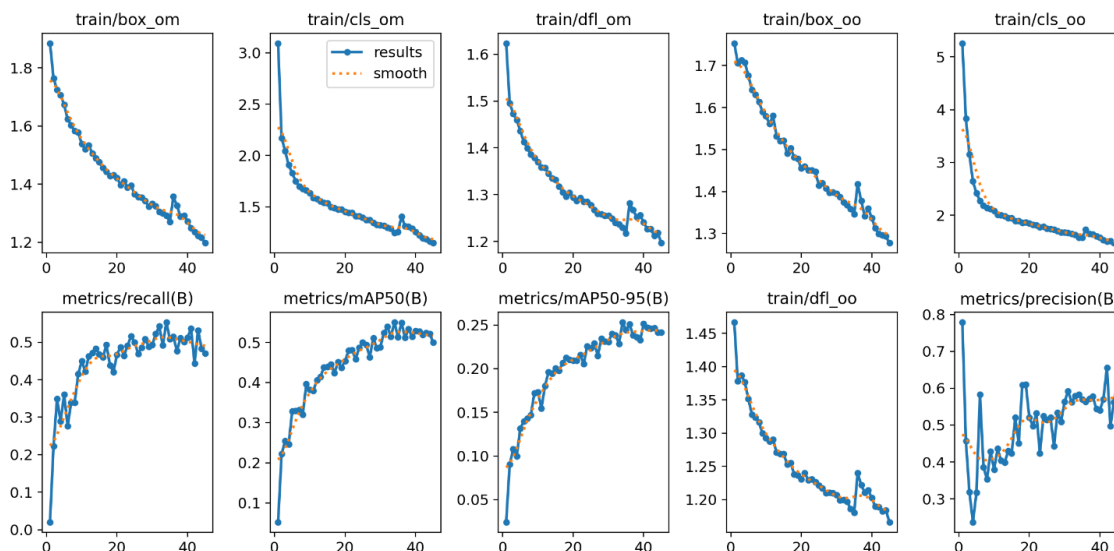


Рис. 4. Метрики навчання

де, **box\_om** - Box Objectness Metric. Це метрика, яка оцінює, наскільки добре модель прогнозує наявність об'єкта в рамці (bounding box). Об'єктність — це впевненість моделі в тому, що в рамці є об'єкт.

**cls\_om** - Class Objectness Metric (можливо Classification Objectness Metric). Це метрика, яка оцінює точність класифікації об'єкта, що знаходиться в рамці.

**dfl\_om** - Distribution Focal Loss Objectness Metric. Це метрика, пов'язана зі статистичним законом розподілу фокальної функції втрат (Focal Loss) для об'єктності.

**box\_oo** - Box Objectness Optimization. Це метрика, яка використовується для оптимізації об'єктності рамки.

**cls\_oo** - Class Objectness Optimization. Це метрика, яка використовується для оптимізації класифікації об'єктів.

**dfl\_oo** - Distribution Focal Loss Optimization. Це метрика, яка використовується для оптимізації закону розподілу Focal Loss.

#### **mAP50**

Mean Average Precision at IoU=0.50.

mAP50 означає середню точність (average precision, AP) при порозі Intersection over Union (IoU) рівному 0.50.

**IoU** — це метрика, що вимірює перекриття між передбаченим рамкою (bounding box) та справжньою рамкою (ground truth box). IoU визначається як площа перетину поділена на площу об'єднання двох рамок.

**AP** розраховується як середнє значення точності (precision) для різних значень recall. mAP (mean AP) — це середнє значення AP для всіх класів.

**mAP50** вказує на те, що для об'єкт вважається правильно детектованим, якщо  $\text{IoU} \geq 0.50$ .

#### **mAP50-95**

Mean Average Precision from IoU=0.50 to IoU=0.95.

mAP50-95 означає середню точність при порогах IoU від 0.50 до 0.95 з кроком 0.05.

Ця метрика більш строга та дає краще уявлення про загальну продуктивність моделі, оскільки враховує різні рівні перекриття.

mAP50-95 розраховується як середнє значення AP для IoU на рівнях 0.50, 0.55, 0.60, ..., 0.95.

На рисунку 5 представлено приклад детектування дефекту типу тріщина.



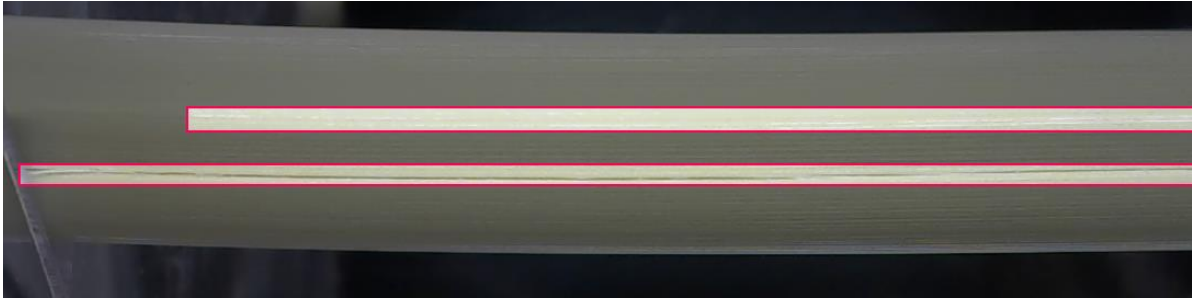


Рис. 5. Дефект типу «break»

Як видно з рисунка 5 нижній дефект виділяється чітко, верхній – важко помітний навіть людині (зображення з навчального набору даних)

### Результати навчання YOLOv10m

45 epochs

YOLOv10m summary (fused): 369 layers, 16537854 parameters, 61.8 GFLOPs

Class	Instances	Box (P	R	mAP50	mAP50-95)
all	388	0.611	0.581	0.583	0.271
break	79	0.539	0.43	0.422	0.139
cut	18	0.812	0.944	0.936	0.528
dirty	207	0.562	0.444	0.471	0.272
hole	61	0.638	0.52	0.5	0.168
plivka	23	0.504	0.565	0.586	0.25

Speed: 0.1ms preprocess, 5.8ms inference, 0.0ms loss, 0.2ms postprocess per image

Матриця помилок та графіки навчання представлені на рис. 6.

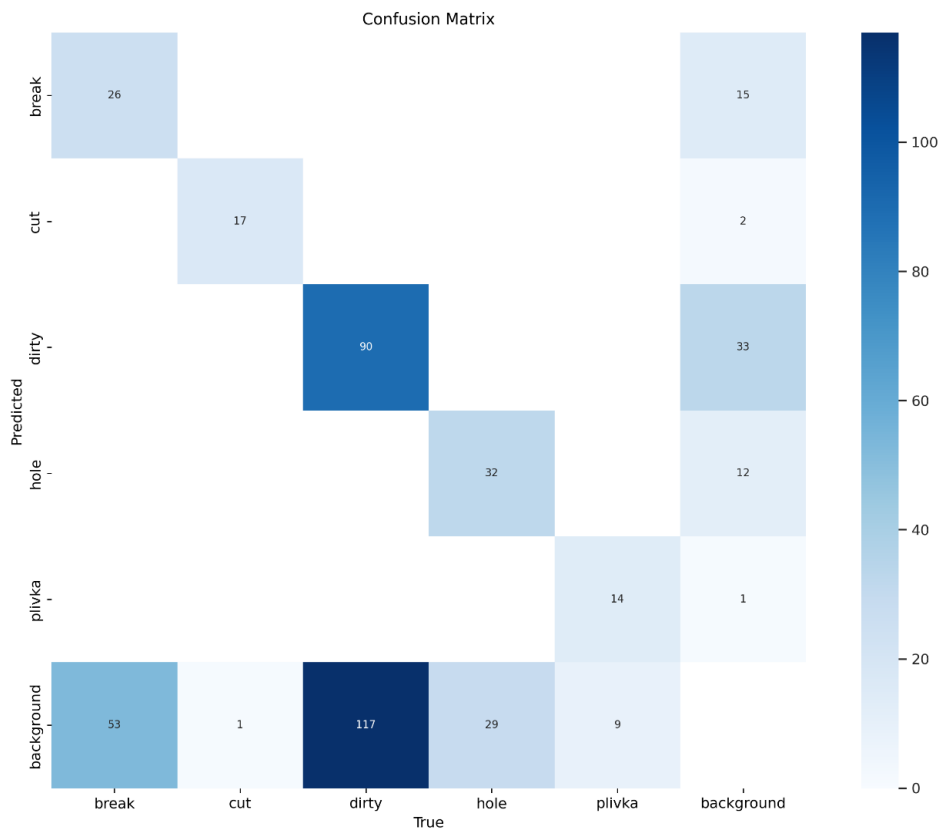


Рис. 6. Матриця помилок та графіки навчання YOLOv10m

Розглянемо приклади детектування дефектів нейронною мережею для кожного з визначених раніше класів таких як плівка, тріщини і півка, що знаходяться поруч та тріщини довгої витягнутої форми рис. 7, 8, 9.



Рис. 7. Дефект типу «hole»

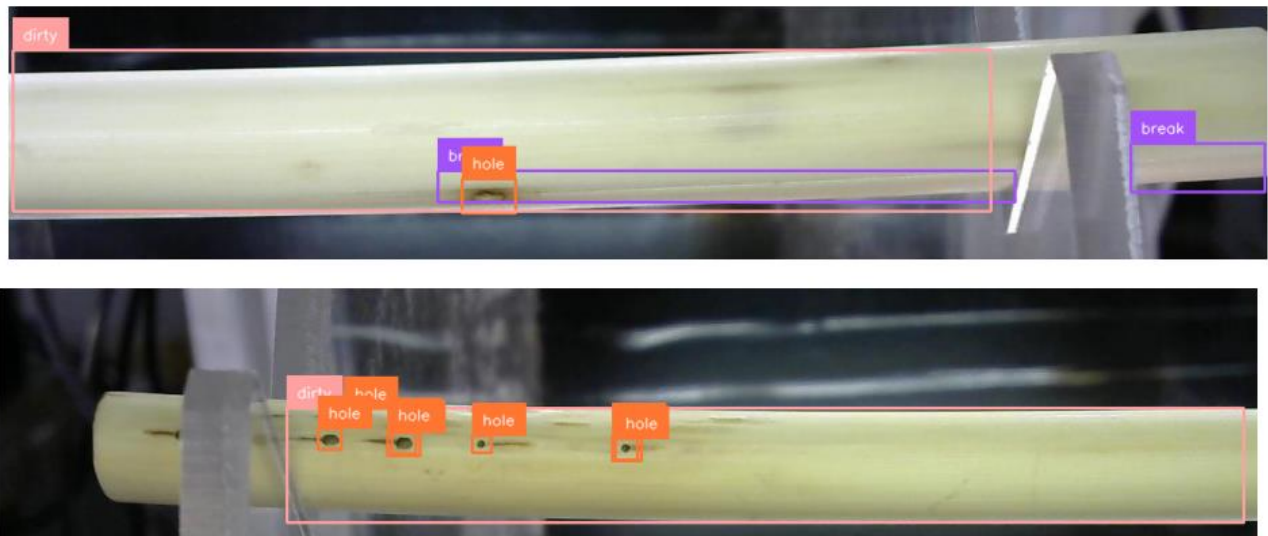


Рис. 8. Дефект типу «break» і «hole» які знаходяться поряд

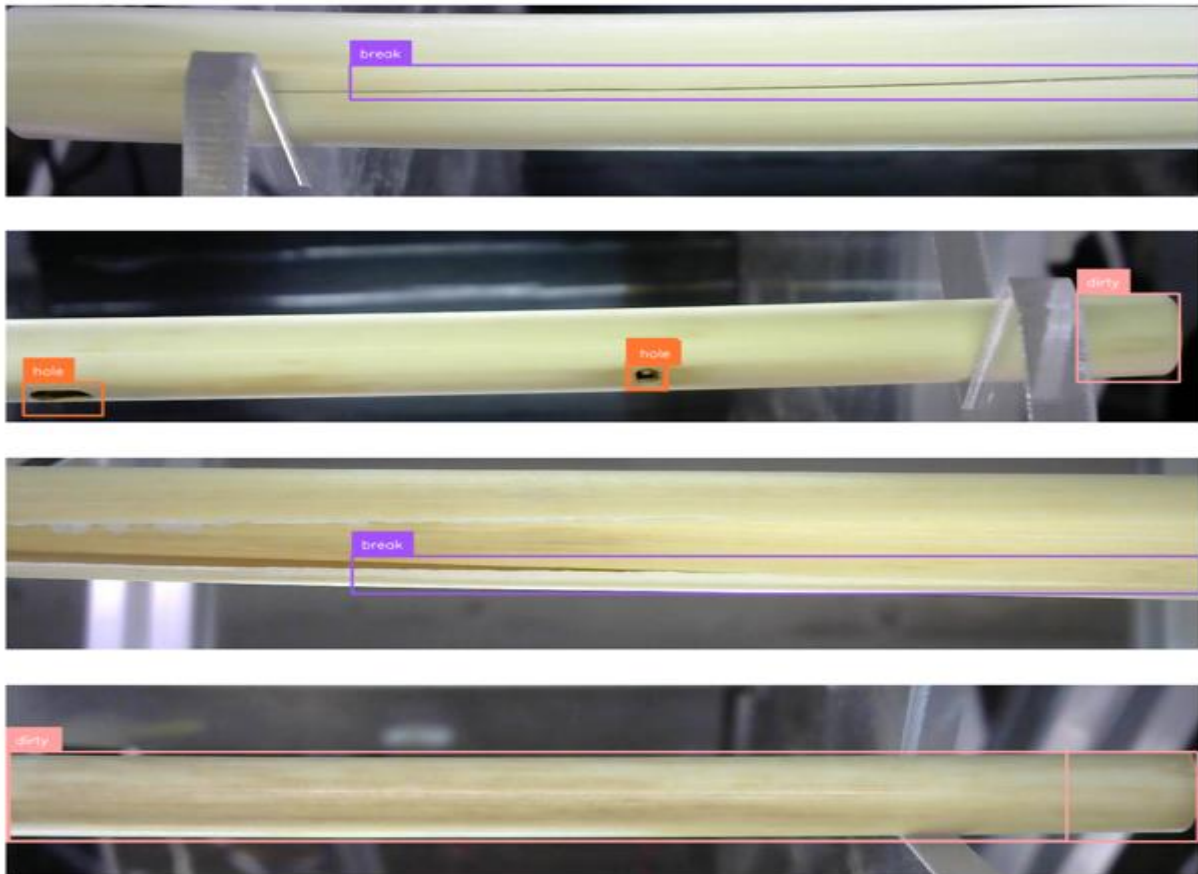


Рис. 9. Дефекти типу «break» які мають витягнуту форму

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У роботі було проаналізовано сучасні можливості виявлення поверхневих дефектів виробів. Розглянуто вироби органічного походження для яких локалізація та класифікація дефектів є ускладненою внаслідок їх природньої структури і унікальності. Для вирішення цього завдання вперше було запропоновано використання нейронних мереж декількох архітектур YOLO. В результаті дослідження було представлено результати навчання двох модифікацій нейронної мережі YOLOv10s та YOLOv10m, які були позитивними. Складність такого детектування здебільшого відображалась у неспроможності нейронної мережі локалізувати дефекти довгої протяжної форми і такі дефекти сегментувались нейронною мережею в окремі прямокутні рамки. Проте результати навчання можна покращити за рахунок збільшення об'єму та репрезентативності навчального набору даних. В перспективі подальших досліджень можна розглянути можливість модифікації моделей нейронної мережі Yolo, які мають враховувати можливість розпізнавання дефектів протяжної форми. А також для вирішення завдання ефективного розпізнавання дефектів виробів органічного походження необхідно виконати дослідження пов'язане із застосуванням інших нейромережових детекторів в результаті чого планується отримати інші результати та порівняти їх обравши найоптимальнішу архітектуру за параметром достовірності розпізнавання mAP.

#### Література

1. N. Stelmakh, I. Mastenko, O. Sulima, & T. Rudyk, “Features of the implementation of computer vision in the problems of automated product quality control”, *Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska*, 2023. 13(1), pp. 38-41, DOI: 10.35784/iapgos.3434
2. Ковалюк, Д. О., Ковалюк, О. О., & Малішевський, В. С. (2024). Реалізація систем контролю якості продукції на основі машинного зору та web-технологій. *Вісник НТУУ “КПІ імені Ігоря Сікорського”*. Серія: Хімічна інженерія, екологія та ресурсозбереження, (1), 28–34. <https://doi.org/10.20535/2617-9741.1.2024.300980>
3. Guo Y. et al.: Rotational projection statistics for 3D local surface description and object recognition. *International journal of computer vision* 105(1), 2013, 63–86.

4. Johnson A. E.: Spin-images: a representation for 3-D surface matching: Diss. Andrew Edie Johnson – Pittsburgh, Pennsylvania, 1997.
5. Hussein N. Al-Jubori1 and Izzat Al-Darrajil, Tools and Process of Defect Detection in Automated Manufacturing Systems. EAI Endorsed Transactions on Scalable Information Systems, pp 1-7. doi: 10.4108/eetsis.4000
6. Steshenko, Y., Momot A., Protasov A., Muraviov O. (2023) Automation of the process of segmentation of images of metal surface defects using the neural network U-Net. Tekh. Diahnost. ta Neruiniv. Kontrol, 2, 34-40 [in Ukrainian]. <https://doi.org/10.37434/tdnk2023.02.05>
7. Momot A. Automated defect detection in printed circuit boards based on the YOLOv5 neural network / A. Momot, V. Kretsul, O. Muraviov, R. Galagan. // The Paton Welding Journal. – 2024. – №4. – pp. 46–52. DOI: <https://doi.org/10.37434/tpwj2024.04.07>
8. Moallem, P., Serajoddin, A., Pourghassem, H., 2017. Computer vision-based apple grading for golden delicious apples based on surface features. Inf. Process. Agric. 4 (1), 33–40
9. Fang, X., Luo, Q., Zhou, B., Li, C., Tian, L. (2020). Research progress of automated visual surface defect detection for industrial metal planar materials. Sensors, 20(18), 5136. DOI 10.3390/s20185136
10. Alzubaidi, Laith, et al. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." Journal of big Data 8 (2021): 1-74.
11. McNeely-White, David, J. Ross Beveridge, and Bruce A. Draper. "Inception and ResNet features are (almost) equivalent." Cognitive Systems Research 59 (2020): 312-318.
12. Redmon J. et al. "You only look once: Unified, real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.
13. Zaidi, S.S., Ansari, M., Aslam, A. et al. (2022) A survey of modern deep learning based object detection models. Digital Signal Processing, 126, 103514. <https://doi.org/10.1016/j.dsp.2022.103514>
14. Jiang, P., Ergu, D., Liu, F. et al. (2022) A review of YOLO algorithm developments. Procedia Computer Sci., 199, 1066- 1073. <https://doi.org/10.1016/j.procs.2022.01.135>
15. Stelmakh N.V. Mastenko I.V. Classification and detection of defects in tubular products using machine vision. Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. Том 35 (74) № 2 2024. с. 167-172. DOI <https://doi.org/10.32782/2663-5941/2024.2/23>

#### References

1. N. Stelmakh, I. Mastenko, O. Sulima, & T. Rudyk, "Features of the implementation of computer vision in the problems of automated product quality control", Informatyka, Automatyka, Pomiar W Gospodarce I Ochronie Środowiska, 2023. 13(1), pp. 38-41, DOI: 10.35784/iapgos.3434
2. Kovalyuk, D.O., Kovalyuk, O.O. i Malishevs'kyy, V.S. 2024. Realizatsiya system kontrolyu yakosti produktsiyi na osnovi mashynnoho zoru ta web-tekhnologiy. Visnyk NTUU "KPI imeni Ihorya Sikors'koho". Seriya: Khimichna inzheneriya, ekolohiya ta resursozberezhennya. 1 (Ber 2024), 28–34. <https://doi.org/10.20535/2617-9741.1.2024.300980>
3. Guo Y. et al.: Rotational projection statistics for 3D local surface description and object recognition. International journal of computer vision 105(1), 2013, 63–86.
4. Johnson A. E.: Spin-images: a representation for 3-D surface matching: Diss. Andrew Edie Johnson – Pittsburgh, Pennsylvania, 1997.
5. Hussein N. Al-Jubori1 and Izzat Al-Darrajil, Tools and Process of Defect Detection in Automated Manufacturing Systems. EAI Endorsed Transactions on Scalable Information Systems, pp 1-7. doi: 10.4108/eetsis.4000
6. Steshenko, Y., Momot A., Protasov A., Muraviov O. (2023) Automation of the process of segmentation of images of metal surface defects using the neural network U-Net. Tekh. Diahnost. ta Neruiniv. Kontrol, 2, 34-40 [in Ukrainian]. <https://doi.org/10.37434/tdnk2023.02.05>
7. Momot A. Automated defect detection in printed circuit boards based on the YOLOv5 neural network / A. Momot, V. Kretsul, O. Muraviov, R. Galagan. // The Paton Welding Journal. – 2024. – №4. – pp. 46–52. DOI: <https://doi.org/10.37434/tpwj2024.04.07>
8. Moallem, P., Serajoddin, A., Pourghassem, H., 2017. Computer vision-based apple grading for golden delicious apples based on surface features. Inf. Process. Agric. 4 (1), 33–40
9. Fang, X., Luo, Q., Zhou, B., Li, C., Tian, L. (2020). Research progress of automated visual surface defect detection for industrial metal planar materials. Sensors, 20(18), 5136. DOI 10.3390/s20185136
10. Alzubaidi, Laith, et al. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions." Journal of big Data 8 (2021): 1-74.
11. McNeely-White, David, J. Ross Beveridge, and Bruce A. Draper. "Inception and ResNet features are (almost) equivalent." Cognitive Systems Research 59 (2020): 312-318.
12. Redmon J. et al. "You only look once: Unified, real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 779-788, 2016.
13. Zaidi, S.S., Ansari, M., Aslam, A. et al. (2022) A survey of modern deep learning based object detection models. Digital Signal Processing, 126, 103514. <https://doi.org/10.1016/j.dsp.2022.103514>
14. Jiang, P., Ergu, D., Liu, F. et al. (2022) A review of YOLO algorithm developments. Procedia Computer Sci., 199, 1066- 1073. <https://doi.org/10.1016/j.procs.2022.01.135>
15. Stelmakh N.V. Mastenko I.V. Classification and detection of defects in tubular products using machine vision. Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки. Том 35 (74) № 2 2024. с. 167-172. <https://doi.org/10.32782/2663-5941/2024.2/23>

<https://doi.org/10.31891/2219-9365-2024-79-16>

УДК 004.9

АХМЕТШИНА Людмила

Дніпровський Національний Університет ім. Олесь Гончара

<https://orcid.org/0000-0002-5802-0907>

e-mail: [akhmlu1@gmail.com](mailto:akhmlu1@gmail.com)

НЕСТЕРЕНКО Антон

Дніпровський Національний Університет ім. Олесь Гончара

e-mail: [an.nestere@gmail.com](mailto:an.nestere@gmail.com)

## АНАЛІЗ АЕРОФОТОЗНІМКІВ З ВИКОРИСТАННЯМ ЗГОРТКОВОЇ НЕЙРОННОЇ МЕРЕЖІ-АВТОКОДУВАЛЬНИКА

Метою цієї роботи було створення автономної системи, здатної автоматично виявляти аномалії на аерофотознімках, отриманих за допомогою безпілотних літальних апаратів. Аномальними у даному контексті вважаються об'єкти та сегменти зображення, які мають нетипові для даних зображень характеристики: яскравість, колір, текстуру або форму (наприклад рукотворні об'єкти та транспортні засоби у загородній місцевості, люди у обмежених зонах тощо). У дослідженні розглядаються інформативні можливості згорткових нейронних мереж (ЗНМ) для вирішення задачі виявлення аномальних об'єктів на аерофотознімках. Запропоновано структуру згорткового автокодувальника з трьома кодуєчими та трьома декодуєчими згортковими нейронними шарами, що забезпечує можливість навчання на немаркованих зображеннях і розпізнавання раніше невідомих типів аномалій. Крім того, ми розглядаємо вплив застосування алгоритмів попередньої обробки на швидкість і ефективність системи. Запропоновано метод локалізації потенційно аномальних сегментів за допомогою алгоритму знаходження контурів. Розроблена система складається з моделі згорткового автокодувальника, двох алгоритмів попередньої обробки, заснованих на виявленні порогів кольору та яскравості, алгоритму виявлення контурів і класифікатора, що зважає на вихідні дані НМ. Розроблена модель НМ була навчена на декількох наборах даних, що містять аерофотознімки без аномалій. Отримана система була протестована на реальних даних, результати цього тестування наведені у даній статті.

Ключові слова: машинне навчання, комп'ютерний зір, згорткова нейронна мережа, автокодувальник, аерофотозйомка.

AKHMETSHINA Ludmila, NESTERENKO Anton

Oles Honchar Dnipro National University

## ANALYSING AERIAL PHOTOGRAPHS USING A CONVOLUTIONAL AUTOENCODER NETWORK

The purpose of this work was to create a functioning autonomous system capable of automatically detecting anomalies in aerial images obtained using unmanned aerial vehicles. Anomalous, in this context, are objects and segments of the image that do not fit with the general picture of the investigated area, such as man-made objects and vehicles in the wilderness, people in restricted areas, etc. The research considers the informative capabilities of convolutional neural networks (CNN) for solving the task of detecting abnormal objects in photographs. The structure of convolutional autoencoder with three encoding and three decoding convolutional neural layers is proposed, providing the capability of learning on unlabeled images and recognizing previously unknown types of anomalies. In addition, we study the impact of applying preprocessing algorithms on the speed and effectiveness of the system. Such algorithms include detection of color and brightness threshold, as well as finding the contours of objects that stand out in the image. We also propose a method of localizing potentially abnormal segments using the combination of these algorithms. We show that this approach allows for a great increase in performance at the cost of a very marginal increase in the number of anomalies missed.

The resulting system consists of a convolutional autoencoder model, two preprocessing algorithms based on detecting thresholds of color and brightness, an algorithm for contour detection and a final classifier, weighting the outputs. The developed NN model was trained on several data sets containing aerial photography images with no anomalies. The resulting system was tested out on real-world data, and the results of this testing are provided here.

Keywords: machine learning, computer vision, convolutional neural network, autoencoder, aerial imagery.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Зображення, зняті з літальних апаратів, зазвичай містять складну картину різних типів земної поверхні, рослинності та рельєфу, а також можуть містити стандартні для місцевості рукотворні об'єкти. У багатьох практичних завданнях важливим етапом їх аналізу є виявлення об'єктів інтересу (аномалій), які мають нетипові для даних зображень характеристики, наприклад, яскравість, колір, текстуру або форму (плавальний засіб на воді, літак у небі, автомобіль на місцевості). Для автоматизації вирішення таких завдань дедалі більшого поширення набувають методи штучного інтелекту. Зокрема для аналізу складних зображень використовуються згорткові нейронні мережі (НМ), які здатні синтезувати багаторівневі карти ознак, що дозволяє ефективно вирішувати завдання класифікації візуальних об'єктів [1].

Однією з численних варіацій НМ є архітектура «автокодувальник», яка базується на реконструюванні вхідного зображення на підставі складеного в процесі навчання набору ознак. Класичні НМ-автокодувальники широко використовуються в дефектоскопії та в аналізі часових рядів [2, 3]. Вони показують високу ефективність у завданнях аналізу простих, однотипних даних, що характеризуються обмеженою кількістю складових структур з незначною мінливістю та варіативністю, у тому числі простих зображень (що не містять складні сцени). Однак, зі збільшенням складності зображень та за наявності на них безлічі різноманітних об'єктів, точність та швидкість їх роботи недостатні.

Ця робота присвячена розробці НМ-автокодувальника, архітектура якого включає згорткові нейронні шари, що забезпечують виявлення аномальних сегментів зображень з високою точністю. Окрім цього використовуються декілька алгоритмів первинного аналізу з метою підвищення швидкості роботи НМ.

### АНАЛІЗ ВІДОМИХ РІШЕНЬ І ПОСТАНОВКА ЗАДАЧІ ДОСЛІДЖЕННЯ

Класичні методи аналізу зображень засновані на виконанні комплексу операцій з метою поліпшення їх візуального сприйняття, або перетворення у форму, зручну для візуального або машинного аналізу. Наприклад, у роботі [4] наводиться порівняння ефективності модифікацій популярного алгоритму Reed-Xiaoli (RX), в основі якого лежить порівняльний аналіз вихідних характеристик яскравості пікселів зображення з усередненим значенням фону у вікні обробки. У статті [5] описується аналогічний підхід, у якому використовуються матриці Кірхгофа для порівняння характеристик сусідніх пікселів.

Нині у напрямках "Комп'ютерний зір" і "Обчислювальний інтелект" з'явилась велика кількість наукових праць, присвячених процесу комп'ютерного отримання інформації (синтезу описів, зразків) із зображень. Наприклад, у галузі машинного навчання застосування НМ-автокодувальників є поширеним підходом для вирішення задач пошуку аномалій у галузях виявлення дефектів та аналізу медичних знімків завдяки можливості їх навчання методом «без вчителя» [6, 7]. З метою виключення надмірності та шумових складових навчальних вихідних даних, використовуються змагальні моделі автокодувальників або генеративні нейронні мережі [8, 9, 10]. У свою чергу, для вирішення завдань класифікації фото-даних активно використовуються «глибокі» згорткові НМ з великою кількістю шарів.

Основною для сучасних архітектур згорткових НМ стала архітектура VGG16, яка була представлена в 2015 році і включала 16 згорткових шарів [11].

На відміну від завдань класифікації об'єктів, задача виявлення аномалій ускладнюється тим, що на етапі створення та навчання моделі, як правило, доступні лише "базові зображення" – тобто такі, що не містять аномалій. Крім цього, «аномаліями», залежно від сфери застосування, можуть вважатися найрізноманітніші об'єкти та явища, які неможливо врахувати під час навчання повною мірою. Виходячи з цього система, яка здійснює пошук аномальних ділянок на зображеннях, повинна мати наступні характеристики:

1. можливість навчання з використанням лише "базових зображень";
2. здатність кваліфікувати об'єкти інтересу, які не зустрічаються в навчальному наборі даних як аномальні;
3. відсутність чи мінімізація кількості хибних спрацьовувань під час роботи з даними, схожими, але не ідентичними навчальним;
4. висока швидкодія, що забезпечує роботу системи у реальному часі.

### МЕТА СТАТТІ

Метою роботи є створення аналітичної системи для виявлення аномалій на фотознімках, знятих з літальних апаратів на основі НМ-автокодувальника, яка, за рахунок використання згорткових нейронних шарів та алгоритмів первинного виявлення потенційно аномальних сегментів дозволяє швидко виявляти складні багаторівневі ознаки об'єктів інтересу.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

У рамках розв'язання задачі аналізу зображень з метою пошуку аномальних об'єктів інтересу була розроблена модель НМ, що являє собою НМ-автокодувальник із шарами згортки. Автокодувальники, як правило, складаються з вхідного та вихідного шарів однакової розмірності, а також набору прихованих шарів, що знаходяться між ними. Для вирішення більшості завдань вони проектуються з використанням повністю пов'язаних шарів, однак у запропонованій архітектурі кожен шар автокодувальника складається з шару згортки та об'єднання.

Основними механізмами роботи згорткових нейронних мереж є процедури згортки та об'єднання. Операція згортки (рис. 1) полягає в покроковому застосуванні операції скалярного добутку до сегментів зображення з покроковим зсувом. У якості множника використовується так званий «фільтр» - невелика матриця чисел фіксованого розміру. Результатом операції згортки є набір матриць, які називають «картами ознак».

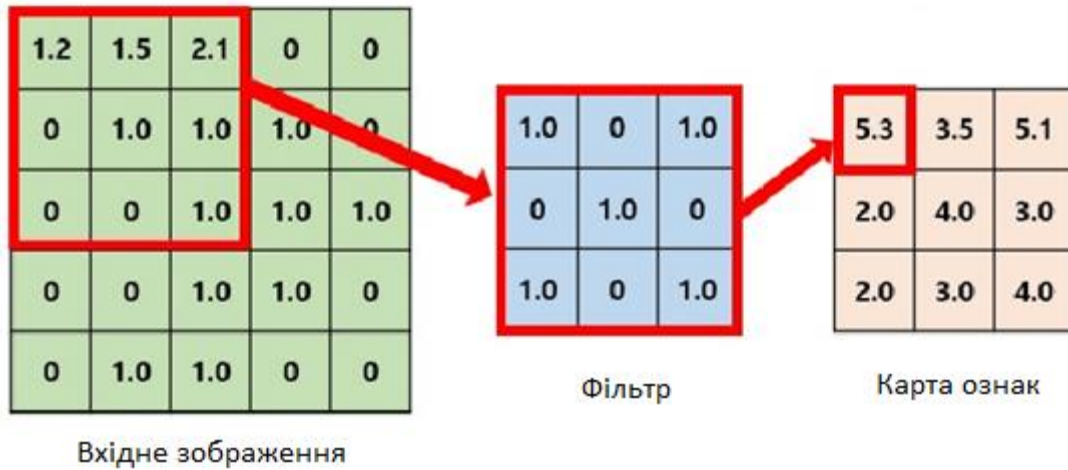


Рис. 1. Операція згортки

Таким чином, для вхідного зображення  $f$  та фільтра  $h$  значення карти ознак можна обчислити за формулою:

$$G[m, n] = \sum_j^s \sum_k^s h[j, k] f[m - j, n - k], \quad (1)$$

де  $m$  і  $n$  – індекси рядів та стовпців результуючої матриці, а  $s$  – розмір фільтра. Після формування карт ознак кожне значення модифікується за допомогою активаційної функції. У згорткових НМ широко використовується функція ReLU (лінійний випрямляч):

$$G_1(x) = \max(0, G), \quad (2)$$

Операція об'єднання проводиться для зниження розмірності карт ознак і полягає у зведенні кількох значень у вхідній матриці одного значення. Найбільш поширеним є використання максимального значення у вікні обробки

$$G_2[m, n] = \max_{s \times s} (G_2), \quad (3)$$

однак є інші варіації, наприклад, знаходження середнього значення.

У класичній архітектурі згорткової НМ за кожним шаром згортки слід шар об'єднання, а потім додаються кілька пов'язаних повністю шарів, які визначають належність вхідного зображення до одного з відомих класів. Цей підхід забезпечує отримання багаторівневих ознак, які представлені у вигляді ієрархічних матриць. На відміну від цього, НМ-автокодувальник включає дві частини, в першій приховані шари (кодувальник) мають розмірність, що послідовно зменшується, а в другій - збільшується (декодувальник). У процесі його навчання на кожному шарі формуються карти ознак, властиві вхідним даним. На рис. 2 показаний приклад карт ознак, які витягуються кожному з шарів під час аналізу зображення.

Шари декодувальника пов'язані з шарами кодувальника і дублюють карти їх ознак для використання в процесі декодування. Коли модель навчається реконструкції вхідного зображення, у якості цільової функції використовується мінімізація середньоквадратичного відхилення (MSE).

$$L = \frac{1}{nm} \sum_j^n \sum_j^m (I'_{ij} - I_{ij})^2 \quad (4)$$

де  $I$  – вхідне зображення, а  $I'$  – результуюче.

У розробленій архітектурі використовується три згорткові шари, для кодування зображення і ще три – для декодування. Кожен шар меншої розмірності має пропорційно більше згорткових фільтрів. Використання такої архітектури дозволяє автокодувальнику отримувати кілька рівнів ознак, які дозволяють більш детально кодувати реальні зображення. Схематично згорткова НМ-автокодувальник представлена на рис. 3.

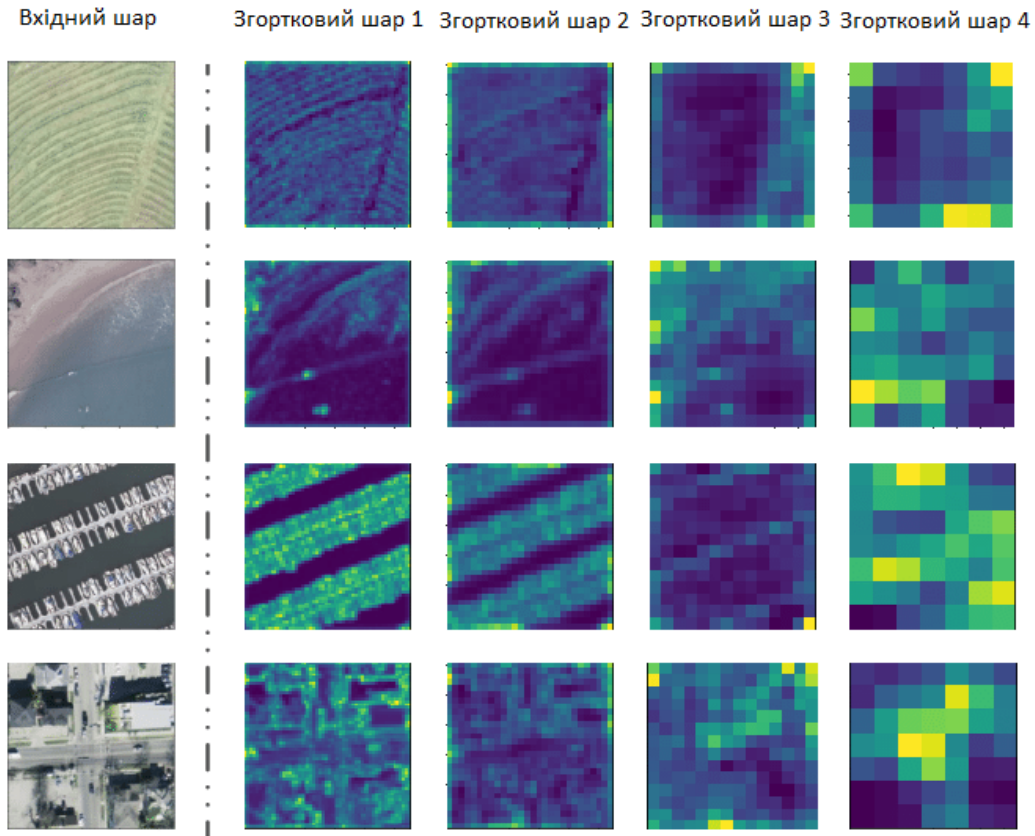


Рис. 2. Візуалізація прихованих шарів згорткової НМ

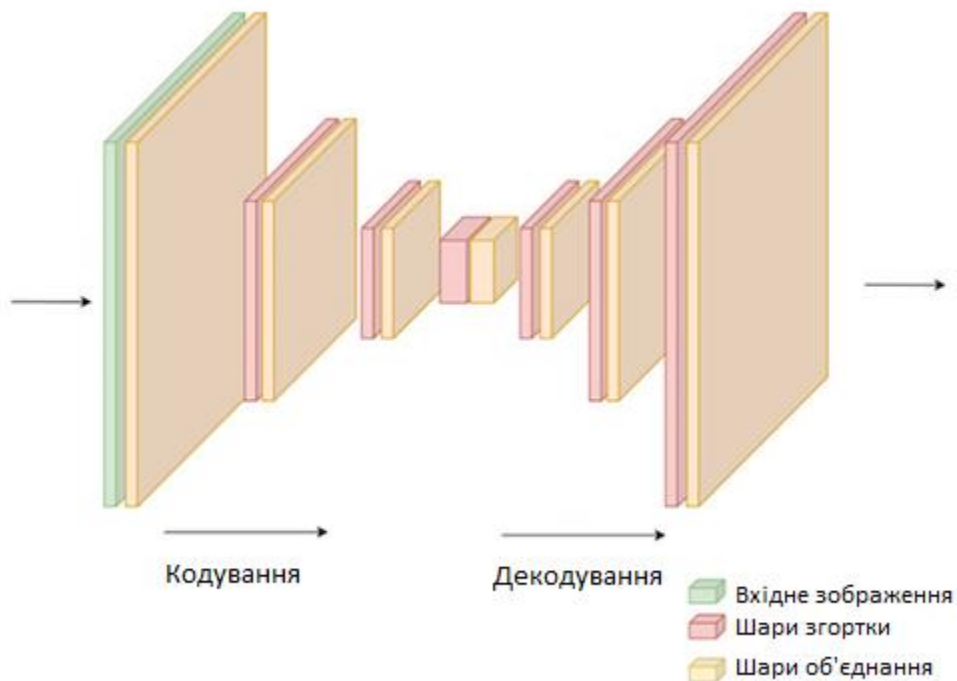


Рис. 3. Схематичне представлення архітектури згорткового НМ-автокодувальника

Для тренування розробленого НМ-автокодувальника використовується набір базових зображень без аномалій, які розбиваються на сегменти. Як метод оптимізації використовувалася варіація методу



стохастичного градієнтного спуску, що описаний у роботі [12]. Далі, до навченої моделі передаються нові зображення. Ті з них, які моделі вдалося реконструювати найменш точно – вважаються аномальними. Для оцінки точності реконструкції використовується функція середньоквадратичного відхилення (MSE), її поріг визначається експериментально.

На додачу до НМ, з метою підвищення швидкодії системи, розроблено алгоритми первинної обробки зображень та визначення на ньому потенційно аномальних ділянок. Цими алгоритмами є аналіз порогів кольору та яскравості, а також пошук закритих контурів. Аналіз перепадів кольору та яскравості виконується за допомогою швидкого та ефективного алгоритму знаходження оператора Лапласа для матриці сусідніх пікселів, що описано у роботі [13].

Наступним етапом аналізу зображення є застосування алгоритму пошуку контурів для попереднього виявлення кластерів пікселів, які можуть являти собою об'єкти, що виділяються. Як вхідні дані цьому алгоритму передаються три варіанти зображень: вихідне, а також два результуючих зображення після фільтрації двома попередніми методами. Алгоритм пошуку контурів реалізує метод, описаний у публікації [14]. В результаті роботи алгоритм і визначає координати виявлених кластерів пікселів, що виділяються, розмір яких перевищує задане порогове значення. Такі кластери потенційно можуть бути об'єктами інтересу. Ці кластери вписуються у квадрати фіксованого розміру і підлягають подальшому аналізу за допомогою нейронної мережі. На рис. 4 показаний результат роботи алгоритму пошуку контурів, в результаті якої були виявлені дві ділянки, що підлягають подальшому аналізу, з яких одна дійсно містить аномалію (надувний човен).

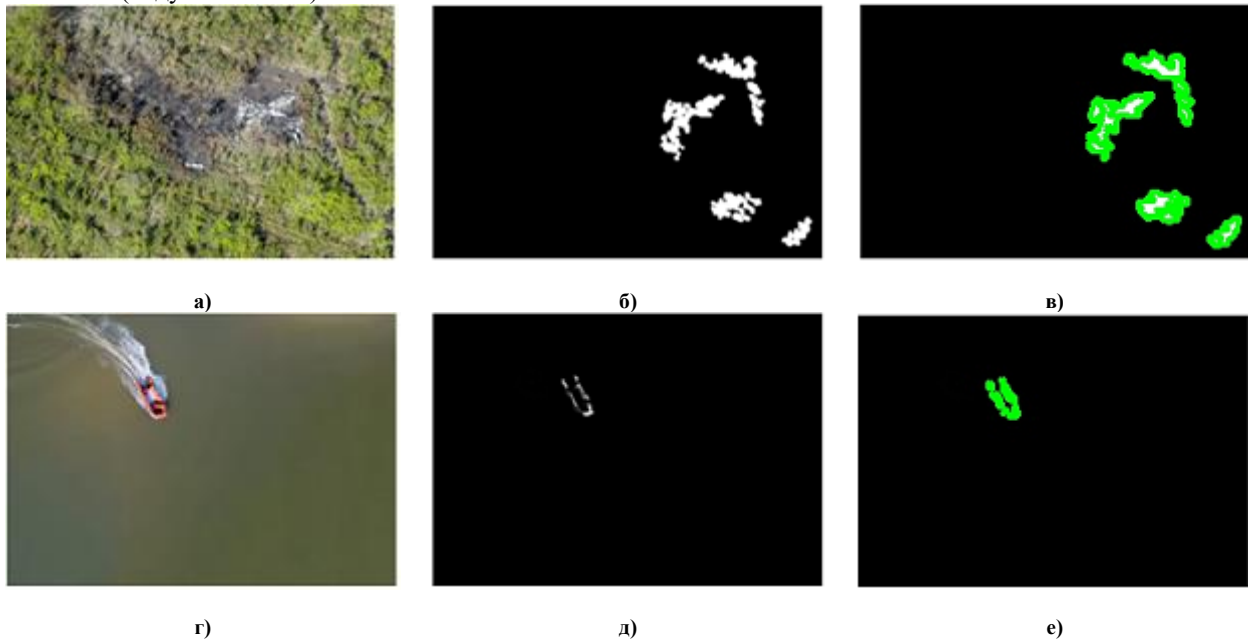


Рис. 4. Приклад роботи алгоритмів первинної обробки. а), г) – оригінальні зображення. б) д) – порогові алгоритми. в) е) – знайдені контури

На рис. 4 представлена загальна схема розробленого програмного забезпечення для аналізу зображень з використанням НМ-автокодувальника із згортковими шарами. Аналітична система реалізована мовою програмування Python із використанням бібліотек OpenCV, Scikit-image, Tensorflow та Keras.

У якості вихідних даних для тренування та тестування були використані кольорові RGB-зображення аерофотозйомки, що містять типове для помірної кліматичної зони місцевість, схожі за типом, але взяті з різних джерел та різних географічних локацій.

Оскільки в більшості сфер потенційного застосування такої системи аномальні ділянки зустрічаються не часто, передбачено використання алгоритмів попередньої локалізації потенційно аномальних ділянок зображення на основі виявлення кольорних і яскравих перепадів, а також пошуку закритих контурів. Лише ті ділянки зображень, які є потенційно аномальними, передаються в нейронну мережу для аналізу. У процесі навчання вихідні дані поділяються на набір навчання, набір валідації та набір тестування, але лише набір тестування містив аномалії. Результатом навчання є вагова модель нейронної мережі, яка надалі використовується для виявлення аномалій у нових даних.

У розробці та тестуванні системи був використаний набір даних із 192 зображення високої роздільної здатності, отриманих у різних географічних локаціях. З усього набору 28 зображень містили аномальні об'єкти, 164 - не містили. Оскільки деякі зображення містили більше одного аномального об'єкта, загальна кількість аномальних об'єктів у тестовому наборі даних становила 36. При тестуванні аномальними

вважалися рукотворні об'єкти нестандартної форми, автомобілі, плавзасоби на водній поверхні тощо. Будинки стандартної форми аномальними об'єктами не вважалися.

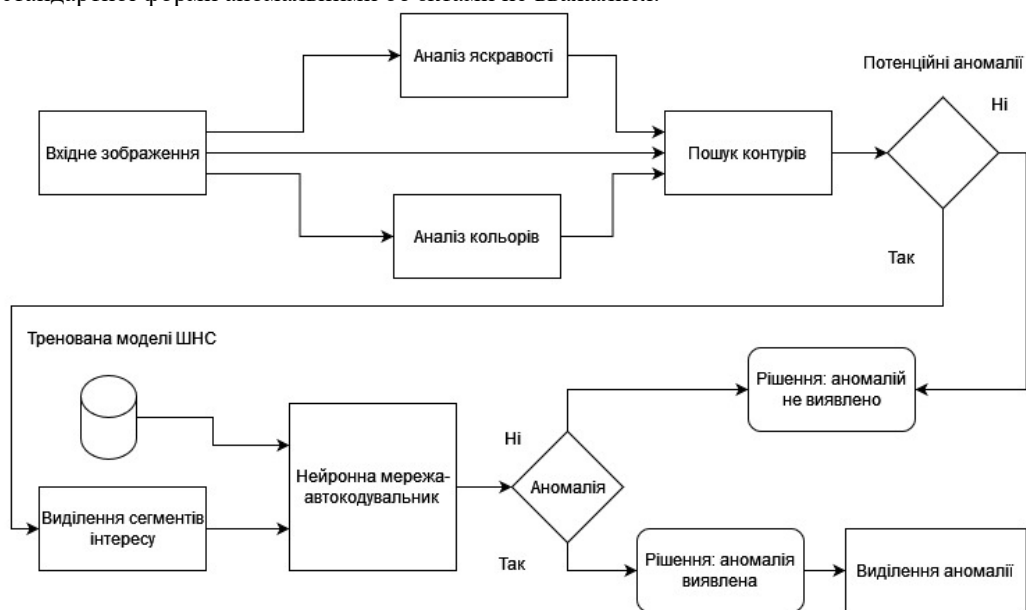


Рис. 5. Схема роботи аналітичної системи

Експериментально було встановлено, що з урахуванням дозволу та деталізації зображень при аналізі доцільно використовувати сегменти даних розміром 128x128 пікселів. Тоді, приховані шари, мають розмірність 64x64, 32x32 і 16x16 відповідно. На першому та другому рівнях шарів використовувалися фільтри, розмірністю 5x5, на третьому та четвертому – 3x3. Операція об'єднання здійснюється за допомогою вікна зсуву 2x2 на кожному із шарів.

Для навчання було використано 100 зображень різних розмірів, розбиті на фрагменти по 128x128 пікселів. Загальна кількість фрагментів склала 46875. Навчання автокодувальника проводилося протягом 40 epoch. На рис. 6 наведено графік сходження функції втрати під час навчання.

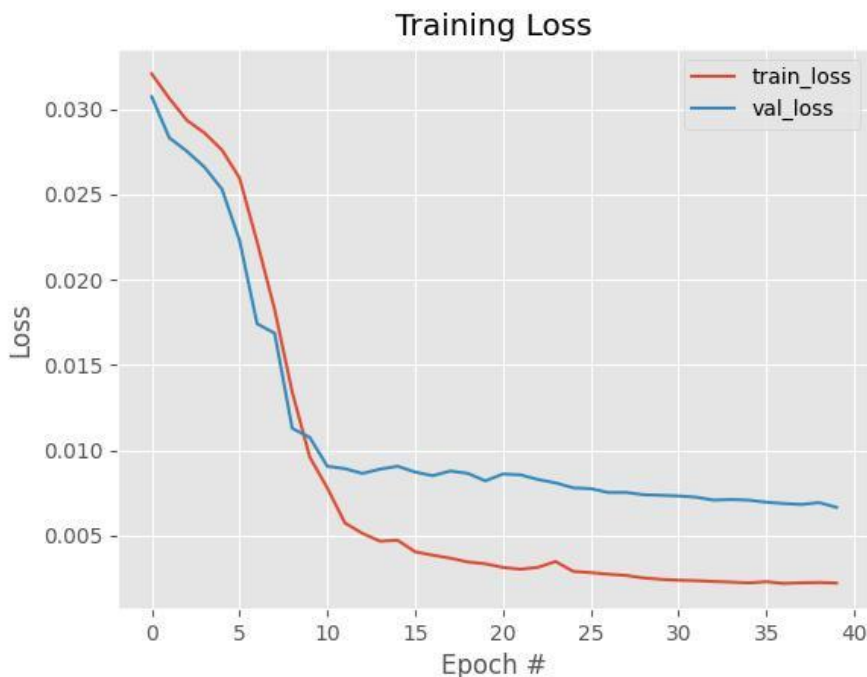


Рис. 6. Графік сходження функції втрати (тренувальний та валідаційний набори даних)

Після навчання моделі, 28 зображень, що містять аномалії, а також 64 зображення, що не містять їх (загалом 92) було використано для тестування системи. За результатами тестування було досягнуто

наступних результатів. З 36 аномальних об'єктів було успішно розпізнано 31, що становить 86% від загальної кількості аномалій. Хибно-позитивні спрацьовування (кількість сегментів зображень, некоректно виділених як аномальні) склали 15 сегментів. Середній час обробки зображень становив 3.2 с, максимальне – ~8 с. На рис. 7 наведено приклад успішно виявленої аномалії (надувний човен) на зображенні, що містить берег річки. Лише сегмент, що містить човен, перевищив поріг функції середньоквадратичного відхилення, що дорівнює 0.05.

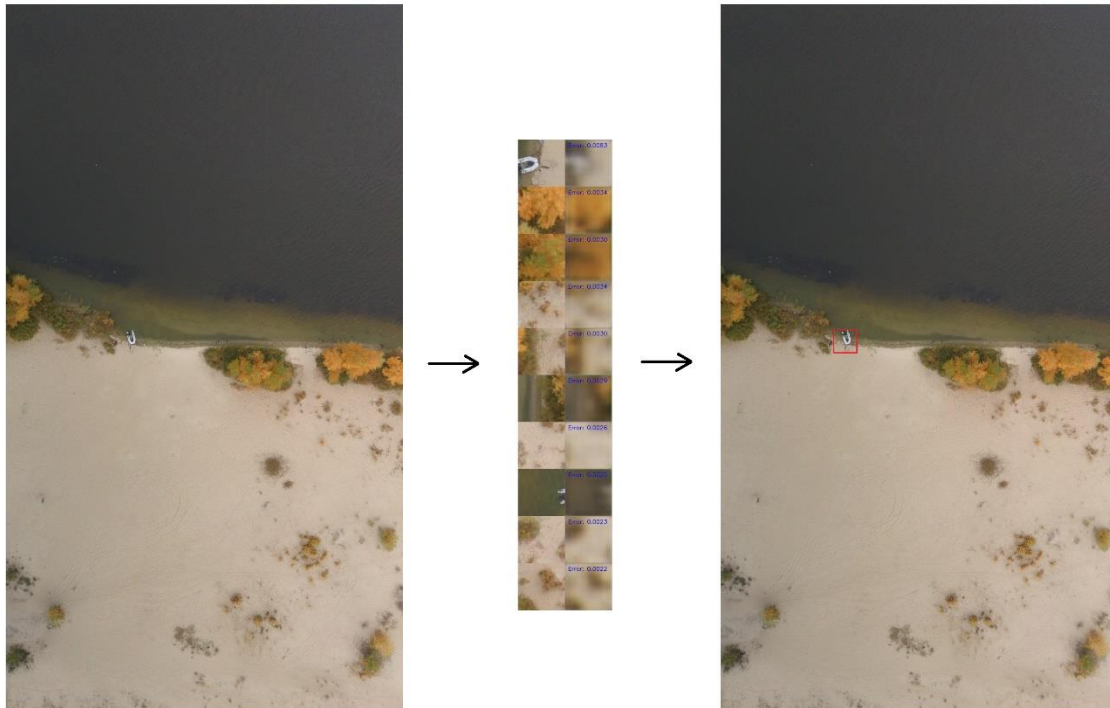


Рис. 7. Приклад успішно розпізнаного аномального сегмента

З метою порівняння ефективності розробленої системи було зроблено порівняння швидкості її роботи з методом Single-Class SVM (Support Vector Machine), а також із використанням автокодувальника без попереднього відсікання зображень. Метод Single-Class SVM продемонстрував надзвичайно високий ступінь хибно-позитивних спрацьовувань зважаючи на його нездатність адаптуватися до комплексних елементів ландшафту, представлених на зображеннях. На 92 тестових зображеннях 326 сегментів було розпізнано як аномальні, при цьому майже 30% справжніх аномалій були пропущені.

Автокодувальник без попереднього аналізу зображень більш простими алгоритмами показав незначний приріст успішно розпізнаних аномалій за рахунок значного зниження швидкодії та підвищення хибно-позитивних спрацьовувань. Це порівняння продемонструвало, що запропонований підхід дозволяє досягти найменшого значення хибно-позитивних спрацьовувань за практично повної відсутності зростання пропущених аномалій. Швидкість роботи системи знаходиться на прийнятному рівні. Візуалізація результатів показана рис. 8.

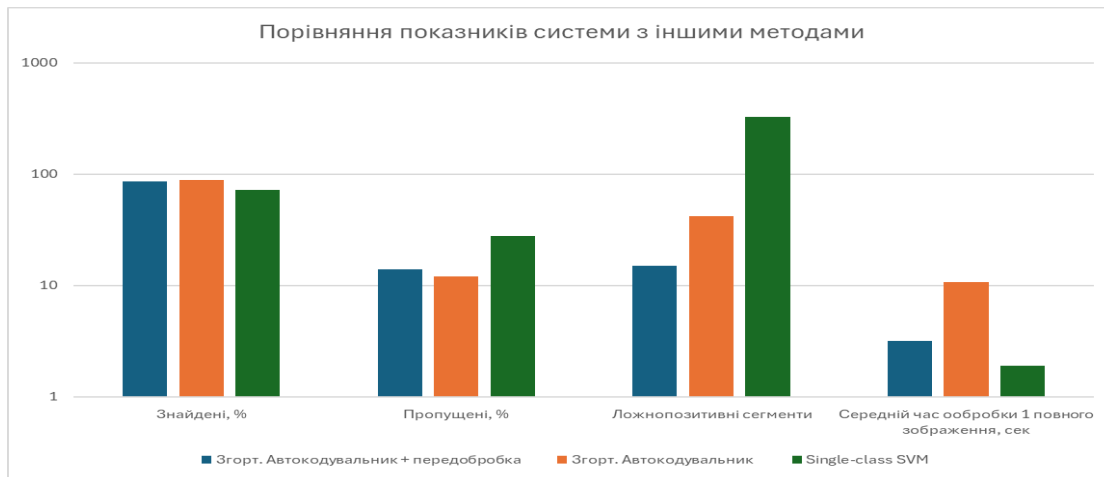


Рис. 8. Порівняння результатів різних підходів (логарифмічна шкала)

## ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШОГО РОЗВИТКУ У ДАНОМУ НАПРЯМІ

Використання в архітектурі НМ-автокодувальника із згортковими шарами забезпечує принципову можливість виявлення аномалій на комплексних, насичених різноманітними об'єктами зображеннях з достатньою для прикладних завдань ефективністю. У порівнянні з іншими методами, не заснованими на використанні згорткових НМ, такий підхід демонструє значно меншу кількість хибних спрацьовувань і високу точність виявлення, що говорить про більш глибоке уявлення системою характеристик досліджуваної місцевості.

Одночасно з цим комбінування НМ-автокодувальника з алгоритмами попереднього аналізу дозволяє значно підвищити швидкодію системи.

Тим не менш, зберігається певна кількість помилкових спрацьовувань, що обумовлено складністю зображень та різноманітністю таких рис, як крони дерев або кам'янистої місцевості. Для вирішення практичних завдань з високою достовірністю доцільним є створення кількох попередньо вивчених моделей та декількох шаблонів налаштувань для алгоритмів попереднього аналізу під кожен тип задачі.

