

*МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЛУЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ*

**КОМП'ЮТЕРНО-ІНТЕГРОВАНІ
ТЕХНОЛОГІЇ:
ОСВІТА, НАУКА, ВИРОБНИЦТВО**

НАУКОВИЙ
ЖУРНАЛ



Головний редактор – професор, д.т.н., Гордєєв О.О.

№55 2024

м. Луцьк

РЕДАКЦІЙНА КОЛЕГІЯ:

Головний редактор:	
професор, д.т.н. Гордєєв О.О.	(м. Луцьк)
Відповідальний секретар:	
доц., к.т.н. Христинець Н.А.	(м. Луцьк)
Члени редакційної колегії:	
проф., д.т.н. Андрущак І.Є.	(м. Луцьк)
проф., д.т.н. Згуровський М.З	(м. Київ)
Affiliate full professor, Avtandil Gagnidze	(Грузія, м. Тбілісі)
д.т.н., доц. Зеленський К.Х.	(м. Київ)
доц., к.т.н. Суринович О.М.	(м. Луцьк)
Affiliate full professor, Iavich Maksim	(Грузія, м. Тбілісі)
проф., д.т.н. Турбал Ю.В.	(м. Рівне)
доц., к.ф.-м.н. Рибицька О.М.	(м. Львів)
PhD. Milosz Marek	(Польща, м. Люблін)
проф., д.т.н. Мельник А.О.	(м. Львів)
проф., д.т.н. Мороз Б.І.	(м. Дніпро)
проф., д.т.н. Тарасенко В.П.	(м. Київ)
проф, PhD. Alison McMillan	(Великобританія, м. Рексем)
проф., д.т.н. Касянчук М.М.	(м. Тернопіль)
проф., д.т.н. Фауре Е.В.	(м. Черкаси)
проф., д.т.н. Олійников Р.В.	(м. Харків)
проф, д.пед.н. Черняшук Н.Л.	(м. Луцьк)
доц., к.т.н. Назаревич О.Б.	(м. Тернопіль)
PhD. Karim Elish	(США, м. Лейкленд)
PhD. Zbigniew Omiotek	(Польща, м. Люблін)
PhD. Dagmar Čagaňová	(Словачина, м. Братишава)
PhD. Paweł Komada	(Польща, м. Люблін)
PhD. José Machado	(Португалія, м. Гімарайш)
проф., д.т.н. Сайко В.Г.	(м. Київ)
доц., к.т.н. Приступа С.О	(м. Луцьк)
PhD. Anna Maria Saniuk	(Польща, м. Зелена Гура)
доц., к.т.н. Ткачук А.А. (заступник головного редактора)	(м.Луцьк)

Адреса редколегії:

Луцький національний технічний університет,
кафедра комп'ютерної інженерії та кібербезпеки
вул. Львівська 75, ауд.141
м.Луцьк, 43018
тел. (0332) 74-61-15
E-mail: cit@lntu.edu.ua,
сайт журналу: cit-journal.com.ua

КОМП'ЮТЕРНО-ІНТЕГРОВАНІ
ТЕХНОЛОГІЇ:
ОСВІТА, НАУКА, ВИРОБНИЦТВО

№55 2024 р.

Зареєстровано Національною радою України з питань телебачення і радіомовлення, як суб'єкт у сфері друкованих медіа (рішення №40 від 11.01.2024 р., ідентифікатор медіа R30-02456)

Рекомендовано до друку Науково-технічною радою Луцького національного технічного університету (протокол №9 засідання від 29.05.2024р.)

Рішенням МОН України наказом №515 від 16.05.2016р, журнал включено в перелік наукових фахових видань

Видання індексується у наукометричних та реферативних баз:

Open Academic Journals Index
Academic Resource Index ResearchBib

Rootindexing
Information Matrix for the Analysis of Journals
Ulrichsweb.

ISSN 2524-0560 (Online)

ISSN 2524-0552 (Print)

ЗМІСТ

АВТОМАТИКА ТА УПРАВЛІННЯ	
Альошин С.В. Реалізація розрахунку навчального навантаження в розрізі кількості студентів на навчальну дисципліну мовою SQL	5
Дубук В.І., Ковівчак Я.В., Свистун Б.І. Розробка автоматизованої системи замовлень мотовелозапчастин	15
ІНФОРМАТИКА ТА ОБЧИСЛЮВАЛЬНА ТЕХНІКА	
Турчин О.Б. Впровадження машинного навчання для підвищення точності прогнозування режимів роботи глибоководних насосних станцій	24
Анісімов В. Г., Кунанець Н.Е. Перехід від монолітної до мікросервісної архітектури: методологія та досвід впровадження	30
Багнюк Н.В., Христинець Н.А., Биков С.О. Розробка мультимедійних 3D моделей для симуляції технологічного процесу	42
Виннишин О.Я. Комплексна методика оптимізації програмного коду VHDL при проектуванні програмованої інтегральної схеми	48
Волинець Л.В., Гарматюк Н.А., Дерев'янка В.С., Дуда О.М. Крамар Т.О., Скалецький П.О. Формування концепту музейного мобільного застосування з елементами доповненої реальності	55
Горгун Р.О., Салтовський Б.Г. Інтерактивна карта тривоги України.	66
Демченко А.В., Федосова І.В. Дослідження інформаційно-комунікаційних платформ для дистанційного навчання	71
Дідус А. В., Терейковський І. А. Формалізація процесу розпізнавання ключових слів у голосовому сигналі.	78
Дмитренко Т.В. Застосування БПЛА з технологією вейвлет-перетворення у військових операціях: розвідка, виявлення цілей та навігація.	87
Добришин Ю.Є. Формалізація процесу проектування систем підтримки прийняття рішень щодо підвищення рівня захисту програмного забезпечення	93
Журавська І. М., Фісун М.Т. Ризики інформаційних витоків з мобільних пристроїв.	100
Коляда К.В., Рожко Д.В. Використання метаевристичних алгоритмів для розв'язку задачі комівояжера	107
Коровій О.С., Терейковський І.А. Концептуальна модель процесу визначення емоційної тональності тексту.	115
Лукашук М.М. Стандартизація веб-доступності інформаційно-комунікаційних технологій як спосіб залучення більшого числа користувачів	124
Маруняк С.Т. Виявлення та пом'якшення вразливостей безпеки в протоколах динамічної маршрутизації: поточні виклики та рішення	130
Марченко О.О. Метод розпаралелення пошуку по дереву методом MCTS.	137
Мельник В.М., Багнюк Н.В., Ройко О.Ю., Бортник К.Я., Кізім С.О. Вплив продуктивності сокетів на інтенсивність обробки даних на кластері віртуальних машин в гетерогенних умовах	143
Мороз Б.І., Круглик А.С. Концептуальна модель обробки інформаційних потоків в системі безпілотних літальних апаратів	155

Морохович В.С., Лях І.М., Хом'як М.І., Морохович Б.В. Прогнозування пасажирів, які вижили під час катастрофи «Титаніка», за допомогою дерева прийняття рішень.	161
Нижник О.О., Яшинський Л.В., Федосов С.А. Принцип роботи і перспективи використання голографічних дисплеїв	167
Пригода А.Я. Оцінка ефективності проєкту розробки та впровадження sgm-системи на основі мікросервісної архітектури	172
Решетняк В.В., Фауре Е.В. Відслідковування погляду як інструмент дослідження поведінки користувача	181
Розломій І.О., Симолюк В.П., Науменко С.В., Михайловський П.В. Модель безпеки взаємопов'язаних обчислювальних пристроїв на основі полегшеної схеми шифрування для IoT.	191
Седінкін О.А., Деркач М.В., Скарга-Бандурова І.С., Матюк Д.С. Система для відстеження руху очей на основі машинного навчання.	199
Філь Н.Ю. Моделі вибору датчика температури та вологості для модуля системи екологічного моніторингу	206
Хамбір В.Р. Автоматизація процесів тестування мобільних застосунків	213
Христинець Н.А., Вікторова Я.А., Жуковська С.Ю., Радіон Ю.А. Розробка файлового менеджера ОС	225
Чибіряк Я.І., Захарова А.М., Лавров Є.А., Сірик О.Є., Баранова І.В., Татарчук Р.М. Комп'ютерне моделювання для медицини катастроф. Підхід до автоматизованого аналізу рішень з організації мобільних медичних шпиталів.	232
Шваліковський Д.М. Дві задачі теплопровідності у двокомпонентних системах.	243
Шенгелія В.О., Жданова О.Г. Методи складання маршрутів польоту для БПЛА з пересувним центром обслуговування	251
Якобчук Б.А., Христинець Н.А. Створення простого голосового помічника у додатку для платформи Android.	260
ТЕЛЕКОМУНІКАЦІЇ ТА РАДІОТЕХНІКА	
Баховський П.Ф., Євсюк М.М. Аналіз енергетичного балансу радіоліній під час обміну даними в системах SAE / EPS.	265
Васильківський М.В., Коломієць А.А., Грабчак Н.В., Грицаюк Д.Ю., Костянін В.Ю. Програмні засоби Python моделювання телекомунікаційних пристроїв.	270
Лишук В.В., Мороз С.А., Заблоцький В.Ю., Євсюк М.М., Гриценюк В.В. Алгоритми керування автономними інверторами напруги перетворювачів частоти	279
Мадінов М.Л. Оптичні лінії зв'язку	286
УПРАВЛІННЯ ПРОЕКТАМИ	
Дейнеко Л.М., Кунанець Н.Е. Роль UML-діаграм у плануванні проєкту інформаційної системи на прикладі системи піших туристичних маршрутів.	293
Садовий Я.С. Порівняльний аналіз розподілених суфіксних дерев і традиційних методів управління даними	301
Димова Г.О., Ларченко О.В. Використання задачі про максимальний потік для бізнес-процесів.	309
Ларченко О. Теоретичні аспекти конфіденційності даних в інфокомунікаційних мережах.	314

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-01>

УДК 004.023

Альошин Сергій Вікторович, старший викладач.

<https://orcid.org/0000-0003-4996-9114>

ДВНЗ «Приазовський державний технічний університет», м. Дніпро/Маріуполь, Україна

РЕАЛІЗАЦІЯ РОЗРАХУНКУ НАВЧАЛЬНОГО НАВАНТАЖЕННЯ В РОЗРІЗІ КІЛЬКОСТІ СТУДЕНТІВ НА НАВЧАЛЬНУ ДИСЦИПЛІНУ МОВОЮ SQL

Альошин С.В. Реалізація розрахунку навчального навантаження в розрізі кількості студентів на навчальну дисципліну мовою SQL. В роботі пропонується математична модель та її реалізація щодо розрахунку штатного розкладу для кафедр закладу вищої освіти, що заснована на обсязі навчального навантаження та на кількості здобувачів закладу освіти. В моделі враховується, що здобувачі розподілені по групах закладу, здобувають вищу освіту на певних рівнях, опанують освітні програми за обов'язковими та вибірковими компонентами. Наведено формалізований опис моделі для розрахунку звітної показника на підставі даних зі сховища інформаційної системи, що використовується в закладі. Приведені основні форми запитів мовою SQL для виконання розрахунку показника та підсумовування за різними рівнями угруповань. Запропонований механізм розрахунку легко впроваджується в інформаційну систему, що вже використовується закладом для автоматизації організації освітнього процесу.

Ключові слова: навчальне навантаження, автоматизована інформаційна система, база даних, звіт, запит мовою SQL.

Aloshyn S.V. Implementation of the educational load calculation in terms of the student count per academic discipline by the SQL language. The work proposes a mathematical model and its implementation for calculating the staffing schedule for departments of a higher education institution, which is based on the volume of the educational load and the student count of the educational institution. The model considers that the students are divided into groups of the institution, acquire higher education at certain levels, take possession of educational programs according to mandatory and optional components. A formalized description of the model for calculating the reporting indicator based on data from the storage of the information system used in the institution is given. The main forms of queries in the SQL language for calculating the indicator and summarizing at different levels of grouping are given. The proposed calculation mechanism is easily implemented in the information system already used by the institution to automate of the educational process management.

Keywords: educational load, automated information system, database, report, SQL query.

Постановка проблеми.

Майже всі державні заклади вищої освіти в даний час стикаються із серйозними фінансовими проблемами, основною причиною яких стало вагоме зменшення контингенту здобувачів вищої освіти на всіх рівнях і формах здобуття. Розділ XII Закону України «Про вищу освіту» [1] регламентує фінансово-економічні відносини у сфері вищої освіти, а саме:

– фінансування державних закладів вищої освіти здійснюється за рахунок коштів державного бюджету на умовах державного замовлення на оплату послуг з підготовки фахівців, наукових і науково-педагогічних кадрів та за рахунок інших джерел, не заборонених законодавством;

– показники державного замовлення на підготовку фахівців з вищою освітою формуються за рівнями вищої освіти та спеціальностями з урахуванням середньострокового прогнозу потреби у фахівцях на ринку праці;

– розміщення державного замовлення здійснюється на конкурсних засадах на принципах добросовісної конкуренції, відкритості та прозорості, рівноправності, об'єктивного та неупередженого оцінювання пропозицій учасників конкурсу.

Зменшення державного замовлення та обсягів фінансування вищої освіти змушує заклади до більш ефективного розподілу своїх бюджетів та пошуку інших джерел фінансування. Запровадження Кабінетом Міністрів України формульного розподілу видатків державного бюджету між закладами вищої освіти на основі показників їх діяльності [2] призвело до використання показника чисельності здобувачів освіти (контингенту студентів) під час планування розподілу видатків бюджету самого закладу вищої освіти.

До того ж, заклад вищої освіти повинен самостійно планувати та розподіляти отримані кошти із будь-яких джерел фінансування внаслідок наданої Законом України автономії в цьому питанні.

Оскільки головною статтею видатків бюджету більшості закладів вищої освіти є фонд заробітної плати науково-педагогічних працівників, пропонується дослідити механізм розрахунку штатного розкладу для кафедр закладу вищої освіти, що заснований на обсязі навчального навантаження та на кількості здобувачів закладу освіти, які розподілені по групах закладу, здобувають вищу освіту на певних рівнях, опанують освітні програми за обов'язковими та вибірковими компонентами.

Метою роботи є розробка математичної моделі для розрахунку обсягу навчального навантаження в розрізі кількості студентів на навчальну дисципліну та реалізація її за допомогою запитів мовою SQL для подальшої побудови звітів у інформаційній системі закладу вищої освіти.

Аналіз останніх досліджень та публікацій.

Використання сучасних інформаційних систем для управління бізнес-процесами в закладі вищої освіти спрямовано на покращення та пришвидшення роботи підрозділів, що направлена на планування та виконання освітнього процесу. В роботах [3-4] автори розглядають шляхи застосування електронного документообігу як в середині, так і між кафедрами закладу вищої освіти, аналізують та описують процеси електронного документообігу для різних видів документації. В роботі [5] автор вважає, що найважливішим фактором використання інформаційної системи в закладі є можливість автоматичних обрахунків документів, які містять навантаження та обсяг годин по всіх дисциплінах.

В роботах [6-8] автори висвітлюють проблеми проектування та впровадження інформаційно-аналітичних систем, що поєднують в собі нормативні, методичні, освітні та інформаційні ресурси. Однією із задач, для рішення якої з давнього часу створюються автоматизовані інформаційні системи, є задача планування навчального навантаження закладу освіти на наступний навчальний рік. Аналіз праць цих та інших науковців засвідчує, що попри накопичення певного досвіду впровадження інформаційно-аналітичних систем управління в освітній процес закладів вищої освіти, залишається багато аспектів, що потребують подальшого вивчення з точки зору їх особливостей та функціональних можливостей.

Сучасні інформаційні системи для автоматизації управління закладом освіти, що присутні на ринку програмного забезпечення, наприклад, [9, 10], надають багато функцій для автоматизації бізнес-процесів закладу вищої освіти. Так, автоматизована система «Деканат», що входить до АСУ «ВНЗ» [10], надає багато функціональних модулів для збору, накопичення та обробки інформації різними структурними підрозділами закладу освіти, а саме, деканати, навчальний відділ, відділ кадрів, тощо. Багато функціоналу надається саме плануванню навчального навантаження для кафедр закладу вищої освіти: закріплення контингенту для навантаження по робочим навчальним планам; розрахунок навчального навантаження кафедри; агрегація та розподілення навантаження між викладачами. Архітектура системи дозволяє закладу освіти не тільки використовувати створені розробниками звітні форми, а й створювати власноруч будь-які звіти з параметрами користувача на основі тієї інформації, що зберігається в базі даних системи, за допомогою мови SQL.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.

Для розрахунку штатного розкладу по кожній кафедрі закладу освіти раніше використовувався тільки сумарний обсяг годин навчального навантаження, що був виділений під час планування навантаження на наступний навчальний рік, та норма Закону України «Про вищу освіту» [1]: «Максимальне навчальне навантаження на одну ставку науково-педагогічного працівника не може перевищувати 600 годин на навчальний рік». В такий спосіб було достатньо розрахувати сумарний обсяг годин навчального навантаження для кафедри закладу освіти за формулою:

$$V_C = \sum_w T_w,$$

де V_C – обсяг навчального навантаження кафедри на наступний навчальний рік в годинах; T_w – обсяг навчального навантаження в годинах за w -ю дисципліною, що буде викладатись кафедрою в наступному навчальному році.

Але в сучасних умовах ведення господарської діяльності, організації освітнього процесу та державного фінансування закладів освіти використовуються інші підходи до розрахунку штатного розкладу. Отже тепер весь обсяг навчального навантаження закладу вищої освіти на навчальний рік необхідно зв'язати з кількістю здобувачів закладу, наявних в той самий навчальний рік, тобто,

$$S = \sum_G \sum_w K_G \cdot T_{w,G}, \quad (1)$$

де S – кількість здобувачів закладу вищої освіти; $T_{w,G}$ – обсяг навчального навантаження в годинах за w -ю дисципліною, що буде викладатись в навчальному році для групи G ; K_G – коефіцієнт приведення навчального навантаження до кількості студентів для групи G , причому він буде різним для кожної групи закладу освіти.

Коефіцієнт приведення навчального навантаження до кількості студентів показує частку студентів навчальної групи на одну годину навчального навантаження групи, тому розрахунок коефіцієнту приведення K_G виконується за такою формулою:

$$K_G = \frac{S_G}{T_G}, \quad (2)$$

де S_G – кількість студентів групи G закладу освіти; T_G – сумарне навчальне навантаження на навчальний рік для групи G . Причому, T_G розраховується, як сума навчального навантаження за обов'язковими (нормативними) дисциплінами для групи та сума навчального навантаження за дисциплінами вільного вибору студентів з урахуванням кількості студентів, які обрали для вивчення дисципліну. Тобто, розрахунок сумарного навчального навантаження за групою студентів можна виразити такою формулою:

$$T_G = \sum_w \frac{S_{G_w}}{S_G} T_{w,G}, \quad (3)$$

де T_G – сумарне навчальне навантаження на навчальний рік для групи G , S_G – кількість студентів групи G закладу освіти; S_{G_w} – кількість студентів групи G , які обрали дисципліну w ; $T_{w,G}$ – обсяг навчального навантаження в годинах за w -ю дисципліною для групи G .

Таким чином, для підрахунку обсягу навчального навантаження та його приведення до кількості студентів групи чи закладу освіти необхідно визначити частку студентів групи за кожною дисципліною, що викладається для студентів групи згідно робочого навчального плану. Для визначення частки студентів групи за дисципліною, виходячи з формул (1) та (2), можна записати такий вираз:

$$t_{w,G} = \frac{S_G \cdot T_{w,G}}{T_G}, \quad (4)$$

де $t_{w,G}$ – частка студентів групи за дисципліною w , що викладається студентам групи G , S_G – кількість студентів групи G закладу освіти; $T_{w,G}$ – обсяг навчального навантаження в годинах за w -ю дисципліною для групи G , T_G – сумарне навчальне навантаження на навчальний рік для групи G , розраховане за формулою (3). Причому, для значень частки студентів за дисциплінами виконуються наступні правила:

- сума всіх $t_{w,G}$ по групі G дорівнює кількості студентів групи G ;
- сума всіх $t_{w,G}$ дорівнює кількості студентів всього закладу освіти (виходячи з формули (1)).

Оскільки будь-яка автоматизована інформаційна система, що використовується в закладі освіти для розрахунку навчального навантаження на наступний навчальний рік, використовує сховище даних у вигляді набору реляційних таблиць, розглянемо походження вихідних даних для їх використання в наведеній моделі, використавши теорію множин.

Навчальне навантаження для всього закладу освіти N можна описати множиною, елементи якої складаються з навчальних навантажень за кожною навчальною дисципліною в навчальному році:

$$N = \{ND_1, ND_2, ND_3, \dots, ND_k\},$$

де ND_i – навчальне навантаження за i -ю дисципліною закладу освіти в навчальному році, $i = \overline{1, k}$. Навчальне навантаження за дисципліною описується шістькою значень, що його характеризують:

$$ND = \langle D, G, S_0, S, C, T \rangle \quad (5)$$

де D – дисципліна, що викладається; G – навчальна група, для якої викладається дисципліна; S_0 – загальна кількість студентів групи, S – кількість студентів групи, для яких викладається дисципліна; C – кафедра, що викладає дисципліну в поточному або наступному навчальному році; T – кількість годин навчального навантаження за дисципліною.

Навчальні дисципліни, що викладаються студентам навчальної групи в навчальному році, формуються за робочим навчальним планом групи. Отже навчальне навантаження для навчальної групи за робочим навчальним планом можна описати множиною, елементи якої складаються із п'ятірки значень:

$$NG = \{ \dots, \langle D, S_0, S, C, T \rangle, \dots \}$$

Кожна кафедра закладу освіти отримує під час планування та виконання навчальне навантаження на навчальний рік у вигляді множини, елементи якої складаються із п'ятірки значень:

$$NC = \{ \dots, \langle D, G, S_0, S, T \rangle, \dots \}$$

Формування множини N , що містить інформацію про навчальне навантаження закладу освіти на навчальний рік, може виконуватись в будь-який спосіб в залежності від рівня використання інформаційних технологій в закладі освіти, наявності автоматизованих систем розрахунку навчального навантаження, тощо. В будь-якому випадку множини N можна представити у вигляді двомірної таблиці (див. рис. 1) як результат виконання вибірки з реляційних таблиць інформаційної системи закладу освіти, де кожен запис буде містити інформацію про навчальне навантаження за однією навчальною дисципліною відповідно до формули (5).

D_1	G_1	S_{1-0}	S_{1-1}	C_1	T_1
D_2	G_1	S_{1-0}	S_{1-2}	C_2	T_2
D_3	G_1	S_{1-0}	S_{1-3}	C_3	T_3
...					
D_4	G_2	S_{2-0}	S_{2-1}	C_1	T_{11}
D_3	G_2	S_{2-0}	S_{2-2}	C_4	T_{12}
D_2	G_2	S_{2-0}	S_{2-3}	C_2	T_{13}
...					
D_n	G_r	S_{j-0}	S_{j-m}	C_p	T_w

Рис. 1. Табличне представлення навчального навантаження

Множина N є тою самою первинною двомірною таблицею, що містить вихідні дані для виконання розрахунку навчального навантаження в розрізі кількості студентів на навчальну дисципліну.

Для реалізації описаної моделі розрахунку навчального навантаження в розрізі кількості студентів необхідно інтегрувати розрахунок частки студентів групи по кожній дисципліні в автоматизовану інформаційну систему, яка використовується в закладі освіти для керування навчальним процесом та допомагає виконувати розрахунок навчального навантаження. Інтеграція можлива за двома напрямками: побудова окремої підсистеми з можливістю розрахунку та перегляду показників навчального навантаження або створення звітів з таблиць БД із можливістю фільтрувати дані за допомогою параметрів. Для отримання звіту із системи, що буде містити потрібні показники навчального навантаження, достатньо побудувати запит для отримання даних з реляційної бази даних, де зберігаються параметри навчального навантаження та розрахований його обсяг за видами навчальних занять.

Розглянемо побудову такого запиту на мові SQL, а саме її реалізацію в системі управління базами даних Oracle – PL/SQL. По-перше, необхідно створити віртуальну таблицю (або представлення) за формою, що наведена на рис. 1, доповнивши її додатковими параметрами.

Представлення можна використовувати в інших запитах як звичайну таблицю для виконання додаткової обробки даних. Створення такого представлення наведено мовою SQL на рис. 2.

```
CREATE OR REPLACE VIEW view_navch_nav AS
SELECT
  study_year,          -- Навчальний рік
  discipline_name,    -- Назва дисципліни
  chair_name,         -- Назва кафедри
  group_name,         -- Назва групи студентів
  student_count,      -- Кількість студентів в групі
  student_count_dis,  -- Кількість студентів, які вивчають дисципліну
  vol_navant         -- Обсяг навантаження в годинах
FROM
...
```

Рис. 2. Створення представлення для отримання первинних даних

У випадку, коли ліцензійні умови використання автоматизованої системи для розрахунку навантаження не дозволяють створювати власних об'єктів бази даних, змінюючи її структуру, можна використати Oracle PL/SQL оператор WITH, який дозволяє дати блоку підзапиту ім'я-псевдонім, на яке можна посилатися у кількох місцях основного SQL-запиту. Ім'я, що присвоєне підзапиту, обробляється так, ніби воно було вбудованим представленням або таблицею. SQL оператор WITH по суті є заміною звичайного підзапиту. Код створення підзапиту для отримання первинних даних наведено на рис. 3).

В наведених на рис. 2-3 лістингах для створення представлення та підзапиту не вказані назви таблиць бази даних, їх зв'язок та походження полів, оскільки в кожній подібній автоматизованій інформаційній системі структура бази даних буде залежати від рівня нормалізації таблиць, способу іменування сутностей та їх атрибутів.

Для розрахунку обсягу сумарного навчального навантаження для групи за формулою (4) необхідно використати код, наведений в лістингу на рис. 4, що реалізує створення підзапиту із використанням віртуальної таблиці з первинними даними.

```
WITH view_navch_nav AS
(
  SELECT
    study_year,          -- Навчальний рік
    discipline_name,    -- Назва дисципліни
    chair_name,         -- Назва кафедри
    group_name,         -- Назва групи студентів
    student_count,      -- Кількість студентів в групі
    student_count_dis,  -- Кількість студентів, які вивчають дисципліну
    vol_navant         -- Обсяг навантаження в годинах
  FROM
  ...
)
```

Рис. 3. Створення підзапиту для отримання первинних даних

```
WITH view_sum_nav AS
(
  SELECT
    nn.study_year,      -- Навчальний рік
    nn.group_name,     -- Назва групи студентів
    sum((nn.student_count_dis / nn.student_count) * nn.vol_navant) as TG
    -- Сумарний обсяг навчального навантаження за групою
  FROM view_navch_nav nn
  GROUP BY
    nn.study_year,
    nn.group_name
)
```

Рис. 4. Створення підзапиту для розрахунку сумарного навчального навантаження для групи

Результат розрахунку частки студентів за кожною дисципліною за формулою (5), що виконується наведеним на рис. 5 кодом мовою SQL, вже можна експортувати до електронної таблиці та виконувати автоматизовану обробку засобами табличного процесору, фільтруючи та сумуючи розраховані частки студентів за дисциплінами: по групах – для перевірки; по кафедрах – для розрахунку кількості студентів, що покривають навантаження кафедр; тощо.

```
SELECT
  vnn.study_year,          -- Навчальний рік (не обов'язкове поле)
  vnn.discipline_name,    -- Назва дисципліни
  vnn.chair_name,         -- Назва кафедри
  vnn.group_name,        -- Назва групи студентів
  vnn.student_count,     -- Кількість студентів в групі
  vnn.student_count_dis, -- Кількість студентів, які вивчають дисципліну
  vnn.vol_navant,        -- Обсяг навантаження в годинах
  (vnn.student_count * vnn.vol_navant / vsn.TG) as twg
  -- Частка студентів групи за дисципліною
FROM view_navch_nav vnn
INNER JOIN view_sum_nav vsn ON
  ((vnn.study_year = vsn.study_year) AND
   (vnn.group_name = vsn.group_name))
WHERE vnn.study_year = :study_year
```

Рис. 5. Розрахунок частки студентів по кожній дисципліні навчального навантаження для первинних даних на навчальний рік

В кодї, наведеному на рис. 5, для фільтрації записів за навчальним роком використовується параметр запиту «:study_year» в предикаті WHERE, що повинен вводитись користувачем перед виконанням запиту та формуванням звіту. Тому перший стовбець предикату SELECT запиту можна виключити зі звіту взагалі.

Наступний крок – формування підсумкового звіту, не такого детального, без великої кількості параметрів щодо навантаження. Для цього запит, наведений на рис. 5, потрібно вважати підзапитом для запиту наступного рівня, що реалізує підведення підсумків та угруповання. Так, для перевірки умови про те, що сума всіх $t_{w,G}$ по навчальній групі G дорівнює кількості студентів цієї початкової групи, необхідно використати запит, наведений на рис. 6. Запит SELECT на вибірку верхнього рівня отримує дані із підзапиту та використовує угруповання з підведенням підсумків. Угруповання виконується за стовбцями підзапиту, переліченими в предикаті GROUP BY зовнішнього запиту на вибірку (в кодї використані стовпці назви навчальної групи та кількості студентів в групі). Для підведення підсумків використовується функція SUM в тексті зовнішнього запиту, в якій вихідним параметром виступає ім'я стовбця із підзапиту (в нашому випадку це стовбець із обсягом навантаження за дисципліною в годинах та стовбець із часткою студентів групи за дисципліною).

```
SELECT
  t.group_name,          -- Назва групи студентів
  t.student_count,      -- Кількість студентів в групі
  sum(t.vol_navant) as nav_group, -- Сумарний обсяг навантаження в годинах
  sum(t.twg) as twg_group -- Сумарна частка студентів за дисциплінами групи
FROM
(
  SELECT
    vnn.study_year,      -- Навчальний рік (не обов'язкове поле)
    vnn.discipline_name, -- Назва дисципліни
    vnn.chair_name,      -- Назва кафедри
    vnn.group_name,      -- Назва групи студентів
    vnn.student_count,   -- Кількість студентів в групі
    vnn.student_count_dis, -- Кількість студентів, які вивчають дисципліну
    vnn.vol_navant,      -- Обсяг навантаження в годинах
    (vnn.student_count * vnn.vol_navant / vsn.TG) as twg
    -- Частка студентів групи за дисципліною
  FROM view_navch_nav vnn
  INNER JOIN view_sum_nav vsn ON
    ((vnn.study_year = vsn.study_year) AND
     (vnn.group_name = vsn.group_name))
  WHERE vnn.study_year = :study_year
) t
GROUP BY
  t.group_name,
  t.student_count
```

Рис. 6. Запит для складання частки студентів за дисциплінами по кожній групі

Отримавши за цим запитом звіт можна проаналізувати та порівняти в кожній групі кількість студентів з сумарною кількістю часток студентів за дисциплінами, що викладаються групі.

Угрупування та складання розрахованого параметру частки студентів групи за кожною навчальною дисципліною можна виконати не тільки за навчальними групами, а й за кафедрами, які викладають навчальні дисципліни студентам. Так, для отримання сумарної кількості часток студентів за дисциплінами по кожній кафедрі закладу освіти необхідно виконати використати запит, що наведений на рис. 7.

```
SELECT
  t.chair_name,          -- Назва кафедри
  sum(t.vol_navant) as nav_chair,
  -- Сумарний обсяг навантаження в годинах
  sum(t.twg) as twg_chair
  -- Сумарна частка студентів за всіма дисциплінами кафедри
FROM
(
  SELECT
    vnn.study_year,      -- Навчальний рік (не обов'язкове поле)
    vnn.discipline_name, -- Назва дисципліни
    vnn.chair_name,      -- Назва кафедри
    vnn.group_name,      -- Назва групи студентів
    vnn.student_count,   -- Кількість студентів в групі
    vnn.student_count_dis, -- Кількість студентів, які вивчають дисципліну
    vnn.vol_navant,      -- Обсяг навантаження в годинах
    (vnn.student_count * vnn.vol_navant / vsn.TG) as twg
    -- Частка студентів групи за дисципліною
  FROM view_navch_nav vnn
  INNER JOIN view_sum_nav vsn ON
    ((vnn.study_year = vsn.study_year) AND
     (vnn.group_name = vsn.group_name))
  WHERE vnn.study_year = :study_year
) t
GROUP BY
  t.chair_name
```

Рис. 7. Запит на складання часток студентів за дисциплінами по кафедрах закладу освіти

Інколи потрібно розділяти навантаження та розрахунок частки студентів групи за дисциплінами робочого навчального плану окремо для різних ступенів освіти, оскільки кількість штатних одиниць залежить не тільки від кількості студентів в групі (за спеціальністю), але й від рівня, який здобувають студенти групи. Для цього потрібно спочатку підзапит, наведений на рис. 2 або 3, доповнити полем «*osvita_level*», розширивши кількість вихідних даних для виконання подальшого розрахунку, а потім використати поле «*osvita_level*» при групуванні записів, як це наведено в лістингу на рис. 8 для отримання сумарних часток студентів по кафедрах з урахуванням ступеня освіти. Приклад результату виконання цього запиту, можна переглянути на рис. 9.

```

SELECT
    t.chair_name,                -- Назва кафедри
    t.osvita_level,             -- Ступінь освіти
    sum(t.vol_navant) as nav_chair, -- Сумарний обсяг навантаження в годинах
    sum(t.twg) as twg_chair
                                -- Сумарна частка студентів за всіма дисциплінами кафедри
FROM
(
    SELECT
        vnn.study_year,         -- Навчальний рік (не обов'язкове поле)
        vnn.discipline_name,    -- Назва дисципліни
        vnn.chair_name,         -- Назва кафедри
        vnn.group_name,        -- Назва групи студентів
        vnn.osvita_level,       -- Ступінь освіти
        vnn.student_count,     -- Кількість студентів в групі
        vnn.student_count_dis, -- Кількість студентів, які вивчають дисципліну
        vnn.vol_navant,         -- Обсяг навантаження в годинах
        (vnn.student_count * vnn.vol_navant / vsn.TG) as twg
                                -- Частка студентів групи за дисципліною
    FROM view_navch_nav vnn
    INNER JOIN view_sum_nav vsn ON
        ((vnn.study_year = vsn.study_year)
        AND (vnn.group_name = vsn.group_name))
    WHERE vnn.study_year = :study_year
) t
GROUP BY
    t.chair_name,
    t.osvita_level
    
```

Рис. 8. Запит на складання часток студентів за дисциплінами по кафедрах закладу освіти з урахуванням ступеня освіти

chair_name	osvita_level	nav_chair	twg_chair
Кафедра 1	Бакалавр	3952,66	61,37404
Кафедра 1	Магістр	552,93	3,892898
Кафедра 2	Бакалавр	1903,3	30,407852
Кафедра 2	Магістр	489,6	3,69593
Кафедра 3	Бакалавр	2361,6	36,527141
Кафедра 3	Магістр	972,1	22,079419
...
Кафедра 17	Бакалавр	2498,8	45,79076

Рис. 9. Приклад результату виконання запиту на отримання кафедрами сумарної частки студентів за освітнім рівнем

Так само можна додати ще декілька параметрів, за якими бажано б було отримувати окремі показники частки студентів закладу для кафедри, наприклад: факультет, форма здобуття освіти, курс навчання, рахунок фінансування навчання. Можна використати фільтрування результатів вибірки та створити угруповання показників для факультетів за групами, для факультетів за кафедрами, тощо.

Висновки та перспективи подальшого дослідження.

В сучасних умовах швидкого розвитку інформаційних технологій неможливо уявити організацію та планування освітнього процесу без сучасних інформаційних систем. Основними задачами планування роботи закладу вищої освіти є розрахунок навчального навантаження та розподіл штатного розкладу між структурними підрозділами закладу. Для удосконалення розподілу штатного розкладу для науково-педагогічних працівників закладу вищої освіти на теперішній час не достатньо використовувати лише обсяг навчального навантаження, що сплановане заздалегідь до початку навчального року. В роботі запропоновано модель розрахунку навчального навантаження за кожною дисципліною в розрізі кількості студентів. Використання такого показника, як частка студентів навчальної групи для кожної дисципліни, дозволяє привести обсяг навчального навантаження будь-якого структурного підрозділу до такого вигляду та значення, що полегшує керівництву закладу приймати рішення та розподіляти отримані кошти для фінансування освітньої діяльності. Запропонований механізм розрахунку легко впроваджується в інформаційну систему, що вже використовується закладом для автоматизації організації освітнього процесу. Реалізація на мові SQL отримання вихідної інформації, виконання розрахунку та подання результату у необхідному вигляді дозволяє будувати звіти без порушення ліцензійних умов використання програмного забезпечення. Задіяння отриманих звітів щодо навчального навантаження із показником частки студентів закладу в плануванні освітнього процесу сприятиме удосконаленню процесів створення навчальних планів, мінімізації необґрунтованих витрат, прозорості розподілу коштів на оплату праці.

Список бібліографічного опису

1. Закон України «Про вищу освіту» (2014) Відомості Верховної Ради України (ВВР), №37-38. С. 2716. URL: <https://zakon.rada.gov.ua/laws/show/1556-18#Text>.
2. Постанова Кабінету Міністрів України від 24 грудня 2019 року № 1146 «Про розподіл видатків державного бюджету між закладами вищої освіти на основі показників їх освітньої, наукової та міжнародної діяльності». Урядовий кур'єр, №251. URL: <https://zakon.rada.gov.ua/laws/show/1146-2019-%D0%BF#Text>.
3. Коцюк Ю.А., Данилець Ю.О. (2016) Системи електронного документообігу у вищому навчальному закладі: тенденції та перспективи впровадження. Наукові записки Національного університету «Острозька академія». Серія «Економіка»: науковий журнал, №1(29). С. 71–75. URL: http://nbuv.gov.ua/UJRN/Nznuoa_2016_1_16.
4. Перехрест Г. (2007) Впровадження електронного документообігу: огляд вітчизняних систем Довідник секретаря та офіс-менеджера, №6. С. 38–44.
5. Мешков Д. Ю. (2001) Електронний документ: обіг, зберігання, використання, законодавчі аспекти. Досвід ФРН (1990-2000). Архіви України, №6. С. 73–87. URL: [http://nbuv.gov.ua/UJRN/ay_2001_6\(248\)_11](http://nbuv.gov.ua/UJRN/ay_2001_6(248)_11).
6. Львов М.С., Співаковський О.В., Щедролюсьєв Д.Є. (2005) Інформаційна система управління вищим навчальним закладом як платформа реалізації управління академічним процесом. Вісник Харківського національного університету. Серія «Математичне моделювання. Інформаційні технології. Автоматизовані системи управління», №1. С. 1-21.
7. Карплюк С.О., Вакалюк Т.А. (2018) Огляд функціональних можливостей програмного забезпечення для управління освітнім процесом закладу вищої освіти. Інформаційні технології і засоби навчання, Том 65, №3. С. 262-276. URL: http://nbuv.gov.ua/UJRN/ITZN_2018_65_3_21.
8. Триус Ю., Стеценко І., Герасименко І., Гриценко В. (2011) Інформаційно-аналітична система управління навчальним процесом ВНЗ. Збірник наукових праць «Інформаційні технології в освіті», №9. С. 40-49. URL: http://nbuv.gov.ua/UJRN/itvo_2011_9_6.
9. Програмне забезпечення для вищих навчальних закладів України, Політек-СОФТ (2017). [Електронний ресурс]. URL: <http://www.politek-soft.kiev.ua/>.
10. АСУ «ВНЗ» [Електронний ресурс]. URL: <https://vuz.osvita.net/>.

References

1. Law of Ukraine "On Higher Education" (2014) The Official Bulletin of the Verkhovna Rada (BVR), 2014, N. 37–38, P. 2716. URL: <https://zakon.rada.gov.ua/laws/show/1556-18#Text>.
2. Resolution to the Cabinet of Ministers of Ukraine dated April 24, 2019 N1146 "On the division of funds of the state budget between deposits based on evidence from their informative, scientific and international and activities". District courier, No. 251. URL: <https://zakon.rada.gov.ua/laws/show/1146-2019-%D0%BF#Text>.
3. Kotsyuk Yu.A., Danilet Yu.O. (2016) Electronic document management systems for major mortgages: trends and prospects for implementation. Scientific notes of the National University "Ostrozka Academy". Series "Economics": scientific journal, No. 1(29), P. 71–75. URL: http://nbuv.gov.ua/UJRN/Nznuoa_2016_1_16.
4. Perehrest G. (2007) Implementation of electronic document management: an overview of financial systems. Adviser to the secretary and office manager, N. 6, P. 38–44.
5. Meshkov D. Yu. (2001) Electronic document: obig, saving, vikoristannya, legislative aspects. Dosvid FRN (1990-2000). Archives of Ukraine, N. 6, P. 73–87. URL: [http://nbuv.gov.ua/UJRN/ay_2001_6\(248\)_11](http://nbuv.gov.ua/UJRN/ay_2001_6(248)_11).

6. Lvov M.S., Spivakovsky O.V., Shchedrolosyev D.Ye. (2005) Information system for managing major initial deposits as a platform for implementing academic process management. Bulletin of Kharkiv National University. Series "Mathematical modeling. Information technologies. Automated control systems", N. 1, P. 1-21.
7. Karplyuk S.O., Vakalyuk T.A. (2018) Review of the functional capabilities of software for controlling the lighting process in the future. Information technologies and science, V. 65, N. 3, P. 262-276. URL: http://nbuv.gov.ua/UJRN/ITZN_2018_65_3_21.
8. Trius Y., Stetsenko I., Gerasimenko I., Gritsenko V. (2011) Information-analytical system for managing the initial process of induction. Collection of scientific works "Information technologies in education", N. 9, P. 40-49. URL: http://nbuv.gov.ua/UJRN/itvo_2011_9_6.
9. Software for most initial mortgages in Ukraine, Politek-SOFT (2017) URL: <http://www.politek-soft.kiev.ua/>.
10. ACS "VNZ". URL: <https://vuz.osvita.net/>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-39>

УДК 004.62, 004.67, 004.91

Дубук Василь Іванович, к.т.н., доцент

<https://orcid.org/0000-0002-6339-1032>

Ковівчак Ярослав Васильович, к.т.н., доцент

<https://orcid.org/0000-0003-3562-4924>

Свистун Богдан Ігорович, студент магістратури

Національний університет "Львівська політехніка", м. Львів, Україна

РОЗРОБКА АВТОМАТИЗОВАНОЇ СИСТЕМИ ЗАМОВЛЕНЬ МОТОВЕЛОЗАПЧАСТИН

Дубук В.І., Ковівчак Я.В., Свистун Б.І. Розробка автоматизованої системи замовлень мотвелозапчастин. У роботі розглянуто проектування та розробку автоматизованої системи замовлень мотвелозапчастин. Приведено обґрунтування актуальності розробки системи. Здійснено аналіз систем-аналогів, приведено їх переваги і недоліки. Побудовано концептуальну модель автоматизованої системи замовлень мотвелозапчастин. Розроблено діаграму прецедентів для користувачів системи. Приведено блок-схему алгоритму роботи системи. Розглянуто функціональну модель системи. Побудовано діаграму класів. Розроблено базу даних. Описано реалізацію компонентів системи, бази даних та інтерфейсу користувача. Запропонована автоматизована система замовлень мотвелозапчастин може знайти успішне використання на підприємствах та в організаціях, що спеціалізуються на технічному обслуговуванні та ремонті транспортних засобів різного типу. Використання розробленої системи підвищить ефективність роботи підприємств, буде сприяти зменшенню логістичних витрат і, як підсумок, зменшить загальні витрати та пришвидшить обслуговування і ремонт транспортних засобів.

Ключові слова: автоматизована система, управління замовленнями, автозапчастини, транспортні засоби, технічне обслуговування, ремонт.

Dubuk V., Kovivchak Ya., Svystun B. Development of an automated system for ordering motorcycle and bicycle parts. The paper considers the design and development of an automated system for ordering of motorcycle and bicycle parts. System development is justified. Analog systems have been analyzed, their advantages and disadvantages are presented. The conceptual model of an automated system of orders of motorcycle and bicycle parts is constructed. A precedent chart for system users has been developed. The block diagram of the system algorithm is given. The functional model of the system is considered. The class chart was constructed. Database has been developed. The implementation of the system components, databases and user interface is described. The proposed automated system of orders of motorcycle and bicycle parts can find successful use at enterprises and organizations specializing in the maintenance and repair of different types of vehicles. The use of the developed system will increase the efficiency of enterprises, will help reduce logistics costs and, as a result, reduce total costs and accelerate the maintenance and repair of vehicles.

Keywords: automated system, order management, auto parts, vehicles, maintenance, repair.

Постановка задачі. Зростання економіки держави неможливе без розвитку та вдосконалення транспортної інфраструктури країни. Розв'язання складних логістичних транспортних завдань з постачання сировини, матеріалів, готової продукції є визначальним фактором економічного росту будь-якого підприємства. Такі завдання можна успішно вирішувати за допомогою використання різних видів транспорту. Серед різних видів, особливе місце належить автомобільному, мотоциклетному і велосипедному транспорту. Саме вказані види транспорту дають змогу найбільш оперативно розв'язати складні логістичні задачі.

Важливу роль індивідуальний автомобільний, мотоциклетний і велосипедний види транспорту відіграють при організації перевезень малогабаритних вантажів, пасажирів, при організації відпочинку та туристичних подорожей.

Для забезпечення надійного функціонування всіх цих видів транспорту необхідно здійснювати їх планове технічне обслуговування та ремонти. Ефективність проведення таких робіт залежить від оперативності та своєчасності постачання необхідних комплектуючих. На сьогодні, здійснити пошук, вибір та організувати постачання потрібних деталей можна за допомогою автоматизованих систем замовлень запчастин. Чим більша кількість подібних систем присутня на ринку одночасно, тим більше можливостей відкривається для проведення успішного планового та позапланового технічного обслуговування різних видів механічних транспортних засобів.

Тому, актуальною є розробка автоматизованих систем призначених для пошуку та здійснення замовлень необхідних комплектуючих для транспортних засобів різних видів.

Сучасний стан розвитку комп'ютерних інформаційних технологій відкриває нові можливості для розробки систем відповідного призначення.

Аналіз останніх досліджень і публікацій. На сьогодні, на ринку присутні різні комерційні системи, призначені для пошуку та замовлення деталей для різних транспортних засобів. Серед них можна виділити наступні спеціалізовані системи "Europarts", "Velogo", "Velik-shop", "Veliki.ua", "ARDIS" - зорієнтовані на здійснення пошуку та замовлення запчастин. Розглянемо їх особливості.

Система "Europarts" [1] розроблена з метою організації продажів запчастин та аксесуарів для мотоциклів і квадроциклів. За її допомогою можна здійснити пошук, аналіз та замовлення понад 11000 комплектуючих для різних видів транспортної техніки. На рис. 1 приведено інтерфейс автоматизованої системи «Europarts».

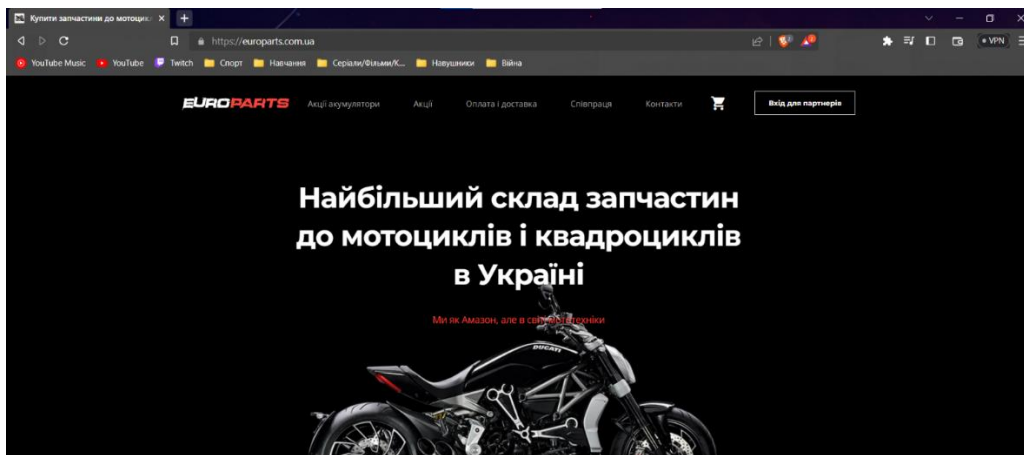


Рис. 1. Інтерфейс системи "Europarts"

До переваг даної системи можна віднести: інтуїтивний інтерфейс; великий вибір комплектуючих; наявність акцій на комплектуючі. Недоліки: відсутність оперативної комунікації з продавцями; неможливість авторизації для звичайних користувачів.

За допомогою системи "Velogo" [2] можна здійснити пошук і замовлення різних видів велосипедів, а також запчастин та аксесуарів до них. У випадку замовлення невідповідних комплектуючих, ця система передбачає можливість обміну товару. При виборі альтернативних комплектуючих, користувачі можуть скористатись функцією порівняння вибраних запчастин за основними характеристиками.

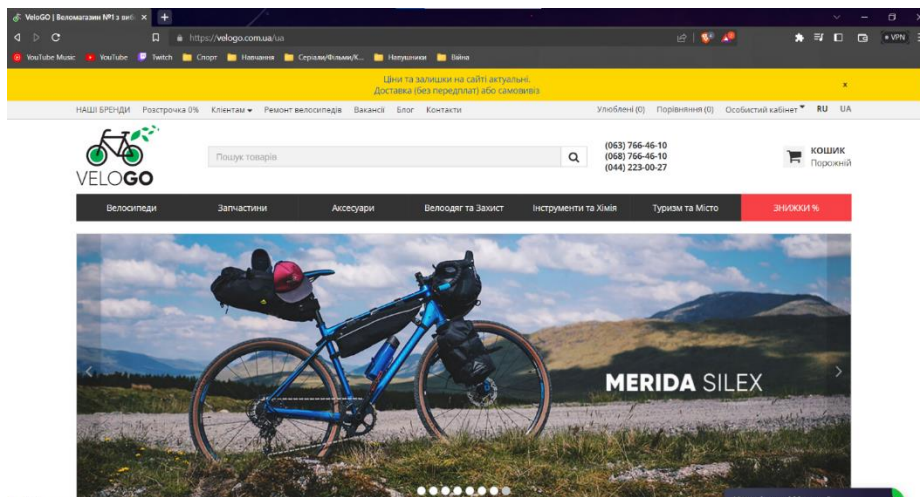


Рис. 2. Інтерфейс системи "Velogo"

При необхідності, користувачі системи можуть отримати оперативну онлайн-консультацію з питань, які виникають.

До переваг системи "Velogo" можна віднести: інтуїтивний інтерфейс; можливість обміну товару; здійснення замовлення без авторизації; наявність онлайн-консультацій. Недоліки: складний пошук комплектуючих по категоріях; затримка у роботі при пошуку даних.

Система "Velik-shop" [3] націлена на користувачів, що активно використовують різного типу велосипеди. Інтерфейс системи приведено на рис. 3.

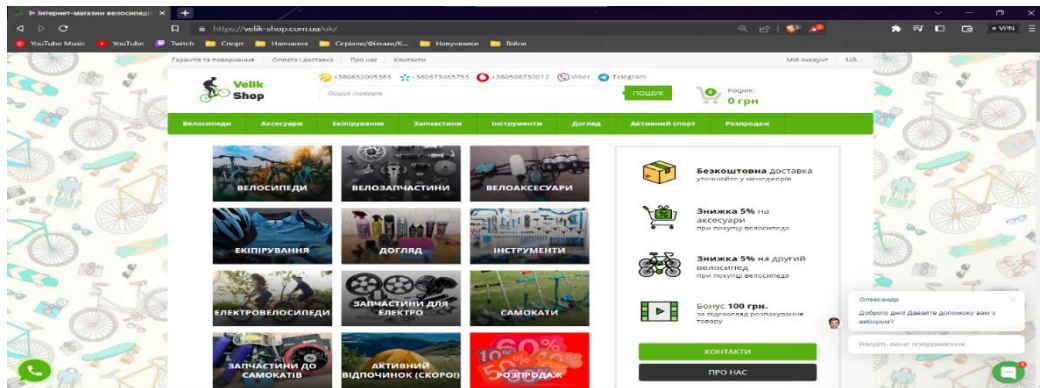


Рис. 3. Інтерфейс системи "Velik-shop"

Як і в більшості подібних систем, за її допомогою можна здійснити пошук, вибір та замовлення необхідних комплектуючих та аксесуарів до велосипедів. Система передбачає роботу як з авторизованими, так і неавторизованими користувачами. Переваги: великий вибір комплектуючих; підтримується різними платформами; формування акційних пропозицій по комплектуючих; можливість використання особистого кабінету. До недоліків можна віднести: складний інтерфейс; онлайн-комунікація не є оперативною.

Подібні недоліки мають системи-аналоги магазинів запчастин [4, 5].

Метою статті є представлення результатів розробки автоматизованої системи замовлень мотовелозапчастин.

Основна частина. У результаті проектування системи на основі підходів [6] було розроблено концептуальну модель автоматизованої системи замовлень мотовелозапчастин. Концептуальна модель системи приведена на рис. 4.

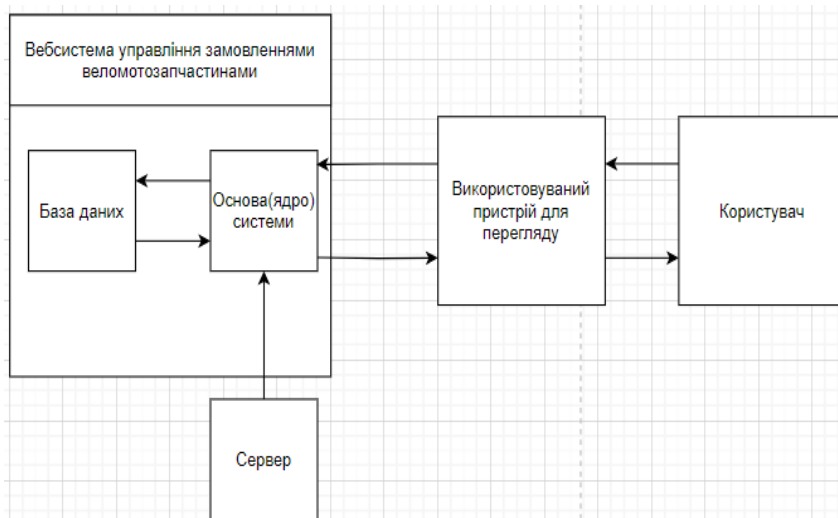


Рис. 4. Концептуальна модель автоматизованої системи замовлень мотовелозапчастин
На концептуальній моделі (рис.4) відображено лише одна зовнішня сутність –

користувач та способи взаємодії користувача із системою.

Також, розроблено діаграму використання для зовнішніх сутностей системи. Діаграма варіантів використання автоматизованої системи замовлень мотобелозапчастин, побудована з використанням рекомендацій, викладених у [7], приведена на рис. 5.

Діаграма варіантів використання системи відображає основні функції системи, які можуть виконувати користувачі та послідовність доступу до них.

У результаті проектування системи розроблено необхідну множину функціональних моделей системи з різними рівнями їх деталізації. При цьому для побудови діаграм, що представляють розроблювану систему на рівні моделей за стандартом IDEF0, було використано інструментальне середовище Ramus Educational [8].

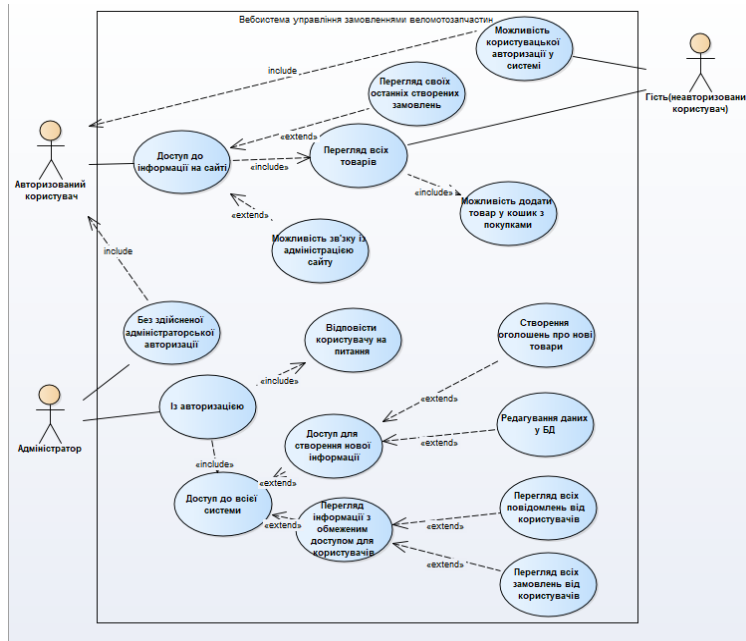


Рис. 5. Діаграма варіантів використання системи

Було розроблено блок-схему алгоритму роботи автоматизованої системи замовлень мотобелозапчастин, результат приведений на рис. 6.

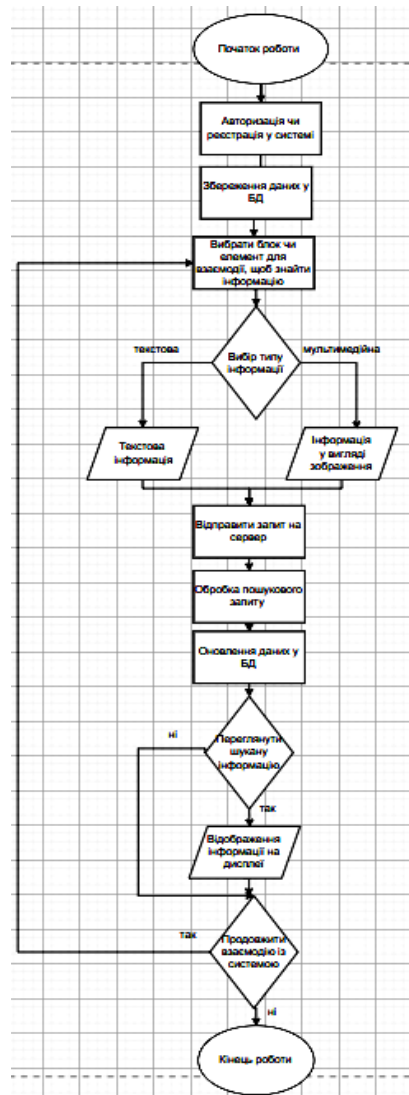


Рис. 6. Блок-схема алгоритму роботи системи

Приведена на рис. 6 блок-схема відображає загальний алгоритм роботи системи при пошуку та здійсненні замовлення необхідних комплектуючих користувачами системи.

Загальна функціональна модель системи розроблена у вигляді контекстної діаграми (рис. 7). На цій діаграмі зображено інформаційні входи та виходи системи, управляючі впливи і механізми, за допомогою яких здійснюється опрацювання вхідних інформаційних потоків. Під час проектування системи деталізація функціональної моделі контекстної діаграми здійснювалась за допомогою діаграм інших рівнів. Як приклад, на рис. 8 приведено діаграму декомпозиції першого рівня.

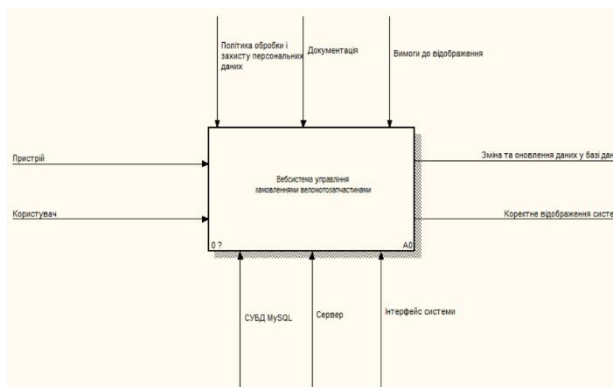


Рис. 7. Контекстна діаграма автоматизова-

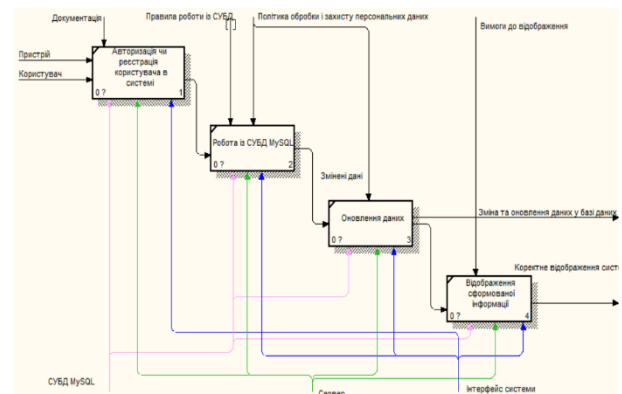


Рис. 8. Діаграма декомпозиції системи

ної системи замовлень мототвелозапчастин

першого рівня

Як можна зауважити з рис. 8, до складу системи будуть входити чотири основні під-системи: підсистема авторизації чи реєстрації користувачів; підсистема роботи з СУБД MySQL; підсистема, яка відповідає за оновлення даних та підсистема відображення інформації. На даній діаграмі (рис.8) вказано необхідні інформаційні та управляючі зв'язки між підсистемами.

Було розроблено діаграму класів, яка відображає класи системи та зв'язки між ними. Результат приведений на рис. 9. У складі системи передбачено наступні класи: користувач; адміністратор; замовлення; кошик; запчастина; реєстрація та відгук. Визначено атрибути для класів та типи зв'язків між ними.

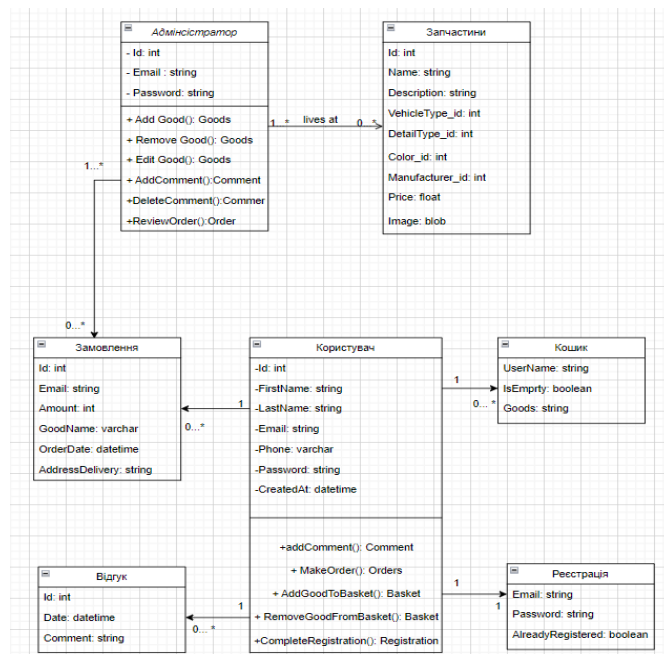


Рис. 9. Діаграма класів автоматизованої системи замовлень мототвелозапчастин

У відповідності до визначеної структури системи та внутрішніх сутностей, розроблено базу даних автоматизованої системи замовлень мототвелозапчастин. Схему бази даних показано на рис. 10.

У результаті розробки системи для роботи з базою даних було вибрано реляційну СУБД MySQL [9], яка відзначається високою продуктивністю, багатозадачністю і користується значною популярністю.

Робота системи побудована на основі клієнт-серверної архітектури. При цьому веб-підсистема реалізована на основі архітектурного шаблону "Модель-Вигляд-Контролер" (Model-View-Controller, MVC) [10], що передбачає поділ даних програми, інтерфейсу користувача і керуючої логіки на три окремих компоненти: модель (М), вигляд (V) та контролер (С) з можливістю незалежної модифікації кожного компонента.

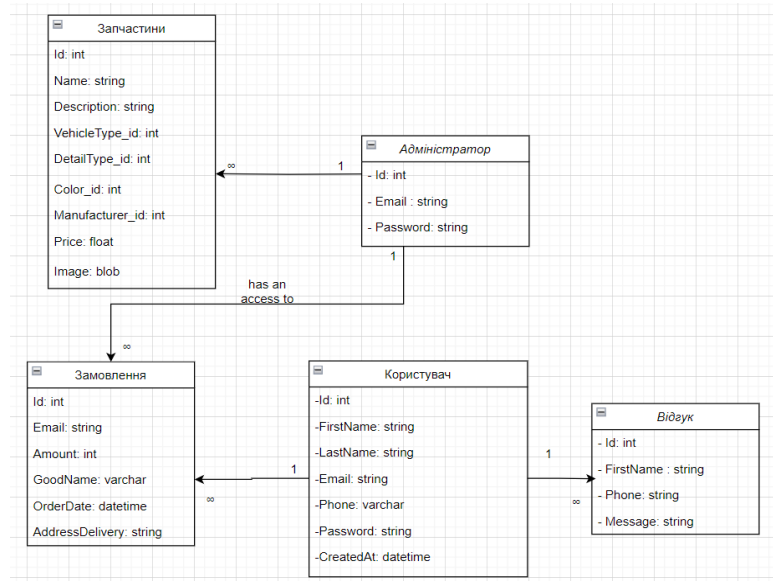


Рис. 10. Схема бази даних системи

При розробці системи було використано фреймворк Spring [11], що забезпечує реалізацію архітектури системи на базі шаблону MVC за допомогою готових шаблонних компонентів. Spring-фреймворк [11] є популярним серед розробників веб-застосунків, оскільки він легко інтегрується з мовою програмування Java.

Як інструментальне середовище для розробки програмного забезпечення системи було вибрано та використано IntelliJ IDEA Ultimate Edition [12]. Для реалізації клієнтської частини системи було вибрано технологію JSP (JavaServer Pages) [13], стандартну бібліотеку тегів JSTL (JavaServer Pages Standard Tag Library) [14], використано HTML (Hyper Text Markup Language) [15], CSS (Cascading Style Sheets) [16] та мову програмування JavaScript [17]. JSP дозволяє динамічно генерувати HTML, XML та різні формати веб-сторінки. Для реалізації серверної частини інформаційної системи було вибрано мову програмування Java та технологію JavaScript [18].

Було здійснено реалізацію інтерфейсу користувача, всіх модулів і компонентів системи, а також проведено її тестування. Як приклад, головну сторінку інтерфейсу користувача приведено на рисунку 11.

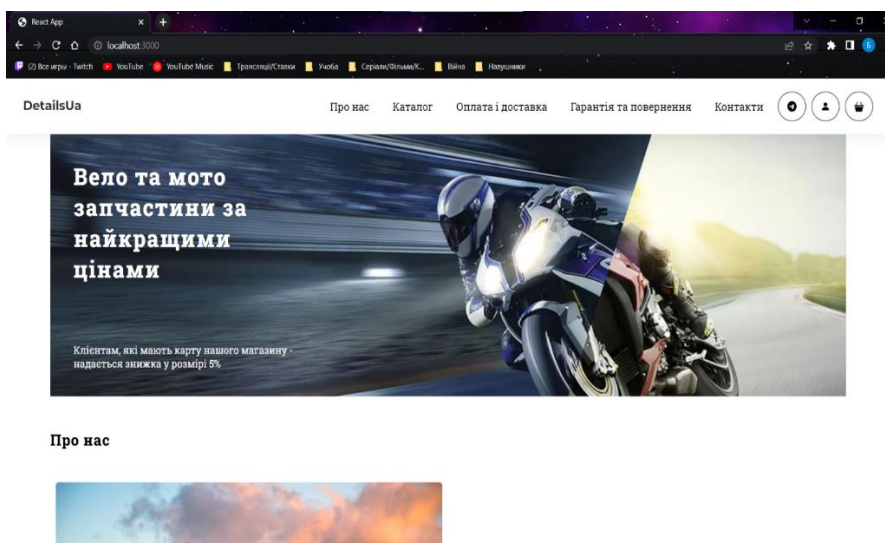


Рис. 11. Головна сторінка інтерфейсу користувача системи

Висновки. У результаті виконання роботи розроблено автоматизовану систему за-
 мовлень мотобелозапчастин. Розглянуто подібні системи, проведено аналіз їх переваг і не-

доліків. Виконано проектування, реалізацію та тестування системи.

Розроблена автоматизована система замовлень мотовелозапчастин може бути успішно застосована на підприємствах та в організаціях, які спеціалізуються на технічному обслуговуванні та ремонті транспортних засобів різного типу. Використання розробленої системи підвищить ефективність роботи ремонтних підприємств та зменшить логістичні витрати на ремонт і технічне обслуговування транспортних засобів.

У перспективі розвитку проекту автоматизованої системи замовлень мотовелозапчастин можна передбачити впровадження модуля інтелектуального аналізу даних з метою прогнозування попиту на різні види запчастин та більш ефективного управління їх запасами.

Список бібліографічного опису

1. Найбільший склад запчастин до мотоциклів і квадроциклів в Україні. Europarts. EuroParts. 2024. URL: <https://europarts.com.ua/> (дата звернення: 12.06.2024 р.)
2. VeloGO. Спеціалізований магазин. Швидко. Зручно. Професійно. Спеціалізована мережа веломагазинів VeloGO. 2024. URL: <https://velogo.com.ua/ua> (дата звернення: 12.06.2024 р.)
3. Velik-Shop. Інтернет магазин велосипедів "velik-shop". Магазин велосипедів у Харкові, Україні. 2024. URL: <https://velik-shop.com.ua/> (дата звернення: 12.06.2024 р.)
4. Офіційний Інтернет-магазин Veliki.ua. Veliki.Ua. 2024. URL: <https://shop.veliki.ua/> (дата звернення: 12.06.2024 р.)
5. ARDIS. ТОВ "Ардіс". 2024. URL: <https://ardis.com.ua/> (дата звернення: 12.06.2024 р.)
6. Olivé A. Conceptual Modeling of Information Systems. Heidelberg: Springer-Verlag Berlin, 2007. 455 p. URL: <https://doi.org/10.1007/978-3-540-39390-0> (date of access: 12.06.2024).
7. Use Case Diagram Tutorial. Visual Paradigm. 2024. URL: <https://online.visual-paradigm.com/diagrams/tutorials/use-case-diagram-tutorial/> (date of access: 12.06.2024).
8. RAMUS Java-based IDEF0 & DFD Modeler. Ramus. URL: <https://ramussoftware.com/> (date of access: 12.06.2024).
9. MySQL and HeatWave Summit. MySQL. Oracle. 2024. URL: <https://www.mysql.com/> (date of access: 12.06.2024).
10. Sheldon R. Model-View-Controller (MVC). TechTarget. 2024. URL: <https://www.techtarget.com/whatis/definition/model-view-controller-MVC> (date of access: 12.06.2024).
11. Why Spring? Spring by VMware Tanzu. URL: <https://spring.io/why-spring> (date of access: 12.06.2024).
12. IntelliJ IDEA – the Leading Java and Kotlin IDE. JetBrains. URL: <https://www.jetbrains.com/idea/> (date of access: 12.06.2024).
13. JavaServer Pages Technology. Oracle. URL: <https://www.oracle.com/java/technologies/jspt.html> (date of access: 12.06.2024).
14. JavaServer Pages Standard Tag Library. Oracle. URL: <https://www.oracle.com/java/technologies/jstl.html> (date of access: 12.06.2024).
15. HTML: HyperText Markup Language. Mozilla. URL: <https://developer.mozilla.org/en-US/docs/Web/HTML> (date of access: 12.06.2024).
16. Lie H.W., Bos B. Cascading Style Sheets: Designing for the Web. Upper Saddle River, NJ: Addison-Wesley, 2005. 392 p.
17. JavaScript. Mozilla. URL: <https://developer.mozilla.org/ru/docs/Web/javascript> (date of access: 12.06.2024).
18. Server-Side JavaScript Guide. Oracle. URL: <https://docs.oracle.com/cd/E19957-01/816-6411-10/contents.htm> (date of access: 12.06.2024).

References

1. Najbilshyj sklad zapchastyn do motocykliv i qvadrocykliv v Ukraini. Europarts. (2024). [The largest composition of spare parts for motorcycles and quadrocycles in Ukraine]. <https://europarts.com.ua/> [in Ukrainian].
2. VeloGO. Spetsializovanyj magazin. Shvydko. Zruchno. Profesijno. Spetsializovana merezha velomagazyniv VeloGO. (2024). [Specialized store. Quickly. Easy. Professionally. Specialized Network of bicycle stores]. <https://velogo.com.ua/ua> [in Ukrainian].
3. Velik-Shop. Internet magazin velosypediv "velik-shop". Magazin velosypediv u Kharkovi, Ukraini. (2024). [Internet store of bicycles "velik-shop". Store of bicycles in Kharkiv, Ukraine]. <https://velik-shop.com.ua/> [in Ukrainian].
4. Ofitsijnyj Internet-magazin Veliki.ua. Veliki.Ua. (2024). [Official Internet store Veliki.ua]. <https://shop.veliki.ua/> [in Ukrainian].
5. ARDIS. LTD "Ardis".(2024). <https://ardis.com.ua/>
6. Olivé, A. (2007). Conceptual Modeling of Information Systems. Heidelberg: Springer-Verlag Berlin. DOI: 10.1007/978-3-540-39390-0
7. Visual Paradigm. (2024). Use Case Diagram Tutorial. <https://online.visual-paradigm.com/diagrams/tutorials/use-case-diagram-tutorial/>
8. Ramus. (2024). RAMUS Java-based IDEF0 & DFD Modeler. <https://ramussoftware.com/>
9. MySQL and HeatWave Summit. MySQL. Oracle. (2024). <https://www.mysql.com/>
10. Sheldon, R. (2024). Model-View-Controller (MVC). TechTarget. <https://www.techtarget.com/whatis/definition/model-view-controller-MVC>
11. Why Spring? Spring by VMware Tanzu. (2024).<https://spring.io/why-spring>

12. IntelliJ IDEA – the Leading Java and Kotlin IDE. JetBrains. (2024). <https://www.jetbrains.com/idea/>
13. JavaServer Pages Technology. Oracle. (2024). <https://www.oracle.com/java/technologies/jspt.html>
14. JavaServer Pages Standard Tag Library. Oracle. (2024). <https://www.oracle.com/java/technologies/jstl.html>
15. HTML: HyperText Markup Language. Mozilla. (2024). <https://developer.mozilla.org/en-US/docs/Web/HTML>
16. Lie, H.W. & Bos, B. (2005). Cascading Style Sheets: Designing for the Web. Upper Saddle River, NJ: Addison-Wesley.
17. JavaScript. Mozilla. (2024). <https://developer.mozilla.org/ru/docs/Web/javascript>
18. Server-Side JavaScript Guide. Oracle. (2024). <https://docs.oracle.com/cd/E19957-01/816-6411-10/contents.htm>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-02>

UDC: 004.9

Turchyn Oleksandr, Postgraduate student

<https://orcid.org/0009-0001-5989-1712>

Ivano-Frankivsk National Technical University of Oil and Gas, Ivano-Frankivsk, Ukraine

IMPLEMENTATION OF MACHINE LEARNING TO INCREASE THE ACCURACY OF FORECASTING THE OPERATING MODES OF DEEP-SEA PUMPING STATIONS

Turchyn O. Implementation of Machine Learning to Increase the Accuracy of Forecasting the Operating Modes of Deep-Sea Pumping Stations. This study endeavors at implementing ML algorithms that are capable of refining the forecast of operating modes of deep-water pumping stations which offshore processes draw their energy from. The classic forecasting methods often do not take into account the complexity of the underwater environment, and so they tend to show suboptimal efficiency, higher maintenance costs and of course wastage of resources. For this study, different ML algorithms including neural networks, support vector machines, random forests, gradient boosting, and linear regression are employed to evaluate how they can imagine operating circumstances under conditions of changes. The rental housing datasets, which contain historical operational data, environmental factors as well as system parameters, are applied to training and validation processes. Data illustrates enhanced capabilities of AI systems with leading candidates being neural network, random forests, and gradient boosting in demonstrating the exact relationships in the sample. The models deliver better performance than the traditional techniques, thereby allowing to assess in-depth the interaction scheme between environmental variables and working modes. These pivotal variables, depth, temperature and pump characteristics are among those that got scrutinized; therefore, insights as to what ought to be embraced for an efficient prediction. Comparative analyses bring forth the tradeoff between the model complexity and interoperability, which state that the algorithm chosen toward application must be thought out very wisely. Ensemble models, which contain a spectrum of different models with each one strong with its own abilities, are seen to be among the balanced way of making precise and useful forecasts. The deep sea pumping stations developed model based on ML(ML) represents an example in practice that sets the framework for increased operational efficiency, reduced maintenance costs, and optimized resource utilization. The findings of this research uncover crucial aspect for engineers, researchers, as well as industry, experts who are prospects of deep-water resource extraction sector. This itself implies a transformation approach toward addressing the problems encountered in dynamic deep-sea environments. With developments in the area of ML, there is a lot of scope for future research ventures to explore new algorithms and real-time techniques which will help to further improve the forecasting capabilities and will certainly result in viable offshore operations. Thus, it can be said that the future of sustainable and resilient offshore operations can to some extent be credited to ML.