## References

- [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, *Commun. ACM* 60, 6 (June 2017), p 84–90. doi:10.1145/3065386
- [2] A. Mujeeb, et al, One Class Based Feature Learning Approach for Defect Detection Using Deep Autoencoders, *Advanced Engineering Informatics*, vol. 42, Oct. 2019. doi:10.1016/j.aei.2019.100933
- [3] J. Pereira and M. Silveira, Unsupervised Anomaly Detection in Energy Time Series Data Using Variational Recurrent Autoencoders with Attention, 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 2018, pp. 1275-1282. doi:10.1109/ICMLA.2018.00207
- [4] S. Küçük and S. E. Yüksel, Comparison of RX-based anomaly detectors on synthetic and real hyperspectral data, 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2015, pp. 1-4. doi:10.1109/WHISPERS.2015.8075504
- [5] A. Legrand, H. Trannois and A. Courmier, Use of Uncertainty with Autoencoder Neural Networks for Anomaly Detection, 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), 2019, pp. 32-35. doi:10.1109/AIKE.2019.00014
- [6] C. Baur, et al. Autoencoders for Unsupervised Anomaly Segmentation in Brain MR Images: A Comparative Study, *Medical Image Analysis*, vol. 69, Apr. 2021, doi:10.1016/j.media.2020.101952
- [7] J. Chow, et al. Anomaly detection of defects on concrete structures with the convolutional autoencoder, *Advanced Engineering Informatics*, vol 45. 2020. doi:10.1016/j.aei.2020.101105.
- [8] C. Zhou, R. C. Paffenroth, Anomaly Detection with Robust Deep Autoencoders, *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining August 2017* pp. 665–674. doi:10.1145/3097983.3098052
- [9] L. Beggel, M. Pfeiffer, B. Bischl, Robust Anomaly Detection in Images Using Adversarial Autoencoders, *Machine Learning and Knowledge Discovery in Databases*, pp. 206–222. doi:10.1007/978-3-030-46150-8\_13
- [10] A. Marco, et al, Generative Adversarial Networks for anomaly detection in aerial images, *Computers and Electrical Engineering*, Volume 106, March 2023. doi:10.1016/j.compeleceng.2022.108470.
- [11] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, *Arxiv pre-print*, 2014. doi:10.48550/arXiv.1409.1556
- [12] L. Bottou, Large-Scale Machine Learning with Stochastic Gradient Descent, *Proceedings of COMPSTAT'2010*, 2010. doi:10.1007/978-3-7908-2604-3\_16
- [13] F. Verdoja, M. Grangetto, Graph Laplacian for Image Anomaly Detection, *Machine Vision and Applications*, vol. 31, no. 1, Feb. 2020. doi:10.1007/s00138-020-01059-4
- [14] Suzuki, S. Abe, K., Topological Structural Analysis of Digitized Binary Images by Border Following. *CVGIP* 30 1, pp 32-46 (1985)

<https://doi.org/10.31891/2219-9365-2024-79-17>

УДК 658

ПАВЛЮК Віталій

Національний університет «Львівська політехніка»

<https://orcid.org/0009-0009-3237-8108>

[vitalii.pavliuk.mknuo.2023@lpnu.ua](mailto:vitalii.pavliuk.mknuo.2023@lpnu.ua)

НАКОНЕЧНИЙ Адриан

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0002-1873-6337>

[adrian.y.nakonechnyi@lpnu.ua](mailto:adrian.y.nakonechnyi@lpnu.ua)

## ІНТЕГРОВАНА АВТОМОБІЛЬНА СИСТЕМА КОМУНІКАЦІЇ ДЛЯ ПОКРАЩЕННЯ БЕЗПЕКИ ТА ЕФЕКТИВНОСТІ ДОРОЖНЬОГО РУХУ

У даній роботі досліджується проблема створення надійних та ефективних комунікаційних протоколів для інтегрованих систем V2X (Vehicle-to-Everything), які сприяють підвищенню безпеки та ефективності дорожнього руху. Актуальність дослідження зумовлена стрімким розвитком транспортної галузі та впровадженням V2X технологій, що забезпечують бездротовий обмін даними між транспортними засобами та інфраструктурою. У ході роботи проаналізовано наявні системи V2X, розглянуто моделі обміну повідомленнями та розроблено комбінований протокол, що поєднує переваги технологій IEEE 802.11p та LTE-V2X. Запропонований підхід дозволяє досягти високої продуктивності, надійності та безпеки в умовах різних сценаріїв використання, забезпечуючи низьку затримку передачі даних та цілісність повідомлень. Особлива увага приділяється розробленню механізмів пріоритетності повідомлень, їхньої цілісності та автентифікації на основі інфраструктури публічних ключів (PKI).

Ключові слова: Vehicle-to-Everything, IEEE 802.11p, LTE-V2X, протокол V2X, V2X комунікація.

PAVLIUK Vitalii, NAKONECHNYI Adrian

Lviv Polytechnic National University

## INTEGRATED VEHICLE COMMUNICATION SYSTEM FOR ENHANCING ROAD SAFETY AND EFFICIENCY

Currently, the transportation industry is undergoing significant technical transformations thanks to the introduction of the new Vehicle-to-Everything (V2X) technology. This technology provides direct wireless data exchange between vehicles, road infrastructure and other road users. Systems built according to this technology are able to significantly increase the safety, efficiency and convenience of road traffic, helping to reduce the number of accidents, optimize traffic and support autonomous driving. In view of this, the relevance of the presented research is determined by the need to develop reliable and effective communication protocols for V2X, which are able to meet modern requirements for speed, reliability and security of data transmission.

This paper addresses the development of reliable and efficient communication protocols for integrated V2X (Vehicle-to-Everything) systems, aimed at enhancing road safety and efficiency. The relevance of this research is driven by the rapid development of the transportation sector and the implementation of V2X technologies, which enable wireless data exchange between vehicles and infrastructure. The study analyzes existing V2X systems, examines messaging models, and develops a combined protocol that leverages the advantages of IEEE 802.11p and LTE-V2X technologies. The proposed approach achieves high performance, reliability, and safety under various usage scenarios, ensuring low data transmission latency and message integrity. Special attention is given to the development of mechanisms for message prioritization, integrity, and authentication based on Public Key Infrastructure (PKI).

Key words: Vehicle-to-Everything, IEEE 802.11p, LTE-V2X, V2X protocol, V2X communication.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

На даний час транспортна галузь зазнає значних технічних перетворень завдяки впровадженню нової технології Vehicle-to-Everything (V2X). Ця технологія забезпечує прямий бездротовий обмін даними між транспортними засобами, дорожньою інфраструктурою та іншими учасниками дорожнього руху. Системи побудовані за даною технологією здатні значно підвищувати безпеку, ефективність та зручність дорожнього руху, сприяючи зменшенню кількості аварій, оптимізації трафіку та підтримці автономного водіння. З огляду на це актуальність представленого дослідження зумовлена необхідністю розроблення надійних і ефективних протоколів комунікації для V2X, які здатні задовольнити сучасні вимоги до швидкості, надійності та безпеки передачі даних.

## ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою дослідження є розроблення протоколу інтегрованої системи V2X комунікації для забезпечення високої продуктивності, надійності та безпеки руху в різних сценаріях використання. Для досягнення поставленої мети в роботі розглядаються наступні питання.

1. Аналіз наявних систем V2X та їхні характеристики.

2. Розгляд основних моделей обміну повідомленнями та протоколів керування доступом до середовища (MAC) у V2X.

3. Розроблення протоколу комунікації для V2X систем на основі вибору протоколу керування доступом до середовища та протоколу транспортного рівня.

4. Реалізація заходів забезпечення цілісності та автентифікації повідомлень у V2X системах.

Таким чином дане дослідження полягає в розробленні комбінованого протоколу комунікації для V2X систем, який би поєднував переваги протоколів IEEE 802.11p та LTE-V2X. Запропонований в роботі підхід дозволяє забезпечити високу швидкість передачі даних, надійність та низьку затримку, що є критичним для застосувань V2X технології. Крім того, в статті розглядається можливість розроблення механізму забезпечення пріоритетів повідомлень та їхньої цілісності й автентифікації за допомогою інфраструктури публічних ключів (PKI).

## ОГЛЯД НАЯВНИХ РІШЕНЬ ПОСТАВЛЕНОЇ ЗАДАЧІ

### Загальний опис систем V2X

Транспортна галузь зазнає значних перетворень із появою технології зв'язку "Vehicle-to-Everything" (V2X). V2X виходить за рамки традиційних систем, орієнтованих на водія, сприяючи створенню спільного середовища на дорозі завдяки прямому бездротовому обміну даними між транспортними засобами, дорожньою інфраструктурою та навіть пішоходами. Така взаємопов'язана екосистема має великий потенціал революціонізувати транспорт, зробивши його безпечнішим, плавнішим та ефективнішим для всіх [19].

Основні компоненти системи V2X складають.

- Бортові пристрої (on-board units - OBU): встановлені в транспортних засобах і пішоходів, OBU служать центрами комунікації. Вони оснащені радіомодулями, давачами (такими як GPS, камери, радары), антенами та процесорами для полегшення обміну даними [16].

- Придорожні пристрої (roadside units - RSU): розгорнуті вздовж доріг та інфраструктури, RSU збирають та передають інформацію про дорожній рух. Вони діють як ретранслятори інформації в мережі V2X, а також служать як окремі учасники системи [1, 17].

- Центр управління (необов'язковий): керує всією мережею V2X, керує потоком даних, надає послуги (наприклад, оновлення дорожнього руху) та оптимізує продуктивність мережі.

Зв'язок у системах V2X здійснюється [13].

- Протоколами зв'язку V2X, які регулюють взаємодію між OBU та RSU, забезпечуючи безперебійний та надійний обмін даними.

- Обміном даними, який охоплює інформацію в реальному часі, критичну для безпеки та управління дорожнім рухом, включаючи:

- Положення, швидкість і напрямок руху транспортного засобу
- Дані датчиків з камер і радарів
- Стан світлофора
- Наявність пішоходів

Зазначена схема зв'язку схематично зображена на рис. 1.

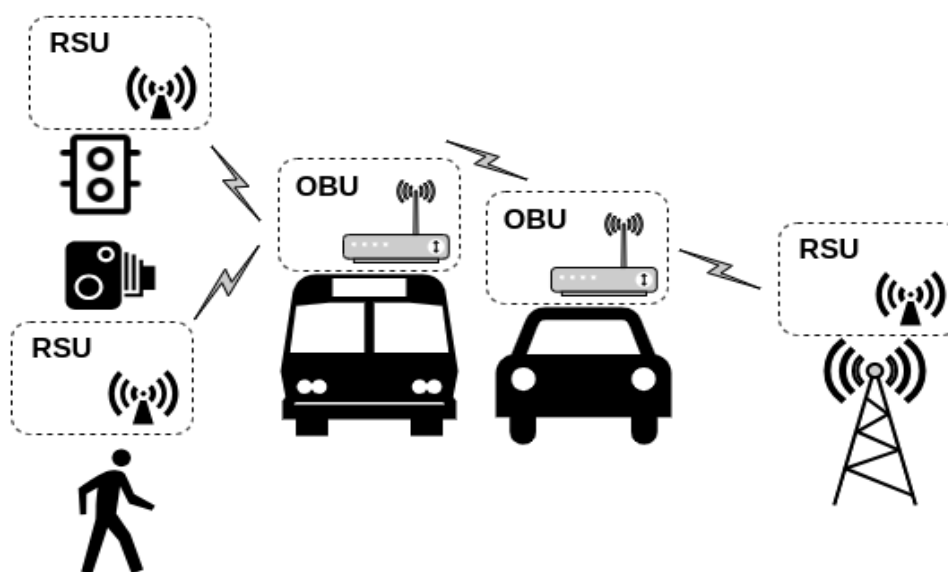


Рис. 1. Структурна схема взаємодії бортових пристроїв (OBU) та придорожніх пристроїв (RSU) у системі V2X

V2X дозволяє використовувати широкий спектр застосувань, які змінюють спосіб нашого пересування дорогами [19] і дозволяють реалізувати:

- Запобігання зіткненням: попереджає водіїв про потенційні небезпеки, дозволяючи їм вживати ухильних дій і запобігати аваріям.
- Спільне маневрування: сприяє узгодженим діям, таким як зміна смуги руху та злиття, покращуючи потік руху.
- Оптимізацію світлофорів: транспортні засоби можуть регулювати швидкість залежно від майбутніх світлофорів, зменшуючи затори.
- Безпеку вразливих учасників дорожнього руху: підвищує безпеку пішоходів і велосипедистів за допомогою оповіщень і попереджень у реальному часі.
- Підтримку автоматизованого водіння: надає дані в реальному часі для того, щоб автономні транспортні засоби могли безпечно та ефективно рухатися.

#### Огляд наявних систем, їх характеристики, переваги та недоліки

Системи типу V2X стають все більш різноманітними і в них використовуються різні протоколи зв'язку та конфігурації інфраструктури. На цей час, найбільш поширеними є наступні системи [4].

- Dedicated Short-Range Communication (DSRC). Дана система використовує стандарт IEEE 802.11p для короткострокового зв'язку (зазвичай до 1000 метрів) між транспортними засобами та інфраструктурою. DSRC широко використовувалася в пілотних проектах, особливо у Сполучених Штатах. Однак її майбутнє залишається невизначеним через появу нових технологій на основі мобільних мереж і наступного покоління з більш широкими можливостями та потенційно нижчими витратами.
- Cellular Vehicle-to-Everything (C-V2X). Система використовує наявні мобільні мережі (LTE) для полегшення зв'язку V2X. Система C-V2X пропонує ширше покриття порівняно з DSRC, охоплюючи райони з обмеженою спеціалізованою інфраструктурою, однак, у C-V2X можливі проблеми з затримкою в перевантажених мобільних мережах, що може вплинути на роботу в режимі реального часу.
- 5G NR V2X. Завдяки використанню технології нового покоління система 5G New Radio (NR), 5G NR V2X обіцяє значні переваги. Вона пропонує надзвичайно малу затримку (що критично важливо для застосунків у режимі реального часу), можливість масового підключення (що підтримує більшу кількість пристроїв) та підвищену надійність. 5G NR V2X має великий потенціал для майбутніх застосунків V2X, особливо для автономних транспортних засобів. Порівняльні характеристики наведених систем наведені в таблиці 1.

Таблиця 1

	DSRC	C-V2X	5G NR V2X
Архітектура	IEEE 802.11p, короткі відстані (до 1 км)	LTE-V2X, Побудована на LTE, підтримка прямих і мережних режимів	Використання 5G для високошвидкісного зв'язку
Технічні характеристики	Частоти 5.9 ГГц, швидкість до 27 Мбіт/с	Частоти 1.8-2.6 ГГц, висока пропускна здатність	Частоти до 52 ГГц, швидкість до кількох Гбіт/с, наднизька затримка
Переваги	Низька затримка, висока надійність	Великий радіус дії, використання наявної інфраструктури LTE	Висока пропускна здатність, підтримка автономного водіння
Недоліки	Обмежений радіус дії, менша пропускна здатність	Вища затримка, можливі проблеми з перевантаженням мережі	Висока вартість впровадження, значні інвестиції в інфраструктуру

Попри свої переваги, наявні системи V2X мають також і ряд недоліків, які необхідно вирішувати при широкому впровадженні.

- Висока вартість: впровадження та підтримка інфраструктури V2X у великих масштабах може бути дорогою. Тут враховується вартість бортових блоків для транспортних засобів, придорожніх блоків та інфраструктури мережі зв'язку (особливо для DSRC).
- Стандартизація: різні протоколи зв'язку та фрагментовані стандарти між DSRC, C-V2X та 5G NR V2X можуть перешкоджати взаємодії між системами.
- Кібербезпека: системи V2X вразливі до кібератак, що викликає занепокоєння щодо безпеки даних і потенційного маніпулювання. Необхідні надійні заходи кібербезпеки, щоб забезпечити цілісність та правдивість даних V2X.

- Конфіденційність: збір та використання даних V2X викликають занепокоєння щодо конфіденційності, особливо щодо поведінки водія та інформації про місцезнаходження. Для розв'язання цих проблем та підвищення довіри до технології V2X необхідні чіткі правила та механізми контролю з боку користувачів.

З наведеного можна зробити висновки, що технологія V2X має потенціал «революціонізувати» транспорт, роблячи його безпечнішим, плавнішим та ефективнішим. Наявні різні види системи V2X, такі як DSRC, C-V2X та 5G NR V2X, пропонують унікальні характеристики та переваги, однак перед широким впровадженням V2X необхідно розв'язувати ряд ключових задач, або врахувати їх при розробці системи. До них в першу чергу потрібно віднести високу вартість, можливість стандартизації, кібербезпеку та конфіденційність. Для успішного розроблення V2X систем, важливо ретельно вивчити характеристики різних архітектур, використати їх переваги та зосередитися на розв'язанні вказаних задач.

## АНАЛІЗ ОСНОВНИХ МОДЕЛЕЙ ТА АЛГОРИТМІВ СИСТЕМИ V2X

### Моделі обміну повідомленнями

Системи V2X використовують різні моделі обміну повідомленнями для забезпечення зв'язку та обміну інформацією між різними учасниками дорожнього руху. До основних моделей можна віднести [14].

- Автомобіль-Інфраструктура (Vehicle to Infrastructure - V2I). Модель передбачає обмін повідомленнями між бортовим пристроєм (OBU) у транспортному засобі та придорожнім пристроєм (RSU) або локальним сервером додатків. RSU зазвичай встановлюються вздовж доріг та служать точками доступу до мережі V2X. Вони можуть збирати дані про дорожній рух, передавати попередження про небезпеку та надавати інші послуги

- Автомобіль-Мережа (Vehicle to Network - V2N). Дана модель передбачає зв'язок OBU з віддаленим сервером додатків через мобільну мережу. Віддалені сервери додатків можуть надавати широкий спектр послуг V2X, таких як оновлення дорожнього руху, прогнозування трафіку та послуги з підтримки водіння.

- Автомобіль-Автомобіль (Vehicle to Vehicle - V2V). Така модель забезпечує прямий обмін повідомленнями між OBU в транспортних засобах. V2V може використовуватися для обміну інформацією про місцезнаходження та динаміку транспортних засобів, що може допомогти запобігти зіткненням та покращити потік руху.

- Автомобіль-Пішохід (Vehicle to Pedestrian - V2P). Модель передбачає передачу інформації між OBU в транспортному засобі та User Equipment (UE) у пішоходів, таких як смартфони. V2P може використовуватися для попередження пішоходів про небезпеку, що може допомогти зменшити кількість ДТП з участю пішоходів.

Структурну схему моделей обміну повідомленнями у системах V2X зображено на Рис. 2.

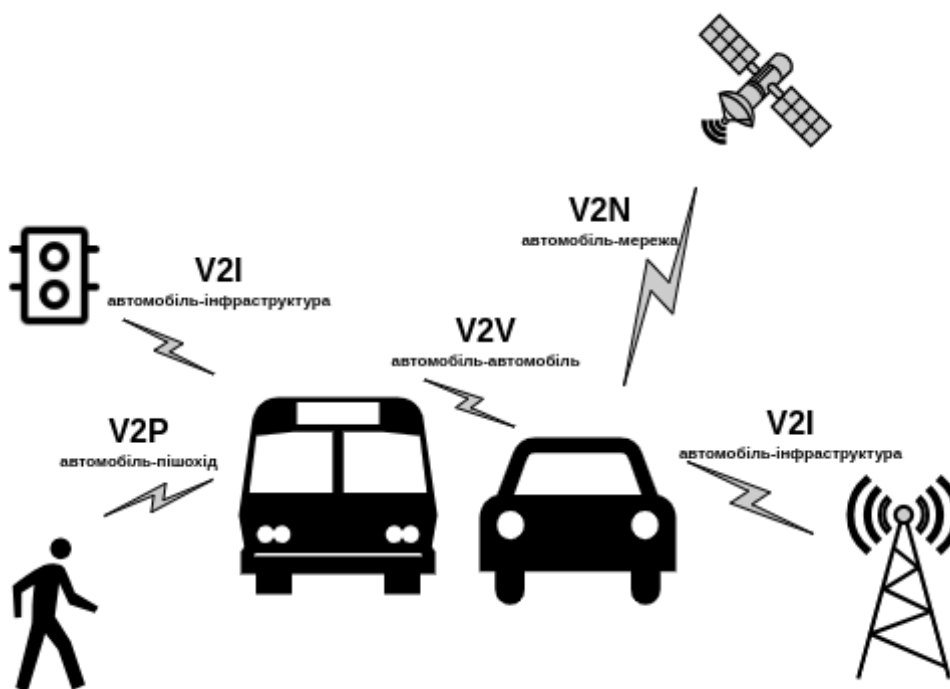


Рис. 2. Структура моделі обміну повідомленнями систем V2X



### Протоколи керування доступом до середовища (MAC) у системах V2X

Рівень керування доступом до середовища (MAC) є фундаментальним у всіх протоколах згідно з моделлю OSI, адже саме він визначає правила доступу до фізичного середовища [8]. У контексті V2X систем, MAC протоколи відіграють важливу роль у забезпеченні ефективного та справедливого розподілу обмежених ресурсів бездротового каналу для передачі даних [2]. Одні з найпоширеніших MAC протоколів, що використовуються у V2X системах:

- **IEEE 802.11p.** Протокол був спеціально розроблений для V2X застосувань. Він використовує модифіковану версію алгоритму CSMA/CA з додатковим каналом безпеки для критичних повідомлень про безпеку [12, 5].
    - Переваги.
    - Низька затримка, що робить його ідеальним для застосунків, які потребують швидкого реагування, таких як запобігання зіткненням.
    - Висока надійність, завдяки стійкості до перешкод і механізмів резервування.
    - Недоліки.
    - Обмежений радіус дії, зазвичай до 1000 метрів.
    - Менша пропускна здатність порівняно з LTE-V2X та 5G NR-V2X.
    - Вимагає розгортання спеціалізованої інфраструктури DSRC
  - **LTE-V2X.** Даний протокол використовує наявну інфраструктуру стільникової мережі LTE для V2X-зв'язку [10].
    - Переваги.
    - Широкий радіус дії. LTE-V2X може охоплювати більші території, порівняно з протоколом DSRC, завдяки ширшому радіусу дії мережі LTE.
    - Висока пропускна здатність. Протокол LTE-V2X пропонує значно вищу пропускну здатність, що робить його придатним для передачі великих обсягів даних, таких як відео та карти високої роздільної здатності.
    - Недоліки:
    - Більша затримка, порівняно з протоколом DSRC, що може вплинути на чутливі до часу V2X застосунків.
    - Можливі проблеми з перевантаженням мережі, особливо в густонаселених районах.
    - Мала надійність у порівнянні з DSRC, адже може бути вразлива до збоїв у роботі стільникової мережі.
  - **5G NR-V2X.** Протокол ґрунтується на принципах LTE-V2X, але пропонує значні покращення у розрізі затримки, пропускної здатності та надійності. Використовує нову технологію стільникового зв'язку 5G [12].
    - Переваги:
    - Має найкращі показники з трьох протоколів по відношенню до затримки, пропускної здатності та надійності.
    - Добре узгоджується з найсучаснішими V2X застосунками.
    - Недоліки:
    - Найдорожчий з трьох протоколів через необхідність розгортання нової інфраструктури 5G.
    - На даний час знаходиться на стадії розроблення і не має широкого доступу.
- Таким чином вибір найбільш ефективного протоколу MAC для систем V2X залежить від поставлених початкових задач, конкретного застосування та доступної інфраструктури.

### РОЗРОБЛЕННЯ СТЕКУ ТЕХНОЛОГІЙ ДЛЯ ПРОТОКОЛУ КОМУНІКАЦІЙ У СИСТЕМІ V2X

#### Вибір керування доступом до середовища (MAC)

Вибір протоколу керування доступом до середовища (MAC) є важливим кроком у розробленні протоколу комунікації для V2X систем. З огляду на це доцільно розглянути деякі основні протоколи MAC, кожен з яких має свої переваги та недоліки.

Хоча 5G NR-V2X є ідеальним варіантом для майбутніх V2X застосунків завдяки його оптимізації для таких систем, наразі ця технологія є занадто новою і не має необхідної інфраструктури.

Тому, враховуючи переваги та недоліки протоколів IEEE 802.11p та LTE-V2X, запропоновано використовувати обидва протоколи у зв'язці для розробленого протоколу.

LTE-V2X буде вважатися основним протоколом завдяки його кращим загальним характеристикам передачі даних. Він буде використовуватися для передачі інформаційних повідомлень, таких як: телеметрія автомобіля, зображення та відео з бортових камер, інформація про дорожній рух, тощо.

Щоб адресувати дані із меншою затримкою у передачі повідомлень та вищою надійністю, IEEE 802.11p буде використовуватися як резервний протокол. Також він буде використовуватися для передачі високопріоритетних та екстрених повідомлень, таких як: інформація про потенційне зіткнення, проїзд транспорту екстрених служб, попередження про небезпеку, тощо.

Переваги використання обох протоколів, забезпечує суттєвий позитивний ефект.

Протокол LTE-V2X працює переважно на нижніх рівнях моделі OSI [10, 15], зокрема на фізичному рівні забезпечує передачу та прийом сигналу, включаючи вибір радіочастоти, методи модуляції та контроль потужності сигналу. На каналному рівні керує кадріванням даних, виявленням і виправленням помилок, а також керуванням доступом до середовища (MAC) для спільного використання каналу зв'язку між транспортними засобами. На мережевому рівні (який частково залучений) у деяких сценаріях зв'язку може використовуватися для виконання таких функцій, як адресація та обмін маршрутною інформацією з мережевою інфраструктурою.

Протокол IEEE 802.11р працює головним чином на фізичному та каналному рівнях моделі OSI [10, 15], зокрема на фізичному рівні обробляє фізичну передачу пакетів даних через бездротове середовище, включаючи такі завдання, як модуляція сигналу, керування доступом до середовища та виявлення помилок. На каналному рівні даний протокол відповідає за кадрівання даних у пакети, додавання заголовків для маршрутизації та перевірки помилок, а також керування доступом до середовища (MAC) для забезпечення ефективного спільного використання бездротового каналу між пристроями.

Таким чином, протоколи LTE-V2X та IEEE 802.11р закривають нижні рівні моделі OSI розроблюваного протоколу, як показано на Рис. 3.

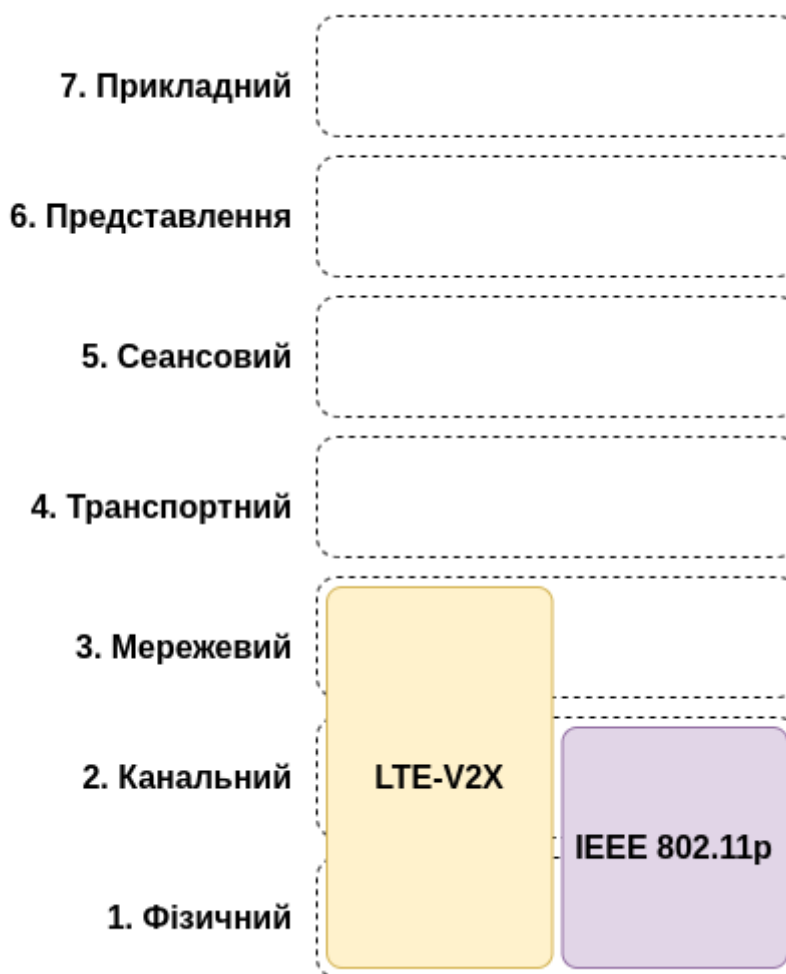


Рис. 3. Модель OSI запропонованого протоколу із закритими нижніми рівнями

#### Забезпечення пріоритетності повідомлень

Для гарантування своєчасної доставки критичних повідомлень у V2X-системах додано механізм пріоритетності трафіку. Така операція досягається шляхом використання комбінації протоколів TCP/UDP та додаткового заголовка з пріоритетом повідомлення (PRIO).

Використання TCP/UDP [9]:

- TCP (Transmission Control Protocol). Даний протокол використовується для надійної та впорядкованої передачі даних, що робить його придатним для інформаційних повідомлень, які потребують гарантії доставки, таких як телеметрія та дані про дорожній рух.

- UDP (User Datagram Protocol). Протокол використовується для швидкої передачі даних, що робить його придатним для критичних повідомлень, які потребують низької затримки, таких як попередження про зіткнення та інформація про екстрені служби.

Додатковий заголовок додається до кожного повідомлення V2X, який містить 1-байтове поле пріоритету. Значення пріоритету визначає важливість повідомлення та впливає на його обробку в системі: 0 - найвищий пріоритет, 255 - найнижчий. Пропонується резервувати  $\frac{1}{4}$  пріоритетів під високо пріоритетні повідомлення (пріоритети 0-63) та  $\frac{3}{4}$  під низько пріоритетні повідомлення (пріоритети 64-255). На основі значення пріоритету повідомлення обирається відповідний протокол нижнього рівня (IEEE 802.11p або LTE-V2X).

Високо пріоритетні повідомлення (пріоритети 0-63) передаються через UDP і відправляються по IEEE 802.11p.

Низько пріоритетні повідомлення (пріоритети 64-255) передаються через TCP і відправляються через LTE-V2X.

Таким чином, транспортний рівень моделі OSI для розроблюваного протоколу використовуватиме TCP/UDP із додатковим заголовком пріоритетності повідомлення, як зображено на Рис. 4.

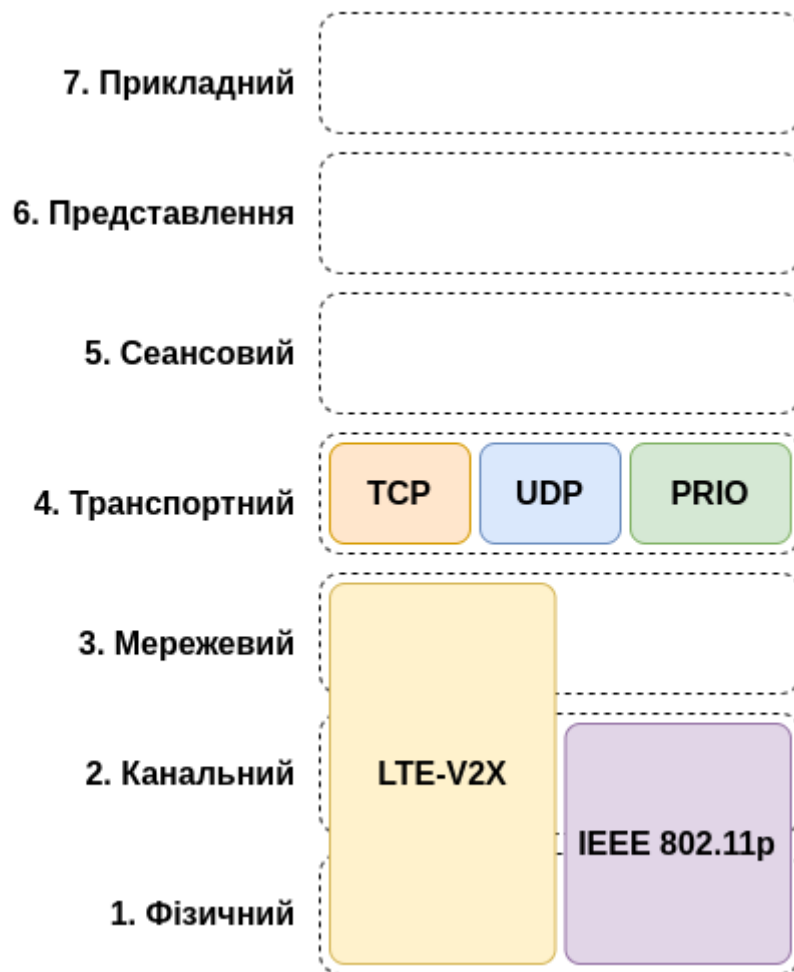


Рис. 4. Поточна модель OSI розроблюваного протоколу із транспортним рівнем

Таким чином, структура формування V2X повідомлення для відправлення із верхніх рівнів виглядатиме, як зображено на Рис. 5. На мові C розроблений код реалізації обгортки повідомлень на транспортному рівні.

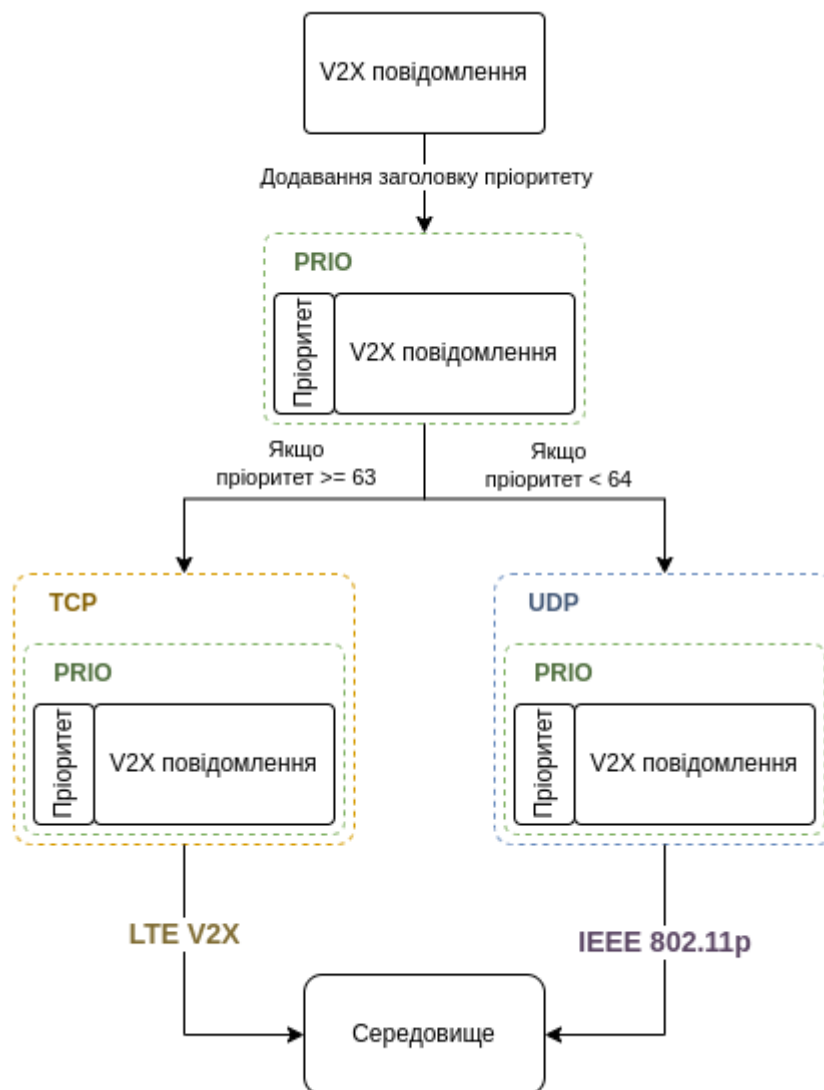


Рис. 5. Структура формування V2X повідомлення для відправлення у фізичне середовище

### Забезпечення цілісності та автентифікації

У системах V2X обмін повідомленнями є надзвичайно чутливим, адже вони можуть нести критичну інформацію, що впливає на безпеку дорожнього руху [11]. Тому захист цих повідомлень від модифікації та підробки є дуже важливим.

Для досягнення цієї мети використано інфраструктуру публічних ключів (Public Key Infrastructure, PKI) [6].

Інфраструктура публічних ключів PKI, є основою довіри в цифровій комунікації. Вона керує створенням, розповсюдженням, зберіганням та відкликанням цифрових сертифікатів. Дані сертифікати діють як електронні паспорти, що зв'язують відкритий ключ із певною сутністю (людина, пристрій, веб-сайт) та гарантують його дійсність за допомогою довіреного центру сертифікації. PKI забезпечує безпечне спілкування шляхом перевірки особи учасників та шифрування передачі даних [18].

Відомо, що цифровий підпис представляє криптографічний код, який дозволяє підтвердити автентичність та цілісність повідомлення. Він генерується шляхом застосування приватного криптографічного ключа до хеш-суми повідомлення [3]. Цифровий сертифікат представляє електронний документ, який зв'язує відкритий криптографічний ключ з ідентифікатором власника (наприклад, VIN автомобіля або ID пристрою інфраструктури). Він видається довіреним центром сертифікації (CA) і підписується цифровим підписом CA [7].

Забезпечення цілісності та автентифікації V2X-повідомлень здійснюється наступним чином.

1. Кожен учасник системи V2X (автомобіль, пристрій інфраструктури тощо) повинен мати цифровий сертифікат. Цей сертифікат може бути виданий виробником пристрою або державним органом.

2. Пул довірених центрів сертифікації (Certificate Authority - CA) має бути погоджений між усіма учасниками системи. Це гарантує, що всі сертифікати, що використовуються, є надійними та можуть бути перевірені.

3. V2X-повідомлення підписується цифровим підписом відправника. Цей підпис генерується за допомогою приватного ключа відправника та хеш-суми повідомлення.

4. Отримувач повідомлення перевіряє сертифікат відправника. Дана операція виконується шляхом перевірки цифрового підпису у сертифікаті та ланцюжка довіри (chain of trust), який веде до кореневого СА.

5. Отримувач використовує відкритий ключ відправника, отриманий із сертифіката, для перевірки цифрового підпису повідомлення. Якщо підпис дійсний, це підтверджує, що повідомлення не було модифіковано під час передачі, а також, що відправник дійсно є той, за кого він себе видає.

На Рис. 6 наведений процес формування підпису для V2X повідомлення, а, на Рис. 7 зображено процес перевірки цілісності отриманого повідомлення та перевірка дійсності сертифіката відправника.

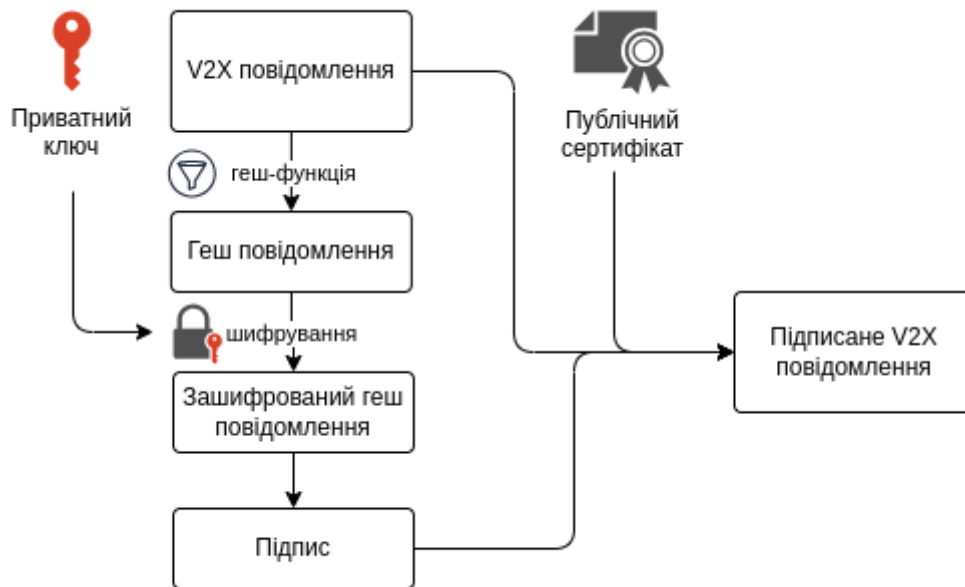


Рис. 6. Схематичне представлення процесу підпису V2X повідомлення цифровим ключем

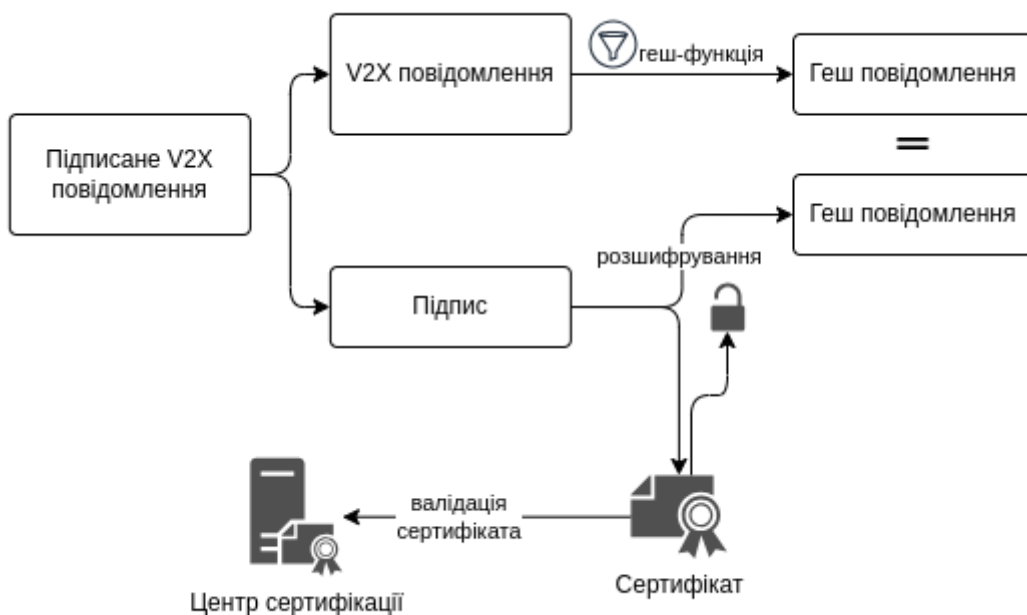


Рис. 7. Схематичне зображення процесу перевірки цілісності отриманого V2X повідомлення та дійсності сертифіката відправника

Остаточний вигляд запропонованого протоколу в моделі OSI наведено на Рис. 8.

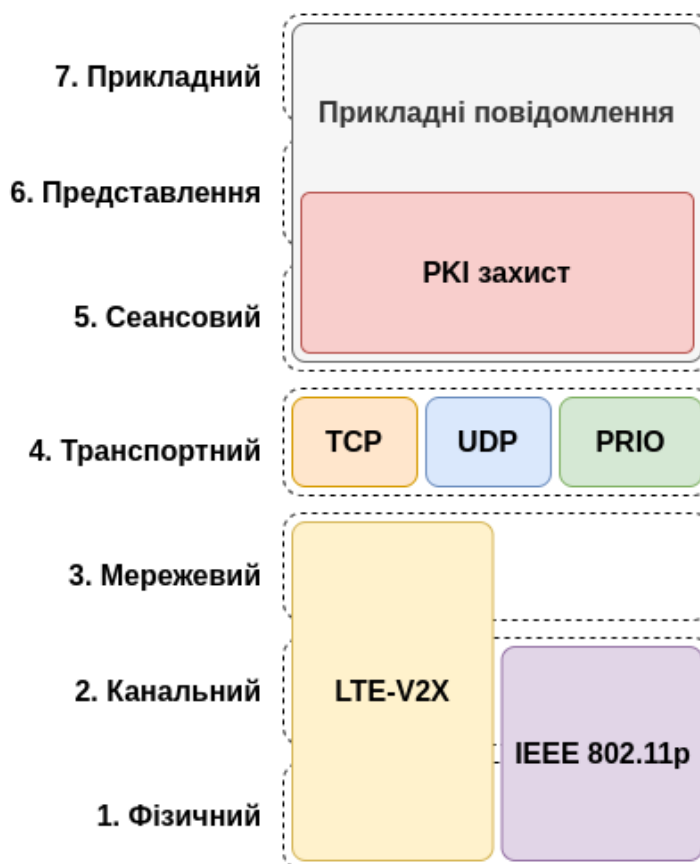


Рис. 8. Модель OSI запропонованого протоколу

Таким чином, можна зробити висновок, що запропоноване поєднання протоколів IEEE 802.11p та LTE-V2X дозволяє на даний час забезпечити високу ефективність та надійність в різних сценаріях, використання комбінації TCP/UDP з додатковим заголовком дозволяє реалізувати пріоритетність своєчасної доставки критичних повідомлень, а використання інфраструктури публічних ключів для цифрового підпису та сертифікації повідомлень, дозволяє запобігти модифікації та підробці.

### ДОСЛІДЖЕННЯ ЕФЕКТИВНОСТІ ПРОПОНОВАНОГО ВИБОРУ ВИКОРИСТАННЯ ДВОХ ПРОТОКОЛІВ ФІЗИЧНОГО РІВНЯ

Наступним важливим етапом є детальне дослідження ефективності вибору двох запропонованих протоколів фізичного рівня, а саме IEEE 802.11p та LTE-V2X. Дане дослідження включає кілька важливих аспектів:

- Аналіз продуктивності. Вивчення продуктивності обраних протоколів у різних умовах трафіку та навантаження на мережу. Дослідження включає вимірювання пропускної здатності, затримок і надійності передачі даних.
- Симуляція та реальні тести. Проведення комп'ютерної симуляції та реальних експериментів для перевірки ефективності протоколів у різних сценаріях, що дозволить оцінити роботу в умовах високої мобільності та динамічного середовища та в загальному доцільність такого вибору.
- Порівняння результатів. Порівняння результатів, отриманих у проведених дослідженнях, з метою визначення найкращого підходу для конкретних випадків використання в системах V2X.

#### Вимоги до апаратного забезпечення

Для забезпечення ефективної роботи розробленого протоколу необхідно чітко сформулювати вимоги до апаратного забезпечення.

- Вибір радіомодулів. Проведення вибору відповідних радіомодулів для підтримки IEEE 802.11p та LTE-V2X, яке включає оцінку доступних на ринку модулів за такими параметрами, як діапазон частот, пропускна здатність, затримка та енергоефективність.

• Контролери та обчислювальні ресурси. Визначення вимог до контролерів і обчислювальних ресурсів для обробки даних у реальному часі, яке включає вибір процесорів, пам'яті та інших компонентів, що забезпечують швидке і надійне виконання необхідних алгоритмів.

#### Імплементация пристроїв для проведення досліджень

Наступний крок полягає у проведенні імплементации прототипів пристроїв, які дозволяють проводити експериментальні дослідження.

• Розроблення прототипів. Розроблення прототипів транспортних засобів та інфраструктурних пристроїв, які підтримують обрані протоколи і включають апаратне та програмне забезпечення, необхідне для тестування V2X комунікацій.

• Проведення експериментів. Проведення експериментів з використанням розроблених прототипів у різних умовах дорожнього руху, що дозволяє зібрати дані для подальшого аналізу та вдосконалення протоколу.

#### Деталізація інфраструктури публічних ключів

З метою забезпечення безпеки V2X комунікацій необхідно деталізувати інфраструктуру публічних ключів (PKI). Для цього необхідно здійснити.

• Визначення процедур. Визначення процедур видачі, зберігання та відкликання цифрових сертифікатів для учасників системи V2X.

• Визначення центрів сертифікації. Визначення процесу створення надійних центрів сертифікації (CA), які відповідають за видачу та управління цифровими сертифікатами.

Таким чином, сформульовані етапи дослідження спрямовані на подальше вдосконалення та підвищення ефективності системи V2X і є ключовими на шляху до створення надійної та ефективної системи комунікації для транспортних засобів.

Запропонований протокол комунікації для V2X систем забезпечує надійний і безпечний обмін даними між транспортними засобами та інфраструктурою, сприяє підвищенню безпеки і ефективності дорожнього руху, а проведення подальших досліджень та впровадження запропонованих рішень дозволять досягти більш високих результатів у цій галузі, що зробить транспортну інфраструктуру більш технологічно розвиненою.

#### References

1. Ahmed Z., Naz S., Ahmed J. Minimizing transmission delays in vehicular ad hoc networks by optimized. *Wireless Networks*. 2020. Vol. 26, no. 4. P. 2905–2914. URL: <https://doi.org/10.1007/s11276-019-02198-x>.
2. An Analysis on Contemporary MAC Layer Protocols in Vehicular Networks: / L. Hota et al. *Future Internet*. 2021. Vol. 13, no. 11. P. 287. URL: <https://doi.org/10.3390/fi13110287>.
3. Andvsilva. Digital Signature with Hash Function – How it works?. *Medium*. 2024. URL: <https://andsilvadrc.medium.com/digital-signature-with-hash-function-how-it-works-f4eed52267f5>.
4. Anwar W., Franchi N., Fettweis G. Physical Layer Evaluation of V2X Communications Technologies: 5G NR-V2X.. 2019. URL: <https://doi.org/10.1109/vtcfall.2019.8891313>.
5. Arena F., Pau G., Severino A. A Review on IEEE 802.11p for Intelligent Transportation Systems. *Journal of Sensor and Actuator Networks*. 2020. Vol. 9, no. 2. P. 22. URL: <https://doi.org/10.3390/jsan9020022>.
6. Bengtsson J. Steering the Course of Cybersecurity in the Automotive Industry with PKI. *Nexusgroup*. URL: <https://www.nexusgroup.com/steering-the-course-of-cybersecurity-in-the-automotive-industry-with-pki/>.
7. Content Manager OnDemand for Multiplatforms 10.5.0. *IBM - United States*. URL: <https://www.ibm.com/docs/en/cmofm/10.5.0?topic=cryptography-digital-certificates-certificate-authorities> (дата звернення: 10.06.2024).
8. Contributors to Wikimedia projects. Medium access control - Wikipedia. *Wikipedia, the free encyclopedia*. URL: [https://en.wikipedia.org/wiki/Medium\\_access\\_control](https://en.wikipedia.org/wiki/Medium_access_control) (дата звернення: 10.06.2024).
9. Gorman B. TCP vs UDP: What's the Difference and Which Protocol Is Better?. *TCP vs UDP: What's the Difference and Which Protocol Is Better?*. URL: <https://www.avast.com/c-tcp-vs-udp-difference>.
10. On the Impact of Multiple Access Interference in LTE-V2X and NR-V2X / A. Rehman et al. *Sensors*. 2023. Vol. 23, no. 10. P. 4901. URL: <https://doi.org/10.3390/s23104901>.
11. Popović I. V2X - Paving the Superhighway for Tomorrow's Smart Mobility | HTEC | HTEC. *HTEC*. URL: <https://htecgroup.com/v2x-paving-the-superhighway-for-tomorrows-smart-mobility/>.
12. Review of V2X-IoT Standards and Frameworks for ITS Applications / K. Kiela et al. *Applied Sciences*. 2020. Vol. 10, no. 12. P. 4314. URL: <https://doi.org/10.3390/app10124314>.
13. Team E. What Is V2X and The Future of Vehicle to Everything Connectivity. *www.emqx.com*. URL: <https://www.emqx.com/en/blog/what-is-v2x-and-the-future-of-vehicle-to-everything-connectivity>.
14. V2X Communication Technologies and Service Requirements for Connected and / E. Cinque et al. 2020. URL: <https://doi.org/10.23919/aetitautomotive50086.2020.9307388>.
15. V2X Network Architecture and Standards System / S. Chen et al. *Wireless networks*. 2023. P. 81–116. URL: [https://doi.org/10.1007/978-981-19-5130-5\\_3](https://doi.org/10.1007/978-981-19-5130-5_3).
16. Waysion. Unveiling the Power of On-Board Units (OBU) in Modern Vehicles - Waysion. *Waysion*. URL: <https://www.waysion.com/blog/on-board-units-obu/>.
17. What is a Roadside Unit (RSU)? RSU Meaning | Isarsoft. *Isarsoft | We make every camera count*. URL: <https://www.isarsoft.com/knowledge-hub/rsu> (дата звернення: 10.06.2024).
18. What is PKI | Public Key Infrastructure | DigiCert. *SSL Digital Certificate Authority | Encryption & Authentication | DigiCert.com*. URL: <https://www.digicert.com/what-is-pki> (дата звернення: 10.06.2024).
19. What Is Vehicle-to-Everything (V2X)?. *Embedded OS, Support, and Services | BlackBerry QNX*. URL: <https://blackberry.qnx.com/en/ultimate-guides/software-defined-vehicle/vehicle-to-everything> (дата звернення: 10.06.2024).

<https://doi.org/10.31891/2219-9365-2024-79-18>

УДК 004.94 (072):004.2

ДЯЧЕНКО Григорій

Національний технічний університет «Дніпровська політехніка»

<https://orcid.org/0000-0001-9105-1951>

[diachenko.g@nmu.one](mailto:diachenko.g@nmu.one)

КРЕМНЬОВ Володимир

Національний технічний університет «Дніпровська політехніка»

<https://orcid.org/0009-0005-1149-2017>

[kremnov.vo.v@nmu.one](mailto:kremnov.vo.v@nmu.one)

## МАТЕМАТИЧНЕ ТА КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ПРОЦЕСУ КОМП'ЮТЕРИЗОВАНОГО ПРЕДИКАТИВНОГО КОНТРОЛЮ РЕЖИМІВ ЗРОШЕННЯ ҐРУНТУ ПІД ЧАС ВИРОЩУВАННЯ ЗЕРНОВИХ КУЛЬТУР

У теперішній в Україні значної актуальності набули питання, що пов'язані з забезпеченням базових продовольчих потреб населення та підтримання національної економіки. Рослинництво відкритого ґрунту є однією з основних галузей національного сільського господарства та відноситься до переліку стратегічних, адже вносить значний вклад у формування національної продовольчої безпеки, економіки та експортного потенціалу. В свою чергу, ця галузь потребує технологічного переоснащення та програмно-технічної модернізації, в тому числі, завдяки розробці й упровадженню високоефективних інформаційних і комп'ютерних технологій контролю агротехнічних процесів. Основна мета статті полягає в обґрунтуванні підходів до підвищення ефективності агротехнічних процесів вирощування зернових культур за рахунок розробки та дослідження математичної й комп'ютерної моделей процесу комп'ютеризованого предикативного контролю режимів зрошення ґрунту під час вирощування зернових культур. Об'єкт дослідження процеси комп'ютерної обробки та аналізу даних щодо динаміки вологи в ґрунті під час вирощування зернових культур. Предмет дослідження методи та моделі побудови комп'ютерних технологій предикативного контролю режимів зрошення ґрунту під час вирощування зернових культур. В цій статті отримано та висвітлено наступні науково-практичні результати: запропоновано та алгоритмізовано математичний опис процесу динаміки вологи в різних типах ґрунту з урахуванням специфічних агрокліматичних зон України; методами комп'ютерного експерименту оцінено параметри і характеристики динаміки вологи в різних типах ґрунту за періоди, що дорівнюють одній годині та добі; доведено, що отримані результати можуть бути використані в якості програмно-алгоритмічного та структурно-функціонального базису під час подальших розробок і досліджень з проектування та тестування інтелектуальної комп'ютерної технології предикативного контролю режимів зрошення зернових культур у різних агрокліматичних зонах України.

Ключові слова: вологість, зернові культури, математична модель, комп'ютерний експеримент, контроль.

DIACHENKO Grygorii, KREMNOV Volodymyr

Dnipro University of Technology

## MATHEMATICAL AND COMPUTER MODELLING OF THE PROCESS OF COMPUTERISED PREDICTIVE CONTROL OF SOIL IRRIGATION MODES DURING GRAIN CROPS CULTIVATION

Nowadays, issues related to ensuring the basic food needs of the population and supporting the national economy have become of significant relevance in Ukraine. Open-air crop production is one of the key areas of national agriculture and is classified as a strategic sector, since it significantly contributes to national food security, economy and export potential. In turn, this production area requires technological re-equipment, as well as software and hardware modernisation, including through the development and implementation of highly efficient information and computer technologies for monitoring agricultural processes. The aim of the article is to substantiate approaches to improving the efficiency of agrotechnical processes of growing grain crops by developing and studying mathematical and computer models of the process of computerised predictive control of soil irrigation modes when cultivating grain crops. The object of the research is the processes of computer processing and analysis of data on the dynamics of soil moisture when cultivating grain crops. The subject of the research is methods and models of creating computer technologies for predictive control of soil irrigation modes when cultivating grain crops. The following scientific and practical results have been obtained and presented in the article: a mathematical description of the process of moisture dynamics in different types of soil has been proposed and algorithmised, taking into account the specific agroclimatic zones of Ukraine; the parameters and characteristics of moisture dynamics in different types of soil for periods equal to one hour and one day have been estimated by computer experiment; it has been proved that the obtained results can be used as a software and algorithmic, as well as a structural and functional basis for further development and research on the design and testing of an intelligent computer technology for predictive control of irrigation modes in different agroclimatic zones of Ukraine.

Keywords: moisture, grain crops, mathematical model, computer experiment, monitoring.

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ

#### ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Через війну в Україні значної актуальності та важливості набувають питання, що пов'язані з забезпеченням базових продовольчих потреб населення та підтримання національної економіки. Сільське господарство (с/г), а саме, рослинництво відкритого ґрунту, є галуззю, що відноситься до переліку



стратегічних, адже вносить значний вклад у формування національної економіки та експортного потенціалу. В той самий час земельні ресурси є обмеженими для цілей масштабування під час вирощування зернових культур, отже зростає потреба в розробці та впровадженні високоефективних підходів до забезпечення необхідних показників обсягів і якості виробництва с/г продукції рослинництва. В свою чергу, глобальна світова практика країн із розвиненими аграрними технологіями доводить, що використання інформаційних і комп'ютерних технологій моніторингу й контролю стану посівів і земельних ресурсів обумовлює значний позитивний ефект провадження с/г діяльності рослинництва відкритого ґрунту у коротко- та середньостроковій перспективі.

В Україні стратегічне значення під час формування національної економіки та забезпечення продовольчої безпеки мають такі типи зернових культур: пшениця, ячмінь, кукурудза, жито й просо [1, 2]. Результати графічної інтерпретації апріорних статистичних показників динаміки обсягів вирощування цих зернових культур в різних агрокліматичних зонах України (з 2018 по 2022 рр.) наведено на рис. 1 (2022 р. – остання актуалізація даних).

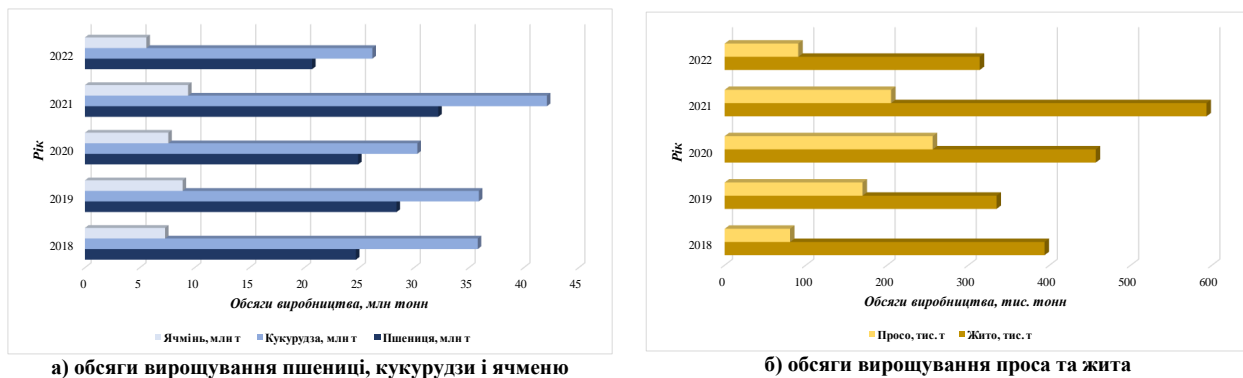


Рис. 1. Графічна інтерпретація статистичних показників вирощування основних зернових культур в Україні з 2018 по 2022 рр.

В результаті аналізу даних, які наведено на рис. 1, встановлено наступне: пшениця (середній обсяг вирощування за період із 2018 по 2022 рр. становить 26,2 млн т), кукурудза (34,1 млн га) і ячмінь (7,8 млн га) вносять основний вклад до загальних обсягів виробництва с/г продукції рослинництва; через повномасштабну війну в Україні обсяги вирощування зазначених с/г культур значно зменшились: як результат порівняння показників за 2022 р. та 2021 р. у відсотках встановлено, що показник зменшення є наступним: просо (–55,6 %), жито (–47,1 %), ячмінь (–40,4 %), кукурудза (–37,8 %) і пшениця (–35,7 %).