Keywords: deep water pumping, Machine learning, Neural network, comparison, prediction.

Турчин О.Б. Впровадження машинного навчання для підвищення точності прогнозування режимів роботи глибоководних насосних станцій. Анотація – Це дослідження спрямоване на реалізацію алгоритмів машинного навчання, які здатні вдосконалити прогноз режимів роботи глибоководних насосних станцій, від яких отримують енергію офшорні операції. Класичні методи прогнозування часто не враховують складність підводного середовища, і тому вони, як правило, показують неоптимальну ефективність, вищі витрати на обслуговування і, звичайно, марнотратство ресурсів. У цьому дослідженні використовуються різні алгоритми машинного навчання, включаючи нейронні мережі, машини опорних векторів, випадкові ліси, градієнтне підсилення та лінійну регресію, щоб оцінити, як вони можуть уявити робочі обставини в умовах змін. Для навчання та валідації застосовуються набори даних про орендоване житло, які містять історичні операційні дані, фактори навколишнього середовища, а також параметри системи. Дані ілюструють розширені можливості систем штучного інтелекту, серед яких провідними кандидатами є нейронні мережі, випадкові ліси та градієнтне підсилення для демонстрації точних взаємозв'язків у вибірці. Моделі демонструють кращі результати, ніж традиційні методи, що дозволяє поглиблено оцінити схему взаємодії між змінними навколишнього середовища та режимами роботи. Ці ключові змінні, глибина, температура і характеристики насосів, є одними з тих, що були ретельно вивчені; отже, з'явилося розуміння того, що має бути враховано для ефективного прогнозування. Порівняльний аналіз дає змогу знайти компроміс між складністю моделі та інтероперабельністю, що свідчить про те, що алгоритм, обраний для застосування, має бути дуже добре продуманий. Ансамблеві моделі, які містять спектр різних моделей, кожна з яких має свої власні можливості, вважаються одним із збалансованих способів створення точних і корисних прогнозів. Розроблена модель глибоководних насосних станцій на основі машинного навчання (МН) є прикладом практичного застосування, що створює основу для підвищення операційної ефективності, зниження витрат на обслуговування та оптимізації використання ресурсів. Результати цього дослідження розкривають важливий аспект для інженерів, дослідників, а також фахівців промисловості, які займаються перспективами глибоководного видобутку ресурсів. Це передбачає трансформаційний підхід до вирішення проблем, що виникають у динамічному глибоководному середовищі. З розвитком у сфері МН відкриваються широкі можливості для майбутніх дослідницьких проєктів з вивчення нових алгоритмів і методів роботи в реальному часі, які допоможуть ще більше поліпшити можливості прогнозування і, безумовно, приведуть до життєздатних морських операцій. Таким чином, можна сказати, що майбутнє стійких і надійних офшорних операцій певною мірою пов'язане з відминанням грошей.

Ключові слова: глибоководні насоси, машинне навчання, нейронна мережа, порівняння, прогнозування.

Introduction. Treating water (sometimes multiple times) and pumping it across great distances are very energy-hungry activities. About 30% perusal of the municipal government expenditures deals with water for drinking and waste disposal. With the largest remaining share of electricity use (more than 10%) are the moving, pumping, and treating of water, use of which is 4% only for the latter (i.e., water conveyance). As the water sector makes growing demands to electricity over time, power intake prediction plays a paramount role in electricity planning for the institutional infrastructure [1]. Besides the municipal water providers and electric grid operators, avoided expenses incurred by either the rate spikes or inconvenience caused by the blackouts that are a result of the sudden surges in energy demand, the smart grid power network benefits a lot from the well-predicted energy use. However, the accurate prediction simplifies energy consumption which in turn reduces under - or an over-estimation [2]. A wrong prediction

about energy use increases the financial impact on electricity supply, so the prices for electricity may grow high if the infrastructure is set up proportionally [3]. With energy consumption being taken for granted the supply of electricity may fall short, power systems may collapse and water supply systems would be disrupted. Thus, on-target energy consumption prediction is a necessary beginning if we hope to understand the energy supply and demand patterns [4].

Machine learning (ML) which is a particular domain of artificial intelligence, helps to train the models by data utilized in the process, so they can be used at specific problems and new information extraction from big data as well [5, 6]. In addition, the computer programming languages and the underlying algorithms have become more adaptive and mature, which have made machine learning more applicable in the technologic applications. The last decade had seen remarkable progress in developing computer technologies which is now increasingly applied in our research for groundwater prediction [7] and the assessment and monitoring of groundwater [8]. Consequentially, an agent of machine learning has emerged as efficient systems for acquiring output(s) from the implied information in groundwater. The role of surface water is dominant in the provision of various socioeconomic benefits. The coefficient for the regression analysis has been measured extensively and the data have been collected. Therefore, artificial intelligence can be used to generate reliable prediction of coefficient. Using ML approaches, it is possible to get to the data point and this is used in forming an expression that relates the coefficient and influencing factors. In that case, it should be noted that ML has not yet been employed to predict the value of electricity which is transferred as a power unit into transported water.

There are several modeling methodologies available for predicting water-related energy usage. We undertake a literature study to determine the developing technique to predicting water sector energy use. Machine learning models have also been shown to be beneficial in modeling the energy consumption of a wastewater treatment plant [9-12] and a distribution system [13]. Several research has used machine learning algorithms to estimate water-related energy demand, however, they have only forecasted the energy usage of a single water plant. Previous research has not examined and compared the model's performance for the full transfer of the water system and its components. ML algorithms might have fared well in estimating the consumption of energy for a certain water plant. However, nothing is known about estimating energy use for a collection of water facilities using machine learning. Inter-basin transfers of water projects such as the State Water Project and the Mokelumne River Aqueduct might involve energy forecasting models, but they are unknown or not publicly available.

The correct operation of water delivery systems is closely related to the population's ability to obtain water. Water and energy are two of the most important resources, and their combined management may yield major economic and environmental advantages in both sectors. In this regard, the United Nations Sustainable Development Goals (SDGs), particularly Goals 6 and 11, have identified the issue of ensuring the availability and long-term sustainability of water for all, addressing specific actions that guarantee the availability of clean water and focusing on resource utilization improvement [14].

Problem statement. Deep-sea pumping stations are vital parts of marine resource extraction operations, acting as important gear in the effective recovery of resources from the ocean floor. However, because the undersea environment is complicated and dynamic, accurately anticipating modes of operation in these deep-sea pumping centers is a substantial difficulty. Previous prediction methods often ignored the interaction between work, environment, and physical activity, resulting in decreased performance, increased maintenance costs, and material waste [15].

The ability of deep water stations to anticipate and adapt to changes in operations is limited by current forecast accuracy, resulting in positive and negative operational inefficiencies. Additionally, inaccurate estimates can cause unnecessary wear and tear on equipment, reducing the overall reliability of the station.

It seems that more and advanced predictive machine learning algorithms will be needed to solve these problems. In many fields, machine learning shows promise in identifying complex patterns and relationships in data. Using machine learning techniques to predict the operating patterns of deep water pumping stations increases the accuracy of predictions, ensures optimum utilization and improves the operation of the whole.

Aim and Objective. The aim of this study is to analyze and develop a machine learning-based prediction system that takes into account the interaction between underperformance, environment, and genetic information. In this way, the research aims to solve the limitations of the estimation of the current deep-sea station; This will ultimately lead to improved operational efficiency, reduced maintenance costs and environmental responsibility of external resources.

The objectives of the research are:

1. A novel ML comparison is created to predict how deep-sea pumping stations will operate.
2. Determining the variables that influence operating mode forecast accuracy in deep-sea environments.
3. Increased forecasting leads to better resource consumption, maintenance scheduling, and system efficiency.

Deep-sea resource extraction engineers, researchers, and industry experts will benefit greatly from the study's findings, which will enhance operating protocols and ensure the long-term viability of deep-sea pumping station operations.

Related work. Water delivery systems have been significantly improved by software and digital technologies. The objective of scheduling and managing water supply systems with software-based solutions is to optimize energy savings, minimize water loss, and save resources during the water distribution process. Although

water treatment techniques can potentially result in significant energy savings, the focus of this section is on water supply efficiency solutions. Conventional hydraulic modeling methods, like EPANET, are frequently employed to analyze distribution network performance and allocate water needs from customers to calculation nodes. The United States Environmental Protection Agency was the first to develop EPANET, a modeling software package for water distribution networks. It simulates water distribution and hydraulics in pressurized pipe water systems. It enhances comprehension of water flow in distribution systems.

These software tools have been utilized in several projects that have developed water distribution systems to meet the constant demand for water in various locations.

Several optimization strategies have been created, except commercial software packages, to boost the efficacy of pump scheduling throughout the day by defining the exact hours of the day that the pump should be switched on. This is because the operation of the system that supplies water has to be made more effective. The implementation of genetic algorithms to reduce pumping operational costs by taking advantage of off-peak electricity rates and space for storage in the water distribution system marked the beginning of efforts toward this method. In order to provide water supply operational methods for reducing costs and using energy, heuristic and meta-heuristic methodologies have also been integrated into software that is easily accessible, such as EPANET, and deployed in actual water distribution networks. The aforementioned research has opened the door for data-driven solutions by achieving notable savings in energy of up to 10% when comparing the consumption of energy to the energy used prior to the execution of the suggested strategies [16].

Other studies have tried to integrate the water level into storage tanks with the pumping operation optimization schedule problem. In order to control the water system, it is essential to manage both the pumping schedules and trigger levels. Different trigger levels should be used at different times of the day in order to lower peak pumping and pumping heads. 20% less energy is being used now that an algorithm based on evolution that incorporates historical data and integer decision criteria has been implemented. Additionally, the fact that specific regulations governing the management of water systems have been resolved. For example, pumping stations can now be controlled under the water levels in different tanks, or tank levels and the time of day can both be taken into consideration to minimize pumping throughout peak tariff periods [17-20].

The amount of data being generated these days is always increasing. Real-time data in the energy domain is produced by Internet of Things technologies. Examples of this data include sensor-based data [21], efficiency investment information [22], smart meters for energy consumption and RES production [23], and grid-based assets like transformer feeders [24]. Additionally, it is now easier to obtain data that may be used, such as power or weather records, which opens up new possibilities for developing models and developing techniques for finding patterns in data. It is also possible to obtain and utilize other data that isn't directly related to the energy sector, including data from water pumps, in the algorithms that are developed. Without a doubt, each of these data sources offers the potential for creating multiple scales and multi-stakeholder strategies through innovative analytics meant to give energy stakeholders more solid and useful information, enhancing decision-making based on data [25].

The rapid progress of deep learning in recent years has led to a paradigm shift in methods of visual analysis. Pumping station equipment pictures may be assessed using deep learning algorithms, which might lead to more dependable and effective technical support for pumping station operations administration and maintenance. The first deep learning method based on neural networks was presented by Ma X et al. [26], marking the start of the deep learning era. The industry responded strongly to Alex Krizhevsky et al.'s creation of an AlexNet [27], which was based on the architecture of the convolutional neural network (CNN) that won the ImageNet recognition competition. To tackle the phenomenon of network models degrading after thorough training, A CNN ResNet built on the shortcut design was presented by He and colleagues [28]. Deep learning models become more expressive and suitable for difficult tasks as a result. Szegedy et al. at Google developed an Inception V4 that reduces the number of variables in CNN and accelerates the algorithm's execution, based on Inception and Residual architecture [22]. The continued progress in the area is leading to the proposal of a wider range of deep learning approaches. Recurrent neural networks (RNNs), which recognize connections in sequence data and maintain the model's retention of prior knowledge, and generative adversarial networks (GANs), which increase models' ability for generalization by increasing sample sizes, are a few examples. Deep learning algorithms perform better in picture segmentation tasks when the U-net network is used [28], Deep learning models may now more effectively capture distant dependencies thanks to the Transformer network's solution to the disappearing and expanding gradient problem [29]. Additionally, LSTM is a special sort of RNN that can learn dependence over time information [27]. Deep learning models may therefore be applied in a wide range of fields, including as image recognition, speech recognition, and data analysis. Deep learning can be used for visual analysis to achieve high precision and resilience while reducing the need for human feature extraction design techniques after training on a large dataset [29].

Methodology. The study methodology's goals are to list, categorize, and evaluate the significant ML and DL models that are applied to energy systems. As per our thorough analysis, implementing search queries using Thomson Reuters Web-of-Science and Elsevier Scopus will guarantee that all papers inside the database satisfy the crucial criteria of originality, high impact, and high h-index. Additionally, we sought to create four distinct categories—single ML models, hybrid models, ensemble models, and DL—for the models utilized in energy systems in order to provide a thorough analysis and comprehension of each modeling approach and its advancement. The initial database of

pertinent articles is found in step 1 of the process utilizing the terms "energy system," "machine learning," "neural network," "support vector", "DT," "MLP," "ELM," "ensemble," and "deep learning." On the other hand, we performed a fresh search query to match each ML approach appropriately. These searches will locate pertinent articles, but they do not specify whether the ML model is part of an ensemble or a hybrid. Furthermore, not all of the articles in the original database may be relevant. One or more single models may be included in a hybrid or ensemble machine learning model. Because of this, the methodology's stages two and three are made to group the ML models into the appropriate groups for the review. Step 4 involves classifying the models into four groups and setting them up in distinct tables for individual inspection.

The research was primarily based on an extensive review of literature available through online databases, IEEE Xplore, PubMed, ScienceDirect, SpringerLink, Google Scholar. Using the following methodology, a quantitative approach was used to perform the research: thorough analysis of the body of research on current studies and web resources. assembling of deep-sea pumping station operating history data that was found in pertinent internet resources. using statistical techniques and exploratory data analysis to choose important traits. An analysis of several machine learning methods, such as linear regression, gradient boosting, random forests, neural networks, and support vector machines. Model selection training, and validation using hyper-parameter adjustments. model performance is compared using related measures.

Result and Discussion. Positive outcomes were obtained when deep-sea pumping station operating modes were predicted using ML techniques; forecasts were significantly more accurate than those made using traditional methods. After being trained on a large dataset comprising historical operational data, environmental factors, and system features, the model demonstrated a high level of adaptability to the dynamic conditions typical of deep-sea ecosystems.

Analysis of non-significant factors affecting the prediction accuracy of the performance model shows that water level, temperature and pump performance have a significant impact on the prediction model. The machine learning system can learn and adapt to the relationship between various components to provide more detailed and context-aware predictions of pump performance. This predictive accuracy has important effects for forces, permitting them to work flexibly and respond quickly to ecological changes.

Compared with traditional approaches, estimation of ML-based prediction procedures clearly shows their benefits in precision and consistency. ML prototypes often overtake traditional predicting techniques, particularly when conservational parameters change unexpectedly or unanticipated working problems occur. This displays that ML has the probable to be a game-changer for the flexibility and sustainability of deep water operations.

The efficiency of the predicting structure is showed by its skill to rise production, lessen maintenance budgets, and advance resource operation. The outcomes of the study illustration how significant it is to use ML techniques in the deep-sea situation to establish a steady and long-term offshore operation in the ocean.

Evaluation and comparison of various ML algorithms to predict deep ocean station model performance:

1. **Neural Networks:**

- *Strengths:* Using NN to predict operational models of deep-sea stations is advantageous due to their ability to capture non-linear associations in data. They can make very precise forecasts by classifying intricate outlines from past data.

- *Considerations:* The results of NN training can be difficult and data intensive to interpret. Careful architecture and hyper-parameter tuning are important to demonstrate performance.

2. **Support Vector Machines (SVM):**

- *Strengths:* SVM can handle high-dimensional data and adapt to new situations. Their ability to make complex boundary decisions makes them ideal for documenting a variety of operations in the deep water region.

- *Considerations:* The choice of the kernel function can have a significant impact on the performance of the SVM and should therefore be evaluated carefully. May be less defined than basic structure.

3. **Random Forest:**

- *Strengths:* An integrated algorithm that can perform many different operations, including random forests. They have the ability to be the best, to be strong, and to catch different trends effectively. They work well in complex dynamic systems that need to be predicted.

- *Considerations:* Random forests can be computationally expensive and their definition increases as the number of trees in the set increases.

4. **Gradient Boosting:**

- *Strengths:* Using gradient boosting algorithms (such as XGBoost or LightGBM) can identify connections and relationships in data. They generally provide excellent accuracy and handle large files well.

- *Considerations:* Hyper-parameters must be maintained carefully and, as usual, overfitting may occur if they are not sufficient. Interpretations may be limited compared to simple models.

5. **Linear Regression:**

- *Strengths:* Linear regression provides interpretation and convenience. This can be useful when there are many relationships between different columns. Serves as a standard model for comparison.

- *Considerations:* Linear regression may have limited effectiveness in estimating complex functional models due to its inability to capture nonlinear relationships.

Ultimately, the unique characteristics of the data, the complexity of the system, and the balance between interpretation and model complexity will determine which machine learning is best for prediction: tower operating models of deep stations. In this difficult situation, combination methods such as gradient boosting or combining random forest results with simple model interpretation can provide the necessary strategies to obtain accurate and efficient results.

Conclusion. Using machine learning algorithms to predict the performance of deep water facilities is a revolutionary technology with the potential to improve safety and efficiency. This study demonstrates the advantages and disadvantages of various machine learning algorithms such as neural networks, support vector machines, random forests, gradient boosting and linear regression in predicting the dynamic behavior of deep stations.

The results demonstrate the effectiveness of learning models (particularly neural networks, random forests, and gradient boosting) for capturing the interaction between operating parameters, environmental variables, and historical data. These advanced algorithms demonstrate their ability to create detailed and adaptive features of deep-water environments, demonstrating the advantages of traditional methods.

The key factors when choosing a machine learning algorithm are interpretability, computational efficiency, and the need for accurate hyper-parameter tuning. Combining the advantages of various models, the integrated system has become a possible way to achieve harmony between interpretation and reality. The importance of this study is demonstrated by the fact that predictive methods developed by machine learning bring specific results in increasing efficiency, reducing maintenance costs and optimizing resource use. The proposed model makes the operation of deep-sea stations more stable and reliable, and this model enables a good decision to adapt to changes.

To further improve prediction, future research in machine learning will examine pruning techniques, data augmentation, and real-time techniques. This research has laid the foundation for using new technologies to solve unique problems arising from the deep ocean environment and opened the door to more powerful and efficient ways of outsourcing.

References

1. Yi S., Kondolf G. M., Sandoval-Solis S., Dale L. Application of machine learning-based energy use forecasting for inter-basin water transfer project. *Water Resources Management*. 2022. Vol. 36. P. 5675–5694.
2. Pasha M. F. K., Weathers M., Smith B. Investigating energy flow in water-energy storage for hydropower generation in water distribution systems. *Water Resources Management*. 2020. Vol. 34. P. 1609–1622.
3. da Silveira A. P. P., Mata-Lima H. Assessing energy efficiency in water utilities using long-term data analysis. *Water Resources Management*. 2021. Vol. 35. P. 2763–2779.
4. Perelman G., Fishbain B. Critical Elements Analysis of Water Supply Systems to Improve Energy Efficiency in Failure Scenarios. *Water Resources Management*. 2022. Vol. 36. P. 3797–3811.
5. Rattan P., Penrice D. D., Simonetto D. A. Artificial intelligence and machine learning: what you always wanted to know but were afraid to ask. *Gastro Hep Advances*. 2022. Vol. 1. P. 70–78.
6. Soori M., Arezoo B., Dastres R. Artificial intelligence, machine learning and deep learning in advanced robotics, A review. *Cognitive Robotics*. 2023.
7. Zhao T., Zhu Y., Ye M., Mao W., Zhang X., Yang J., et al. Machine-Learning Methods for Water Table Depth Prediction in Seasonal Freezing-Thawing Areas. *Groundwater*. 2020. Vol. 58. P. 419–431.
8. Chen C., Zhang H., Shi W., Zhang W., Xue Y. A novel paradigm for integrating physics-based numerical and machine learning models: A case study of eco-hydrological model. *Environmental Modelling & Software*. 2023, Vol. 163, P. 105669.
9. Bagherzadeh F., Nouri A. S., Mehrani M.-J., Thennadil S. Prediction of energy consumption and evaluation of affecting factors in a full-scale WWTP using a machine learning approach. *Process Safety and Environmental Protection*. 2021. Vol. 154. P. 458-466.
10. Das A., Kumawat P. K., Chaturvedi N. D. A Study to Target Energy Consumption in Wastewater Treatment Plant using Machine Learning Algorithms. In *Computer Aided Chemical Engineering*. Elsevier. 2021. Vol. 5. P. 1511-1516.
11. Li J., Tang W. Z. Improved Unit Energy Efficiency and Reduced Cost by Innovative Industrial Wastewater Treatment Systems. *Environmental Processes*, 2021, Vol. 8. P. 1433-1454.
12. Zhang S., Wang H., Keller A. A. Novel machine learning-based energy consumption model of wastewater treatment plants. *ACS ES&T Water*. 2021. Vol. 1. P. 2531-2540.
13. Salvino L. R., Gomes H. P., Bezerra S. d. T. M. Design of a Control System Using an Artificial Neural Network to Optimize the Energy Efficiency of Water Distribution Systems. *Water Resources Management*. 2022. Vol. 36. P. 2779-2793.
14. Carlsen L., Bruggemann R. The 17 United Nations' sustainable development goals: A status by 2020. *International Journal of Sustainable Development & World Ecology*. 2022. Vol. 29. P. 219-229.
15. Martinsen G., Liu S., Mo X., Bauer-Gottwein P. Optimizing water resources allocation in the Haihe River basin under groundwater sustainability constraints. *Journal of Geographical Sciences*. 2019. Vol. 29. P. 935-958.
16. Sarmas E., Spiliotis E., Marinakis V., Tzanes G., Kaldellis J. K., Doukas H. ML-based energy management of water pumping systems for the application of peak shaving in small-scale islands. *Sustainable Cities and Society*, 2022, Vol. 82. P. 103873.

17. Wang Y., Sun S., Chen X., Zeng X., Kong Y., Chen J., et al. Short-term load forecasting of industrial customers based on SVM and XGBoost. *International Journal of Electrical Power & Energy Systems*. 2021. Vol. 129. P. 106830.
18. Oreshkin B. N., Dudek G., Peřka P., Turkina E. N-BEATS neural network for mid-term electricity load forecasting. *Applied Energy*. 2021. Vol. 293. P. 116918.
19. Nespoli A., Ogliaari E., Leva S., Massi Pavan A., Mellit A., Lughii V., et al. Day-ahead photovoltaic forecasting: A comparison of the most effective techniques. *Energies*. 2019. Vol. 12. P. 1621.
20. Abbasi R. A., Javaid N., Ghuman M. N. J., Khan Z. A., Rehman S. Ur, Amanullah. Short term load forecasting using XGBoost. In *Web, Artificial Intelligence and Network Applications: Proceedings of the Workshops of the 33rd International Conference on Advanced Information Networking and Applications (WAINA-2019)*. 2019. Vol. 33. P. 1120-1131.
21. Li C. Designing a short-term load forecasting model in the urban smart grid system. *Applied Energy*. 2020. Vol. 266. P. 114850.
22. Sarmas E., Spiliotis E., Marinakis V., Koutselis T., Doukas H. A meta-learning classification model for supporting decisions on energy efficiency investments. *Energy and Buildings*. 2022. Vol. 258. P. 111836.
23. Saadi M., Noor M. T., Imran A., Toor W. T., Mumtaz S., Wuttisittikulij L. IoT enabled quality of experience measurement for next generation networks in smart cities. *Sustainable Cities and Society*. 2020. Vol. 60. P. 102266.
24. Kong X., Liu C., Shen Y., Hu W., Ma T. Power supply reliability evaluation based on big data analysis for distribution networks considering uncertain factors. *Sustainable Cities and Society*. 2020. Vol. 63. P. 102483.
25. Makridakis S., Spiliotis E., Assimakopoulos V. The M4 Competition: 100,000 time series and 61 forecasting methods. *International Journal of Forecasting*. 2020. Vol. 36. Pp. 54-74.
26. Ma X., Wu J., Xue S., Yang J., Zhou C., Sheng Q. Z., et al. A comprehensive survey on graph anomaly detection with deep learning. *IEEE Transactions on Knowledge and Data Engineering*. 2021.
27. Shorten C., Khoshgoftaar T. M. A survey on image data augmentation for deep learning. *Journal of Big Data*. 2019. Vol. 6. P. 1-48.
28. Liu Z., Mao H., Wu C.-Y., Feichtenhofer C., Darrell T., Xie S. A convnet for the 2020s. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. P. 11976-11986.
29. Khan S., Naseer M., Hayat M., Zamir S. W., Khan F. S., Shah M. Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*. 2022. Vol. 54. P. 1-41.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-03>

УДК 004:02

Анісімов Віктор Геннадійович, магістр

<https://orcid.org/0009-0006-0146-081X>

Кунанець Наталія Едуардівна, д.н.с.к., професор

<https://orcid.org/0000-0003-3007-2462>

Національний університет «Львівська політехніка», м. Львів, Україна

ПЕРЕХІД ВІД МОНОЛІТНОЇ ДО МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ: МЕТОДОЛОГІЯ ТА ДОСВІД ВПРОВАДЖЕННЯ

Анісімов В. Г., Кунанець Н.Е. Перехід від монолітної до мікросервісної архітектури: методологія та досвід впровадження. У статті розглядається проблема переходу від монолітної архітектури до мікросервісної з розподіленою базою даних на прикладі інформаційної системи для розвиваючого контролю використання мобільних пристроїв. Монолітна архітектура часто стає неефективною при активному розвитку системи, створюючи проблеми з масштабованістю, гнучкістю та управлінням даними. У роботі запропоновано метод поступової декомпозиції монолітної системи на незалежні мікросервіси з власними базами даних. Проаналізовано етапи аналізу початкової структури, визначення потенційних проблем та обмежень монолітної архітектури, а також кроки переходу до розподіленої архітектури. Особливу увагу приділено питанням проектування розподіленої бази даних, зокрема, розподілу даних між регіонами з урахуванням вимог до захисту персональних даних, забезпечення узгодженості та цілісності даних у розподіленому середовищі. Отриманий досвід та рішення можуть бути корисними для розробників та архітекторів при модернізації інформаційних систем зі схожими викликами.

Ключові слова: монолітна архітектура, мікросервісна архітектура, розподілена база даних, декомпозиція, цілісність даних, системний аналіз, інформаційна система.

Anisimov V., Kunanets N. Transition from Monolithic to Microservice Architecture: Methodology and Implementation Experience. The article addresses the problem of transitioning from a monolithic architecture to a microservice architecture with a distributed database, using the example of an information system for the developmental controlling of mobile devices. Monolithic architecture often becomes inefficient with the active development of the system, creating issues with scalability, flexibility, and data management. The study proposes a method of gradual decomposition of the monolithic system into independent microservices with their own databases. It describes the stages of analyzing the initial structure, identifying potential problems and limitations of the monolithic architecture, and the steps for transitioning to a distributed architecture. Special attention is paid to the design of the distributed database, particularly the distribution of data across regions considering personal data protection requirements, and ensuring data consistency and integrity in a distributed environment. The experience and solutions obtained may be useful for developers and architects in modernizing information systems facing similar challenges.

Keywords: monolithic architecture, microservice architecture, distributed database, decomposition, data integrity, system analysis, information system.

Постановка наукової проблеми. У сучасному світі інформаційних технологій, де системи постійно розвиваються та ускладнюються, питання ефективної архітектури та управління даними стає все більш актуальним. Переважна більшість інформаційних систем починають своє життя як монолітна архітектура з єдиною централізованою базою даних. Такий підхід часто є простим та зручним на початкових етапах розроблення, коли функціональність інформаційної системи обмежена, а обсяг даних не надто великий. Однак, з плином часу, коли інформаційна система активно розвивається і до неї постійно додається новий функціонал, монолітна архітектура створює певні проблеми, такі як складність розгортання та оновлення окремих компонентів, обмеження у горизонтальному масштабуванні, труднощі в управлінні великою монолітною базою даних та інші.

Постає наукова проблема дослідження методів та підходів до переходу від монолітної архітектури до більш гнучкої та масштабованої архітектури, зокрема, мікросервісної архітектури з розподіленою базою даних. Необхідно розробити ефективні стратегії декомпозиції монолітних систем на незалежні сервіси, механізми забезпечення узгодженості та цілісності даних у розподіленому середовищі, а також методи проектування та управління розподіленими базами даних для забезпечення високої продуктивності, доступності та відмовостійкості системи.

Аналіз досліджень та публікацій. Проблеми масштабованості, гнучкості та складності розгортання і підтримки монолітних архітектур інформаційних систем широко висвітлені в науковій літературі. Багато дослідників та практиків в галузі інформаційних технологій вивчають переваги та виклики переходу до мікросервісних архітектур та розподілених баз даних.

У дослідженні [1] проведено ґрунтовний аналіз концепцій, характеристик та переваг мікросервісних архітектур. Автори розглянули різні аспекти проектування, розроблення та розгортання мікросервісів, а також виділили ключові виклики, такі як забезпечення узгодженості

даних та міжсервісної комунікації. Це дослідження є важливим для розуміння основних принципів та особливостей мікросервісної архітектури, а також для виявлення потенційних проблем та шляхів їх вирішення. У праці [2], автори детально описали принципи та підходи до декомпозиції монолітних систем на мікросервіси. Вони запропонували методи виділення функціональних областей, визначення меж сервісів та організації взаємодії між ними. Ці методи є ключовими для успішного переходу від монолітної до мікросервісної архітектури, оскільки дозволяють систематично розділити систему на незалежні компоненти та забезпечити їх ефективну взаємодію.

Дослідження [3] зосереджено на шаблонах та практиках проектування мікросервісних архітектур, зокрема питаннях розподілу даних між сервісами, забезпечення узгодженості та цілісності даних у розподіленому середовищі. Автор розглядає різні підходи до організації даних у мікросервісах та надає рекомендації щодо вибору відповідних стратегій в залежності від вимог системи. Ці знання є важливими для проектування ефективної та надійної мікросервісної архітектури з розподіленою базою даних.

Щодо розподілених баз даних, у дослідженні [4] розглянуто різні стратегії розподілу даних, такі як реплікація, шардування та створення федеративних баз даних. Автори проаналізували переваги та недоліки кожного підходу з точки зору продуктивності, масштабованості та доступності даних. Це дослідження надає цінну інформацію для вибору оптимальної стратегії розподілу даних у мікросервісній архітектурі в залежності від специфічних вимог та обмежень системи.

У роботі [5], автор сформулював теорему CAP, яка визначає фундаментальні обмеження розподілених систем щодо узгодженості, доступності та стійкості до розділення. Ця теорема є важливим фактором при проектуванні розподілених баз даних та виборі моделі узгодженості даних. Вона допомагає зрозуміти компроміси, які необхідно враховувати при розробленні розподілених систем, та прийняти обґрунтовані рішення щодо балансу між узгодженістю, доступністю та стійкістю до розділення.

Незважаючи на значну кількість досліджень у галузі побудови мікросервісних архітектур, процес переходу від монолітної до мікросервісної системи залишається унікальним викликом для кожної окремої інформаційної системи. Це пов'язано зі специфічними вимогами, обмеженнями та бізнес-логікою, притаманними кожній інформаційній системі [3, 6].

Згідно з дослідженнями [1, 6, 7], перехід до розподіленої архітектури та бази даних може бути необхідним у наступних випадках:

1. Масштабованість. Коли обсяг даних або кількість користувачів зростає настільки, що одного вузла бази даних стає недостатньо для опрацювання навантаження.

2. Висока доступність. Розподілена база даних з репліками даних забезпечує вищу доступність системи, оскільки у разі збою одного з вузлів, інші вузли можуть продовжувати обслуговувати запити.

3. Географічна розподіленість. Якщо користувачі системи розподілені географічно, розподілена база даних може забезпечити кращу продуктивність за рахунок зменшення затримок при обробці запитів.

4. Відмовостійкість. Розподілена база даних з репліками даних є більш відмовостійкою, оскільки у разі збою одного з вузлів, дані залишаються доступними на інших вузлах.

Дана робота спрямована на те, щоб поділитися досвідом здійснення переходу від монолітної до мікросервісної архітектури на прикладі інформаційної системи розвиваючого контролю використання мобільних пристроїв. Метою є запропонувати практичний метод такого переходу, який враховує специфічні вимоги та обмеження подібних систем. Цей метод може слугувати орієнтиром та основою для інших розробників та архітекторів, які стикаються з подібними викликами при модернізації своїх інформаційних систем.

Мета роботи. Запропонувати методологію трансформації архітектури інформаційної системи, що активно використовується, від монолітної до мікросервісної з розподіленою базою даних, дослідити процес трансформації, виділити ключові етапи, виклики та запропонувати ефективні рішення для успішної модернізації системи без переривання її роботи.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. В сучасному світі інформаційних технологій, де системи постійно розвиваються та ускладнюються, питання ефективної архітектури та управління даними стає все більш актуальним. Багато систем починають своє життя як монолітна архітектура з єдиною централізованою базою даних. Такий підхід часто є простим та зручним на початкових етапах розроблення, коли функціональність системи обмежена, а обсяг даних не надто великий.

Однак, з плином часу, коли система активно розвивається і до неї постійно додається новий функціонал, монолітна архітектура починає створювати певні проблеми. База даних стрімко розширюється, що може призвести до зниження продуктивності та ускладнення масштабування системи. Крім того, монолітна архітектура часто обмежує можливості незалежного розгортання та оновлення окремих компонентів системи, що уповільнює процес розроблення та ускладнює впровадження нових функцій [6].

У певний момент стає зрозуміло, що для забезпечення подальшого розвитку та ефективного функціонування інформаційної системи необхідно змінювати її архітектуру. Одним із популярних підходів є перехід до мікросервісної архітектури з розподіленою базою даних. Такий перехід дозволяє розбити монолітну систему на менші, незалежні сервіси, кожен з яких відповідає за певну функціональність та має власну базу даних. Це забезпечує кращу масштабованість, гнучкість та стійкість системи до збоїв.

Для розгляду процесу переходу від монолітної архітектури та монолітної бази даних до мікросервісної архітектури з розподіленою базою даних, ми використаємо приклад інформаційної системи для розвиваючого контролю використання мобільних пристроїв. Основна функція даної інформаційної системи полягає у можливості тимчасово блокувати доступ до певних мобільних додатків, поки користувач не вирішить поставлене перед ним завдання або головоломку. Ця функціональність дозволяє користувачам розвивати свої когнітивні здібності, логічне мислення та вирішувати цікаві завдання, перш ніж отримати доступ до бажаних додатків.

Однією з ключових особливостей системи є можливість для користувачів самостійно створювати та додавати нові завдання до бази даних. Це дозволяє інформаційній системі постійно розвиватися та розширюватися за рахунок внеску спільноти користувачів. Користувачі можуть проявляти свою креативність, розробляючи унікальні та захоплюючі завдання, якими потім зможуть користуватися інші учасники системи.

Завдяки цій функціональності, база даних інформаційної системи постійно збільшується та збагачується новим контентом. Це приваблює нових користувачів, які зацікавлені в різноманітних та складних завданнях, а також стимулює існуючих користувачів до активної участі в розвитку системи.

На рис.1 наведена спрощена схема бази даних для кращого розуміння функціоналу даної інформаційної системи.

Наше дослідження складалося з кількох етапів.

Аналіз інформаційної системи. Першим кроком було проведення аналізу початкової монолітної структури системи. Ми детально вивчили існуючу архітектуру, визначили основні компоненти системи та їх взаємозв'язки. Це дало нам чітке розуміння поточного стану системи та дозволило виявити проблеми та обмеження, які існують на даний момент.

Обов'язковою частиною даного етапу стало визначення основних потенційних проблем та обмежень, які можуть виникнути при активному розвитку даної інформаційної системи при збереженні монолітної архітектури. Ми проаналізували, як зростання функціональності та обсягу даних впливає на продуктивність, масштабованість та гнучкість системи.

Були виявлені такі потенційні проблеми, притаманні більшості монолітних інформаційних систем:

1. Складність розгортання та оновлення окремих компонентів. Монолітна архітектура ускладнює процес оновлення окремих частин системи. Будь-які зміни в одному компоненті вимагають перезбирання та розгортання всієї системи, що збільшує час та ресурси, необхідні для внесення змін, а це в свою чергу збільшує час, протягом якого сервіс є недоступним для користувачів.

2. Обмеження у горизонтальному масштабуванні та висока вартість масштабування. Монолітна архітектура має значні обмеження щодо горизонтального масштабування, тобто додавання нових екземплярів системи для опрацювання зростаючого навантаження. Кожен новий інстанс монолітної системи повинен мати достатньо ресурсів для забезпечення ефективного функціонування всіх компонентів інформаційної системи, що робить горизонтальне масштабування досить витратним. Вартість одного інстанса монолітної інформаційної системи є високою, оскільки він має бути достатньо потужним для опрацювання всіх запитів та забезпечення прийнятної продуктивності. Це призводить до неефективного використання ресурсів та значних фінансових витрат при необхідності збільшення масштабів інформаційної системи.

Однією з ключових вимог стало забезпечення відповідності регіональним нормативним актам щодо зберігання персональної інформації користувачів. Різні країни мають свої специфічні вимоги до збору, опрацювання та зберігання персональних даних. Наприклад, «Загальний регламент про захист даних (GDPR)» в Європейському Союзі вимагає, щоб персональні дані європейських користувачів зберігалися та опрацьовувалися в межах ЄС. Аналогічні вимоги існують і в інших регіонах світу.

Для забезпечення відповідності цим нормативним актам, необхідно було розробити архітектуру, яка дозволяла б зберігати персональні дані користувачів у тих географічних регіонах, де вони знаходяться. Це означає, що дані користувачів з Європи повинні зберігатися на серверах в Європі, дані користувачів зі США - на серверах у США і так далі. Така архітектура вимагає розподілу бази даних на окремі регіональні сегменти та забезпечення синхронізації даних між ними.

Крім того, географічний розподіл бази даних сприяє покращенню продуктивності та швидкості доступу до даних для користувачів з різних регіонів. Розміщення даних ближче до користувачів зменшить затримки в мережі та забезпечить більш швидке завантаження та оновлення інформації в додатку.

Якщо зробити звичайну повну копію середовища в різних локаціях, ми зіткнемося з проблемою роздільного використання бази даних завдань у цих локаціях. Кожна копія середовища матиме свою окрему базу даних завдань, які будуть розвиватися незалежно одна від одної. Це означає, що нові завдання, створені в одній локації, не будуть доступні в інших локаціях. Синхронізація та перенесення завдань між регіонами стане основним викликом для команди розробників.

Підтримка цілісності та консистентності бази даних завдань у такій розподіленій конфігурації буде надзвичайно складною операцією. Будь-які зміни, внесені в одній локації, доведеться вручну синхронізувати з іншими локаціями, що може призвести до конфліктів версій та втрати даних. Такий підхід також ускладнить процес розроблення та тестування нових завдань. Розробникам доведеться враховувати відмінності між базами даних завдань у різних локаціях, що може призвести до помилок та неузгодженостей. Тестування нових завдань доведеться проводити окремо для кожної локації, що збільшить час та ресурси, необхідні для випуску оновлень.

Тепер, після визначення ключових цілей та драйверів змін, ми можемо перейти до опису кроків, здійснених для успішного переходу від моноліту з єдиною базою даних до мікросервісної архітектури з вертикально розподіленою базою даних.

Визначення меж мікросервісів. Для декомпозиції інформаційної системи на менші, незалежні мікросервіси, ми проаналізували її функціональність та структуру бази даних (рис. 1). Для визначення меж мікросервісів доцільно використовувати методологію, засновану на понятті "обмежених контекстів" (bounded contexts) з предметно-орієнтованого проектування (domain-driven design, DDD) [2].

Обмежений контекст - це явна межа, в рамках якої існує модель домену. Всередині обмеженого контексту всі терміни та поняття мають специфічне значення, і модель відображає узгоджене уявлення про предметну область в рамках цього контексту.

Кожен обмежений контекст зазвичай відповідає певній частині бізнесу або підрозділу організації, що займається конкретним аспектом інформаційної системи. Обмежені контексти часто мають власні моделі даних, які можуть відрізнятися від моделей в інших контекстах.

Кожен мікросервіс повинен реалізовувати функціональність в рамках одного обмеженого контексту і не виходити за його межі. Це забезпечує високу зв'язність всередині сервісу та низьку зв'язність між сервісами.

На основі проведеного аналізу, ми виділили наступні "обмежені контексти", які можуть існувати незалежно один від одного як мікросервіси:

1. User Service - відповідає за управління користувачами інформаційної системи, їх аутентифікацію, авторизацію та збереження персональних даних. Цей сервіс відповідає за взаємодію з таблицями user, role, user_role.

2. Device Service - відповідає за управління мобільними пристроями користувачів, їх реєстрацію в системі та збереження інформації про встановлені додатки. Взаємодіє з таблицями device та application.

3. Task Service - відповідає за управління завданнями та головоломками в системі, їх створення, редагування, категоризацію та зберігання. Цей сервіс працює з таблицями task, category, task_tag, tag, version.

4. Solution Service - відповідає за процес розв'язання завдань користувачами, перевірку правильності відповідей та розблокування додатків. Взаємодіє з таблицями solution, session, unlock.

5. Feedback Service - відповідає за збір та збереження відгуків користувачів про завдання, їх оцінки та коментарі. Працює з таблицею feedback.

6. Recommendation Service - відповідає за підбір оптимальних завдань для користувачів на основі їх активності, рівня знань та налаштувань складності. Аналізує дані з різних сервісів та формує персоналізовані рекомендації.

7. Analytics Service - відповідає за збір та аналіз статистичних даних про активність користувачів, популярність завдань, ефективність розблокування додатків тощо. Агрегує дані з різних сервісів та надає аналітичну інформацію для прийняття рішень.

8. Notification Service - відповідає за надсилання сповіщень та повідомлень користувачам про нові завдання, досягнення, оновлення системи тощо.

Ітераційний процес декомпозиції. Декомпозиція функціональності монолітної інформаційної системи на мікросервіси - це поступовий та ітеративний процес (рис. 2). Не рекомендується намагатися переробити одразу весь моноліт на мікросервіси, оскільки це може призвести до хаосу та труднощів в управлінні процесом.

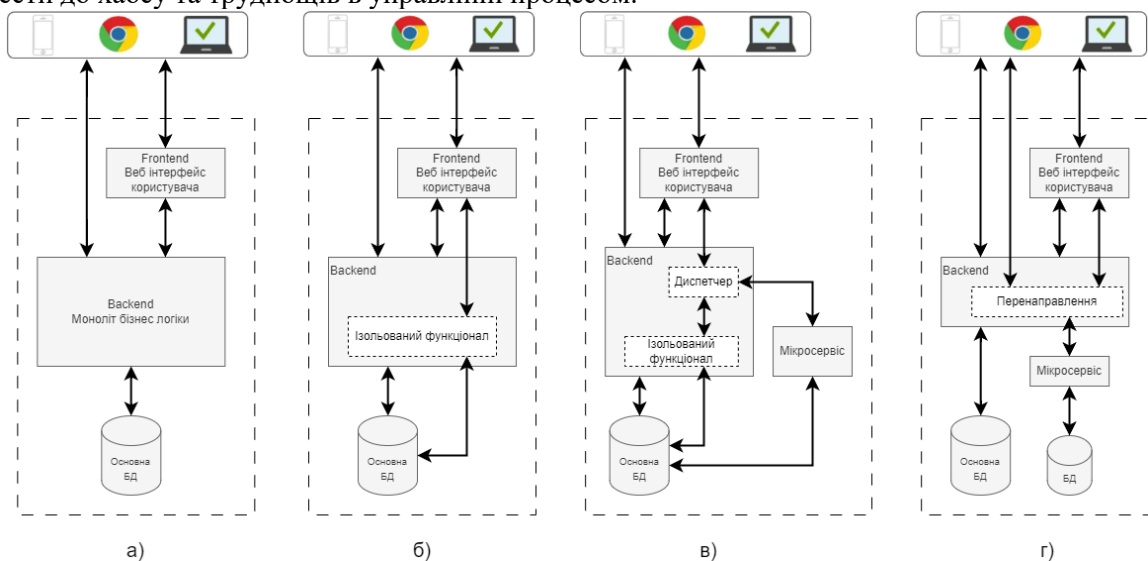


Рис. 2. Ітераційний процес декомпозиції частини функціоналу на мікросервіси:
а) початковий стан; б) ізоляція функціоналу; в) створення мікросервісу;
г) декомпозиція частини бази даних.

При виборі функціональності для здійснення декомпозиції, ми починаємо з найбільш незалежної та ізольованої частини інформаційної системи. У нашому випадку, ми вибрали Feedback Service як кандидата для першого мікросервісу.

Ізоляція функціоналу. Якщо код монолітної системи з самого початку був правильно спроектований з використанням принципів об'єктно-орієнтованого програмування (ООП), то вибрана частина функціональності буде вже достатньо ізольованою. Однак, якщо це не так, необхідно провести рефакторинг коду, щоб максимально ізольовати цю функціональність як незалежний модуль або аплікацію в межах моноліту (рис. 2б). Цей модуль повинен взаємодіяти з іншими частинами інформаційної системи тільки через чітко визначений інтерфейс. Саме цей інтерфейс згодом буде визначати інтерфейс нашого мікросервісу.

У нашому випадку, ми провели рефакторинг коду, пов'язаного з Feedback Service, та виділили його в окремий пакет Python разом з усіма необхідними тестами. Загальні утиліти та допоміжні функції, які використовувалися по всій системі, також були винесені в окремий пакет. Це дозволило нам пізніше легко перенести цей код в окремий мікросервіс.

Після завершення цього етапу, ми ретельно протестували монолітну систему, щоб переконатися, що вона функціонує належним чином і всі зміни не призвели до появи нових помилок або збоїв.

Створення мікросервісу. Наступним кроком стала побудова мікросервісу для вибраної функціональності. Оскільки необхідна функціональність була ізольована, ми легко перенесли код з

монолітної системи в окремий сервіс, зберігаючи при цьому зв'язок з базою даних моноліту. На цьому етапі важливо уникати внесення змін в існуючий функціонал моноліту, щоб запобігти виникненню функціональних розбіжностей між монолітом та новим мікросервісом.

Оскільки мікросервіс це незалежна частина нашої інформаційної системи, ми проводимо його повне незалежне тестування, щоб переконатися в коректності роботи та відповідності очікуваній поведінці.

Перевірка. На наступному етапі, ми переробляємо інтерфейсні методи в моноліті таким чином, щоб вони стали диспетчерами (dispatchers). Ці диспетчери дозволяють перемикаєти реалізацію функціональності між внутрішніми модулями моноліту та зовнішнім мікросервісом в залежності від налаштувань конфігурації системи. Це дає нам можливість поступово переходити на використання мікросервісу, не порушуючи роботу існуючої системи (рис. 2в).

Протягом певного періоду часу, функціональність буде існувати одночасно в двох варіантах - як мікросервіс та як внутрішній модуль моноліту. Це дозволить нам, у випадку виявлення будь-яких недоліків або проблем з мікросервісом, швидко повернутися до попередньої версії функціональності в моноліті, поки ми вносимо необхідні виправлення та покращення в мікросервіс.

Коли всі необхідні тести успішно завершені і ми переконалися, що мікросервіс функціонує коректно та надійно, ми можемо остаточно видалити відповідний функціонал з монолітної системи. При цьому, ми залишаємо тільки інтерфейсні методи, які перетворюють всі функціональні виклики на запити до мікросервісу. Це означає, що будь-які звернення до даного функціоналу в моноліті тепер будуть автоматично перенаправлятися до мікросервісу для опрацювання.

Після цього етапу, мікросервіс стає повноцінною частиною нашої інформаційної системи, а моноліт виступає в ролі маршрутизатора, який перенаправляє виклики на мікросервіс. Однак, він все ще не є незалежним, оскільки використовує спільну базу даних з монолітом. Наступним кроком є декомпозиція частини бази даних, за яку відповідає мікросервіс.

Декомпозиція домену бази даних. При декомпозиції, важливо забезпечити відповідність обмеженому контексту. База даних мікросервісу повинна містити тільки ті дані, які безпосередньо стосуються функціональності сервісу та його предметної області. Необхідно уникати включення даних, які виходять за межі відповідальності сервісу, щоб зберегти його автономність та незалежність. Це дозволяє мікросервісу зосередитися на своїй основній функції та мінімізувати залежності від інших частин інформаційної системи.

В мікросервісній архітектурі допускається денормалізація даних для забезпечення автономності сервісів. На відміну від традиційного підходу до проектування бази даних, де дані нормалізуються для уникнення дублювання, кожен мікросервіс може мати свою власну оптимізовану схему бази даних. Це означає, що деякі дані можуть бути продубльовані в різних сервісах, якщо це необхідно для забезпечення незалежності та продуктивності. Денормалізація даних дозволяє мікросервісу швидко отримувати необхідну інформацію без звернення до інших сервісів, що зменшує затримки та підвищує автономність.

Під час декомпозиції частини бази, ми стикнемося з новим викликом - забезпечення цілісності даних у розподіленій системі. У монолітній архітектурі, цілісність бази даних зазвичай забезпечується засобами самої системи управління базами даних (СУБД) за допомогою зовнішніх ключів. Однак, при переході до мікросервісної архітектури з незалежними базами даних, ми змушені реалізувати власні механізми для підтримки цілісності та узгодженості даних. Замість цього, рекомендується використовувати асинхронні механізми комунікації, такі як черги повідомлень або подій, для підтримки узгодженості даних між сервісами [3]. Для прикладу, коли з інформаційної системи видається користувач, ця подія повинна бути поширена на всі мікросервіси, які зберігають дані, пов'язані з цим користувачем. Кожен мікросервіс повинен отримати відповідне повідомлення і видалити всі дані користувача, за які він відповідає, зі своєї бази даних. Це може бути реалізовано за допомогою шини повідомлень (message bus) або інших механізмів асинхронної комунікації між сервісами.

Забезпечення цілісності даних у розподіленій інформаційній системі вимагає ретельного проектування та реалізації додаткових механізмів синхронізації та узгодженості даних. Це може включати використання розподілених транзакцій, компенсуючих операцій (compensating transactions), подійно-орієнтованої архітектури (event-driven architecture) тощо.

Після успішного розділення бази даних і реалізації механізмів забезпечення цілісності даних, наш мікросервіс стає повністю незалежним і самодостатнім компонентом системи. Він має

власну базу даних, яка оптимізована під його потреби, і може розвиватися та масштабуватися окремо від інших частин інформаційної системи (рис. 2г).

Ітерація завершена - переходимо до декомпозиції наступного функціонального блоку. Процес повторюється, доки моноліт не буде повністю розділений на мікросервіси. У колишньому коді моноліту залишаться лише методи, які перенаправляють виклики на відповідні мікросервіси. На цьому етапі моноліт можна замінити на API Gateway сервіс (рис. X), який буде слугувати єдиною точкою входу для клієнтських запитів та забезпечувати маршрутизацію до відповідних мікросервісів (рис. 3).

Кожен з мікросервісів в кінцевому результаті отримує свою власну базу даних, яка буде містити тільки ті таблиці та дані, які безпосередньо стосуються функціональності даного сервісу. Це означає, що кожен мікросервіс буде реалізовувати лише частину загальної схеми бази даних, яка відповідає його потребам та обов'язкам. Фізично це може бути окремий сервер бази даних, віртуальна машина або контейнер, залежно від обраної технології та інфраструктури. Кожен інстанс бази даних буде мати свої власні ресурси, налаштування та засоби управління.

Така архітектура забезпечує високий рівень незалежності між мікросервісами. Кожен сервіс має повний контроль над своєю базою даних і може вільно змінювати схему даних, оптимізувати запити та масштабувати базу даних відповідно до своїх потреб, не впливаючи на інші сервіси.

Це також дозволяє кожному мікросервісу обрати найбільш підходящу для нього технологію баз даних. Наприклад, один сервіс може використовувати реляційну базу даних, таку як PostgreSQL, для зберігання структурованих даних, тоді як інший сервіс може обрати документо-орієнтовану базу даних, таку як MongoDB, для зберігання неструктурованих даних.

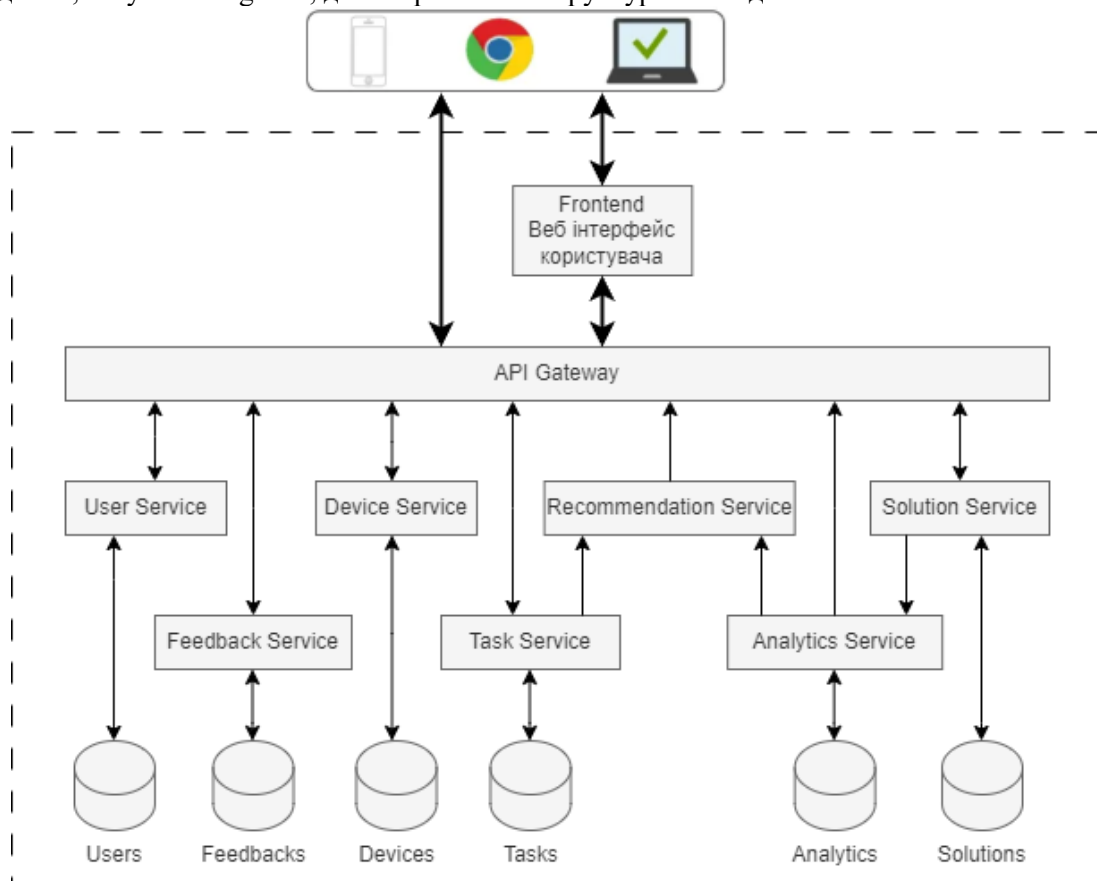


Рис. 3. Схема мікросервісної архітектури з розподіленою базою даних.

Незалежність баз даних мікросервісів також спрощує процес розгортання та масштабування системи. Кожен сервіс може бути розгорнутий та масштабований окремо, без необхідності координації з іншими сервісами. Це дозволяє більш ефективно використовувати ресурси та адаптуватися до змін навантаження.

Під час переходу до мікросервісної архітектури, ми провели ретельний аналіз та проектування API та протоколів обміну даними між сервісами. Для кожного мікросервісу ми визначили набір API-ендпоінтів, які дозволяють іншим сервісам та клієнтським додаткам

взаємодіяти з ним та отримувати необхідні дані або виконувати певні операції. Ми задокументували ці ендпоінти за допомогою специфікацій OpenAPI (Swagger) [8], що забезпечує чітке розуміння доступних методів, параметрів та форматів даних.

Для обміну даними між мікросервісами ми обрали формат JSON (JavaScript Object Notation), який дозволяє структуровано та ефективно передавати дані через мережу, незалежно від мови програмування чи технологічного стеку кожного сервісу.

Ми також реалізували додаткові протоколи та механізми для забезпечення безпечної та надійної комунікації між мікросервісами:

1. Ми використали HTTPS (Hypertext Transfer Protocol Secure) для шифрування та захисту даних під час передачі через мережу.

2. Для аутентифікації та авторизації запитів між сервісами ми впровадили використання JWT (JSON Web Tokens) [8].

3. Для асинхронної та подійно-орієнтованої взаємодії між сервісами ми використали Message Broker Apache Kafka. Це дозволяє сервісам публікувати події та підписуватися на них, забезпечуючи механізми обміну управляючою інформацією між різними мікросервісами.

В результаті проведеної роботи, ми успішно перейшли від монолітної архітектури до мікросервісної з розподіленою базою даних. Однак, незважаючи на досягнутий прогрес, ми все ще не досягли нашої основної мети - можливості використання спільної бази даних завдань у різних регіонах.

Географічний розподіл бази даних. Завдяки ізольованості кожного мікросервісу, ми можемо легко переконфігурувати Tasks Service таким чином, щоб він використовував не локальну базу даних, а спільну віддалену базу, яка знаходиться в іншому регіоні. Це дозволить нам досягти нашої цілі: завдання будуть спільними для всіх регіонів, тоді як приватні дані користувачів зберігатимуться у відповідній локації.

Проте, перш ніж розпочати таке об'єднання баз даних завдань з різних регіонів, необхідно ретельно продумати механізми забезпечення узгодженості даних. Фактично, ми плануємо об'єднати кілька незалежних баз даних в єдину горизонтально розподілену базу, що може призвести до виникнення певних проблем з цілісністю даних.

Розглянемо, наприклад, записи в таблиці завдань, які містять зовнішні ключі, що посилаються на користувачів, які завантажили ці завдання. В кожному регіоні є свій власний набір користувачів, і при об'єднанні баз даних можуть виникнути конфлікти ідентифікаторів. Два різних користувача з різних регіонів можуть мати однаковий ідентифікатор, що призведе до невідповідності зовнішніх ключів та порушення цілісності даних.

Для забезпечення коректної роботи об'єднаної бази даних, нам необхідно гарантувати унікальність ідентифікаторів для всіх сутностей незалежно від локації. Одним із можливих підходів є використання концепції multi-tenancy, де кожен ідентифікатор складається з двох частин: tenant ідентифікатора (ідентифікатора регіону в нашому випадку) та ідентифікатора сутності [9]. При цьому, всі основні та зовнішні ключі перетворюються на комплексні ключі, що складаються з двох значень.

Такий підхід надає значну перевагу - завжди можна легко визначити, якому регіону належить кожен запис у базі даних. Це дозволяє реалізувати ефективний роутинг запитів до інформаційної системи на основі ідентифікатора регіону. Коли надходить запит, інформаційна система аналізує ідентифікатор регіону в ключах і автоматично перенаправляє запит до відповідного регіонального сервера.

Завдяки цьому механізму, ми можемо забезпечити прозорість розподілу даних між регіонами для користувачів та інших компонентів системи. З точки зору клієнта, система виглядає як єдине ціле, в той час як внутрішньо дані розподілені та опрацьовуються у відповідних регіонах.

Якщо використання комплексних ідентифікаторів, з якоїсь причини не бажано, як альтернативу, можна запропонувати використання універсальних унікальних ідентифікаторів (UUID) для всіх сутностей у системі. UUID є 128-бітним значенням, яке генерується за допомогою спеціального алгоритму. Це стандартний формат ідентифікаторів, який гарантує унікальність не тільки в межах однієї інформаційної системи, але й у глобальному масштабі [10].

Змінивши тип всіх ідентифікаторів на UUID, ми можемо забезпечити унікальність ідентифікаторів для всіх сутностей у нашій розподіленій системі незалежно від регіону. Кожен мікросервіс зможе генерувати ідентифікатори для своїх сутностей самостійно, без необхідності координації з іншими сервісами.

На даному етапі в інформаційній системі залишається важливий недолік. Розміщення бази даних завдань лише в одному, умовно головному, регіоні, до якого віддалено підключаються інші регіони, може призвести до певних проблем з продуктивністю та масштабованістю системи.

По-перше, необхідно враховувати затримки при підключенні до географічно віддаленої локації. Якщо база даних завдань знаходиться в одному регіоні, а користувачі з інших регіонів звертаються до неї через мережу, це може призвести до значних затримок при опрацюванні запитів.

По-друге, концентрація великої кількості запитів в єдиній базі даних може перетворити її на "вузьке місце" (bottleneck) в інформаційній системі. Це може призвести до перевантаження бази даних, збільшення часу на отримання відгуку та потенційних збоїв. У результаті, затримки при опрацюванні запитів можуть стати неприйнятними для користувачів.

Крім того, з точки зору розгортання та підтримки інфраструктури, бажано, щоб усі регіони розгорталися однаково. Якщо кожен регіон має ідентичну структуру та конфігурацію, це спрощує процес розгортання нових регіонів та управління ними. Однакова інфраструктура полегшує моніторинг, оновлення та масштабування інформаційної системи, оскільки всі регіони мають однакові компоненти та налаштування.

Мікросервіс завдань на даному етапі виконує дві незалежні функції: повернення метаданих завдань для успішного функціонування мобільного додатку, функцію поповнення і модерації бази даних завдань. Операції редагування метаданих завдань не є критичними з точки зору швидкодії, оскільки виконуються набагато рідше, ніж операції читання. Тому було вирішено розділити Task Service на два мікросервіси: перший дозволяє лише читати дані про завдання (Task Reader Service), інший відповідає за адміністрування бази даних завдань (Task Admin Service).

Task Admin Service, розгортається разом з базою даних завдань в окремому центральному регіоні. Цей сервіс відповідає за додавання нових завдань, редагування існуючих завдань, модерацію контенту та інші адміністративні операції. Зміни, внесені в базу завдань, автоматично реплікуються засобами СУБД до локальних баз даних завдань у кожному регіоні, забезпечуючи глобальну актуальність даних.

Task Reader Service розгортається в кожному регіоні та взаємодіє з локальною read-only реплікою бази завдань. Цей сервіс опрацьовує запити від мобільних додатків та інших компонентів інформаційної системи, які потребують доступу до метаданих завдань. Завдяки локальному розміщенню бази даних, Task Reader Service забезпечує швидкий доступ до даних та мінімізує затримки при опрацюванні запитів.

Для забезпечення безпеки та авторизації доступу до Task Admin Service використовується механізм JWT авторизації. JWT токен генерується в тому регіоні, де авторизується адміністратор або модератор, після успішної перевірки їх облікових даних та прав доступу.

Кінцевий вигляд розподіленої бази даних для даної інформаційної системи зображено на рис.1.

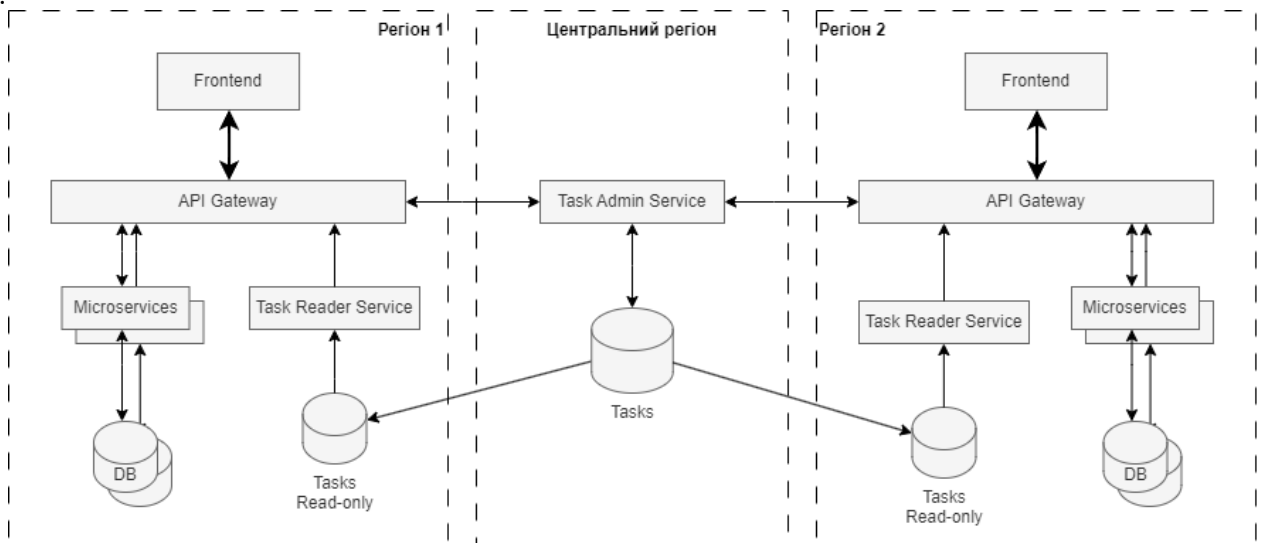


Рис. 4. Схема розподіленої інформаційної системи.

В результаті проведеної роботи, інформаційна система для розвиваючого контролю роботи мобільних пристроїв була успішно переведена з монолітної архітектури на мікросервісну з розподіленою базою даних. Система тепер відповідає всім поставленим цілям і вимогам:

1. Забезпечена можливість зберігання та опрацювання персональних даних користувачів у відповідності до регіональних нормативних вимог. Дані користувачів зберігаються в окремих базах даних у кожному регіоні, що дозволяє дотримуватися законодавства про захист персональних даних.

2. Забезпечена можливість використання спільної бази завдань у всіх регіонах.

3. Досягнута висока доступність та відмовостійкість інформаційної системи завдяки розподіленій архітектурі.

4. Досягнута висока продуктивність та низька затримка опрацювання запитів завдяки локальному розгортанню сервісів у кожному регіоні. Запити користувачів опрацьовуються в найближчому географічному регіоні, що мінімізує затримки та покращує швидкість відгуку інформаційної системи.

5. Забезпечена масштабованість та гнучкість інформаційної системи завдяки мікросервісній архітектурі. Це дозволить ефективно справлятися зі зростаючим навантаженням та швидко додавати нові функції.

Висновки. У статті запропоновано та досліджено метод переходу від монолітної архітектури до мікросервісної з розподіленою базою даних на прикладі інформаційної системи для розвиваючого контролю роботи мобільних пристроїв. Розроблений метод включає в себе етапи поступової декомпозиції монолітної системи на незалежні мікросервіси, розподілу бази даних між регіонами, а також реалізації механізмів для підтримки узгодженості та цілісності даних у розподіленій системі.

Головною перевагою запропонованого методу є можливість здійснювати всі зміни короткими ітераціями на працюючій інформаційній системі, без необхідності зупинки розроблення основної функціональності. Це дозволяє поступово та керовано переходити від монолітної до мікросервісної архітектури, не порушуючи роботу системи та не призупиняючи її розвиток.

Дослідження показали, що перехід до мікросервісної архітектури забезпечує високу модульність, масштабованість та адаптивність системи до змін бізнес-вимог. Отриманий досвід може бути корисним для розробників та архітекторів, які стикаються з подібними викликами при розробленні інформаційних систем.

Подальші дослідження будуть спрямовані на всебічний аналіз ефективності переходу до мікросервісної архітектури з розподіленою базою даних. Одним з ключових аспектів є аналіз економічної доцільності та обґрунтованості такого переходу для різних типів інформаційних систем та бізнес-вимог. Це включатиме детальну оцінку витрат на розроблення, розгортання та підтримку мікросервісної архітектури, а також порівняння цих витрат з потенційними перевагами та вигодами, які може принести така трансформація.

Також важливо дослідити вплив мікросервісної архітектури на процеси розроблення, тестування та розгортання програмного забезпечення. Потрібно проаналізувати, як перехід до мікросервісів впливає на швидкість розроблення, частоту релізів, якість коду та можливості автоматизації процесів. Це дозволить оцінити, наскільки ефективною є мікросервісна архітектура з точки зору підвищення продуктивності розроблення та скорочення часу виходу на ринок.

Окрім економічних факторів, важливо дослідити технологічну ефективність мікросервісної архітектури. Необхідно провести емпіричні дослідження та збір даних з реальних систем, щоб оцінити, наскільки ефективно мікросервісна архітектура справляється з різними навантаженнями та сценаріями використання.

Для формального обґрунтування ефективності переходу до мікросервісної архітектури з розподіленою базою даних необхідно розробити математичні моделі та провести кількісний аналіз. Це може включати моделювання продуктивності системи, аналіз масштабованості та надійності з використанням теорії масового обслуговування [11], а також застосування методів оптимізації для визначення оптимальної конфігурації та розподілу ресурсів у мікросервісній архітектурі.

Список бібліографічного опису

1. Драгоні Н. Мікросервіси: вчора, сьогодні та завтра / Н. Драгоні, С. Джіаллоренцо, А. Л. Лафунте, М. Маззара, Ф. Монтезі, Р. Мустафін, Л. Сафіна // Present and ulterior software engineering. – 2017. – С. 195-216. - DOI: <https://doi.org/10.1145/343477.343502>

2. Фаулер М. Мікросервіси: визначення нового архітектурного терміну [електронний ресурс] / М. Фаулер, Дж. Льюїс // martinfowler.com. – 2014. – Режим доступу: <https://martinfowler.com/articles/microservices.html>
3. Ньюмен С. Будівництво мікросервісів: проектування дрібнозернистих систем / С. Ньюмен. – O'Reilly Media, 2015. – 280 с. – ISBN 978-1491950357.
4. Прамод Д. Огляд стратегій розподілу та реплікації даних у розподілених базах даних / Д. Прамод, К. Венкатарамана, С. Фані Кумар // *International Journal of Advanced Computer Research*. – 2018. – Том 8, № 36. – С. 80-90. DOI: <http://dx.doi.org/10.30534/ijatcse/2019/117852019>
5. Брюер Е. На шляху до надійних розподілених систем / Е. Брюер // *Матеріали дев'ятого щорічного симпозиуму ACM з принципів розподілених обчислень*. – 2000. – С. 7. – DOI: <https://doi.org/10.1145/343477.343502>.
6. Ньюмен С. Від моноліту до мікросервісів: еволюційні шаблони для перетворення вашого моноліту / С. Ньюмен. – O'Reilly Media, 2019. – 272 с. – ISBN 978-1492047841.
7. Гролінгер К. Виклики великих даних: технології, моделі та парадигми / К. Гролінгер, В. А. Хігашіно, А. Тіварі, М. А. Капрец // *Управління даними в хмарних, грід та P2P системах*. – Springer, Berlin, Heidelberg, 2013. – С. 1-15. – DOI: https://doi.org/10.1007/978-3-642-40053-7_1.
8. Амундсен М. Проектування та створення чудових веб-API: надійних, стабільних та стійких / М. Амундсен. – Pragmatic Bookshelf, 2020. – 275 с. – ISBN 978-1680506808.
9. Вайлдер Б. Шаблони хмарної архітектури: використання Microsoft Azure / Б. Вайлдер. – O'Reilly Media, 2020. – 622 с. – ISBN 978-1492040675.
10. Пачеко Ф. Практична архітектура програмного забезпечення з C# 10 та .NET 6: проектування програмних рішень з використанням мікросервісів, DevOps та шаблонів проектування для Azure, 2-е видання / Ф. Пачеко. – Packt Publishing, 2022. – 650 с. – ISBN 978-1800566040.
11. Гаутам Н. Аналіз черг: методи та застосування / Н. Гаутам. – CRC Press, 2012. – 802 с. – ISBN 978-1439806586.

References

1. Dragoni N. Microservices: yesterday, today, and tomorrow / N. Dragoni, S. Giallorenzo, A. L. Lafuente, M. Mazzara, F. Montesi, R. Mustafin, L. Safina // *Present and ulterior software engineering*. – 2017. – P. 195-216. – DOI: <https://doi.org/10.1145/343477.343502>
2. Fowler M. Microservices: a definition of this new architectural term [internet resource] / M. Fowler, J. Lewis // martinfowler.com. – 2014. – Access mode: <https://martinfowler.com/articles/microservices.html>
3. Newman S. *Building Microservices: Designing Fine-Grained Systems* / S. Newman. – O'Reilly Media, 2015. – 280 p. – ISBN 978-1491950357.
4. Pramod D. A survey on data partitioning and replication strategies in distributed databases / D. Pramod, K. Venkataramana, S. Phani Kumar // *International Journal of Advanced Computer Research*. – 2018. – Vol. 8, No. 36. – P. 80-90. DOI: <http://dx.doi.org/10.30534/ijatcse/2019/117852019>
5. Brewer E. Towards robust distributed systems / E. Brewer // *Proceedings of the nineteenth annual ACM symposium on Principles of distributed computing*. – 2000. – P. 7. – DOI: <https://doi.org/10.1145/343477.343502>.
6. Newman S. *Monolith to Microservices: Evolutionary Patterns to Transform Your Monolith* / S. Newman. – O'Reilly Media, 2019. – 272 p. – ISBN 978-1492047841.
7. Grolinger K. Challenges in Big Data: Technologies, Models, and Paradigms / K. Grolinger, W. A. Higashino, A. Tiwari, M. A. Capretz // *Data Management in Cloud, Grid and P2P Systems*. – Springer, Berlin, Heidelberg, 2013. – P. 1-15. – DOI: https://doi.org/10.1007/978-3-642-40053-7_1.
8. Amundsen M. *Design and Build Great Web APIs: Robust, Reliable, and Resilient* / M. Amundsen. – Pragmatic Bookshelf, 2020. – 275 p. – ISBN 978-1680506808.
9. Wilder B. *Cloud Architecture Patterns: Using Microsoft Azure* / B. Wilder. – O'Reilly Media, 2020. – 622 p. – ISBN 978-1492040675.
10. Pacheco F. *Hands-On Software Architecture with C# 10 and .NET 6: Architecting software solutions using microservices, DevOps, and design patterns for Azure, 2nd Edition* / F. Pacheco. – Packt Publishing, 2022. – 650 p. – ISBN 978-1800566040.
11. Gautam N. *Analysis of Queues: Methods and Applications* / N. Gautam. – CRC Press, 2012. – 802 p. – ISBN 978-1439806586.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-04>

УДК 004.942

Багнюк Наталія Володимирівна, к.т.н., доцент

<https://orcid.org/0000-0002-7120-5455>

Христинець Наталія Анатоліївна, к.т.н., доцент

<https://orcid.org/0000-0002-4836-7632>

Биков Сергій Олегович, здобувач вищої освіти

кафедри комп'ютерної інженерії та кібербезпеки

Луцький національний технічний університет, м. Луцьк, Україна

РОЗРОБКА МУЛЬТИМЕДІЙНИХ 3D МОДЕЛЕЙ ДЛЯ СИМУЛЯЦІЇ ТЕХНОЛОГІЧНОГО ПРОЦЕСУ

Багнюк Н.В., Христинець Н.А., Биков С.О. Розробка мультимедійних 3D моделей для симуляції технологічного процесу. Задачі моделювання поведінки частинок сумішей є важливими для різних галузей промисловості, включаючи будівництво, матеріалознавство, хімічну промисловість та фармацевтику. У будівництві, наприклад, моделювання засипки частинок може бути корисним для дослідження властивостей бетонних сумішей або композитних матеріалів. У матеріалознавстві, моделювання взаємодії різних частинок може допомогти в розробці нових матеріалів з покращеними механічними або хімічними властивостями. У роботі були розглянуті комп'ютерні методи проектування 3D моделей частинок сталі та кераміки, що підкреслює широкий спектр можливостей застосування цих технологій в різних галузях. Особлива увага приділяється моделюванню текстур та полігонів, оскільки ці аспекти важливі для досягнення реалістичного візуального представлення частинок. Коректне відтворення текстур і деталей полігонів дозволяє забезпечити точність та достовірність симуляційних результатів. Використання Blender у поєднанні з налаштованим світлом, камерою та контейнером відкриває широкі можливості для створення інтерактивної інструментальної системи. Ця система може бути використана для вивчення та оптимізації технологічних процесів у промисловості, що сприяє підвищенню ефективності та якості виробництва. Такий підхід дозволяє не лише зменшити витрати та час на дослідження, але й стимулює інновації та покращення в галузях, де моделювання технологічних процесів має ключове значення.

Ключові слова: Blender, симулятор, технологічний процес, об'єкти, полігони, текстура, 3D модель

Bahniuk N., Khrystynets N., Bykov S. Development of multimedia 3D models for technological process simulation.