Отже, основна мета статті полягає в обґрунтуванні підходів до підвищення ефективності агротехнічних процесів вирощування зернових культур за рахунок розробки та дослідження математичної й комп'ютерної моделей процесу комп'ютеризованого предикативного контролю режимів зрошення ґрунту під час вирощування зернових культур. Об'єкт дослідження процесу комп'ютерної обробки та аналізу даних щодо динаміки вологи в ґрунті під час вирощування зернових культур. Предмет дослідження методи та моделі побудови комп'ютерних технологій предикативного контролю режимів зрошення ґрунту під час вирощування зернових культур. Таким чином, можна констатувати факт того, що науково-практична задача розробки, дослідження та впровадження високоефективних комп'ютерних і інформаційних технологій агротехнічного призначення є важливою і актуальною в Україні і світі, а також її розв'язання дозволить обґрунтувати техніко-функціональні підходи до раціонального використання земельних та матеріальних, у тому числі, водних ресурсів під час ведення с/г діяльності рослинництва відкритого ґрунту.

### ПІДХОДИ ТА МЕТОДОЛОГІЯ ДОСЛІДЖЕНЬ

Процес розробки та подальшого використання комп'ютерної технології інтелектуального контролю режимів зрошення зернових культур потребує синтезу, дослідження та валідації відповідної математичної моделі, що описує динаміку вологи в різних типах ґрунтів на різних глибинах. В свою чергу, основним призначенням цієї моделі є обґрунтування функціонально-алгоритмічного базису програмних компонент досліджуваної комп'ютерної технології. Інформаційно-алгоритмічною основою зазначеної моделі є формалізований опис явища гідрологічного балансу, що ґрунтується на закономірностях динаміки об'ємного вмісту вологи в шарах ґрунту протягом аналізованого періоду часу. Гідрологічний баланс ґрунту обумовлено фізичними процесами надходження (примусове зрошення, підйом ґрунтових вод, опади та ін.) та відтоку (випаровування, поверхневий стік, інфільтрація та ін.) ґрунтової вологи [3, 4], а також біологічними явищами споживання вологи рослинами з ґрунту. На підставі вищенаведеного опису було запропоновано схему гідрологічного балансу ґрунту, як у графічному вигляді наведено на рис. 2.

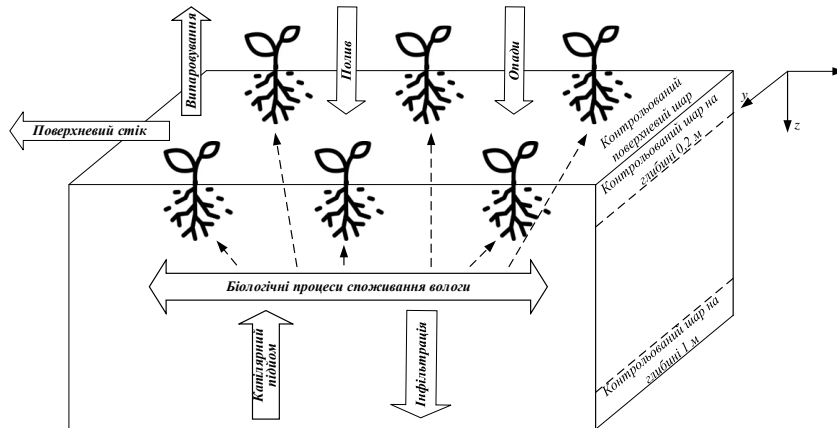


Рис. 2. Графічна інтерпретація процесу гідрологічного балансу ґрунту

Рівняння водного балансу ґрунту в загальному вигляді з урахуванням основних впливних біологічних і фізичних процесів (див. рис. 2) має наступний вигляд [3, 4]:

$$W_{start} + W_{gw} + W_{pr} + W_{irr} = W_{end} + W_{ev} + W_{inf} + W_{dr} + W_{crop}, \quad (1)$$

де  $W_{start}$  – запас вологи в шарі ґрунту, що аналізується, на момент початку детектування;  $W_{gw}$  – волога, яка надходить через процес капілярного підйому підземних ґрунтових вод;  $W_{pr}$  – волога, яка надходить через опади;  $W_{irr}$  – притік вологи через полив;  $W_{end}$  – запас вологи в шарі ґрунту, що аналізується, на момент закінчення детектування;  $W_{ev}$  – волога, яка випаровується;  $W_{inf}$  – волога, яка втрачається через процес інфільтрації ґрунту;  $W_{dr}$  – волога, яка втрачається через поверхневий стік;  $W_{crop}$  – волога, яка споживається рослинами.

Під час досліджень варто прийняти до уваги той факт, що фізико-математичною основою формалізованого опису процесу розповсюдження вологи в аналізованих шарах ґрунту є рівняння Річардса в одновимірній формі (за глибиною), яке визначає закономірності перенесення вологи в аераційній зоні ґрунту з обліком джерел надходження й відтоку води (див. рис. 2) [5–7]:

$$\frac{\partial W_{soil}}{\partial t} = \frac{\partial}{\partial z} \left( K(W_{soil}) \frac{\partial \psi(W_{soil})}{\partial z} \right) + S_{in}(z, t) - S_{out}(z, t), \quad (2)$$

де  $W_{soil}$  – вологість ґрунту (об'ємна);  $z$  – аналізована координата (глибина);  $t$  – час;  $K(W_{soil})$  – гідрравлічна провідність ґрунту;  $\psi(W_{soil})$  – потенціал ґрунтової вологи;  $S_{in}$  – функція джерел надходження вологи;  $S_{out}$  – функція джерел відтоку (споживання та випаровування) вологи.

Гідрравлічна провідність ґрунту, яка входить до формули (2), з обліком масштабованої області застосування від максимальної гігроскопічності до повного насичення ґрунту водою обчислюється за такою формулою [7]:

$$K(W_{soil}) = K_f \left( \frac{W_{soil} - W_{max}}{m - W_{max}} \right)^5, \quad (3)$$

де  $K(W_{soil})$  – гідрравлічна провідність ґрунту;  $K_f$  – коефіцієнт, що визначає фільтраційні властивості ґрунту;  $W_{soil}$  – вологість ґрунту;  $W_{max}$  – максимальна гігроскопічність ґрунту;  $m$  – параметр, що відповідає за пористість ґрунту.

Потенціал вологи в ґрунті  $\psi(W_{soil})$  може змінюватись у значному динамічному діапазоні, який обумовлено поточною вологістю ґрунту й типом ґрунту: від значно негативного (-1500 кПа – для посушливих умов) до позитивного потенціалу (+10 кПа – для ґрунтів, які значно перенасичено вологою) [8, 9].

З урахуванням зазначеного вище розв'язання рівняння (2) виконано методами комп'ютерного моделювання на основі чисельного інтегрування в онлайн-середовищі Jupyter Notebook із використанням розробленої програми мовою Python. Під час розв'язання диференційного рівняння (2) враховано чисельні характеристики, умови та обмеження, які зазначено в табл. 1.

Таблиця 1.

Характеристики, параметри, умови та обмеження під час комп'ютерного моделювання динаміки  
вологи в ґрунті

Враховані характеристики, параметри, умови та обмеження	Прийняті значення
Шари ґрунту, що аналізуються	Координата глибини ( $z$ ) варіюється від 0 м до 1 м з кроком 0,01 м.
Період часу аналізу	Параметр $t$ (час) варіюється в діапазоні від 0 год. до 1 год. з кроком 1 хв.
Початкові умови	Дискретизований діапазон значень об'ємної вологи $W_{soil}(z,0)$ варіюється від 50 % до 100 % з кроком 10 %.
Граничні умови	На верхній межі (поверхневий шар ґрунту $z=0$ м): дискретизований діапазон даних від 50 % до 100 %; на нижній межі (кореневий шар ґрунту $z=1$ м): в початковий момент часу вологість дорівнює 0 %.
Джерела надходження вологи до ґрунту	Полив, капілярний підйом вологи та опади. Враховується через чисельні значення граничних і початкових умов: $S_m(t,z)=0$ .
Джерела втрат вологи з ґрунту	Випаровування в умовах вологості, напівпосушливої й посушливої зон, відповідно рівні: 0,08 мм/год; 0,16 мм/год та 0,32 мм/год [10].
Провідність ґрунту (гідравлічна)	Обчислюється за формулою (3) з обліком різних типів ґрунту: 8,1 мм/год. (чорнозем потужний), 10,8 мм/год. (чорнозем звичайний) та 12,6 мм/год. (чорнозем південний) [11] з показником пористості аерації для різного гранулометричного складу (осереднено дорівнює 0,3) [12].
Потенціал вологи в ґрунті	Варіюється в динамічному діапазоні від 0 до -100 кПа.

Таким чином, математичний опис динаміки об'ємної вологи ґрунту на основі рівнянь (1)–(3) та прийняті умови й обмеження (див. табл. 1) дозволили провести серію комп'ютерних експериментів із обґрунтування вимог до комп'ютерної технології предикативного контролю режимів зрошення зернових культур.

РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ

Отримані методами комп'ютерного моделювання результати розв'язання диференційного рівняння (2) щодо аналізу динаміки вологи в ґрунті за період часу 1 год. з урахування прийнятих чисельних параметрів, обмежень і умов, а також визначених агрокліматичних зон і типів ґрунту (див. табл. 1) наведено на рис. 3–5.

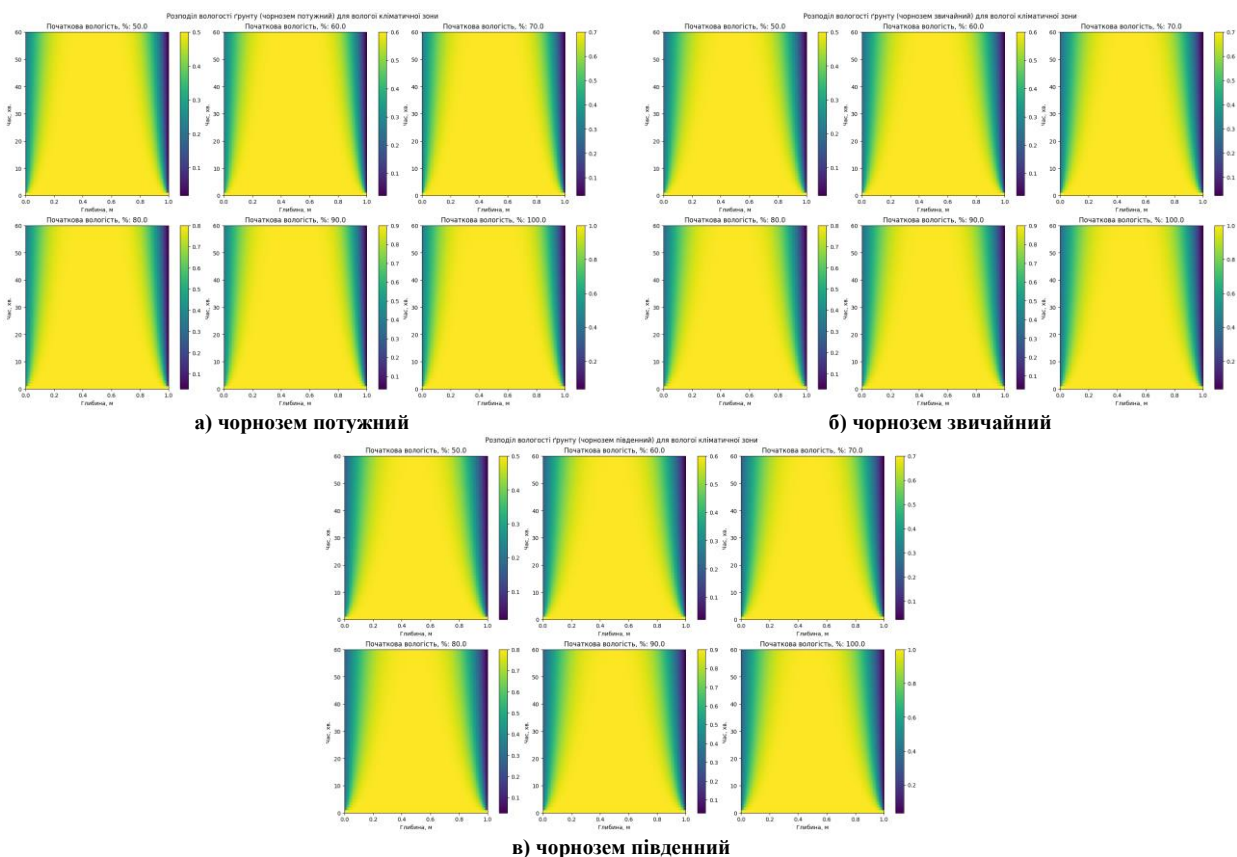


Рис. 3. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах вологості агрокліматичної зони за годину

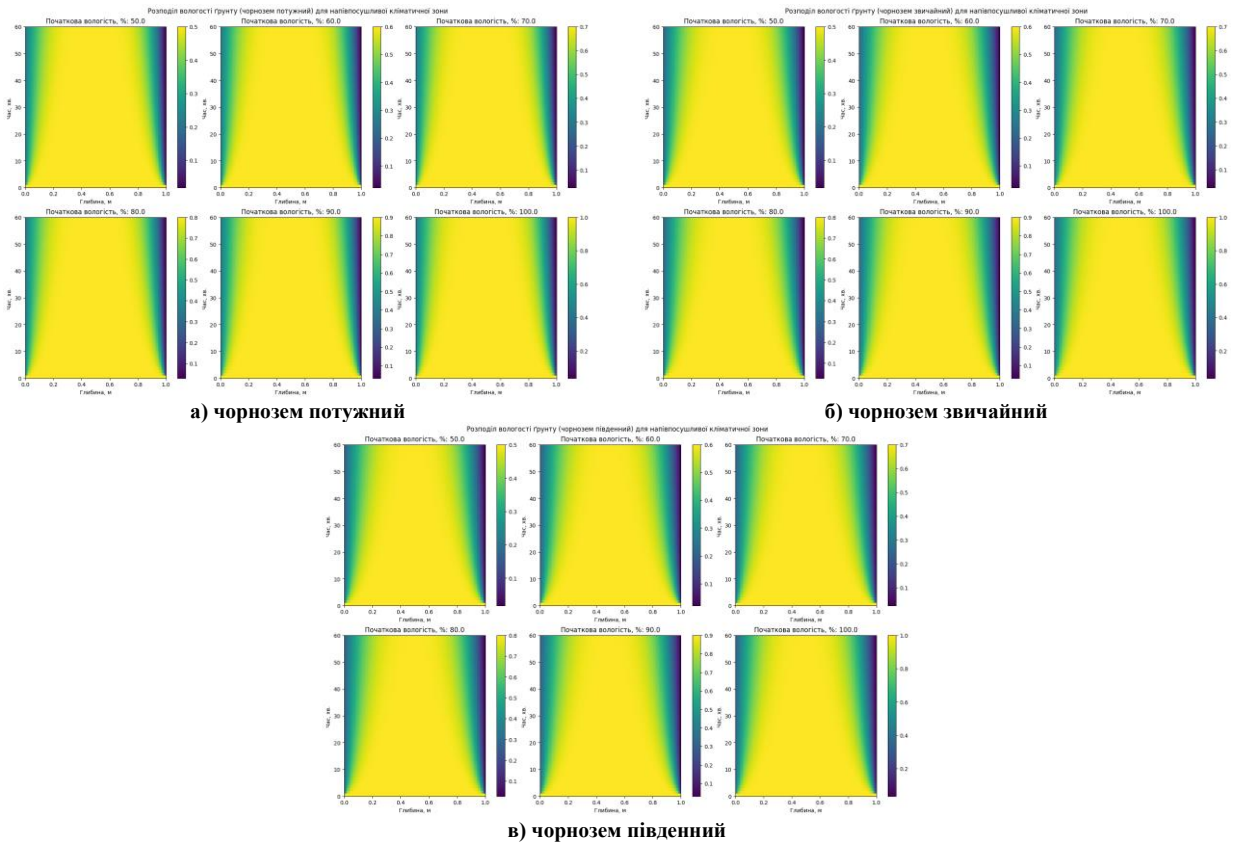


Рис. 4. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах напівпосушливої агрокліматичної зони за годину

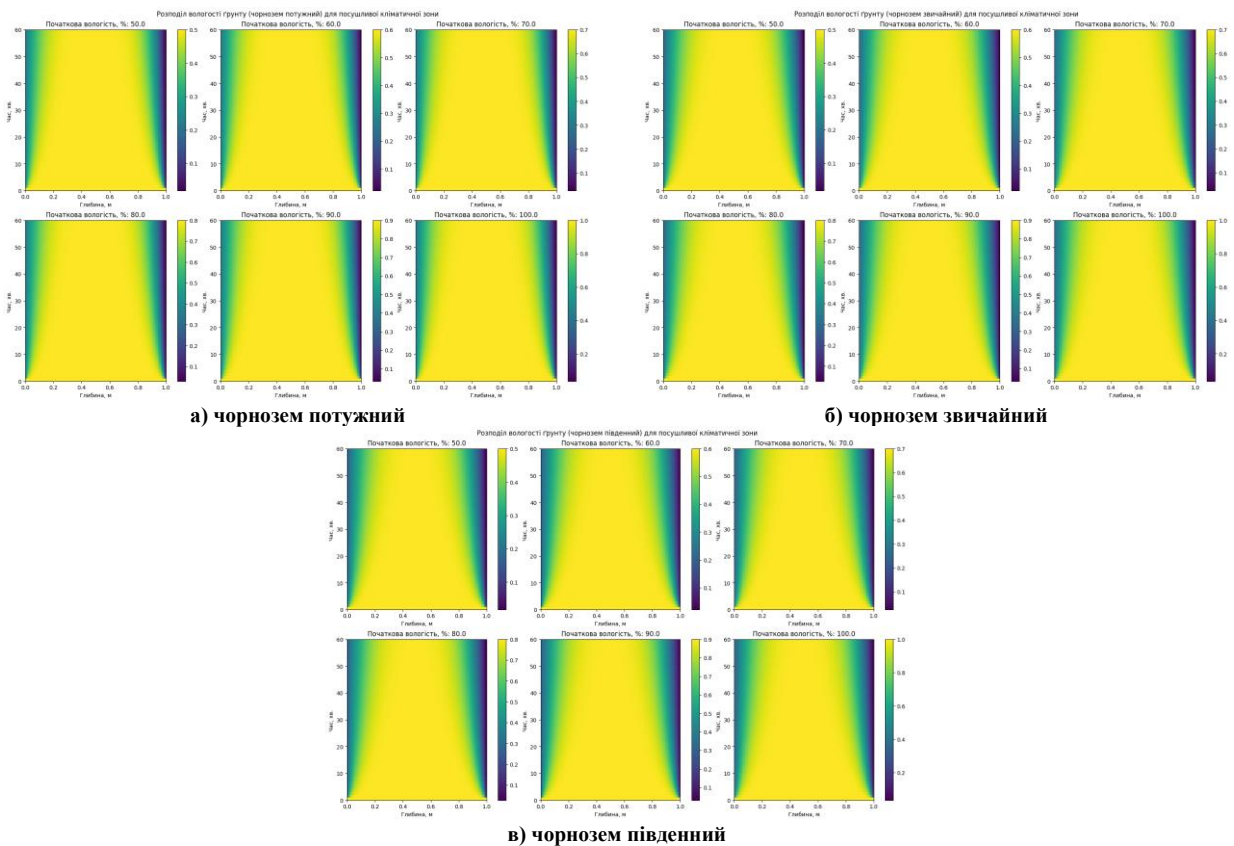


Рис. 5. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах посушливої агрокліматичної зони за годину

Отримані результати досліджень (див. рис. 3–5) дозволили оцінити якісні та кількісні показники процесу розповсюдження ґрунтової вологи в залежності від аналізованого шару на короткотривалому (1

год.) проміжку часу: для чорнозему потужного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,17 м до 0,83 м для всіх агрокліматичних зон, що аналізуються; для чорнозему звичайного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,22 м до 0,78 м для всіх агрокліматичних зон, що аналізуються; для чорнозему південного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,26 м до 0,74 м для всіх агрокліматичних зон, що аналізуються.

Варто зазначити, що отримані результати досліджень (див. рис. 3–5) не дозволили проаналізувати вплив процесів споживання та відтоку ґрунтової вологи через недостатній проміжок часу аналізу. Таким чином, авторами було встановлено необхідність проведення додаткових комп'ютерних експериментів для масштабованого періоду часу, рівного одній добі. Отримані графічні результати комп'ютерного експерименту наведено на рис. 6–8.

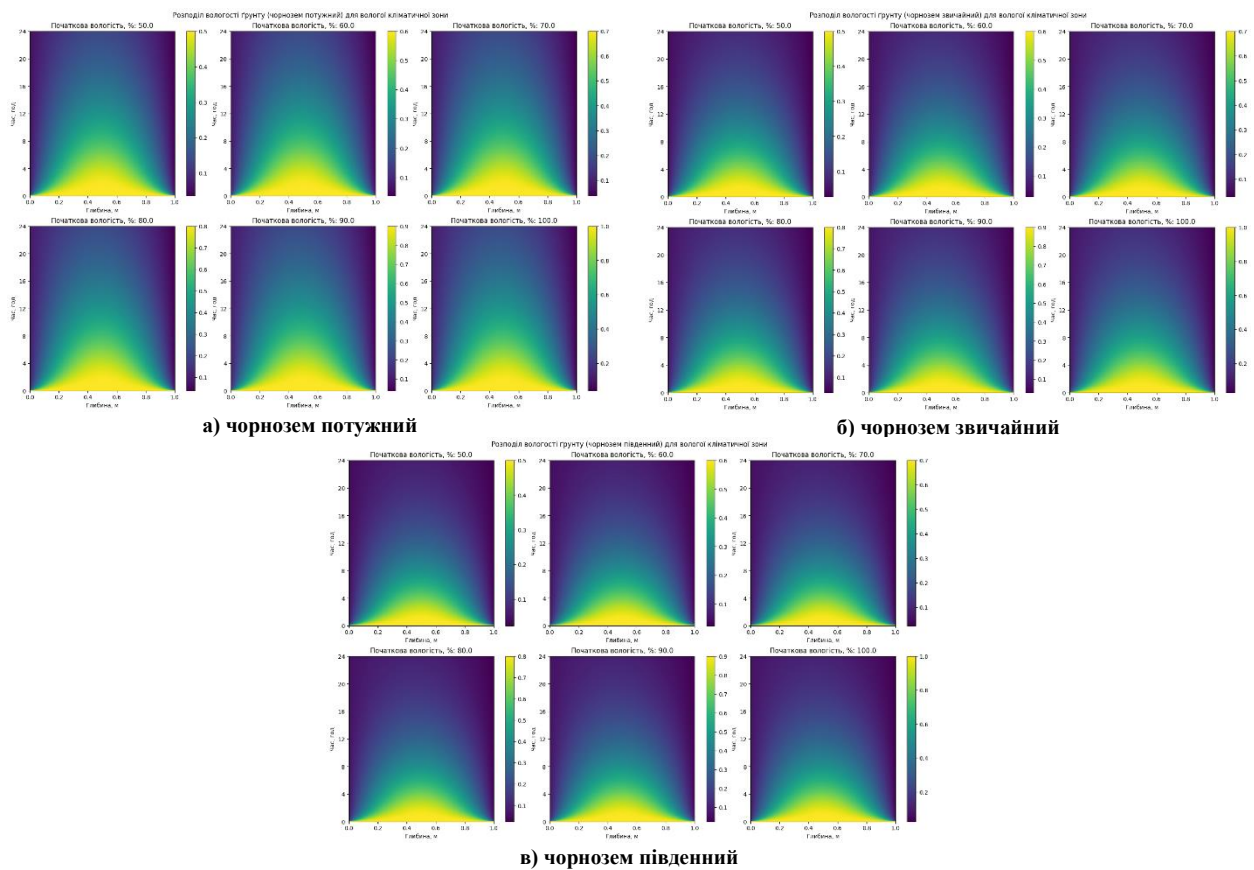
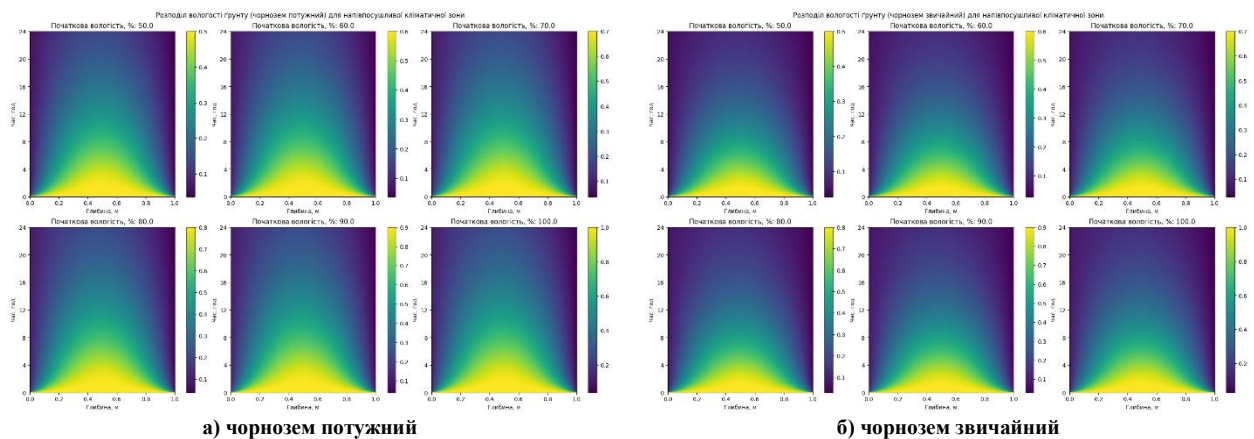
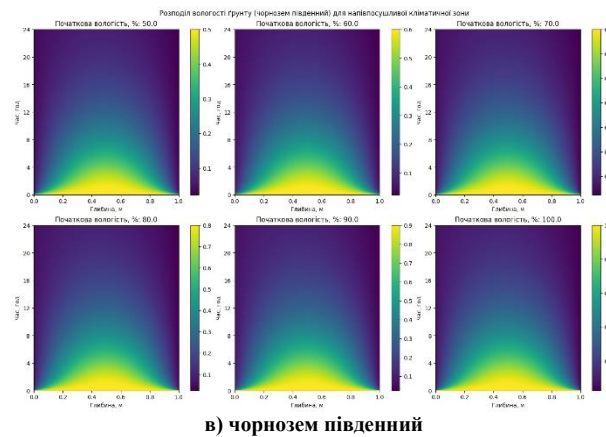


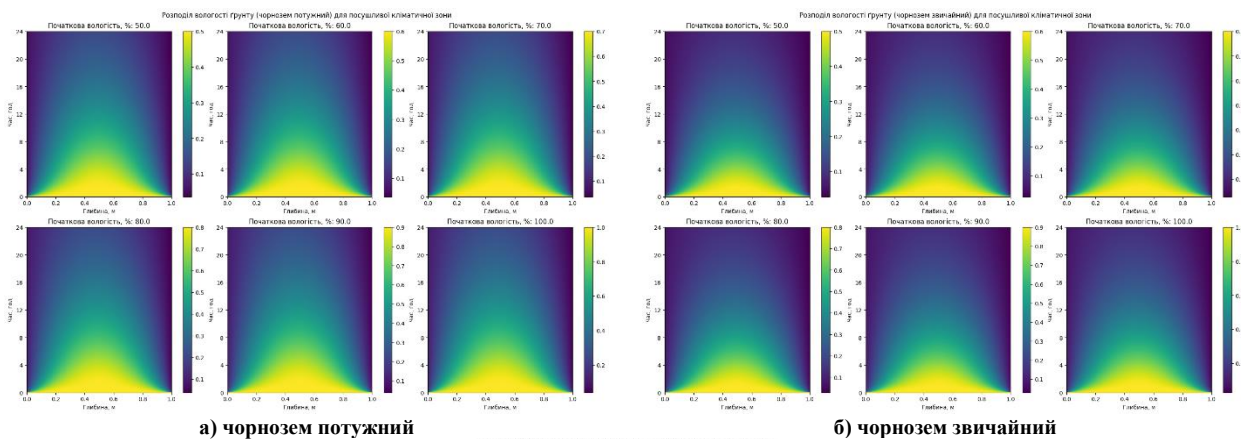
Рис. 6. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах вологій агрокліматичної зони за добу





в) чорнозем південний

Рис. 7. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах напівпосушливої агрокліматичної зони за добу



а) чорнозем потужний

б) чорнозем звичайний

в) чорнозем південний

Рис. 8. Графічна інтерпретація розповсюдження вологи в ґрунті в умовах посушливої агрокліматичної зони за добу

Одержані результати математичного та комп'ютерного моделювання динаміки вологи в ґрунті на проміжку часу, що дорівнює одній добі (див. рис. 6–8), дозволили встановити значний вплив факторів типу ґрунту та агрокліматичної зони вирощування на вологісний режим вирощування с/г культур, а саме: граничний показник об'ємної вологості для чорнозему потужного як функція початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 3,5–4,0 год. для вологої агрокліматичної зони й через 2,5–3,0 год. для посушливої та напівпосушливої агрокліматичних зон; граничний показник об'ємної вологості для чорнозему звичайного як функція початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 2,0–2,5 год. для всіх аналізованих агрокліматичних зон; граничний показник об'ємної вологості для чорнозему південного як функція початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 2,0 год. для вологої агрокліматичної зони й через 1,5–2,0 год. для посушливої та напівпосушливої агрокліматичних зон.

Отже, одержані результати досліджень щодо годинної (див. рис. 3–5) і добової (див. рис. 6–8) динаміки розповсюдження вологи в різних типах ґрунту і агрокліматичних умовах можуть бути використані

в якості функціонально-алгоритмічного базису під час подальших перспективних досліджень із розробки, проєктування та валідації програмних компонент комп'ютерної інтелектуальної технології предикативного контролю режимів зрошення зернових культур.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

На основі проведених досліджень в цій статті було досягнуто її основу мету, а саме, розроблено та валідовано математичну та комп'ютерну моделі процесу комп'ютеризованого предикативного контролю режимів зрошення ґрунту під час вирощування зернових культур у агрокліматичних умовах України. Основні кількісні та якісні результати розробки та досліджень статті полягають у наступному:

1. Запропоновано та алгоритмізовано математичний опис процесу динаміки вологи в різних типах ґрунту з урахуванням специфічних агрокліматичних зон України, у яких може бути використана інтелектуальна комп'ютерна технологія предикативного контролю режимів зволоження. Одержані результати досліджень дозволили оцінити характеристики та параметру процесу розповсюдження вологи за шарами ґрунту в часовому діапазоні, що дорівнює одній годині, а саме: для чорнозему потужного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,17 м до 0,83 м для всіх агрокліматичних зон, що аналізувалися; для чорнозему звичайного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,22 м до 0,78 м для всіх агрокліматичних зон, що аналізувалися; для чорнозему південного значення об'ємної вологості, яке було встановлено в початковий момент аналізу на поверхневому шарі, через годину спостерігається на глибині від 0,26 м до 0,74 м для всіх агрокліматичних зон, що аналізувалися.

2. Методами комп'ютерного експерименту оцінено параметри й характеристики процесу добової динаміки вологи в різних типах ґрунту, що дозволило обґрунтувати необхідність врахування факторів агрокліматичної зони вирощування зернових культур та типу ґрунту, а саме: граничне значення об'ємної вологості для чорнозему потужного як функції початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 3,5–4,0 год. для вологої агрокліматичної зони й через 2,5–3,0 год. для посушливої та напівпосушливої агрокліматичних зон; граничне значення об'ємної вологості для чорнозему звичайного як функції початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 2,0–2,5 год. для всіх аналізованих агрокліматичних зон; граничне значення об'ємної вологості для чорнозему південного як функції початкової вологості поверхневого шару ґрунту (варіюється від 50 % до 100 %) детектується на глибині 0,5 м через 2,0 год. для вологої агрокліматичної зони й через 1,5–2,0 год. для посушливої та напівпосушливої агрокліматичних зон.

3. Отримані результати математичного та комп'ютерного моделювання довели, що вони можуть бути використані в якості програмно-алгоритмічного та структурно-функціонального базису під час подальших розробок і досліджень з проєктування та тестування інтелектуальної комп'ютерної технології предикативного контролю режимів зрошення зернових культур у різних агрокліматичних зонах України.

### ВПРОВАДЖЕННЯ І ВИКОРИСТАННЯ ОТРИМАНИХ РЕЗУЛЬТАТІВ

Дослідження статті виконані в рамках науково-дослідної теми «Розвиток програмно-апаратного забезпечення інтелектуальних технологій для сталого вирощування сільськогосподарських культур у воєнний та повоєнний час» (номер держреєстрації 0124U000289).

### Література

1. FAOSTAT: Food and agriculture organization of the united nations. URL: <https://www.fao.org/faostat/en/#data/QCL> (дата звернення 05.06.2024).

2. АГРОСЕПМАШ: Зернові культури – основні види, популярні сорти та умови вирощування. URL: <https://agrosepmash.ua/uk/zemovi-kulturi-osnovni-vidi-populyarni-sorti-ta-umovi-viroshhuvannya/> (дата звернення 05.06.2024).

3. Лозовіцький П.С. Водні та хімічні меліорації ґрунтів. Навчальний посібник. Київ: Видавничо-поліграфічний центр “Київський університет”, 2010. 276 с.

4. Таврійський державний агротехнологічний університет імені Дмитра Моторного: Лекція. Водні властивості та водний режим ґрунту. URL: <http://www.tsatu.edu.ua/rosl/wp-content/uploads/sites/20/lekcija-9.vodni-vlastyvoli-ta-vodnyj-rezhym-hruntu.pdf> (дата звернення 25.07.2024).

5. Tracy F.T. Analytical and Numerical Solutions of Richards' Equation with Discussions on Relative Hydraulic Conductivity. *Hydraulic Conductivity – Issues, Determination and Applications*. InTech, 2011. P. 203–222.

6. Колесников В.А., Ляшко С.І. Стійкість розв'язків рівняння Річардса–Клюта. *Доповіді Національної академії наук України*. 2023. № 6. С. 12–18.

7. Солоня О.В., Ковбаса В.П. Обґрунтування параметрів робочих органів для укладання

внутрішньогрунтових зрошувачів: монографія. Вінниця, 2020. 155 с.

8. Brady N.C., Weil R.R. *The Nature and Properties of Soils: 14<sup>th</sup> edition*. Columbus: Pearson Education, 2008. 965 p.

9. Allen R.G., Pereira L.S., Raes D., Smith M. *Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements*. Rome: FAO, 1998. 300 p.

10. EOS DATA ANALYTICS: Евапотранспірація як процес та методи її вимірювання. URL: <https://eos.com/uk/blog/evapotranspiratsiia/> (дата звернення 31.07.2024).

11. Карти України: Фільтрація ґрунтів. URL: <https://geomap.land.kiev.ua/soil-6.html> (дата звернення 31.07.2024).

12. Велика Українська Енциклопедія: Аерація ґрунту. URL: [https://vue.gov.ua/Аерація\\_ґрунту](https://vue.gov.ua/Аерація_ґрунту) (дата звернення 31.07.2024).

### References

1. FAOSTAT: Food and agriculture organization of the united nations. Retrieved from: <https://www.fao.org/faostat/en/#data/QCL>
2. AGROSEPMASH: Cereals - main types, popular varieties and growing conditions. Retrieved from: <https://agrosepmash.ua/uk/zernovi-kulturi-osnovni-vidi-populyarni-sorti-ta-umovi-viroshhuvannya/> [in Ukrainian].
3. Lozovitskyi, P.S. (2010). *Vodni ta khimichni melioratsii gruntiv. Navchalnyi posibnyk [Water and chemical soil reclamation. Textbook]*. Kyiv. 276 p.
4. Tavriiskiy derzhavnyi ahrotekhnolohichnyi universytet imeni Dmytra Motornoho: Lektsiia. *Vodni vlastyvoli ta vodnyi rezhym gruntu [Dmytro Motorny Tavria State Agrotechnological University: Lecture. Water properties and water regime of soil]*. Retrieved from: <http://www.tsatu.edu.ua/ros/ wp-content/uploads/sites/20/lekcija-9.vodni-vlastyvoli-ta-vodnyj-rezhym-hruntu.pdf> [in Ukrainian].
5. Tracy, F.T. (2011). Analytical and Numerical Solutions of Richards' Equation with Discussions on Relative Hydraulic Conductivity. *Hydraulic Conductivity – Issues, Determination and Applications*, 203–222.
6. Kolesnykov, V.A., & Liashko, S.I. (2023). Stiiikist rozviazkiv rivniannia Richardsa–Kliuta [Stability of solutions to the Richards-Kluyt equation]. *Reports of the National Academy of Sciences of Ukraine*, 6, 12–18.
7. Solona, O.V., & Kovbasa, V.P. (2020) *Obgruntuvannia parametriv robochykh orhaniv dlia ukladannia vnutrishnogruntovykh zroshuvachiv: monohrafiia [Substantiation of parameters of working bodies for laying in-soil irrigators: monograph]*. Vinnytsia. 155 p.
8. Brady, N.C., & Weil, R.R. (2008). *The Nature and Properties of Soils: 14<sup>th</sup> edition*. Columbus. 965 p.
9. Allen, R.G., Pereira, L.S., Raes, D., & Smith, M. (1998). *Crop Evapotranspiration – Guidelines for Computing Crop Water Requirements*. Rome. 300 p.
10. EOS DATA ANALYTICS: Evapotranspiratsiia yak protses ta metody yii vymiriuvannia [Evapotranspiration as a process and methods of its measurement]. Retrieved from: <https://eos.com/uk/blog/evapotranspiratsiia/> [in Ukrainian].
11. Karty Ukrainy: Filtratsiia gruntiv [Maps of Ukraine: Soil filtration]. Retrieved from: <https://geomap.land.kiev.ua/soil-6.html> [in Ukrainian].
12. Velyka Ukrainska Entsyklopediia: Aeraatsiia gruntu [The Great Ukrainian Encyclopedia: Soil aeration]. Retrieved from: [https://vue.gov.ua/Аерація\\_ґрунту](https://vue.gov.ua/Аерація_ґрунту) [in Ukrainian].



<https://doi.org/10.31891/2219-9365-2024-79-19>

УДК 681.324

МУСІЄНКО Андрій

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

<https://orcid.org/0000-0002-1849-6716>

e-mail: [mysienkoandrey@gmail.com](mailto:mysienkoandrey@gmail.com)

ВОРВУЛЬ Данило

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

e-mail: [vorvul.danylo@gmail.com](mailto:vorvul.danylo@gmail.com)

БЕЗСМЕРТНИЙ Владислав

Київський національний університет імені Тараса Шевченка

e-mail: [v.y.bessmertniy@gmail.com](mailto:v.y.bessmertniy@gmail.com)

## ОЦІНКА ПРОДУКТИВНОСТІ ГОТОВИХ РІШЕНЬ ПОШУКОВО-ДОПОВНЕНОЇ ГЕНЕРАЦІЇ ДЛЯ НАДАННЯ РЕКОМЕНДАЦІЇ З ВИБОРУ НАЙКРАЩОГО РІШЕННЯ

В даній статті представлено процес розробки застосунка для надання комплексної оцінки існуючих рішень з розширеним пошуком (Retrieval-Augmented Generation) з акцентом на їхню ефективність у фінансовій сфері, зокрема для компаній, що займаються злиттям та поглинанням. Системи RAG, які поєднують механізми пошуку з можливостями генерації, продемонстрували перспективність у наданні точних і контекстно-релевантних відповідей завдяки використанню великих обсягів даних, специфічних для конкретної галузі. Ми створили програмний код, який автоматично оцінює відповіді та дозволяє створити таблицю лідерів для порівняння цих рішень за тринадцятьма ключовими показниками ефективності. Було зібрано датасет з питань та відповідей від дійсних експертів галузі, оцінено здатність кожного рішення обробляти структуровані та неструктуровані фінансові дані. Також порівняно оцінки від експертів та великих мовних моделей, зроблено висновок про ефективність дослідження. Отримані результати висвітлюють сильні і слабкі сторони існуючих систем RAG, надають уявлення про їхню застосовність та потенціал для покращення процесів прийняття рішень у фінансовому секторі. Це дослідження має на меті надання допомоги організаціям у виборі найбільш відповідного рішення RAG для їхніх потреб, а також надати інформацію про майбутні зміни в цій сфері, що швидко розвивається.

Ключові слова: Пошуково-доповнена генерація, оцінка ефективності, штучний інтелект, великі мовні моделі, фінансова сфера, інформаційні технології

MUSIENKO Andrii, VORVUL Danylo, BEZSMERTNIY Vladyslav

National Technical University of Ukraine "Igor Sikorsky Polytechnic Institute", Taras Shevchenko National University of Kyiv

## EVALUATING THE PERFORMANCE OF OFF-THE-BOX RETRIEVAL AUGMENTED GENERATION SOLUTIONS TO PROVIDE RECOMMENDATIONS ON CHOOSING THE BEST SOLUTION

This article presents the process of developing an application to provide a comprehensive assessment of existing Retrieval-Augmented Generation (RAG) solutions with a focus on their effectiveness in the financial sector, particularly for M&A companies. RAG systems, which combine search mechanisms with generation capabilities, have shown promise in providing accurate and contextually relevant answers using large amounts of industry-specific data. We created software code that automatically scores the answers and creates a league table to compare these solutions across thirteen key performance indicators. A dataset of questions and answers from actual industry experts was collected, and the ability of each solution to process structured and unstructured financial data was assessed. We also compared the scores from experts and large language models, and concluded on the effectiveness of the study. The findings highlight the strengths and weaknesses of existing RAG systems, provide insight into their applicability and potential to improve decision-making in the financial sector. This study aims to assist organisations in choosing the most appropriate RAG solution for their needs, as well as provide information on future developments in this rapidly evolving field.

The study demonstrates the significant potential of advanced search systems in the field of mergers and acquisitions. The results obtained allow organizations to choose the most suitable RAG solution for their needs, increasing the accuracy and relevance of answers to complex queries, which in turn improves decision-making processes in the financial sector.

This study makes a significant contribution to the development of knowledge about RAG technologies, especially in the context of their application in the financial sector. The results are an important step in understanding and implementing these technologies, which will help provide more accurate, relevant and useful answers to complex financial queries.

Keywords: Search-Augmented Generation, performance evaluation, Artificial Intelligence, Large Language Models, financial sector.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Поява передових технологій штучного інтелекту зробила революцію в багатьох галузях, серед яких фінанси є одним із секторів, що зазнали найбільшого впливу. Однією з передових методологій ШІ, що з'явилася [1-6], є RAG, яка поєднує в собі сильні сторони пошуку інформації та генерації тексту. Системи

RAG використовують специфічні для конкретної галузі набори даних для надання високоточних, контекстно-релевантних відповідей, що робить їх безцінними для завдань, які вимагають тонкого розуміння та точного отримання інформації.

У фінансовій сфері, зокрема у сфері злиття та поглинання, застосунки RAG можуть значно покращити процеси прийняття рішень. Ці системи можуть швидко отримувати та генерувати інформацію з великих масивів фінансових даних, що має вирішальне значення для таких видів діяльності, як аналіз ринку, оцінка ризиків та дотримання нормативних вимог. У контексті злиття та поглинання, де точність і своєчасність мають першорядне значення, системи RAG можуть допомогти в проведенні комплексної перевірки, оціночного аналізу та інтеграційного планування, синтезуючи складну фінансову інформацію, надаючи дієві висновки.

Незважаючи на перспективність рішень RAG, існує нагальна потреба систематично оцінювати їхню ефективність у різних сферах застосування у фінансовій галузі. Існуючі дослідження [7, 8] часто зосереджуються на широких реалізаціях, які не можуть бути застосовані під конкретні юзкейси. Ця стаття має на меті заповнити цю прогалину, розробивши програмну платформу, яка вимірює продуктивність декількох рішень RAG в сфері злиття та поглинання. Створивши стандартизовану таблицю лідерів, ми надаємо порівняльний аналіз цих рішень, виділяючи їхні сильні сторони та сфери, що потребують вдосконалення.

Наше дослідження фокусується на тринадцяти ключових показниках ефективності. Ці показники мають вирішальне значення для оцінки практичної корисності систем RAG у реальних фінансових додатках.

Мета цього дослідження є подвійною. По-перше, ми прагнемо надати чітке, засноване на даних порівняння існуючих рішень RAG, щоб допомогти організаціям вибрати найбільш відповідну систему для їхніх конкретних потреб. По-друге, ми прагнемо виявити обмеження та потенційні можливості для вдосконалення існуючих технологій RAG, тим самим надаючи інформацію для майбутніх досліджень і розробок у цій галузі, що швидко розвивається.

Таким чином, цей документ робить внесок у зростаючий обсяг знань про технології RAG, пропонуючи комплексну систему оцінки, яка охоплює сферу компаній з фокусом на злиття та поглинання. Наші висновки допоможуть зацікавленим сторонам приймати обґрунтовані рішення щодо впровадження та розвитку штучного інтелекту, що в кінцевому підсумку розширить можливості та підвищить ефективність рішень RAG у фінансовому секторі.

## АНАЛІЗ ЛІТЕРАТУРНИХ ДАНИХ ТА ПОСТАНОВКА ПРОБЛЕМИ

Щоденні задачі фінансових експертів полягають у пошуку інформації про компанії та їхній стан, узагальненні цієї інформації та її аналізі. Ефективність цієї роботи безпосередньо впливає на прийняття інвестиційних рішень, розробку фінансових стратегій. Вибір системи для інтеграції в бізнес-процеси компанії є надзвичайно важливим кроком і не може відбуватися без належного етапу тестування. У фінансовій сфері навіть передові великі мовні моделі не можуть досягти точності понад 90% [8]. Тому багато компаній проводять дослідження для своїх специфічних випадків використання. Наразі відсутні дослідження, які б оцінювали ефективність RAG систем у сфері злиттів та поглинань компаній. Проте існує багато досліджень які міряють ефективність великих мовних моделей у відповідях на питання. Зазвичай ці бенчмарки базуються на неспеціалізованих датасетах в конкретній області [9, 10, 11].

Було проаналізовано останнє дослідження, яке максимально близьке до нашого - FinanceBench [8]. Це перший у своєму роді набір тестів для оцінювання ефективності великих мовних моделей з відповіді на фінансові питання. В їх бенчмарку представлено 2 493 питання які передбачають виключно вилучення інформації (28%), 5 897 питань передбачають числові міркування (66%) та 518 (6%) передбачають логічні міркування. На основі цього дослідження можна зробити висновок, що великі мовні моделі набагато гірше справляються з фінансовими питаннями, якщо не мають доступ до додаткової інформації (20% точності проти 79% для GPT-4Turbo). Тому наше дослідження буде базуватись лише на RAG системах. Також FinanceBench дослідження підкреслює лімітації великих мовних моделей при міркуванні з цифрами. Тому при зборі датасету для нашого дослідження було прийнято рішення зосередитись на питаннях з основною задачею пошуку інформації та нескладних логічних міркувань.

Таблиця 1

### Порівняння інструментів для оцінки ефективності систем RAG

Фреймворк	Час	Оцінка Метрик	Пошук	Генерація
LangChain Benchmark [16]	Листопад 2023	Точність, Правдивість, Час Виконання, Косинусна Відстань Вбудовування	Точність	ШІ як Оцінювач
Databricks Assessment [17]	Грудень 2023	Коректність, Читальність, Повнота, Релевантність Контексту, Релевантність Відповіді, Обґрунтованість	-	ШІ як Оцінювач
TruEra RAG Group [18]	Жовтень 2023	Релевантність Контексту, Релевантність Відповіді, Обґрунтованість	ШІ як Оцінювач	ШІ як Оцінювач

Існує багато інструментів та бенчмарків для оцінки ефективності систем RAG. Зазвичай, оцінювання ефективності поділяється на автоматичне та експертне (проводиться людьми). Автоматичний підхід зазвичай передбачає використання великої мовної моделі як оцінювача, причому цей метод демонструє високу кореляцію з уподобаннями людей (80%) [12, 13]. Основною перевагою автоматичного оцінювання є його здатність до масштабування, оскільки неможливо залучати експертів для оцінки нової версії системи після кожного покращення.

Проте не можна нехтувати оцінкою від експертів, особливо в нашому випадку, де попередніх досліджень ще не проводилося. Це дозволить зрозуміти, наскільки добре система відповідає потребам експертів.

Для автоматичної оцінки було обрано кастомну реалізацію фреймворку RGAR (Retrieval, Generation, Additional Requirement) [13] завдяки його найширшій варіативності використаних метрик. Цей підхід забезпечує всебічний аналіз продуктивності системи, дозволяючи охопити різноманітні аспекти її роботи та забезпечити більш точну та комплексну оцінку рис.1.

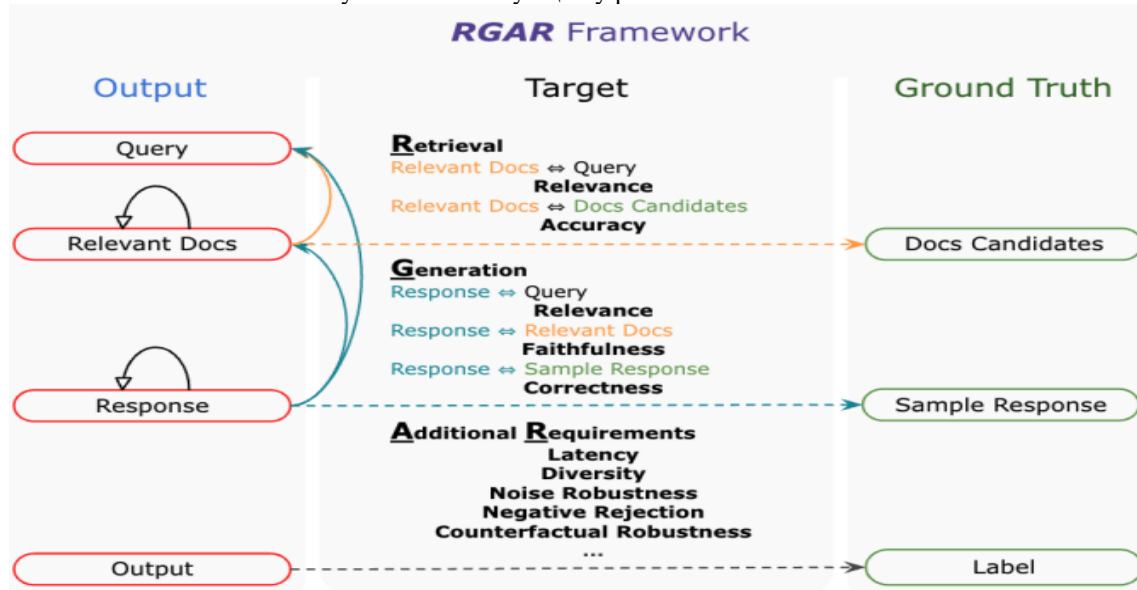


Рис. 1. Схема реалізації RGAR фреймворку

Як було встановлено, у відкритому доступі відсутні дослідження, присвячені доменній області компаній, що займаються злиттям та поглинанням. Зазвичай такі компанії при виборі рішень для своїх бізнес-процесів змушені орієнтуватися на продуктивність великих мовних моделей або RAG систем у відповідях на загальні та фінансові питання.

На основі аналізу наявних матеріалів можна окреслити проблему, яка вирішує наше дослідження: відсутність досліджень ефективності RAG систем у сфері діяльності компаній, що займаються злиттям та поглинанням.

#### Цілі та завдання дослідження

Цілями даного дослідження є:

1. Аналіз можливостей систем Retrieval-Augmented Generation (RAG) у наданні точних відповідей на запитання шляхом використання великих мовних моделей для вилучення інформації (включаючи добування певних даних або текстового контенту з файлів) та здійснення нескладних логічних міркувань у домені компаній, що займаються злиттям та поглинанням.
2. Розробка рейтингової таблиці (лідерборду) існуючих рішень RAG. Для досягнення цілей було поставлено наступні завдання:
  1. Збір датасету типових запитань та еталонних відповідей для компаній, що займаються злиттям та поглинанням.
  2. Налаштування досліджуваних систем, та запуск процесу генерації відповідей на питання.
  3. Імплементція фреймворку для оцінювання результатів відповідей системи.
  4. Створення лідерборду, де кращі відповіді отримують вищі оцінки.

#### Матеріали та методи дослідження

Для вирішення поставлених задач було створено комплексний додаток для оцінювання відповідей, який інтегрується з наявним програмним забезпеченням на Python. Цей додаток здатний автоматично

оцінювати продуктивність різних систем з розширеним пошуком. Нижче наведено детальну діаграму послідовностей роботи системи:

sequenceDiagram

учасник FinancialExpert як Фінансовий експерт

учасник Dataset як Набір даних

учасник RAG як RAG

учасник DB як База даних

учасник EvalApp як Додаток для оцінювання

Фінансовий експерт->Набір даних: Задати питання

loop for every система RAG

Набір даних->RAG: поставити запитання

RAG->База даних: відповідь у журналі, отримані фрагменти,

затримка

База даних->Додаток для оцінювання: Надання журналів

Фінансовий експерт->База даних: оцінка згенерованих

відповідей

Додаток для оцінювання->База даних: оцінити отримані

фрагменти за метриками P@K, Релевантність, Точність, Джерела

Додаток для оцінювання->База даних: оцінити згенеровані

відповіді за показниками Стійкість до контрфактів, Негативне відхилення, Інтеграція інформації, Контрфактична стійкість, Узгодженість, Релевантність

Додаток для оцінювання->База даних: Створити звіт на основі

зібраних метрик

кінець

### Збір датасету

Для збору датасету було залучено трьох осіб з досвідом роботи в Private Equity фондах, кожен з яких має відповідний досвід у сфері компаній, що займаються злиттям та поглинанням. Основна діяльність фінансових експертів у цільових компаніях полягає в обробці нотаток щодо угод. Угоди включають декілька стадій; у нашому випадку було розглянуто дві найбільш значущі стадії:

- Approval - остання нотатка перед придбанням компанії, яка використовується для прийняття рішення.

- Closing - стадія, на якій компанія була придбана.

Було зібрано 32 файли з нотатками про угоди купівлі п'яти різних компаній. Для кожної угоди купівлі компанії було зібрано один файл зі стадії Approval, один зі стадії Closing та декілька супутніх матеріалів із додатковою фінансовою інформацією про діяльність компанії. На основі інформації, що розміщена у файлах, експертами було створено 40 питань та надано еталонні відповіді на ці питання.

### Налаштування досліджуваних систем та запуск процесу генерації відповідей на питання

У цьому розділі ми розглянемо кілька передових RAG-інструментів, зосереджуючи увагу на їхній ефективності у фінансовій сфері. До них належать Azure + AI Search з GPT-3.5/GPT-4, AWS Q, Amazon Bedrock Agents, Plaibox та Nuclia. Кожна з цих систем надає можливості для створення RAG-систем, які слугують базовими рішеннями для подальшої кастомізації та оптимізації під конкретні випадки використання. У кожній з цих систем було створено базу знань та завантажено всі наявні документи. Крім того, написано скрипт, який зчитує датасет із запитаннями, задає їх кожній системі та зберігає відповіді. Розроблений скрипт також зберігає посилання на документи, які були використані для генерації відповідей, та чанки тексту.

### Імплементація фреймворку оцінювання

Для оцінки якості Retrieval частини було розроблено застосунок який може оцінити відповіді за наступними метриками:

- Міра точності при К (P@K) є одним з показників, що використовується в інформаційному пошуку для оцінки релевантності результатів, які були отримані за запитом або рекомендаційною системою серед перших К позицій. Цей показник є простим, але потужним засобом оцінювання, що дозволяє оцінити точність системи лише на перших К результатах, надаючи інформацію про ефективність системи у виявленні релевантних елементів серед найвищих позицій у її вихідних даних.

Формула для обчислення точності при K (Precision at K, P@K) виглядає наступним чином:

$$P @ k(q) = \# \{ \text{relevant items on the top } k \text{ positions} \} \div k \quad (1)$$

- Релевантність (Релевантні документи / Запит) оцінює, наскільки знайдені документи відповідають необхідній інформації, вираженій у запиті. Він вимірює точність та специфічність процесу пошуку.

- Точність (Accuracy) є показником, що відображає пропорцію правильних результатів (як істинно позитивних, так і істинно негативних) серед загальної кількості розглянутих випадків. Вона показує, наскільки правильно модель класифікує випадки в різних категоріях, та використовується для оцінки загальної ефективності моделі.

Формула для обчислення точності виглядає наступним чином:

$$Accuracy = \frac{\{TP + TN\}}{\{TP + TN + FP + FN\}} \quad (2)$$

де:

- (TP) (True Positives) - кількість випадків, коли модель правильно визначила позитивний результат.

- (TN) (True Negatives) - кількість випадків, коли модель правильно визначила негативний результат.

- (FP) (False Positives) - кількість випадків, коли модель помилково визначила позитивний результат.

- FN (False Negatives) - кількість випадків, коли модель помилково визначила негативний результат.

- Джерела - наявність посилання на джерело, звідки отримана інформація, щоб можна було швидко перевірити її точність.

Для оцінки Generation частини було обрано наступні метрики спираючись на запропоновані RGAR фреймворку [12]:

- Стійкість до шуму [14, 15]. Цей показник оцінює здатність RAG системи зберігати продуктивність у присутності шуму, нерелевантної або оманливої інформації під час фази пошуку. Модель з високою стійкістю до шуму може ефективно фільтрувати та ігнорувати таку інформацію, зосереджуючись на релевантних даних для генерації точних та зрозумілих відповідей. Стійкість до шуму є критично важливою для забезпечення того, щоб використання зовнішніх механізмів пошуку не знижувало якість результатів моделі.

- Відхилення негативних результатів [14, 15]. Цей показник вимірює здатність RAG системи виявляти та виключати помилково позитивні або неправильно отримані дані, які не відповідають суті запиту. Ця здатність є надзвичайно важливою для запобігання поширенню помилок або дезінформації в згенерованому контенті, забезпечуючи, щоб процес доповнення покращував, а не погіршував якість результатів моделі.

- Інтеграція інформації [14, 15]. Цей показник вимірює, наскільки добре RAG система може синтезувати та включати отриману інформацію з різних джерел у зв'язну та контекстуально відповідну відповідь. Цей показник підкреслює здатність моделі ефективно поєднувати різномірні дані, що є важливим для створення всебічних та нюансованих відповідей, які відображають глибоке розуміння запиту.

- Стійкість до контрфактів [14, 15]. Стійкість до контрфактів відноситься до здатності LLM обробляти ситуації, що вимагають розуміння та застосування знань, які суперечать загальноприйнятим або очікуваним. Це тестує здатність моделі міркувати та генерувати результати, що відповідно відповідають гіпотетичним або контрфактичним ситуаціям, що свідчить про високий рівень когнітивної гнучкості та креативності.

- Узгодженість [12]. Узгодженість оцінює здатність великої мовної моделі (LLM) створювати зв'язні та логічні відповіді, що є зрозумілими та послідовними. Модель з високим рівнем узгодженості забезпечує, що кожне речення гармонійно вписується в загальний контекст відповіді, не суперечачи одне

одному і утворюючи цілісну картину. Узгодженість є важливою для створення змістовних та зрозумілих відповідей.

- Відповідність [12]. Відповідність вимірює, наскільки добре модель може генерувати інформацію, що безпосередньо стосується заданого запиту. Висока відповідність означає, що модель надає релевантну інформацію, яка точно відповідає на поставлені питання або задовольняє інформаційні потреби користувача. Відповідність критична для забезпечення корисності та точності відповідей.

- Обґрунтованість [12]. Обґрунтованість оцінює, наскільки добре модель базує свої відповіді на достовірних джерелах інформації. Це включає в себе перевірку фактів та надання підтверджувальних доказів для своїх тверджень. Модель з високою обґрунтованістю здатна зменшити ризик поширення дезінформації, забезпечуючи точність та надійність створених відповідей.

- Оцінка людиною. Оцінка людиною передбачає оцінювання відповідей моделі експертами або звичайними користувачами для визначення якості, зрозумілості та точності. Цей метод забезпечує суб'єктивну перевірку продуктивності моделі, враховуючи людське сприйняття і досвід. Оцінка людиною є важливим інструментом для вдосконалення моделі та забезпечення її відповідності очікуванням користувачів. Нашим експертам було запропоновано переглянути отримані відповіді і оцінити їх якість від 1 до 5.

- Затримка. Вимірювалась в секундах в залежності від швидкості конкуруючих систем. Усі оцінки, за винятком Оцінки людиною, Затримки та оцінок Retrieval, були здійснені за допомогою великих мовних моделей. Для кожної метрики було розроблено відповідний запит, у якому моделі пропонувалося оцінити, наскільки добре оцінка відповідає критерію, за шкалою від 1 до 5.

#### Приклад промпту:

System:

You are an AI assistant. You will be given the definition of an evaluation metric for assessing the quality of an answer in a question-answering task. Your job is to compute an accurate evaluation score using the provided evaluation metric.

User:

Relevance measures how well the answer addresses the main aspects of the question, based on the context. Consider whether all and only the important aspects are contained in the answer when evaluating relevance. Given the context and question, score the relevance of the answer between one to five stars using the following rating scale:

One star: the answer completely lacks relevance

Two stars: the answer mostly lacks relevance

Three stars: the answer is partially relevant

Four stars: the answer is mostly relevant

Five stars: the answer has perfect relevance

This rating value should always be an integer between 1 and 5. So the rating produced should be 1 or 2 or 3 or 4 or 5.

context: Marie Curie was a Polish-born physicist and chemist who pioneered research on radioactivity and was the first woman to win a Nobel Prize.

question: What field did Marie Curie excel in?

answer: Marie Curie was a renowned painter who focused mainly on impressionist styles and techniques.

stars: 1

context: The Beatles were an English rock band formed in Liverpool in 1960, and they are widely regarded as the most influential music band in history.

question: Where were The Beatles formed?

answer: The band The Beatles began their journey in London, England, and they changed the history of music.

stars: 2

context: The recent Mars rover, Perseverance, was launched in 2020 with the main goal of searching for signs of ancient life on Mars. The rover also carries an experiment called MOXIE, which aims to generate oxygen from the Martian atmosphere.

question: What are the main goals of Perseverance Mars rover mission?

answer: The Perseverance Mars rover mission focuses on searching for signs of ancient life on Mars.

stars: 3

context: The Mediterranean diet is a commonly recommended dietary plan that emphasizes fruits, vegetables, whole grains, legumes, lean proteins, and healthy fats. Studies have shown that it offers numerous health benefits, including a reduced risk of heart disease and improved cognitive health.

question: What are the main components of the Mediterranean diet?

answer: The Mediterranean diet primarily consists of fruits, vegetables, whole grains, and legumes.

stars: 4

context: The Queen's Royal Castle is a well-known tourist attraction in the United Kingdom. It spans over 500 acres and contains extensive gardens and parks. The castle was built in the 15th century and has been home to generations of royalty.

question: What are the main attractions of the Queen's Royal Castle?

answer: The main attractions of the Queen's Royal Castle are its expansive 500-acre grounds, extensive gardens, parks, and the historical castle itself, which dates back to the 15th century and has housed generations of royalty.

stars: 5

context: {{context}}  
question: {{question}}  
answer: {{answer}}  
stars:

Your response must include following fields and should be in json format:

score: Number of stars based on definition above

reason: Reason why the score was given

### Створення лідерборду

Для створення лідерборду були використані логи роботи системи. Всі зібрані дані були зведені та обраховані за допомогою пакета pandas, що дозволило зручно організувати інформацію та створити формули для оцінки місця учасників на лідерборді. Цей підхід забезпечив можливість автоматичного оновлення та точного обчислення результатів на основі визначених критеріїв. Кожна метрика має максимальне значення 5 і мінімальне 0. Єдине виключення - метрики точності, але перед підрахунком рядка Totals їх було помножено на 5, щоб всі метрики мали однакову розмірність. Рішення відсортовані по максимальному балу зліва направо див. Таблиця 2. Порівняння кінцевих результатів та порівняння по метриках наведені на рис. 2 та рис. 3 відповідно.

Таблиця 2

### Фінальний лідерборд

Competitors	Azure GPT4	AWS Q	Plaibox (GPT4)	Nuclia (GPT4)	Azure GPT3.5	AWS Bedrock Agents (Claude 2.1)
Counterfactual Robustness	4,66	3,87	3,96	3,79	2,99	3,52
Negative rejection	3,74	3,92	4,29	4,2	4,65	3,83
Information Integration	4,71	4,48	3,04	3,11	4,1	3,57
Accuracy %	0,91	0,83	0,83	0,48	0,73	0,66
P@K	0,92	0,8	0,68	0,48	0,7	0,64
Coherence	4,36	3,87	2,87	2,87	3,07	2,98
Relevance	4,41	3,88	3,69	3,24	3,34	2,97
Groundedness	4,5	4,35	4,05	3,06	4,05	3,74
Human eval	4,43	3,39	3,91	3,48	3,48	4,17
Sources	4,27	2,92	2,13	4,05	3,83	3,14
Latency	3,64	2,73	4,55	4,55	1,82	0,91
<b>Total</b>	<b>47,88</b>	<b>41,53</b>	<b>40,03</b>	<b>38,85</b>	<b>38,48</b>	<b>35,35</b>

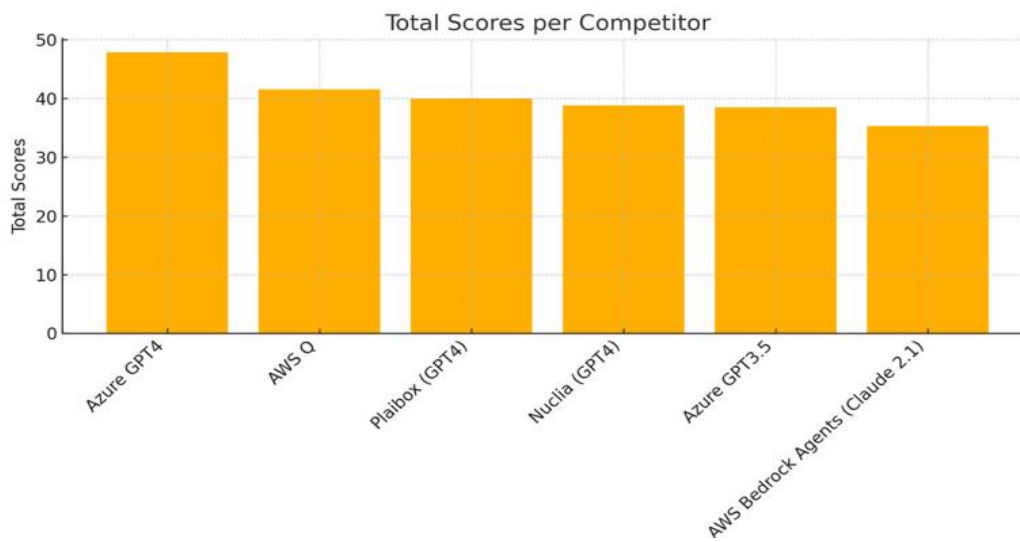


Рис. 2. Порівняння кінцевих результатів

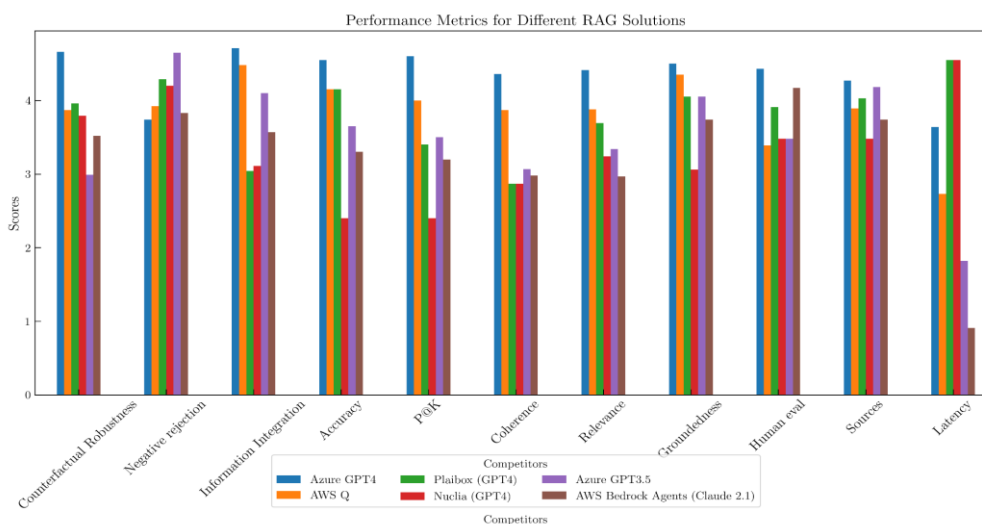


Рис. 3. Порівняння по метрикам

Було розраховано кореляцію Пірсона між середнім балом та оцінкою відповідей людиною:

$$r = \frac{\{n(\sum xy) - (\sum x)(\sum y)\}}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (3)$$

Ця формула обчислює коефіцієнт кореляції, який коливається від -1 до 1. Значення 1 вказує на ідеальну позитивну кореляцію, -1 вказує на ідеальну негативну кореляцію, а 0 вказує на відсутність кореляції. Кореляція між "Людською оцінкою" та "Середнім балом" становить приблизно 0,382. Це вказує на помірну позитивну кореляцію між цими двома змінними, якщо зібрати більше питань, ми впевнені що кореляція буде тільки збільшуватись.

### Обговорення результатів дослідження

Отримані результати дослідження свідчать, що серед протестованих рішень найкращі результати показала RAG система Azure OpenAI GPT-4 + Azure AI Search з відривом 15% від другого місця. Всі інші системи працювали на майже однаковому рівні. Ось ключові аспекти її ефективності:

1. Точність: Azure OpenAI GPT-4 + Azure AI Search має найвищий показник точності, що свідчить про її високу здатність надавати правильну інформацію. Це підтверджує її ефективність у фінансовій сфері, де точність даних є критично важливою.



2. Стійкість до контрафактів та інтеграція інформації: Оцінки 4.66 та 4.71 відповідно показують її силу в обробці гіпотетичних сценаріїв та синтезі інформації. Це підкреслює здатність системи справлятися з комплексними завданнями та надавати релевантні відповіді навіть у випадках з наявністю оманливої інформації.

3. Обґрунтованість: Оцінка 4.5 бала відображає ефективність системи у наданні відповідей, підкріплених знайденою інформацією. Це важливо для забезпечення довіри користувачів до отриманих відповідей.

4. Людська оцінка та затримка: Система має хороші показники людської оцінки (4.43) і затримки (3.64), що робить її надійним вибором для додатків, що працюють у режимі реального часу. Це свідчить про те, що результати, надані системою, були високо оцінені експертами, що підвищує її привабливість для кінцевих користувачів.

Також важливо зазначити, що ми отримали помірну кореляцію між людською оцінкою та середнім балом від оцінки великими мовними моделями. Це підтверджує застосовуваність автоматичного підходу до оцінки ефективності RAG систем, дозволяючи ефективно масштабувати процес оцінювання без значних втрат в якості.

Хоча вибірка експертів та питань не є достатньо великою, щоб вважати дослідження фундаментальним для сфери компаній, що займаються злиттям і поглинанням, отримані результати дозволяють застосовувати розроблену методологію для збору більшого датасету і компіляції більш фундаментальних результатів. Це відкриває можливості для подальших досліджень і розширення нашого розуміння ефективності RAG систем у фінансовому секторі. Цей крок був необхідний щоб зумовити необхідність виділення ресурсів на наступну стадію, бо залучення експертів до таких досліджень обходиться дуже дорого.

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

### I ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Дослідження демонструє значний потенціал систем з розширеним пошуком у сфері злиття та поглинання компаній. Отримані результати дозволяють організаціям обирати найбільш відповідне рішення RAG для своїх потреб, підвищуючи точність та релевантність відповідей на складні запити, що, в свою чергу, покращує процеси прийняття рішень у фінансовому секторі.

Основні результати:

1. Покращення прийняття рішень: Системи RAG можуть швидко отримувати та генерувати інформацію з великих масивів даних, що є критичним для таких видів діяльності, як аналіз ринку, оцінка ризиків та дотримання нормативних вимог. Це особливо важливо у контексті злиття та поглинання, де точність і своєчасність інформації є ключовими.

2. Стандартизована оцінка ефективності: Створення лідерборду на основі тринадцяти ключових показників ефективності дозволяє чітко порівнювати різні системи RAG. Такий підхід забезпечує комплексну оцінку їх практичної корисності у реальних фінансових додатках, що допомагає виявити сильні та слабкі сторони кожної системи.

3. Інтеграція нових технологій: Дослідження підкреслює важливість інтеграції систем RAG з фінансовими процесами для забезпечення конкурентних переваг на ринку. Використання таких систем дозволяє зменшити час на обробку інформації та підвищити її точність, що є критичним у швидкозмінних ринкових умовах.

4. Можливості для покращення: Виявлені обмеження існуючих систем RAG вказують на напрямки для майбутніх досліджень та розробок. Це стосується покращення обробки специфічних фінансових запитів, точності вилучення інформації, а також інтеграції складних логічних міркувань.

Практичне застосування:

1. Вибір відповідної системи: Результати дослідження надають організаціям цінну інформацію для вибору найбільш відповідного рішення RAG, що підвищує ефективність фінансових аналітиків.

2. Підготовка до майбутніх викликів: Оскільки технології RAG продовжують розвиватися, компанії мають можливість використовувати отримані знання для підготовки до майбутніх викликів та покращення своєї конкурентоспроможності.

Це дослідження робить значний внесок у розвиток знань про технології RAG, особливо в контексті їх застосування у фінансовому секторі. Отримані результати є важливим кроком у розумінні та впровадженні цих технологій, що допоможе забезпечити більш точні, релевантні та корисні відповіді на складні фінансові запити.

### Література

1. Kwiatkowski, T., et al. (2019). Natural Questions: A Benchmark for Question Answering Research. Transactions of the Association for Computational Linguistics, 7, 453-466.
2. Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv preprint arXiv:2005.11401.
3. Shao, Z., Gong, Y., Shen, Y., Huang, M., Duan, N., & Chen, W. (2023). Enhancing Retrieval-Augmented Large Language Models with Iterative Retrieval-Generation Synergy (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2305.15294>.
4. He, X., Tian, Y., Sun, Y., Chawla, N. V., Laurent, T., LeCun, Y., Bresson, X., & Hooi, B. (2024). G-Retriever: Retrieval-Augmented Generation for Textual Graph Understanding and Question Answering (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2402.07630>.
5. Mao, Y., Dong, X., Xu, W., Gao, Y., Wei, B., & Zhang, Y. (2024). FIT-RAG: Black-Box RAG with Factual Information and Token Reduction (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2403.14374>
6. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020, May 22). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv.org. <https://arxiv.org/abs/2005.11401>
7. Gao Y, Xiong Y, Gao X, et al. Retrieval-Augmented Generation for Large Language Models: A survey. arXiv.org. <https://arxiv.org/abs/2312.10997>. Published December 18, 2023.
8. Islam, P., Kannappan, A., Kiela, D., Qian, R., Scherrer, N., & Vidgen, B. (2023). FinanceBench: A New Benchmark for Financial Question Answering (Версія 1). arXiv. <https://doi.org/10.48550/ARXIV.2311.11944>
9. Daniel Khashabi, Sewon Min, Tushar Khot, Ashish Sabharwal, Oyvind Tafjord, Peter Clark, and Hannaneh Hajishirzi. 2020. UNIFIEDQA: Crossing Format Boundaries with a Single QA System. In Findings of the Association for Computational Linguistics: EMNLP 2020, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, Online, 1896–1907. <https://doi.org/10.18653/v1/2020.findings-emnlp.171>
10. Matthew Dunn, Levent Sagun, Mike Higgins, V. Ugur Guney, Volkan Cirik, and Kyunghyun Cho. 2017. SearchQA: A New Q&A Dataset Augmented with Context from a Search Engine. arXiv:1704.05179 [cs.CL]
11. Joshua Robinson and David Wingate. 2023. Leveraging Large Language Models for Multiple Choice Question Answering. In The Eleventh International Conference on Learning Representations. <https://openreview.net/forum?id=yKbprarjc5B>
12. Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena (Version 4). arXiv. <https://doi.org/10.48550/ARXIV.2306.05685>
13. Yu, H., Gan, A., Zhang, K., Tong, S., Liu, Q., & Liu, Z. (2024). Evaluation of Retrieval-Augmented Generation: A Survey (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2405.07437>
14. Chen, M., et al. (2021). Evaluating Large Language Models Trained on Code. arXiv preprint arXiv:2107.03374.
15. Chen, J., Lin, H., Han, X., Sun, L.: Benchmarking large language models in retrieval augmented generation (Sep 2023). <https://doi.org/10.48550/ARXIV.2309.01431>
16. LangChain: Evaluating rag architectures on benchmark tasks (Nov 2023), [https://langchain-ai.github.io/langchain-benchmarks/notebooks/retrieval/langchain\\_docs\\_qa.html](https://langchain-ai.github.io/langchain-benchmarks/notebooks/retrieval/langchain_docs_qa.html)
17. Leng, Q., Uhlenhuth, K., Polyzotis, A.: Best Practices for LLM Evaluation of RAG Applications (Dec 2023), <https://www.databricks.com/blog/LLM-auto-eval-best-practices-RAG>
18. TruLens: TruLens (2023), [https://www.trulens.org/trulens\\_eval/getting\\_started/quickstarts/quickstart/](https://www.trulens.org/trulens_eval/getting_started/quickstarts/quickstart/)

### References

1. Kwiatkowski, T., et al. (2019). Natural Questions: A Benchmark for Question Answering Research. Transactions of the Association for Computational Linguistics, 7, 453-466.
2. Lewis, P., et al. (2020). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv preprint arXiv:2005.11401.
3. Shao, Z., Gong, Y., Shen, Y., Huang, M., Duan, N., & Chen, W. (2023). Enhancing Retrieval-Augmented Large Language Models with Iterative Retrieval-Generation Synergy (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2305.15294>.
4. He, X., Tian, Y., Sun, Y., Chawla, N. V., Laurent, T., LeCun, Y., Bresson, X., & Hooi, B. (2024). G-Retriever: Retrieval-Augmented Generation for Textual Graph Understanding and Question Answering (Version 3). arXiv. <https://doi.org/10.48550/ARXIV.2402.07630>.
5. Mao, Y., Dong, X., Xu, W., Gao, Y., Wei, B., & Zhang, Y. (2024). FIT-RAG: Black-Box RAG with Factual Information and Token Reduction (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2403.14374>
6. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W. T., Rocktäschel, T., Riedel, S., & Kiela, D. (2020, May 22). Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. arXiv.org. <https://arxiv.org/abs/2005.11401>
7. Gao Y, Xiong Y, Gao X, et al. Retrieval-Augmented Generation for Large Language Models: A survey. arXiv.org. <https://arxiv.org/abs/2312.10997>. Published December 18, 2023.
8. Islam, P., Kannappan, A., Kiela, D., Qian, R., Scherrer, N., & Vidgen, B. (2023). FinanceBench: A New Benchmark for Financial Question Answering (Версія 1). arXiv. <https://doi.org/10.48550/ARXIV.2311.11944>

9. Daniel Khashabi, Sewon Min, Tushar Khot, Ashish Sabharwal, Oyvind Tafjord, Peter Clark, and Hannaneh Hajishirzi. 2020. UNIFIEDQA: Crossing Format Boundaries with a Single QA System. In Findings of the Association for Computational Linguistics: EMNLP 2020, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, Online, 1896–1907. <https://doi.org/10.18653/v1/2020.findings-emnlp.171>
10. Matthew Dunn, Levent Sagun, Mike Higgins, V. Ugur Guney, Volkan Cirik, and Kyunghyun Cho. 2017. SearchQA: A New Q&A Dataset Augmented with Context from a Search Engine. arXiv:1704.05179 [cs.CL]
11. Joshua Robinson and David Wingate. 2023. Leveraging Large Language Models for Multiple Choice Question Answering. In The Eleventh International Conference on Learning Representations. <https://openreview.net/forum?id=yKbprarjc5B>
12. Zheng, L., Chiang, W.-L., Sheng, Y., Zhuang, S., Wu, Z., Zhuang, Y., Lin, Z., Li, Z., Li, D., Xing, E. P., Zhang, H., Gonzalez, J. E., & Stoica, I. (2023). Judging LLM-as-a-Judge with MT-Bench and Chatbot Arena (Version 4). arXiv. <https://doi.org/10.48550/ARXIV.2306.05685>
13. Yu, H., Gan, A., Zhang, K., Tong, S., Liu, Q., & Liu, Z. (2024). Evaluation of Retrieval-Augmented Generation: A Survey (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2405.07437>
14. Chen, M., et al. (2021). Evaluating Large Language Models Trained on Code. arXiv preprint arXiv:2107.03374.
15. Chen, J., Lin, H., Han, X., Sun, L.: Benchmarking large language models in retrieval augmented generation (Sep 2023). <https://doi.org/10.48550/ARXIV.2309.01431>
16. LangChain: Evaluating rag architectures on benchmark tasks (Nov 2023), [https://langchain-ai.github.io/langchain-benchmarks/notebooks/retrieval/langchain\\_docs\\_qa.html](https://langchain-ai.github.io/langchain-benchmarks/notebooks/retrieval/langchain_docs_qa.html)
17. Leng, Q., Uhlenhuth, K., Polyzotis, A.: Best Practices for LLM Evaluation of RAG Applications (Dec 2023), <https://www.databricks.com/blog/LLM-auto-eval-best-practices-RAG>
18. TruLens: TruLens (2023), [https://www.trulens.org/trulens\\_eval/getting\\_started/quickstarts/quickstart/](https://www.trulens.org/trulens_eval/getting_started/quickstarts/quickstart/)

<https://doi.org/10.31891/2219-9365-2024-79-20>

УДК 621.396.969.1

СЕМЕНКО Анатолій

Європейський університет

<https://orcid.org/0000-0002-7043-7801>

e-mail: [setel@ukr.net](mailto:setel@ukr.net)

БОЙКО Юлій

Хмельницький національний університет

<https://orcid.org/0000-0003-0603-7827>

e-mail: [boiko\\_julius@ukr.net](mailto:boiko_julius@ukr.net)

ШПУР Ольга

Національний університет «Львівська політехніка»

<https://orcid.org/0000-0001-8156-8017>

e-mail: [olha.m.shpur@lpnu.ua](mailto:olha.m.shpur@lpnu.ua)

## КОМПЛЕКСНА ОПТИМІЗАЦІЯ ТЕЛЕКОМУНІКАЦІЙНОЇ СИСТЕМИ НА ОСНОВІ МОДИФІКОВАНОЇ ПВП ГОЛДА І RAKE-ПРИЙМАЧІВ

При проектуванні телекомунікаційних систем важливим завданням є забезпечення мінімальної ймовірності бітової помилки приймання сигналу при даній потужності передавача. Для вирішення даної задачі прагнуть зменшити ширину смуги пропускання каналу аби зменшити рівень шумів, які неминуче надходять до каналу із зовнішнього середовища. В роботі запропонована і досліджена телекомунікаційна система (ТКС) з широкосмуговим сигналом, що передається в 2-х окремих антенних каналах. В системі здійснюється кодове розділення сигналів каналів з використанням модифікованої псевдовипадкової послідовності (ПВП) Голда і оптимальна обробка сигналів за допомогою RAKE-приймачів. При цьому досягається мінімальна ширина смуги пропускання каналу та максимальне відношення сигнал/шум.

Ключові слова: телекомунікаційна система, ймовірність бітової помилки, відношення сигнал/шум, псевдовипадкова послідовність, RAKE-приймач

SEMENKO Anatoliy

European University

BOIKO Juliy

Khmelnitskyi National University

SHPUR Olga

Lviv Polytechnic National University

## COMPREHENSIVE OPTIMIZATION OF TELECOMMUNICATION SYSTEM BASED ON MODIFIED PRS GOLD AND RAKE RECEIVERS

In designing telecommunication systems, the fundamental task is to ensure a low probability of bit error in signal reception at a given transmitter power. To solve this problem, they try to reduce the channel bandwidth in order to reduce the noise level that inevitably enters the channel from the external environment. In this paper, a telecommunication system (TS) with a wideband signal transmitted in 2 separate antenna channels is proposed and investigated. The system implements code division of channel signals using a modified pseudo-random sequence (PRS) of Gold and optimal signal processing using RAKE receivers. In this case, the minimum channel bandwidth and maximum signal-to-noise ratio are achieved. The main directions of optimization of the TCS based on the studied solutions are based on the formation of new modified Gold sequences with better correlation characteristics and a lower level of cross-correlation to improve the signal quality in communication systems. Quite promising is the approach that covers the development of algorithms for optimizing the operation of RAKE receivers, including adaptive tuning for working with various modified Gold sequences. The study of the efficiency of RAKE reception in combination with modified Gold sequences was carried out to reduce the negative impact of Multipath Propagation. A search for optimization of signal processing algorithms for RAKE receivers was carried out to reduce computational costs and ensure their effective application in real time. In order to synthesize optimal Gold sequences, cross-correlation plots are obtained showing how two realizations of the same Gold sequence (shifted by one element) interact with each other after they have been modulated using BPSK and QPSK. A RAKE receiver with BPSK type signals was simulated. The resulting graphs show signals with delays for each finger of the RAKE receiver, as well as the signal structure obtained as a result of processing the RAKE type receiver.

Keywords: telecommunication system, bit error rate, signal-to-noise ratio, pseudo-random sequence, RAKE receiver

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Функціонування сучасних телекомунікаційних систем пов'язане з численними викликами, такими як обмежена пропускна здатність, вплив інтерференції та зменшення якості сигналу через відбиття, а також явища коли сигнали, що накладаються один на одного можуть посилюватися чи гаситися в залежності від фазових відмінностей (англ. Multipath Propagation) і що в кінцевому рахунку може призводити до інтерференції за рахунок багатопроменевого поширення сигналів (БППС) [1]. Один з підходів до покращення ефективності таких систем полягає в застосуванні вдосконалених методів кодування та

прийому сигналів. В цьому контексті модифіковані послідовності Голда та RAKE-приймачі представляють собою перспективні технології для вирішення цих проблем [2].

Послідовності Голда, які є детермінованими ПВП [3-5], використовуються для покращення характеристик систем зв'язку, таких як покращення якості сигналу і зменшення впливу БППС. Модифікація класичних послідовностей Голда дозволяє підвищити їх ефективність у специфічних умовах застосування телекомунікаційних систем, адаптуючи їх до конкретних вимог системи. RAKE-приймачі є ефективним методом для боротьби з БППС в бездротових системах зв'язку. Вони дозволяють обробляти відбиті сигнали з різними затримками, що допомагає підвищити якість прийому сигналу. Оптимізація RAKE-приймачів у поєднанні з модифікованими послідовностями Голда може суттєво покращити продуктивність системи.

В цьому контексті можна виділити основні напрямки оптимізації ТКС на основі досліджуваних рішень. Зокрема, важливо сконцентруватись на розробці нових модифікованих послідовностей Голда, які мають кращі характеристики кореляції та менший рівень перехресної кореляції для покращення якості сигналу в системах зв'язку. Досить перспективним є підхід який охоплює розробку алгоритмів для оптимізації роботи RAKE-приймачів, що включають адаптивне налаштування для роботи з різними модифікованими послідовностями Голда і безперечно важливо провести вивчення ефективності RAKE-приймачів у поєднанні з модифікованими послідовностями Голда для зменшення негативного впливу БППС.

Стаття сконцентрована на описі запропонованої схеми ТКС у якій застосовуються підходи на основі використання модифікованих послідовностей Голда, а обробка сигналів реалізується RAKE-приймачами з метою комплексної оптимізації телекомунікаційної системи в умовах БППС.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Аналіз методів предметної області досліджень [6, 7] показує, що модифіковані послідовності Голда і RAKE-приймачі можуть бути інтегровані в системи мобільного зв'язку, Wi-Fi, і супутникові системи для покращення їх продуктивності [8, 9]. В літературі [10-13] описано перспективні напрямки використання модифікованих послідовностей Голда для підвищення точності та надійності радіолокаційних систем. Окремі аспекти інтеграції RAKE-приймачів для покращення точності систем глобального позиціонування (GPS) в умовах БППС обговорюються в роботі [14]. В статті [15] автори описують методику застосування технології RAKE як методу боротьби з багатопроменевістю в системах мобільного зв'язку. В статті [16] автори сконцентрувались на розгляді еволюції систем з кодовим розділенням каналів. Окремо досліджено принципи побудови та роботи RAKE-приймача, а також зміни, які вносять у приймач нові стандарти зв'язку.

Таким чином, аналіз літературних джерел відомих авторів дозволяє намитити певні тренди у подальших пошуках предметної області. Важливо розробити механізми синтезу алгоритмів для автоматизованої генерації та оптимізації модифікованих послідовностей Голда з урахуванням специфічних умов системи. Крім того, аналіз джерел дозволяє ставити питання пошуку засобів оптимізації алгоритмів обробки сигналів RAKE-приймачів для зменшення обчислювальних витрат та забезпечення їх ефективного застосування в реальному часі. Важливо, також, здійснити розробку сумісних інтерфейсів та протоколів для інтеграції модифікованих послідовностей Голда і RAKE-приймачів у вже існуючі системи без потреби значних змін у їх архітектурі. Реалізація таких надважливих задач неможлива без проведення комплексних тестів та моделювання для оцінки впливу нових технологій на існуючі системи та визначення оптимальних параметрів для їх спільної роботи. Таким чином, модифіковані послідовності Голда та RAKE-приймачі мають потенціал для значного покращення ТКС. Однак, для успішного впровадження цих технологій необхідно враховувати і вирішувати проблеми, пов'язані з їх реалізацією та інтеграцією в існуючі системи.

### Особливості формування та кроскореляції послідовностей Голда

Для візуалізації послідовностей Голда ми провели синтез та моделювання такої послідовності. В цьому моделюванні було використано лінійний зсувний регістр (ЛЗР) для генерації послідовностей Голда. По суті ЛЗР є цифровим генератором випадкових чисел, який використовує поліном для визначення послідовності виходів. В ЛЗР, бітовий регістр зсувається на одну позицію кожен раз коли генерується новий біт, а новий біт обчислюється як XOR (функція mod2) виходів деяких бітів регістру. В нашому дослідженні поліном визначає структуру ЛЗР і використовується для обчислення нового біту на основі поточного стану регістру. Тобто ми визначали поліном як масив коефіцієнтів, де коефіцієнти поліному є бітами (0 або 1).

Для генерації послідовності Голда був використаний поліном типу:  $x^5 + x^2 + 1$  який представлено масивом [1, 0, 0, 1, 1], а також поліном:  $x^5 + x^3 + x^2 + 1$  де в дослідженні ми отримали масив [1, 1, 0, 1, 1]. Причому ЛЗР містить регістр довжиною 5 бітів. Такі поліноми забезпечують специфічну структуру генерації послідовності Голда, де XOR виходи з обох ЛЗР утворюючи кінцеву послідовність яку за результатами моделювання представлено на рис. 1.

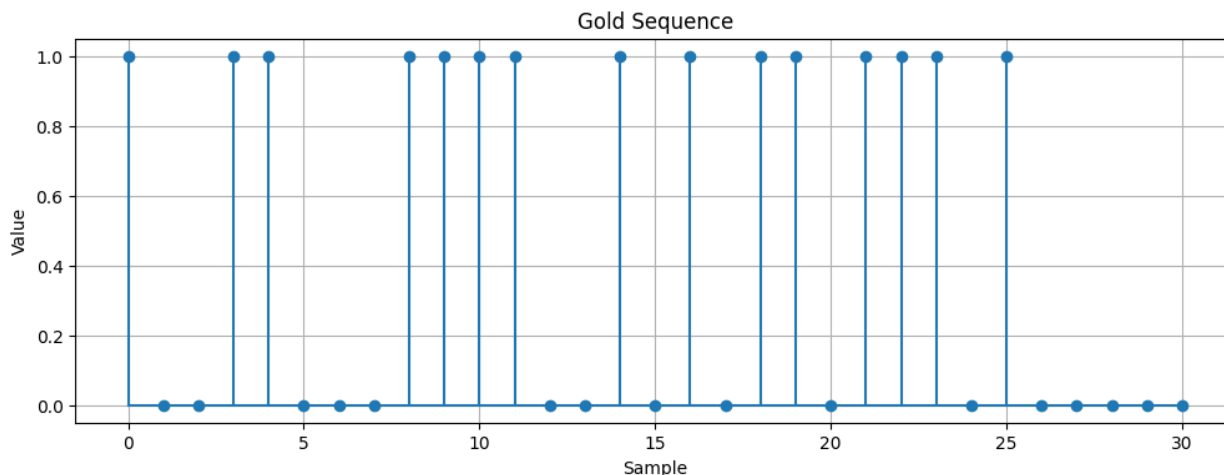


Рис.1. Згенерована послідовність Голда: Sample - це окреме значення або точка даних у послідовності (на графіку це відповідно до індексу вибірки і її значення (яке може бути або 0, або 1 для послідовності Голда))

З точки зору застосування послідовностей Голда у ТКС нас буде цікавити аспекти кроскореляції таких послідовностей. Зокрема ми отримали графіки кроскореляції які повинні показати, як дві реалізації однієї послідовності Голда (зсунутої на один елемент) взаємодіють між собою, після того як їх було модульовано за допомогою BPSK і QPSK (ФМ-2 і ФМ-4). Ми використовували сигнали фазової маніпуляції ФМ-2 та ФМ-4 при побудові схеми ТКС, рис. 2.

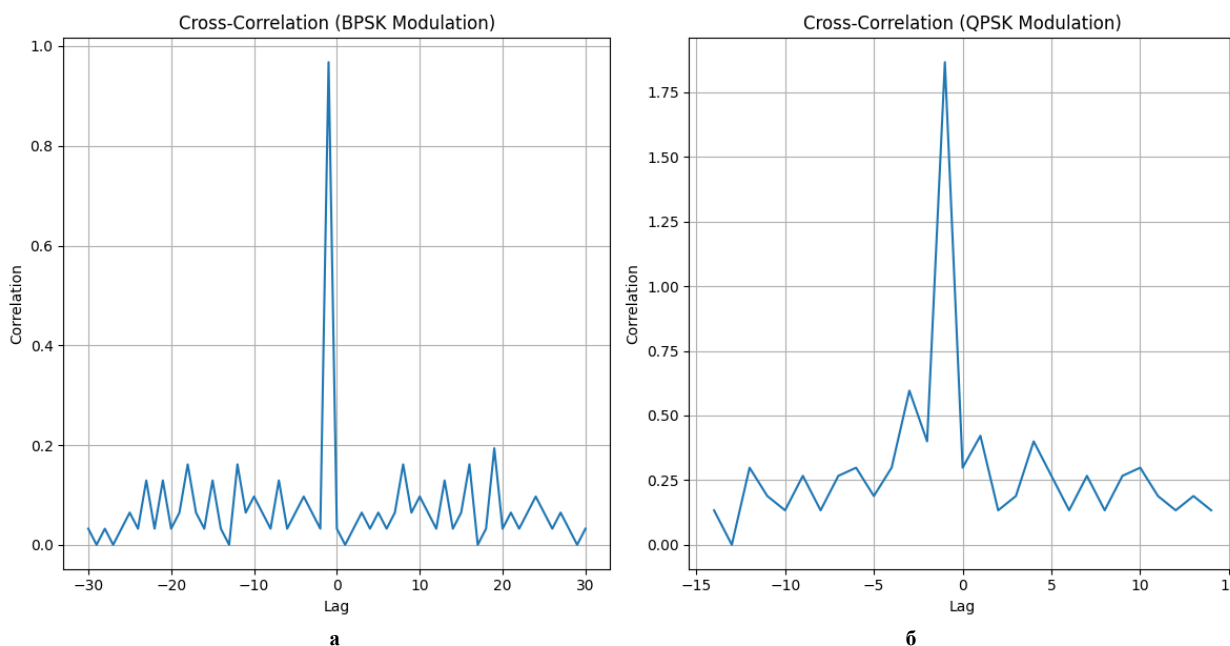


Рис.2. Згенеровані графіки кроскореляції які показують як дві реалізації однієї послідовності Голда (зсунутої на один елемент) взаємодіють між собою, після того як їх було модульовано за допомогою BPSK -а і QPSK -б (ФМ-2 і ФМ-4): Lag – зсув між двома послідовностями в часі або у просторі, який використовується для аналізу їхньої взаємозалежності

Аналіз отриманих графіків дозволяє зробити наступні висновки. Графіки показують амплітуду кроскореляційної функції для двох ФМ-2 та ФМ-4 модульованих послідовностей. Пік на графіку вказує на те, що ці послідовності дуже схожі або однакові на певному зсуві. В контексті кроскореляції, зсув означає на скільки елементів одна послідовність зсувається відносно іншої, щоб обчислити їх взаємну кореляцію. У випадку послідовностей Голда, кореляція звичайно має пік на центральному зсуві (при зсуві 0), що вказує на їх хороші автокореляційні властивості, а інші значення близькі до нуля. Відмітимо, що у випадку ФМ-4 (рис. 2 б) графік показує кроскореляцію для двох ФМ-4-модульованих послідовностей, де послідовність розбита на реальну та уявну частини. Подібно до ФМ-2, пік на графіку відповідає максимальній схожості послідовностей на певному зсуві. Через складнішу структуру ФМ-4 модульованих сигналів (реальна та уявна частини), кореляційна функція має більш складну форму, але центральний пік все ще вказує на схожість сигналів. За підсумками проведеного моделювання робимо висновок, що високі значення

кроскореляції на центральному зсуві (0) означають, що послідовності мають хороші автокореляційні властивості. Крім того, низькі значення кроскореляції на інших зсувах показують, що послідовності добре відокремлені одна від одної при будь-якому іншому зсуві, що є важливою властивістю для використання цих послідовностей в системах зв'язку для мінімізації взаємних завад. Таким чином отриманий результат підтверджує, що послідовності Голда мають хороші кореляційні властивості, що робить їх корисними для застосування в широкосмугових системах зв'язку і зокрема під час вирішення завдання комплексної оптимізації телекомунікаційної системи на основі модифікованої ПВП Голда і RAKE-приймачів які ми вирішуємо в представленому дослідженні.

### Розробка функціональної схеми ТКС із використанням модифікованої ПВП Голда і RAKE-приймачів

На рис. 3 наведена схема запропонованої ТКС.

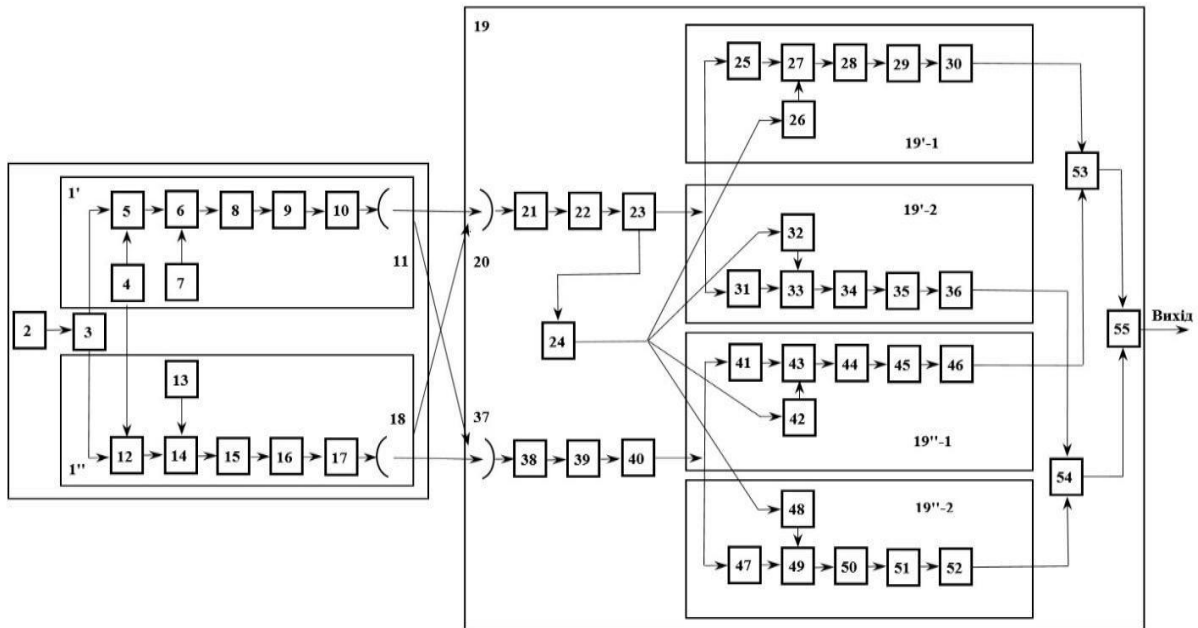


Рис.3. Функціональна схема ТКС

Таким чином на рис. 3 представлено 2-х каналну систему, що загалом представляє собою систему типу MIMO 2x2, в якій, в каналах з однаковою шириною смуги пропускання передаються різні сигнали.

Опишемо цю систему в позначеннях. На рис.3 введені позначення: 1-передавач; 2-джерело інформаційного сигналу; 3-дільник сигналу; 4-генератор сигналу проміжної частоти; 5,12-модулятори; 6, 14-пристрої створення шумоподібного сигналу ФМ-2; 7,13-генератори модифікованого коду Голда; 8,15-пристрої створення сигналів з ФМ-4; 9, 16-перетворювачі частоти; 10,17-підсилювачі потужності; 11, 18, 20, 37-антени; 19-приймач; 21,38-фільтр; 22, 39-малощумливі підсилювачі; 23, 40- перетворювачі частоти; 25,31,41,47-підсилювачі; 26, 32, 42, 48-генератори опорних радіосигналів ФМ 4; 24-синхронізатор; 27,33, 43, 49-RAKE-приймачі; 28,34, 44, 50-вирішувачі пристрої; 29, 35, 45, 51-детектори; 30, 36, 46, 52-формувачі відеосигналу; 53, 54, 55-суматори; 1', 1''-підсистеми передавача, 19'-1, 19'-2, 19''-1, 19''-2- підсистеми приймача.

Звичайно обробка сигналів в приймачі від 2-х антенних каналів в системах з MIMO здійснюється доволі складними способами [17].

В даній роботі пропонується прийнятний за складністю спосіб обробки сигналів в приймачі.

В системі здійснюється комплексна оптимізація системи за критерієм енергетичної ефективності в 3 етапи:

1. Зменшення смуги пропускання радіоканалу шляхом зменшення в 2 рази швидкості передачі сигналу.
2. Подальше зменшення смуги пропускання радіоканалу в 2 рази шляхом використання чотирирохпозиційного фазоманіпульованого сигналу ФМ-4.
3. Оптимальна обробка сигналів в RAKE-приймачах.

Передумовою розробки схеми ТКС є передача широкосмугового сигналу в радіоканалі з мінімальною смугою пропускання з огляду на дефіцит частотного ресурсу та зменшення потужності зовнішніх завад, що надходять до радіоканалу із навколишнього середовища.

Для цього, швидкість сигналу зменшена в 2 рази, тобто тривалість імпульсу інформаційного сигналу  $T_1$  збільшена вдвоє, що забезпечує використання радіоканалу із зменшеною вдвічі смугою пропускання.

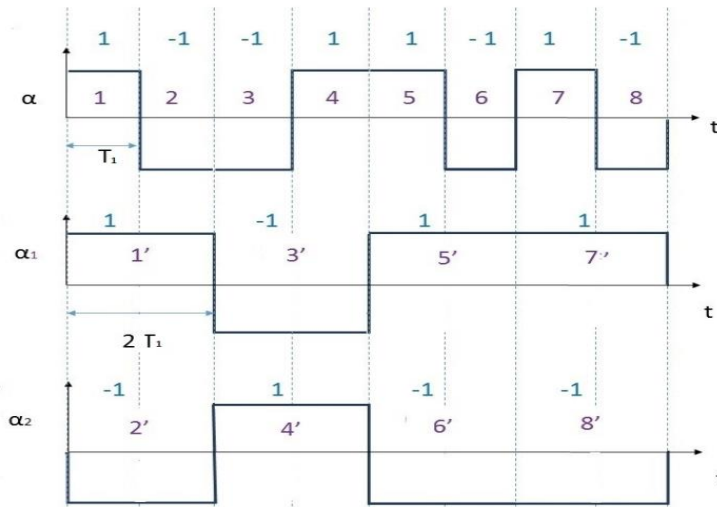


Рис.4. Структура вхідного сигналу  $\alpha$  і сигналів в каналах  $\alpha_1$  і  $\alpha_2$

Тоді при уявній модуляції високочастотного сигналу даним сигналом і збереженні потужності передавачів відношення сигнал/шум буде:

$$\gamma_1 = \frac{A^2 T_1}{N_0} = 2\gamma, \quad (1)$$

де  $A$  – амплітуда імпульсу;  $T_1$  – тривалість імпульсу інформаційної послідовності;  $N_0$  – спектральна щільність шумів в радіоканалі;  $\gamma$  – відношення сигнал шум при уявній модуляції первинним інформаційним сигналом високочастотного сигналу.

Шумоподібний сигнал зазвичай створюється методом прямого розширення спектру за допомогою перетворення імпульсів інформаційного сигналу в послідовність коротких імпульсів тривалістю  $\tau$ , наприклад, за ПВП Голда. При модуляції таким сигналом високочастотного сигналу створюється складний псевдошумовий сигнал ФМ-2 з базою  $B = 2T_1/\tau$ . Тоді, відношення сигнал/шум буде зменшене в  $B$  разів:

$$\gamma_2 = \frac{A^2 \tau}{2N_0} = \frac{A^2 T_1}{N_0 B} = \frac{2\gamma}{B} \quad (2)$$

Для подальшого зменшення смуги пропускання радіоканалу в 2 рази слід використати фазоманіпульований сигнал ФМ-4 з тривалістю імпульсу  $2\tau$ . Для створення широкосмугового сигналу з ФМ-4 класична ПВП Голда не придатна, так як вона має непарну кількість бітів. Для створення системи з кодовим розділенням каналів із сигналом ФМ-4 доцільно використати модифіковану ПВП Голда з парною кількістю бітів шляхом додавання до класичної послідовності 1-го біта, наприклад, в кінці послідовності.

Як показали дослідження з використанням програмування в системі MATLAB, взаємкореляційна функція 2-х реалізацій модифікованої ПВП Голда практично не погіршується в порівнянні з використанням класичної ПВП Голда (див. рис. 2).

Тоді відношення сигнал/шум в системі з ФМ-4 буде покращено ще в 2 рази [18]:

$$\gamma_3 = \frac{A^2 2T}{N_0} = \frac{4\gamma}{B} \quad (3)$$

Створені в передавачі сигнали передаються до антен, з яких сигнали надходять до радіоканалу. На виході кожної антени приймача відношення сигнал/шум буде:

$$\gamma_a = \frac{P_c}{[P_m(1+m)]}, \quad (4)$$

де  $P_m$  – власні шуми передавача;  $m$  – відношення завада/власний шум передавача.



$$m = \frac{P_z}{P_m}, \quad (5)$$

де  $P_z$  – сумарна потужність завад, що надходять до радіоканалу.

Сигнали надходять до RAKE-приймачів 27, 33, 43, 49 (див. рис. 3), де автоматично формуються промені сигналів з визначенням величини їх запізнення та з використанням опорних ФМ-4 сигналів від генераторів 26, 32, 42, 48 (див. рис. 3), здійснюється кореляційна обробка сигналів [19]. При цьому, доцільно використати виділення головних пелюсток сигналів стробами, що дозволить видалити інші пелюстки меншої величини як завади.

В системі використовується кадрова синхронізація з формуванням в передавачі маркерного сигналу. В синхронізаторі створюються синхросигнали з використанням автокореляційного приймання сигналу проміжної частоти, які надходять до генераторів опорних сигналів кореляторів RAKE-приймачів.

Враховуючи, що при обробці сигналів в кореляторах RAKE-приймачів відношення сигнал/шум на виході корелятора буде збільшене в  $B$  разів, в системі буде відновлене відношення сигнал/шум, яке було погіршене в  $B$  разів при створенні шумоподібного сигналу. RAKE-приймач складається з кількох «пальців» або «каналів», кожен з яких обробляє сигнал, що надійшов через певний шлях або з певною затримкою. Кожен палець (канал) коригує і декодує сигнал для свого конкретного шляху. Результати обробки з усіх пальців комбінуються для відновлення вихідного сигналу. Це дозволяє покращити загальну якість прийому, компенсуючи затримки та ослаблення сигналу, які можуть бути присутніми в окремих каналах. RAKE-приймач може коригувати затримки сигналу, щоб забезпечити точний збір сигналів від різних шляхів. Отже, такий приймач використовує кореляцію для виявлення та розділення сигналів, що надійшли через різні канали. Кожен палець приймача призначений для обробки сигналу з певною затримкою, а результати обробки сумуються для покращення якості прийому.

Для представлення принципів роботи RAKE-приймача ми провели моделювання та отримали вигляд сигналів на вході та виході такого приймача за умови використання послідовностей Голда промодульованих сигналом ФМ-2 (див. рис. 2 а). Результати представлено на рис. 5.

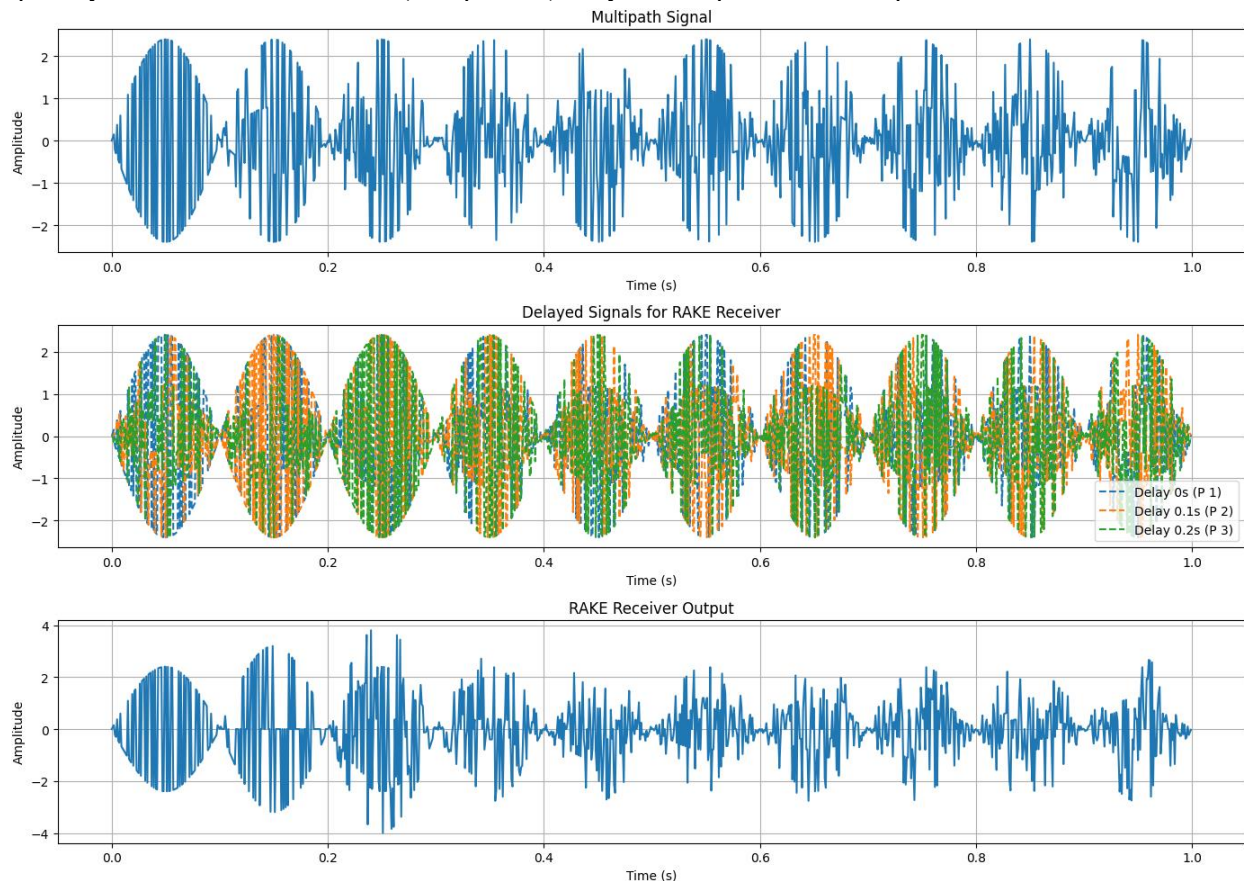


Рис.5. Структура сигналів на вході та виході RAKE-приймача

Згідно до рис. 5 сигнал типу ФМ-2 представляється у вигляді бітової послідовності де кожен біт у послідовності перетворюється у фазу сигналу [20]. В наслідок дії багатопроменовості (Multipath Propagation) утворюються сумарні маніпульовані сигнали типу ФМ-2 з різними затримками і амплітудами (дивитись рис.

5). За рахунок використання RAKE-приймача відбувається обробка сигналу утвореного продуктами явища багатоприменності і для цього використовуються кілька “пальців” RAKE-приймача з різними затримками. Тобто на рис. 5 зверху до низу маємо спочатку маємо вигляд сумарного сигнал з усіх шляхів. Далі на наступному графіку сигнали з затримками для кожного пальця RAKE-приймача. Тоді далі на нижньому графіку представлено результат обробки RAKE-приймачем.

Подача на вхід RAKE-приймачів збільшеного значення відношення сигнал/шум у порівнянні з вхідним значенням дозволяє одержати підвищене відношення сигнал/шум на виході системи в  $B$  разів завдяки обробці сигналів в кореляторах. В даній системі на вході RAKE-приймачів забезпечується в 4 рази більше відношення сигнал/шум в порівнянні із значенням на вході.

При об'єднанні сигналів променів за принципом максимального відношення сигнал/шум зважуються їх відношення сигнал/шум і потім складаються. При цьому необхідно досягти синфазності сигналів. Складання з максимальним відношенням сигнал/шум дає середнє значення сигнал/шум на виході кожного RAKE-приймача, що дорівнює сумі окремих відношень сигнал/шум.

Доведено, що на виході кожного RAKE-приймача відношення сигнал/шум буде [7]:

$$y_{\Sigma} = y_0 + \psi \sum_{i=0}^3 y_i, \quad (6)$$

де  $y_0$  – відношення сигнал/шум сигналу основного напрямку;  $y_i$  – відношення сигнал/шум  $i$ -ого променю ( $i_{\max} = 3$ );  $\psi$  – відношення суми значень променів сигнал/шум до значення сигнал/шум від сигналу основного напрямку:

$$\psi = (\sum_{i=0}^3 y_i) / y_0. \quad (7)$$

Звичайно  $\psi < 1$ . Слід зазначити, що в реальних випадках можуть бути відсутніми окремі промені ( $i = 0$ ), тоді буде здійснюватися обробка тільки сигналу від основного напрямку  $y_0$  і  $\psi = 0$ .

В пристроях 14, 17, 27, 32 (див. рис. 3) здійснюється відновлення сигналів тривалістю  $T_1$  і знаходяться за розробленим методом значення окремо суми потужностей корисних сигналів від 4-х каналів:

$$P_c = P_{c1} + P_{c2} + P_{c3} + P_{c4}, \quad (8)$$

де  $P_{c1}, P_{c2}, P_{c3}, P_{c4}$  - потужності сигналів від каналів  $19^{1-1}, 19^{11-1}, 19^{1-2}, 19^{11-2}$ , до складу яких входять сигнали основних каналів і променів (див. рис. 3), а також окремо сума потужностей шумів від 4-х каналів:

$$P_{\text{ш}} = P_{\text{ш}1} + P_{\text{ш}2} + P_{\text{ш}3} + P_{\text{ш}4}, \quad (9)$$

де  $P_{\text{ш}1}, P_{\text{ш}2}, P_{\text{ш}3}, P_{\text{ш}4}$  - потужності шумів від каналів  $19^{1-1}, 19^{11-1}, 19^{1-2}, 19^{11-2}$ , до складу яких входять шуми основних каналів і променів.

Після складення потужностей сигналів та шумів в суматорах 53, 54, 55 (див. рис. 3) на виході системи одержується відношення сигнал/шум:

$$y = \frac{(P_{c1} + P_{c2} + P_{c3} + P_{c4})}{(P_{\text{ш}1} + P_{\text{ш}2} + P_{\text{ш}3} + P_{\text{ш}4})}, \quad (10)$$

За графіком рис. 6 можна знайти величину ймовірності помилки приймання сигналу за одержаним значенням відношення сигнал/шум. Якщо ймовірність помилки приймання сигналу (BER) [21] буде кращою за потрібну, можна зменшити відношення сигнал/шум, тобто зменшити потужність передавачів. Отриманий графік (рис. 6) демонструє залежність ймовірності бітової помилки (BER) від відношення сигналу до шуму ( $E_b/N_0$ ) для модуляцій BPSK (ФМ-2) та QPSK (ФМ-4) на виході RAKE-приймача. Як для ФМ-2, так і для ФМ-4, із зростанням відношення  $E_b/N_0$  ймовірність бітової помилки зменшується. Це очікуваний результат, оскільки при вищих значеннях  $E_b/N_0$  сигнал стає більш стійким до впливу шуму, що призводить до меншої кількості помилок при прийомі. Обидві маніпуляції мають схожі залежності BER від  $E_b/N_0$ , але для ФМ-4 значення BER трохи гірші при тому ж значенні  $E_b/N_0$ . Це пов'язано з тим, що при ФМ-4 передається два біти на символ, що робить її більш чутливою до шуму в порівнянні з ФМ-2, яка передає лише один біт на символ. При низьких значеннях  $E_b/N_0$  (від 0 до 5 dB) спостерігається більш різке зменшення BER для обох маніпуляцій, але все ж таки ФМ-4 має трохи вищу ймовірність помилки у цьому діапазоні. Це свідчить про те, що в умовах низького відношення сигнал/шум ФМ-2 забезпечує трохи кращу стійкість. При значеннях  $E_b/N_0$  більше 10 дБ обидві модуляції досягають дуже низьких значень BER ( $\sim 10^{-6}$  і нижче), що є свідченням їх високої ефективності при хороших умовах прийому.