The problems of modeling the behavior of particles in mixtures are important for various industries, including construction, materials science, chemical industry, and pharmaceuticals. In construction, for example, the modeling of particle filling can be useful for investigating the properties of concrete mixes or composite materials. In materials science, modeling the interaction of different particles can help in the development of new materials with improved mechanical or chemical properties. The paper considered computer methods for designing 3D models of steel and ceramic particles, which emphasizes the wide range of possibilities of application of these technologies in various fields. Particular attention is paid to the modeling of textures and polygons, as these aspects are important to achieve a realistic visual representation of particles. Correct reproduction of textures and details of polygons allows to ensure accuracy and reliability of simulation results. Using Blender in combination with customized lights, camera and container opens up a wide range of possibilities to create an interactive tool system. This system can be used to study and optimize technological processes in industry, which contributes to increasing the efficiency and quality of production. This approach not only reduces research costs and time, but also drives innovation and improvement in industries where process modeling is key.

Keywords: Blender, simulator, technological process, objects, polygons, texture, 3D model

Постановка наукової проблеми. Моделювання та симуляції технологічних процесів є актуальними науковими проблемами, які поєднують в собі питання математичного моделювання, фізики, інформатики та інженерії [1]. Такого роду завдання є надзвичайно популярними у різних галузях науки і техніки: методами моделювання проектуються складні системи від космічних та літальних апаратів до мікро- та нано-частинок дисперсних середовищ. Задачі моделювання поведінки частинок сумішей актуальні для галузей будівництва, матеріалознавства, хімічної промисловості, фармацевтики. В роботі розглянуто комп'ютерні методи проектування 3D моделей частинок сталі і кераміки. Ці композиційні матеріали використовуються при проектуванні деталей машин та механізмів, фільтрів, будівельних елементів, сорбційних мембран тощо. Однією з основних задач симулювання поведінки таких сумішей є створення алгоритмів для реалістичного формування динаміки частинок під час засипки, враховуючи їх взаємодію та поведінку у вузьких об'ємах засипки.

Інша важлива складова проблеми полягає в інтеграції мультимедійних технологій для візуалізації симуляції, що дозволяє користувачам спостерігати процес в реальному часі. Використання сучасних графічних процесорів та технологій 3D-моделювання є необхідним для

досягнення високої якості візуалізації та швидкодії системи як з економічної точки зору, так і з екологічних питань. Тобто, окрім того, що процеси симуляції значно здешевлюють процес виготовлення пробних заготовок, вони ще й дозволяють не лише виявити проблеми та помилки, а й програмно вирішити такі питання. Розроблена система емулявання має бути гнучкою та масштабованою, щоб дозволяти налаштування параметрів процесу під різні типи матеріалів та розміри контейнерів.

Важливою науковою задачею є розробка методів валідації та перевірки точності симуляційної моделі, що включає порівняння результатів моделювання з експериментальними даними. Оптимізація алгоритмів для обробки великих обсягів даних також є невід'ємною частиною дослідження, оскільки процес засипки часто включає в себе мільйони частинок.

Аналіз досліджень. З огляду на літературні закордонні та українські джерела [2-4], навчальні матеріали, та інтернет-ресурси з'ясовано, що питання розробки мультимедійної системи симуляції технологічного процесу засипки частинок в циліндричний контейнер є предметом активних досліджень. Закордонні джерела надають детальні алгоритми моделювання таких систем і підходи до реалізації ефективних обчислювальних методів. Українські дослідники також вносять вагомий вклад, зосереджуючи увагу на оптимізації процесів та адаптації існуючих моделей до специфічних умов вітчизняної промисловості [5].

Сучасні навчальні матеріали висвітлюють основні принципи комп'ютерного моделювання та симуляції, що є основою для розробки таких систем [6]. Інтернет-ресурси надають велику кількість інформації щодо програмного забезпечення та інструментів, які можуть бути використані для створення мультимедійних симуляцій. В роботі [7] зазначено, що, наприклад, платформи Unity або Unreal Engine дозволяють створювати візуалізації високої якості, що є важливим для точного відображення процесу засипки частинок. Зокрема, літературні джерела [8-9] часто описують фізичні моделі, такі як методи дискретних елементів (DEM), які використовуються для симуляції взаємодії частинок.

Мета роботи полягає у створенні інтегрованих програмних елементів системи, що дозволить моделювати і візуалізувати процес засипки частинок у реальному часі з високим ступенем точності.

Вклад основного матеріалу. Вибір ПЗ Blender для розробки 3D моделей частинок обумовлений кількома вагомими причинами. По-перше, Blender є потужним і безкоштовним програмним забезпеченням з відкритим вихідним кодом, що робить його доступним для широкого кола користувачів та дослідників. По-друге, Blender має розвинені інструменти для 3D-моделювання та анімації, які дозволяють створювати реалістичні візуалізації процесів. Третє, можливості фізичної симуляції в Blender дозволяють точно моделювати поведінку частинок, враховуючи їх взаємодію та динаміку. Важливо також зазначити, що Blender підтримує інтеграцію з іншими програмними продуктами, що розширює його функціональні можливості для комплексних симуляцій, оскільки в подальшому проєкт реалізується на платформі Unity.

Базовою основою на початку проєктування 3D моделі будь-якого об'єкта в Blender є полігони частинок (рис. 1).

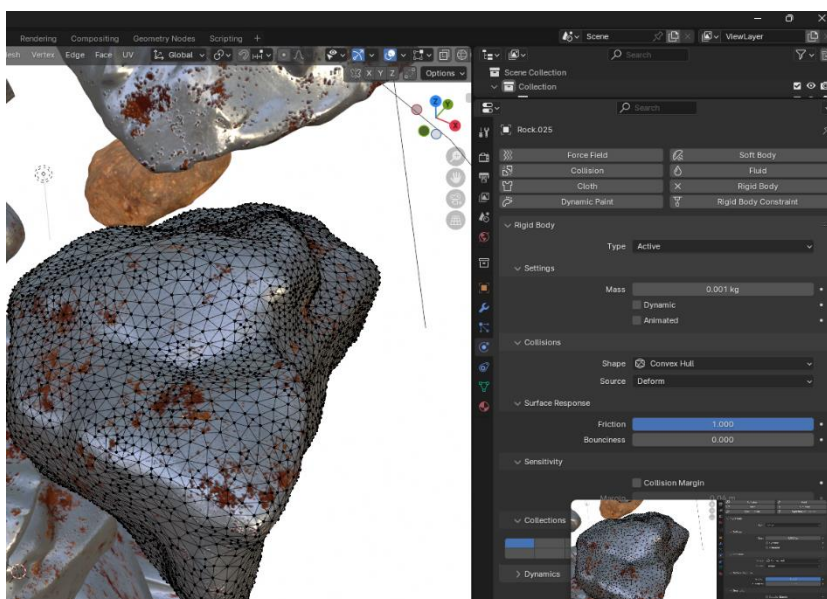


Рис. 1 – Проектування полігонів в Blender

Формуються вони шляхом почергового створення базової геометрії об'єкта за допомогою інструментів моделювання – кубів, сфер тощо. або ручне моделювання вершин, ребер і граней. Далі, для створення системи частинок, до моделі застосовується модифікатор Particle System, який дозволяє налаштувати параметри, такі як кількість частинок, їх розмір, швидкість та напрямок руху. Важливою частиною є використання модифікатора Collision, який дозволяє частинкам реагувати на поверхні об'єктів, з якими вони стикаються. Для більш реалістичної візуалізації частинки можна налаштувати за допомогою матеріалів і текстур, які задають їхній зовнішній вигляд.

Взаємозв'язки спроектованих текстур відтворюються схемою вузлів Node Editor (рис. 2), яка надає гнучкий інтерфейс для створення складних матеріалів і текстур. За допомогою вузлів можна комбінувати різні текстури, шейдери та ефекти, налаштовуючи їх взаємодію для досягнення бажаного результату. Кожен вузол виконує певну функцію, наприклад, задає кольори, додає блиск або визначає рельєф поверхні об'єкта.

Зв'язуючи виходи одного вузла з входами іншого, було створено частинки різних розмірів та текстур, що реагують на освітлення та інші умови сцени. Через цю схему можна інтерактивно налаштовувати текстури, миттєво бачити результати змін у вікні перегляду, що значно полегшує процес 3D моделювання.

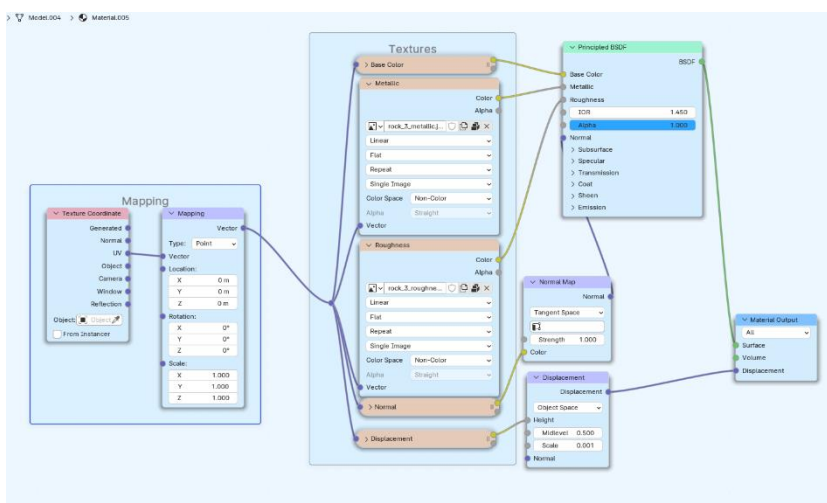


Рис. 2 – Схема вузлів Node Editor

Полігони та розгортки в розроблених 3D моделях використані для мінімізації швів і розтягувань, що забезпечує більш реалістичний і привабливий зовнішній вигляд часток. Це

виконано через інструментарій UV Mapping в Blender, що дозволяє точно проектувати 2D текстури на 3D об'єкти. Процес полягає в розгортці моделі, що означає розділення поверхні 3D об'єкта на плоску 2D сітку. В Blender існують різні методи розгортки, такі як Smart UV Project, Unwrap, Cube Projection та інші, які вибираються в залежності від форми об'єкта і складності текстури (рис. 3).

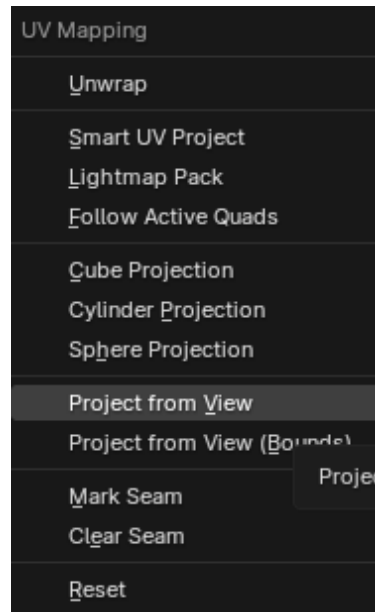


Рис. 3 – Панель UV Mapping та типи розгортки

В роботі використано розгортку Smart UV Project (рис. 4), яка є автоматизованим методом створення UV розгортки, що значно спрощує процес текстурування. Ця розгортка працює шляхом автоматичного розділення моделі на дрібні ділянки і їх оптимального розташування на UV карті з мінімальними спотвореннями. Smart UV Project особливо корисна для об'єктів зі складною геометрією, в тому числі для дендритних і пластинчастих неоднорідних структур, де ручне розгортання було б трудомістким та складним. Після застосування Smart UV Project, є можливість коригувати результуючу UV карту, якщо необхідно, для досягнення більш точного і бажаного вигляду.

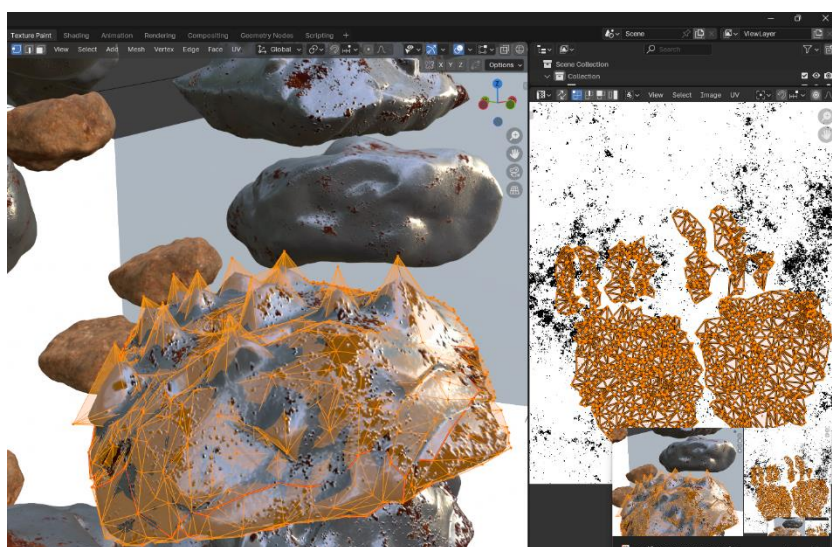


Рис. 4 – Розгортка Smart UV Project

Ключовим моментом симуляції є процес засипки суміші – проектування активів розроблених частинок. Даний процес включає налаштування камери сцени, світла сцени та параметрів самої системи частинок. Камера повинна бути розташована так, щоб забезпечити

оптимальний огляд процесу засипки, дозволяючи користувачам бачити взаємодію частинок з контейнером у реальному часі (рис. 5).

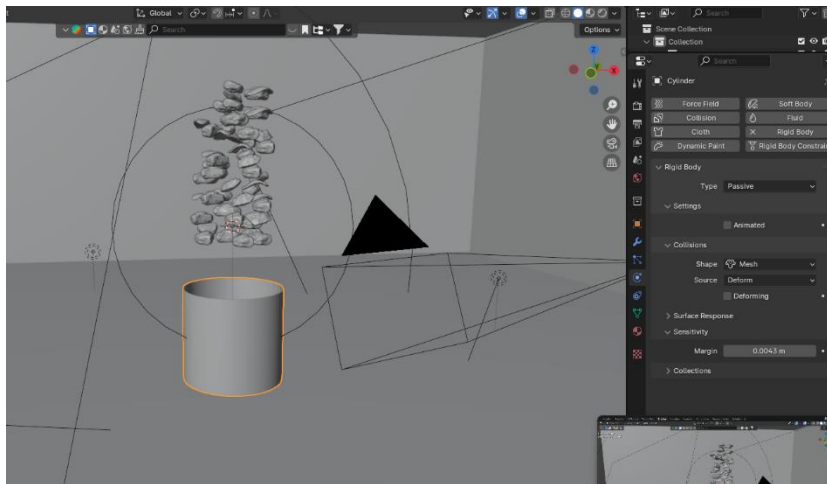


Рис. 5 – Налаштування камери та світла

Налаштування освітлення сцени є критичним для досягнення реалістичної візуалізації, підкреслюючи текстуру і динаміку частинок. Крім того, параметри системи частинок, такі як розмір, форма, швидкість і взаємодія між частинками, повинні бути точно налаштовані для достовірного відтворення фізичних властивостей суміші.

Висновки.

В роботі успішно реалізовано побудову 3D моделей. По-перше, за допомогою Blender були створені текстури та полігони частинок, що дозволило досягти бажаного зовнішнього вигляду для симуляції процесу засипки. Далі, було проведено налаштування світла та камери сцени, забезпечуючи оптимальне освітлення та огляд для користувачів під час перегляду симуляції. Також, для емуляції процесу засипки спеціально розроблений та інтегрований циліндричний контейнер, що дозволило відтворити реалістичні умови дослідження. Ці кроки сприяли створенню мультимедійної системи, яка через інтеграцію з Unity може ефективно моделювати технологічний процес засипки частинок, забезпечуючи користувачам зручний і реалістичний інтерфейс для спостереження та аналізу.

Список бібліографічного опису

1. The genetic evolution behavior of carbides in powder metallurgy FGH96 Ni-based superalloys / Liu Y. et al. Journal of Materials Science. №58. 2023. URL: <https://doi.org/10.1007/s10853-023-09161-4> (дата звернення: 01.03.2024).
2. Zabolotnyi O., Pasternak V., Andrushchak I., Ilchuk N., Svirzhevskiy K. Numerical Simulation of the Microstructure of Structural-Inhomogeneous Materials. Lecture Notes in Mechanical Engineering. 2020. P. 562-571.
3. Continuum percolation of the realistic nonuniform ITZs in 3D polyphase concrete systems involving the aggregate shape and size differentiation / J. Lin et al. Science and Engineering of Composite Materials. 2024. P. 161-182. <https://doi.org/10.1515/secm-2022-0237>
4. Microstructural evolution and mechanical properties of Inconel 718 superalloy thin wall fabricated by pulsed plasma arc additive manufacturing/ K. Wang et al. Journal of Alloys and compounds, 2020. P. 79-91 <https://doi.org/10.1016/j.jallcom.2019.152936>
5. MATLAB. MathWorks. Version R2024a. URL: <https://www.mathworks.com> (date of access: 27.02.2024).
6. Advanced Simulation Library. Avtech Scientific. Version 0.1.7. URL: <https://asl.org.il> (date of access: 27.02.2024).
7. John W. Eaton, Torsten Liljed. GNU Octave. GNU General Public License. Version 9.1.0. Free Software Foundation. URL: <https://www.octave.org> (date of access: 23.02.2024).
8. Долгов О. М. Композиційні матеріали. Дніпро : Дніпр. політехніка, 2022. 67 с. URL: <http://surl.li/swvwq> (дата звернення: 16.03.2024).
9. Алгоритм оцінювання якості 3D моделі для адитивного виробництва / В. Марчук та ін. Матеріали та технології в інженерії (МТІ-2023) : 36. наук. доп. міжнар. конф., м. Луцьк, 16-18 трав. 2023 р. Луцьк, 2023. С. 178-180. URL: <http://surl.li/swwge> (дата звернення: 13.02.2024).

References

1. The genetic evolution behavior of carbides in powder metallurgy FGH96 Ni-based superalloys / Liu Y. et al. Journal of Materials Science. №58. 2023. URL: <https://doi.org/10.1007/s10853-023-09161-4> (дата звернення: 01.03.2024).

2. Zabolotnyi O., Pasternak V., Andrushchak I., Ilchuk N., Svirzhevskiy K. Numerical Simulation of the Microstructure of Structural-Inhomogeneous Materials. Lecture Notes in Mechanical Engineering. 2020. P. 562-571.
3. Continuum percolation of the realistic nonuniform ITZs in 3D polyphase concrete systems involving the aggregate shape and size differentiation / J. Lin et al. Science and Engineering of Composite Materials. 2024. P. 161–182. <https://doi.org/10.1515/secm-2022-0237>
4. Microstructural evolution and mechanical properties of Inconel 718 superalloy thin wall fabricated by pulsed plasma arc additive manufacturing/ K. Wang et al. Journal of Alloys and compounds, 2020. P. 79-91 <https://doi.org/10.1016/j.jallcom.2019.152936>
5. MATLAB. MathWorks. Version R2024a. URL: <https://www.mathworks.com> (date of access: 27.02.2024).
6. Advanced Simulation Library. Avtech Scientific. Version 0.1.7. URL: <https://asl.org.il> (date of access: 27.02.2024).
7. John W. Eatond, Torsten Lilged. GNU Octave. GNU General Public License. Version 9.1.0. Free Software Foundation. URL: <https://www.octave.org> (date of access: 23.02.2024).
8. Долгов О. М. Композиційні матеріали. Дніпро : Дніпр. політехніка, 2022. 67 с. URL: <http://surl.li/swvwq> (дата звернення: 16.03.2024).
9. Алгоритм оцінювання якості 3D моделі для адитивного виробництва / В. Марчук та ін. Матеріали та технології в інженерії (МТІ-2023) : 36. наук. доп. міжнар. конф., м. Луцьк, 16–18 трав. 2023 р. Луцьк, 2023. С. 178–180. URL: <http://surl.li/swwge> (дата звернення: 13.02.2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-05>

УДК 004.021:621.391.8

Виннишин Олег Ярославович, спеціаліст

<http://orcid.org/0009-0006-7113-1501>

Київський національний університет імені Тараса Шевченка, м. Київ, Україна

КОМПЛЕКСНА МЕТОДИКА ОПТИМІЗАЦІЇ ПРОГРАМНОГО КОДУ VHDL ПРИ ПРОЕКТУВАННІ ПРОГРАМОВАНОЇ ІНТЕГРАЛЬНОЇ СХЕМИ

Виннишин О.Я. Комплексна методика оптимізації програмного коду VHDL при проектуванні програмованої інтегральної схеми. Проведено аналіз підходів, що використовуються при оптимізації інформаційної системи на основі програмованої інтегральної схеми, що описується програмними алгоритмами, розробленими з застосуванням засобів мови опису апаратури VHDL. Основні принципи оптимізації визначались через представлення логіки проекту та функціонального моделювання програмованої інтегральної схеми у відповідності до декларованого функціоналу та умов експлуатації. Перехід від високорівневого до низькорівневого представлення при проектуванні логічних блоків програмованої інтегральної схеми, що проводиться на етапі компіляції, дозволяє визначити кількість логічних елементів, а отже завдяки функціональному моделюванню оцінюється архітектура схеми та формуються методичні рекомендації. Застосування комплексної методики оптимізації на основі проведеного аналізу надає можливість збільшити швидкодію виконання поставленої задачі з застосування розробленої програмованої інтегральної схеми, уникнути ризики надмірного генерування логіки, спростити арифметичні операції та логічні вирази, видалити блоки програмного коду, що не виконуються внаслідок модифікації проекту, застосувати засоби паралелізації процесів, а також виконати структурування блоків програмного коду з метою зменшення ризику появи помилок при його компіляції. Наведений у рамках дослідження інструментарій включав у себе введення у програмний код додаткових змінних і констант, зменшення об'єму програмного коду у тілі циклу, видалення зайвих блоків програмного коду в умовних конструкціях, структурування арифметичних операцій та логічних виразів, а також належний поділ блоків програмного коду. Окрема задача полягає у оптимізації блоку арифметико-логічного пристрою програмованої інтегральної схеми, що надає можливість зменшити ресурсоемність, скоротити кількість логічних вентилів, зменшити ризик виникнення помилок за умови виникнення затримки при подачі сигналів операційного коду, а також налагодити спільне використання ресурсів для окремих складових логічних блоків. Розроблена методика, таким чином, стала основою для формування цілісного набору підходів по формуванню логічної схеми через оптимізацію програмного коду VHDL у відповідності до вимог, поставлених при проектуванні програмованих інтегральних схем

Ключові слова: програмована інтегральна схема, мова опису апаратури, функціональна модель, алгоритми оптимізації, арифметико-логічний пристрій, паралелізація процесів.

Vynnyshyn O. A Comprehensive Methodology for Optimizing Vhdl Program Code when Designing a Programmable Integrated Circuit. An analysis of approaches used in optimizing information systems based on field-programmable gate arrays described by software algorithms developed using VHDL is conducted. The fundamental principles of optimization were determined through the representation of the project logic and functional modeling of the FPGA in accordance with the declared functionality and operating conditions. The transition from high-level to low-level representation in the design of logical blocks of the FPGA, carried out at the compilation stage, allows for the determination of the number of logic elements. Consequently, functional modeling assesses the architecture of the circuit and forms methodological recommendations. The application of a comprehensive optimization methodology based on the conducted analysis enables the increase of task execution speed using the developed FPGA, avoids the risks of excessive logic generation, simplifies arithmetic operations and logical expressions, removes blocks of code that are not executed due to project modifications, applies process parallelization tools, and structures code blocks to reduce the risk of errors during compilation. The tools introduced within the study included the addition of variables and constants to the code, reduction of code volume within loop bodies, removal of unnecessary code blocks in conditional constructions, structuring of arithmetic operations and logical expressions, as well as appropriate partitioning of code blocks. A separate task lies in the optimization of the arithmetic logic unit block of the FPGA, which allows for the reduction of resource consumption, the reduction of the number of logic gates, the mitigation of error risks due to signal delay in opcode delivery, and the establishment of resource sharing among individual components of logical blocks. The developed methodology, thus, serves as the basis for forming a comprehensive set of approaches for generating a logical circuit through VHDL code optimization in accordance with the requirements set during FPGA design.

Keywords: FPGA, hardware description language, functional model, optimization algorithms, arithmetic logic unit, process parallelization.

Постановка проблеми. На сьогоднішній день високою актуальністю характеризується завдання організації комп'ютерних систем, що відповідає вимогам по інформаційній стійкості та рівню захищеності від зовнішніх загроз [1-4]. Зазначається, найбільш ефективно відповідне завдання може бути вирішено через проектування інформаційної системи на основі програмованої логічної інтегральної схеми (Programmable Logic Device; PLD). Необхідно зазначити, що впровадження PLD при організації масштабованих комп'ютерних систем характеризується адаптивністю до зміни набору задач завдяки перепрограмування базових функцій та відповідне конфігурування інформаційної системи [1, 2]. Це надає можливість побудувати гнучку систему

захисту, що базується на зміні алгоритмів обробки поточкових даних та, зокрема, впровадженні криптографічних процедур [3, 4]. Водночас, стійкість системи і оптимізація розподілу навантаження на обчислювальний ресурс забезпечується шляхом паралелізації процесів при роботі у режимі реального часу [5, 6], впровадженні протоколів спільного використання логічних блоків [7, 8], а також інтеграції логічних блоків з іншими компонентами PLD [9, 10]. У якості окремого класу PLD можна виділити програмовані логічні інтегральні схеми, архітектура яких визначається як набір програмованих користувачем вентиляльних матриць (Field-Programmable Gate Arrays; FPGA). Можна вказати, що FPGA характеризуються високою складністю структури, що відповідає широкому набору функціональних можливостей як окремого логічного пристрою, так і загальної інформаційної системи, що організується на її основі [11, 12]. Це включає у себе високі показники адаптивності, пов'язані з програмованістю повного набору логічних функцій засобами FPGA [11], а також ефективні підходи по паралелізації процесів передачі і обробки поточкових даних, що оптимізує інформаційну систему за цільовими показниками швидкодії і навантаження на обчислювальний ресурс при роботі у режимі реального часу [12].

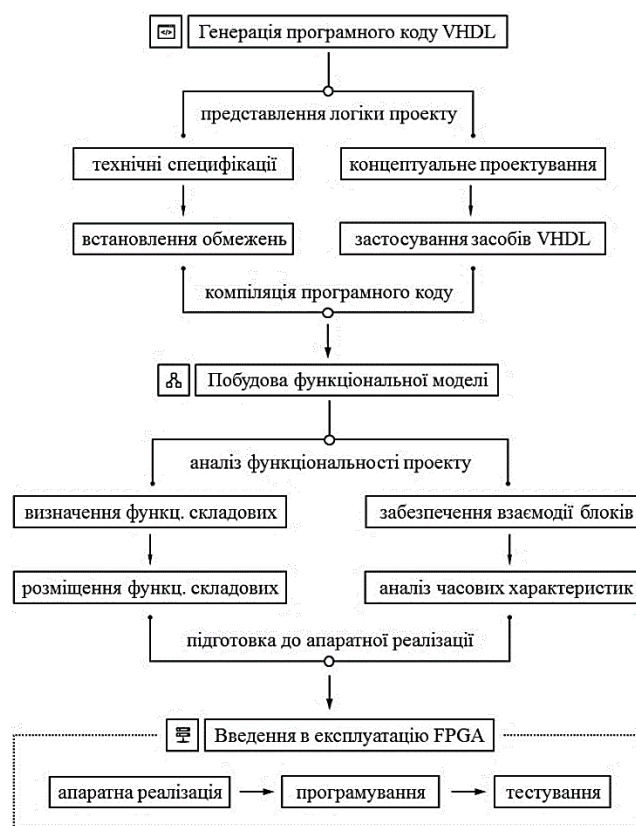


Рис. 1. Базова схема проектування FPGA

Як показав проведений аналіз, основні переваги PLD та, в особливості, FPGA, як окремого і найбільш функціонального класу PLD полягають у програмованості відповідних логічних пристроїв. Формалізація зазначеного процесу полягає у визначенні інструментарію мови опису апаратури (Hardware Description Language; HDL) та, зокрема, високошвидкісної мови опису апаратури, як то VHDL (Very-High-Speed Integrate Circuit Hardware Description Language), як спеціалізованої мови програмування дизайну FPGA, синтаксис якої буде розглянуто у даному дослідженні [13-15]. Впровадження інструментарію VHDL при проектуванні FPGA надає широкі можливості для проведення автоматичного аналізу та імітації робочих режимів експлуатації відповідно до експлуатаційних вимог. VHDL характеризується модульністю структури, що надає можливість реорганізації і спільного використання окремих блоків програмного коду, належний опис на основі VHDL забезпечує паралельне виконання набору завдань, а математичні моделі, що формуються засобами VHDL реалізують ефективне прогнозування результатів роботи у відповідності до заданих сценаріїв. Також важливою перевагою VHDL є стандартизація

програмного коду, що забезпечує можливість його портування без внесення змін та сумісність з інструментарієм, що використовується при організації широкого набору інформаційних систем.

Зазначені переваги формування PLD на основі наборів програмованих користувачем вентильних матриць з застосуванням інструментарію мови опису апаратури вказують на високу актуальність задачі побудови комплексної методики оптимізації програмного коду VHDL при проектуванні FPGA.

Аналіз останніх досліджень і публікацій. Найбільш актуальним підходом при проектуванні архітектури FPGA є формування комплексної методики, що базується на визначенні технічних специфікацій і функціональних можливостей відповідного логічного пристрою [14, 15]. При цьому комплексна методика проектування FPGA включає у себе виконання наступних етапів (рис. 1):

- представлення логіки проекту побудови FPGA на рівні генерації програмного коду VHDL;
- функціональне моделювання FPGA з визначенням продуктивності виконання операцій для різних режимів роботи;
- побудова FPGA на апаратному рівні, програмування логічного пристрою та тестування.

Таким чином, перехід від високорівневого до низькорівневого представлення при проектуванні логічних блоків FPGA, що проводиться на етапі компіляції, дозволяє визначити кількість логічних елементів, а завдяки функціональному моделюванню оцінюється архітектура схеми та формуються методичні рекомендації. На етапі функціонального моделювання оцінюється ефективність алгоритмів обробки та захисту поточних даних [3, 4, 15], визначаються підходи по паралелізації процесів, що виконуються логічними блоками FPGA [5, 6, 15], формуються протоколи спільного використання та інтеграції логічних блоків [7-10, 15]. При цьому невирішеною частиною загального дослідження є задача побудови комплексної методики оптимізації архітектури FPGA, що надає можливість сформулювати набір підходів по проектуванню схеми з застосуванням засобів VHDL.

Формулювання цілей статті. У відповідності до проведеного аналізу метою дослідження є визначення підходів по оптимізації програмного коду VHDL при проектуванні FPGA та формування на основі зазначених підходів цілісної методології проектування стійкої і масштабованої інформаційної системи.

Виклад основного матеріалу. У якості початкового прикладу оптимізації архітектури FPGA через співставлення високорівневого та низькорівневого представлення проекту пропонується розглянути реалізацію на основі програмованої інтегральної логічної схеми завдання розрахунку суми двох з величин з набору у залежності від контрольного сигналу. Формалізуємо завдання наступним чином: на основі набору поточних даних $\{x_1; x_2; x_3\}$, що поступають на вхід логічного пристрою у відповідності до контрольного двійкового значення x необхідно розрахувати вихідне значення, як $y = x_1 + x_2$ при $x = 1$ або $y = x_2 + x_3$ при $x = 0$. Засобами VHDL відповідна задача вирішується через створення проекту «P1» з чотирма вхідними значеннями «X_IN: in std_logic», «X1_IN: in std_logic», «X2_IN: in std_logic» і «X3_IN: in std_logic» та одним вихідним значенням «Y_OUT: out std_logic», а також застосування умовного оператора «if-then-else» у тілі проекту. Слід зазначити, що з метою спрощення програмного коду умовний оператор співставляє двійкову величину X_IN лише зі значенням «1», альтернативне значення вважається за «0» і для нього використовується команда «else».

На рис. 2 наведено співставлення лістингу відповідного програмного коду з функціональною моделлю, що генерується в результаті його компіляції. Аналіз функціональної схеми надає можливість відзначити надмірне генерування логічних елементів (один мультиплексор і два суматори), що не відповідає простоті як поставленої задачі, так і генерованого програмного коду. Процедура оптимізації при цьому має базуватись на модифікації високорівневого представлення FPGA.

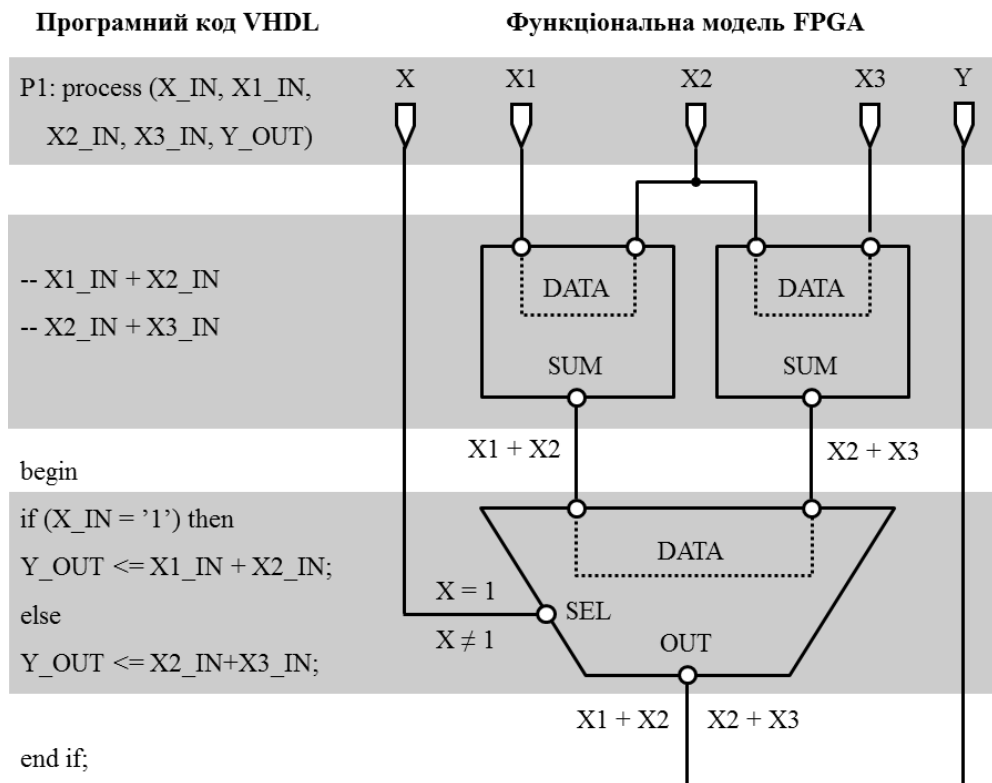


Рис. 2. Базова функціональна модель FPGA з суматорами і мультиплексором

Розглянемо модифікований проект «P1M», у рамках якого умовний оператор в залежності від контрольного значення X_IN обирає величину $X1_IN$ або $X3_IN$, яка вже надалі поза тіла умовного оператора додається до величини $X2_IN$. Очевидно, що результат вирішення поставленого завдання при виконанні програмного коду «PM» і «P1M» не відрізняється, але компіляція проекту «P1M» надає можливість отримати функціональну схему, яка містить лише один суматор (рис. 3), що дозволяє скоротити кількість логічних елементів у півтори рази.

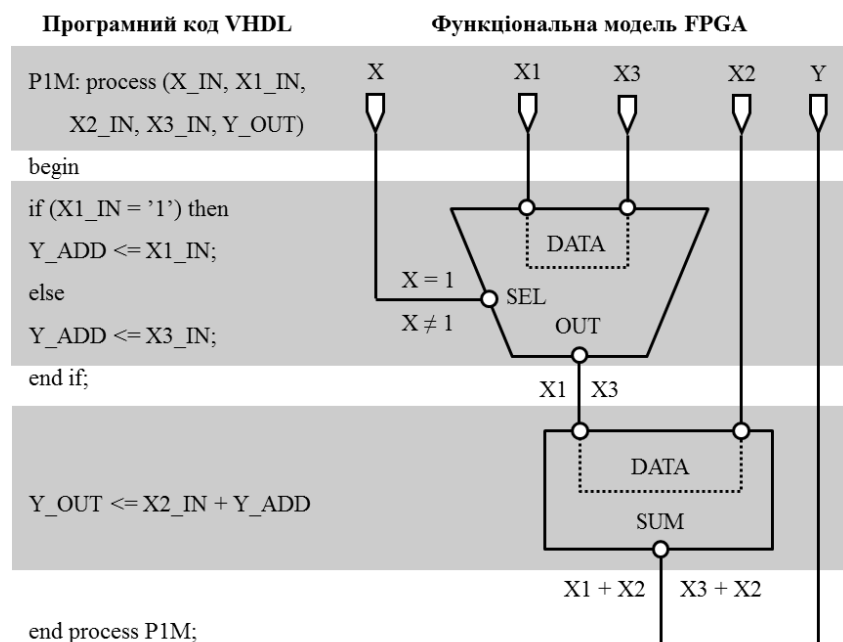


Рис. 3. Оптимізована функціональна модель FPGA з суматором і мультиплексором

Для складних архітектур FPGA відповідний підхід надає можливість суттєво зменшити загальну площу логічного пристрою, енергоспоживання та ризик виникнення помилок при його експлуатації.

Іншим прикладом, що пропонується розглянути у дослідженні, є оптимізація арифметико-логічного пристрою (Arithmetic Logic Unit; ALU) через паралелізацію процесів на рівні модифікації арифметичних операцій у програмному коді. Розглянемо задачу побудови функціонального блоку FPGA, що виконує задачу розрахунку вихідної величини на основі набору вхідних величин $\{x_1; x_2; x_3; x_4\}$ як $y = x_1 + x_2 - x_3 - x_4$. Задачу можна розглядати як тривіальну, що засобами VHDL вирішується через створення проекту «P2» з чотирма вхідними значеннями «X1_IN: in std_logic», «X2_IN: in std_logic», «X3_IN: in std_logic» і «X4_IN: in std_logic» та одним вихідним значенням «Y_OUT: out std_logic», а також застосуванням блоків суматорів і субтракторів, причому виконання арифметичних операцій описується як «Y_OUT <= X1_IN + X2_IN - X3_IN - X4_IN».

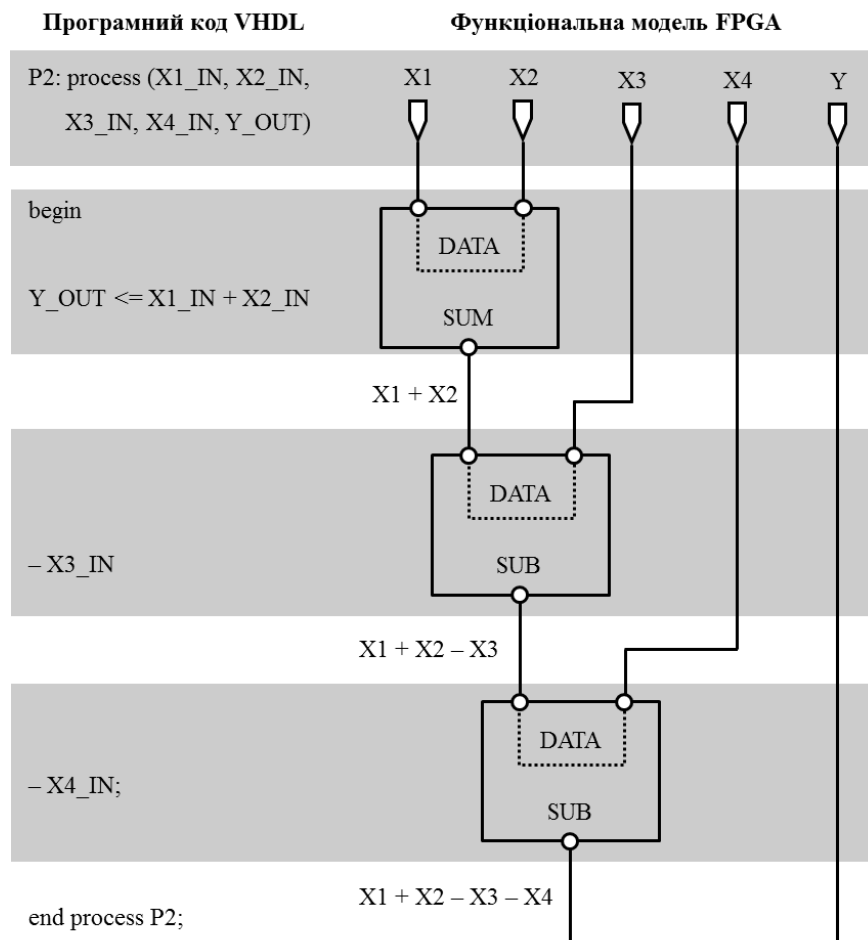


Рис. 4. Базова функціональна модель блоку ALU

Тим не менш, компіляція програмного коду надає можливість сформувати функціональну модель (рис. 4), яка свідчить про те, що відповідна арифметична операція виконується у три етапи, у той час як паралелізація надала б можливість суттєво збільшити швидкість виконання операцій блоком ALU. Засобами VHDL можна спростити вираз як «Y_OUT <= X1_IN + X2_IN - (X3_IN + X4_IN)». Компіляція програмного коду відповідного проекту «P2M» вказує на реорганізацію функціональної схеми таким чином, що два з трьох процесів виконуються паралельно, що збільшує швидкість пристрою у півтори рази.

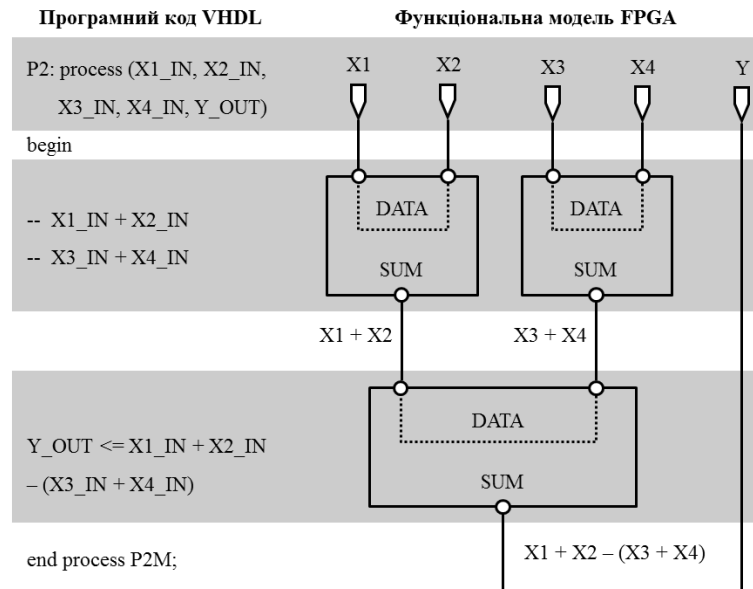


Рис. 5. Модифікована функціональна модель блоку ALU

Отже, можна зазначити, що співвіднесення низькорівневого і високорівневого представлення при проектуванні архітектури FPGA надає можливість суттєво збільшити продуктивність роботи логічного пристрою за цільовими показниками представленими на кількісному рівні, як то загальна площа схеми, кількість логічних елементів, швидкодія виконання операцій, а також ризик виникнення помилок. Застосування комплексної методики оптимізації програмного коду VHDL при проектуванні FPGA можливе через введення додаткових змінних та констант з метою спрощення операцій і виразів, виключення надмірного генерування логіки через зменшення об'єму програмного коду у тілі циклу та видалення блоків коду в умовних конструкціях, що не використовуються, а також структурування блоків програмного коду.

Висновки. Проведений аналіз надав можливість визначити набір підходів по оптимізації програмного коду VHDL при проектуванні FPGA та формування на основі зазначених підходів цілісної методології проектування стійкої і масштабованої інформаційної системи. У дослідженні розглянуто ряд підходів по спрощенню архітектури програмованої логічної інтегральної схеми та збільшенню швидкодії обробки поточкових даних та виконання операцій зазначених у специфікаціях FPGA.

Список використаних джерел

1. Alsabbagh, W., & Langendörfer, P. (2023). Security of programmable logic controllers and related systems: Today and Tomorrow. *IEEE Open Journal of the Industrial Electronics Society*, 4, 659–693. <https://doi.org/10.1109/ojies.2023.3335976>.
2. Chizek, M. (2019). Programmable logic device (PLD) safety design approach. *Journal of System Safety*, 55(1), 32–41. <https://doi.org/10.56094/jss.v55i1.54>.
3. Zhou, Y. (2016). Security analysis of a mutual-authentication cryptographic protocol based on Strand Space Model theory. *Revista Tecnica De La Facultad De Ingenieria Universidad Del Zulia*. <https://doi.org/10.21311/001.39.3.8>.
4. Almeida, J. B., Barbosa, M., Barthe, G., & Dupressoir, F. (2016). Verifiable side-channel security of cryptographic implementations: Constant-time Mee-CBC. *Fast Software Encryption*, 163–184. https://doi.org/10.1007/978-3-662-52993-5_9.
5. Timchenko, L. I., Petrovski, M. S., Kokriatskaia, N. I., & Denysova, A. E. (2014). Laser beam images classification methods with the use of parallel-hierarchical network running on PLD. *2014 X International Symposium on Telecommunications (BIHTEL)*. <https://doi.org/10.1109/bihTEL.2014.6987644>.
6. Kelly, M. S., Mayes, K., & Walker, J. F. (2017). Characterising a CPU fault attack model via run-time data analysis. *2017 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. <https://doi.org/10.1109/hst.2017.7951802>.
7. Babu, H. Md. (2022). Generic Complex Programmable Logic Device Board. *VLSI Circuits and Embedded Systems*, 299–310. <https://doi.org/10.1201/9781003269182-24>.
8. Kalkhof, T., & Koch, A. (2022). Direct device-to-device physical page migrations in Multi-FPGA shared virtual memory systems. *2022 32nd International Conference on Field-Programmable Logic and Applications (FPL)*. <https://doi.org/10.1109/fpl57034.2022.00043>.

9. Communication Networks, S. and. (2023). Retracted: Hierarchical network security measurement and optimal proactive defense in cloud computing environments. *Security and Communication Networks*, 2023, 1–1. <https://doi.org/10.1155/2023/9808414>.
10. Li, R., Decocq, B., Fang, Y., Zeng, Z., & Barros, A. (2022). A petri net-based model to study the impact of traffic changes on 5G network resilience. *Book of Extended Abstracts for the 32nd European Safety and Reliability Conference*. https://doi.org/10.3850/978-981-18-5183-4_s28-02-493.
11. Sunkavilli, S., & Yu, Q. (2022). Security threats and countermeasure deployment using partial reconfiguration in FPGA CAD Tools. *2022 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. <https://doi.org/10.1109/host54066.2022.9839731>.
12. Kumar Saha, S., & Bobda, C. (2020). FPGA accelerated embedded system security through hardware isolation. *2020 Asian Hardware Oriented Security and Trust Symposium (AsianHOST)*. <https://doi.org/10.1109/asianhost51057.2020.9358258>.
13. Sunkavilli, S., Zhang, Z., & Yu, Q. (2023). FPGA security: Security threats from untrusted FPGA CAD toolchain. *Frontiers of Quality Electronic Design (QED)*, 551–573. https://doi.org/10.1007/978-3-031-16344-9_14.
14. Ekonde Sone, M. (2020). A new cross-layer FPGA-based security scheme for wireless networks. *Computer and Network Security*. <https://doi.org/10.5772/intechopen.82390>.
15. Taraate, V. (2017). Erratum to: Pld Based Design with VHDL. *PLD Based Design with VHDL*. https://doi.org/10.1007/978-981-10-3296-7_12

References

1. Alsabbagh, W., & Langendörfer, P. (2023). Security of programmable logic controllers and related systems: Today and Tomorrow. *IEEE Open Journal of the Industrial Electronics Society*, 4, 659–693. <https://doi.org/10.1109/ojies.2023.3335976>.
2. Chizek, M. (2019). Programmable logic device (PLD) safety design approach. *Journal of System Safety*, 55(1), 32–41. <https://doi.org/10.56094/jss.v55i1.54>.
3. Zhou, Y. (2016). Security analysis of a mutual-authentication cryptographic protocol based on Strand Space Model theory. *Revista Tecnica De La Facultad De Ingenieria Universidad Del Zulia*. <https://doi.org/10.21311/001.39.3.8>.
4. Almeida, J. B., Barbosa, M., Barthe, G., & Dupressoir, F. (2016). Verifiable side-channel security of cryptographic implementations: Constant-time Mee-CBC. *Fast Software Encryption*, 163–184. https://doi.org/10.1007/978-3-662-52993-5_9.
5. Timchenko, L. I., Petrovski, M. S., Kokriatskaia, N. I., & Denysova, A. E. (2014). Laser beam images classification methods with the use of parallel-hierarchical network running on PLD. *2014 X International Symposium on Telecommunications (BIHTEL)*. <https://doi.org/10.1109/bihotel.2014.6987644>.
6. Kelly, M. S., Mayes, K., & Walker, J. F. (2017). Characterising a CPU fault attack model via run-time data analysis. *2017 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. <https://doi.org/10.1109/hst.2017.7951802>.
7. Babu, H. Md. (2022). Generic Complex Programmable Logic Device Board. *VLSI Circuits and Embedded Systems*, 299–310. <https://doi.org/10.1201/9781003269182-24>.
8. Kalkhof, T., & Koch, A. (2022). Direct device-to-device physical page migrations in Multi-FPGA shared virtual memory systems. *2022 32nd International Conference on Field-Programmable Logic and Applications (FPL)*. <https://doi.org/10.1109/fpl57034.2022.00043>.
9. Communication Networks, S. and. (2023). Retracted: Hierarchical network security measurement and optimal proactive defense in cloud computing environments. *Security and Communication Networks*, 2023, 1–1. <https://doi.org/10.1155/2023/9808414>.
10. Li, R., Decocq, B., Fang, Y., Zeng, Z., & Barros, A. (2022). A petri net-based model to study the impact of traffic changes on 5G network resilience. *Book of Extended Abstracts for the 32nd European Safety and Reliability Conference*. https://doi.org/10.3850/978-981-18-5183-4_s28-02-493.
11. Sunkavilli, S., & Yu, Q. (2022). Security threats and countermeasure deployment using partial reconfiguration in FPGA CAD Tools. *2022 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*. <https://doi.org/10.1109/host54066.2022.9839731>.
12. Kumar Saha, S., & Bobda, C. (2020). FPGA accelerated embedded system security through hardware isolation. *2020 Asian Hardware Oriented Security and Trust Symposium (AsianHOST)*. <https://doi.org/10.1109/asianhost51057.2020.9358258>.
13. Sunkavilli, S., Zhang, Z., & Yu, Q. (2023). FPGA security: Security threats from untrusted FPGA CAD toolchain. *Frontiers of Quality Electronic Design (QED)*, 551–573. https://doi.org/10.1007/978-3-031-16344-9_14.
14. Ekonde Sone, M. (2020). A new cross-layer FPGA-based security scheme for wireless networks. *Computer and Network Security*. <https://doi.org/10.5772/intechopen.82390>.
15. Taraate, V. (2017). Erratum to: Pld Based Design with VHDL. *PLD Based Design with VHDL*. https://doi.org/10.1007/978-981-10-3296-7_12

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-06>

УДК 004.5, 004.9

Волинець Леонід Васильович, магістрант

<https://orcid.org/0009-0007-2006-4465>

Гарматюк Назар Андрійович, аспірант

<https://orcid.org/0009-0005-2780-2032>

Дерев'яно Володимир Сергійович, магістрант

<https://orcid.org/0009-0005-9300-2455>

Дуда Олексій Михайлович, к.т.н, доцент

<https://orcid.org/0000-0003-2007-1271>

Крамар Тарас Олександрович, асистент

<https://orcid.org/0000-0001-8060-0169>

Скалецький Петро Олегович, аспірант

<https://orcid.org/0009-0007-4543-3714>

Тернопільський національний технічний університет, м. Тернопіль, Україна

ФОРМУВАННЯ КОНЦЕПТУ МУЗЕЙНОГО МОБІЛЬНОГО ЗАСТОСУНКУ З ЕЛЕМЕНТАМИ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

Волинець Л.В., Гарматюк Н.А., Дерев'яно В.С., Дуда О.М. Крамар Т.О., Скалецький П.О. Формування концепту музейного мобільного застосунку з елементами доповненої реальності. В епоху стрімкого технологічного розвитку, музеї мають унікальні можливості для вдосконалення способів подання інформації та залучення уваги широкого кола відвідувачів. Одним з інноваційних шляхів досягнення цієї мети є створення мобільних застосунків. Вони стають ключовим інструментом для збагачення досвіду відвідувачів, відкриваючи нові горизонти для взаємодії між відвідувачами та музейними експозиціями, вдосконалюють роботу музеїв, даючи змогу досліджувати культурну та історичну спадщину більш пізнавальними та інтерактивними способами. Мобільні застосунки сприяють створенню персоналізованих траєкторій та маршрутів відвідувачів, поглибленню вивчення історії експонатів, залученню технологій доповненої реальності для формування унікальних вражень. Крім того, вони можуть спростити адміністративні завдання, надаючи працівникам музею інструменти для ефективнішого управління та оновлення інформації про експозиції. Такий інтегрований підхід до мобільних технологій сприяє зближенню музеїв із сучасною аудиторією, роблячи культурне просвітництво захопливим та доступним. У цьому дослідженні проаналізовано інноваційні мобільні застосунки провідних музеїв світу, визначено їхні переваги та недоліки. На основі отриманих даних запропоновано концепт мобільного застосунку з елементами доповненої реальності, спрямованого на покращення досвіду відвідувачів музею та оптимізацію роботи персоналу. Використання у концепті інноваційних інформаційних технологій сприяє створенню цілісного та новаторського підходу до вивчення та збереження культурної спадщини у музейному середовищі.

Ключові слова: бази даних, діаграма класів, доповнена реальність, інформаційні технології, культурна спадщина, мобільний застосунок, музеї, проєктування, прототип.

Volynets L., Harmatiuk N., Derevianko V., Duda O., Kramar T. Skaletskyi P. Development of a museum mobile application concept with elements of augmented reality. In the era of rapid technological development, museums have unique opportunities for improving the ways of presenting information and attracting the attention of a wide range of visitors. One of the innovative ways to achieve this goal is the creation of mobile applications. They become a key tool for enriching the experience of visitors, opening new horizons for the interaction between visitors and the exhibition, improving the work of the museum, enabling the exploration of cultural and historical heritage in more informative and interactive ways. Mobile applications contribute to the creation of personalized trajectories and routes, deepening the study of the history of exhibits, and the involvement of augmented reality technologies to create unique experiences. In addition, they can simplify administrative tasks by giving museum staff the tools to more effectively manage and update exhibit information. This integrated approach to mobile technology brings museums closer to modern audiences, making cultural education engaging and accessible. This study analyzed the innovative mobile applications of the world's leading museums, identified their advantages and disadvantages. Based on the obtained data, a concept of a mobile application with elements of augmented reality aimed at improving the experience of museum visitors and optimizing the work of the staff is proposed. The use of innovative information technologies in the concept contributes to the creation of a holistic and innovative approach to the study and databases, class diagram, augmented reality, information technology, cultural heritage, mobile application, museums, design, prototype preservation of cultural heritage in the museum environment.

Key words:

Постановка наукової проблеми. На даний час переважна більшість музеїв стикається із складнощами впровадження, використання та оптимізації інформаційних та комунікаційних технологій для покращення взаємодії з відвідувачами. Нерівномірний розподіл можливостей доступу громадян до сучасних мобільних застосунків призводить до диференціації досвіду та зменшення інклюзивності. Водночас, технічні обмеження, висока вартість розробки та супроводу ускладнюють впровадження мобільних застосунків для невеликих музейних установ. Тому на даний

час сформована потреба розробки стратегій ефективного впровадження та використання інформаційних технологій та доповненої реальності у музейному середовищі.

Аналіз досліджень. Впродовж останнього періоду часу відбувається активна цифрова трансформація, що впливає на різні аспекти суспільного життя, включаючи музеї. [1] В умовах формування інформаційного суспільства [2] музеї змушені впроваджувати новаторські підходи до взаємодії з відвідувачами та змінювати методи подання та поширення інформації щодо своїх колекцій. Сучасний ринок пропонує різноманітні програмно-алгоритмічні комплекси та застосунки для музеїв, сприяючи впровадженню інновацій у їхню роботу [3].

Прогрес у царині інформаційних та комунікаційних технологій та зростаюче використання Інтернету спрощує установам культури доступ до цільової аудиторії. Водночас музеї стали чутливішими до визначення цільової аудиторії та подання експозицій широким колам відвідувачів, активніше використовуючи інформаційні та комунікаційні технології для підвищення їх доступності та зрозумілості [4].

Сучасні інформаційні та комунікаційні технології, зокрема проєктори, інтерактивні застосунки та симулятори, що використовуються для музейних експозицій, зумовлюють необхідність цифрової трансформації демонстраційних площ музеїв. Створення високопродуктивної комунікаційної інфраструктури, поширення безперешкодного доступу до Wi-Fi та впровадження спеціалізованого програмного забезпечення допомагає відвідувачам активно досліджувати музеї та виставки [5].

У процесі розвитку інформаційних та комунікаційних технологій і програмно-алгоритмічних комплексів, відбувається активна трансформація мобільних технологій на туристичні путівники [6]. Завдяки мобільним застосункам з елементами доповненої реальності, відвідувачі матимуть можливість інтерактивно взаємодіяти з музейними експозиціями у новий, захоплюючий спосіб, збагачуючи свій культурний досвід [7]. В перспективі інтерактивні інструменти можуть стати ключовим елементом підвищення експозиційної привабливості музеїв та залучення нових груп відвідувачів.

Музеї стикаються з широким переліком нових потреб, що пов'язані зі збільшенням відвідувачів та інформаційних запитів. Це сприяє вдосконаленню інформаційних технологій з використанням віртуальної та доповненої реальності у мобільних застосунках. Технології доповненої реальності надають унікальні можливості для інтеграції віртуального та реального середовища, даючи відвідувачам музеїв змогу насолоджуватися покращеним віртуальним досвідом при взаємодії з історичними об'єктами [8].

Мета роботи. Метою дослідження є формування комплексного концепту мобільного музейного застосунку з розширеними характеристиками зручності, інтерактивності та персоналізації на основі аналізу існуючих застосунків провідних світових музеїв. Зокрема, акцент зроблено на використанні елементів доповненої реальності для формування інтерактивного досвіду відвідувачів, формуванні рекомендацій щодо функціональних особливостей та вимог до мобільного застосунку, зручності використання інтерфейсу та оновлення інформації про експонати працівниками музею.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. В ході проведеного дослідження проаналізовано мобільні застосунки, що використовуються у провідних музеях світу.

Застосунок Art of London: augmented gallery (Лондон, Велика Британія) [9] використовує технологію доповненої реальності (англ. Augmented Reality, AR) для відображення картин та скульптур у галереї. Зокрема він дає змогу користувачам переглядати 3D-анімації та взаємодіяти з інтерактивними елементами, які доповнюють твори мистецтва. Відвідувачі галереї можуть прослухати аудіо чи переглянути відео, які розкривають цікаві факти та історії про твори мистецтва та художників. Art of London: augmented gallery перетворює ознайомлення, вивчення та розуміння сучасного мистецтва в захопливий та поглиблений процес.

Застосунок The British Museum (Лондон, Велика Британія) [10] дає змогу користувачам досліджувати колекції музею за темою, періодом або місцем походження. Присутні функції пошуку експонатів на карті музею, а також за темою, назвою або часовим періодом. Застосунок надає можливість формувати маршрути подорожей по музею. Користувачі можуть прослуховувати аудіогіди для широкого переліку експонатів музею, а також відвідати віртуальні виставки та переглянути віртуальні експонати. Застосунок пропонує користувачам ознайомитися зі статтями про експонати та події, або переглянути календар майбутніх подій.

Explorer (Нью-Йорк, США) [11] – це віртуальний провідник по Американському музею природознавства. Він пропонує інтерактивну карту музею, що полегшує процеси пошуку експонатів та музейних зручностей. Застосунок надає текстові описи, фотографії та відео супровід експонатів, а також функціональні можливості тематичного пошуку за назвою або історичним періодом. Є можливість віртуальних відвідин виставок та перегляду експонатів музею. Застосунок Explorer має функціональні можливості рекомендувати відвідувачам експонати та виставки, ґрунтуючись на попередніх результатах пошуку та переглядів. Застосунок пропонує інтерактивний календар музейних подій, зокрема лекцій, семінарів та виставок. Користувачі можуть віртуально взаємодіяти з деякими експонатами музею, наприклад, обертати 3D-моделі або здійснювати інтерактивні симуляції. Застосунок Explorer пропонує доступ до фото-, аудіо- та відеоматеріалів про експонати та виставки, а також навчальні ресурси для дорослих та дітей, наприклад статті, ігри та вікторини.

Застосунок Louvre Buddy (Париж, Франція) [12] є комплексним мобільним гідом по Лувру, надаючи розширені можливості для ознайомлення з музейними колекціями. Він дає користувачам можливість досліджувати колекції за різними критеріями, формувати власні маршрути відвідування та прослухати аудіосупровід для широкого переліку експонатів. Завдяки функціональним можливостям віртуальних турів, відвідувачі можуть ознайомитись з виставками, експонатами, отримати доступ до статей та інформацію про події в Луврі. Louvre Buddy є важливим інструментом для планування візиту до музею, збагачуючи досвід відвідувачів в процесі вивчення мистецтва та історії музею.

Застосунок Natural History Museum (Лондон, Велика Британія) [13] забезпечує можливість досліджувати колекції музею за темою, періодом або місцем походження. Можна здійснювати пошук експонатів на віртуальній карті музею. Застосунок допомагає формувати маршрути по музею для оптимізації часу відвідувачів. У застосунку доступний функціонал для прослуховування аудіогідів для широкого набору експонатів музею, а також для відвідин віртуальних виставок. Застосунок пропонує ігрові сценарії, які, в інтерактивній формі, допомагають більше дізнатися про музей, експозиції та артефакти.

Застосунок Skin and Bones (Смітсонівський національний музей, Вашингтон, США) [14] використовує доповнену реальність (AR), щоб "оживити" скелети тварин. Даний функціонал дає користувачам можливість побачити, як би виглядали тварини з м'язами, шкірою та хутром. Застосунок дає змогу взаємодіяти з 3D-моделями тварин, наприклад, обертати або збільшувати масштаб, отримати інформацію про анатомію тварин та особливості їх адаптації до навколишнього середовища. Увагу застосунку акцентовано на експонатах Смітсонівського національного музею, що пов'язані з анатомією та еволюцією тварин.

Порівняльну характеристику функціональних можливостей проаналізованих мобільних музейних застосунків подано в таблиці 1.

Таблиця 1 – Функціональні особливості музейних застосунків.

Застосунок	Карта експонатів	Мультимедійна інформація про експонати, галереї	AR-функції	Віртуальні екскурсії	Загальна інформація про музей чи проєкт	Інформація щодо придбання квитків	Можливість залишити відгук	Рекомендації відповідно до інтересів користувачів
Art of London: augmented gall	+	+	+	-	+	-	-	-
The British Museum	+	+	-	+	+	+	+	-
Explorer	+	+	-	+	+	+	+	+
Louvre Buddy	+	+	-	+	+	+	+	-
Natural History Museum	+	+	-	+	+	+	+	-
Skin and Bones	+	+	+	-	+	-	-	-

Проаналізувавши зібрані в таблиці 1 дані можемо зробити висновки, що:

- У кожному з проаналізованих застосунків є карта експонатів, мультимедійна інформація про експонати та галереї, а також загальна інформація про музей чи проєкт.
- Оскільки застосунки з AR-функціоналом в основному сфокусовані саме на відображенні цифрових елементів засобами доповненої реальності, тому в них доступні обмежені функціональні можливості порівняно з іншими застосунками.
- Лише в одному із проаналізованих застосунків наявні рекомендації за інтересами користувачів. Проте на це слід звернути увагу, оскільки дані функціональні можливості значно покращують досвід використання застосунків [15].
- Інформація про квитки та екскурсії є в більшості проаналізованих застосунків, що вказує на важливість даних функцій.

На основі проведеного аналізу особливостей подамо перелік функціональних можливостей концептуального прототипу музейного застосунку з елементами доповненої реальності у виді множини.

$$F(MP, MI, AR, RC, RS),$$

де MP – функції віртуальної карти музею для пошуку експонатів, визначення локації користувача, прокладання маршрутів. MI – функції інформування відвідувачів про експонати, експозиції та мультимедійний супровід об'єктів музею. AR – набір функцій з елементами доповненої реальності для відображення віртуальних елементів. RC – множина рекомендаційних функцій на основі пошукових запитів чи індивідуальних траєкторій відвідувачів. RS – функціональний набір щодо роботи установи: квитки, відпочинкові зони, зручності тощо.

Після аналізу застосунків і їх функціоналу можна перейти до формування загального концепту музейного застосунку з елементами доповненої реальності

Діаграму використання концепту музейного застосунку елементами доповненої реальності у форматі use-case актора «Користувач» подано на рис 1.

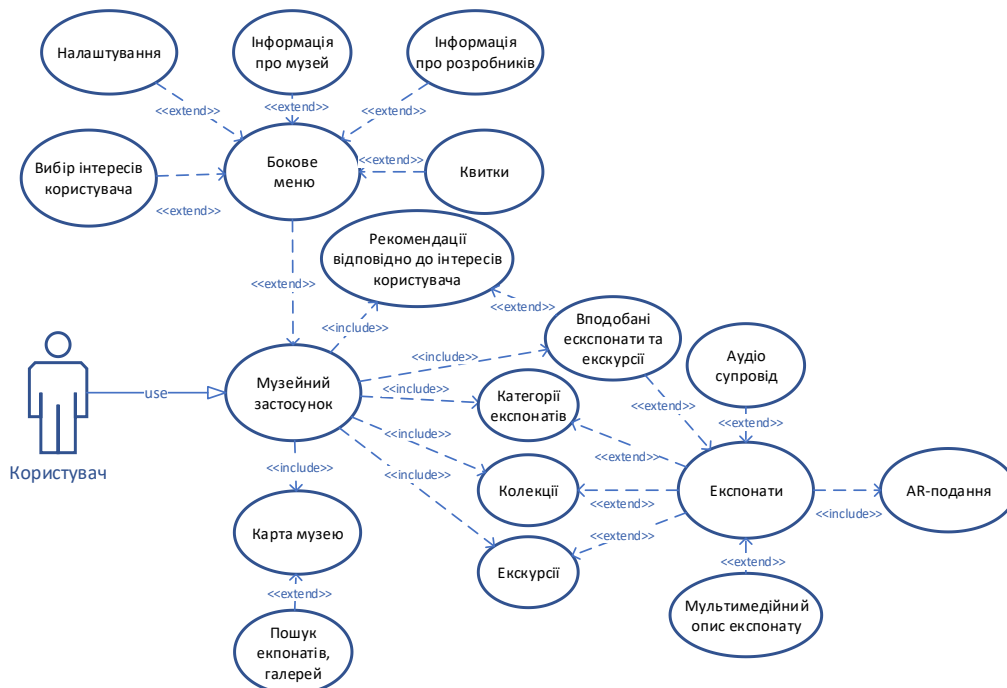


Рис.1 – Use-case діаграма актора «Користувач» концепту музейного мобільного застосунку з елементами доповненої реальності

Взаємодія з актором «Користувач» починається з головного меню, де відображені основні елементи навігації та рекомендації щодо екскурсій за замовчуванням. До основних елементів навігації належать: перегляд категорій, колекцій та експонатів, список вподобаних екскурсій та експонатів, карта музею. При перегляді експонатів актор «Користувач» може отримати інформацію про конкретний експонат у вигляді текстового опису, зображень та аудіосупроводу. AR-подання надає можливість взаємодіяти з віртуальними елементами та досліджувати експонати в

інтерактивному режимі. Актор «Користувач» може додавати екскурсії та експонати до списку вподобань, щоб повторно не шукати їх. Сформований список вподобань використовується для надання рекомендацій. За допомогою карти музею можна орієнтуватися між галереями та експонатами, формувати траєкторії переходів, знаходити відпочинкові зони та зручності тощо.

Діаграма використання концепту музейного мобільного застосунку з елементами доповненої реальності у форматі use-case актора «Працівник музею» подана на рис.2.



Рис.2 – Use-case діаграма актора «Працівника музею» концепту музейного мобільного застосунку з елементами доповненої реальності

Актори «Працівник музею» мають можливість отримувати звіти та ініціювати процедури аналітичного опрацювання даних про дії користувачів, зокрема, які експонати, галереї та маршрути користуються найбільшою популярністю, як користувачі переміщуються по музею, скільки часу вони проводять біля експонатів. Актори «Працівник музею» можуть редагувати інформацію про експонати, додавати та видаляти об'єкти на карті. В концептуальному прототипі мобільного музейного застосунку з елементами доповненої реальності доцільно створити функціональні можливості для взаємодії з користувачами. При цьому «Працівник музею» може переглядати та відповідати на відгуки користувачів, надсилати їм електронні листи з новинами про музей, майбутні подіями, новими експонатами тощо, а також давати відповіді на запитання громадян, які можуть виникнути під час використання мобільного застосунку.

На рис. 3 подано узагальнену структуру бази даних концепту мобільного застосунку з елементами доповненої реальності.

Таблиця Mob_AR_User містить інформацію щодо персональних даних користувачів, зокрема, універсальний ідентифікатор облікового запису користувача, його ім'я, адресу електронної пошти, хеш пароля тощо. До пов'язаних таблиць відносяться:

- Mob_AR_Visit – зберігає інформацію про візит користувача до музею, зокрема, дату, час і тривалість візиту.
- Mob_AR_Viewed_exhibits – містить відомості про експонати, які переглянули користувачі, час перегляду конкретного експонату та ID користувача, що здійснював перегляд.

- Mob_AR_Favorite_exhibits – містить список вподобаних користувачами експонатів та відомості про користувача, якому належить конкретний список.



Рис.3 – Узагальнена структура бази даних концепту музейного мобільного застосунку з елементами доповненої реальності

Крім того таблиця Mob_AR_User пов'язана з таблицями:

- Mob_AR_User_feedback – зберігає інформацію про відгуки, надіслані користувачем. Основними полями є зміст повідомлення та оцінки, які залишив користувач, тип відгуку та дата надсилання.
- Mob_AR_User_trajectory – зберігає дані про траєкторії користувачів та час витрачений на проходження конкретної траєкторії.
- Mob_AR_Routes – містить відомості про шляхи, які обирали користувачі під час їхнього відвідування музею, популярність різних шляхів та середній час проходження конкретного шляху.

- Mob_AR_Coordinates – містить набори координат, які використовуються для визначення обраних користувачами шляхів.

Таблиця Mob_AR_Exhibits зберігає інформацію про експонати відповідно до рекомендацій загальноєвропейської колекції 3D-моделей спадщини Europeana [16], зокрема, унікальний ідентифікатор експоната, опис, фізичні характеристики, AR-елементи, супровідні матеріали тощо. Ці атрибути дають змогу детально описати кожен експонат та його характеристики. До пов'язаних таблиць відносяться:

- Mob_AR_Collections – містить дані щодо приналежних до колекції експонатів. Записи характеризуються назвою, описом, кількістю експонатів та ідентифікатором працівника музею, який курує колекції. Curator (Staff_ID).

- Mob_AR_Exposition – містить відомості про назву, опис та тематику експозиції, час її початку та завершення, а також список колекцій, які виставлені для демонстрації в конкретній експозиції.

- Mob_AR_Locations – зберігає дані про локації музею, де знаходяться експозиції музею, їх місткість та рівень доступності для людей з обмеженими можливостями.

Таблиця Mob_AR_Museum_staff містить інформацію про працівників музею, зокрема унікальний ідентифікатор, його ім'я, електронну пошту, пароль, номер телефону, адресу та відомості щодо входу в застосунок.

З цією таблицею Mob_AR_Museum_staff пов'язані таблиці:

- Mob_AR_Roles – зберігає дані про ролі працівників музею, їх права, привілеї, та відділ, до якого належить працівник музею.

- Mob_AR_Feedback_reply – містить інформацію про сформовані працівниками відповіді на відгуки користувачів.

- Mob_AR_Events – зберігає відомості про події, які будуть відбуватися в музеї, зокрема, опис, тип, бюджет та ID відповідального за організацію працівника музею.

- Mob_AR_News – містить дані інформацію про автора, зміст та публікацію новин.

- Mob_AR_Analytics_report, яка відображає дані аналітичних звітів, що були згенеровані працівниками музею, а саме вміст, дату генерації, тип звіту, метод генерації звіту, короткий зміст звіту коментарі та ID працівника, що згенерував звіт.

Узагальнена діаграма класів концепту музейного мобільного застосунку з елементами доповненої реальності умовно розділена на дві частини. Класи для обслуговування експонатів і класи для обслуговування облікових записів та дій користувачів. На рис.4 подано множину класів для обслуговування експонатів концепту мобільного застосунку з елементами доповненої реальності. Основним батьківським класом є клас Exhibits. Він містить атрибути, що описують загальні характеристики експонатів музею: ID, назву, автора, опис, категорію, мультимедіа матеріали, рік створення та стан. Методи цього класу використовуються для додавання, оновлення, видалення експонатів та перегляду деталей експонатів.

Від основного класу походять два дочірніх класи:

- Viewed_exhibits – містить атрибути view_duration та view_date для опису тривалості перегляду експонату та дати і часу перегляду. Методи в даному класі реалізують можливості для додавання, редагування, перегляду та видалення переглянутих експонатів.

- Favorite_exhibits – використовується для реалізації методів додавання, редагування та видалення експонатів в списку вподобань.

Кожен експонат належить до колекції. Методи та атрибути пов'язані колекціями реалізовано в класі Collection. Окрім атрибутів імені та опису наявний атрибут Number_of_exhibits, який використовується для збереження числа експонатів, що належать до конкретної колекції. Методи даного класу реалізують можливості для створення, редагування, перегляду та видалення колекцій, а також методи для додавання чи видалення експонатів в колекції. Колекції виставляються для демонстрації в межах експозиції. Методи та атрибути необхідні для роботи з експозиціями реалізовано в класі Exposition. До унікальних атрибутів даного класу можна віднести:

- collection_list – список колекцій, які є частиною експозиції;
- theme – тематика експозиції;
- start_date – дата початку експозиції;
- end_date – дата завершення експозиції.

Методи даного класу реалізують можливості для створення, редагування, перегляду та видалення експозицій, а також методи для додавання чи видалення колекцій в експозиції.

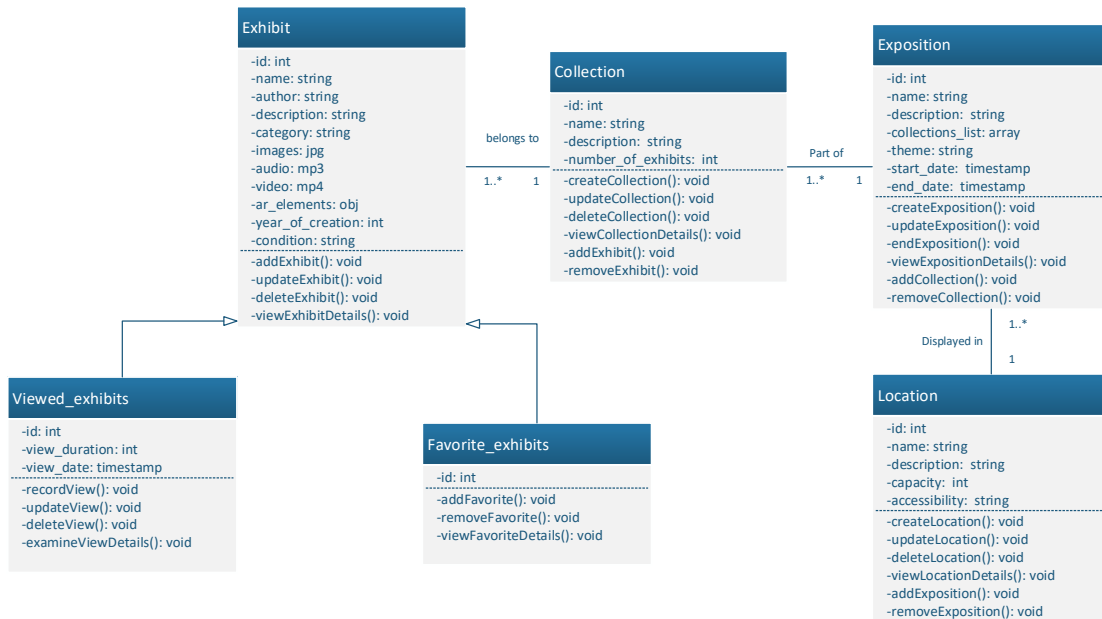


Рис.4 – Множина класів для обслуговування експонатів концепту мобільного застосунку з елементами доповненої реальності

Експозиції виставляються для демонстрації в певній локації музею. Методи та атрибути пов'язані з локаціями реалізовано в класі Location. Унікальними атрибутами цього класу є: *capacity*, який містить інформацію про місткість локацій та *accessibility*, що описує рівень доступності локації для осіб з обмеженими можливостями. Методи класу реалізують можливості для створення, редагування, перегляду та видалення інформації про локації, а також методи для додавання чи видалення експозицій на локації.