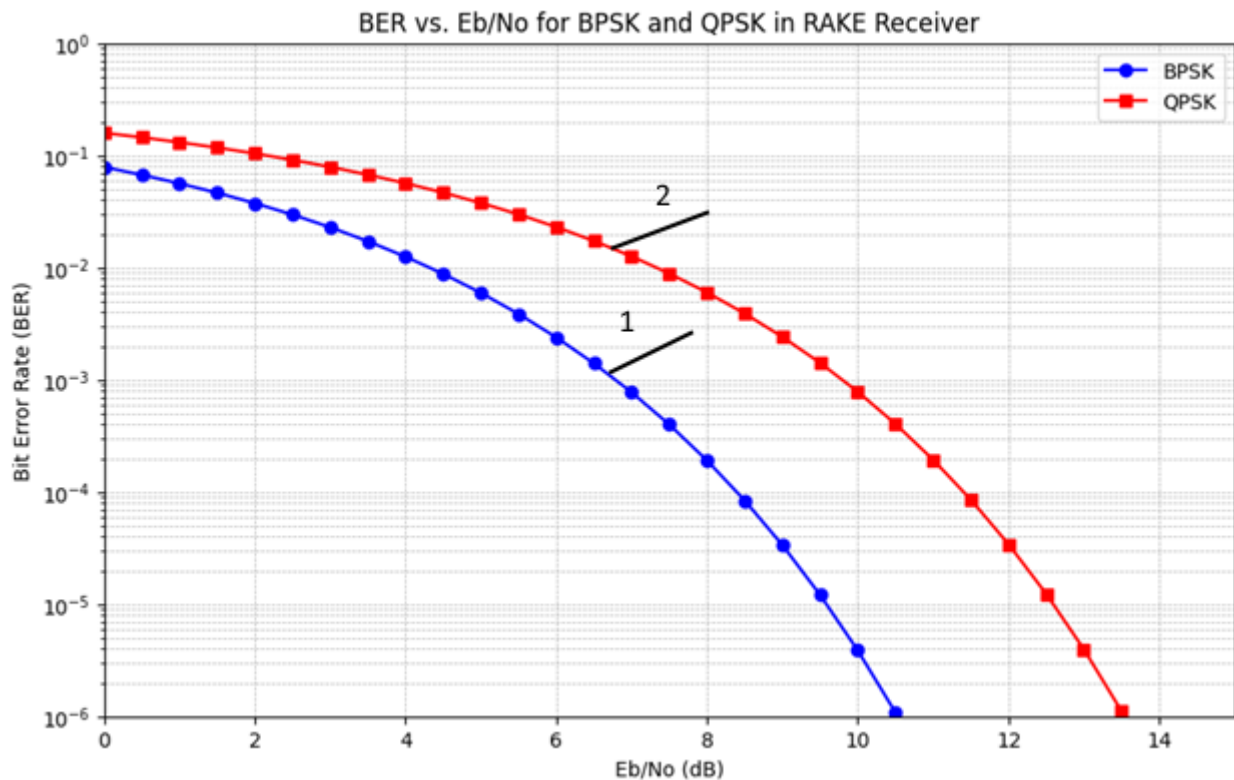


Рис.6. Графік залежностей BER для BPSK (ФМ-2) та QPSK (ФМ-4) при різних значеннях відношення  $E_b/N_0$  у ТКС з RAKE-приймачем: 1 – використання модуляції ФМ-2 (BPSK); 2 – використання модуляції ФМ-4 (QPSK)

Потрібно наголосити, що у практичних ТКС вибір між ФМ-2 та ФМ-4 може бути обумовлений компромісом між вимогами до спектральної ефективності та стійкістю до шуму. ФМ-2 може бути обраний у системах, де критично важлива низька ймовірність помилки при складних умовах прийому, тоді як ФМ-4 може бути використаний для підвищення пропускної здатності за умов відносно високого рівня сигналу. Отже, досить важливо забезпечити адаптивний формат модуляції за умов використання запропонованих рішень в сучасних ТКС.

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

### І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

У даній роботі запропоновано комплексну оптимізацію ТКС на основі модифікованої ПВП Голда та RAKE-приймачів. Проведені дослідження показали, що використання модифікованих послідовностей Голда разом із RAKE-приймачами забезпечує ефективне зниження впливу БППС, що суттєво покращує якість приймання сигналів.

Отримані графіки крос-кореляції для двох реалізацій однієї послідовності Голда (зсунутої на один елемент), модульованих за допомогою ФМ-2 та ФМ-4, продемонстрували високу ступінь схожості сигналів при певних зсувах. Це свідчить про хороші кореляційні властивості модифікованих послідовностей, що є важливим для підвищення ефективності роботи ТКС в умовах впливу шумів та інтерференції.

Запропонована функціональна схема ТКС відзначається високою ефективністю у передачі даних, забезпечуючи оптимальний баланс між швидкістю передачі та енергетичною ефективністю. Розроблена методика визначення відношення сигнал-шум дозволяє здійснювати точний аналіз роботи системи в умовах різних рівнів завад, що важливо для забезпечення надійності та якості зв'язку в сучасних ТКС.

Перспективи подальших досліджень у даному напрямку полягають у розробці нових модифікованих послідовностей Голда з покращеними кореляційними характеристиками, що дозволить ще більше підвищити якість сигналів у ТКС. Крім того, перспективним є подальше вдосконалення алгоритмів RAKE-приймачів, включаючи адаптивне налаштування для роботи з різними послідовностями, що може знизити обчислювальні витрати та підвищити ефективність у реальних умовах експлуатації.

### Література

1. A. Hamza, H. AlShammary, C. Hill and J. F. Buckwalter, "A Full-Duplex Rake Receiver Using RF Code-Domain Signal Processing for Multipath Environments," in IEEE Journal of Solid-State Circuits, vol. 56, no. 10, pp. 3094-3108, Oct. 2021.

2. A. F. Molisch, "Spread Spectrum Systems," in *Wireless Communications*, IEEE, 2011, pp.387-416.
3. Y.-H. Lee and S.-J. Kim, "Sequence acquisition of DS-CDMA systems employing Gold sequences," in *IEEE Transactions on Vehicular Technology*, vol. 49, no. 6, pp. 2397-2404, Nov. 2000.
4. A. I. Семенко, Н. І. Бокла, "Створення модифікованих псевдовипадкових послідовностей Голда для телекомунікаційних систем з кодовим розподілом каналів," *Телекомунікаційні та інформаційні технології*, №3, с. 10-14, 2014.
5. A. I. Semenکو, N. I. Bokla, "Effektivnost telekommunikatsionnikh sistem s ispolzovaniem modifitsirovannikh psevdosluchainikh posledovatelnosti Golda," *Elektrosvyaz*, №3, pp. 37-40, 2014.
6. X. Peng, J. Chen, Y. Li and Y. Song, "Single Subcarrier Gold Sequences Modulated Timing Synchronization for Upstream OFDMA-PON," in *IEEE Photonics Journal*, vol. 9, no. 1, pp. 1-7, Feb. 2017, Art no. 7200207.
7. U. Grob, A. L. Welti, E. Zollinger, R. Kung and H. Kaufmann, "Microcellular direct-sequence spread-spectrum radio system using N-path RAKE receiver," in *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 5, pp. 772-780, June 1990.
8. D. Dharmappa, M. V. Mandi and S. Ramesh, "Analysis of Binary Sequences with better correlation performance compared to Gold code under different doppler conditions for GNSS applications," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018, pp. 209-215.
9. J. -H. Yan, C. -W. Wang, K. -H. Lin and K. -M. Feng, "A Timing-Synchronization-Free WDM-Compatible Colorless DRoF Uplink System for 5G Mobile Fronthaul Employing Gold Sequence Multiplexing," 2019 Conference on Lasers and Electro-Optics (CLEO), San Jose, CA, USA, 2019, pp. 1-2.
10. A. Semenکو, N. Bokla, A. Shokotko and V. Lusenکو, "Construction of microwave link using channel code division based on modified pseudorandom Gold sequences," 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), Lviv, Ukraine, 2016, pp. 634-637.
11. Y. Jun and Z. Xiang-Li, "A new design of Gold-like binary sequences," 2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), Wuhan, China, 2017, pp. 118-124, doi: 10.1109/ACIRS.2017.7986077
12. K. Liu, Z. Zhou, A. R. Adhikary and C. Tang, "Large Sets of Binary Spreading Sequences With Low Correlation and Low PAPR via Gold Functions," in *IEEE Transactions on Information Theory*, vol. 70, no. 7, pp. 5309-5322, July 2024.
13. L. Shufeng, H. Shiyao and W. Hongda, "Analysis of DS-CMDA system using orthogonal gold sequence," 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), Macau, China, 2017, pp. 520-523.
14. D. Dharmappa, M. V. Mandi and S. Ramesh, "Generation of Binary Sequences Having Better Odd Correlation and Linear Complexity Property Compared to Gold Codes for use in Global Navigation Satellites Systems (GNSS) Applications," 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, India, 2017, pp. 155-160.
15. Н.І. Кунах, О.М. Ткаленко, "Застосування технології RAKE як методу боротьби з багатопроменевою в системах мобільного зв'язку," *Сучасна спеціальна техніка*, №3(50), с. 34-41, 2017.
16. А.О. Кахно, С.М. Дяченко, "RAKE-приймачі в сучасних системах телекомунікації," *Вісник Національного технічного університету України "КПІ" Серія – Радіотехніка. Радіоапаратобудування*, №34, с. 92-96, 2007.
17. Zhen Gao, Yikun Mei, Li Qiao, "Sparse Signal Processing for Massive MIMO Communications," Springer: Singapore, pp. 217, 2023.
18. А.І. Семенко, Н.І. Бокла, "Метод двократного зменшення необхідної смуги пропускання каналу зв'язку в телекомунікаційній системі з кодовим розподілом каналів," *Зв'язок*, №4, с.23-25, 2011.
19. J. Boiko, L. Karpova, "Increasing Efficiency in the Correlation Processing of Information Signals for Radar," *Proceedings of Fourth International Conference on Communication, Computing and Electronics Systems. Lecture Notes in Electrical Engineering*, vol 977. Springer: Singapore, pp. 549–569, 2023.
20. Ю. Бойко, А. Семенко, І. П'ятін, "Особливості формування кодової надлишковості у каналах передачі інформації," *Інфокомунікаційні та комп'ютерні технології*, №04, Т. 2, с.12-25, 2023.
21. J. Boiko, I. Pyatin, L. Karpova, O. Eromenko, "Study of the Influence of Changing Signal Propagation Conditions in the Communication Channel on Bit Error Rate," in *Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies*, vol 69, Springer: Cham, pp. 79–103, 2021.

## References

1. A. Hamza, H. AlShammary, C. Hill and J. F. Buckwalter, "A Full-Duplex Rake Receiver Using RF Code-Domain Signal Processing for Multipath Environments," in *IEEE Journal of Solid-State Circuits*, vol. 56, no. 10, pp. 3094-3108, Oct. 2021.
2. A. F. Molisch, "Spread Spectrum Systems," in *Wireless Communications*, IEEE, 2011, pp.387-416.

3. Y.-H. Lee and S.-J. Kim, "Sequence acquisition of DS-CDMA systems employing Gold sequences," in *IEEE Transactions on Vehicular Technology*, vol. 49, no. 6, pp. 2397-2404, Nov. 2000.
4. A. I. Semenko and N. I. Bokla, "Stvorennia modyfikovanykh psevdovypadkovykh poslidovnostei Holda dlia telekomunikatsiinykh system z kodovym rozpodilom kanaliv," *Telekomunikatsiini ta informatsiini tekhnologii*, №3, s. 10-14, 2014.
5. A. I. Semenko and N. I. Bokla, "Effektivnost telekomunikatsionnykh sistem s ispolzovaniem modifitsirovannykh psevdosluchainykh posledovatelnosti Golda," *Elektrosvyaz*, №3, pp. 37-40, 2014.
6. X. Peng, J. Chen, Y. Li and Y. Song, "Single Subcarrier Gold Sequences Modulated Timing Synchronization for Upstream OFDMA-PON," in *IEEE Photonics Journal*, vol. 9, no. 1, pp. 1-7, Feb. 2017, Art no. 7200207.
7. U. Grob, A. L. Welti, E. Zollinger, R. Kung and H. Kaufmann, "Microcellular direct-sequence spread-spectrum radio system using N-path RAKE receiver," in *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 5, pp. 772-780, June 1990.
8. D. Dharmappa, M. V. Mandi and S. Ramesh, "Analysis of Binary Sequences with better correlation performance compared to Gold code under different doppler conditions for GNSS applications," 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018, pp. 209-215.
9. J.-H. Yan, C.-W. Wang, K.-H. Lin and K.-M. Feng, "A Timing-Synchronization-Free WDM-Compatible Colorless DRoF Uplink System for 5G Mobile Fronthaul Employing Gold Sequence Multiplexing," 2019 Conference on Lasers and Electro-Optics (CLEO), San Jose, CA, USA, 2019, pp. 1-2.
10. A. Semenko, N. Bokla, A. Shokotko and V. Lusenko, "Construction of microwave link using channel code division based on modified pseudorandom Gold sequences," 2016 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET), Lviv, Ukraine, 2016, pp. 634-637.
11. Y. Jun and Z. Xiang-Li, "A new design of Gold-like binary sequences," 2017 2nd Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), Wuhan, China, 2017, pp. 118-124, doi: 10.1109/ACIRS.2017.7986077
12. K. Liu, Z. Zhou, A. R. Adhikary and C. Tang, "Large Sets of Binary Spreading Sequences With Low Correlation and Low PAPR via Gold Functions," in *IEEE Transactions on Information Theory*, vol. 70, no. 7, pp. 5309-5322, July 2024.
13. L. Shufeng, H. Shiyao and W. Hongda, "Analysis of DS-CMDA system using orthogonal gold sequence," 2017 7th IEEE International Conference on Electronics Information and Emergency Communication (ICEIEC), Macau, China, 2017, pp. 520-523.
14. D. Dharmappa, M. V. Mandi and S. Ramesh, "Generation of Binary Sequences Having Better Odd Correlation and Linear Complexity Property Compared to Gold Codes for use in Global Navigation Satellites Systems (GNSS) Applications," 2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC), Mysore, India, 2017, pp. 155-160.
15. N.I. Kunakh and O.M. Tkalenko, "Zastosuvannia tekhnologii RAKE yak metodu boroty z bahatopromenevistiuv u systemakh mobilnoho zviazku," *Suchasna spetsialna tekhnika*, №3(50), s. 34-41, 2017.
16. A.A. Kahno and S.M. Dyachenko "Rake-receivers in the modern telecommunication systems," *Bulletin of National Technical University of Ukraine «Kyiv Polytechnic Institute»* №34, pp. 92-96, 2007.
17. Z. Gao, Y. Mei and L. Qiao, "Sparse Signal Processing for Massive MIMO Communications," Springer: Singapore, pp. 217, 2023.
18. A.I. Semenko, N.I. Bokla, "Metod dvokratnoho zmnshennia neobkhidnoi smuhy propuskannia kanalu zviazku v telekomunikatsiini systemi z kodovym rozpodilom kanaliv," *Zviazok*, №4, c.23-25, 2011.
19. J. Boiko and L. Karpova, "Increasing Efficiency in the Correlation Processing of Information Signals for Radar," *Proceedings of Fourth International Conference on Communication, Computing and Electronics Systems. Lecture Notes in Electrical Engineering*, vol 977. Springer: Singapore, pp. 549-569, 2023.
20. J. Boiko, A. Semenko and I. Pyatin, "Features of code redundancy formation in information transmission channels", *Infocommunication and computer technologies*, vol. 2, no. 04, pp. 12-25, 2023.
21. J. Boiko, I. Pyatin, L. Karpova and O. Eromenko, "Study of the Influence of Changing Signal Propagation Conditions in the Communication Channel on Bit Error Rate," in *Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies*, vol 69, Springer: Cham, pp. 79-103, 2021.

<https://doi.org/10.31891/2219-9365-2024-79-21>

УДК 621.376.3

ШПАТАР Петро

Чернівецький національний університет імені Юрія Федьковича  
<https://orcid.org/0000-0003-4088-1458>  
[p.shpatar@chnu.edu.ua](mailto:p.shpatar@chnu.edu.ua)

НІЧИЙ Богдан

Чернівецький національний університет імені Юрія Федьковича  
<https://orcid.org/0009-0001-4855-2053>  
[b.nichyi@chnu.edu.ua](mailto:b.nichyi@chnu.edu.ua)

НІЧИЙ Сергій

Чернівецький національний університет імені Юрія Федьковича  
<https://orcid.org/0000-0003-2662-9694>  
[s.nichyi@chnu.edu.ua](mailto:s.nichyi@chnu.edu.ua)

## ДОСЛІДЖЕННЯ ЕНЕРГОЕФЕКТИВНОСТІ ТЕХНОЛОГІЙ LORA В БЕЗПРОВІДНИХ МОНІТОРИНГОВИХ МЕРЕЖАХ

В даній роботі розглядається моделювання функціонування кінцевих пристроїв (end point) компонентів безпроводних моніторингових мереж, які можна використати для побудови мережі типу "точка-точка" або "зірка" із використанням технології фізичного рівня LoRa. Технологія LoRa володіє достатньо високими параметрами по дальності. Енергоефективність даної технології досягається за рахунок накладання обмеження кількості передачі даних протягом доби. В цьому випадку досягається довготривале автономне функціонування (не менше 5 років) кінцевих пристроїв при використанні малогабаритних джерел живлення. Мета моделювання – дослідити можливість функціонування передавачів LoRa на фізичному рівні для забезпечення роботи безпроводних моніторингових мереж з максимальною періодичністю 60 с, з тривалістю автономного функціонування в межах 2,5 - 3 роки при використанні малогабаритних джерел живлення.

Ключові слова: безпроводні сенсорні мережі, LoRa, автономне функціонування, RFM95/96/97/98(W).

SHPATAR Petro, NICHYI Bohdan, NICHYI Sergiy

Chernivtsi National University named after Yury Fedkovich

## STUDY OF ENERGY EFFICIENCY OF LORA TECHNOLOGIES IN WIRELESS MONITORING NETWORKS

This work considers the modeling of the functioning of end devices (end point) components of wireless monitoring networks, which can be used to build a "point-to-point" or "star" type network using LoRa physical layer technology. LoRa technology has sufficiently high range parameters. The energy efficiency of this technology is achieved by limiting the number of data transfers during the day. In this case, long-term autonomous functioning (declared up to 5 years) of end devices is achieved when using small-sized power sources. The purpose of the simulation is to investigate the possibility of functioning of LoRa transmitters at the physical level to ensure the operation of wireless monitoring networks with a maximum frequency of 60 s, with the duration of autonomous operation within 2,5 - 3 years when using small-sized power sources. The following parameters of the final transmitters were investigated: the transmitter output power range was 4–8 dBm; minimum battery life of at least 3 years (1095 days). The data field must contain the minimum information: the identifier of the end device, the parcel number, the type of device, the condition (status) of the device. That is, the minimum length of the data field should contain 4 bytes, which will allow receiving the minimum information from the end device of the monitoring network. Simulation results for packages with parameters that allow to obtain an autonomous operation time of at least 900 days and a package period of 30 seconds at a radio signal power of 4 and 8 dBm with a power source capacity of 3200 mAh. The paper proves that when designing a wireless network based on LoRa technology, it is necessary to coordinate the parameters of the spectrum width BW and the spectrum expansion factor SF to obtain satisfactory performance characteristics of the network. The simulation results were compared with the results on mock-ups using PIC18LF4520 microcontrollers and the RFM95W (REV 1.2) LoRa module based on the RF96 chip (Semtech SX1276 analogue). As the results of this work show, LoRa technology can be used to build wireless monitoring networks with much more frequent data transmissions for premises with transmitter output power of 2-8 dBm and short messages.

Keywords: wireless sensor networks, LoRa, autonomous operation, RFM95/96/97/98(W).

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ

### ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Безпроводні моніторингові мережі призначені для використання збору даних, моніторингу стану фізичних об'єктів, навколишнього середовища та ін. Такий тип мереж відносять до безпроводних сенсорних мереж. Залежно від призначення та алгоритму роботи функціональна база таких мереж є різноманітною [1]. Побудова безпроводних сенсорних мереж із використанням радіопередачі даних по технології LoRa все більше знаходить застосування. Запатентований принцип модуляції дозволяє отримати досить пристойні (до 15 км) відстані передачі на відкритій місцевості. Технологія LoRa розроблялась із метою використання в безпроводних сенсорних мережах LoRaWAN із максимальною швидкістю 50 кбіт/с. Наразі використання протоколу LoRaWAN розповсюджене в системах: "розумне місто", автоматизації контролю енергетичних

систем, інтернет речей та ін. Дана мережа передбачає функціонування приймачів та передатчиків в протоколі LoRaWAN [2]. Перспективність розвитку даної технології зумовлюється також нарощуванням випуску мікросхем прийомо-передавачів LoRa. Зокрема на ринку присутній широкий вибір таких мікросхем та модулів, на основі яких можливо побудувати пристрої LoRa:

- WLR089U0 від MICROCHIP TECHNOLOGY [3];
- RFM95/96/97/98(W) від HOPE MICROELECTRONICS [4];
- SX1276/77/78/79 від SEMTECH CORPORATION [5].

Для створення енергоефективних безпроводних моніторингових мереж в приміщеннях використовуються мережі, побудовані на основі використання стандарту IEEE802.15.4 [6], яскравим прикладом якого є мережа на основі протоколу ZigBee. Існують пристрої, які володіють великою енергоефективністю, автономність роботи яких при потужності радіосигналу 4 dBm і періоді передачі 30 с доходить до 3-х років. Швидкість передачі в них дорівнює 250 кбіт/с, в діапазоні 2,4 ГГц. Наприклад, SoC CC2530, із вбудованими приймачами-передавачами, формують MAC рівень стандарту IEEE802.15.4 на апаратному рівні і це дає можливість створювати бюджетні моніторингові системи. Проте, з досвіду авторів, такі системи мають малий радіус дії (до 100-150 м). В приміщенні, де існують поглинаючі матеріали (стіни, меблі), даний радіус зменшується до 35 м. Щоб розширити дальність передачі даних від кінцевих пристроїв (end point) до точок доступу (Access Point) в таких мережах використовують технологію mesh-мереж, яка також має свої недоліки (складність налаштувань, а при збільшенні компонентів ускладнюється маршрутизація повідомлень). Для багатьох безпроводних моніторингових використовуються мережі типу "зірка" або деревоподібні мережі.

Технологія передачі LoRa дозволяє будувати мережі типу "зірка" без розгортання мережі LoRaWAN. В такому випадку використовується фізичний та MAC рівень системи OSI. Решта рівнів (транспортний, прикладний) розробляються розробниками системи, що дозволить розробити енергоефективні мережеві системи під конкретні параметри. Так при розробці автономних автоматизованих моніторингових систем необхідно враховувати наступні критичні моменти:

- 1) максимально дозволена потужність випромінювання згідно санітарних норм, крім цього даний фактор впливає на п.2.
- 2) мінімальний час автономного функціонування (для прикладу пожежні радіоканальні системи моніторингу повинні мати термін не менше 3 років [7]).
- 3) вимоги пунктів 1 і 2 визначають формфактор пристрою, на який значний вплив мають розміри автономних джерел живлення.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Технологія LoRa - це метод модуляції з розширенням спектра, в якій дані кодується широкосмуговими ЛЧМ (Chirp Spread Spectrum) імпульсами з частотою, що збільшується (up-chirp), або зменшується (down-chirp) на деякому часовому інтервалі. Математично ЛЧМ сигнал подається у вигляді виразу:

$$x = A_0 \cos \left( \omega_0 t + \frac{\mu}{2} t^2 \right), \text{ причому } \frac{T_{sym}}{2} \leq t < \frac{T_{sym}}{2} \quad (1)$$

де  $\omega_0 = 2\pi f_0$  - центральна частота радіосигналу;

$\mu = BW/T_{sym}$  - швидкість зміни частоти радіосигналу;

$T_{sym} = 2^{SF}/BW$  - тривалість радіосигналу;

$BW$  - ширина спектру радіосигналу, яка визначає нижню та верхню границю смуги робочих частот  $f_H = (f_0 - BW/2)$  -  $f_B = (f_0 + BW/2)$ ;

$SF$  - коефіцієнт розширення спектру (7 - 12), який визначає базу радіосигналу  $B = BW \cdot T_{sym} = 2^{SF}$ .

База радіосигналу визначає умови кодування байта даних [8], швидкість передачі символу.

Кількість символів в блоці даних фізичного рівня визначається наступним виразом [9]:

$$N_{sym} = 8 + \max \left( \text{ceil} \left( \frac{8PL - 4SF + 28 + 16CRC - 20H}{4(SF - 20DE)} \right) (CR + 4), 0 \right) \quad (2)$$

де,  $\text{max}(\text{ceil} \dots)$  - операція (функція) округлення до найбільшого додатного цілого;  $PL$  - довжина поля даних в байтах;  $SF$  - коефіцієнт розширення спектру;  $CRC$  - включення контрольної суми (два байта  $CRC$  можна включати в довжину поля даних - *payload*), якщо  $CRC = 1$  поле контрольної суми включається в пакет;  $H$  - індикатор поля заголовку, якщо  $H = 1$  поле заголовку в пакет не включається;  $DE$  - оптимізація передачі даних при малих значень швидкостей переді даних ( $T_{sym} > 16$  мс), якщо  $DE = 0$  оптимізація виключена;  $CR$  - швидкість виправлення помилок ( $1 \div 4$ ), яка рекомендована до використання, коли в мережі передбачаються великі імпульсні завади .

Загальний формат пакету MAC рівня LoRa має вигляд на рис 1. Пакет включає поля преамбули, заголовка із контрольною сумою, самих даних та контрольної суми поля даних.

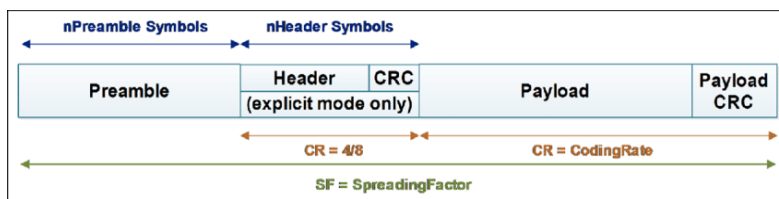


Рис. 1. Формат пакету MAC рівня LoRa

Енергетичні параметри фізичного рівня та наповненість пакету MAC рівня при передачі даних визначають наступні характеристики безпроводної системи:

- Завадостійкість передачі даних, що включає достовірність переданої інформації, яка передається, та кількість втрачених переданих пакетів.
- Витрати енергії на передачу пакету і, як наслідок, тривалість автономної роботи передавача.

Завдання різних параметрів для формул (1) і (2) приводить до різної структури пакету, який передається кінцевими пристроями (ED) до точок доступу (AP), наслідком чого є довжина пакету, час передачі, завадостійкість пакету, споживана енергоємність передачі. Цілий ряд наукових публікацій присвячений математичному моделюванню процесів передачі та функціонуванню пристроїв LoRa [10]. Моделювання та прикладні дослідження функціонування пристроїв LoRa висвітлюють можливості експлуатації в мережах LoRaWAN [11,12]. В таких роботах визначаються параметри для досягнення максимальної дальності ( $\geq 10 - 15$  км) із досягненням максимальної енергоефективності [13].

#### ФОРМУВАННЯ ЦІЛЕЙ СТАТТІ

Метою дослідження є проведення математичного моделювання функціонування передавачів LoRa на MAC рівні мережі та визначити можливість отримання оптимальних енергетичних і часових параметрів, які забезпечують роботу безпроводної моніторингової мережі з періодичністю передачі даних один раз в 30 с, та тривалістю автономного функціонування в межах 2,5 - 3 роки, при використанні малогабаритних джерел живлення, наближених до типу CR123A (3В 1600 мА\*год).

#### ВИКЛАДЕННЯ ОСНОВНОГО МАТЕРІАЛУ

Нами пропонується розглянути результати досліджень по встановленню оптимальних енергетичних властивостей пристроїв, які передають дані за технологією LoRa, для побудови мережі типу “зірка”. Мета моделювання - вибрати оптимальні параметри фізичного рівня для технології LoRa для таких мереж, які за енергетичними характеристиками подібні до мереж з використанням IEEE802.15.4. Вихідними параметрами кінцевих передавачів є наступні вимоги: діапазон вихідної потужності передавача 4 –8 dBm; мінімальний час автономної роботи не менше 3 років (1095 днів). Поле даних повинно містити мінімальну інформацію: ідентифікатор кінцевого пристрою, номер посилки, тип пристрою, стан (статус) пристрою. Тобто мінімальна довжина поля даних повинна містити 4 байта, яка дасть змогу отримувати мінімальну інформацію від кінцевого пристрою моніторингової мережі. Так як передбачається, що пристрої працюють з наперед заданими параметрами, то поле заголовку (*Header*) ми виключили.

Для отримання енергетичних, часових характеристик нами використано формулу (2), та інші формули розрахунку параметрів передачі пакета згідно технічним паспортам на вироби (dataset) SX1276/77/78/79, і RFM95/96/97/98 (формули ідентичні) та програмний пакет “LoRaModemCalculatorTool”, від фірми “Semtech Corporation” для проектування мереж з використанням технології LoRa, на базі мікросхем SX1272/1273/1276/1277 [14].

Програмний пакет “LoRaModemCalculatorTool” дає змогу отримати теоретичну чутливість приймача при різних параметрах фізичного рівня пакета (рис. 2).



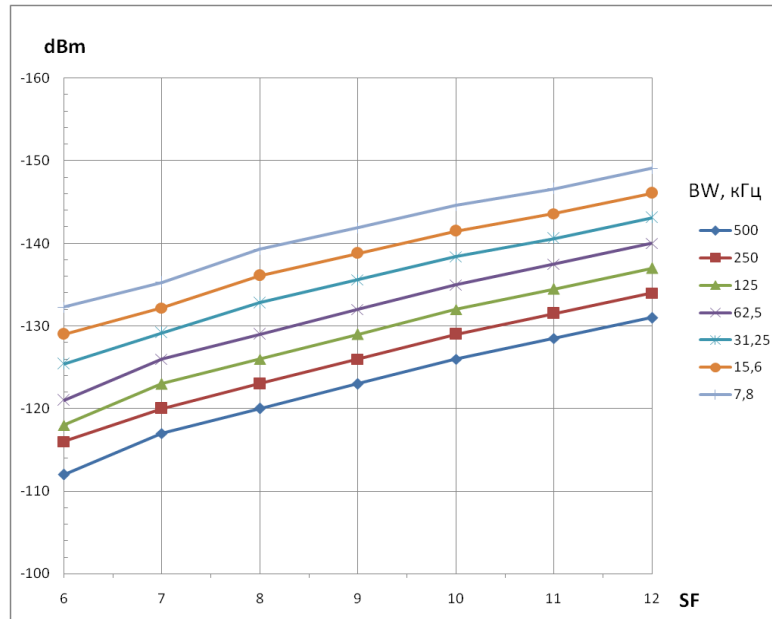


Рис 2. Залежність чутливості приймача LoRa від ширини спектру BW і коефіцієнта розширення спектру SF

Значення величини ширини спектру  $BW$  і коефіцієнта розширення спектру  $SF$  впливає не тільки на чутливість приймача (бюджет каналу), але й на швидкість передачі даних, відповідно, і на час радіофіру, що в сукупності визначає тривалість автономної роботи кінцевого пристрою.

Для розрахунку бітової швидкості використовується формула

$$R_b = SF * \frac{BW}{2^{SF}} * \frac{4}{4+CR} \quad (3)$$

Використовуючи дані формули, отримуються значення, які наведені в таблицях 1 і 2.

Таблиця 1.

**Залежність швидкості передачі бітів  $R_b$  (біт/сек) від  $BW$  і  $SF$**

SF\BW	500 кГц	250 кГц	125 кГц	62,5 кГц	31,25 кГц	15,6 кГц	7,8 кГц
6	37500	18750	9375,00	4687,50	2343,75	1170,00	585,00
7	21875	10937,5	5468,75	2734,38	1367,19	682,50	341,25
8	12500	6250	3125,00	1562,50	781,25	390,00	195,00
9	7031,25	3515,63	1757,81	878,91	439,45	219,38	109,69
10	3906,25	1953,13	976,56	488,28	244,14	121,88	60,94
11	2148,44	1074,22	537,11	268,56	134,28	67,03	33,52
12	1171,88	585,94	292,97	146,49	73,24	36,56	18,28

При передачі даних за технологією LoRa краще враховувати тривалість передачі символу  $T_{sym}$  – час, протягом якого передається  $SF$  бітів даних. Крім того, слід враховувати кількість бітів, які включаються в пакет для корекції помилок. Тоді протягом  $T_{sym}$  передається  $SF * \frac{4}{4+CR}$  інформаційних бітів. В таблиці 2 наведені результати розрахунків тривалості передачі символу в залежності від значень параметрів  $BW$  і  $SF$ , які вираховуються за формулою

$$T_{sym} = \frac{2^{SF}}{BW} \quad (4)$$

Таблиця 2.

**Залежність тривалість передачі символу  $T_{sym}$  (мкс) від  $BW$  і  $SF$**

SF\BW	500 кГц	250 кГц	125 кГц	62,5 кГц	31,25 кГц	15,6 кГц	7,8 кГц
6	0,13	0,26	0,51	1,02	2,05	4,1	8,21
7	0,26	0,51	1,02	2,05	4,1	8,21	16,41
8	0,51	1,02	2,05	4,1	8,21	16,41	32,82
9	1,02	2,05	4,1	8,19	16,38	32,82	65,64
10	2,05	4,1	8,19	16,38	32,77	65,64	131,28
11	4,1	8,19	16,38	32,77	65,54	131,28	262,56
12	8,19	16,38	32,77	65,64	131,07	262,56	525,13

Графіки залежності швидкості передачі бітів  $R_b$  (біт/сек) і тривалості передачі символу  $T_{sym}$  (мкс) від  $BW$  і  $SF$  на основі таблиць 1 і 2, зображені на рис. 3.

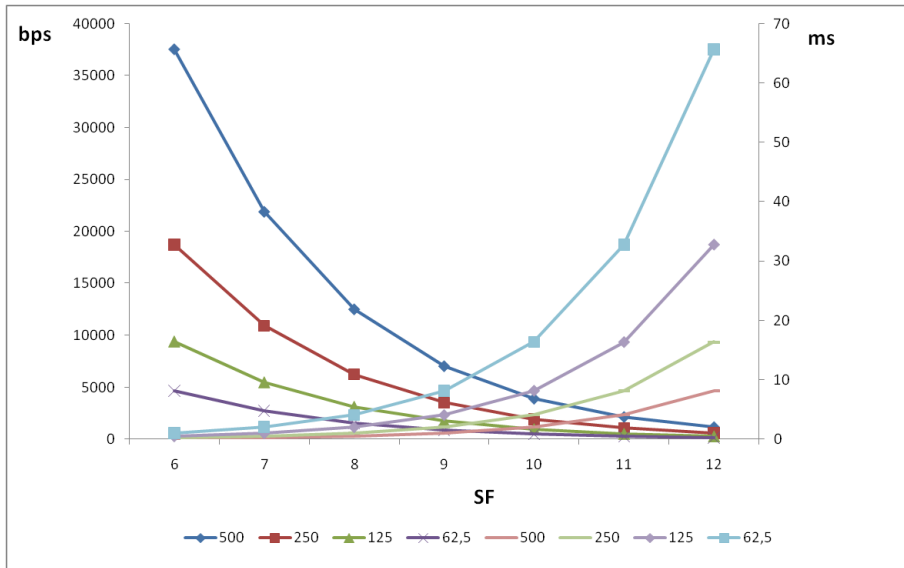


Рис. 3. Залежність швидкості передачі бітів  $R_b$  (біт/сек) і тривалості передачі символу  $T_{sym}$  (мкс) від  $BW$  і  $SF$

Аналіз графіків на рисунках 2 і 3 вказує, що при проектуванні безпроводної мережі на базі технології LoRa необхідно узгоджувати між собою параметри  $BW$  і  $SF$  для отримання задовільних характеристик роботи мережі. Збільшення параметра  $SF$  приводить до зростання порогу чутливості приймача та часу тривалості передачі символу і, відповідно, збільшення енергозатрат. В той же час зростання параметру  $BW$  приводить до зменшення порогу чутливості приймача та часу тривалості передачі символу. Підбираючи дані параметри, можна отримати необхідні енергетичні характеристики.

Біти даних за технологією LoRa передаються символами. В одному символі кодується  $SF$

бітів. Згідно формули (2), для передачі даних мінімальна кількість символів дорівнює 8. В таблиці 3 наведені результати розрахунків кількості символів, які необхідні для кодування байтів поля даних.

Таблиця 3.

Кількості символів, які необхідні для кодування байтів поля даних

SF/ Байт	CR = 1							CR = 4						
	6	7	8	9	10	11	12	6	7	8	9	10	11	12
4	13	13	13	13	8	8	8	16	16	16	16	8	8	8
5	13	13	13	13	13	13	8	16	16	16	16	16	16	8
6	18	13	13	13	13	13	13	24	16	16	16	16	16	16
7	18	18	13	13	13	13	13	24	24	16	16	16	16	16
8	18	18	18	13	13	13	13	24	24	24	16	16	16	16
9	23	18	18	18	13	13	13	32	24	24	24	16	16	16
10	23	23	18	18	18	13	13	32	32	24	24	24	16	16
11	23	23	18	18	18	18	13	32	32	24	24	24	24	16
12	28	23	23	18	18	18	18	40	32	32	24	24	24	24

З таблиць випливає, що кількість переданих символів є незмінною при зміні байтів в полі даних (горизонтальні стовбці). Значення величини  $CR$  суттєво впливає на значення кількості символів, необхідних для кодування поля даних (наведені дані для  $CR = 1$  та  $CR = 4$ ). На рис. 4. зображені графіки залежність кількості символів для поля даних, при значеннях  $SF = 8$  і  $SF = 10$  ( $CR=1$ ). Графіки побудовані для двох вибірок з метою кращої наглядності.

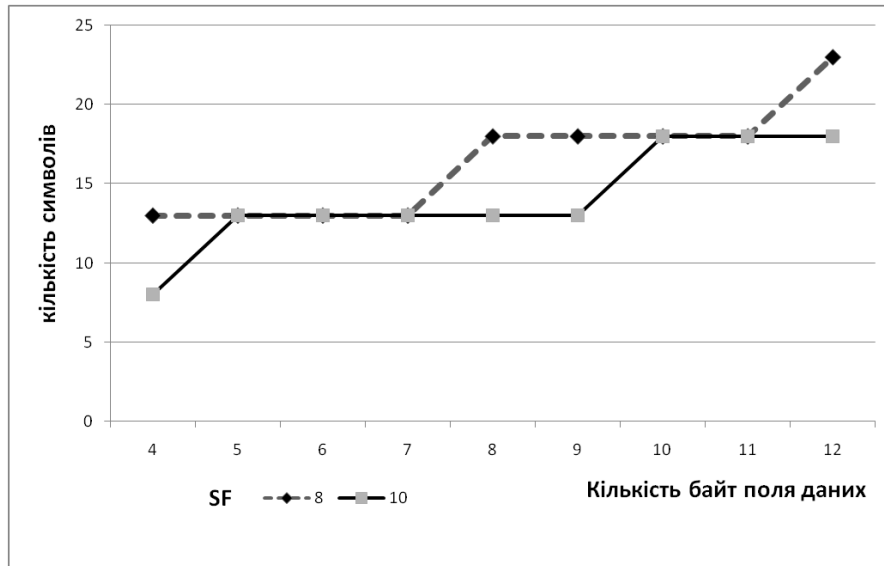


Рис. 4. Залежність кількості символів для поля даних при двох значеннях SF (CR=1)

Загальний час передачі пакету (рис. 1) визначається наступною формулою:

$$T_{packet} = T_{preamble} + T_{payload} \quad (5)$$

де

$$T_{payload} = N_{sym} T_{sym} \quad (6)$$

$$T_{preamble} = (n_{preamble} + 4,25) T_{sym} \quad (7)$$

$n_{preamble}$  – кількість символів, які задаються програмно.

З аналізу графіків рис. 4 можна зробити висновок, що, якщо вибрати поле даних з 5 ÷ 9 байтів для SF=10, кращою енергоефективністю буде володіти пакет із 9 байтів. Проте слід зауважити, що  $T_{sym}$  має залежність від значень BW і SF. Програмний пакет “LoRaModemCalculatorTool” розраховує величини струму (таб. 4) під час передачі для відповідної вихідної потужності передавача.

Таблиця 4

**Залежність величина споживаного струму від потужності радіосигналу**

$P_{вих}$ передачі	$I_{ж}$ передачі
1 dBm	23 мА
2 dBm	24 мА
3 dBm	24 мА
4 dBm	24 мА
5 ÷ 8 dBm	25 мА

В таблицях 5 наведені основні результати моделювання для пакетів з параметрами, що дозволяють отримати автономний час функціонування не менше 900 діб та періодом посилок 30 сек при потужності радіосигналу 4 і 8 dBm. Ємність джерела живлення 3200 мА\*год. В таблицю не включені результати розрахунків для  $T_{sym} > 16$  мс. Це дає змогу виключити, відповідно до рекомендації виробників, оптимізацію передачі даних для малих значень швидкостей, що дозволяє зменшити кількість символів при передачі даних відповідно до формули (2). Дані в таблиці викладені для наступних параметрів: довжина преамбули дорівнює 8,25 (4+4,25) символів; довжина поля даних – 8 байт; байти контрольної суми виключені. Поле заголовку в пакет не включається; оптимізація передачі даних для малих значень швидкостей виключена; CR – швидкість виправлення помилок дорівнює 1 (4/5).

Присутність однакових результатів розрахунків при різних параметрах BW і SF є наслідком формули (2), яка вказує, що однаковою кількістю символів можна передати різну кількість бітів.

Результати моделювання порівнювались з результатами на макетах із використанням мікроконтролерів PIC18LF4520 і модуля LoRa RFM95W (REV 1.2) на базі мікросхеми RF96 (аналог Semtech SX1276), згідно схеми на рис. 5 з використанням радіопередаючого тракту з максимальною вихідною потужністю 14 dBm. Програмне забезпечення передавача та приймача створено на мові C. Алгоритм функціонування передбачав 8-ми байтну посилку в полі даних. Метою експериментальних вимірювань було встановлення кореляції розрахункових та реальних енергетичних параметрів макетного зразка.

Розрахункова кількість днів автономної роботи

BW	SF					
	Вихідна потужність 4 dBm					
	7	8	9	10	11	12
500	24253	12262	7606	3816	1911	956
250	12262	6165	3816	1911	956	
125	6165	3091	1911	958		
62,50	3091	1547	956			
	Вихідна потужність 8 dBm					
	7	8	9	10	11	12
500	23304	11777	7304	3664	1835	918
250	11777	5920	3664	1835	918	
125	5920	2968	1835	918		
62,50	2868	1486	918			

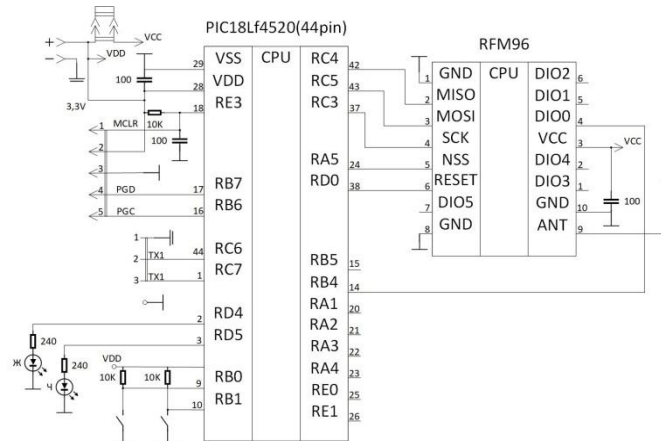


Рис.5. Принципова електрична схема пристрою

Передача 8-ми байтів в полі даних здійснювалась з наступними параметрами:  $BW = 125$ ,  $SF = 9$ , запрограмована вихідна потужність сигналу 4 dBm, напруга живлення 3 В. В цьому випадку, згідно з теоретичними розрахунками, бюджет каналу становить 133 dB,  $T_{sum}=4,1$  мс,  $T_{посилки}=87,04$  мс. Під час вимірювань модуль RFM95W між передачами пакетів був запрограмований в режим *Standby*. Струм споживання, в режимі *Sleep* визначали окремо. Вимірювання споживаного струму і часу передачі пакета проводилися за допомогою цифрового осцилографа SmartDS5932E. Осцилограма падіння напруги на резисторі ( $R=6,4$  Ом), включеного в коло живлення RFM95W, зображено на рис. 6 (скріншот екрана осцилографа).

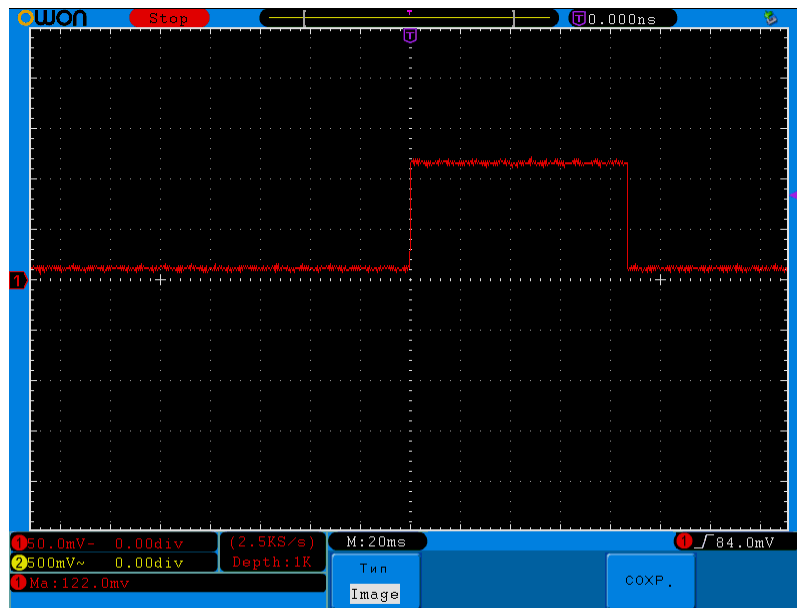


Рис. 6. Типова часова залежність напруги на резисторі в колі колектора.  
 $BW = 125$ ,  $SF = 9$ ,  $R=6,4$  Ом

З осцилограм розраховані значення струму споживання під час передачі ( $I=U/R=U/6,4$  Ом) і тривалість передачі пакету. В таблиці 6 представлені результати змодельованих та визначених з вимірювань параметрів роботи передавача RFM95W.

Таблиця 6

Порівняльна таблиця параметрів передавача

Параметр	Розрахунки "LoRaModemCalculatorTool"	Визначені з осцилограм
Струм режиму <i>Sleep</i>	100 нА	~ 200 нА (окреме вимірювання)
Струм режиму <i>Standby</i>	1,6 – 1,8 мА (для RF96W)	~ 2 мА
Струм режиму передачі	24 мА	~ 19 мА
Тривалість передачі	87,04 мс	~ 88 – 90 мс
Автономний час роботи	1911 день	1731 день

З результатів, наведених в таблиці 6, випливає, що є можливість використання технології LoRa для розгортання безпроводних сенсорних мереж із низьким енергоспоживанням. Як показали виміри передавач на висоті 1,5 м від підлоги, з налаштованими параметрами:  $BW = 125$ ,  $SF = 9$ , вихідна потужність сигналу 4 dBm, дає змогу отримати стабільний прийом в структурі "точка-точка", на відстані до 250 м в середині приміщення (кількість цегляних стін не менше 10, кімнати не порожні).

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

В порівнянні із протоколом IEEE802.15.4, технологія LoRa має меншу швидкість передачі корисних даних і є більш енергозатратною. Проте, як показують результати даної роботи, технологія LoRa може використовуватись для побудови безпроводних моніторингових мереж із значно частішими передачами даних для приміщень із вихідною потужністю передавача 2 – 8 dBm та короткими повідомленнями. Важливо при проектуванні мережі знаходити оптимальні значення  $BW$  і  $SF$ . Зменшення  $BW$  приводить до збільшення часу радіопередачі, збільшення енергоспоживання, але збільшується чутливість приймача. При зменшенні  $SF$  зменшується час передачі, але й зменшується чутливість приймача. Дані в таблиці 5 вказують, що існує можливість обирати оптимальні значень  $BW$  і  $SF$  в залежності від умов експлуатації мережі без суттєвої втрати енергоефективності.

### Література

1. Степанов М. Зв'язок елементів сенсорної мережі системи моніторингу/ М. Степанов, В.Лаврінченко // Вимірювальна та обчислювальна техніка в технологічних процесах. – 2023. – №4. – С.199-206
2. LoRa Alliance Technical committee. LoRaWAN™ 1.0.2 Regional Parameters. Revision: B. [Електронний ресурс]. – Режим доступу: [https://lora-alliance.org/wp-content/uploads/2020/11/lorawan\\_regional\\_parameters\\_v1.0.2\\_final\\_1944\\_1.pdf](https://lora-alliance.org/wp-content/uploads/2020/11/lorawan_regional_parameters_v1.0.2_final_1944_1.pdf).
3. WLR089U0. – Режим доступу : <https://www.microchip.com/en-us/product/wlr089u0>
4. RFM95W. Режим доступу: <https://www.hoperf.com/modules/lora/RFM95W.html>
5. SX1276 Режим доступу <https://www.semtech.com/search?q=%E2%80%A209SX1276>
6. Галкин П.В. Особенности имитационного моделирования беспроводных сенсорных сетей стандарта IEEE 802.15.4/ П.В. Галкин // Телекоммуникаційні та інформаційні технології. – 2015. – No 2. –С. 67-79.
7. ДСТУ CEN/TS 54-14:2021 Системи пожежної сигналізації та оповіщення. Частина 14. – [Чинний від 2021-07-01]. – ДП «Український науково-дослідний і навчальний центр проблем стандартизації, сертифікації та якості» (ДП «УкрНДНЦ»)
8. AN1200.22. LoRa™ Modulation Basics. Revision 2. [Електронний ресурс]. – Режим доступу: <https://www.frugalprototype.com/wp-content/uploads/2016/08/an1200.22.pdf>
9. SX1272/3/6/7/8. LoRa Modem Design Guide. AN1200.13. Revision 1 [Електронний ресурс]. – Режим доступу: [https://www.openhacks.com/uploadsproductos/loradesignguide\\_std.pdf](https://www.openhacks.com/uploadsproductos/loradesignguide_std.pdf)
10. Md. Noor-A-Rahim. Hybrid Chirp Signal Design for Improved Long-Range (LoRa) Communications// Md. Noor-A-Rahim, M. Omar Khyam, Apel Mahmud, Xinde Li, Dirk Pesch, H. Vincent Poor. // *Signals* 2022, 3, 1–10. Режим доступу: <https://doi.org/10.3390/signals3010001>.
11. Congduc Pham. Dense Deployment of LoRa Networks: Expectations and Limits of Channel Activity Detection and Capture Effect for Radio Channel Access/ Congduc Pham, Muhammad Ehsan // *Sensors* 2021, 21, 825. Режим доступу: <https://doi.org/10.3390/s21030825>.
12. D.Monica Satyavathi. Real-time hidden data transmission using lora./ D.Monica Satyavathi, B.Vandana Mala, Ch.Veera Vamsi, Ch.Chiranjeevi, Ch.Neeraj// *EPR International Journal of Research and Development*. – 2022. –v.4. –№9. –p.90-94.

13. Husam Rajab. Evaluation of Energy Consumption of LPWAN Technologies/ Husam Rajab, Tibor Cinkler, Taoufik Bouguera // Режим доступу: <https://www.researchsquare.com/article/rs-343897/v1>
14. SX1272/3/6/7/8: LoRa Modem. Low Energy Consumption Design. AN1200.17. Revision 1 [Електронний ресурс]. — Режим доступу: [https://www.openhacks.com/uploadsproductos/loralowenergydesign\\_std.pdf](https://www.openhacks.com/uploadsproductos/loralowenergydesign_std.pdf)

### References

1. Stepanov M. Zviazok elementiv sensornoi merezhi systemy monitorynhu/ M. Stepanov, V.Lavrinenko // Vymiriuvalna ta obchysliuvalna tekhnika v tekhnolohichnykh protsesakh. – 2023. - №4. – S.199- 206
2. LoRa Alliance Technical committee. LoRaWAN™ 1.0.2 Regional Parameters. Revision: B. [Електронний ресурс]. — Режим доступу: [https://lora-alliance.org/wp-content/uploads/2020/11/lorawan\\_regional\\_parameters\\_v1.0.2\\_final\\_1944\\_1.pdf](https://lora-alliance.org/wp-content/uploads/2020/11/lorawan_regional_parameters_v1.0.2_final_1944_1.pdf).
3. WLR089U0. – Режим доступу: <https://www.microchip.com/en-us/product/wlr089u0>
4. RFM95W. Режим доступу: <https://www.hoperf.com/modules/lora/RFM95W.html>
5. SX1276 Режим доступу <https://www.semtech.com/search?q=%E2%80%A2%09SX1276>
6. Halkyn P.V. Osobennosti umytatsyonnoho modelyrovannya besprovodnykh sensornykh setei tandarta IEEE 802.15.4/ P.V. Halkyn // Telekomunikatsiini ta informatsiini tekhnolohii. – 2015. – No 2. –S. 67-79.
7. DSTU CEN/TS 54-14:2021 Systemy pozhezhnoi syhnalizatsii ta opovishchuvannia. Chastyna 14. – [Chynnyi vid 2021–07–01]. — DP «Ukrainskyi naukovo-doslidnyi i navchalnyi tsentr problem standartyzatsii, sertyfikatsii ta yakosti» (DP «UkrNDNTs»).
8. AN1200.22. LoRa™ Modulation Basics. Revision 2. [Електронний ресурс]. — Режим доступу: <https://www.frugalprototype.com/wp-content/uploads/2016/08/an1200.22.pdf>
9. SX1272/3/6/7/8. LoRa Modem Design Guide. AN1200.13. Revision 1 [Електронний ресурс]. — Режим доступу: [https://www.openhacks.com/uploadsproductos/loradesignguide\\_std.pdf](https://www.openhacks.com/uploadsproductos/loradesignguide_std.pdf)
10. Md. Noor-A-Rahim. Hybrid Chirp Signal Design for Improved Long-Range (LoRa) Communications// Md. Noor-A-Rahim, M. Omar Khyam, Apel Mahmud, Xinde Li, Dirk Pesch, H. Vincent Poor. // *Signals* 2022, 3, 1–10. Режим доступу: <https://doi.org/10.3390/signals3010001>.
11. Congduc Pham. Dense Deployment of LoRa Networks: Expectations and Limits of Channel Activity Detection and Capture Effect for Radio Channel Access/ Congduc Pham, Muhammad Ehsan // *Sensors* 2021, 21, 825. Режим доступу: <https://doi.org/10.3390/s21030825>.
12. D.Monica Satyavathi. Real-time hidden data transmission using lora./ D.Monica Satyavathi, B.Vandana Mala, Ch.Veera Vamsi, Ch.Chiranjeevi, Ch.Neeraj// EPRA International Journal of Research and Development. – 2022. –v.4. –№9. –p.90-94.
13. Husam Rajab. Evaluation of Energy Consumption of LPWAN Technologies/ Husam Rajab, Tibor Cinkler, Taoufik Bouguera // [Електронний ресурс] — Режим доступу: <https://www.researchsquare.com/article/rs-343897/v1>
14. SX1272/3/6/7/8: LoRa Modem. Low Energy Consumption Design. AN1200.17. Revision 1 [Електронний ресурс]. — Режим доступу: [https://www.openhacks.com/uploadsproductos/loralowenergydesign\\_std.pdf](https://www.openhacks.com/uploadsproductos/loralowenergydesign_std.pdf)

<https://doi.org/10.31891/2219-9365-2024-79-22>

УДК 681.324

БАРАБАШ Олег

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

<https://orcid.org/0000-0003-1715-0761>

[bar64@ukr.net](mailto:bar64@ukr.net)

МАКАРЧУК Андрій

Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"

<https://orcid.org/0000-0002-6422-7488>

[makarchukandriy1999@gmail.com](mailto:makarchukandriy1999@gmail.com)

## НОВИЙ МЕТОД ДЕШИФРУВАННЯ ДІАГНОСТИЧНОЇ ІНФОРМАЦІЇ НА ОСНОВІ ІНТЕРВАЛЬНИХ ОЦІНОК

*Активний розвиток обчислювальних систем та, як наслідок, збільшення сфер їх застосування ставить перед науково-технічним прогресом ряд нових задач, однією з найважливіших з яких є організація діагностування систем. Оскільки габарити та складність обчислювальних систем постійно зростає, то логічною є спроба організувати їх діагностування в автономному режимі, при мінімальному, а то й взагалі без участі людини.*

*На даний момент, такий спосіб діагностування можна реалізувати, наприклад, програмно в два етапи: збір діагностичної інформації та її дешифрування. З деякими проблемами можна стикнутися саме на етапі дешифрування, оскільки багато запропонованих на даний момент дешифрування діагностичної інформації вимагають досить багато обчислень. В силу того, що сучасні обчислювальні системи можуть містити дуже велику кількість елементів, об'єм діагностичної інформації може бути досить великим. В наслідок цього застосування деяких існуючих методів дешифрування діагностичної інформації може займати багато часу. Це пов'язано з тим, що, на відміну від збору діагностичної інформації, розроблені на даний момент методи її дешифрування вимагають велику кількість обчислень: деякі з них мають експоненційну складність відносно кількості елементів системи. А даний фактор може бути критичним при розробці, наприклад, великих інформаційних систем чи систем критичної інфраструктури. Це породжує необхідність в розробці нових, більш швидких методів дешифрування.*

*В даній роботі пропонується новий метод дешифрування діагностичної інформації, який є дещо швидшим від багатьох існуючих. Даний факт досягається завдяки поєднанню інтервальних оцінок та розподілу Бернуллі, які не вимагають надмірної кількості обчислень. Також коротко описано особливості даного методу, одна з яких наглядно демонструється на двох конкретних прикладах інформаційних систем.*

*Ключові слова: інформаційні системи, діагностування, самодіагностування, обчислювальна складність, дешифрування діагностичної інформації, програмне забезпечення методів діагностування.*

BARABASH Oleg, MAKARCHUK Andriy

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

## A NEW METHOD OF DIAGNOSTIC INFORMATION DESCRIPTION BASED ON INTERVAL ASSESSMENTS

*The active development of computer systems and, as a result, the increase in the scope of their application presents scientific and technical progress with a number of new tasks, one of the most important of which is the organization of system diagnostics. Since the dimensions and complexity of computer systems are constantly growing, it is logical to try to organize their diagnostics in an autonomous mode, with minimal, or even without human intervention.*

*At the moment, such a method of diagnostics can be implemented, for example, by software in two stages: collection of diagnostic information and its decryption. Some problems can be encountered precisely at the decryption stage, since many currently proposed decryption of diagnostic information require quite a lot of calculations. Due to the fact that modern computer systems can contain a very large number of elements, the volume of diagnostic information can be quite large. As a result of this application of some existing methods, diagnostic information decryption can take a long time. This is due to the fact that, in contrast to the collection of diagnostic information, the currently developed methods of its decryption require a large number of calculations: some of them have an exponential complexity relative to the number of system elements. And this factor can be critical when developing, for example, large information systems or critical infrastructure systems. This creates the need to develop new, faster decryption methods.*

*This paper proposes a new method of diagnostic information decryption, which is somewhat faster than many existing ones. This fact is achieved thanks to the combination of interval estimation and the Bernoulli distribution, which do not require an excessive number of calculations. The features of this method are also briefly described, one of which is clearly demonstrated on two specific examples of information systems.*

*Keywords: information systems, diagnostics, self-diagnostics, computational complexity, diagnostic information decryption, diagnostic methods software.*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Під час проектування інформаційних систем іноді виникає необхідність в забезпечення в її автономній роботі. Цей фактор зазвичай породжує ряд додаткових задач, які слід вирішити до введення

системи в експлуатацію. Однією з таких задач є задача самодіагностування, тобто щоб використовувана в подальшому система могла в подальшому виявляти несправність елементів самостійно, без стороннього втручання.