На рис.5 подано множину класів для обслуговування облікових даних користувачів концепту мобільного застосунку з елементами доповненої реальності. В даній множині класів основним батьківським є клас User. Він містить атрибути, що описують загальні характеристики користувачів музейного застосунку: ID, ім'я, email, пароль, дату реєстрації, стать, дату народження і час останнього входу в систему.

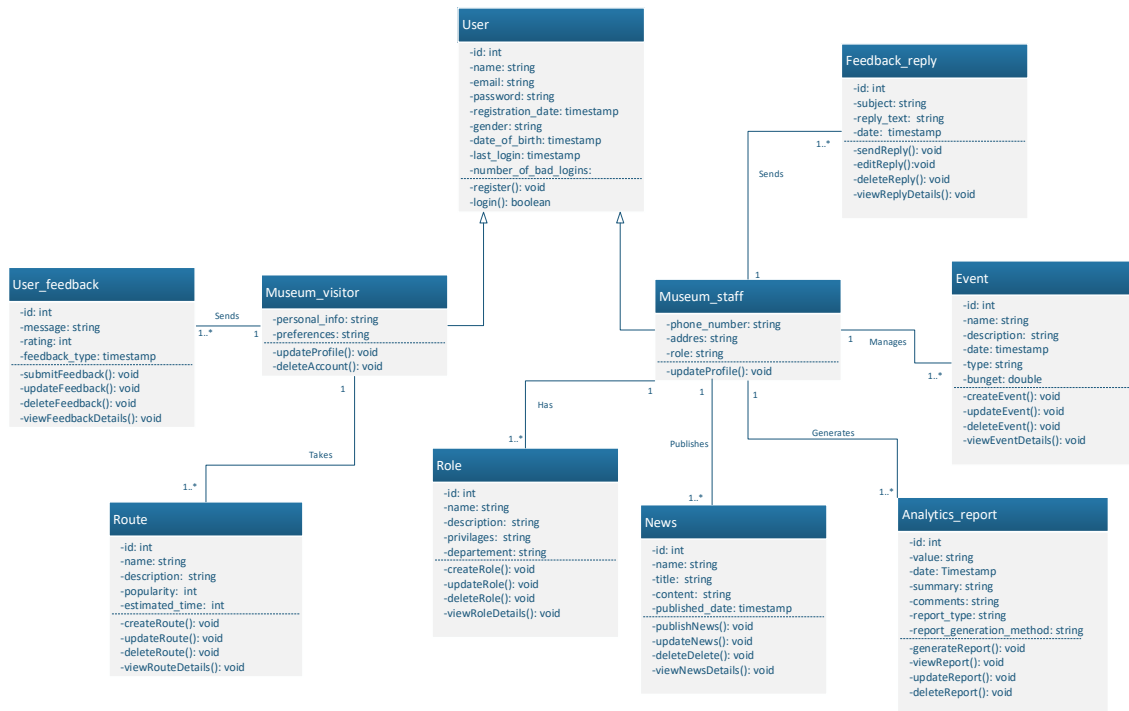


Рис.5 – Класи для облікових даних користувачів концепту мобільного застосунку з елементами доповненої реальності

Методи цього класу передбачають можливості реєстрації та аутентифікації в застосунку. Одним з дочірніх класів є клас *Museum_visitor*. Окрім методів і атрибутів успадкованих від батьківського класу він містить атрибути, що описують персональні відомості про користувачів-відвідувачів музею та їхні вподобання. Методи класу реалізують можливості для оновлення даних облікового запису та його видалення. З даним класом пов'язаний клас *User_feedback*. Він містить набори атрибутів та методів, що реалізують функціональні можливості щодо надсилання відгуків користувачами застосунку.

Клас *Route* містить атрибути, що описують маршрути користувачів в музеї: назву, опис, популярність маршруту, час його проходження. Методи цього класу реалізують можливості для створення, редагування, перегляду та видалення маршрутів. Іншим дочірнім класом є клас *Museum_staff*. Окрім успадкованих атрибутів і методів, він містить унікальні атрибути, що зберігають телефон, адресу та роль працівника музею. Методи цього класу надають функціональні можливості для оновлення даних про працівника музею. З класом *Museum_staff* пов'язані наступні класи *Role*, *Feedback_reply*, *Event*, *Analytic_report* та *News*. Клас *Role* містить дані, що описують ролі працівників музею. Методи цього класу включають можливості для створення, редагування, перегляду та видалення ролей. Клас *Feedback_reply* містить атрибути, що описують відповіді на відгуки користувачів. Методи цього класу дають змогу додавати, оновлювати, переглядати та видаляти відповіді на відгуки. Клас *Event* містить атрибути, що описують події в музеї. Методи цього класу забезпечують можливості для створення, редагування, перегляду та видалення подій.

Клас *Analytic_report* містить атрибути, що описують загальні характеристики аналітичних звітів. Методи цього класу надають можливості для генерування, редагування, перегляду та видалення звітів. *News* – містить атрибути, що описують новини, які створюють працівники музею. Методи цього класу дають змогу створювати, редагувати, переглядати та видаляти новини.

Висновки та перспективи подальших досліджень. В епоху стрімкого технологічного прогресу музеї мають унікальну можливість трансформувати свій підхід до взаємодії з відвідувачами та інтерактивного подання експозицій. Впровадження інформаційних та комунікаційних технологій, зокрема мобільних застосунків із елементами доповненої реальності, може збагатити досвід відвідувачів, залучити нові групи громадян та покращити процеси управління в музеях. В результаті аналізу провідних музейних застосунків було сформовано низку функціональних можливостей, які потрібно закласти при концептуальному проектуванні.

У статті запропоновано діаграми варіантів використання для опису користувацького та адміністративного функціоналу концепту застосунку з елементами доповненої реальності. Подано

узагальнену структуру бази даних та взаємозв'язки всередині інформаційної системи. Вона представляє спосіб організації та взаємозв'язку різних компонентів, визначаючи загальну модель та архітектуру бази даних. Сформовано діаграми класів, що візуально представляють структуру за шляхом моделювання класів, їхніх атрибутів, методів та взаємозв'язків між ними.

Запропонований у дослідженні концепт музейного мобільного застосунку з елементами доповненої реальності, рекомендаційною системою, інтерфейсом для користувачів та функціоналом для працівників комбінує у собі рішення різних застосунків з передових музеїв світу. Такий комплексний підхід здатний підвищити привабливість музеїв, залучити широку аудиторію та забезпечити якісний досвід вивчення культурної спадщини. В процесі подальших досліджень доцільно зосередитися на розробці та тестуванні запропонованого концепту музейного мобільного застосунку з елементами доповненої реальності серед цільової аудиторії для отримання зворотного зв'язку та вдосконалення функціоналу згідно з їхніми потребами та побажаннями. Важливим аспектом є розробка детальних технічних специфікацій для реалізації концепту з урахуванням сучасних тенденцій у галузі мобільних технологій, доповненої реальності та дизайну користувацьких інтерфейсів. Також слід приділити увагу питанням безпеки та конфіденційності даних користувачів, захисту інтелектуальної власності музеїв, а також економічним аспектам розробки, впровадження та підтримки мобільних застосунків. Перспективним напрямком є інтеграція запропонованого концепту з іншими цифровими ініціативами музеїв для створення комплексного цифрового досвіду відвідувачів.

Список бібліографічного опису

1. Kamariotou, V., Kamariotou, M., Champipi, E., & Kitsios, F. (2021). Moving towards museum digital strategy: a transformational framework. In *Business Intelligence and Modelling: Unified Approach with Simulation and Strategic Modelling in Entrepreneurship 8th* (pp. 397-402). Springer International Publishing.
2. United Nations Group on the Information Society. – URL: <https://sustainabledevelopment.un.org/index.php?page=view&type=30022&nr=2487&menu=3170> (date of access: 30.04.2024)
3. Teslyuk, T., Teslyuk, V., Lypak, H., Kunanets, N., & Veretennikova, N. (2020). A Mobile Museum Guide Application. In *MoMLeT+ DS* (pp. 314-326).
4. Cristobal-Fransi, E., Ramón-Cardona, J., Daries, N., & Serra-Cantalalops, A. (2021). Museums in the digital age: An analysis of online communication and the use of e-commerce. *Journal on Computing and Cultural Heritage (JOCCH)*, 14(4), 1-21.
5. Philippopoulos, P. I., Drivas, I. C., Tselikas, N. D., Koutrakis, K. N., Melidi, E., & Kouis, D. (2024). A Holistic Approach for Enhancing Museum Performance and Visitor Experience. *Sensors*, 24(3), 966.
6. Luther, W., Baloian, N., Biella, D., & Sacher, D. (2023). Digital twins and enabling technologies in museums and cultural heritage: An overview. *Sensors*, 23(3), 1583.
7. Agustini, K., Wahyuni, D. S., Mertayasa, I. N. E., Ratminingsih, N. M., & Ariadi, G. (2023). The effect of augmented reality mobile application on visitor impact mediated by rational hedonism: Evidence from Subak Museum. *International Journal of Advanced Computer Science and Applications*, 14(1), 77-88.
8. Bachiller, C., Monzo, J. M., & Rey, B. (2023). Augmented and virtual reality to enhance the didactical experience of technological heritage museums. *Applied Sciences*, 13(6), 3539.
9. Augmented Reality Art Gallery. URL: <https://artoflondon.co.uk/events/augmented-reality-art-gallery> (date of access: 30.04.2024).
10. The British Museum. URL: https://play.google.com/store/apps/details?id=org.britishmuseum.tpartme&hl=en_US (date of access: 30.04.2024).
11. Explorer App. URL: <https://www.amnh.org/plan-your-visit/explorer> (date of access: 30.04.2024).
12. Louvre Museum Buddy. URL: <https://play.google.com/store/apps/details?id=air.com.lvr.paris.vusiem&hl=uk&gl=US>. (date of access: 30.04.2024)
13. Natural History Museum. URL: https://play.google.com/store/apps/details?id=air.com.nhm.london.vusiem&hl=en_US (date of access: 30.04.2024)
14. Skin & Bones. URL: <https://apps.apple.com/us/app/skin-bones/id929733243> (date of access: 30.04.2024)
15. Davidavičienė, V., Raudeliūnienė, J., & Viršilaitė, R. (2021). Evaluation of user experience in augmented reality mobile applications. *Journal of Business Economics and Management*, 22(2), 467-481. <https://doi.org/10.3846/jbem.2020.13999>
16. Twin it! A pan-European collection of heritage 3D models. URL: <https://www.europeana.eu/en/galleries/15694-twin-it-a-pan-european-collection-of-heritage-3-d-models> (date of access: 30.04.2024)

References

1. Kamariotou, V., Kamariotou, M., Champipi, E., & Kitsios, F. (2021). Moving towards museum digital strategy: a transformational framework. In *Business Intelligence and Modelling: Unified Approach with Simulation and Strategic Modelling in Entrepreneurship 8th* (pp. 397-402). Springer International Publishing.
2. United Nations Group on the Information Society. – URL: <https://sustainabledevelopment.un.org/index.php?page=view&type=30022&nr=2487&menu=3170> (date of access: 30.04.2024)
3. Teslyuk, T., Teslyuk, V., Lypak, H., Kunanets, N., & Veretennikova, N. (2020). A Mobile Museum Guide Application. In *MoMLeT+ DS* (pp. 314-326).

4. Cristobal-Fransi, E., Ramón-Cardona, J., Daries, N., & Serra-Cantalops, A. (2021). Museums in the digital age: An analysis of online communication and the use of e-commerce. *Journal on Computing and Cultural Heritage (JOCCH)*, 14(4), 1-21.
5. Philippopoulos, P. I., Drivas, I. C., Tselikas, N. D., Koutrakis, K. N., Melidi, E., & Kouis, D. (2024). A Holistic Approach for Enhancing Museum Performance and Visitor Experience. *Sensors*, 24(3), 966.
6. Luther, W., Baloian, N., Biella, D., & Sacher, D. (2023). Digital twins and enabling technologies in museums and cultural heritage: An overview. *Sensors*, 23(3), 1583.
7. Agustini, K., Wahyuni, D. S., Mertayasa, I. N. E., Ratminingsih, N. M., & Ariadi, G. (2023). The effect of augmented reality mobile application on visitor impact mediated by rational hedonism: Evidence from Subak Museum. *International Journal of Advanced Computer Science and Applications*, 14(1), 77-88.
8. Bachiller, C., Monzo, J. M., & Rey, B. (2023). Augmented and virtual reality to enhance the didactical experience of technological heritage museums. *Applied Sciences*, 13(6), 3539.
9. Augmented Reality Art Gallery. URL: <https://artoflondon.co.uk/events/augmented-reality-art-gallery> (date of access: 30.04.2024).
10. The British Museum. URL: https://play.google.com/store/apps/details?id=org.britishmuseum.tapartme&hl=en_US (date of access: 30.04.2024).
11. Explorer App. URL: <https://www.amnh.org/plan-your-visit/explorer> (date of access: 30.04.2024).
12. Louvre Museum Buddy. URL: <https://play.google.com/store/apps/details?id=air.com.lvr.paris.vusiem&hl=uk&gl=US> (date of access: 30.04.2024)
13. Natural History Museum. URL: https://play.google.com/store/apps/details?id=air.com.nhm.london.vusiem&hl=en_US (date of access: 30.04.2024)
14. Skin & Bones. URL: <https://apps.apple.com/us/app/skin-bones/id929733243> (date of access: 30.04.2024)
15. Davidavičienė, V., Raudeliūnienė, J., & Viršilaitė, R. (2021). Evaluation of user experience in augmented reality mobile applications. *Journal of Business Economics and Management*, 22(2), 467-481. <https://doi.org/10.3846/jbem.2020.13999>
16. Twin it! A pan-European collection of heritage 3D models. URL: <https://www.europeana.eu/en/galleries/15694-twin-it-a-pan-european-collection-of-heritage-3-d-models> (date of access: 30.04.2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-07>

УДК 004.353.254.5

Горгун Руслан Олександрович, студент

<https://orcid.org/0009-0008-5550-0009>

Салтовський Борис Григорович, старший викладач

<https://orcid.org/0009-0006-8296-201X>

Чорноморський національний університет ім. Петра Могили, м. Миколаїв, Україна

ІНТЕРАКТИВНА КАРТА ТРИВОГ УКРАЇНИ

Горгун Р.О., Салтовський Б.Г. Інтерактивна карта тривоги України. Представлений варіант інформування населення про небезпеки, пов'язані з бойовими діями, у вигляді інтерактивної карти України. Основною задачею є створення недорогої, вандалостійкої та наочної автономної інформаційної системи. Інформер призначений для розміщення в укриттях, ситуаційних кімнатах, інших приміщеннях та на транспорті. Сама карта України надрукована на 3D-принтері і поділена на окремі комірки у формі областей. Система складається з системи керування, побудованій на мікроконтролері ESP32, адресних світлодіодів, які відсвітлюють окремі області на карті України, та звукової підсистеми, що звуковими сигналами повідомляє про початок, тип тривоги (повітряна тривога, загроза артилерії, хімічна тривога, радіаційна небезпека, вуличні бої) та її закінчення. Кожна область на карті має задне підсвітлення своїм окремим адресним світлодіодом, колір світлодіодів та звукові сигнали різняться в залежності від видів небезпеки. Після закінчення тривоги відповідний світлодіод одну хвилину світить зеленим світлом, а потім вимикається для економії енергії. Данні для обробки беруться через бездротову мережу Wi-Fi за допомогою API сайта ALERTS.IN.UA. Система живиться від зовнішнього блока живлення напругою 5 вольт з током 2,5 ампера, але передбачена можливість використання аварійного джерела напруги (пауербанка, приладу безперебійного живлення, сонячних панелей та ін.). Одним з напрямів вдосконалення системи може бути отримання інформації про небезпеки через системи стільникового зв'язку GSM.

Ключові слова: тривоги, інформер, інтерактивна карта, ESP32, WS2812, I2S.

Horhun R., Saltovskiy B. Interactive Alarm Map of Ukraine. The presented variant of informing the population about dangers related to combat actions, in the form of an interactive map of Ukraine, is aimed at creating an inexpensive, vandal-resistant and visually clear autonomous information system. The informer is intended for placement in shelters, situational rooms, other premises and transportation. The map is 3D-printed and divided into separate cells in the form of regions. The system consists of a ESP32 microcontroller board, addressable LEDs illuminating areas on the map, and an audiosubsystem that notifies about the beginning, type of alarm (air, artillery, chemical, radiation, streetfighting), and its end. Each area on the map has backlighting with its own addressable LED, the color of LEDs, and sound signals varying depending on the types of danger. After the alarm ends, the corresponding LED lights up green for minute and then turns off. Data for processing are obtained via Wi-Fi using the API of the ALERTS.IN.UA website. The system is powered by a power supply 5V/2.5A, but the possibility of using emergency power sources (power banks, solar panels, etc.) is provided. One of the directions for improving the system could be obtaining information through GSM.

Keywords: alerts, informer, interactive map, ESP32, WS2812, I2S.

Постановка наукової проблеми. В зв'язку з російською агресією постає питання інформування населення про різні тривоги, що загрожують регіонам нашої країни. Нажаль, інформування сиренами не є ефективним, тому що воно не повідомляє про вид загрози та час її закінчення. Це особливо важливо при інформуванні людей, що знаходяться в укриттях, на транспорті або в інших приміщеннях (органах влади, ситуаційних центрах та ін.).

Аналіз досліджень. На даний момент існує застосунок для смартфонів [1] та веб-сайт [2], зображення з якого можна вивести на великий екран. Недоліком першого варіанту є невеликий розмір екрана смартфона або планшета, для другого варіанта необхідний обчислювальний пристрій (наприклад, персональний комп'ютер), до якого під'єднується великий екран. Крім того, обидва варіанти є дуже коштовними та вандално нестійкими, що обмежує їх широке застосування.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Для інформування населення в укриттях та інших приміщеннях суспільного користування пропонується створення автономного апаратного інформера у вигляді інтерактивної карти України. Такий пристрій має наступні переваги:

- дешевизна – його вартість значно менше ніж інших варіантів інформерів;
- інформативність – пристрій наочно показує небезпечну територію та вид загрози, звуком повідомляє про початок та закінчення загрози;
- вандалостійкість – так як пристрій не вимагає технічного обслуговування, то його можна розмістити у недоступному місці: наприклад, на стіні під стелею.

Інформер зроблений у вигляді карти України, надрукованої на 3D-принтері (рис.1). Кожна

окрема комірка карти представляє собою окрему область країни.

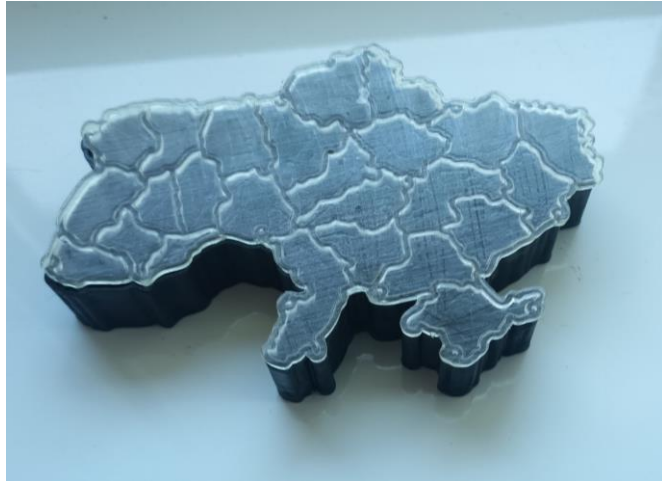


Рис. 1 – Карта України

В кожній комірці знаходиться адресний світлодіод WS2812 (рис. 2). Одна з найпоширеніших технологій управління адресними світлодіодами - це технологія WS2812, також відома як "NeoPixel" [3]. Кожен світлодіод має власну унікальну адресу, що дозволяє індивідуально керувати кожним світлодіодом у матриці. Це дозволяє створювати різноманітні ефекти світла, анімації та зображення. Керування відбувається через один дріт, що дозволяє економити апаратні ресурси мікроконтролера. Крім сигнального дроту, використовується дріт живлення +5 вольт та нульовий дріт. Всі світлодіоди з'єднуються послідовно, що дозволяє окремо керувати світінням кожного світлодіоду.

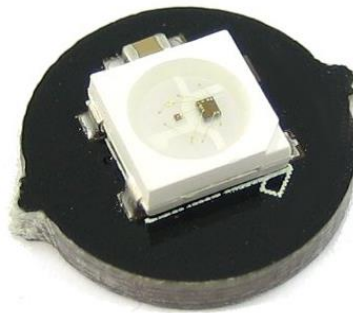


Рис. 2 – Адресний світлодіод WS2812

Перевагами таких світлодіодів є:

- дешевизна;
- можливість відображати всю палітру кольорів RGB;
- керування всіма світлодіодами лише через один дріт.

Різними кольорами можна відображати різні тривоги.

Візуальний сигнал дублюється звуком через звукову плату MAX98357 (рис. 3), яка керується через протокол I2S [4].



Рис. 3 – Звукова плата MAX98357

Протокол I2S (Inter-IC Sound) - це стандарт передачі аудіосигналів між електронними пристроями. Його основне призначення - передача цифрових аудіоданих від мікрофонів, аудіо-кодеків, DSP (цифрових сигнальних процесорів) до аудіо-зберігаючих пристроїв, таких як аудіо-процесори, звукові карти або цифрові підсилювачі.

Протокол I2S передає дані у форматі PCM (Pulse Code Modulation), де аналоговий сигнал конвертується у цифровий шляхом квантування й кодування амплітуди аналогового сигналу. Кожен кадр даних містить одне або декілька аудіосемплів, залежно від бітової глибини і кількості каналів.

Основні переваги протоколу I2S включають високу якість передачі аудіосигналів і простоту взаємодії між пристроями, що підтримують цей стандарт. Він часто використовується в аудіоапаратурі, вбудованих системах, мобільних пристроях і комп'ютерах для передачі аудіоданих від одного пристрою до іншого без втрати якості. До недоліків такого рішення можна віднести необхідність зберігання звукових файлів у внутрішній пам'яті мікроконтролера, що обмежує ці файли як по якості так і по розміру.

Як альтернативу, для відтворення звуку можливо використовувати плату DFPlayer Mini MP3 player (рис. 4).

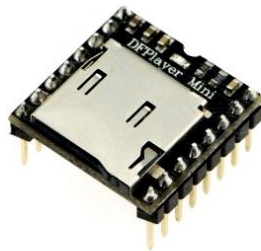


Рис. 4 – Плата DFPlayer Mini MP3 player

На цій платі всі звукові сигнали записуються на MicroSD картку у форматі MP3. Вибір необхідного файлу здійснюється по протоколу UART [5]. Ця плата підтримує частоти дискретизації 8/11.025/12/16/22.05/24/32/44.1/48 кГц, має 24bit-розрядний ЦАП, підтримка динамічного діапазону: 90 дБ, співвідношення SNR: 85dB. Повністю підтримує FAT16, FAT32 файлову систему, підтримка MicroSD карти до 32 Гб. Перевагою такої реалізації є низька ціна та простота програмування вибору звукових сигналів. До недоліків можна віднести підвищену складність загального проєкту та необхідність купівлі MicroSD картки.

Звук виводиться через динамік потужністю 0,5 Вт з опором 8 Ом (рис. 5).



Рис. 5 – Динамік

Обидва варіанти реалізації виведення звуку не вимагають додаткового підсилювача.

Для керування пристроєм використовується плата ESP32 development board (рис. 6). Дана плата базується на мікроконтролері ESP32, який має вбудований Wi-Fi та Bluetooth, що робить його ідеальним для створення систем з підключенням до мережі Інтернет [6]. Цей мікроконтролер має наступні особливості:

- два 32-розрядних ядра Xtensa LX7 @ до 240 МГц
- співпроцесор RISC-V Ultra Low Power
- 2,4 ГГц Wi-Fi - 802.11b/g/n
- Bluetooth 5, BLE + Mesh
- 16 МБ QSPI Flash



Рис. 6 – Плата ESP32 development board

Наявність на цій платі можливості під'єднуватись до мережі через Wi-Fi дозволяє отримувати поточну інформацію через Інтернет, велика кількість програмованих ввідів-виводів загального призначення – керувати адресними світлодіодами та виводити звук.

Крім того, потрібно джерело живлення з напругою 5 вольт та силою току 2,5 А.

Алгоритм функціонування цього пристрою є наступним: спочатку, він отримує інформацію про атаки через API з Інтернету. Потім, світлодіоди в областях, де відбуваються ці атаки, активуються відповідними кольорами. Окремими кольорами позначаються наступні тривоги [7]:

- повітряна тривога;
- загроза артилерії;
- хімічна тривога;
- радіаційна небезпека;
- вуличні бої;
- відміна тривоги.

У випадку тривоги у територіальній громаді, де встановлений пристрій, відтворюється попереджувальний сигнал відповідно до тривоги. Якщо тривогу скасовано відповідно до областей, світлодіоди на одну хвилину миготять зеленим світлом, а в територіальній громаді додається сигнал відбою. Оскільки передбачено, що пристрій буде фіксованим, інформація про територіальну громаду вбудовується в пристрій на етапі його прошивки.

Але, крім переваг, даний інформер також має недоліки:

- Необхідність зовнішнього джерела доступу до мережі Інтернет; але більшість укриттів на

поточний момент, як правило, обладнані таким доступом. Також є можливість дообладнати схему 4G-модулем доступу до мобільного інтернету.

- Необхідність зовнішнього живлення. Це вирішується за допомогою пауербанків або систем безперебійного живлення.

Висновки та перспективи подальшого дослідження. Нажаль, існує необхідність розповсюдження пристроїв такого призначення і наведений інформер – один з варіантів їх реалізації. Напрямами подальшого дослідження можуть бути можливості безпосереднього отримання необхідної інформації через мобільну мережу. Може бути використаний як готовий модем, що приєднується за протоколом USB, так і окрема плата, що передає данні за допомогою UART або I2C. По характеристикам підходить, наприклад, плата SIMcom SIM7600E-H (рис. 7), яка крім мобільного зв'язку дозволяє через супутникову систему місцезнаходження GPS автоматично визначати територіальну громаду, для якої актуальна інформація про тривоги [8].



Рис 7 – Плата SIMcom SIM7600E-H

Також є актуальним пошук рішення енергетичної незалежності інформера від електромережі. У простішому випадку – використання пауербанків, але можливе застосування пристроїв безперервного живлення, що може суттєво підвищити вартість пристрою.

Список бібліографічного опису

1. Застосунок «Мапа тривог України». URL: <https://play.google.com/store/apps/details?id=org.ukrzen.alertsinua> (дата звернення: 25.04.2024).
2. Карта повітряних тривог України. URL: <https://alerts.in.ua/> (дата звернення: 25.04.2024).
3. WS2812. Intelligent control LED integrated light source. URL: <https://cdn-shop.adafruit.com/datasheets/WS2812.pdf> (дата звернення: 25.04.2024).
4. Lady Ada. Adafruit MAX98357 I2S Class-D Mono Amp. Adafruit Learning System. URL: <https://learn.adafruit.com/adafruit-max98357-i2s-class-d-mono-amp/overview> (дата звернення: 25.04.2024).
5. DFPlayer Mini MP3 player. URL: https://wiki.dfrobot.com/DFPlayer_Mini_SKU_DFR0299 (дата звернення: 25.04.2024).
6. Espressif ESP32 Development Board - Developer Edition. Adafruit Industries, Unique & fun DIY electronics and kits. URL: <https://www.adafruit.com/product/3269> (дата звернення: 25.04.2024).
7. Ukraine Alert API. URL: <https://api.ukrainealarm.com/swagger/index.html> (дата звернення: 25.04.2024).
8. Simcom SIM7600E-H Datasheet URL: <https://octopart.com/datasheet/sim7600e-h-simcom-89998097> (дата звернення: 25.04.2024).

References

1. Application "Map of alarms of Ukraine". URL: <https://play.google.com/store/apps/details?id=org.ukrzen.alertsinua> (last accessed: 25.04.2024).
2. Map of alarms of Ukraine. URL: <https://alerts.in.ua/> (last accessed: 25.04.2024).
3. WS2812. Intelligent control LED integrated light source. URL: <https://cdn-shop.adafruit.com/datasheets/WS2812.pdf> (Last accessed: 25.04.2024).
4. Lady Ada. Adafruit MAX98357 I2S Class-D Mono Amp. Adafruit Learning System. URL: <https://learn.adafruit.com/adafruit-max98357-i2s-class-d-mono-amp/overview> (last accessed: 25.04.2024).
5. DFPlayer Mini MP3 player. URL: https://wiki.dfrobot.com/DFPlayer_Mini_SKU_DFR0299 (last accessed: 25.04.2024).
6. Espressif ESP32 Development Board - Developer Edition. Adafruit Industries, Unique & fun DIY electronics and kits. URL: <https://www.adafruit.com/product/3269> (last accessed: 25.04.2024).
7. Ukraine Alert API. URL: <https://api.ukrainealarm.com/swagger/index.html> (last accessed: 25.04.2024).
8. Simcom SIM7600E-H Datasheet URL: <https://octopart.com/datasheet/sim7600e-h-simcom-89998097> (last accessed: 25.04.2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-08>

УДК 004.05

Демченко Андрій Володимирович, магістр

<https://orcid.org/0009-0008-8180-9042>

Федосова Ірина Василівна, д.п.н., професор

<https://orcid.org/0000-0003-3923-8270>

Приазовський державний технічний університет, м. Дніпро/Маріуполь, Україна

ДОСЛІДЖЕННЯ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ПЛАТФОРМ ДЛЯ ДИСТАНЦІЙНОГО НАВЧАННЯ

Демченко А.В., Федосова І.В. Дослідження інформаційно-комунікаційних платформ для дистанційного навчання. В роботі досліджено платформи дистанційного навчання «Classroom» та «Microsoft Team», які навчальні заклади України почали активно впроваджувати в освітній процес наприкінці 2019 року із запровадженням надзвичайного стану, причиною якого стала всесвітня пандемія Covid-19. Також проаналізовано платформи відеотелефоного зв'язку: «Discord», «Zoom», «Meet», «Microsoft Teams», створено порівняльні таблиці основних критеріїв цих програмних продуктів, розглянуто тарифні плани платформ дистанційного навчання та програм відеотелефоного зв'язку. Приділено увагу Українським проєктам «Класна оцінка» та «На урок». Підсумовано чинники важливості створення власної платформи дистанційної освіти для України, як нового етапу розвитку освітнього процесу та розвитку електронного документообігу в державних навчальних установах.

Ключові слова: дистанційне навчання; інформаційно-комунікаційні платформи; Covid-19; інтернет; Microsoft Team; Classroom; Zoom; Discord; Meet.

Demchenko A., Fedosova I. Research of information and communication platforms for distance learning. The work examines the distance learning platforms «Classroom» and «Microsoft Team», which educational institutions of Ukraine began to actively implement in the educational process at the end of 2019 with the introduction of a state of emergency caused by the worldwide Covid-19 pandemic. The platforms of video-telephone communication: «Discord», «Zoom», «Meet», «Microsoft Teams» were also analyzed, comparative tables of the main criteria of these software products were created, tariff plans of distance learning platforms and video-telephone communication programs were considered. Attention was paid to the Ukrainian projects «Class assessment» and «For a lesson». The factors of the importance of creating an own distance education platform for Ukraine as a new stage in the development of the educational process and the development of electronic document management in state educational institutions are summarized.

Keywords: distance learning; information and communication platforms; Covid-19; Internet; Microsoft Team; Classroom; Zoom; Discord; Meet.

Постановка наукової проблеми. З початком всесвітньої пандемії, спричиненою вірусом COVID-19, усі навчальні заклади України перейшли на дистанційний формат освіти. Хоча загроза пандемії вже не є актуальною, але дистанційне навчання продовжується в більшості областей країни і донині через початок повномасштабного вторгнення країни-агресора в Україну. Метою впровадження такої форми навчання є забезпечення безпечних умов, як для тих, хто навчається, так і для тих, хто навчає. Також слід зауважити, що спостерігається тенденція розвитку дистанційного навчання і без потреби забезпечення безпеки, а просто для комфортного процесу навчання, наприклад, різноманітні освітні курси або заочна форма навчання.

Дистанційне навчання — це форма навчання з використанням комп'ютерних і телекомунікаційних технологій, які забезпечують інтерактивну взаємодію викладачів та здобувачів освіти на різних етапах навчання і самостійну роботу з матеріалами інформаційної мережі [1].

Дистанційне навчання – сукупність наступних заходів:

- засоби надання учбового матеріалу здобувачу освіти;
- засоби контролю успішності здобувача освіти;
- засоби консультації здобувача програмою-викладачем;
- засоби інтерактивної співпраці викладача і здобувача;
- можливість швидкого доповнення курсу новою інформацією, коригування помилок [1].

Оскільки наразі в Україні не існує централізованого рішення для дистанційного процесу навчання, а засоби які пропонуються приватними компаніями не можуть покрити більшої частини потреб, як здобувачів освіти так і викладачів, вважаємо, що дослідження критеріїв та функцій існуючих інформаційно-комунікаційних платформ є важливим етапом для створення узагальненої платформи для освітнього процесу в Україні.

Метою роботи є виокремлення проблем платформ дистанційного навчання та додаткових програмних рішень, котрі наразі використовуються у закладах вищої освіти України. Завданням

дослідження є проектування та реалізація програмного продукту, який буде надавати інформаційно-комунікаційні послуги та автоматизувати процеси введення обліку успішності учнів.

Аналіз досліджень. В Україні найбільш використовуваними сервісами є «Classroom», «Zoom», «Discord», «Microsoft Teams», «Meet», «На урок», «Класна оцінка». Розглянемо кожен цей сервіс детальніше.

«Classroom». Була презентована компанією Google 12 серпня 2014 року. Створений спеціально для дистанційного навчання, освітніх закладів США. Оскільки розробником є компанія Google в цей сервіс інтегровані всі інші продукти від цієї компанії такі, як: Meet, Документи, Таблиці, Презентації, Календар, Диск (хмарне сховище), Форми, Gmail, YouTube. Таким чином формується доволі зручний сервіс для дистанційного навчання, який базується на роботі з текстовими документами, відео зустрічами, та обліком виконання робіт здобувачів освіти. Є можливість коментувати завдання, як здобувачам освіти так і викладачам, є можливість створювати стрічку з новинами.

Серед недоліків можна виділити наступні:

1. Перший проблема стосується навчання в школах, це стосується системи оцінювання і запису оцінок. У програмі автоматично вводиться сто бальна система і немає можливості створювати електронний журнал з оцінками та темами. Немає можливості створити журнал відвідування здобувачів освіти уроків та іншого документа обігу, який стосується навчального процесу.

2. Друга проблема стосується обмеження робочого місця на Google диску, всі роботи здобувачів освіти, а також завдання вчителів зберігаються на ньому. Внаслідок чого диск може переповнюватися, що призводить до неприємних наслідків. Ця проблема вирішується або підпискою на збільшення простору хмарного сховища, або видалення застарілих файлів.

3. Третя проблема стосується неможливості створити акаунти для батьків, щоб вони могли бачити сповіщення на результати виконання робіт своїх учнів.

Також слід зазначити, що в багатьох школах «Classroom» використовують не правильно. Замість створення окремого класу для кожного предмету створюють клас, де розміщуються всі предмети і завдання до них. Це спричиняє великі складнощі у використанні сервісом. Наприклад, один викладач може випадково видалити завдання іншого викладача і в «Classroom» не має можливості відновити видалене завдання.

«Meet». Програма створена компанією Google в 2017 році, є сервісом відеотелефонного зв'язку, використовується як окрема платформа так і в сервісі «Classroom». Серед мінусів можна назвати необхідність платної підписки для проведення конференцій 100+ людей, що може вплинути на деякі процеси навчання у закладах вищої освіти. Не має можливості зручно виставляти та здавати завдання і не має сервісу оцінювання робіт.

«Discord». Створена компанією Discord Inc. 13 травня 2015 року, є сервісом відеотелефонного зв'язку, також має можливість вести інформативні канали, форуми та чати. Найчастіше використовується, як спосіб проведення онлайн конференцій (занять). Не має можливості зручно виставляти та здавати завдання і не має сервісу оцінювання робіт.

«Zoom». Створено компанією Zoom Video 25 січня 2013 року, є сервісом відеотелефонного зв'язку, стала популярною під час всесвітньої пандемії Covid-19. Використовується для проведення онлайн конференцій (занять). Має обмежену кількість користувачів зустрічі у безкоштовній версії. Також є обмеження в часі – тільки 40 хвилин може тривати розмова у безкоштовному тарифі, що робить дуже не зручним використання програми у освітньому процесі.

«Microsoft Teams». Розроблена корпорацією Microsoft 14 березня 2017 року. Платформа створювалася, як сервіс для спільного робочого простору в хмарі із використанням продуктів Microsoft. Із початком пандемії Covid-19 платформа почала розвиватися у напрямку дистанційного навчання, що з часом призвело до створення модифікованої версії саме для навчальних закладів. Даний сервіс має широкий спектр функцій для закладів вищої освіти. Наприклад, є можливість створювати завдання та оцінювати їх прямо в застосунку, можна створювати розклад занять, і він буде відображатися, як у викладачів так і у здобувачів вищої освіти. Сервіс має можливості проводити онлайн конференції без обмежень у кількості користувачів та за часом проведення. Мінусом є те, що треба звертатися до технічної підтримки і надсилати запит на створення навчальних акаунтів. Без цього на даний момент програма не буде надавати доступу до функцій, які перераховані вище, а буде обмежена функціоналом програми «Discord».

«На урок». Це український приватний проєкт, який ставить за мету створити освітню цифрову екосистему для викладачів і здобувачів освіти. На платформі розміщуються матеріали, створені викладачами (презентації, завдання, тести), а також він має функцію автоматичної перевірки тестових завдань, що спрощує роботу викладачу і надає одразу результат перевірки здобувачу освіти. Використовується саме для надання тестів здобувачам освіти і пошуку матеріалів для занять, і більше не має ніякого іншого функціоналу.