Одним із найпопулярніших методів організації самодіагностування системи є його організація у вигляді взаємоконтролю елементів системи, абстрагуючись від їх внутрішньої структури, яка свого часу була запропонована Препаратом [1, 2]. При такому підході покладається розгляд пар машин системи  $(v_i, v_j)$ , у яких перша машина проводить контроль другої і називається перевіряючою, а друга – тією, що перевіряється. В даній ситуації результатом такого контролю (або, як іноді кажуть, елементарної перевірки) буде число  $r_{ij}$ , яке рівне 0, якщо машини  $v_i$  та  $v_j$  справні, 1 – якщо машина  $v_i$  справна, а  $v_j$  – ні, і або 0, або 1 з однаковою ймовірністю, якщо несправна машина  $v_i$ .

Маючи набір результатів перевірок (який прийнято називати діагностичною інформацією або, іноді, синдромом) одних елементів системи другими можна проводити так званий процес дешифрування цього набору результатів, який полягає в остаточному визначенні того, вважати той чи інший елемент системи справним чи ні. Нині розроблені методи дешифрації синдрому, як от описаний в [3], дозволяють якісно вирішувати дану задачу. Однак, головною їх проблемою є висока обчислювальна складність, особливо, з урахуванням, що елемент  $v_i$  може перевіряти елемент  $v_j$  неодноразово, внаслідок чого в синдромі конкретній парі машин може відповідати не один результат перевірки (0 або 1), а їх набір. На основі цього виникає необхідність розробки нового методу дешифрації діагностичної інформації, можливо, не такого достовірного, але, на протитягу існуючим, швидшого.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Актуальність та, відповідно, інтерес до діагностування систем та її елементів почала зростати зі зростанням застосування обчислювальної техніки. Так, наприклад, в [1, 4] автор описує підходи до діагностування елементів обчислювальних систем. Однак, в [4] акцент робиться на технічній стороні даного питання, а в [1] розглядаються ідеї більш загального характеру, застосовні й до самодіагностування. В [3] автор розширює спектр питань, пов'язаних з самодіагностуванням інформаційних систем, ще й питаннями, що стосуються забезпечення їх функціональної стійкості. Деякі концепції, що описуються в [4], можна зустріти і, наприклад, [5]. Однак, в більшості робіт, що стосуються безпосередньо самодіагностування, часто на етапі збору діагностичної інформації [3] або її дешифрування використовують елементи байєсівського аналізу [3], модульної алгебри [6], теорії динамічних систем [1, 7], тощо.

Коли ми говоримо безпосередньо про методи дешифрації діагностичної інформації, то багато цих методів, розроблених на даний момент, є досить складними в плані кількості обчислень. Так, наприклад, метод, описаний в [3], має експоненційну обчислювальну складність, а алгоритм, описаний в [8] – кубічну. Оскільки зараз розміри реальних обчислювальних та інформаційних систем є досить немалими, то, відповідно, й об'єм діагностичної інформації може бути достатньо великий [9]. В таких випадках результуюча обчислювальна складність може бути небажаною або взагалі неприпустимою, навіть, на фоні обчислювальних потужностей сучасних приладів. Зокрема, значимість обчислювальної складності методів дешифрації може грати важливу роль при проектуванні різних електронних [10] та радіоелектронних [11] пристроїв, де обчислювальні потужності можуть бути невеликими, а використання таких пристроїв може широко використовуватися в самих різних напрямках [12-14]. На основі цього постає проблема в розробці альтернативних, можливо, не таких точних, але, все ж таки, швидших методів дешифрації діагностичної інформації.

### Основна ідея методу самодіагностування.

Нехай структура розглядуваної інформаційної системи представлена у вигляді неорієнтованого графа  $G = G(V, E)$ , де  $V$  – множина вершин графа, що відповідають машинам в системі, а  $E$  – множина ребер, що відповідають лініям зв'язку між цими машинами. В задачі самодіагностування цікавим є орієнтований граф  $\tilde{G}$ , який прийнято називати діагностичним [1].

Нехай машина  $v_i$  перевіряє  $iN$  машин  $v_{k_1}, v_{k_2}, \dots, v_{k_{iN}}$ . З другої сторони, якщо машина  $v_i$  перевіряє машину  $v_j$ , то поставимо у відповідність цій парі машин  $(v_i, v_j)$  множину проведених елементарних перевірок  $\rho_{ij} = \{r_{ij}^1, r_{ij}^2, \dots\}$ . Вважатимемо, що елементарні перевірки проводилися по схемі



Препарата, тобто результат  $k$ -ої перевірки  $j$ -ої машини  $i$ -ою  $r_{ij}^k$  рівний 0, якщо  $i$ -а та  $j$ -а машини справні, рівний 1, якщо  $i$ -а машина справна, а  $j$ -а машина – ні, або ж 0 чи 1 з однаковою ймовірністю, якщо несправна  $i$ -а машина. Позначимо як  $\bar{\rho}_{ij}$  середні арифметичні відповідних множин елементарних перевірок  $\rho_{ij}$ . Побудувавши довірчий інтервал для кожного середнього  $\bar{\rho}_{ij}$  з рівнем довіри  $\alpha$  [15], поставимо у відповідність парі машин  $(v_i, v_j)$  число  $\omega_{ij}$ , яке рівне 1, якщо число 1 входить в побудований довірчий інтервал для середнього  $\bar{\rho}_{ij}$ , і рівне 0 в протилежному випадку.

Нехай  ${}^+N$  – кількість  $\omega_{ji}$ , рівних одиниці, а  ${}^-N$  – кількість  $\omega_{ji}$ , рівних нулю. Очевидно, що має місце рівність  ${}^+N = {}^+N + {}^-N$ . Тоді, спираючись на розподіл Бернуллі, вважатимемо, що ймовірність несправності машини  $v_i$ , яку позначимо як  $p_i^-$ , буде обчислюватися по формулі

$$p_i^- = C_{iN}^+ \alpha^{+N} (1-\alpha)^{-N}, \quad (1)$$

а ймовірність справності  $p_i^+$ , відповідно, по формулі

$$p_i^+ = 1 - p_i^- = 1 - C_{iN}^+ \alpha^{+N} (1-\alpha)^{-N}. \quad (2)$$

#### Приклади використання описаного методу.

Розглянемо даний метод на прикладі інформаційної системи, яку можна представити наступним графом

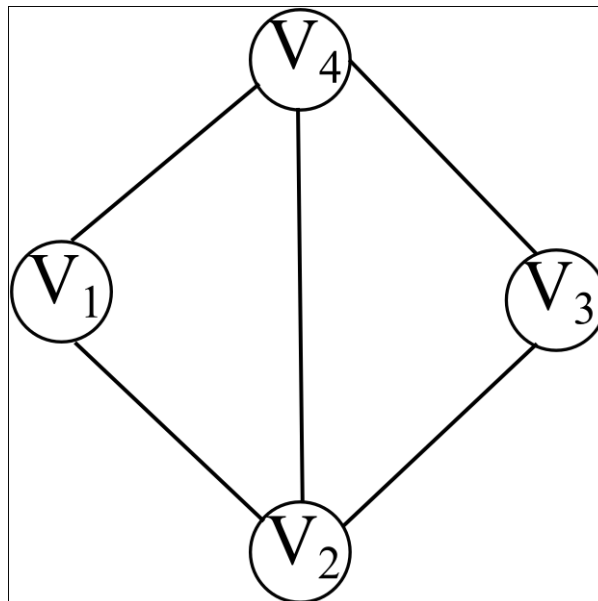


Рис. 1. Структура інформаційної системи, для якої застосовується описаний метод самодіагностування

Вважатимемо, що діагностичний граф даної системи матиме вигляд

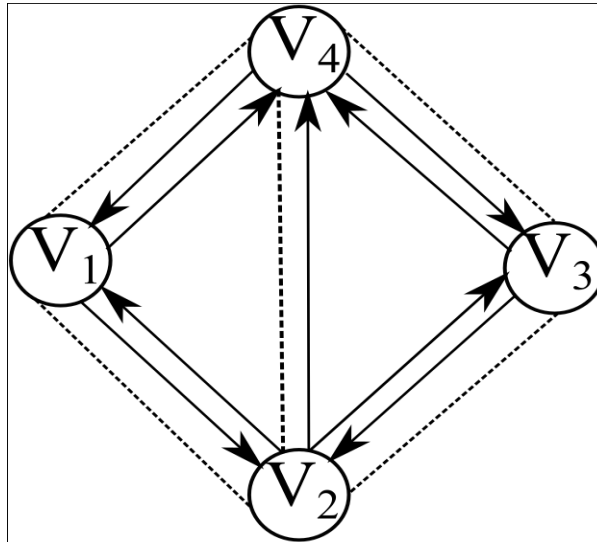


Рис. 2. Діагностичний граф інформаційної системи, представленій на рис. 1

Нехай на момент початку самодіагностування мається діагностична інформація, представлена в таблиці 1.

Таблиця 1.

Діагностична інформація для інформаційної системи, представленій на рис. 1

Множина елементарних перевірок	Набір значень
$\rho_{12}$	0, 0, 0, 1, 0
$\rho_{14}$	1, 0, 1, 1
$\rho_{21}$	1, 0
$\rho_{23}$	1, 1, 0
$\rho_{24}$	0, 1
$\rho_{32}$	1, 0
$\rho_{34}$	0, 1
$\rho_{41}$	0, 0, 1, 0
$\rho_{43}$	1, 0

На основі таблиці 1 обчислимо всі значення  $\omega_{ji}$ . Для кращої наочності побудуємо таблицю, в якій кожен  $i$ -ий рядок відповідає машині, яку діагностують, кожен  $j$ -ий стовбець відповідає машині, яка діагностує. Кожній клітинці, якій відповідає пара вершин  $(v_i, v_j)$ , запишемо значення відповідне цій парі число  $\omega_{ji}$ , якщо  $i$ -а машина перевіряє  $j$ -у, і залишимо незаповненою в протилежному випадку.

Таблиця 2.

Значення  $\omega_{ji}, \forall i, j = 1, 2, 3, 4$  на основі таблиці 1

$i \backslash j$	$v_1$	$v_2$	$v_3$	$v_4$
$v_1$	-	0	-	0
$v_2$	0	-	0	0
$v_3$	-	0	-	0
$v_4$	0	-	0	-

Тепер, використовуючи  $i$ -ий рядок та використовуючи описаний вище алгоритм, можна обчислити ймовірність справності машини  $v_i$ . Таким чином, для випадків, коли  $\alpha$  рівне 0,05 та 0,01, отримаємо ймовірності справності всіх машин (див. табл. 3).

Таблиця 3.

**Ймовірність справності машин системи, представленої на рис. 1, за допомогою (2)**

Номер машини	Ймовірність справності при	
	$\alpha = 0,05$	$\alpha = 0,01$
1	0,0975	0,02
2	0,905	0,02
3	0,0975	0,02
4	0,142625	0,03

Як видно з результатів, представлених в таблиці 3, при використанні описаного вище методу дешифрації діагностичної інформації, можуть спостерігатися досить різочі результати, навіть, при невеликій зміні параметра  $\alpha$ . До найпростіших причин такого явища можна віднести наступні: специфіка довірчих інтервалів як інтервального методу або ж замала кількість діагностичної інформації. У випадку, якщо причиною такої колосальної різниці в результатах є саме специфіка довірчих інтервалів, то описаний вище метод є некоректним, а тому виникає необхідність у його доопрацюванні. Якщо ж проблема в недостатньому об'ємі діагностичної інформації, то це вдасться виявити на прикладі, де її буде більше.

Так, для прикладу, розглянемо систему не з чотирьох, а з шести машин, структуру якої представлено на рис. 3.

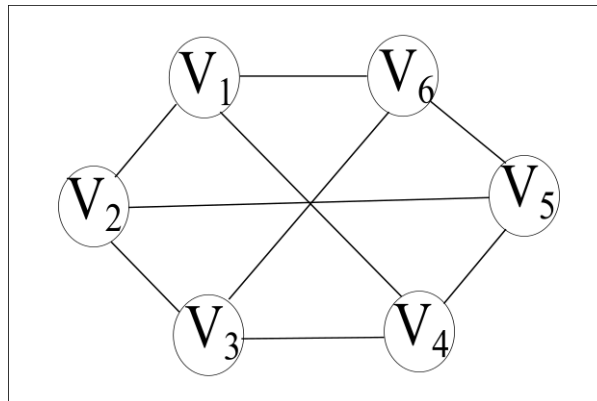


Рис. 3. Приклад структури іншої інформаційної системи, для якої тестується описаний метод самодіагностування

Вважатимемо, що діагностичний граф цієї системи має вигляд, представлений на рис. 4.

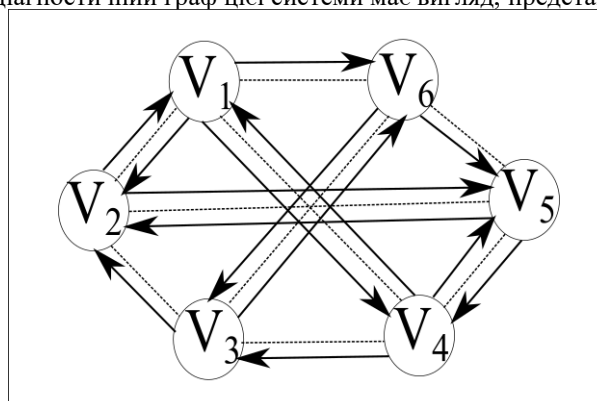


Рис. 4. Діагностичний граф системи, структура якої представлена на рис. 2

Нехай в процесі проведення елементарних перевірок була зібрана діагностична інформація, яку, розбивши на множини  $\rho_{ij}$ , можна представити так, як на таблиці 4.

Таблиця 4.

**Діагностична інформація, отримана перед самодіагностуванням системи на рис. 3**

Множина елементарних перевірок	Набір значень
$\rho_{12}$	0, 0, 1, 0
$\rho_{14}$	1, 0, 0, 0, 1, 0, 0
$\rho_{16}$	0, 0, 1, 0, 0
$\rho_{21}$	0, 0, 1, 0, 0, 1, 0, 0, 1
$\rho_{25}$	1, 0, 0, 0, 0, 0, 0, 0, 0
$\rho_{32}$	1, 1, 1, 0, 1, 1, 1
$\rho_{36}$	0, 0, 1, 0, 0, 0, 0, 0
$\rho_{41}$	0, 1, 0
$\rho_{43}$	1, 0, 0, 0, 0, 0, 0
$\rho_{45}$	1, 0, 1, 0, 0, 1
$\rho_{52}$	0, 0, 0, 1, 0
$\rho_{54}$	0, 1, 0, 0, 1, 0, 0
$\rho_{63}$	0, 0, 0, 0, 0, 1, 1, 0
$\rho_{65}$	0, 1, 0, 1, 1, 0, 0, 1

Провівши обчислення, аналогічні до обчислень в попередньому прикладі, обчислимо ймовірності справності кожної з машин для випадків  $\alpha = 0,05$  та  $\alpha = 0,01$ . Отримані результати представлені в таблиці 5.

Таблиця 5.

**Ймовірності справності машин системи, представлені на рис. 3, за допомогою описаного методу, на основі таблиці 4**

Номер машини	Ймовірність справності при	
	$\alpha = 0,05$	$\alpha = 0,01$
1	0,905	0,02
2	0,865	0,03
3	0,9975	0,9999
4	0,9975	0,9999
5	0,993	0,97
6	0,9975	0,98

Отже, як можна бачити з наведених прикладів, для достатньо точного висновку описаний вище метод самодіагностування вимагає достатньо багато діагностичної інформації з однієї сторони та те, щоб кожна з машин діагностувала максимальна кількість інших. В протилежному випадку результуючі ймовірності справності машин в інформаційній системі будуть занадто грубими, що, в свою чергу, може вимагати використання інших методів дешифрації інформаційної інформації.

**Особливості застосування описаного методу.**

Перед використанням того чи іншого методу вирішення поставленої задачі завжди важливо розуміти специфіку цього методу. Якщо ж говорити про описаний вище метод самодіагностування, то варто відмітити наступні його особливості:

- На точність обчислених ймовірностей справності машин впливає об'єм діагностичної інформації. Коли машина  $v_i$  перевіряє машину  $v_j$ , то більша кількість елементарних перевірок останньої забезпечить точніший кінцевий висновок машини  $v_i$  щодо її справності з імовірністю  $1 - \alpha$ .
- Кількість машин, що перевіряють розглядувану, явно впливає на точність обчисленої ймовірності справності. Неважко побачити, що формула (2) явним чином залежить від загальної кількості

машин, що перевіряли розглядувану (параметр  $i^N$ ). Це (в силу математичних особливостей розподілу Бернуллі) гарантує, що збільшення цієї кількості підвищує точність обчислення ймовірності справності машини. Звідси випливає, що чим більша за габаритами інформаційна система, тим більша вірогідність достатньо точно обчисленої ймовірності справності кожної з машин.

• Описаний метод досить швидкий. Описаний вище метод самодіагностування є достатньо швидким: його обчислювальна складність в середньому становить  $O(nV \log n)$ , де  $n$  – це кількість машин в розглядуваній системі,  $V$  – об'єм діагностичної інформації.

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

#### І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

В роботі описано новий метод дешифрації діагностичної інформації, який базується на поєднанні теорії статистичних гіпотез та теорії ймовірностей. В процесі дослідження було математично обґрунтовано ідею методу та на кількох конкретних методах показано застосування.

При вивченні описаного методу дешифрації було виявлено його головні переваги та недоліки. Головною перевагою даного методу є його швидкість, яка, на відміну від деяких раніше запропонованих методів, є поліноміальною. Однак, попри швидкість та відносну простоту в реалізації він має й недоліки. Основним недоліком є висока залежність від об'єму діагностичної інформації та розмірів інформаційної системи, що може робити його досить неточним.

#### Література

1. Машков В.А. Самодіагностування бортових обчислювальних систем. Київ : КВВАІУ, 1989. 82 с.
2. Preparata F.P., Metzger G., Chien R.T. On the Connection Assignment Problem of Diagnosable Systems IEEE Transactions on Electronic Computers, EC. 1967. Vol.16, No.6. pp.848–854.
3. Барабаш О.В. Побудова функціонально стійких розподілених інформаційних систем. Київ, 2004, 226 с.
4. Машков В.А. Контроль та діагностування та контроль бортових цифрових обчислювальних пристроїв. Київ : КВВАІУ, 1991. 84 с.
5. Машков О.А., Барабаш О.В. Топологічні критерії та показники функціональної стійкості складних ієрархічних систем. Моделювання та інформаційні технології: Збірник наукових праць, 2003. № 25: С. 29-35.
6. Mashkov V.A., Mashkov O.A. Interpretation of diagnosis problem of system level self-diagnosis (2015) Mathematical Modeling and Computing, 2015, Vol. 2, No. 1, pp. 71 – 76.
7. Пічкур В.В., Капустян О.В., Собчук В.В. Теорія динамічних систем. Луцьк. Вежа-Друк, 2020. 348 с.
8. Albin L.C.P., Duarte E.P., Ziwich R.P. A generalized model for distributed comparison-based system-level diagnosis Journal of the Brazilian Computer Society, 2004, Vol. 10, No.3, pp.42–54.
9. Raskin L., Karpenko V., Ivanchykhin Y., Sokolov D. Diagnosis of systems under conditions of small initial data sampling. Advanced Information Systems, 2023. Vol.7, No.3, pp. 39– 43.
10. Вишнівський В.В., Савран В.О. Розробка діагностичної моделі цифрового пристрою для спектрального методу на основі енергодинамічних перехідних процесів в шині живлення. Наукові записки Державного університету інформаційно-комунікаційних технологій. 2017. Т. 2, № 46. С. 12–17.
11. Вишнівський В.В., Жердев М.К., Креденцер Б.П., Кузавков В.В., Редзюк Є.В. Безконтактний індукційний метод діагностування радіоелектронних блоків. Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. 2013. № 43. С. 17–23.
12. Жердев М.К., Вишнівський В.В., Пампуха І.В., Скуйбіда О.Ю. Напрями розвитку систем контролю технічного стану і діагностування складних технічних систем. Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. 2006. Т. 3. С. 22–25.
13. Глухов С., Сакович Л., Рижов Є., Бабій О., Гальоса А. Перспективні напрями підвищення надійності радіоелектронної техніки на основі фізичного діагностування з використанням інформаційних технологій. Військово-технічний збірник. 2022. № 26, С. 42–48.
14. Вишнівський В.В., Гахович С.В., Катін П.Ю., Круценко В.В. Пристрій для діагностування цифрових ТЕЗ з використанням параметрів енергодинамічного процесу. Вісник Київського національного університету імені Тараса Шевченка. Військово-спеціальні науки. 2003. № 6. С. 71–75.
15. Siegel A.F. Practical business statistics. 4th ed. Washington : Williams, 2000. 800 p.

---

**References**

1. Mashkov V. . Self-diagnosis of on-board computer systems. Kyiv: KVAIU, 1989. 82 p.
2. Preparata F.P., Metzger G., Chien R.T. On the Connection Assignment Problem of Diagnosable Systems (1967) IEEE Transactions on Electronic Computers, EC-Vol.16 No.6, pp.848–854.
3. Barabash O.V. Construction of functionally stable distributed information systems. Kyiv, 2004, 226 p.
4. Mashkov V.A. Control and diagnostics and control of on-board digital computing devices. Kyiv: KVAIU, 1991. 84 p.
5. Mashkov O.A., Barabash O.V. Topological criteria and indicators of functional stability of complex hierarchical systems. Modeling and information technologies: Collection of scientific papers, 2003. No. 25: P. 29-35.
6. Mashkov V.A., Mashkov O.A. Interpretation of diagnosis problem of system level self-diagnosis Mathematical Modeling and Computing. 2015. Vol. 2 No. 1, pp. 71 – 76.
7. Pichkur V.V., Kapustyan O.V., Sobchuk V.V. Theory of dynamic systems. Lutsk. Tower-Druk, 2020. 348 p.
8. Albini L.C.P., Duarte E.P., Ziwich R.P. A generalized model for distributed comparison-based system-level diagnosis Journal of the Brazilian Computer Society, 2004, Vol. 10, No.3, pp.42–54.
9. Raskin L., Karpenko V., Ivanchykhin Y., Sokolov D. Diagnosis of systems under conditions of small initial data sampling. Advanced Information Systems, 2023. Vol.7, No.3, pp. 39–43.
10. Vyshnivskiy V.V., Savran V.O. Development of a diagnostic model of a digital device for the spectral method based on energy-dynamic transients in the power bus. Scientific notes of the State University of Information and Communication Technologies. 2017. Vol. 2, No. 46. P. 12–17.
11. Vyshnivskiy V.V., Zherdev M.K., Kredentser B.P., Kuzavkov V.V., Redzyuk E.V. Contactless induction method of diagnosing radioelectronic units. Collection of scientific works of the Military Institute of Taras Shevchenko Kyiv National University. 2013. No. 43. P. 17–23.
12. Zherdev M.K., Vyshnivskiy V.V., Pampukha I.V., Skuibida O.Yu. Directions of development of technical condition control systems and diagnostics of complex technical systems. Collection of scientific works of the Military Institute of Taras Shevchenko Kyiv National University. 2006. Vol. 3. P. 22–25.
13. Glukhov S., Sakovich L., Ryzhov E., Babii O., Galiosa A. Prospective directions for increasing the reliability of radio-electronic equipment based on physical diagnostics using information technologies. Military and technical collection. 2022. No. 26, pp. 42–48.
14. Vyshnivskiy V.V., Gakhovich S.V., Katin P.Yu., Krutsenko V.V. The device for diagnosing digital thermal power plants using the parameters of the energy-dynamic process. Bulletin of Taras Shevchenko Kyiv National University. Military special sciences. 2003. No. 6. P. 71–75.
15. Siegel A.F. Practical business statistics. 4th ed. Washington : Williams, 2000. 800 p.

<https://doi.org/10.31891/2219-9365-2024-79-23>

UDC 004

MUSTAFAYEV Kamran

Institute of Control Systems  
Ministry of Science and Education Republic of Azerbaijan  
<https://orcid.org/0009-0000-1762-7457>  
e-mail: [k.mustafa91@bk.com](mailto:k.mustafa91@bk.com)

## INFORMATION TECHNOLOGIES IN TRANSPORT

(Industry 4.0; digital transport corridors and platforms; modeling and simulation transport systems, etc.)

*The aim of the article is to consider the significant importance of the use of information technology in modern society as the main technical tool for the development of transport systems. It also addresses the expansion of economic relations among countries and the acceleration of integrative processes, emphasizing the transportation system's role as a fundamental pillar of each country's national economy and socio-economic essence of transport services.*

*In the rail transport it has been investigated using smart trains, smart locomotives and smart wagons in the connection of rail transport with other types of transport for safety of trains, transmission capacity of the lines, speed of shipping control and also increasing employee productivity through computer systems of automated control systems. It has been explored the importance of monitoring the level of service more smoothly and qualitatively with the mutual integration of these systems as well.*

*The novelty of the article the novelty of the article is that information technologies in transport are integrated into transport systems of the 4th Industrial revolution, researching the development of an intelligent management system for modeling digital transport corridors and platforms, and achieving optimization of traffic flow and traffic management of the application of intelligent transport. In conclusion, in the future, smart locomotives should be integrated with smart cars, they are local controls for all components of the car – braking equipment; electrical equipment, service equipment, etc. satellite technology should be used to locate moving objects and monitoring systems. These technologies should be used in conjunction with radio communication and radar sensing of railway facilities from satellites and determine their coordination. In the future, connected management should become the basis of a single transport complex for all transport sectors and all forwarders for operational monitoring and forecasting of situations for all mobile units and each link involved in the technological process. The application of intelligent transport will provide by predicting traffic flow and optimizing traffic management.*

*Key words: transportation system, national economy, productivity, monitoring, computer systems, train safety,*

МУСТАФАЙВ Камран

Інститут систем управління Міністерство науки та освіти Республіки Азербайджан

## ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ НА ТРАНСПОРТІ

(Індустрія 4.0; цифрові транспортні коридори та платформи; моделювання та моделювання транспортних систем і т.)

*Мета статті розглядати значну важливість використання інформаційних технологій у суспільстві як основного технічного інструменту розвитку транспортних систем. У ньому також розглядаються питання розширення економічних зв'язків між країнами та прискорення інтеграційних процесів, наголошуючи на ролі транспортної системи як фундаментальної опори національної економіки кожної країни та соціально-економічної сутності транспортних послуг. На залізничному транспорті досліджувалося використання «розумних» поїздів, «розумних» локомотивів та «розумних» вагонів при з'єднанні залізничного транспорту з іншими видами транспорту для забезпечення безпеки поїздів, пропускної спроможності ліній, швидкості контролю перевезень, а також підвищення продуктивності праці працівників за допомогою комп'ютера. системи автоматизованих систем керування. Досліджено важливість більш плавного та якісного моніторингу рівня обслуговування при взаємній інтеграції цих систем. Новизна статті полягає в тому, що інформаційні технології на транспорті інтегровані в транспортні системи 4-ї промислової революції, досліджуються розробки інтелектуальної системи керування для моделювання цифрових транспортних коридорів та платформ, а також досягнення оптимізації транспортних потоків та керування дорожнім рухом із застосуванням інтелектуального транспорту. На закінчення зазначається, що у майбутньому «розумні» локомотиви мають бути інтегровані з «розумними» вагонами, вони є локальне управління всіма компонентами вагона – гальмівним устаткуванням; електрообладнання, сервісне обладнання і т. д. Для визначення розташування об'єктів, що рухаються, і систем моніторингу слід використовувати супутникові технології. Ці технології повинні використовуватися спільно з радіозв'язком та радіолокаційним зондуванням залізничних об'єктів із супутників та визначати їх координату. В майбутньому пов'язане управління має стати основою єдиного транспортного комплексу для всіх транспортних галузей та всіх експедиторів для оперативного моніторингу та прогнозування ситуацій для всіх мобільних одиниць та кожної ланки, задіяної у технологічному процесі. Застосування інтелектуального транспорту забезпечить прогнозування транспортних потоків та оптимізацію керування дорожнім рухом.*

*Ключові слова: транспортна система, національна економіка, продуктивність, моніторинг, комп'ютерні системи, безпека поїздів*

**Information technology** ("Information technologies" in English) - the term information is derived from the Latin word "informatio" and means reporting the object, event and fact, interpretation, explanation. The word technology formed from the combination of Greek word "techno" (mastery, skill) and "logos" (study, cognition). It expresses a body of knowledge about methods and means for carrying out production processes and those processes themselves.

Information technology is the process of receiving information (information product) about the state of the object, process and event which investigated using methods and means for data collection, transmission and processing.

### **Purpose in information technology**

Purpose in information technology is information production for being analyzed by a human and making decision based on it. Information technology has passed several evolutionary stages determined by the creation of new technical means for the development of scientific and technical progress and information processing.

In modern society the main technical means of information processing technology is considered computers that affected seriously to the concept of development and use of technological processes and also quality of information. Application of computers in the information environment and use of telecommunication means brought to the new stage of development of information technology. So, new information technology stage has started.

The development of the transportation system plays a crucial role for the Republic of Azerbaijan.



**Pic. 1. Highways intersections**

Besides the infrastructures of state economy, such as energy, communication, education, and health care, the transport also plays an important role for achieving social, economic, foreign policy and other public priorities by ensuring initial demands of community life.

The dynamic development gained in all spheres of public life in our country in recent years increased significantly the geopolitical significance of our republic.

Along with all of these, important tasks such as active participation in socio-economic processes in the global economic space, effective access to foreign markets, and the forming of a competitive national industry have become a necessity.

The issues such as globalization, rapid integration and sustainable development etc. which are typical for modern era have also specific importance for our country.

The expansion of economic connections among the countries of the world and the acceleration of integrative processes have confirmed that the transport system is one of the main pillars of the national economy of each country.

Researchers who are engaged with clarification of the socio-economic nature of transport services proved that this system is an important condition for large-scale reproduction.

From this view, during the large-scale radical reforms in the economy, it is significant to formulate and implement a policy that takes into account the specifics of transport, its role in economic and social processes.

That's why the sustainable development of the transport system requires a large amount of capital investment, the construction and reconstruction of a large number of major transport hubs.





Fig. 2. Crossing roads through the bridge over the river

Currently, along with all spheres of life in the country, special attention is paid to the systematic and comprehensive development of the road industry. Consequently, today the renewal and modernization of the road industry of our republic are expanding and construction programs aimed at increasing the strength and power of our state, improving the welfare of our people are successfully continuing.

During the last years, all planned infrastructure projects and social programs in Azerbaijan were completed successfully. In addition, projects such as the reconstruction of the road sector, construction of new roads, an overhaul of existing roads, which are important areas of social infrastructure, have been successfully continued. Numerous overpasses, bridges, tunnels, underpasses and overpasses were built in the capital and regions.

Currently, Azerbaijan is implementing infrastructure projects to create the North-South and East-West international transport corridors. Work is underway to complete the following important projects to turn Azerbaijan into an international transport hub.

According to the report of the World Economic Forum, in terms of the state of roads, Azerbaijan ranking first in the CIS and has shown the highest result. Azerbaijan is ranked 10th on the Asian continent and 24th among 141 countries (2021).

Currently, in the liberated territories, work is underway to build road infrastructure.

"Zafar Yolu" (The "Road to Victory") corridor extends from the Hajigabul-Minivan-Zangezur corridor to the city of Shusha that founded on November 16, 2020. The length of the road is 101 kilometers. Zangilan-Horadiz highway which is another important project and has great strategic importance, construction process continues. The total length of this road is 124 kilometers and consists of 4-6 lanes. The construction of the Hadrut-Jabrayil-Shukurbayli highway also started. The beginning of this road starts from Hadrut settlement of Khojavend region and passes through Shukurbayli village and Jabrayil region. The length of the road is 43 kilometers. One of the road infrastructure projects implemented in the liberated areas is the Barda-Agdam highway. This road starts from the city of Barda. The length of this road, called Barda-Agdam, is 44.5 kilometers. One of the road infrastructure projects implemented in the liberated areas is the Khudaferin-Gubadli-Lachin and Khanlig-Gubadli highways.

The length of the Khudaferin-Gubadli-Lachin highway will be 66 kilometers. This road will be separated from the Khanlig settlement and a new 17-kilometer highway will be built in Gubadli. The total length of the designed roads is 83 km.

In addition, the 29-kilometer-long Tartar-Chayli-Sugovushan-Talysh highway was restored.



**Pic. 3. Transit transportation by rail transport**

Railways are one of the important areas of the developed transport system in our country.

Currently, our republic, located in a favorable geographical area, is actively involved in the implementation of all international and regional projects.

In recent years, intensive work has been done to improve the railway transport system in terms of its expansion, modernization, and renewal of the railway with international standards. This process aimed to improve the country's railway network. Therefore, Azerbaijan's active participation in the implementation of international and regional transport corridors, Baku-Tbilisi-Kars and North-South projects has also been ensured.



**Pic. 4. Sea transport in international economic relations**

It is impossible to imagine international economic relations without the participation of maritime transport. Comparing with other roads, maritime transport is the cheapest transport system.

In order to get advantages from maritime transport, investments are being made in a number of countries around the world. The Baku International Sea Trade Port started its operation in 2018. Along with port operations in Azerbaijan and abroad, the modernity and high carrying capacity of ships are some of the important contexts.

Currently, along with the transport fleet, Azerbaijan Caspian Shipping Company CJSC includes specialized fleet and ship repair plants. Our transport fleet consists of 51 ships, including 20 tankers, 13 ferries, 15 universal dry cargo ships, 2 Ro-Ro ships, as well as 1 technical ship and 1 floating workshop. The specialized fleet consists of 204 ships.

One of the important areas of the transport system is air transport. In the last years, 7 airports have been built or reconstructed in the country. Currently, there are six international airports (Baku, Nakhchivan, Ganja, Lankaran, Gabala, Zagatala) in the country. Yevlakh airport is a local airport in Azerbaijan. Currently, 3 more airports are planned to build in the liberated territories (Fuzuli, Zangilan and Lachin).

### **Automated control systems in railway transport**

Automated system for operational control of transport was prepared to create and maintain a data model for forecasting and continuously planning of the transport process, operational work of road enterprises.

Automated control system sorting station, cargo station and container station – according to automated control systems for cargo yard, cargo station, container station - is designed for planning the entry of trains at the station, dispersing them, forming and dispatching trains, processing technological and train documents.

Information provides to maintain wagon model and organize information and information services for station employees, reduce idle time of trains and wagons, improve conditions of automated control systems in railway transport and increase labor productivity of station employees. Automated system for keeping executed movement schedule is an automated dispatcher control center - designed for centralized and operational management of the transport process in the railway network as a whole or separately in the training ground.

Using network polygons, information boards, controlled computers, graphic color displays, etc. it provides improved conditions, productivity and transportation quality of staff equipped with collective and individual means that reflect the current state of the transportation process.

Automated control systems that regulate the movement of trains on a railway line.

Automated management systems in railway transport rely on the use of workstations with all automated systems and the use of automated workplaces related to transport process management and information provision. Operators of station technological centers, station attendants, locomotive and car warehouses, shunting dispatchers, as well as train dispatchers and other operational heads of traffic, engineering staff and each workstation is equipped with a special computing complex or computer center, personal computers and screens related to communication.

Passenger trains (except suburban trains) often pass sorting stations without stopping. So, stations where a lot of work is done with wagons and where technical routes, field trains, collecting and transmission trains are drawn up, sorted and destroyed based on the train layout plan are called sorting stations. Unemployed transit freight trains from sorting stations are accepted to special transit parks, but working transit freight trains, field, collector and transmission trains are accepted to the reception park.

Automated control systems in railway transport are sent from the station after locomotive crews or locomotives of transit trains are changed. After the technical and commercial inspection of the trains which arrive for sorting and preparation of the sorting sheet, they are sent to the hill and from here they are released on the tracks of the sorting park. After the train crew is fully assembled in the sorting park, it is transferred to the dispatch park. Here, after passing technical and commercial inspections again, the train locomotive takes this train to the destination station.

The functional composition of the set of applied tasks includes:

- processing of information about the train arriving at the station, the train being set up;
- information and reference system;
- analysis, accounting and report on operation of the station;
- planning of station works.

Automation of operational data entry for automated control systems is very complex in rail transport. Therefore, issuing a full sheet for the train to dispatched is limited by the process of manually adjusting the composition information based on the results of the natural extraction of car numbers.

In railway transport, automated control systems are introduced as part of the train, with carriage numbers coming from a keyboard via a teletype connected to a computer at the train's flat stations.

The computer-generated sort list is compared with a pre-made manual sort list in terms of filling accuracy and precision. In addition, the "machine" in the sorting list provides information on the control numbers of the head and tail wagons of each trailer, the total number of physical wagons directed to each joint venture track, clearly visible signs of conventional signs for trains. If necessary, re-purchase of sorting sheet for jobs is given on special request.

When automated control systems in rail transportation are transferred from joint venture to software with trains, the person on duty in the park informs the operator about the cancellation post and the operator writes down the numbers of the cars as when canceling the train at the station entrance.

After the train leaves the station, the operator in the departure park informs about the train numbers and departure time. Information about the actual departure of the train is compiled based on the their and received information and entered to the computer.

It can be accessed automated control systems through a telegraph station or an automatic telephone exchange in railway transport. Remote data processing and control of data transmission over communication channels is carried out by means of data processing, including computer interface devices with data transmission equipment, data transmission devices, signal conversion devices (modems), subscriber station equipment.

Automated railway traffic control systems cover all levels of railway management. At the same time, it makes up 72.2% of freight and 41.8% of passenger transport from all types of transport.

Railway transport provides 62.5% of all cargo transportation within the country. The railways of our country carry out more than half of the cargo turnover of the global railway network. There is an established operating system, a system of norms and rules based on technological processes related to planning of transport and operational activities, organization of vehicle flows, stations, schedules of operational trains of warehouses and other units, technical regulation and operational planning of operational works.

Automated control systems in railway transport consist of optimal planning and maintenance of railway network operation mode, ensuring the best use of transportation means, high economic indicators and labor productivity.

### **Various automated control in railway transport technologies**

Automated control systems in railway transport are in a new stage of today's technical and economic development of Azerbaijan railways through smart passenger train. In economics engineering and in technology scientific progress results are used.

In recent years, a large majority increase in volume of transportation by rail transport has been obtained with the application of modern and promising science workings related to the organization technologies, services, equipment, automated management systems and transportation process.

Today, new technologies are required for providing stable activity of industry in a market environment.

The automated railway control system was designed to organize all production activities based on the wide spread of methods of the centralized automated management and modern analysis, management, modeling, logistics and forecasting, as well as computer and information technologies on Russian railway lines. The system in all industry should integrate automated control systems into a single complex. System was designed to manage fixed schedule execution, safety of trains, capacity of lines and vehicles, to increase speed of shipping control and employees productivity. Here includes following basic intellectual levels of software-technical and organizational and technological means:

- dispatcher control;
- train safety and automated train control;
- intelligent train and management level of production and business activity of Railways in Azerbaijan.

The main functions of the system are:

- train movement planning;
- conducting production and economic activities of Azerbaijan Railways in an automated mode;
- state of trains and state of technological processes in real time of automated control systems in railway transport;
- operational management in regular and emergency situations;
- operational diagnosis of the state of all technological processes and equipment;
- predicting risks and taking operational measures to eliminate them;
- to work when centralized management "Peregrine Falcon" cabin on emergency recovery is a risk.

### **Development of an intelligent management system**

In the article, an attempt is made to develop a paradigm for the development of the intellectual management system of the transport-logistics complex by determining the breaks within the current socio-economic situation. This study is based on rigorous results presented in several sources. Digitalization of transport and logistics is one of the leading trends in the development of the industry. Today, despite the high level of economic volatility, the digital services market logistics sector is developing rapidly. Thus, in 2020, its volume was 89.4 billion rubles, and by the end of the decade it is expected to increase seven times to 627 billion rubles.

The expected effect of the digitization of the industry by 2030 involves a 20% increase in productivity. In Russia, in 2021, the Ministry of Transport developed a Digital Strategy, and high-level factors were revealed to increase the Transformation of the Transport Industry until 2030.

In relation to the employment effects of Industry 4.0, researchers point to two possible trends: massive technological unemployment in the future, and a shift of the workforce to new jobs with digital tasks and skills. The first approach is referred to in the literature as substitution-digitization, which forecasts the inevitability of massive technological unemployment due to the impact of new digital technologies and the loss of significant employment. The second alternative approach is called task-digitalization [1, pg.249–273].

Digitalization does not eliminate entire occupations but specific tasks within jobs, eliminating some jobs and creating new ones, thus increasing global employment figures. It is necessary to underline that both scenarios coincide with the emergence of losers and winners resulting from digitalization.

An important question for our future, then, is how well-prepared countries' education systems are to meet the significantly changing needs of the workforce as the contents of professions are transformed. Based on predictions, technological spending is expected to rise by 50% between 2015 and 2030, resulting in the creation of 20 to 50 million new jobs worldwide, which would also help to offset the effects of Industry 4.0. These would be new and novel high-end jobs, such as Big Data analysts and machine learning specialists [2, pg.448–456].

#### **Operational connection with other modes of transport.**

Automated control systems in rail transport include smart trains, smart locomotives and smart cars. The local computer network of a smart train connects all devices and systems in a train using a common data bus and controls locomotive and carriage equipment.

Locomotive equipment includes control systems of traction and auxiliary electrical circuits:

- ensuring traffic safety by performing optimal energy consumption functions and automatic control of schedule execution (car driver);
- diagnostic and data recording systems, digital communication systems; longitudinal dynamic force determination systems;
- distributed braking control systems;
- devices that determine the completeness of a train, etc.

It is an indisputable fact that Industry 4.0 is a broad concept and its content is constantly changing. Industry 4.0 encompasses a diversity of technologies, systems, and processes, and it aims to make manufacturing processes more flexible, autonomous and dynamic and to incorporate these into a network. It uses digital and cybernetic resources in production and industrial management environments.

Fundamentally, integrated manufacturing consists of 3D printing technologies, automation, and artificial intelligence. Intelligent manufacturing aggregates the Industrial Internet of Things (IIoT), cyber–physical systems (CPS) and virtual and augmented realities, and creates smarter and more adaptive processes through better use of production resources.

The definitions presented agree that Industry 4.0 will transform the entire corporate value chain, with impacts that will go beyond the organizational framework and have an impact on the global job market as well.

Industry 4.0 cannot be described in one word; however, the literature often uses the terms digitalization or digital transformation [3].

Today, Info Communication Technologies (ICT) is one of the most outstanding examples of this, but Industry 4.0 itself incorporates a wealth of smart technologies that, while not as far-reaching as the examples mentioned earlier, may prove essential to the development of companies.

Industry 4.0 has six main principles: virtualization (virtual replicas of physical tasks as digital data), interoperability (connecting machines via the internet, enabling them to communicate with each other, and with humans), decentralization (autonomous decisionmaking ability of interconnected systems), real-time capacity (simultaneous flow of information between systems, for better and quicker decision making), service orientation (the ability of the system to provide services and functions to stakeholders), and modularity (adaptation of systems to possible changes or errors, addition or replacement of operational modules) [4, pg.925–953].

Industrial revolutions cannot be separated from each other at a specific point in time; rather, the current industrial revolution is a continuation of the previous one. The dataSustainability 2021, 13, 7703 11 of 19 also show that the economic and financial crisis of 2008 had a significantly greater impact on employment than the fourth industrial revolution, which started in 2011. It is easy to see from global labor market data that crises can trigger or exacerbate different economic processes. Employment in the productive sectors (industry, agriculture) is likely to continue its downward trend in the coming decade, despite the conjuncture of the 2010s.

This change further reinforces the question of how labor that became redundant in the productive sectors can be reinstated in the future, and what absorption capacity the service sector has to offer to assimilate industrial labor surpluses. Nevertheless, we can assume that digitalization will have a number of positive effects on human life, especially on quality of life. In spite of all this, the number of jobs replaced by digitalization, and the economical and mental effects of this phenomenon, are also not negligible.

It is clear that the latest technological developments in digitalization solutions—cyber–physical systems, Big Data, Internet of Things, cloud computing and robotics—are having a negative impact on people's employment opportunities [5, pg.27–40].

The manufacturing sector, which is currently the industry most effected by digitalization, suffered a loss of 1.7 million jobs worldwide between 2010 and 2016. If this trend continues until 2030, as much as 20 million jobs could disappear, thanks to digitalization, which means an 8.5% loss of jobs in the sector. The effects of lost jobs will vary from country to country, but it will affect less educated workers and the poorer countries the most, meaning a twofold greater loss compared to richer countries.

#### **References**

1. Sánchez, A. L. Digitalization, Robotization, Work and Life: Cartographies, Debates and Practices. Cuad. Relac. Laborales 2019, 37 (2), Pages 249–273;
2. Forum, W. E. The Future of Jobs Report 2018; Centre for the New Economy and Society: Geneva, 2018;
3. Cerezo-Narváez, A.; Otero-Mateo, M.; Rodríguez-Pecci, F.; Pastor-Fernández, A. Digital Transformation of Demands in Industry 4.0: The Case of Offshore Platforms. Dyna 2018, 93 (4), Pages 448–456;
4. Manavalan, E.; Jayakrishna, K. An Overview of Internet of Things (IoT) Embedded Sustainable Supply Chain for Industry 4.0 Requirements. Calculation. Ind. Eng. 2019, 127, Pages 925–953;
5. Liu, Y.; Peng, Y.; Wang, B.; Yao, S.; Liu, Z. Review of Cyber-Physical Systems. IEEE/CAA J. Winter. 2017, 4 (1), Pages 27–40.

<https://doi.org/10.31891/2219-9365-2024-79-24>

УДК 004

КОРНАГА Ярослав  
Національний Технічний Університет України “Київський політехнічний інститут імені Ігоря Сікорського”  
[y.kornaga@kpi.ua](mailto:y.kornaga@kpi.ua)

ГУБАРЕВ Олександр  
Національний Технічний Університет України “Київський політехнічний інститут імені Ігоря Сікорського”  
<https://orcid.org/0000-0002-0924-4103>  
[gubarev.alexandr@gmail.com](mailto:gubarev.alexandr@gmail.com)

## АЛГОРИТМ ПЕРЕХОДУ ВІД МОНОЛІТНОЇ ДО МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ ІЗ ЗАСТОСУВАННЯМ МЕТОДІВ КЛАСТЕРИЗАЦІЇ ГРАФІВ

У даній статті представлений підхід до трансформації монолітної архітектури в мікросервісну з використанням *bottom-up* стратегії декомпозиції. В даному підході використовується статичний аналіз коду для представлення монолітної системи у вигляді зв'язного графа. На другому етапі, використовуючи алгоритми кластеризації, відбувається кластеризація графа. На третьому — оптимізація отриманих кластерів.

Ключові слова: мікросервіси, декомпозиція системи, кластеризація графу, комбінаторна оптимізація.

KORNAGA Yaroslav, GUBAREV Oleksandr  
National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”

## ALGORITHM FOR TRANSITION FROM MONOLITHIC TO MICROSERVICE ARCHITECTURE USING GRAPH CLUSTERING METHODS

The transition from a monolithic architecture to a microservice architecture using graph clustering methods is an effective approach that allows you to divide the system into logically connected components.

Static analysis is the first step in making the transition from a monolithic to a microservice architecture. It allows you to highlight components (nodes) and establish relationships between them, as well as the weight of these relationships, depending on the type of components it connects.

After constructing a connected graph based on static analysis, the next important step is graph clustering. Graph clustering provides an automated approach to partitioning the monolith, which allows for the detection of natural boundaries between components. The use of clustering algorithms, such as Louvain, Spectral Clustering or Girvan-Newman, simplifies and speeds up the process of module identification.

The next important stage in the process of transforming a monolithic architecture into a microservice is the optimization of the resulting clusters. After completing the practical task of clustering a graph with 5 vertices, 2 clusters were obtained, the first included 2 nodes, and the second - 3. But after performing cluster optimization, a different result was obtained - one of the graph vertices was moved from the second to the first cluster. This happened due to the reduction of intercluster dependencies, which ensured the efficient distribution of components and ultimately increased the performance and scalability of the system.

This article presents an approach to the transformation of a monolithic architecture into a microservice one using a *bottom-up* decomposition strategy. This approach uses static code analysis to represent a monolithic system in the form of a connected graph. At the second stage, the graph is clustered using clustering algorithms. The third is optimization of the obtained clusters.

Key words: microservices, system decomposition, graph clustering, combinatorial optimization.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Перехід від монолітної архітектури до мікросервісної є актуальним напрямком розвитку сучасних програмних систем, що дозволяє підвищити гнучкість, масштабованість і швидкість розробки та впровадження програмного забезпечення. У цьому процесі великі монолітні додатки розбиваються на невеликі, автономні мікросервіси, кожен з яких відповідає за окрему функціональність або бізнес-процес [1].

Основними мотивуючими факторами для переходу до мікросервісної архітектури є:

1. Гнучкість і швидкість вдосконалення: Мікросервіси дозволяють здійснювати окреме вдосконалення і впровадження нових функцій без необхідності розгортання всього додатку.
2. Масштабованість: Кожен мікросервіс може бути масштабований незалежно від інших, що дозволяє оптимізувати використання ресурсів і забезпечити високу доступність системи.
3. Ізоляція помилок: Проблеми в одному мікросервісі зазвичай не впливають на інші, що полегшує відладку і підтримку системи.
4. Технологічна гетерогенність: Різні мікросервіси можуть використовувати різні технології і стеки, що дозволяє використовувати найкращі інструменти для кожного конкретного завдання.

Перехід від монолітної архітектури до мікросервісної, хоча і має багато переваг, зазвичай супроводжується рядом складнощів і викликів, які варто урахувати:

1. Системна складність: Управління багатьма невеликими сервісами потребує ефективного моніторингу, координації і управління версіями.
2. Комунікація між сервісами: Необхідність ефективної організації міжсервісних взаємодій і забезпечення консистентності даних.
3. Тестування і відладка: Складність тестування, особливо інтеграційного, через розподілену природу системи [2].

Існують різні стратегії рефакторингу монолітної системи в мікросервісну архітектуру. Їх можна класифікувати за трьома категоріями, так званими стратегіями декомпозиції: top-down, bottom-up та hybrid підходи.

Однією з головних проблем, з якою доводиться стикатися при використанні будь-якої стратегії декомпозиції є ідентифікація мікросервісів із високою цілісністю та слабким зв'язком.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою даної статті є представлення алгоритму декомпозиції моноліта у мікросервіси із високою цілісністю і слабким зв'язком з використанням методів кластеризації графів і методів для комбінаторної оптимізації утворених кластерів.

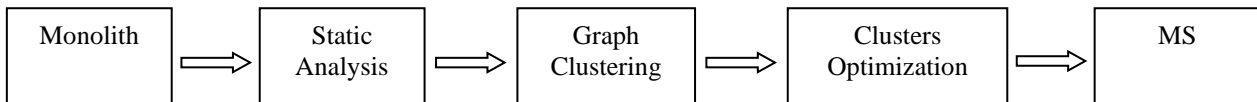


Рис.1.1 Алгоритм переходу від моноліту до мікросервісів

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

#### 1. Стратегії рефакторингу монолітної системи

*Top-Down підхід (Рис.1.2):* орієнтується на початкове визначення високорівневих бізнес-функцій і вимог до системи, які потім розбиваються на менші компоненти - мікросервіси. Цей підхід спрощує процес розробки, покращує масштабованість та забезпечує гнучкість системи. Основними етапами цього підходу є: аналіз бізнес-функцій і вимог (починаючи з високорівневого аналізу, визначаються основні бізнес-функції та вимоги до системи. Цей етап включає розуміння бізнес-процесів, що виконуються системою, та ідентифікацію ключових бізнес-компонентів), розбиття на функціональні області(бізнес-функції і вимоги поділяються на окремі функціональні області, які можуть стати кандидатами для мікросервісної реалізації. Це може включати ідентифікацію окремих процесів або бізнес-правил, які можуть бути реалізовані як окремі сервіси), визначення мікросервісних границь(на основі функціональних областей визначаються мікросервісні границі - інтерфейси та контракти, які визначають, як сервіси взаємодіють між собою та з іншими системами. Цей етап також включає визначення інтерфейсів API, які дозволяють комунікацію між мікросервісами), реалізація мікросервісів(на основі визначених мікросервісних границь розробляються і реалізуються окремі мікросервіси. Кожен мікросервіс відповідає за виконання конкретної функціональної області і має чітко визначені внутрішні правила і логіку), інтеграція і тестування (Після реалізації мікросервісів їх необхідно інтегрувати та виконати комплексне тестування, включаючи інтеграційне тестування для визначення коректності взаємодії між сервісами), моніторинг і підтримка(після впровадження мікросервісів в експлуатацію, необхідно налаштувати моніторинг та забезпечити підтримку, щоб впевнитися в їхній надійності та ефективності).

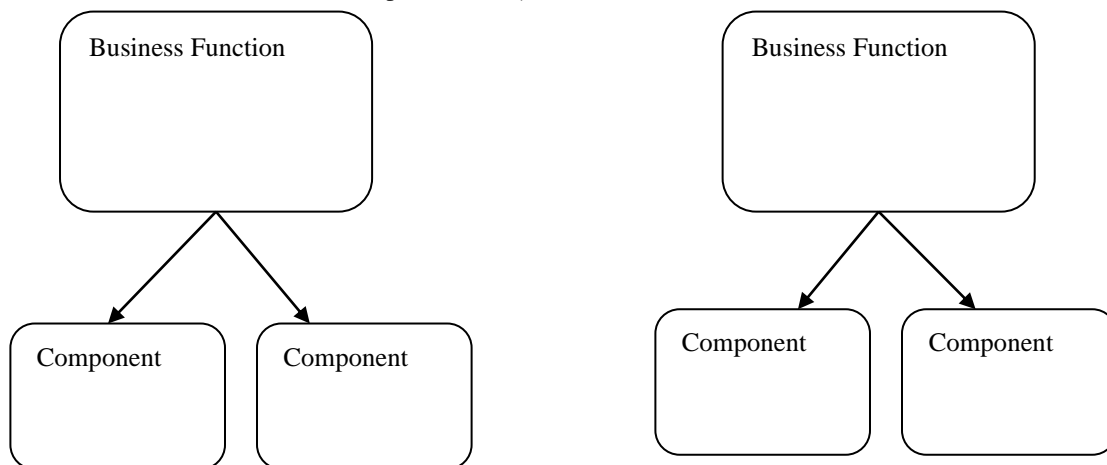


Рис.1.2 Top-Down підхід



До переваг цього підходу можна віднести:

- Простота управління: Цей підхід спрощує процес розробки, оскільки він дозволяє концентруватися на високорівневих бізнес-вимогах і функціях.
- Гнучкість і масштабованість: Завдяки чіткому визначенню мікросервісних границь, система стає більш гнучкою і масштабованою, що дозволяє ефективно відповідати на зміни в бізнес-потребах.

Недоліками є:

- Складність визначення мікросервісних границь: Не завжди просто точно визначити, які функції повинні бути реалізовані як окремі мікросервіси, особливо в складних системах з великою кількістю функціональності.
- Комплексність інтеграції: Інтеграція різних мікросервісів може виявитися складною задачею, особливо при потребі забезпечення консистентності даних і ефективної комунікації.

Top-Down підхід до розробки мікросервісної архітектури є корисним інструментом для організації та створення складних систем, проте його успішність залежить від глибини і точності аналізу бізнес-вимог і функцій.

*Bottom-Up підхід (Рис.1.3):* розглядається як протилежність Top-Down підходу. Він полягає в початковому розробленні невеликих компонентів або сервісів, які пізніше об'єднуються для створення більш складної системи. Основна ідея полягає у тому, щоб перш ніж визначити високорівневі бізнес-функції і вимоги, розробляти індивідуальні сервіси і поступово їх інтегрувати у єдину систему.

Основними етапами цього підходу є: розробка незалежних мікросервісів: початково розробляються невеликі, самодостатні сервіси, які реалізують обмежену функціональність або певний бізнес-процес. Кожен сервіс має власний інтерфейс, базується на внутрішніх правилах і має власну базу даних (якщо необхідно); інтеграція і комунікація: після створення незалежних сервісів вони інтегруються один з одним із використанням механізмів комунікації, таких як API, шини повідомлень або інші технології. Цей етап включає встановлення зв'язків із сервісами, які можуть бути необхідними для спільного вирішення бізнес-завдань; розширення та масштабування: Після успішної інтеграції і випробувань сервісів можна розглядати їх розширення і масштабування. Цей етап включає оптимізацію та реорганізацію коду, а також впровадження моніторингу та управління версіями; ітерації і підтримка: Bottom-Up підхід сприяє ітеративному покращенню і розвитку системи, оскільки нові функції можуть додаватися шляхом створення нових мікросервісів або розширенням існуючих. Підтримка системи включає надання підтримки, виправлення помилок та впровадження оновлень.

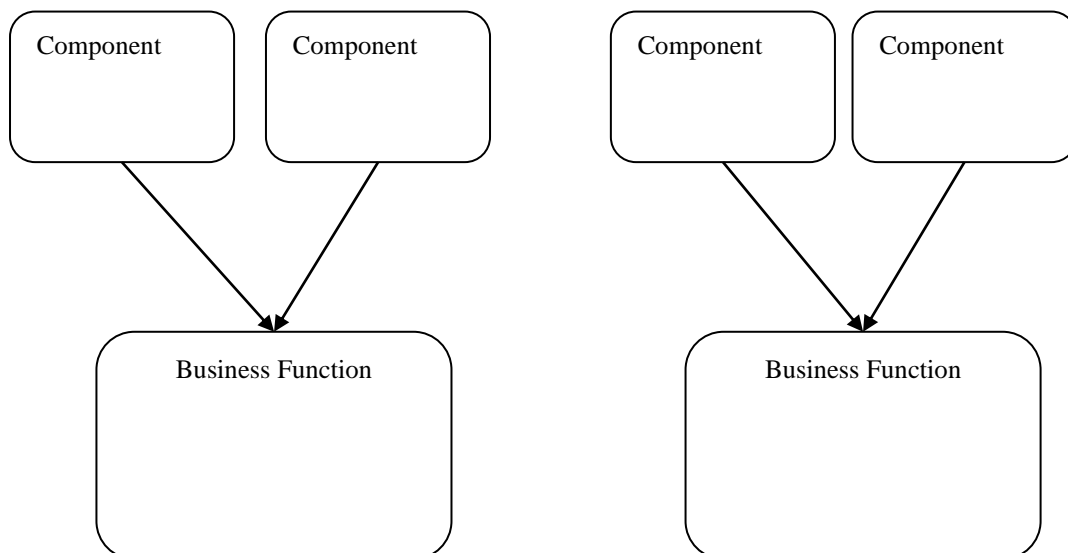


Рис.1.3 Bottom-Up

До переваг цього підходу можна віднести:

- Гнучкість і швидкість розробки: Розробка мікросервісів може відбуватися незалежно один від одного, що сприяє швидкому впровадженню нових функцій та експериментам.
- Масштабованість: Кожен сервіс може масштабуватися незалежно від інших, що дозволяє легше управляти завантаженням і ресурсами.

Недоліками є:

- Складність інтеграції: Інтеграція багатьох незалежних сервісів може бути складною

задачею, особливо при потребі забезпечення консистентності даних та ефективної комунікації.

- Управління залежностями: Важливо правильно керувати залежностями між мікросервісами, щоб уникнути зайвої складності і зберегти стабільність системи.

Bottom-Up підхід є корисним для організацій, які шукають швидкі результати і можливість експериментувати з новими функціями, однак він вимагає уважного управління залежностями і інтеграцією, щоб досягти стабільної і ефективної мікросервісної архітектури.

*Hybrid підхід(Рис 1.4):* поєднує в собі елементи як Top-Down, так і Bottom-Up підходів. Його основна ідея полягає в тому, щоб використовувати переваги обох підходів для досягнення оптимальних результатів при переході від монолітної до мікросервісної архітектури. Основними етапами цього підходу є: аналіз бізнес-функцій і вимог: Починаючи з високорівневого аналізу бізнес-функцій і вимог до системи, визначаються ключові функціональні області і компоненти, які потрібно реалізувати; розробка початкових мікросервісів: На цьому етапі розробляються початкові мікросервіси, які вже відповідають за окремі функціональні області чи бізнес-процеси. Ці мікросервіси можуть створюватися в рамках Bottom-Up підходу; визначення мікросервісних границь: На основі аналізу бізнес-функцій і початкових мікросервісів визначаються мікросервісні границі і інтерфейси між ними. Цей етап також включає визначення механізмів комунікації між сервісами; ітеративна розробка та інтеграція: Після визначення мікросервісних границь розробка продовжується ітеративно, додаванням нових мікросервісів та їх інтеграцією з існуючими сервісами; тестування і впровадження: Після розробки і інтеграції мікросервісів проводяться комплексне тестування для визначення коректності їх роботи в складі системи. Після успішного тестування мікросервіси впроваджуються в експлуатацію; моніторинг і оптимізація: Після впровадження необхідно налаштувати моніторинг сервісів та проводити їх оптимізацію для забезпечення високої доступності та продуктивності.

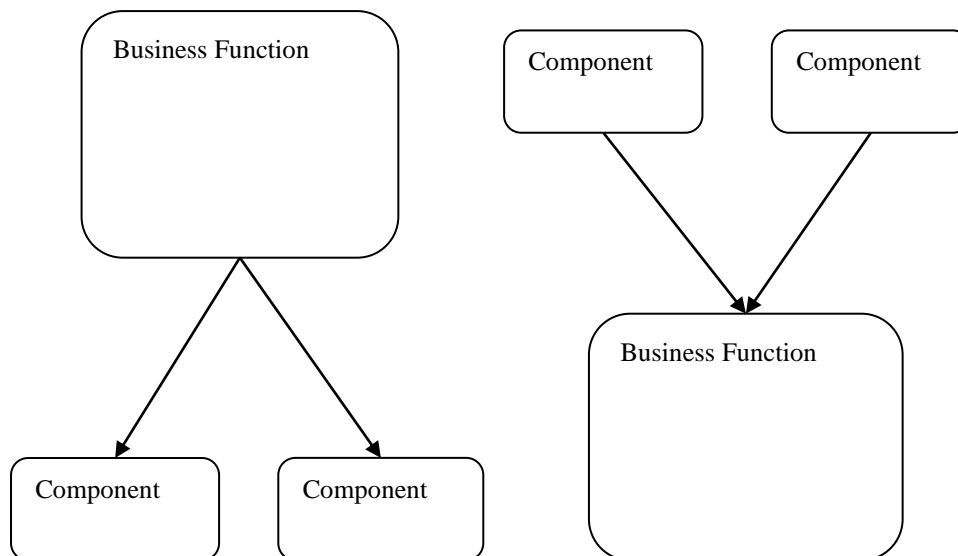


Рис 1.4 Hybrid підхід

До переваг цього підходу можна віднести:

- Гнучкість і адаптивність: Гібридний підхід дозволяє використовувати найкращі аспекти як Top-Down, так і Bottom-Up підходів, залежно від конкретних потреб проекту.
- Ефективність розробки: Розробка починається з швидкої ітерації Bottom-Up, що дозволяє швидко отримувати результати, а потім переходить до більш структурованого визначення мікросервісних границь у вигляді Top-Down підходу.

Недоліками є:

- Складність керування залежностями: Під час інтеграції різних елементів, що були розроблені різними методами, можуть виникати проблеми з керуванням залежностями і сумісністю.
- Потреба в збалансованості: Для успішної реалізації гібридного підходу потрібно правильно збалансувати використання Top-Down і Bottom-Up методів, щоб уникнути перекриття чи недоліків.

Гібридний підхід до розробки мікросервісної архітектури є потужним інструментом для організацій, які шукають компроміс між швидкістю розробки і структурованістю дизайну. Він дозволяє ефективно використовувати переваги обох підходів для досягнення успішного переходу до мікросервісної архітектури.

2. Автоматизований підхід переходу на мікросервісну архітектуру на основі кластеризації графів

Перехід від монолітної до мікросервісної архітектури можна автоматизувати за допомогою підходу, який використовує кластеризацію графів і комбінаторну оптимізацію. Цей підхід зосереджений на аналізі залежностей у монолітній системі та визначенні оптимальних меж для мікросервісів.

Розглянемо покроковий перехід від моноліту до мікросервісів з використанням кластеризації графів.

### 2.1 Статичний аналіз коду (Static Code Analysis)

Статичний аналіз - аналіз вихідного коду монолітного додатка для виділення класів, методів і їх взаємозв'язків. Цей аналіз може включати визначення залежностей між класами, методами, а також взаємодії з базою даних. В ході аналізу можуть бути використані анотації із фреймворку, такого як SpringFramework `@Controller`, `@Service`, `@Repository`, `@Entity`. Ці анотації вказують на технічні ролі класів (наприклад, контролери, сервіси, репозиторії, сутності), що є цільовими для виокремлення в окремі мікросервіси.

Далі анотації можуть бути використані для визначення границь між різними мікросервісами. Наприклад, класи з анотацією `@Controller` можуть стати окремими сервісами, які надають API для зовнішніх клієнтів. Класи з анотацією `@Service` можуть представляти бізнес-логіку, яку можна розділити на окремі мікросервіси з різними функціональними обов'язками. Класи, які мають анотації `@Entity` або відповідають доменним об'єктам, таким як моделі бази даних, можуть бути відокремлені в окремі мікросервіси, які управляють певними частинами доменної моделі.

### 2.2 Побудова графового представлення (GraphRepresentation)

Побудова графового представлення - створення графової структури, де вузлами будуть представлені класи і методи програми, а ребра відобразять взаємозв'язки між ними, такі як виклики методів, використання різних класів тощо. Це дає уявлення про структурні зв'язки в монолітному додатку [3,4].

Вузол графа може бути створений для: кожного класу рівня логіки та репозиторію та для кожного класу сутності. Зв'язки між класами мають бути поміщені на граф у вигляді спрямованих дуг. Також слід визначити типи дуг і їхні ваги. Ваги визначають релевантність зв'язків між вузлами: чим вища вага дуги, тим більша ймовірність, що сутності в її кінцевих точках слід помістити в той самий мікросервіс [5,6]. Після аналізу ваги призначаються відповідно до типу зв'язку.

`repository` → `persistence_entity` (w=5),  
`service` → `repository` (w=4)  
`service` → `service` (w=3)  
`service` → `entity` (w=2)  
`entity` → `entity` (w=1)

Зв'язок `repository` → `persistence_entity` має найвищу вагу, оскільки `persistence_entity` є моделями бази даних, для зв'язків `service` → `repository` та `service` → `service` встановлена нижча вага і вага зв'язків `service` → `entity` і `entity` → `entity` нижча, щоб сутності домену було розподілено між різними мікросервісами.

### 2.3 Кластеризація графа

Метою кластеризації є знаходження груп вузлів (класів, методів), які мають високу внутрішню схожість (висока когезія) і низьку зовнішню залежність (низька зв'язаність). Ця задача математично формулюється через оптимізаційні критерії або евристики, залежно від обраного алгоритму [7].

#### 2.3.1 Алгоритми кластеризації графів

*Спектральна кластеризація* використовує спектральні властивості (власні значення і вектори) матриці Лапласіана графа для розбиття графа [10].

Алгоритм:

1. Побудувати матрицю Лапласіана  $L=D-A$ , де  $D$  — діагональна матриця ступенів,  $A$  — матриця суміжності.
2. Обчислити  $k$  найменших власних векторів Лапласіана, де  $k$  — кількість кластерів.
3. Використовувати  $k$  власних векторів для кластеризації вершин за допомогою алгоритму  $k$ -середніх ( $k$ -means).

Формула:  $L=D-A$

*Метод Лувена* максимізує модулярність графа, що вимірює густину ребер всередині кластерів порівняно з очікуваною густиною у випадковому графі [8].

Алгоритм:

1. Початково кожен вузол розглядається як окремий кластер.
2. Повторювано переміщуються вузли між кластерами для збільшення модулярності.

3. Злиття вузлів у супер-вузли та повторення процесу на новому рівні абстракції до досягнення оптимальної модулярності.

Формула Модулярності:

$$Q = \frac{1}{2 \cdot m} \sum_{ij} \left[ A_{ij} - \frac{k_i \cdot k_j}{2 \cdot m} \right] \cdot \delta(C_i, C_j) \quad (2.1)$$

де:

$A_{ij}$  — елемент матриці суміжності,

$k_i, k_j$  — ступені вузлів  $i$  та  $j$ ,

$m$  — загальна кількість ребер,

$\delta(C_i, C_j)$  — функція Кронекера.

Алгоритм Гірвана-Ньюмана знаходить спільноти в графі шляхом послідовного видалення ребер з високою центральною між кластерами [9].

Алгоритм:

1. Обчислити Центральність Бетвіна для всіх ребер.
2. Видалити ребро з найбільшою центральною.
3. Повторювати кроки 1 і 2 до досягнення бажаного числа кластерів.

Формула Центральності Бетвіна:

$$C_b(e) = \sum_{\text{allpairs}(s,t)} \frac{\sigma_{st}(e)}{\sigma_{st}}, \quad (2.2)$$

де:

$\sigma_{st}$  — загальна кількість найкоротших шляхів між  $s$  і  $t$ ,

$\sigma_{st}(e)$  — кількість шляхів, що проходять через ребро  $e$

Метод поширення міток призначає мітки вузлам графа на основі міток їхніх сусідів, припускаючи, що сусідні вузли схильні належати до одного і того ж кластера [11].

Алгоритм:

1. Початково кожному вузлу присвоюється унікальна мітка.
2. Повторювано оновлюються мітки вузлів відповідно до найпоширенішої мітки серед сусідів.
3. Процес повторюється до збіжності (тобто мітки перестають змінюватися).

Кластеризація графа допомагає автоматично розбивати монолітні додатки на логічні групи компонентів, які можуть бути виділені в окремі мікросервіси. Використання алгоритмів, таких як Louvain, дозволяє знайти оптимальні кластери, зменшуючи міжкластерні взаємодії і збільшуючи внутрішньокластерну зв'язаність.

### 3. Оптимізація кластерів

Мета оптимізації - мінімізувати міжкластерні взаємодії та покращити внутрішньокластерну зв'язаність, що дозволяє досягти більшої автономності та ефективності кожного мікросервісу.

Приклад оптимізації кластеризації

1. Створення початкових кластерів

Розглянемо граф з вузлами та ребрами:

- Вузли: A, B, C, D, E
- Залежності: A -> B, A -> C, B -> D, C -> D, C -> E

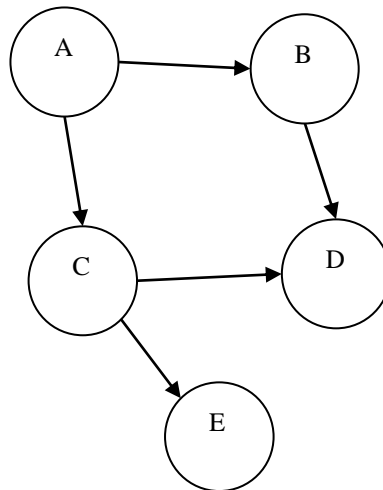


Рис.3.1. Граф з вузлами та ребрами

Для простоти припустимо, що всі залежності мають однакову вагу.

```
import networkx as nx
import community as community_louvain

# Створення графа
G = nx.DiGraph()
edges = [('A', 'B'), ('A', 'C'), ('B', 'D'), ('C', 'D'), ('C', 'E')]
G.add_edges_from(edges)

# Виконання кластеризації з використанням алгоритму Louvain
partition = community_louvain.best_partition(G.to_undirected())

# Виведення початкових кластерів
for node, cluster in partition.items():
    print(f"Node {node} is in cluster {cluster}")
```

Результат роботи програми:

```
Node A is in cluster 0
Node B is in cluster 0
Node C is in cluster 1
Node D is in cluster 1
Node E is in cluster 1
```

### 3.1. Формулювання оптимізаційної задачі

Мета - мінімізувати кількість міжкластерних залежностей.

Змінні:  $x_{ij}$  – двійкова змінна, яка приймає значення 1, якщо вузол  $i$  призначено кластеру  $j$ , і 0 в іншому випадку.

Цільова функція:

$$\text{Minimize } \sum_{(i,k) \in E} \sum_{j \neq l} x_{ij} \cdot x_{kl}, \quad (3.1)$$

Обмеження:

Кожен вузол призначається тільки одному кластеру:

$$\sum_j x_{ij} = 1 \forall i, \quad (3.2)$$

Кількість кластерів обмежена:

$$\sum_i x_{ij} \leq \max \quad \forall j, \quad (3.3)$$

### 3.2. Розв'язання задачі з використанням ІЛР

Використовуємо бібліотеку PuLP для розв'язання ІЛР:

```
!pip install pulp
from pulp import LpProblem, LpMinimize, LpVariable, lpSum, LpBinary, value

# Вузли і ребра
nodes = ['A', 'B', 'C', 'D', 'E']
edges = [('A', 'B'), ('A', 'C'), ('B', 'D'), ('C', 'D'), ('C', 'E')]

# Кластери
clusters = [0, 1] # Припустимо, що ми хочемо 2 кластери

# Створення ІЛР задачі
problem = LpProblem("ClusterOptimization", LpMinimize)

# Змінні
x = LpVariable.dicts("assignment", [(i, j) for i in nodes for j in clusters], cat=LpBinary)

# Допоміжні змінні для мінімізації міжкластерних залежностей
y = LpVariable.dicts("inter_cluster", [(i, k) for i, k in edges], cat=LpBinary)

# Обмеження: кожен вузол в одному кластері
for i in nodes:
    problem += lpSum(x[(i, j)] for j in clusters) == 1

# Обмеження: найбільше max_nodes_per_cluster вузлів у кластері (опційно)
max_nodes_per_cluster = 3
for j in clusters:
    problem += lpSum(x[(i, j)] for i in nodes) <= max_nodes_per_cluster

# Обмеження для допоміжних змінних
for (i, k) in edges:
    for c1 in clusters:
        for c2 in clusters:
            if c1 != c2:
                problem += y[(i, k)] >= x[(i, c1)] + x[(k, c2)] - 1

# Цільова функція
problem += lpSum(y[(i, k)] for (i, k) in edges)

# Розв'язання задачі
problem.solve()

# Виведення результатів
cluster_assignment = {i: j for i in nodes for j in clusters if x[(i, j)].varValue == 1}
print("Cluster assignment:")
for node, cluster in cluster_assignment.items():
    print(f"Node {node} is in cluster {cluster}")

print("Objective value:", value(problem.objective))

Результат роботи програми:
Cluster assignment:
Node A is in cluster 0
Node B is in cluster 1
```

Node C is in cluster 0  
Node D is in cluster 1  
Node E is in cluster 0  
Objective value: 2.0

### ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Перехід від монолітної архітектури до мікросервісної з використанням методів кластеризації графів є ефективним підходом, який дозволяє розділити систему на логічно зв'язані компоненти.

Статичний аналіз є першим етапом здійснення переходу від монолітної до мікросервісної архітектури. Він дозволяє виділити компоненти (вузли) та встановити зв'язки між ними, а також вагу цих зв'язків, в залежності від типу компонентів, які він з'єднує.

Після побудови зв'язного графа на основі статичного аналізу, наступним важливим кроком є кластеризація графів. Кластеризація графів забезпечує автоматизований підхід до розбиття моноліту, що дозволяє виявити природні межі між компонентами. Використання алгоритмів кластеризації, таких як Louvain, Spectral Clustering або Girvan-Newman, спрощує і прискорює процес ідентифікації модулів.

Наступним важливим етапом у процесі перетворення монолітної архітектури в мікросервісну є оптимізація отриманих кластерів. Виконавши практичну задачу кластеризації графа з 5 вершинами, було отримано 2 кластери, перший включав 2 вузли, а другий – 3. Але після виконання кластерної оптимізації було отримано інший результат – одна з вершин графу була переміщена з другого до першого кластеру. Відбулося це завдяки зменшенню міжкластерних залежностей, що забезпечило ефективний розподіл компонентів і в кінцевому результаті підвищило продуктивність і масштабованість системи.

До переваг підходу можна віднести:

1. Аналіз графу залежностей дозволяє виявити приховані зв'язки між компонентами, які можуть бути неочевидними в традиційному аналізі. Це допомагає більш точно визначити межі мікросервісів та забезпечити їх ефективну взаємодію.
2. Кластеризація графів допомагає розбити систему на менші, незалежні модулі з високою внутрішньою зв'язаністю і мінімальними зовнішніми залежностями.
3. Автоматизоване визначення кластерів зменшує складність ручного аналізу і розбиття моноліту, забезпечуючи більш об'єктивний і точний розподіл.

Виклики та обмеження:

1. Якість даних:
  - Успішність кластеризації графів залежить від якості даних про залежності між компонентами. Неповні або неточні дані можуть призвести до некоректного розбиття.
  - Необхідно забезпечити точний збір і аналіз метрик залежностей.
2. Складність налаштування:
  - Вибір відповідного алгоритму кластеризації і його параметрів може бути складним завданням, яке вимагає спеціальних знань і досвіду.
  - Потрібно експериментувати з різними підходами і параметрами, щоб досягти оптимального результату.
3. Інтеграція з існуючою системою:
  - Процес переходу від моноліту до мікросервісів може вимагати значних змін в існуючій інфраструктурі і організаційних процесах.
  - Важливо забезпечити безперервність роботи системи під час переходу та мінімізувати вплив на користувачів.

Розбиття монолітної архітектури на мікросервіси за допомогою кластеризації графів є потужним і ефективним підходом, який дозволяє автоматизувати процес ідентифікації модулів і зменшити складність ручного аналізу. Це допомагає підвищити модульність системи, знизити ризик помилок і забезпечити більш ефективний перехід до мікросервісної архітектури.

Незважаючи на виклики, пов'язані зі складністю аналізу та розподілу, правильна стратегія та інструменти дозволяють досягти значних переваг і успішно впровадити мікросервісну архітектуру.

### Література

1. Newman S. Building Microservices: Designing Fine-Grained Systems. O'Reilly Media, Inc.; 2015.
2. S. Hassan, R. Bahsoon, and R. Kazman, Microservice transition and its granularity problem: A systematic mapping study, *Software: Practice and Experience*, vol. 50, no. 9, pp. 1651–1681, 2020.
3. Trabelsi, M. Abdellatif, A. Abubaker, N. Moha, S. Mosser, S. Ebrahimi-Kahou, and Y.-G. Gue'he'neuc, From legacy to microservices: A type-based approach for microservices identification using machine learning and semantic analysis, *Journal of Software: Evolution and Process*, 2022.
4. M. Abdellatif, R. Tighilt, N. Moha, H. Mili, G. El Boussaidi, J. Privat, Y.-G. Gue'he'neuc, A type-sensitive service identification approach for legacy-to-soa migration, *International Conference on Service-Oriented Computing*, pp. 476–491, Springer, 2020.

5. Selmadji, A.-D. Seriai, H. L. Bouziane, R. O. Mahamane, P. Zaragoza, and C. Dony, From monolithic architecture style to microservice one based on a semi-automatic approach, ICISA. IEEE, 2020, pp. 157–168
6. Levcovitz A, Terra R, Valente MT. Towards a technique for extracting microservices from monolithic enterprise systems, 2016.
7. M. E. Newman, Modularity and community structure in networks, Proceedings of the national academy of sciences, vol. 103, no. 23, pp. 8577–8582, 2006.
8. Scott Emmons, Stephen Kobourov, Mike Gallant, Katy Börner Analysis of Network Clustering Algorithms and Cluster Quality Metrics at Scale URL: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0159161>
9. Lj. Despalatović, T. Vojković, D. Vukičević Community structure in networks: improving the Girvan-Newman algorithm
10. Scott White, Padhraic Smyth A Spectral Clustering Approach To Finding Communities in Graphs
11. Kamal Berahmanda, Sogol Haghanib, Mehrdad Rostamic, Yuefeng Lia A new attributed graph clustering by using label propagation in complex networks, Journal of King Saud University- Computer and Information Sciences, vol 34, May 2022, Pages 1869-1883



<https://doi.org/10.31891/2219-9365-2024-79-25>

УДК: 004.8

БЕСЕДОВСЬКИЙ Олександр

Харківський національний економічний університет імені Семена Кузнеця

<https://orcid.org/0000-0002-9161-4061>

[oleksii.besedovskiy@hneu.net](mailto:oleksii.besedovskiy@hneu.net)

## ТЕОРЕТИЧНІ ЗАСАДИ РОЗРОБКИ ШІ ЧАТ-БОТ

В статті розглянуто сучасні тенденції розвитку штучного інтелекту (ШІ) та його застосування в різних галузях. Метою дослідження є систематизація та узагальнення наявних теоретичних аспектів розробки ШІ чат-ботів та обґрунтування процесу розробки ШІ чат-боту. Проаналізовано три основних типів ШІ: вузькоспеціалізованого, загального та надінтелекту, їх можливостям і потенційним напрямкам розвитку. У дослідженні представлено детальний огляд застосування ШІ в контексті чат-ботів, які швидко стають незамінними інструментами для автоматизації бізнес-процесів, підвищення ефективності взаємодії з клієнтами та покращення обслуговування. Обґрунтовано, що ШІ чат-боти, як комп'ютерні програми, що імітують та обробляють людські розмови, відіграють важливу роль у підвищенні ефективності взаємодії між користувачами та цифровими системами. Вони можуть бути як простими програмами для виконання базових запитів, так і складними цифровими асистентами, які використовують штучний інтелект для навчання та персоналізації взаємодії з користувачами. Наведено статистичні дані, що підтверджують актуальність впровадження чат-ботів у бізнес, зокрема в електронній комерції. Згідно з дослідженням, до 2024 року понад 91% онлайн-магазинів у США вже впровадили інтелектуальні системи, що дозволяє суттєво знизити витрати на людські ресурси та оптимізувати процеси обслуговування клієнтів. Виявлено, що інтеграція чат-ботів значно підвищує продуктивність підприємств, забезпечуючи цілодобову доступність сервісу та персоналізовану взаємодію з користувачами. Дослідження також дозволило виявити ключові етапи розробки ШІ чат-бота, включаючи вибір мови програмування, фреймворків та платформи, організацію бази даних і налаштування моделей машинного навчання для обробки запитів користувачів. Обґрунтовано важливість використання хмарних обчислень та технологій обробки природної мови (NLP) для підвищення гнучкості та ефективності системи. Наведено порівняльний аналіз чат-ботів з іншими суміжними технологіями, що дозволяє оцінити їхню унікальність і переваги. Обґрунтовано, що впровадження ШІ чат-ботів є критично важливим кроком для розвитку сучасних підприємств, оскільки вони допомагають оптимізувати бізнес-процеси, покращити комунікацію з клієнтами та зменшити витрати, забезпечуючи економічну ефективність.

Ключові слова: штучний інтелект, чат-боти, автоматизація бізнес-процесів, обробка природної мови, технології

BESEDOVSKYI Oleksii

Simon Kuznets Kharkiv National University of Economics

## THEORETICAL ASPECTS OF AI CHATBOT DEVELOPMENT

The article discusses modern trends in the development of artificial intelligence (AI) and its application across various industries. The aim of the study is to systematize and summarize the existing theoretical aspects of AI chatbot development and to justify the process of AI chatbot creation. The paper analyzes three main types of AI: narrow, general, and superintelligence, their capabilities, and potential areas of development. The research provides a detailed overview of AI usage in the context of chatbots, which are rapidly becoming indispensable tools for automating business processes, enhancing customer interaction, and improving service quality. It is justified that AI chatbots, as computer programs that simulate and process human conversations, play an essential role in improving the efficiency of interactions between users and digital systems. These chatbots can range from simple programs handling basic requests to complex digital assistants using artificial intelligence for learning and personalizing user interactions. The article presents statistical data supporting the relevance of chatbot implementation in business, particularly in e-commerce. According to the study, by 2024, more than 91% of online stores in the U.S. will have adopted intelligent systems, significantly reducing labor costs and optimizing customer service processes. It was found that chatbot integration significantly boosts business productivity, ensuring 24/7 service availability and personalized user interaction. The research also identifies the key stages of AI chatbot development, including the selection of programming language, frameworks, and platform, the organization of a database, and the configuration of machine learning models to process user queries. The importance of using cloud computing and natural language processing (NLP) technologies to enhance system flexibility and efficiency is substantiated. A comparative analysis of chatbots with other related technologies is provided, allowing an assessment of their uniqueness and advantages. It is argued that the implementation of AI chatbots is a critical step for the development of modern enterprises, as they help optimize business processes, improve customer communication, and reduce costs, ensuring economic efficiency.

Keywords: artificial intelligence, chatbots, business process automation, natural language processing, technologies.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Стрімкий розвиток технологій автоматизації та цифровізації бізнес-процесів, а також зростаючі вимоги до підвищення ефективності взаємодії з користувачами зумовлюють необхідність дослідження теоретичних засад розробки штучного інтелекту (ШІ) чат-ботів. Сучасні чат-боти, що використовують штучний інтелект, стають невід'ємною частиною багатьох сфер діяльності, включаючи підтримку клієнтів, управління персоналом, маркетинг та електронну комерцію. Завдяки можливостям машинного навчання та обробки природної мови, чат-боти зі штучним інтелектом здатні ефективно імітувати людське спілкування, забезпечуючи своєчасні та релевантні відповіді на запити користувачів.

Теоретична база розробки таких систем охоплює питання алгоритмів навчання, розпізнавання та аналізу текстової інформації, а також моделей взаємодії людина-комп'ютер. Це вимагає поглибленого вивчення питань адаптації чат-ботів до специфіки різних галузей, культурних відмінностей та мовних особливостей. Крім того, з огляду на зростаючу роль ШІ у вирішенні бізнес-завдань, критичним є розуміння етичних та правових аспектів використання таких технологій, зокрема щодо конфіденційності даних та безпеки інформації.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Вивченню сутності штучного інтелекту, зокрема теоретичним засадам функціонування ШІ чат-ботів присвячено праці таких вчених як: Корчевський М. [6], Кумар В. [7], Шулянська Є. [2], Мірджалілі В. [3], Іванченко А. [4], Худолій Ю.С. [9], Косолапенко В.С. [9] та інших вітчизняних та закордонних науковців. Незважаючи на велику кількість досліджень із заданої тематики питання розробки ШІ чат-ботів потребують більш ґрунтового вивчення.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Актуальність дослідження та необхідність вивчення теоретичних аспектів формують мету дослідження, а саме – систематизація та узагальнення наявних теоретичних аспектів розробки ШІ чат-ботів та обґрунтування процесу розробки ШІ чат-боту.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Штучний інтелект можна визначити як область комп'ютерних наук, яка займається симуляцією людського мислення в машині. Це включає створення систем, здатних виконувати завдання, які зазвичай вимагають людського інтелекту, такі як розпізнавання мови, вивчення, планування, проблемне вирішення і перцепція [1].

ШІ пропонує значні переваги для людства та допомагає вирішувати безліч складних проблем у різних галузях. Починаючи від турботи про здоров'я та медицину, де ШІ може прогнозувати захворювання та допомагати в діагностиці, до освіти, де він може пропонувати персоналізоване навчання, і до бізнесу, де він може аналізувати витрати та прогнозувати тренди.

В наш час ШІ активно використовується практично у всіх галузях промисловості та науки. Він допомагає автоматизувати виробничі лінії, покращувати системи прогнозування погоди, розпізнавати та класифікувати зображення та тексти, створювати рекомендаційні системи для сервісів споживання контенту, розробляти інтелектуальні системи керування домашніми пристроями та багато іншого [2].

ШІ може бути розділений на кілька типів, в основному, на основі їх здібностей та області застосування [3]:

1. Штучний вузькоспеціалізований (обмежений) інтелект (Narrow AI): Цей тип ШІ розроблений для виконання конкретного завдання і може працювати в рамках структурованого середовища з певним набором правил. До таких систем можна віднести більшість вже існуючих ШІ, включаючи, наприклад, рекомендаційні системи, аналізатори текстів, системи розпізнавання голосу та нашого RAG.

2. Штучний загальний інтелект (AGI): Це тип ШІ, який має здатність виконувати будь-яке завдання, яке людина може виконати, завдяки своїй здатності розуміти, вчитися та застосовувати знання у широкому спектрі різних областей. AGI настільки гнучкий та адаптивний, що може аналізувати та вирішувати складні проблеми в рамках неструктурованих середовищ та навчатися без чітко заданих правил. AGI поки що залишається в основному в рамках теорії та активних досліджень.

3. Надінтелект (Superintelligent AI): Це концептуальний тип ШІ, який не тільки в змозі виконувати те, що людина може зробити, але і перевершує людський інтелект у всіх його аспектах, включаючи креативність, загальну мудрість та соціальні навички. Надінтелект нині залишається предметом наукової фантастики та філософських дебатів.

Різні типи ШІ являють собою різні рівні здібностей і потенціалу та застосовуються залежно від поставлених цілей та завдань. Саме штучний інтелект лежить в основі чат-ботів.

Сучасні чат-боти швидко завойовують популярність і стають незамінними інструментами в багатьох сферах життєдіяльності людини. Їх функціональні можливості постійно розширюються, дозволяючи автоматизувати рутинні завдання, покращувати взаємодію з користувачами та надавати широкий спектр послуг - від надання інформації про погоду чи курс валют до запису клієнтів та виконання функцій віртуальної бібліотеки. Завдяки розвитку штучного інтелекту, чат-боти можуть інтегруватися в різні сфери, підвищуючи ефективність та зручність комунікації. У майбутньому їхня роль буде лише зростати, надаючи ще більше можливостей для бізнесу та суспільства.

Чат-боти штучного інтелекту справляють значний вплив на ринок електронної комерції, виступаючи потужним інструментом підвищення ефективності бізнесу та його масштабування. Статистичні дані від ChatInsight свідчать, що на початку 2024 року близько 91% онлайн-магазинів у США впровадили інтелектуальні чат-боти, здатні самостійно обробляти понад 70% запитів клієнтів, що дозволяє суттєво

знизити навантаження на людські ресурси [5]. У глобальному масштабі це забезпечує економію близько 2,5 мільярда робочих годин на рік, що вказує на надзвичайну ефективність цієї технології. Таким чином, впровадження ШІ чат-ботів є важливим кроком до оптимізації бізнес-процесів, а вибір найкращого сервісу є ключовим фактором для подальшого розвитку та успіху підприємств [6].

Кумар В. зазначає, що ШІ чат-боти, як комп'ютерні програми, що імітують та обробляють людські розмови, відіграють важливу роль у підвищенні ефективності взаємодії між користувачами та цифровими системами. Вони можуть бути як простими програмами для виконання базових запитів, так і складними цифровими асистентами, які використовують штучний інтелект для навчання та персоналізації взаємодії з користувачами. Це дає змогу компаніям не лише масштабувати свої операції, але й одночасно персоналізувати та проактивно обслуговувати клієнтів, що неможливо забезпечити виключно за рахунок людських ресурсів. За рахунок автоматизації комунікацій чат-боти здатні значно зменшити навантаження на персонал, підвищуючи швидкість та якість обслуговування великої кількості клієнтів [7]. Варто зазначити, впровадження чат-ботів сприяє оптимізації бізнес-процесів, зменшенню витрат та підвищенню економічної ефективності підприємств, що робить їх важливим інструментом в умовах сучасної конкуренції та цифровізації бізнесу.

ШІ чат-бот - це програма, яка використовує алгоритми машинного навчання і штучного інтелекту, щоби виконувати різноманітні задачі, оптимізувати процеси та полегшувати взаємодію користувачів з цифровими сервісами. Принцип роботи АШІ чат-бота заснований на обробці вхідних повідомлень від користувача і формуванні відповідей на них. Цей процес включає розпізнавання природної мови (NLP), що є ключовим компонентом майже всіх сучасних чат-ботів. Завдяки NLP, чат-боти здатні розуміти велику кількість варіацій слів і фраз, що використовуються людьми, і відповідно реагувати на них [4].

Щодо суміжних технологій, то однією з них є обмін повідомленнями, оскільки ШІ чат-боти в основному використовують технології обміну повідомленнями для взаємодії з користувачами. Іншою важливою технологією є хмарні обчислення, оскільки багато сучасних ботів зберігають і обробляють дані в хмарі.

Чат-боти мають чимало переваг. Вони здатні працювати цілодобово, 7 днів на тиждень, що забезпечує немиттєве реагування на запит користувача без перерв. Окрім того, вони здатні обслуговувати багато користувачів одночасно, що значно підвищує продуктивність роботи. І, нарешті, чат-боти можуть бути навчені опрацьовувати громіздкі дані та підтримувати глибокий рівень персоналізації у взаємодії з користувачами. Порівняння чат-ботів з суміжними технологіями наведено у табл. 1.

Таблиця 1.

**Порівняння чат-ботів з суміжними технологіями**

Технологія	Опис	Схожість з чат-ботами	Відмінності від чат-ботів
Віртуальні помічники	Програми, які можуть виконувати завдання, такі як установка будильників, відтворення музики, пошук інформації.	Можуть взаємодіяти з користувачами природною мовою. Можуть виконувати завдання на запит.	Зазвичай мають більш обмежений набір функцій. Не завжди можуть вести бесіди.
Голосові помічники	Програми, які можуть взаємодіяти з користувачами за допомогою голосових команд.	Можуть взаємодіяти з користувачами природною мовою. Можуть виконувати завдання на запит.	Обмежені голосовим інтерфейсом. Не можуть надавати інформацію в текстовому вигляді.
Системи машинного перекладу	Програми, які можуть перекладати текст або мову з однієї мови на іншу.	Можуть обробляти і розуміти природну мову.	Не призначені для ведення бесід. Не можуть виконувати завдання.
Системи FAQ	Системи, які можуть відповідати на часто задавані питання.	Можуть надавати інформацію у відповідь на запити користувачів.	Не можуть вести бесіди. Не можуть виконувати завдання.

Джерело: складено автором на основі [7, 3, 4]

Факт того, що чат-боти дуже популярні зараз, підтверджується статистичними даними. Згідно зі звітом Business Insider отримано досить вражаючі статистичні дані: до 2022 року 80% підприємств вже використовували або планували використовувати чат-боти. Особливо активно цю технологію застосовують у секторі обслуговування клієнтів [8]. Кількість чат-ботів стрімко зростає, з 0.3 мільйона в 2017 році до 5.5 мільйона в 2023 році. Тенденція зростання популярності чат-ботів протягом останніх років наведено на рис. 1.

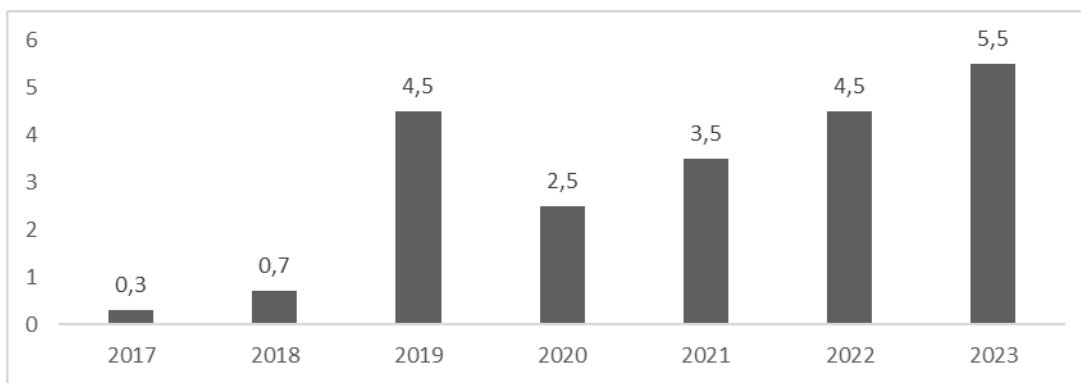


Рис. 1. Кількість чат-ботів, млн од.

Джерело: складено автором на основі [8]

Грунтуючись на актуальності дослідження та теоретико-методичних засадах представлених в літературі [6, 7, 2 - 4, 9] в роботі пропонується процес розробки ШІ чат бота. Етапи розробки ШІ чат-бота включають кілька ключових кроків, спрямованих на побудову ефективної системи взаємодії з користувачами (рис 2).

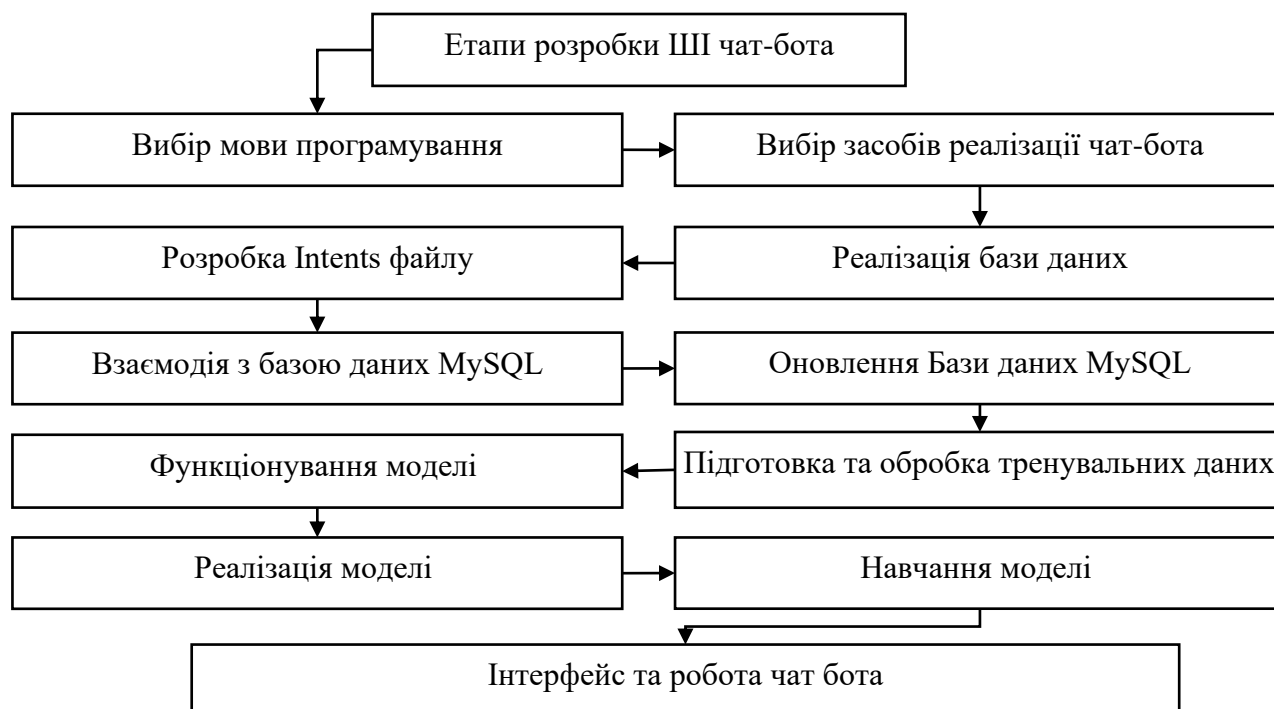


Рис. 2. Процес розробки ШІ чат-бота

Джерело: розроблено автором

Розробка ШІ чат-бота є складним та багатограним процесом, який вимагає дотримання кількох ключових етапів. Першим кроком є вибір мови програмування, яка буде використовуватися для реалізації чат-бота. Зазвичай, для цієї мети обирають мови, такі як Python, JavaScript або C++, адже вони забезпечують зручну інтеграцію з бібліотеками для обробки природної мови (NLP) та роботи з базами даних. Правильний вибір мови програмування є важливим для забезпечення ефективності розробки та подальшого обслуговування чат-бота.

Наступним етапом є вибір засобів реалізації чат-бота. На цьому етапі розробник визначає, яка платформа або фреймворк буде використано для створення чат-бота. Серед популярних варіантів можна виділити Google Dialogflow, Microsoft Bot Framework та Rasa. Ці платформи надають потужні інструменти для обробки запитів користувачів, інтеграції з іншими системами та автоматизації навчання, що значно спрощує процес розробки.

Після вибору інструментів слідує реалізація бази даних. На цьому етапі створюється структура бази даних, що дозволяє зберігати інформацію про користувачів, сеанси спілкування та історію взаємодій. Використання реляційних баз даних, таких як MySQL або PostgreSQL, забезпечує зручний доступ до даних і

можливість їх швидкої обробки. Правильна реалізація бази даних є критично важливою для функціонування чат-бота, оскільки вона дозволяє ефективно управляти інформацією.

Паралельно з цим розробляється Intents файл, у якому визначаються наміри користувачів, що взаємодіють з чат-ботом. Intents – це ключові елементи, які відображають наміри та запити користувачів, а також відповідні до них відповіді. Цей етап включає створення прикладів запитів, що допоможуть системі правильно ідентифікувати наміри під час взаємодії.

Подальшим кроком є взаємодія чат-бота з базою даних MySQL. Розробники реалізують механізми, які дозволяють чат-боту отримувати та зберігати дані, обробляти запити та підтримувати актуальність інформації в базі даних. Це дозволяє зберігати важливу інформацію про користувачів та їхню взаємодію з ботом, що є основою для подальшого аналізу та вдосконалення сервісу.

Важливим етапом є також оновлення бази даних MySQL. Чат-бот повинен мати можливість автоматично оновлювати базу даних, додаючи нові записи або оновлюючи існуючі під час взаємодії з користувачами. Це забезпечує актуальність даних для аналізу та прогнозування, що сприяє покращенню взаємодії з користувачами.

Після цього настає етап обробки даних та підготовки тренувальних даних. На цьому етапі розробники збирають і підготовлюють дані для навчання моделі. Вони можуть включати приклади запитів користувачів і очікувані відповіді, що дозволяє моделі навчитися розпізнавати шаблони взаємодій. Правильна підготовка даних є критично важливою, оскільки від цього залежить якість навчання моделі.

Функціонування моделі є наступним етапом. Чат-бот використовує модель машинного навчання для розпізнавання намірів і генерації відповідей на запити користувачів. На цьому етапі важливо, щоб модель функціонувала ефективно, забезпечуючи швидку і точну обробку запитів у реальному часі. Це дозволяє покращити досвід користувачів та підвищити їхню задоволеність.

Після реалізації моделі відбувається навчання. Модель проходить навчання на підготовлених даних, що дозволяє їй адаптуватися до нових запитів і формувати відповідні реакції. Після первинного навчання проводиться перевірка на тестових даних, що допомагає оцінити якість моделі та налаштувати її параметри для покращення продуктивності.

Останнім етапом є розробка інтерфейсу та реалізація роботи чат-бота. Розробники створюють інтерфейс, який дозволяє користувачам взаємодіяти з чат-ботом через різні канали, такі як вебсайти, месенджери (наприклад, Telegram або Facebook Messenger) або корпоративні системи. Важливо забезпечити зручність взаємодії та підтримку основних функцій, таких як відповідь на запити, ініціювання розмов та обробка повідомлень у реальному часі.

Таким чином, процес розробки чат-бота складається з багатьох етапів, що потребують ретельного планування та реалізації. Кожен з цих етапів має своє значення і впливає на загальну ефективність роботи чат-бота, що, в свою чергу, забезпечує якісну взаємодію з користувачами та відповідає сучасним вимогам бізнесу.

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

Штучний інтелект є потужним інструментом, який відкриває нові можливості для людства в багатьох сферах, включаючи медицину, освіту та бізнес. Завдяки своїй здатності виконувати складні завдання, ШІ істотно полегшує життя, автоматизуючи процеси та покращуючи ефективність. ШІ чат-боти, зокрема, продемонстрували свою незамінність у сучасному бізнес-середовищі, надаючи підприємствам можливість підвищити продуктивність, знизити витрати та оптимізувати обслуговування клієнтів.

Процес розробки ШІ чат-ботів є складним і багатогранним, що включає етапи від вибору мови програмування до реалізації інтерфейсу для користувачів. Кожен з цих етапів має критичне значення для створення ефективної системи, яка відповідає потребам бізнесу і забезпечує якісну взаємодію з клієнтами. Актуальність та важливість впровадження ШІ чат-ботів в бізнесі підкреслюється статистичними даними, що свідчать про їх стрімке зростання та інтеграцію в різні сфери. У майбутньому, з розвитком технологій, роль чат-ботів буде лише зростати, забезпечуючи ще більші можливості для автоматизації, персоналізації та підвищення ефективності бізнес-процесів. Отже, розвиток штучного інтелекту, зокрема чат-ботів, відкриває нові горизонти для бізнесу та суспільства, зумовлюючи необхідність їх впровадження для підтримки конкурентоспроможності в умовах швидко змінюваного ринкового середовища.

### Література

1. Що таке штучний інтелект: історія, види та складові. GigaCloud. URL: <https://gigacloud.ua/blog/navchannja/scho-take-shtuchnij-intelekt-istorija-vidi-ta-skladovi>
2. Шуляньська Є. Що таке штучний інтелект простими словами? БізнесТренд. 2024. URL: <https://biznestrendy.com.ua/shcho-take-shtuchnyy-intelekt-prostymy-slovamy/>
3. Mirjalili V. Python machine learning: machine learning and deep learning with python, scikit-learn, and tensorflow 2. 3<sup>rd</sup> edition. Birmingham. 2019. 742 p.

4. Іванченко А. Що таке ChatGPT та чому чат боти ШІ це майбутнє? Acer.Blog. 2024. URL: <https://blog.acer.com/ua/discussion/1701/scho-take-chatgpt-ta-chomu-chat-boti-shi-ce-maybutnye>
5. 50 Chatbot Statistics You Should Follow in 2024. Chatinsight. 2024. URL: <https://www.chatinsight.ai/chatbots/chatbot-statistics/>
6. Корчевський М. 7 найкращих чат-ботів з ШІ у 2024 році. Claspо. 2024 URL: <https://claspo.io/ua/blog/7-best-ai-chatbots-in-2024/>
7. Кумар В. Очікується, що на ринку чат-ботів очікується значне зростання. Північна Америка очолює зростання! Univdatos. 2022. URL: <https://univdatos.com/uk/%D1%80%D0%>
8. Businessinsider. 2020. URL: <https://www.businessinsider.com>
9. Худолій Ю.С., Косолапенко В.С. Особливості застосування чат-ботів на основі штучного інтелекту у фінансовій сфері. Економіка і регіон, 2023 № 3 (90). С. 97-103. DOI: 10.26906/EiR.2023.3(90).3036

### References

1. Shcho take shtuchnyi intelekt: istoriia, vydy ta skladovi. GigaCloud. URL: <https://gigacloud.ua/blog/navchannja/scho-take-shtuchnij-intelekt-istorija-vidi-ta-skladovi>
2. Shulianska Ye. Shcho take shtuchnyi intelekt prostymy slovamy? BiznesTrend. 2024. URL: <https://biznestrendy.com.ua/shcho-take-shtuchnyy-intelekt-prostymy-slovamy/>
3. Mirjalili V. Python machine learning: machine learning and deep learning with python, scikit-learn, and tensorflow 2. 3rd edition. Birmingham. 2019. 742 p.
4. Ivanchenko A. Shcho take ChatGPT ta chomu chat boty ShI tse maibutnie? Acer.Blog. 2024. URL: <https://blog.acer.com/ua/discussion/1701/scho-take-chatgpt-ta-chomu-chat-boti-shi-ce-maybutnye>
5. 50 Chatbot Statistics You Should Follow in 2024. Chatinsight. 2024. URL: <https://www.chatinsight.ai/chatbots/chatbot-statistics/>
6. Korchevskiy M. 7 naikrashchykh chat-botiv z ShI u 2024 rotsi. Slaspо. 2024 URL: <https://claspo.io/ua/blog/7-best-ai-chatbots-in-2024/>
7. Kumar V. Ochikuietsia, shcho na rynku chat-botiv ochikuietsia znachne zrostannia. Pivnichna Ameryka ocholyt zrostannia! Univdatos. 2022. URL: <https://univdatos.com/uk/%D1%80%D0%>
8. Businessinsider. 2020. URL: <https://www.businessinsider.com>
9. Khudolii Yu.C., Kosolapenko V.C. Osoblyvosti zastosuvannia chat-botiv na osnovi shtuchnogo intelektu u finansovii sferi. Ekonomika i rehion, 2023 № 3 (90). S. 97-103. DOI: 10.26906/EiR.2023.3(90).3036

<https://doi.org/10.31891/2219-9365-2024-79-26>

УДК 004

ЗАГОРУЛЬКО Олександр

Приватний вищий навчальний заклад «Європейський університет»

<https://orcid.org/0009-0008-7626-3874>

## ТЕХНОЛОГІЇ ПОКРАЩЕННЯ ЗАХИСТУ ІНФОРМАЦІЇ КОМП'ЮТЕРНИХ МЕРЕЖ

*Стаття розглядає сучасні технології забезпечення безпеки інформації у комп'ютерних мережах, аналізує методи підвищення захищеності даних від кіберзагроз та оцінює ефективність різних систем захисту. Запропоновано підхід до розробки інтегрованої системи захисту, що включає засоби криптографії, аутентифікації, контролю доступу, а також методи виявлення та запобігання загрозам на основі штучного інтелекту (ШІ). Особливо розглянуто важливість апаратного та програмного захисту, а також питання безпеки в хмарних середовищах.*

*Ключові слова: захист інформації, комп'ютерні мережі, кібербезпека, криптографія, контроль доступу, штучний інтелект.*

ZAHORULKO Oleksandr

Private Higher Educational Institution «European University»

## TECHNOLOGIES FOR ENHANCING INFORMATION SECURITY IN COMPUTER NETWORKS

*This article examines modern technologies for ensuring information security in computer networks, analyzing methods to strengthen data protection against cyber threats and assessing the effectiveness of various security systems. It proposes an integrated approach to developing a comprehensive protection system that includes cryptography, authentication, and access control, along with threat detection and prevention techniques powered by artificial intelligence (AI). Special attention is given to both hardware and software security solutions, as well as the unique challenges and strategies for maintaining security in cloud environments.*

*The study begins by exploring foundational concepts in network security, such as the importance of cryptographic methods for safeguarding sensitive data, including encryption and hashing techniques. By encoding information, cryptographic systems make data accessible only to authorized users with the correct decryption keys, offering a robust initial layer of defense. Authentication methods, such as multi-factor authentication (MFA) and biometrics, are also examined as essential measures for verifying user identity, thereby preventing unauthorized access.*

*The article highlights the role of access control systems, which enforce policies determining user permissions within a network. Role-Based Access Control (RBAC) and Attribute-Based Access Control (ABAC) are discussed in detail, showcasing how these approaches minimize the risk of unauthorized data access by aligning permissions with organizational roles or specific attributes.*

*A significant focus is placed on AI-driven threat detection and prevention systems. Machine learning algorithms can analyze network traffic patterns to detect anomalies, which may indicate potential security breaches. This proactive method of identifying suspicious behavior allows systems to react to threats in real time, often mitigating risks before they cause harm.*

*Finally, the article addresses the growing importance of cloud security. As organizations increasingly rely on cloud services, they must contend with new security considerations, such as shared responsibility between cloud providers and clients. Effective cloud security involves robust encryption, stringent access management, and continuous monitoring to ensure that sensitive data remains protected in a virtualized, highly accessible environment.*

*Keywords: information security, computer networks, cybersecurity, cryptography, access control, artificial intelligence, cloud security*

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Безпека інформації стала одним із ключових викликів сучасного інформаційного суспільства. З кожним роком кількість кіберзлочинів неухильно зростає, а методи й технології атак стають дедалі складнішими й витонченішими. Це створює нагальну потребу в розробці комплексних і адаптивних підходів до захисту даних, що циркулюють у комп'ютерних мережах та використовуються як в особистих, так і в корпоративних цілях.

Надійний захист інформації включає не лише забезпечення захисту від зовнішніх атак, а й створення умов для безпечного зберігання, обробки та передачі даних в межах мережі, незалежно від її масштабу. Особливо важливим є контроль за привілеями доступу користувачів, регулярне оновлення програмного забезпечення, яке мінімізує можливості атак, та впровадження системи моніторингу, що виявляє підозрілу активність у режимі реального часу.

Питання кібербезпеки охоплюють широкий спектр технологій, які мають забезпечувати цілісність, доступність і конфіденційність даних у різних середовищах, включаючи хмарні технології, мобільні платформи та розподілені мережі. Особливу увагу приділяють захисту інформації у хмарних середовищах, де виникає потреба в нових підходах до забезпечення безпеки через високу доступність даних для

користувачів і динамічні умови зберігання. У результаті комплексна система кіберзахисту сьогодні включає засоби шифрування, автентифікації, контроль доступу, а також інтелектуальні методи виявлення та нейтралізації загроз, що дозволяють адаптуватися до змінюваних умов сучасного кіберсередовища.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

Існує велика кількість досліджень у сфері кібербезпеки, що стосуються як програмних, так і апаратних рішень. Розглянуто комплексний підхід до розробки системи захисту, яка використовує методи мультифакторної аутентифікації та динамічний контроль доступу. Загалом потрібно зосереджуватися на питаннях шифрування даних та алгоритмах, що забезпечують більш ефективне шифрування з меншими витратами ресурсів. Згідно з дослідженням, застосування ШІ та машинного навчання дозволяє виявляти аномалії в мережевих трафіках і запобігати більшості атак у режимі реального часу.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою цієї роботи є аналіз технологій підвищення захисту інформації в комп'ютерних мережах та розробка рекомендацій щодо інтеграції сучасних рішень, зокрема, мультифакторної аутентифікації, методів контролю доступу та запобігання загрозам. Стаття також описує вплив використання ШІ та технологій блокчейн для покращення безпеки в хмарних середовищах.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Захист інформації в комп'ютерних мережах сьогодні є одним із найважливіших завдань у сфері кібербезпеки. З розвитком цифрових технологій збільшується ризик кіберзагроз, які можуть призвести до втрати конфіденційної інформації, порушення роботи критичних інфраструктур та значних фінансових втрат. Для забезпечення надійного захисту комп'ютерних мереж розроблено численні технології, кожна з яких має свої переваги та недоліки. До найпоширеніших технологій, що забезпечують кібербезпеку в сучасних мережах, належать мультифакторна аутентифікація, контроль доступу, криптографія, інтелектуальні системи виявлення загроз, апаратні та програмні засоби для захисту мережі, технології блокчейн та хмарні рішення з інтегрованою безпекою.

Мультифакторна аутентифікація (MFA) є важливим компонентом у захисті мереж від несанкціонованого доступу. Традиційна аутентифікація за допомогою лише пароля має значні вразливості, оскільки паролі можуть бути перехоплені або зламані. MFA ж використовує кілька рівнів перевірки, що забезпечує значно вищий рівень захисту. Основні компоненти MFA включають знання (пароль), володіння (мобільний пристрій, на який надсилається код підтвердження), та біометричні фактори (відбиток пальця, розпізнавання обличчя). Наприклад, можна зобразити процес MFA у вигляді діаграми, де кожен етап представляє один із факторів підтвердження.

MFA вимагає інтеграції кількох технологій, що забезпечують одночасно зручність та безпеку. Успішна реалізація цієї технології забезпечує баланс між зручністю використання та захистом, зменшуючи ймовірність того, що зловмисник отримає доступ до мережі навіть у разі компрометації одного з факторів. Наприклад, на етапі біометричної аутентифікації застосовуються технології розпізнавання відбитків пальців або обличчя, а також система безпечного зберігання біометричних шаблонів.

Наприклад:

Псевдокод процесу MFA для авторизації в мережі має наступний вигляд:

IF (пароль == вірний) THEN

    Запросити другий фактор аутентифікації (код через SMS чи додаток)

    IF (код == вірний) THEN

        Запросити третій фактор (біометрія)

        IF (біометрія == вірна) THEN

            Доступ надано

        ELSE

            Відмовити в доступі

        ENDIF

    ELSE

        Відмовити в доступі

    ENDIF

ELSE

    Відмовити в доступі

ENDIF

Контроль доступу є наступною важливою технологією кібербезпеки. Він забезпечує управління правами доступу до даних та ресурсів мережі на основі ролей та привілеїв користувачів. Система Role-Based Access Control (RBAC) дозволяє адмініструвати права доступу до ресурсів, мінімізуючи ризик



несанкціонованого доступу. Наприклад, у корпоративній мережі права доступу можуть бути розподілені відповідно до посадових обов'язків користувачів, забезпечуючи захист конфіденційних даних. Модель доступу, орієнтована на ролі, може бути представлена діаграмою, де кожна роль має визначений рівень доступу до певних ресурсів.

Для реалізації RBAC застосовуються спеціальні правила, які обмежують дії користувачів на основі їх ролей. Наприклад, псевдокод для надання доступу в системі на основі RBAC виглядає наступним чином:

ROLE адміністратора:

Доступ до всіх ресурсів

ROLE співробітника:

Доступ до ресурсів рівня співробітника

IF (користувач має роль адміністратора) THEN

Доступ до всіх ресурсів

ELSE IF (користувач має роль співробітника) THEN

Доступ до ресурсів рівня співробітника

ELSE

Відмовити в доступі

ENDIF

Криптографія є основою для захисту даних під час їх передачі та зберігання. Алгоритми симетричного шифрування, такі як AES, використовуються для швидкої та ефективної передачі зашифрованих даних. У той же час, асиметричне шифрування, зокрема RSA або ECC, забезпечує високий рівень безпеки для аутентифікації та передачі ключів. Наприклад, шифрування симетричним алгоритмом AES можна представити формулою:

$$C = E_k(M)$$

де  $C$  — зашифрований текст,  $M$  — початкове повідомлення,  $E$  — алгоритм шифрування, а  $k$  — ключ. У свою чергу, для асиметричних алгоритмів використовується публічний і приватний ключі для шифрування та розшифрування.

Інтелектуальні системи виявлення та запобігання загрозам (IDS/IPS) забезпечують виявлення аномальних дій у мережевому трафіку та запобігають потенційним загрозам у режимі реального часу. Алгоритми машинного навчання аналізують поведінкові шаблони, що дозволяє виявляти навіть невідомі загрози. IDS та IPS автоматично блокують небезпечний трафік, запобігаючи витоків даних та несанкціонованому доступу. Наприклад, для класифікації трафіку в IDS можна використати алгоритми кластеризації, які дозволяють виявляти аномалії. Функція виявлення аномалій описується так:

$$I, \& \text{\{якщо\}} x \text{\{ — аномалія\}} \setminus 0, \& \text{\{інакше\}} \end{\text{cases}}$$

де  $A(x)$  — функція, що визначає, чи є трафік аномальним (1 — аномалія, 0 — нормальний трафік).

Апаратне та програмне забезпечення для безпеки мережі також є важливим елементом захисту. Апаратні засоби, такі як міжмережеві екрани, шлюзи безпеки та захищені модулі шифрування (HSM), забезпечують додатковий рівень безпеки та можуть інтегруватися з програмними рішеннями. Це дозволяє створити комплексну систему захисту, яка забезпечує більш ефективний захист від кіберзагроз. Наприклад, при використанні VPN забезпечується безпечний тунель для передачі даних, який шифрує інформацію між кінцевими точками.

Технології блокчейн пропонують новий рівень захисту для комп'ютерних мереж, особливо у сфері збереження логів дій та подій. Блокчейн, як незмінний реєстр даних, може використовуватися для безпечного зберігання інформації про доступи, транзакції та дії у мережі, забезпечуючи їхню автентичність і захист від несанкціонованих змін. Це має критичне значення для забезпечення цілісності даних, адже блокчейн дозволяє автоматично фіксувати всі мережеві події та дії користувачів у форматі, який неможливо видалити чи змінити без згоди всіх учасників мережі.

Блокчейн стає потужним інструментом для моніторингу мережевої активності та кіберзахисту. Зберігання даних про доступ та аутентифікацію в блокчейні дозволяє знизити ризик внутрішніх загроз, оскільки жоден користувач не може змінити запис без відповідної верифікації. Це особливо корисно у великих мережах, де потрібно підтримувати надійний контроль доступу та простежувати дії користувачів. Блокчейн допомагає ідентифікувати аномалії та потенційні загрози в режимі реального часу, оскільки всі зміни у мережі відстежуються на блокчейн-рівні, і будь-яка спроба втручання миттєво стає видимою для адміністраторів мережі.

У корпоративних комп'ютерних мережах блокчейн також забезпечує стійкість до зловмисного втручання, особливо при обробці конфіденційних даних. Наприклад, завдяки децентралізованій структурі блокчейну, компанії можуть захистити важливі дані, такі як логи доступів до мережі чи конфігурації систем,

від знищення чи підробки. Децентралізація дозволяє мережам продовжувати працювати навіть у разі часткових атак або збоїв, оскільки інформація зберігається на багатьох вузлах, а не в одному центральному місці.

Інтеграція блокчейн-технологій у комп'ютерні мережі відкриває нові можливості для забезпечення безпеки даних у хмарних середовищах. Для зберігання та передачі інформації в хмарі блокчейн надає додатковий рівень захисту, забезпечуючи незмінність і верифікацію даних без необхідності повністю довіряти стороннім постачальникам хмарних послуг. Блокчейн дозволяє мережевим адміністраторам мати прозорий огляд всіх дій і подій, що відбуваються в хмарних сервісах, підвищуючи загальну безпеку та захист від зовнішніх і внутрішніх кіберзагроз.

Хмарні рішення з інтегрованою безпекою також стають важливою складовою сучасних комп'ютерних мереж. Хмарні провайдери пропонують рішення з інтегрованими інструментами захисту, що включають засоби шифрування, контроль доступу, регулярний аудит та моніторинг загроз. Ці заходи дозволяють зменшити ризики, пов'язані з кібербезпекою у хмарних середовищах. Наприклад, політики безпеки для кожного типу даних можна контролювати на рівні доступу до хмарного середовища, забезпечуючи надійний захист даних у багатокористувачьких платформах.

Перспективи подальших досліджень у сфері кібербезпеки передбачають інтеграцію адаптивних систем захисту, що здатні навчатися на основі нових загроз та автоматично адаптуватися до них. Розвиток технологій штучного інтелекту та розподілених обчислень дозволяє створювати інтегровані рішення, які забезпечують більш ефективний захист комп'ютерних мереж. Іншою важливою областю досліджень є стандартизація протоколів для IoT-пристроїв, що забезпечить ефективне управління та захист інформації у великих мережах, які включають різні пристрої з обмеженими ресурсами.

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ

### І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

З огляду на зростаючу складність кіберзагроз, подальші дослідження повинні зосередитися на розробці адаптивних систем захисту, здатних навчатися та пристосовуватися до нових типів загроз у режимі реального часу. Використання технологій штучного інтелекту (ШІ) та машинного навчання відкриває можливості для створення інтелектуальних систем, які можуть автоматично виявляти та реагувати на аномалії у мережевому трафіку, визначаючи потенційні загрози ще до того, як вони зможуть завдати шкоди. Інтеграція ШІ з блокчейн-технологіями також може забезпечити високий рівень прозорості та надійності, надаючи нові засоби для відстеження та верифікації дій у мережі.

Розвиток розподілених обчислень сприяє створенню ефективних рішень для захисту комп'ютерних мереж, зокрема в умовах розподілених середовищ, таких як хмарні платформи та інфраструктура Інтернету речей (IoT). Подальші дослідження в цій сфері можуть бути спрямовані на розробку архітектур, які дозволять швидко обробляти великі обсяги даних з різних джерел, одночасно забезпечуючи цілісність і конфіденційність інформації навіть при передаванні через незахищені канали.

Ще однією важливою областю досліджень є стандартизація безпекових протоколів для хмарних платформ і IoT-пристроїв. Ці платформи стають усе популярнішими, але залишаються вразливими через відсутність єдиних стандартів захисту. Дані, що передаються через хмарні платформи та IoT-пристрої, часто проходять через загальнодоступні або недостатньо захищені канали, що підвищує ризик перехоплення та несанкціонованого доступу. Впровадження єдиних стандартів для шифрування, автентифікації та контролю доступу в цих системах сприятиме зменшенню вразливостей, підвищенню довіри до таких технологій та покращенню їхньої взаємодії.

Також перспективним напрямом є розробка технологій для підвищення кіберстійкості мереж — здатності швидко відновлюватися після кіберінцидентів і автоматично мінімізувати збитки. Такі системи, що передбачають самовідновлення, можуть мати особливе значення для критичних інфраструктур, де безперервна робота має вирішальне значення для безпеки та економіки.

Комплексний підхід до захисту інформації, який включає сучасні технології автентифікації, шифрування, контролю доступу та системи виявлення загроз, є необхідною умовою для забезпечення безпеки у комп'ютерних мережах. Використання ШІ та блокчейну створює додаткові можливості для покращення кіберзахисту, а хмарні рішення з інтегрованими інструментами захисту дозволяють забезпечити високий рівень безпеки у сучасних інфраструктурах. Удосконалення цих технологій відкриває нові перспективи для розробників та дослідників у сфері кібербезпеки.

## Література

1. Андрієнко І., Кравчук О. Методи захисту інформації в комп'ютерних мережах / І. Андрієнко, О. Кравчук // Системи захисту інформації. – 2022. – Т. 3. – № 24. – С. 110-120. DOI: <https://doi.org/10.20535/1560-8956.24.2022.264104>
2. Офіційна документація OpenSSL : [вебсайт]. – Режим доступу : <https://www.openssl.org/>

3. Романов М.М., Поляков С.В. Криптографія та мережеві протоколи / М.М. Романов, С.В. Поляков // Харків: Харківський національний університет радіоелектроніки. – 2023. – 312 с. – Режим доступу : <https://repository.nure.ua/bitstream/123456789/910/1/%D0%9A%D1%80%D0%B8%D0%BF%D1%82%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D1%96%D1%8F.pdf>
4. Wireshark: офіційний вебсайт : [вебсайт]. – Режим доступу : <https://www.wireshark.org/>
5. Metasploit Framework : [вебсайт]. – Режим доступу : <https://www.metasploit.com/>
6. Nmap: довідка та документація : [вебсайт]. – Режим доступу : <https://nmap.org/>
7. Snort : [вебсайт]. – Режим доступу : <https://www.snort.org/>
8. Cisco Security Documentation : [вебсайт]. – Режим доступу : <https://www.cisco.com/c/en/us/support/docs/security/>
9. OWASP: рекомендації щодо веббезпеки : [вебсайт]. – Режим доступу : <https://owasp.org/>

### References

1. Andriyenko I., Kravchuk O. Information Protection Methods in Computer Networks / I. Andriyenko, O. Kravchuk // Information Security Systems. – 2022. – Vol. 3. – № 24. – P. 110-120. DOI: <https://doi.org/10.20535/1560-8956.24.2022.264104>
2. Official OpenSSL Documentation : [website]. Access mode : <https://www.openssl.org/>
3. Romanov M.M., Polyakov S.V. Cryptography and Network Protocols / M.M. Romanov, S.V. Polyakov // Kharkiv: Kharkiv National University of Radioelectronics. – 2023. – 312 p. Access mode : <https://repository.nure.ua/bitstream/123456789/910/1/%D0%9A%D1%80%D0%B8%D0%BF%D1%82%D0%BE%D0%B3%D1%80%D0%B0%D1%84%D1%96%D1%8F.pdf>
4. Wireshark: Official Website : [website]. Access mode : <https://www.wireshark.org/>
5. Metasploit Framework : [website]. Access mode : <https://www.metasploit.com/>
6. Nmap: Documentation and Help : [website]. Access mode : <https://nmap.org/>
7. Snort : [website]. Access mode : <https://www.snort.org/>
8. Cisco Security Documentation : [website]. Access mode : <https://www.cisco.com/c/en/us/support/docs/security/>
9. OWASP: Web Security Guidelines : [website]. Access mode : <https://owasp.org/>

<https://doi.org/10.31891/2219-9365-2024-79-27>

УДК 004

СЕРДЮК Юрій

Приватний вищий навчальний заклад «Європейський університет»  
<https://orcid.org/0009-0004-5483-1261>

## ІНТЕРПРЕТАЦІЯ РІШЕНЬ ШТУЧНОГО ІНТЕЛЕКТУ: ЯК ЗАБЕЗПЕЧИТИ ПРОЗОРИСТІТЬ ТА ЗРОЗУМІЛІСТЬ ДЛЯ КІНЦЕВИХ КОРИСТУВАЧІВ

У статті розглянуто проблеми інтерпретації та прозорості рішень систем штучного інтелекту (ШІ), що набувають критичного значення у сферах із високою відповідальністю, таких як медицина, фінанси, правосуддя та управління ризиками. З огляду на складність алгоритмів, зокрема нейронних мереж, рішення ШІ часто є непрозорими для кінцевих користувачів, що може знижувати рівень довіри та ускладнювати їх практичне застосування. У статті проаналізовано основні методи пояснення рішень ШІ, зокрема локальні моделі (LIME), глобальні підходи, а також метод SHAP, що базується на теорії ігор. Особливу увагу приділено ролі візуалізації для покращення інтерпретації результатів та доступності моделей для кінцевих користувачів. У статті також обговорюються переваги та обмеження існуючих методів, перспективи їх удосконалення та інтеграції в реальні системи, що дозволить підвищити прозорість, зрозумілість і довіру до алгоритмів ШІ.

Ключові слова: штучний інтелект, інтерпретованість, прозорість, пояснювальні моделі, LIME, SHAP, візуалізація, довіра, кінцевий користувач, пояснення рішень

SERDYUK Yuriy

Private Higher Educational Institution «European University»

## INTERPRETATION OF ARTIFICIAL INTELLIGENCE DECISIONS: HOW TO ENSURE TRANSPARENCY AND COMPREHENSIBILITY FOR END USERS

This article delves into the pressing issues of interpretation and transparency in artificial intelligence (AI) decision-making systems, particularly in critical domains such as healthcare, finance, legal decision-making, and risk management. As AI models become increasingly complex—especially with the rise of neural networks and deep learning models—the decisions produced by these systems are often opaque to end users. This opacity can lead to decreased trust, limited adoption, and challenges in practical implementation, particularly in sectors where decisions directly impact human lives and well-being. Ensuring that users can understand, evaluate, and, if necessary, question AI decisions is thus becoming a vital part of responsible AI development.

The article provides an in-depth review of foundational methods for achieving AI interpretability. Key techniques covered include Local Interpretable Model-agnostic Explanations (LIME), which offers localized insights into specific model outputs, and SHapley Additive exPlanations (SHAP), a game-theoretic approach that quantifies the contribution of each feature to a given decision. In addition to these, the article examines global interpretive methods that aim to reveal overarching patterns and rules that guide the AI model's behavior.

Moreover, the article highlights the pivotal role of visualization tools—such as feature importance charts, heatmaps, and temporal visualizations—that help make complex AI processes more understandable and intuitive for non-technical users. Visualization techniques are instrumental in conveying how certain features or data points influence model outcomes, thus bridging the gap between technical model outputs and user-friendly insights.

The article also critically assesses the strengths and limitations of current interpretative approaches, such as the computational demands of SHAP and the localized nature of LIME, which may not generalize across datasets. Finally, the article discusses future directions for advancing interpretability, including the development of real-time interpretive methods, hybrid approaches combining local and global explanations, and interactive user interfaces that allow end users to query AI models for explanations. By addressing these challenges and evolving interpretative frameworks, AI systems can be better aligned with the values of transparency, accountability, and trustworthiness, making them more suitable for sensitive, high-stakes applications.

Keywords: artificial intelligence, interpretability, transparency, explainable AI, LIME, SHAP, visualization, accountability, trust, end-user comprehension, decision-making transparency

### ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОК ІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

Застосування штучного інтелекту (ШІ) швидко поширюється в різних сферах життя, таких як медицина, фінанси, право, освіта та управління ризиками. Алгоритми ШІ, зокрема нейронні мережі та методи глибокого навчання, забезпечують високу точність і ефективність у вирішенні багатьох складних завдань, але їх складність часто робить процес прийняття рішень непрозорим. Багато з цих моделей функціонують як «чорні ящики», в яких кінцевий користувач не може оцінити, як і на основі яких факторів алгоритм приймає певне рішення.

Ця непрозорість створює значні проблеми, особливо коли ШІ застосовується в критичних сферах, де від його рішень залежать життя, здоров'я, фінансовий стан або навіть безпека користувачів. Відсутність розуміння механізмів роботи ШІ може призвести до зниження довіри з боку користувачів, зокрема медичних працівників, фінансових аналітиків та правозахисників, які покладаються на рекомендації цих систем у повсякденній роботі. Крім того, така непрозорість може створювати юридичні та етичні проблеми, оскільки неможливо перевірити обґрунтованість рішень ШІ або відповідально їх обґрунтувати.

Таким чином, постає гостра потреба у створенні інтерпретованих і прозорих моделей, які б дозволили користувачам не лише отримувати результат, а й розуміти логіку прийняття рішення. Це особливо актуально для забезпечення відповідальності та надійності систем ШІ в умовах, де помилки можуть призвести до серйозних наслідків. Розробка таких моделей допоможе кінцевим користувачам отримати уявлення про принципи роботи ШІ, а також підвищить довіру та впевненість у їх використанні.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

У галузі інтерпретованості ШІ активно досліджуються такі методи, як пояснювальні моделі, локальні та глобальні інтерпретації, моделі-агенти для пояснення рішень тощо. Деякі роботи фокусуються на розробці спеціалізованих підходів для інтерпретації складних моделей, зокрема методів LIME (Local Interpretable Model-agnostic Explanations) та SHAP (SHapley Additive exPlanations). Існує велика кількість досліджень, присвячених оцінці ефективності цих методів у різних галузях застосування, зокрема у медицині, де зрозумілість рішень ШІ є критично важливою для лікарів і пацієнтів.

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою цієї роботи є аналіз методів, які дозволяють забезпечити інтерпретацію рішень ШІ, а також рекомендації для розробників та дослідників щодо впровадження прозорих моделей. Особливу увагу буде приділено методам, які допомагають кінцевим користувачам розуміти, як і чому були прийняті ті чи інші рішення, що особливо важливо в критично значущих системах.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

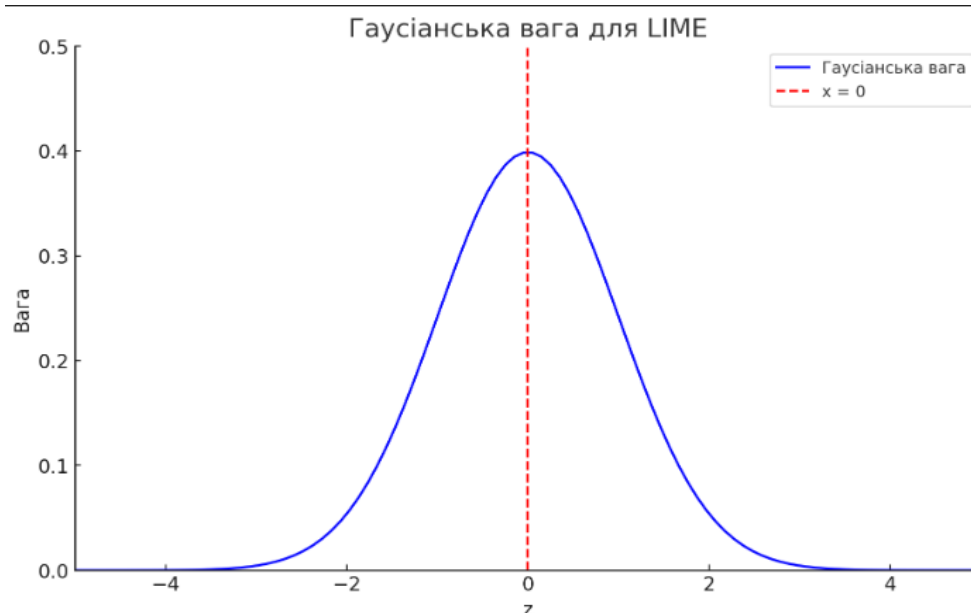
Інтерпретованість моделей ШІ стала однією з ключових вимог у сучасних системах, де результат обчислень впливає на критичні рішення. У таких сферах, як медицина, фінанси, управління ризиками та правосуддя, алгоритми ШІ часто мають значний вплив на життя людей. Кінцеві користувачі, наприклад лікарі або банківські аналітики, потребують не лише точного результату, але й розуміння логіки, яка стоїть за конкретним передбаченням. Це важливо для того, щоб мати можливість обґрунтовано довіряти алгоритму, коригувати його роботу в разі виявлення неточностей та відповідати на запитання користувачів.

Крім того, інтерпретованість стає важливою з точки зору етики та відповідальності. Складні «чорні ящики», наприклад, глибокі нейронні мережі, можуть не враховувати специфіку індивідуального випадку або ж діяти, спираючись на певні приховані упередження в даних. Розробка моделей з інтерпретованими рішеннями дозволяє створювати справедливі та відповідальні алгоритми, які можна пояснити та верифікувати з боку експертів і користувачів.

#### Локальні пояснювальні моделі

Локальні пояснювальні моделі розглядають рішення моделі в контексті конкретного випадку, або обчислювального результату. Одним із найпопулярніших методів є LIME (Local Interpretable Model-agnostic Explanations), який надає можливість створювати локальні сурогатні моделі для пояснення рішень будь-якої алгоритмічної системи.

Ось графік гаусіанської ваги, яка використовується в LIME. Графік показує, як вага  $w(x)$  змінюється в залежності від значення  $x$  навколо точки  $x=0$  (позначено червоною пунктирною лінією).



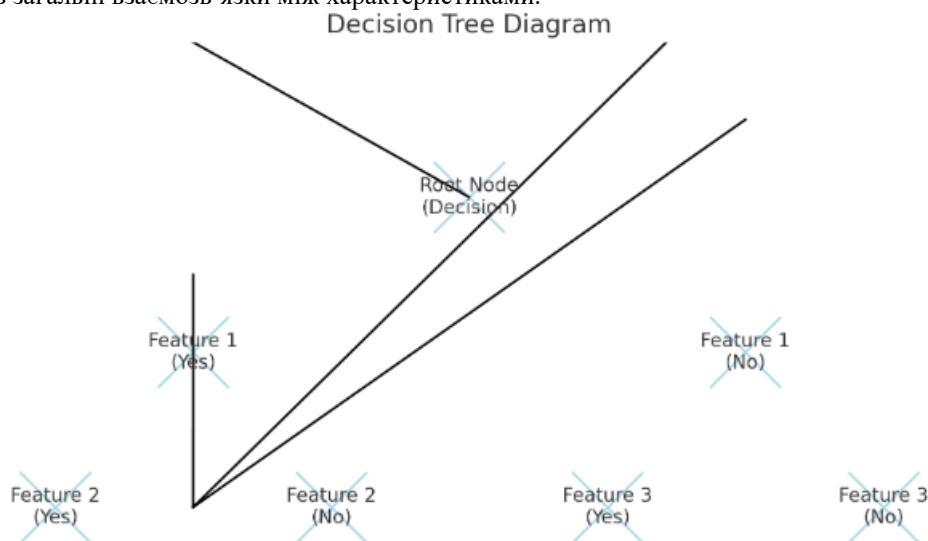
Вага максимальна в точці xxx і зменшується, коли ми віддаляємося від цієї точки. Це дозволяє зосередитися на зразках, які найближчі до обраного прикладу, для створення локальної інтерпретаційної моделі.

LIME працює шляхом генерації варіацій початкового випадку з незначними змінами в даних, а потім спостерігає, як модель змінює своє передбачення у відповідь на ці варіації. Отримані результати використовуються для побудови лінійної або іншої простої моделі, яка пояснює вплив кожного з факторів на конкретне рішення.

Приклад застосування LIME можна розглянути в контексті медичної діагностики, де алгоритм ШІ передбачає наявність захворювання. Застосовуючи LIME, лікар може побачити, які симптоми або фактори найбільше вплинули на передбачення, що дозволяє підвищити довіру до результату та прийняти обґрунтоване рішення.

#### Глобальні пояснювальні моделі

На відміну від локальних, глобальні моделі пояснюють загальні закономірності, які використовує алгоритм для прийняття рішень. Одним з таких методів є побудова дерев рішень або правил, які відображають загальні взаємозв'язки між характеристиками.



Діаграма дерева рішень, яка ілюструє загальні взаємозв'язки між характеристиками. У цьому дереві: Коренева нода (Root Node) представляє початкове рішення.

Від кореневої ноди йдуть два шляхи на основі Перемінної 1: "Так" та "Ні".

Від кожної з цих нод можна далі розглядати Перемінну 2 та Перемінну 3.

Ця структура показує, як глобальні моделі можуть використовуватися для прийняття рішень, спираючись на різні характеристики.

Приклад глобальної пояснювальної моделі можна знайти в системах кредитного скорингу, де важливо зрозуміти, які фактори найчастіше впливають на позитивне чи негативне рішення про надання кредиту. Використання дерев рішень або більш комплексних моделей для побудови сурогатної моделі допомагає представити основні чинники, на яких базується скорингова система, що дозволяє кінцевим користувачам зрозуміти загальні принципи роботи алгоритму.

#### SHAP (SHapley Additive exPlanations)

SHAP є одним із найпоширеніших інструментів для пояснення рішень складних моделей, який використовує концепцію значень Шеплі з теорії ігор. Значення Шеплі показує внесок кожної характеристики до фінального результату, що дозволяє визначити, які фактори найбільше вплинули на конкретне передбачення.

Застосування SHAP у фінансовій сфері: якщо алгоритм ШІ оцінює ризик дефолту клієнта, то SHAP дозволяє аналітику побачити, які з характеристик клієнта (наприклад, рівень доходу, кредитна історія) вплинули на підвищення або зниження ризику. Такий підхід забезпечує високий рівень прозорості для аналітиків, допомагає пояснити модель регуляторним органам і підвищує довіру клієнтів.

#### Сурогатні моделі

Сурогатні моделі є спрощеними версіями основної моделі, які надають наближене пояснення логіки її роботи. Це можуть бути лінійні моделі, дерева рішень або інші інтерпретовані моделі, що підбираються так, щоб відображати основні закономірності в даних.

Сурогатні моделі часто використовуються, коли головна модель надто складна для прямої інтерпретації. Наприклад, у випадку з глибокою нейронною мережею, яка прогнозує поведінку користувачів у маркетингових кампаніях, сурогатна модель на основі дерева рішень дозволить маркетологам краще розуміти основні фактори, що впливають на інтерес клієнтів до продуктів, і використовувати ці знання для корекції стратегії.

#### **Роль візуалізації у покращенні інтерпретації**

Візуалізація даних є важливим інструментом для донесення складних рішень моделей до користувачів. Зокрема, популярними підходами є:

✓ Теплові карти: Застосовуються у випадках, коли важливо виявити ключові області на зображеннях, що впливають на рішення моделі. Наприклад, у медичній діагностиці теплові карти дозволяють лікарям побачити, які частини рентгенівського знімка призвели до діагнозу, що робить рішення ШІ більш прозорим і легким для оцінки.

✓ Діаграми важливості ознак: Дозволяють користувачам побачити, які характеристики найбільше вплинули на рішення моделі. Це особливо корисно в галузі фінансів, де можна показати, які фактори враховувалися під час оцінки кредитоспроможності клієнта.

✓ Візуалізація часових рядів: У задачах прогнозування на основі часових рядів, наприклад, у системах моніторингу промислових процесів, важливо відстежувати, як зміни у вхідних даних впливають на прогнозовані значення. Візуалізація допомагає користувачам побачити взаємозв'язок між минулими подіями та передбаченими результатами.

#### **Переваги та обмеження існуючих методів**

Незважаючи на важливі переваги методів інтерпретації, кожен із них має свої обмеження. Наприклад:

✓ LIME забезпечує локальне пояснення, але не може відображати загальні закономірності в усіх даних, що обмежує його застосування в задачах, де важлива глобальна інтерпретація.

✓ SHAP надає точні оцінки внеску кожної характеристики, але потребує значних обчислювальних ресурсів, що може бути проблемою для великих моделей або при обробці в реальному часі.

✓ Сурогатні моделі добре підходять для пояснення складних моделей, але не завжди точно передають логіку роботи оригінальної моделі, що може призводити до неправильних інтерпретацій.

#### **Розвиток та перспективи інтерпретованих ШІ**

На основі викладених методів можна зробити висновок, що забезпечення прозорості та зрозумілості рішень ШІ є складною, але надзвичайно важливою задачею. Подальші розробки можуть бути спрямовані на інтеграцію інтерпретованих алгоритмів в інтерфейси користувача, створення інтерактивних моделей, де користувачі можуть самостійно досліджувати вплив кожного параметра, а також на створення легких методів, що дозволяють використовувати пояснення в реальному часі.

## **ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ**

### **I ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ**

Подальші дослідження в сфері інтерпретованості ШІ можуть бути зосереджені на створенні інтерактивних систем, де кінцевий користувач зможе задавати питання щодо рішень ШІ та отримувати пояснення у зрозумілому форматі. Розвиток комбінованих підходів, які об'єднують кілька методів пояснення, може забезпечити кращу зрозумілість та адаптивність моделей.

Успішне впровадження пояснювальних моделей є важливим кроком для підвищення довіри та розуміння рішень ШІ користувачами. Поєднання методів локальної та глобальної інтерпретації, а також використання візуалізації, може суттєво підвищити прозорість алгоритмів ШІ. Подальші розробки в цій галузі сприятимуть створенню ефективних, адаптивних і зрозумілих систем, що робить ШІ більш доступним для кінцевих користувачів та сприяє його етичному використанню у критично важливих сферах.

#### **Література**

1. Лисенко О., Ковальчук П. Прозорість рішень штучного інтелекту / О. Лисенко, П. Ковальчук // Наукові записки. – 2023. – Т. 5. – № 1. – С. 23-35. DOI: 10.1234/sciencenotes.2023.001
2. Мельниченко С. Етичні аспекти штучного інтелекту: забезпечення прозорості / С. Мельниченко. – Київ: Видавництво КПІ, 2022. – 250 с.
3. Doshi-Velez, F., Kim, P. Towards a rigorous science of interpretable machine learning / F. Doshi-Velez, P. Kim // arXiv preprint arXiv:1702.08608. – 2017. – Режим доступу: <https://arxiv.org/abs/1702.08608>
4. Ribeiro, M. T., Singh, S., Guestrin, C. "Why should I trust you?" Explaining the predictions of any classifier / M. T. Ribeiro, S. Singh, C. Guestrin // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – С. 1135-1144. DOI: 10.1145/2939672.2939778
5. Mohseni, S., et al. Transparency and Explainability in Artificial Intelligence: A Survey / S. Mohseni, A. L. K. P. S. Andreev, A. A. et al. // IEEE Access. – 2021. – Т. 9. – С. 123456-123472. DOI: 10.1109/ACCESS.2021.3059234

6. Кучеренко О. Штучний інтелект для людини: пояснення та довіра / О. Кучеренко. – Одеса: ОНУ, 2023. – 180 с.

7. EU Guidelines on Trustworthy AI: [website]. Access mode: <https://ec.europa.eu/digital-strategy/our-policies/eu-guidelines-trustworthy-ai>

#### References

1. Lysenko O., Kovalchuk P. Transparency in Artificial Intelligence Decisions / O. Lysenko, P. Kovalchuk // Scientific Notes. – 2023. – Vol. 5. – No. 1. – P. 23-35. DOI: 10.1234/sciencenotes.2023.001

2. Melnychenko S. Ethical Aspects of Artificial Intelligence: Ensuring Transparency / S. Melnychenko. – Kyiv: KPI Publishing, 2022. – 250 p.

3. Doshi-Velez, F., Kim, P. Towards a Rigorous Science of Interpretable Machine Learning / F. Doshi-Velez, P. Kim // arXiv preprint arXiv:1702.08608. – 2017. – Access mode: <https://arxiv.org/abs/1702.08608>

4. Ribeiro, M. T., Singh, S., Guestrin, C. "Why Should I Trust You?" Explaining the Predictions of Any Classifier / M. T. Ribeiro, S. Singh, C. Guestrin // Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. – 2016. – P. 1135-1144. DOI: 10.1145/2939672.2939778

5. Mohseni, S., et al. Transparency and Explainability in Artificial Intelligence: A Survey / S. Mohseni, A. L. K. P. S. Andreev, A. A. et al. // IEEE Access. – 2021. – Vol. 9. – P. 123456-123472. DOI: 10.1109/ACCESS.2021.3059234

6. Kucherenko O. Artificial Intelligence for Humans: Explanation and Trust / O. Kucherenko. – Одеса: ONU Publishing, 2023. – 180 p.

7. EU Guidelines on Trustworthy AI: [website]. Access mode: <https://ec.europa.eu/digital-strategy/our-policies/eu-guidelines-trustworthy-ai>



<https://doi.org/10.31891/2219-9365-2024-79-28>

УДК 681.5:629.3

СПОРИЙ Олександр

Хмельницький національний університет

МАЙДАН Павло

Хмельницький національний університет

<https://orcid.org/0000-0003-3319-8730>

[maidanps@gmail.com](mailto:maidanps@gmail.com)

МАКАРИШКІН Денис

Хмельницький національний університет

<https://orcid.org/0000-0003-3447-811X>

[makaryshkinde@khnmu.edu.ua](mailto:makaryshkinde@khnmu.edu.ua)

СОКОЛАН Юлія

Хмельницький національний університет

<https://orcid.org/0000-0002-0273-5719>

[sokolan.julia@gmail.com](mailto:sokolan.julia@gmail.com)

## МОДЕЛЮВАННЯ ОСНОВНИХ ЕЛЕМЕНТІВ КОНСТРУКЦІЇ НА ПРИКЛАДІ РОБОМОБІЛЯ «JUNIOR»

Загально відомо, що використання симуляторів допомагає здешевити та прискорити розробку робото-технічних систем, є безпечнішим інструментом первинної верифікації ідей та алгоритмів, а також надасть можливість виконувати тестування нових концепцій та алгоритмів, незалежно від наявності необхідного обладнання у наукової групи. Моделювання може бути використане для всіх типів робототехнічних систем та завдань на початкових етапах розвитку проекту.

Моделювання робомобілю було проведене в середовищі ROS Gazebo. Підхід до моделювання був наступний: врахування фізичних характеристик робомобілю, врахування давачів та характеристик, використання CAD-моделей робомобілю для моделювання. При створенні моделі використовувалися середовища для програмування: ROS, Gazebo. Для того, щоб модель візуально відповідала реальному робомобілю, були використані частини для даних елементів та деяких давачів, САПР-модель є у відкритому доступі. Усі зчленування моделі були задані у пропорціях, що відповідають пропорціям та розмірам реального робомобілю. Щоб адаптувати інерційні характеристики робомобілю до моделі, було розраховано спрощений інерційний тензор на зчленування.

В роботі виконано моделювання роботи робомобіля для переміщення виробів, а саме на прикладі робомобіля «Junior» фірми AURORA, створена модель відповідає реальним розмірам, його габаритним характеристикам, наявності встановлених давачів. Необхідність розробки базового програмного забезпечення полягає в тому, що компанія-виробник надає мінімальний набір ROS-пакетів, які дають змогу лише отримувати дані із давачів і посилати команди на електродвигуни робомобіля.

Ключові слова: робомобіль, Junior, Ackermann drive, симулятор Gazebo, моделювання, давачі.

SPORYI Oleksandr, MAIDAN Pavlo, MAKARYSHKIN Denys, SOKOLAN Iuliia  
Khmelnitskyi National University

## MODELING OF THE MAIN STRUCTURAL ELEMENTS ON THE EXAMPLE OF THE ROBOT VEHICLE «JUNIOR»

It is well known that the use of simulators helps to reduce the cost and speed up the development of robotic systems, is a safer tool for the initial verification of ideas and algorithms, and will allow testing of new concepts and algorithms, regardless of the availability of the necessary equipment for the research team. Modeling can be used for all types of robotic systems and tasks in the initial stages of project development.

The robot vehicle was modeled in the ROS Gazebo environment. The approach to modeling was as follows: taking into account the physical characteristics of the robot vehicle, taking into account sensors and characteristics, using CAD models of the robot vehicle for modeling. When creating the model, the following programming environments were used: ROS, Gazebo. To make the model visually consistent with a real robot vehicle, parts for these elements and some sensors were used; the CAD model is publicly available. All the joints of the model were set in proportions that correspond to the proportions and dimensions of the real robot vehicle. To adapt the inertial characteristics of the robot vehicle to the model, a simplified inertial tensor for the joints was calculated.

In this work, we model the operation of a robot vehicle for moving products, namely, on the example of the AURORA Junior robot vehicle, the created model corresponds to the real size, its overall characteristics, and the presence of installed sensors. The need to develop basic software is due to the fact that the manufacturer provides a minimal set of ROS packages that only allow you to receive data from sensors and send commands to the robot's electric motors.

Keywords: robot vehicle, Junior, Ackermann drive, Gazebo simulator, modeling, sensors.

## ПОСТАНОВКА ПРОБЛЕМИ У ЗАГАЛЬНОМУ ВИГЛЯДІ ТА ЇЇ ЗВ'ЯЗОКІЗ ВАЖЛИВИМИ НАУКОВИМИ ЧИ ПРАКТИЧНИМИ ЗАВДАННЯМИ

В останні десятиліття широке використання симуляцій стало невіддільною частиною досліджень у галузі робототехніки [1-4]. Використання симуляторів допомагає здешевити та прискорити розробку робото-

технічних систем (РТС), є безпечнішим інструментом первинної верифікації ідей та алгоритмів, а також надає можливість виконувати тестування нових концепцій та алгоритмів, незалежно від наявності необхідного обладнання у наукової групи. Моделювання може використовуватись для всіх типів РТС та завдань на початкових етапах розвитку проекту.

Симулятори стали більш прогресивними інструментами, які надають можливість виконати відтворення складних середовищ та використовувати фізику, яку визначив сам розробник, що, в свою чергу, дає можливість створювати моделі РТС, які поведуться максимально близько до реальної поведінки робомобілів (РМ). Gazebo є одним із найпопулярніших 3D-симуляторів РМ, який успішно використовується для симуляції БПЛА [5] та безпілотного наземного робота (БНР) [6-9], та для проведення різноманітних експериментів, включно з тестуванням базових рухів РМ, плануванням траєкторії шляху та колаборативною взаємодією з іншими РМ [10], експериментами з маніпуляторами [11] та моделюванням сценаріїв пошуково-рятувальних операцій [12]. Інші спеціалізовані симулятори (наприклад, UWSim [13], V-REP [14]) дають змогу моделювати водне середовище для автономних підводних роботів.

### АНАЛІЗ ДОСЛІДЖЕНЬ ТА ПУБЛІКАЦІЙ

РМ «Junior» є розробкою фірми «AURORA ROBOTICS», загальний вигляд, якого представлено на рисунку 1. За своєю зовнішньою будовою РМ «Junior» є класичною машиною із Ackermann drive або системою приводів та рульовою рейкою для здійснення керування рухом машини [15]. Даний РМ було розроблено з метою навчання студентів принципів автономного руху БНР. Такий навчально-налагоджувальний комплекс є зручним та оптимальним рішенням для роботи студентів: невеликі габарити та загальна вага машини (відносно повнорозмірних машин, які використовуються для розробки автономного функціонування компанії Uber чи Google), відкрита програмна система на базі ОС Linux Ubuntu 16.04 LTS та ROS, а також відкрита апаратна система.

Відкрита апаратна система передбачає вільний фізичний доступ до бортових давачів РМ, влаштування електричного ланцюга, можливості зміни апаратного складу конструкції, вдосконалення та заміни будь-яких компонентів (як давачів, і базових елементів – акумулятора, коліс, мотор-колес тощо) [16].



Рис. 1. Загальний вигляд РМ «Junior» фірми AURORA ROBOTICS

З РМ передбачається використання спеціального полігону, побудованого на кшталт реальної проїжджої частини міста з відповідними елементами: дорожніми знаками, світлофорами, тунелем, дорожньою розміткою та, відповідно, пішохідним переходом. Завдяки всьому комплексу даного навчального комплексу студенти можуть розробляти алгоритми автономного руху РМ у штучному середовищі дорожнього оточення [16].

### ФОРМУЛЮВАННЯ ЦІЛЕЙ СТАТТІ

Метою роботи є: розроблення алгоритмів переміщення та автономної навігації для РМ «Junior», для чого побудовано модель у симуляторі Gazebo і проведено тестування алгоритмів керування.

### ВИКЛАД ОСНОВНОГО МАТЕРІАЛУ

Моделювання РМ було проведене в середовищі ROS Gazebo. Підхід до моделювання був наступний: врахування фізичних характеристик РМ, врахування давачів РМ та характеристик, використання CAD-моделей РМ для моделювання. При створенні моделі РМ використовувалися наступні середовища для програмування: ROS, Gazebo [17].

Основними елементами РМ «Junior», що представляють модель, були обрані наступні: зовнішній каркас РМ - base link, переднє праве колесо - right steer wheel, переднє ліве колесо - left steer wheel, права частина рульової рейки, частина майбутнього механізму системи підвіски РМ згідно Ackermann drive - right\_steer\_link, ліва частина рульової рейки, частина майбутнього механізму системи підвіски РМ згідно з Ackermann drive - drive\_right\_drive\_left\_steer\_link, заднє ліве колесо - left\_drive\_wheel.

Для того, щоб модель РМ візуально відповідала реальному РМ, були використані частини для даних елементів та деяких давачів, розглянутих нижче (САПР-модель є у відкритому доступі) (рис. 2).

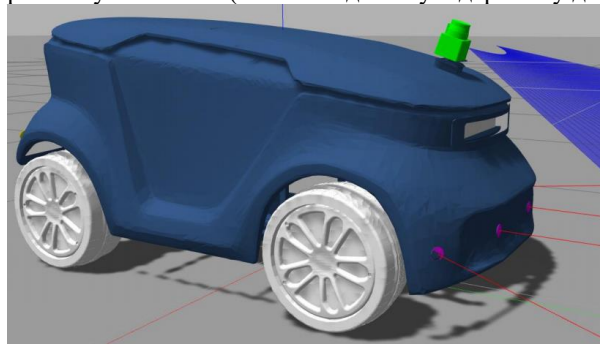


Рис. 2. Загальний вигляд моделі РМ «Junior»

Усі зчленування моделі були задані у пропорціях, що відповідають пропорціям та розмірам реального РМ. Щоб адаптувати інерційні характеристики РМ до моделі, було розраховано спрощений інерційний тензор на зчленування. У процесі було замінено складну структуру каркаса РМ (base link) суцільним кубоїдним об'єктом (solid cuboid); структури рульових коліс РМ (right\_steer\_wheel, left\_steer\_wheel), задніх коліс замінені на суцільний циліндр (solid cylinder) для обчислення наближених тензорів інерції зчленувань моделі. Колізії зчленувань були зазначені з урахуванням структури всіх зчленувань і в тому ж розмірі, що й зчленування.

```
#include "ros/ros.h"
#include "std_msgs/Float64.h"
#include <utility>
#include <cmath>
#include <sstream>
#include <string>
#include "ackermann_msgs/AckermannDriveStamped.h"
ros::Publisher left_pub;
ros::Publisher right_pub;
std_msgs::Float64 fi_r;
std_msgs::Float64 fi_l;
double const lateral_distance = 0.7;
double const front_distance = 0.54;
double const pi2 = 1.5708;
double radiusCalc (double angle)
{ return lateral_distance/tan(angle); }
double formulaCalc (double radius, bool sign)
{ if (sign)
  { return lateral_distance / (radius + (front_distance/2)); }
  else
  { return lateral_distance / (radius - (front_distance/2)); } }
void steeringCallback(const ackermann_msgs::AckermannDriveStamped::ConstPtr &msg)
{ double radius = radiusCalc(msg->drive.steering_angle);
  ROS_INFO("---Calculated radius : %.4f,", radius);
  if (msg->drive.steering_angle >= 0 && msg->drive.steering_angle <= 1.20)
  { fi_r.data = atan(formulaCalc(radius, false));
    fi_l.data = atan(formulaCalc(radius, true));
    ROS_INFO("---Calculated angles : left wheel %.4f, right wheel %.4f", fi_r.data, fi_l.data);
    left_pub.publish(fi_r)
    right_pub.publish(fi_l); }
  if (msg->drive.steering_angle <= 0 && msg->drive.steering_angle >= -1.20)
  { ROS_INFO("---Calculated angles : right wheel %.4f, left wheel %.4f", fi_r.data, fi_l.data);
    fi_r.data = atan(formulaCalc(radius, true));
    fi_l.data = atan(formulaCalc(radius, false));
    left_pub.publish(fi_l);
    right_pub.publish(fi_r); }
  if (msg->drive.steering_angle > 1.20)
  { double radius_exception(1.2);
    ROS_INFO("---Calculated radius exception : %.4f,", radius);
```

```
fi_r.data = atan(formulaCalc(radius_exception, false));
fi_l.data = atan(formulaCalc(radius_exception,true));
ROS_INFO("---Calculated angles : left wheel %.4f, right wheel %.4f", fi_r.data,fi_l.data);
left_pub.publish(fi_r);
right_pub.publish(fi_l); }
if (msg->drive.steering_angle < -1.2)
{ double radius_exception(-1.2);
ROS_INFO("---Calculated radius exception : %.4f,", radius);
fi_r.data = atan(formulaCalc(radius_exception, false));
fi_l.data = atan(formulaCalc(radius_exception,true));
ROS_INFO("---Calculated angles : left wheel %.4f, right wheel %.4f", fi_r.data,fi_l.data);
left_pub.publish(fi_r);
right_pub.publish(fi_l); }
}
int main(int argc, char **argv)
{ ros::init(argc, argv, "steer_remap");
ros::NodeHandle n;
std::string lw_topic, rw_topic, central_steer_topic;
central_steer_topic = "/junior_car/ackermann_cmd";
lw_topic = "/junior_car/SteerLeft_controller/command";
rw_topic = "/junior_car/SteerRight_controller/command";
if (n.getParam("left_wheel_steer_topic", lw_topic))
{ ROS_INFO("Got param: %s", lw_topic.c_str());}
if (n.getParam("right_wheel_steer_topic", rw_topic))
{ ROS_INFO("Got param: %s", rw_topic.c_str()); }
if (n.getParam("right_wheel_steer_topic", central_steer_topic))
{ ROS_INFO("Got param: %s", central_steer_topic.c_str()); }
left_pub = std::move(n.advertise<std_msgs::Float64>(lw_topic, 1024));
right_pub = std::move(n.advertise<std_msgs::Float64>(rw_topic, 1024));
ros::Subscriber sub = n.subscribe(central_steer_topic, 1024, steeringCallback);
ros::spin();
return 0;
}
```

Рис. 3. Приклад програми контролю моделі РМ «Junior»

Моделювання суглобів, моторів та трансмісія РМ.

URDF формат дозволяє описувати наступні типи суглобів: призматичний (prismatic) – ковзний суглоб, що зміщується вздовж вказаної осі та володіє обмеженим діапазоном руху, заданими верхньою та нижньою границями, обертальний без лімітів (continuous) – обертальний суглоб (або шарнірний), який обертається щодо вказаної осі та не має верхніх та нижніх границь, плаваючий (floating) – суглоб, що володіє рухом у всіх шести ступенях свободи, обертальний з лімітами (revolute) – обертальний суглоб (шарнірний), який обертається щодо заданої осі та має обмежений діапазон руху, заданий верхньою та нижньою границею і нерухомий (fixed) – фіксований суглоб, який насправді не є суглобом. Усі ступеня свободи даного типу суглоба заблоковані, він вимагає вказівки осі руху, границь. Даний тип суглобу використовується для жорсткого зчеплення зчленування РМ між собою, а також приєднання давачів РМ до його зчленування.

Обертальний тип суглоба без обмежень був використаний до суглобів передніх та задніх коліс (left\_steer\_wheel\_joint, left\_drive\_wheel\_joint, right\_steer\_wheel\_joint, right\_drive\_wheel\_joint). Обертальний тип суглоба із обмеженням був використано до кермових механізмів моделі (left\_steer\_joint and right\_steer\_joint), для подальшої розробки системи підвіски та керування моделлю згідно з принципом Ackermann drive [15].

Таким чином, для здійснення обчислень величини кута повороту передніх коліс ( $\alpha_1$  та  $\alpha_0$ ), необхідно було знати лінійні геометричні параметри РМ: відстань між передніми колесами ( $T = 0,54\text{ м}$ ) і бічну відстань між переднім та заднім колесом ( $L = 0,70\text{ м}$ ). Дані параметри були використані як константи при розрахунку величин кутів повороту коліс (рис. 4).

Так як на реальному РМ використовуються два мотори для виконання руху задніх коліс та один мотор для організації рульового механізму, були налаштовані наступні ROS-контролери для суглобів: SteerRight\_controller і SteerLeft\_controller - контролери положення (або ROS JointPositionController), EffortDriveRight\_Control - контролери зусиль.

Контролери положення відносяться до суглобів кермового механізму, контролери зусиль відносяться до суглобам задніх коліс (оскільки на реальному РМ моторами оснащені два задніх колеса). Рисунок 5 демонструє фінальну схему роботи контролерів середовища ROS.

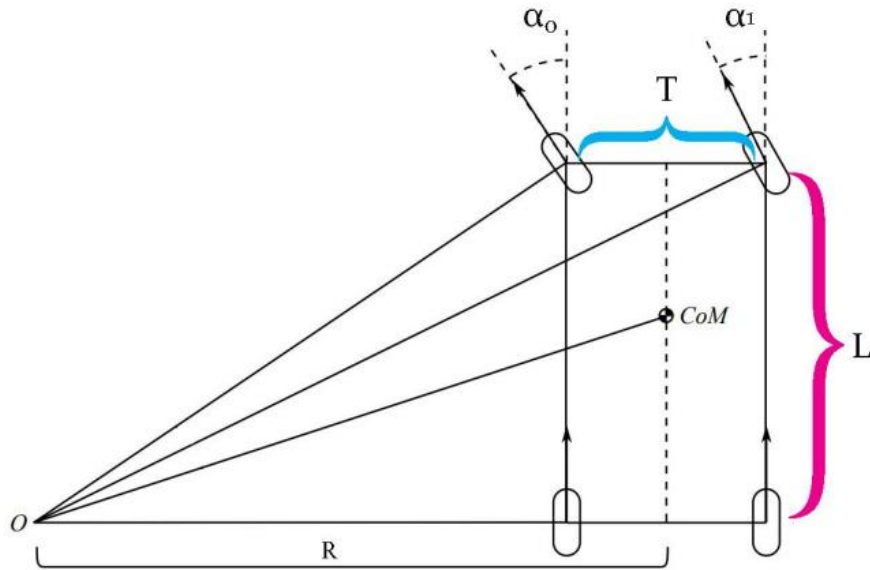


Рис. 4. Загальний вигляд геометрії Ackermann drive для ПМ [15]:

L - бічна відстань між серединами переднього та заднього коліс ПМ; R - відстань від поздовжньої осі ПМ до центру повороту; T - відстань між серединою передньої пари коліс

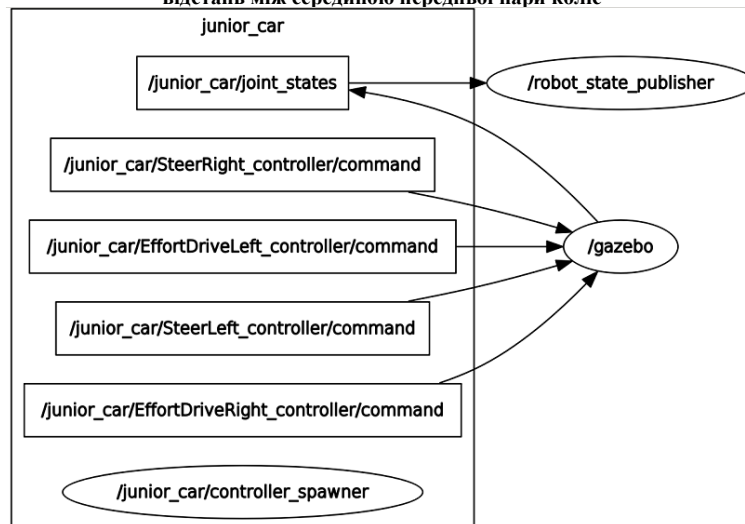


Рис. 5. Загальний вигляд візуалізації в rqt\_graph тем повідомлень моделі ПМ «Junior» (теми повідомлень давачів не відображені)

Модельювання давачів наведено табл. 1 та представляє список давачів ПМ «Junior» та їх імплементація в програмних середовищах ROS та Gazebo.

Таблиця 1.

Давачі ПМ та їх плагіни в середовищах ROS, Gazebo

Давачі ПМ та їх плагіни	Пакет в ROS	Пакет в Gazebo
Лідар	hokuyo_node	laser_controller
Microsoft Kinect	freenect_stack	camera_plugin
GPS	ur_nmea_driver	novatel_gps_sim
УЗ	ur_rangefinder_driver	sonar
Вимірювальний інерційний пристрій	myahrs_driver	imu_plugin

Моделі давачів були поміщені в тому ж положенні та орієнтації, як і на реальному ПМ, для отримання точних даних із давачів. Всі плагіни давачів були налаштовані згідно з технічним описом існуючих давачів, для виконання візуалізації давачів використовувалося ПЗ Rviz (входить у фреймворк ROS).

## ВИСНОВКИ З ДАНОГО ДОСЛІДЖЕННЯ І ПЕРСПЕКТИВИ ПОДАЛЬШИХ РОЗВІДОК У ДАНОМУ НАПРЯМІ

В роботі виконано моделювання роботи робомобіля для переміщення виробів, а саме на прикладі робомобіля «Junior» фірми AURORA, створена модель відповідає реальним розмірам, його габаритним характеристикам, наявності давачів.

Необхідність розробки базового програмного забезпечення полягає в тому, що компанія-виробник надає мінімальний набір ROS-пакетів, які дають змогу лише отримувати дані із давачів і посылати команди на електродвигуни робомобіля.

Далі для моделі робомобілю буде налаштовано систему навігації (локальний планувальник, глобальний планувальник, конфігурації costmap), проведено первинні дослідження в симуляції. Для апробації поставленого завдання буде також створено віртуальний світ з імітацією міської забудови і дев'ятьма різними кінцевими цілями. Запропоноване рішення може продемонструвати концепт автоматизованого завдання останнього кілометра і може бути розвинене в якості повноцінного проекту на реальному робомобілі «Junior».

### Література

1. Павленко І.І., Мажара В.А. Роботизовані технологічні комплекси: Навчальний посібник. – Кіровоград: КНТУ, 2010. – 392 с.
2. Робототехнічні системи: проектування і моделювання [Електронний ресурс]: навч. Посіб. для студ. спеціальності 126 «Інформаційні системи та технології» / М. М. Поліщук, М.М. Ткач; КПІ ім. Ігоря Сікорського. Електронні текстові дані (1 файл: 41,6 Мбайт). Київ: КПІ ім. Ігоря Сікорського, 2021. - 112 с.
3. Синтез робототехнічних систем в машинобудуванні: Підручник / Л. Є. Пелевін, К. І. Почка, О. М. Гаркавенко, Д. О. Міщук, І. В. Русан. – К.: ТОВ «НВП «Інтерсервіс»», 2016. – 258 с.
4. Основи мехатроніки [Електронний ресурс] : навчальний посібник для студентів спеціальності 141 «Електроенергетика, електротехніка та електромеханіка» / КПІ ім. Ігоря Сікорського ; уклад.: С. М. Пересада, М. В. Пушкар. – Електронні текстові дані (1 файл: 1,87 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2020. – 137 с.
5. Towards DJI Phantom 4 Realistic Simulation with Gimbal and RC Controller in ROS/Gazebo Environment // Sagitov A., Gerasimov Y. - 10th International Conference on Developments in eSystems Engineering (DeSE). 2017. pp. 262-266.
6. ROS-based SLAM for a Gazebo-simulated mobile robot in image-based 3D model of indoor environment // Afanasyev I., Sagitov A., Magid E. - International Conference on Advanced Concepts for Intelligent Vision Systems. 2015. pp. 273-283.
7. Modelling a crawler-type UGV for urban search and rescue in Gazebo environment // Sokolov M., Afanasyev I., Lavrenov R., Sagitov A., Sabirova L., Magid E. - Artificial Life and Robotics (ICAROB 2017), International Conference on. 2017. pp. 360-362.
8. Robot simulation physics validation // Pepper C., Balakirsky S., Scrapper C. - Proceedings of the 2007 Workshop on Performance Metrics for Intelligent Systems. 2007. pp. 97-104.
9. Fast simulation of vehicles with nondeformable tracks // Pecka M., Zimmermann K., Svoboda T. - 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2017. pp. 6414-6419.
10. A simulation system based on ros and gazebo for robocup middle size league // Yao W., Dai W., Xiao J., Lu H., Zheng Z.- IEEE international conference on robotics and biomimetics (ROBIO). 2015. pp. 54-59.
11. Manipulation task simulation using ROS and Gazebo // Qian W., Xia Z., Xiong J., Gan Y., Guo Y., Weng S., Deng H., Hu Y., Zhang J. - IEEE International Conference on Robotics and Biomimetics (ROBIO 2014). 2014. pp. 2594-2598.
12. Robocuprescue 2014-robot league team hector darmstadt (Germany) // Kohlbrecher S., Meyer J., Graber T., Petersen K., Von Stryk O., Klingauf U. RoboCupRescue 2014.
13. A dynamic simulator for underwater vehicle-manipulators // Kermorgant O. - International Conference on Simulation, Modeling, and Programming for Autonomous Robots. 2014. pp. 25-36.
14. V-REP: A versatile and scalable robot simulation framework // Rohmer E., Singh S.P., Freese M. - 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2013. pp. 1321 -1326.
15. Ackermann steering geometry [Електронний ресурс]. – Режим доступу: [https://en.wikipedia.org/wiki/Ackermann\\_steering\\_geometry](https://en.wikipedia.org/wiki/Ackermann_steering_geometry)
16. Aurora Driver for Freight —coming to Texas in 2024 [Електронний ресурс]. – Режим доступу: <https://aurora.tech/>
17. Simulate before you build [Електронний ресурс]. – Режим доступу: <https://gazebosim.org/home>

### References

1. Pavlenko I.I., Mazhara V.A. Robotyzovani tekhnolohichni komplekxy: Navchalnyi posibnyk. – Kirovohrad: KNTU, 2010. – 392 s.
2. Robototekhnichni systemy: proektuvannia i modeliuwannia [Elektronnyi resurs]: navch. Posib. dlia stud. spetsialnosti 126 «Informatyini systemy ta tekhnolohii» / M. M. Polishchuk, M.M. Tkach; KPI im. Ihoria Sikorskoho. Elektronni tekstovi dani (1 fail: 41,6 Mbait). Kyiv: KPI im. Ihoria Sikorskoho, 2021. - 112 s.

3. Syntez robototekhnichnykh system v mashynobuduvanni: Pidruchnyk / L. Ye. Pelevin, K. I. Pochka, O. M. Harkavenko, D. O. Mishchuk, I. V. Rusan. – K.:TOV «NVP «Interservis»», 2016. – 258 s.
4. Osnovy mekhatroniky [Elektronnyi resurs] : navchalnyi posibnyk dlia studentiv spetsialnosti 141 «Elektroenerhetyka, elektrotehnika ta elektromekhanika» / KPI im. Ihoria Sikorskoho ; ukklad.: S. M. Peresada, M. V. Pushkar. – Elektronni tekstovi dani (1 fail: 1,87 Mbait). – Kyiv : KPI im. Ihoria Sikorskoho, 2020. – 137 s.
5. Towards DJI Phantom 4 Realistic Simulation with Gimbal and RC Controller in ROS/Gazebo Environment // Sagitov A., Gerasimov Y. - 10th International Conference on Developments in eSystems Engineering (DeSE). 2017. pp. 262-266.
6. ROS-based SLAM for a Gazebo-simulated mobile robot in image-based 3D model of indoor environment // Afanasyev I., Sagitov A., Magid E. - International Conference on Advanced Concepts for Intelligent Vision Systems. 2015. pp. 273-283.
7. Modelling a crawler-type UGV for urban search and rescue in Gazebo environment // Sokolov M., Afanasyev I., Lavrenov R., Sagitov A., Sabirova L., Magid E. - Artificial Life and Robotics (ICAROB 2017), International Conference on. 2017. pp. 360-362.
8. Robot simulation physics validation // Pepper C., Balakirsky S., Scrapper C. - Proceedings of the 2007 Workshop on Performance Metrics for Intelligent Systems. 2007. pp. 97-104.
9. Fast simulation of vehicles with nondeformable tracks // Pecka M., Zimmermann K., Svoboda T. - 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2017. pp. 6414-6419.
10. A simulation system based on ros and gazebo for robocup middle size league // Yao W., Dai W., Xiao J., Lu H., Zheng Z.- IEEE international conference on robotics and biomimetics (ROBIO). 2015. pp. 54-59.
11. Manipulation task simulation using ROS and Gazebo // Qian W., Xia Z., Xiong J., Gan Y., Guo Y., Weng S., Deng H., Hu Y., Zhang J. - IEEE International Conference on Robotics and Biomimetics (ROBIO 2014). 2014. pp. 2594-2598.
12. Robocuprescue 2014-robot league team hector darmstadt (Germany) // Kohlbrecher S., Meyer J., Graber T., Petersen K., Von Stryk O., Klingauf U. RoboCupRescue 2014.
13. A dynamic simulator for underwater vehicle-manipulators // Kermorgant O. - International Conference on Simulation, Modeling, and Programming for Autonomous Robots. 2014. pp. 25-36.
14. V-REP: A versatile and scalable robot simulation framework // Rohmer E., Singh S.P., Freese M. - 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2013. pp. 1321 -1326.
15. Ackermann steering geometry [Електронний ресурс]. – Режим доступу: [https://en.wikipedia.org/wiki/Ackermann\\_steering\\_geometry](https://en.wikipedia.org/wiki/Ackermann_steering_geometry)
16. Aurora Driver for Freight —coming to Texas in 2024 [Електронний ресурс]. – Режим доступу: <https://aurora.tech/>
17. Simulate before you build [Електронний ресурс]. – Режим доступу: <https://gazebosim.org/home>

<https://doi.org/10.31891/2219-9365-2024-79-29>

<https://doi.org/10.31891/2219-9365-2024-79-30>

<https://doi.org/10.31891/2219-9365-2024-79-31>

<https://doi.org/10.31891/2219-9365-2024-79-32>

<https://doi.org/10.31891/2219-9365-2024-79-33>



**За зміст повідомлень редакція відповідальності не несе**



---

Підп. до друку 28.03.2024. Ум. друк. арк. – 12,88. Обл.-вид. арк. – 11,36.  
Формат 30×42/4, папір офсетний. Друк різнографією.  
Наклад 100, зам. № 018/23

---

Тиражування здійснено з оригінал-макета редакційно-видавничим відділом  
Хмельницького національного університету.  
29016, м. Хмельницький, вул. Інститутська, 7/1, тел. (0382) 77-33-63.  
Свідоцтво про внесення в Державний реєстр, серія ДК № 4489 від 18.02.2013 р.