«Класна оцінка». Український проєкт, який надає послуги створення електронного документообігу (створення журналів, табелів, розкладу). Має додаткові функції: створення простих сайтів для навчальних закладів, створення тестів, створення блогу, форуму. Але на практиці його використовують саме для ведення електронного журналу. Цей проєкт має як безкоштовну версію, так і розширену платну версію сайту. Для використання звичайним навчальним закладом необхідно купувати платну версію, оскільки безкоштовна підтримує до 10 осіб у класі, що для закладів освіти України є недостатньою кількістю. Також у безкоштовній версії немає всього необхідного функціоналу для проведення освітнього процесу, але є реклама, що робить сервіс менш привабливим.

Слід також зауважити, що в освітньому процесі активно використовують месенджери Viber, Telegram, WhatsApp. Їх застосовують для швидкого зв'язку між викладачами та здобувачами освіти, також в них дублюють важливу інформацію з платформ дистанційного навчання.

Загалом програмні продукти які використовуються в Україні дублюють один одного. З унікальних можна виділити тільки «На урок» і «Класна оцінка». Побудуємо порівняльну таблицю запропонованих програмних продуктів.

Проаналізуємо платформи «Classroom» та «Microsoft Teams» для дистанційного навчання за вказаними критеріями (див. табл. 1).

Таблиця 1 – Порівняння платформ для дистанційного навчання

Критерії	Classroom	Microsoft Teams
Україномовний інтерфейс	+	+
Можливість відео конференції	+	+
Можливість давати завдання	+	+
Можливість оцінювати роботи	+	+
Розклад занять	-	+
Стрічка новин	+	+
Вільний доступ до платформи	+	Для отримання навчальної версії платформи необхідно зв'язуватися із менеджером
Автоматична перевірка завдань	-	-
Електронний документообіг	-	-
Створення батьківського акаунту	-	-
Коментувати завдання	+	+
Індивідуальний чат із викладачем	+	-
Інтеграція з іншими програмами розробників	+	Тільки у платній версії
Країна походження	США	США
Цільова аудиторія продукту	Заклади освіти	Заклади освіти
Автоматичні звіти про присутність на онлайн конференціях	Тільки у платній версії	Тільки у платній версії
Служба підтримки	+	+
Володіння українською мовою, працівників служби підтримки	+	+

Розглянемо звичайну версію сервісів відеотелефонного зв'язку: «Discord», «Zoom», «Meet», «Microsoft Teams» (див. Табл 2).

Таблиця 2 – Порівняння сервісів відеотелефонного зв'язку

Критерії	Discord	Zoom	Meet	Microsoft Teams
----------	---------	------	------	-----------------

Максимальна кількість користувачів (безкоштовна версія)	50 чоловік	100 чоловік	100 чоловік	100 чоловік
Україномовний інтерфейс	+	-	+	+
Обмеження в часі	24 години	40 хв	24 години	60 хв
Обмеження в часі (платні версії)	24 години (тривалість однієї зустрічі)	30 годин на місяць	24 години (тривалість однієї зустрічі)	30 годин на місяць
Кількість учасників у платних тарифах	Обговорюється індивідуально	Pro 100 чоловік	Business Starter 100 чоловік	Microsoft Teams базовий 300 учасників
		Business 300 чоловік	Business Standard 150 чоловік	Microsoft 365 Business базовий 300 учасників
		Business Plus 300 чоловік	Business Plus 500 чоловік	Microsoft 365 Бізнес Стандарт 300 учасників
		Enterprise 1000 чоловік	Enterprise 1000 чоловік	
Хмарне сховище (платні версії)	Можливість відправляти файли до 5 ГБ	Pro 5 ГБ	Business Starter 30 ГБ	Microsoft Teams базовий 10 ГБ
		Business 5 ГБ	Business Standard 2 ТБ	Microsoft 365 Business базовий 1 ТБ
		Business Plus 10 ГБ	Business Plus 5 ТБ	Microsoft 365 Бізнес Стандарт 1 ТБ
		Enterprise немає обмеження	Enterprise 5 ТБ (можна збільшити при обговоренні оплати)	
Ціни платних тарифів (ціни вказані за одного користувача на місяць)	Спеціальна умов для навчальних закладів. Обговорюється індивідуально	Pro 14.99\$	Business Starter 5.4\$	Microsoft Teams базовий 4\$
		Business 19.99\$	Business Standard 10.8\$	Microsoft 365 Business базовий 6\$
		Business Plus обговорюється індивідуально	Business Plus 18\$	Microsoft 365 Бізнес Стандарт 12.5\$
		Enterprise обговорюється індивідуально	Enterprise обговорюється індивідуально	

Проаналізувавши платформи для дистанційного навчання, які зараз використовуються у закладах освіти в Україні, автори дійшли висновку, що основними критеріями цих платформ є:

1. Централізація всіх необхідних функцій в одному програмному рішенні.

2. Програмне забезпечення має розроблятися, фінансуватися, модернізуватися та підтримуватися коштами МОН, щоб усі заклади освіти могли на безоплатній основі використовувати всі програмні функції.

3. Має бути електронний документообіг. Можливість вести електронний журнал, автоматичне створення табелів, атестатів та іншої документації.

4. Створення розкладу занять.

5. Можливість викладачам викладати завдання та оцінювати їх, а здобувачам освіти здавати завдання та дізнаватися свої оцінки.

6. Можливість швидкої комунікації між викладачами та здобувачами.

7. Для закладів освіти є важливим функціонал взаємодії навчальної платформи із батьками, можливість з ними зв'язатися, надіслати результати оцінювання дитини і відвідування уроків.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.

Кластерний аналіз успішності здобувачів освіти, дозволяє викладачам створити більш досконалу освітню програму з орієнтування на досягненнях здобувачів та створенні для них оптимальних умов. У нових світових тенденціях за таким видом даних створюють окремі класи, де всі здобувачі приблизно одного освітнього рівня, що створює зручні умови, як для здобувачів так і для викладачів.

Аналіз успішності учнів та побудова прогнозу на основі вже існуючих оцінок та рівня складності завдань дає змогу викладачу створити завдання, які найбільш підходять для рівня здобувачів. Для створення кластеру успішності учнів, було створено три правила:

- початковий рівень (Кластер 1), якщо середній бал здобувача освіти знаходиться в проміжку від 1 балу і менше 7.
- середній рівень (Кластер 2), якщо середній бал здобувача освіти знаходиться в проміжку від 7 балів і менше 10.
- високий рівень (Кластер 3), якщо середній бал здобувача освіти знаходиться в проміжку від 10 до 12 балів.

Для побудови середнього балу здобувача береться окремо сторінка предмету в онлайн журналі та розраховується середній бал всіх оцінок.

Після того як середній бал здобувача освіти був врахований і його віднесли до певного класу, розраховується відсоткове значення кожного кластеру для класу. На основі цих даних викладач може відрегулювати рівень складності завдань для здобувачів. Також слід вказати, що на сторінці кластеризації відображаються різноманітні графіки, які демонструють відсоток учнів в тому чи іншому кластері.



Рис. 1 – Вивід діаграм за даними оцінок в журналі

На основі отриманих даних розраховується приблизний середній бал наступних робіт здобувачів і по цих даних виводяться кластери успішності здобувачів. Це призводить до того, що викладач по цих даних може зорієнтуватися, які результати мають бути отримані здобувачами, після чого відкоригувати складність робіт.

У роботі було створено онлайн платформу успішності здобувачів, котра створює умови переходу викладачів на електронний документообіг та максимально автоматизувати одноманітні задачі для покращення освітнього процесу, застосовуючи кластерний аналіз успішності здобувачів. Для реалізації онлайн платформи було використано програмне забезпечення: Visual Studio Code; ХАМРР; PhpMyAdmin; Браузер Chrome; Adobe Photoshop; Git і мови програмування: HTML; CSS;

JavaScript; PHP; SQL. Платформа складається з двох частин: база даних і реалізація, вебдодаток для виконання необхідних запитів до бази даних. Клієнтами онлайн платформи є викладачі, здобувачі й адміністратор платформи, котрий створює основу на якій працюють користувачі.

Експерименти проводилися для пункту меню «Прогнозування». Для цього було запрошено троє експертів (викладачів із сумарним стажем роботи більше 50 років). Для розрахунків використовувались реальні результати успішності здобувачів. Після чого ці дані були надані програмі та експертам для аналізу та прогнозуванню успішності кожного із здобувачів.

Експеремен Експеремен										
Середнє значення рівня складності 1.4										
Середнє значення рівня складності майбутніх робіт 2										
id	suriname	name	user_id	Оцінка1	Оцінка2	Оцінка3	Оцінка4	average_658731bfabcd9	Середній бал	Середній бал наступних робіт
1	Учень	№1	113	5	6	6	7	6.00	6	5
2	Учень	№2	114	10	11	11	9	10.25	10.25	10.25
3	Учень	№3	115	5	8	9	9	7.75	7.75	6.75
4	Учень	№4	116	12	12	12	12	12.00	12	12
5	Учень	№5	117	4	5	4	5	4.50	4.5	3.5
6	Учень	№6	118	9	9	9	9	9.00	9	8
7	Учень	№7	119	6	8	8	11	8.25	8.25	7.25
8	Учень	№8	120	8	8	10	10	9.00	9	8
9	Учень	№9	121	11	12	11	11	11.25	11.25	11.25
10	Учень	№10	122	7	8	10	9	8.50	8.5	7.5

Відсотки:
Кластер 1: 20%
Кластер 2: 50%
Кластер 3: 30%

Відсотки нові:
Кластер 1: 30%
Кластер 2: 40%
Кластер 3: 30%

Рис. 2 – Результати прогнозування

Результати прогнозування викладачів та розробленого алгоритму наведені нижче.



Рис. 3 – Результати експеременту

Висновки та перспективи подальшого дослідження. Здійснено комплексне порівняння платформ дистанційного навчання, які активно використовуються в Україні. Досліджено сильні та слабкі сторони цих платформ. Розроблено Web-додаток для вирішення проблем автоматизації процесів введення документації щодо оцінювання здобувачів освіти та формування документів про їх успішність. Експериментальна частина роботи пройшла апробацію в навчальних закладах м. Дніпра: середня загальноосвітня спеціалізована школа №141 із поглибленим вивченням іноземних мов та Дніпровський ліцей №145. У подальшому створену платформу можна розширювати новим функціоналом і масштабувати на різні заклади освіти, що призведе до оптимізації освітнього процесу.

Список бібліографічного опису

1. Дистанційна освіта. – Режим доступу: <http://vnz.org.ua/dystantsijna-osvita/pro>.
2. Halah Ahmed Alismail. Teachers' perspectives of utilizing distance learning to support 21st century skill attainment for K-3 elementary students during the COVID-19 pandemic era. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2405844023064836#sec8>.
3. Markus Wolfgang Hermann Spitzer, Korbinian Moeller. Performance increases in mathematics during COVID-19 pandemic distance learning in Austria: Evidence from an intelligent tutoring system for mathematics. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2211949323000066#abs0001>.
4. Karen Kwaning, Karen Kwaning, Ayman Ullah, Christopher Biely, Nicholas Jackson, Kulwant K. Dosanjh, Arzie Galvez, Guadalupe Arellano, Rebecca Dudovitz. Adolescent Feelings on COVID-19 Distance Learning Support: Associations With Mental Health, Social-Emotional Health, Substance Use, and Delinquency. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S1054139X22010461#abs0010>.
5. Frank van der Wouden, Hyejin Youn. The impact of geographical distance on learning through collaboration. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S0048733322002190#ab0010>.
6. Miltiadis D. Lytras, Andreea Claudia Serban, Miguel Jesus Torres Ruiz, Stamatios Ntanos, Akila Sarirete. Translating knowledge into innovation capability: An exploratory study investigating the perceptions on distance learning in higher education during the COVID-19 pandemic – the case of Mexico. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2444569X22000932#abs0001>.
7. Ministry of Education and Science of Ukraine. – Режим доступу: <https://mon.gov.ua/ua>.
8. Education.ua. – Режим доступу: <http://osvita.ua/index.html/>.
9. Laith M. Almomani, Niveen Halalsheh, Hanadi Al-Dreabi, Leena Al-Hyari, Raed Al-Quraan. Self-directed learning skills and motivation during distance learning in the COVID-19 pandemic (case study: The university of Jordan). – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2405844023072262>.
10. Ahmed Bossman, Samuel Kwaku Agyei. Technology and instructor dimensions, e-learning satisfaction, and academic performance of distance students in Ghana. – Режим доступу: <https://www.sciencedirect.com/science/article/pii/S2405844022004881>.
11. About Classroom. – Режим доступу: <https://support.google.com/edu/classroom/answer/6020279?hl=en#>.
12. Choose your Google Workspace pricing plan. – Режим доступу: <https://workspace.google.com/intl/en/pricing.html>.
13. Discord. – Режим доступу: <https://discord.com/college>.
14. Zoom, Plans & Pricing. – Режим доступу: <https://zoom.us/pricing>.
15. Microsoft, Education. – Режим доступу: <https://www.microsoft.com/en-us/education>.
16. «НА УРОК». – Режим доступу: <https://naurok.com.ua/>.
17. «Класна оцінка». – Режим доступу: <https://klasnaocinka.com.ua/uk/site/page/view#k15-block8>.

References

1. Distance education. Available at: <http://vnz.org.ua/dystantsijna-osvita/pro>.
2. Halah Ahmed Alismail. Teachers' perspectives of utilizing distance learning to support 21st century skill attainment for K-3 elementary students during the COVID-19 pandemic era. Available at: <https://www.sciencedirect.com/science/article/pii/S2405844023064836#sec8>.
3. Markus Wolfgang Hermann Spitzer, Korbinian Moeller. Performance increases in mathematics during COVID-19 pandemic distance learning in Austria: Evidence from an intelligent tutoring system for mathematics. Available at: <https://www.sciencedirect.com/science/article/pii/S2211949323000066#abs0001>.
4. Karen Kwaning, Karen Kwaning, Ayman Ullah, Christopher Biely, Nicholas Jackson, Kulwant K. Dosanjh, Arzie Galvez, Guadalupe Arellano, Rebecca Dudovitz. Adolescent Feelings on COVID-19 Distance Learning Support: Associations With Mental Health, Social-Emotional Health, Substance Use, and Delinquency. Available at: <https://www.sciencedirect.com/science/article/pii/S1054139X22010461#abs0010>.
5. Frank van der Wouden, Hyejin Youn. The impact of geographical distance on learning through collaboration. Available at: <https://www.sciencedirect.com/science/article/pii/S0048733322002190#ab0010>.
6. Miltiadis D. Lytras, Andreea Claudia Serban, Miguel Jesus Torres Ruiz, Stamatios Ntanos, Akila Sarirete. Translating knowledge into innovation capability: An exploratory study investigating the perceptions on distance learning in higher education during the COVID-19 pandemic – the case of Mexico. Available at: <https://www.sciencedirect.com/science/article/pii/S2444569X22000932#abs0001>.
7. Ministry of Education and Science of Ukraine. Available at: <https://mon.gov.ua/ua>.
8. Education.ua. Available at: <http://osvita.ua/index.html/>.
9. Laith M. Almomani, Niveen Halalsheh, Hanadi Al-Dreabi, Leena Al-Hyari, Raed Al-Quraan. Self-directed learning skills and motivation during distance learning in the COVID-19 pandemic (case study: The university of Jordan). Available at: <https://www.sciencedirect.com/science/article/pii/S2405844023072262>.
10. Ahmed Bossman, Samuel Kwaku Agyei. Technology and instructor dimensions, e-learning satisfaction, and academic performance of distance students in Ghana. Available at: <https://www.sciencedirect.com/science/article/pii/S2405844022004881>.
11. About Classroom. Available at: <https://support.google.com/edu/classroom/answer/6020279?hl=en#>.
12. Choose your Google Workspace pricing plan. Available at: <https://workspace.google.com/intl/en/pricing.html>.
13. Discord. Available at: <https://discord.com/college>.
14. Zoom, Plans & Pricing. Available at: <https://zoom.us/pricing>.
15. Microsoft, Education. Available at: <https://www.microsoft.com/en-us/education>.
16. TO THE LESSON. Available at: <https://naurok.com.ua/>.
17. Cool grade. Available at: <https://klasnaocinka.com.ua/uk/site/page/view#k15-block8>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-09>

УДК 004.02

Дідус Андрій Володимирович, аспірант

<https://orcid.org/0009-0004-2235-6742>

Терейковський Ігор Анатолійович, д.т.н., професор

<https://orcid.org/0000-0003-4621-9668>

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

ФОРМАЛІЗАЦІЯ ПРОЦЕСУ РОЗПІЗНАВАННЯ КЛЮЧОВИХ СЛІВ У ГОЛОСОВОМУ СИГНАЛІ

Дідус А. В., Терейковський І. А. Формалізація процесу розпізнавання ключових слів у голосовому сигналі.

У даній роботі було здійснено формалізацію процесу розпізнавання ключових слів у голосовому сигналі, що стало основою для розробки деталізованої концептуальної моделі. Запропонована модель надає комплексний і систематизований опис процесу створення ефективних засобів розпізнавання. Модель охоплює критичні компоненти такі як виділення ознак, методологію моделювання, стратегію навчання, процедури тестування, та механізми використання. Значна увага приділяється впровадженню передових методів глибокого навчання, включаючи нейронні мережі, марковські моделі та техніки аугментації даних, що сприяє підвищенню точності розпізнавання ключових слів. Ця робота враховує численні виклики, пов'язані з варіативністю голосових сигналів, стійкістю до шумів, обмеженістю навчальних даних та вимогами до продуктивності в режимі реального часу. Запропонований формалізований підхід дозволяє оптимізувати взаємодію між компонентами та підвищити загальну ефективність системи розпізнавання ключових слів у голосовому сигналі.

Ключові слова: розпізнавання ключових слів, формалізація процесу, голосовий сигнал, концептуальна модель, глибоке навчання, нейронні мережі, марковські моделі

Didus A., Tereikovskiy I. Formalization of the keyword recognition process in speech signal. In this work, the formalization of the keyword recognition process in speech signals has been carried out, which has laid the foundation for the development of a detailed conceptual model. The proposed model provides a comprehensive and systematic description of the process for creating effective recognition tools. The model covers critical components such as feature extraction, modeling methodology, training strategy, testing procedures, and usage mechanisms. Significant attention is given to the implementation of advanced deep learning methods, including neural networks, Markov models, and data augmentation techniques, which contribute to enhancing the accuracy of keyword recognition. This work takes into account numerous challenges associated with the variability of speech signals, noise resilience, limited training data, and real-time performance requirements. The proposed formalized approach allows for the optimization of interaction among components and enhances the overall efficiency of the keyword recognition system in speech signals.

Keywords: keyword recognition, process formalization, speech signal, conceptual model, deep learning, neural networks, Markov models.

Постановка наукової проблеми.

В сучасних умовах ефективне розпізнавання ключових слів у голосовому сигналі є важливою та актуальною задачею в області обробки природної мови та голосових технологій. Застосування засобів розпізнавання ключових слів у голосовому сигналі знаходить широке застосування у різноманітних галузях, таких як голосове керування пристроями, системи розпізнавання мови, інтерфейси "людина-пристрій" тощо.

В останні роки при розробці таких систем активно використовуються найбільш сучасні досягнення в галузі глибокого навчання, зокрема нейронні мережі, марковські моделі та різного виду технології обробки даних, що значно підвищує точність та адаптивність систем до різноманітних умов використання.

В контексті швидкого розвитку глибинного навчання, ця задача залучає значну академічну та комерційну увагу, оскільки відкриває нові можливості для розуміння та взаємодії з голосовими даними у різних сферах. Ця наукова робота присвячена формалізації процесу розпізнавання ключових слів у голосовому сигналі, з акцентом на методах глибинного навчання та алгоритмічних підходах, що стає особливо важливим у сучасних умовах розвитку технологій. Важливість теми підтверджується результатами досліджень [1, 4, 5], де вказано, що розв'язання проблеми ефективного розпізнавання ключових слів у голосовому сигналі належить до актуальних завдань у галузі обробки мовленнєвих даних, розробки систем розпізнавання мовлення та створення голосових інтерфейсів. Дана проблема безпосередньо пов'язана з державними науковими та практичними програмами, спрямованими на розвиток сучасних інформаційних технологій, забезпечення кібербезпеки та захисту інформації в умовах зростаючих обсягів даних та ризиків кібератак. [2, 3]

Аналіз досліджень.

В сучасних науково-прикладних роботах дотичних до теми даного наукового дослідження розглядаються різні підходи до розпізнавання ключових слів з використанням прихованих марківських моделей (НММ), згорткових нейронних мереж (CNN), рекурентних нейронних мереж (RNN), трансформерів, багатозадачного навчання та інших сучасних методів машинного навчання. Зокрема:

1. Застосування НММ в поєднанні з мел-частотними кепстральними коефіцієнтами (MFCC) для виділення ознак голосового сигналу та побудови акустичних моделей є класичним підходом [5].
2. Використання CNN, RNN, їх комбінацій та архітектур на кшталт ResNet, TransformerNet для кінцевого розпізнавання ключових слів без проміжної генерації фону [8].
3. Багатозадачне навчання для одночасного розпізнавання ключових слів та визначення, чи належить голос користувачеві пристрою чи сторонній особі [7].
4. Застосування методів аугментації даних, зокрема голосової конверсії, для розширення навчальних вибірок [7].
5. Дослідження аудіо-візуальних підходів із використанням відеопотоку артикуляції губ для покращення розпізнавання [8].

В результаті проведеного аналізу науково-практичних робіт в області розпізнавання ключових слів у голосовому сигналі можливо стверджувати про відсутність формалізованого та детального опису процесу розробки засобів розпізнавання ключових слів у голосовому сигналі, що ускладнює можливість визначення перспективних шляхів підвищення їх ефективності.

Адже це могло б відкрити широкий спектр подальших досліджень і розробок. Важливо розробити концептуальну модель підходу до розпізнавання ключових слів у голосовому сигналі, яка б враховувала низку викликів та обмежень, зокрема:

1. Виклик розпізнавання в умовах шуму: Голосові сигнали часто змішуються з фоновими шумами, що ускладнює процес розпізнавання. Необхідно розробити методи, які ефективно враховують цей аспект.
2. Варіативність голосових сигналів: Голосові сигнали можуть значно відрізнятися від особи до особи через вікові, статеві, етнічні та інші характеристики. Модель повинна бути здатна адаптуватися до цієї варіативності.
3. Обмеженість даних для навчання: Для ефективного навчання моделей глибокого навчання потрібні великі набори даних. Однак, в реальному світі, доступ до великих наборів даних може бути обмеженим.
4. Виклики реалізації в реальному часі: Для багатьох застосувань, таких як голосове керування пристроями, необхідно, щоб розпізнавання ключових слів відбувалося в реальному часі. Це вимагає оптимізації як самої моделі, так і процесу обробки сигналів.
5. Проблема перенавчання: При використанні складних моделей, таких як глибокі нейронні мережі, існує ризик перенавчання, коли модель стає занадто специфічною для навчального набору даних і втрачає здатність узагальнювати на нових даних.

Незважаючи на значні досягнення в області розпізнавання ключових слів в голосових сигналах (ГС) завдяки застосуванню різноманітних алгоритмів та моделей, у науковій літературі спостерігається недостатність у формалізації самого процесу розпізнавання, що у свою чергу ускладнює визначення перспективних шляхів удосконалення відомих технологій розпізнавання.

У світлі існуючих викликів та обмежень, важливо розробити детальну концептуальну модель для розпізнавання ключових слів у голосових сигналах. Такий формалізований опис повинен охоплювати всі критичні компоненти процесу, включно з виділенням ознак, методами моделювання, стратегіями навчання, процедурами тестування та механізмами використання.

Відповідно до загальноприйнятої методології розробки засобів розпізнавання [11, 12] результатом досліджень в області формалізації процесу розпізнавання ключових слів має бути концептуальна модель.

Мета роботи.

Метою даної наукової роботи являється розробка концептуальної моделі розпізнавання ключових слів у голосовому сигналі, що за рахунок детального формалізованого опису означено процесу забезпечує базис методології створення ефективних засобів розпізнавання ключових слів, їх використання.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.

В загальному випадку концептуальна модель - це абстрактне, логічне уявлення реального процесу або системи, яке визначає основні компоненти, їх взаємозв'язки та принципи функціонування. [11]

Концептуальна модель розпізнавання ключових слів може бути представлена як інтеграція низки взаємопов'язаних понять, які використовуються для опису даної проблематики, разом з їх властивостями та характеристиками. Концептуальна модель відображає основну ідею, спрямовуючу концепцію для систематизації даних. З огляду на необхідність розробки ефективних рішень для розпізнавання емоцій, використання дефініцій із сфери комп'ютерної і програмної інженерії, а також теорії розпізнавання образів є цілком виправданим. [7, 8, 9]

Розглядаючи комплексну оцінку ефективності концептуальної моделі підходу до розпізнавання ключових слів у голосовому сигналі, необхідно звернути увагу на цілий ряд взаємопов'язаних параметрів. [4, 5, 6] Ефективність в даному контексті охоплює не тільки технічні аспекти виконання програмної системи, але й оцінку ефективного використання ресурсів, а також адаптацію системи до заданих стандартів і умов експлуатації. Основні характеристики, які допомагають оцінити цю ефективність, включають:

1. Часова ефективність. Відображає здатність системи швидко реагувати на вхідні сигнали, забезпечуючи високу швидкість обробки і скорочення загального часу виконання задач.
2. Оптимізація ресурсів. Оцінює раціональність використання обчислювальних ресурсів та пам'яті, а також забезпечення енергоефективності системи.
3. Стандартизація та адаптивність. Визначає ступінь відповідності системи встановленим нормам і стандартам, її гнучкість у різних умовах використання та надійність у розпізнаванні ключових слів.

У процесі розвитку та оптимізації концептуальної моделі підходу до розпізнавання ключових слів у голосовому сигналі, ключовим елементом є створення єдиного термінологічного базису. Така стандартизація спрямована на включення в модель останніх досягнень у сфері обробки мовлення та аудіо сигналів, що гарантує високу адаптивність та відповідність до сучасних наукових реалій. З урахуванням різноманітності існуючих підходів та алгоритмів, слід визначити наступні основні поняття:

1. Ключові слова - це певні слова або фрази в голосовому сигналі, які мають важливе значення для певного контексту або завдання та які необхідно ідентифікувати та розпізнавати системою.
2. Голосовий сигнал (ГС) - складний акустичний сигнал, що генерується голосом людини, який може включати мову, інтонацію, емоційні нюанси та інші звукові характеристики. Голосовий сигнал зазвичай має діапазон від 75 Гц до 8000 Гц, що охоплює різноманітність голосових ефектів, від мовленнєвих тонів до емоційного колориту.
3. Мовний сигнал - це підтип голосового сигналу, присвячений передачі виключно мовної інформації. Мовний сигнал зосереджується на використанні мови, що включає слова, фрази та лінгвістичні структури, і зазвичай локалізується в діапазоні від 80 Гц до 2600 Гц, що оптимізовано для чіткого розуміння мовлення.
4. Акустичні характеристики - параметри, що описують звукові атрибути голосового сигналу, включаючи частоту, інтенсивність, тембр та інші звукові властивості.
5. Гаусівські змішані моделі (GMM) - це статистична модель, яка представляє сукупність гаусівських розподілів ймовірностей, кожен з яких відповідає окремій компоненті (або кластеру) в сукупності даних. У контексті обробки мови, GMM часто використовують для моделювання розподілу акустичних характеристик голосу, дозволяючи ефективно розпізнавати мовні фонемні чи інші одиниці.
6. Нейронна мережа (НМ) - система обробки інформації, заснована на принципах організації та функціонування біологічних нейронних мереж, зокрема мозку.
7. Марковські моделі - це клас статистичних моделей, що використовуються для прогнозування послідовності подій, де ймовірність кожної події залежить лише від стану, досягнутого в попередньому події. Приховані марковські моделі (НММ) є розширенням цієї ідеї, де спостережувані події залежать від внутрішніх факторів, що не можуть бути прямо спостережуваними.

8. Модель марковських ланцюгів (n-грами) - це статистичний метод моделювання мови, який передбачає, що ймовірність появи слова залежить лише від n-1 попередніх слів. Ця модель часто використовується для лінгвістичних моделей в розпізнаванні мови, щоб оцінити ймовірності послідовностей слів або фраз.

9. Модель глибокого навчання - спеціалізована архітектура нейронної мережі, що складається з багатьох шарів, здатна виявляти складні шаблони та закономірності у великих наборах даних.

10. Алгоритми машинного навчання - набір методів і технік, які використовуються для навчання моделей на основі даних, щоб здійснювати прогнозування або приймати рішення.

11. Акустична модель (AM) - це один з основних компонентів систем розпізнавання мови. AM відповідає за перетворення аудіосигналу в послідовність фонем або інших найменших одиниць мовлення. Вона використовує статистичні методи, такі як гаусівські змішані моделі (GMM) або глибокі нейронні мережі (DNN), для визначення ймовірності кожної фонемі в аудіосигналі.

12. Лінгвістична модель (LM) - це модель, який визначає ймовірність послідовності слів або фраз в мовленні. Вона використовує статистичні методи, такі як модель марковських ланцюгів (n-грами), для оцінки ймовірності різних словосполучень.

Ці поняття становлять основу для розуміння та розробки ефективних методів розпізнавання ключових слів, дозволяючи моделі адаптуватися до різних умов та потреб використання.

Для оцінювання ефективності процесу розпізнавання ключових слів у голосовому сигналі, важливо використовувати метрики що відображають особливості піддослідного процесу.[6] Ось перелік ключових метрик, які можуть бути застосовані у дослідницькому процесі для повноцінного аналізу ефективності:

1. Точність (Accuracy): визначає відсоток випадків, коли ключове слово було правильно розпізнане.

$$A = \frac{TP + TN}{TP + FN + FP + TN}, \quad (1)$$

де A - точність, TP - True Positives (правильно розпізнані ключові слова), FN - False Negatives (неправильно пропущені ключові слова), FP - False Positives (неправильно розпізнані ключові слова), TN - True Negatives (правильно не розпізнані ключові слова).

2. Повнота (Recall): міра того, скільки реальних випадків ключових слів було виявлено.

$$R = \frac{TP}{TP + FN}, \quad (2)$$

де R - повнота, TP - правильно розпізнані ключові слова, FN - неправильно пропущені ключові слова.

3. Прецизійність (Precision): відсоток випадків, коли виявлене ключове слово у голосовому сигналі, було дійсно ключовим

$$P = \frac{TP}{TP + FP}, \quad (3)$$

де P - повнота, TP - правильно розпізнані ключові слова, FP - неправильно розпізнані ключові слова.

4. F-міра (F-Measure): Гармонійне середнє між точністю та відновленням, що допомагає збалансувати ці дві метрики.

$$F = 2 \times \frac{P \times R}{P + R}, \quad (4)$$

де F - F-міра, P - точність виявлення, R - повнота.

5. Оцінка помилок (Error Rate): відсоток випадків, коли ключове слово було неправильно розпізнане або пропущене.

$$ER = \frac{FN + FP}{TP + FN + FP + TN}, \quad (5)$$

де ER - оцінка помилок, TP - True Positives (правильно розпізнані ключові слова), FN - False Negatives (неправильно пропущені ключові слова), FP - False Positives (неправильно розпізнані ключові слова), TN - True Negatives (правильно не розпізнані ключові слова).

6. Час відгуку (Response Time): час, який системі потрібно для розпізнавання ключового слова після його вимови. Цей показник важливий для систем реального часу.

$$RT = t_{end} - t_{start}, \quad (6)$$

де RT - час відгуку системи, t_{start} - момент початку обробки вхідного сигналу, t_{end} - момент завершення розпізнавання ключового слова та надання відповіді.

Використання цих метрик дозволяє провести всебічну оцінку моделі, враховуючи як її здатність правильно розпізнавати ключові слова, так і реагувати на помилкові спрацьовування. Це забезпечує комплексний підхід до оцінки точності та ефективності моделі.

В процесі створення і удосконалення концептуальної моделі підходу до розпізнавання ключових слів у голосовому сигналі необхідно здійснити аналіз та систематизацію компонентів цієї моделі. Цей процес включає визначення основних елементів, їхніх взаємозв'язків та впливу на загальну продуктивність системи. Нижче наведено загальний опис процесу:

1. Попередня обробка голосових даних: цей крок передбачає підготовку вхідних голосових сигналів до подальшої обробки, включаючи очищення від шумів та нормалізацію.

2. Формування параметрів додаткових даних: на цій стадії відбувається визначення та вибір ознак, які будуть використовуватися для навчання моделі.

3. Формування навчальної та тестової вибірки: розробка та структурування наборів даних, які будуть застосовуватися для навчання моделі та оцінювання її ефективності.

4. Вибір та налаштування моделі: вибір конкретної моделі чи набору моделей для розпізнавання ключових слів і налаштування їх параметрів.

5. Оцінювання ефективності моделі: проведення тестування для оцінки точності та надійності обраної моделі в розпізнаванні ключових слів.

6. Розпізнані слова: кінцевий результат процесу, де модель використовується для виявлення та класифікації ключових слів у реальному аудіосигналі.

Зазначені вище етапи відображають послідовність дій, які необхідно виконати для створення та валідації ефективної моделі розпізнавання ключових слів. Кожен крок, від підготовки даних до фінальної оцінки моделі, є важливою частиною процесу, що веде до точного ідентифікування ключових слів з голосових даних. Всі ці компоненти та їх взаємозв'язки детально зображені на рисунку нижче.

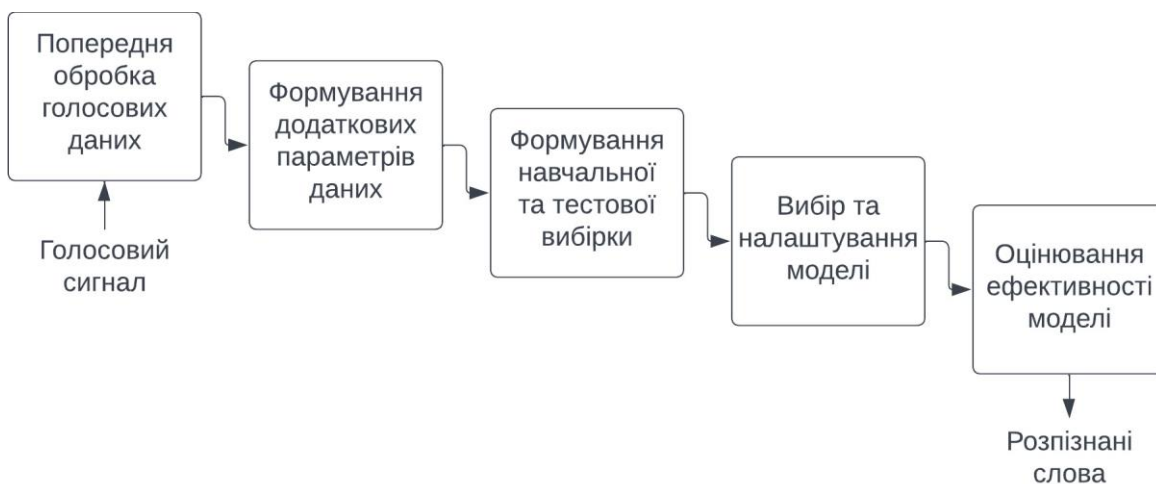


Рис. 1 – Діаграма декомпозиції

Слідом за визначенням основних етапів у створенні концептуальної моделі, було розроблено схему взаємодії компонентів системи розпізнавання ключових слів у голосовому сигналі, яка показана на рисунку 2. Вона ілюструє методіку взаємодії складових у процесі розробки даної системи, формалізуючи взаємодію з важливими складовими, як навчальні та тестові набори даних, вибір оптимальних параметрів моделі та подальшої оцінки її ефективності в контексті розпізнавання ключових слів.

Для розробки схеми взаємодії компонентів концептуальної моделі, яка використовує нейронні мережі, Марковські моделі або алгоритми машинного навчання для розпізнавання ключових слів у голосовому сигналі, можна використовувати таку структуру:

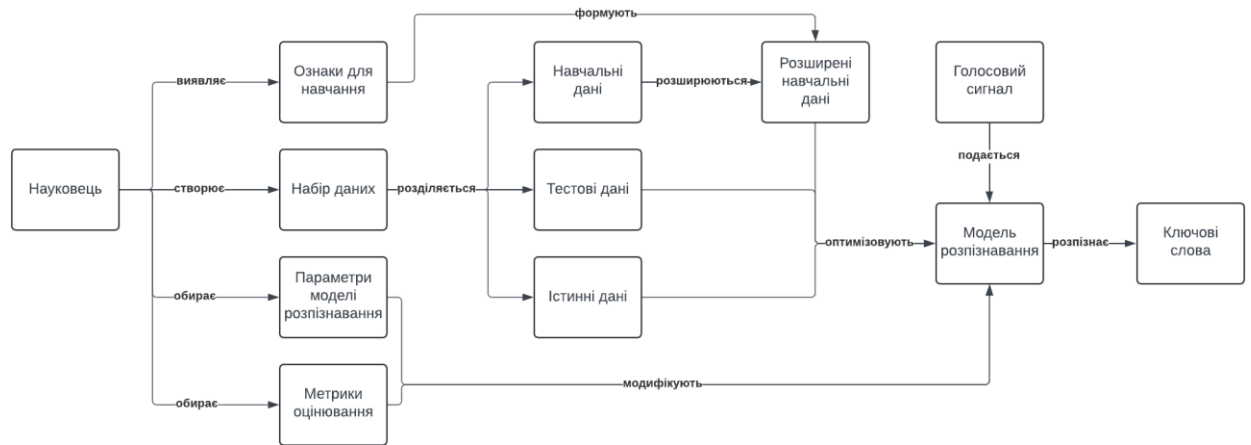


Рис. 2 – Схема взаємодії компонентів системи розпізнавання ключових слів у голосовому сигналі

На схемі, показаній на рисунку 2, детально проілюстровано процес взаємодії елементів системи, починаючи від вибору і формування навчальних даних, до оптимізації параметрів моделювання та врешті - оцінки ефективності моделі. Ці кроки критично важливі для створення ефективного інструменту розпізнавання ключових слів, здатного функціонувати у різноманітних застосуваннях, включаючи, але не обмежуючись, голосовими асистентами та системами автоматичного перекладу. Також у схемі враховано особливості реалізації концептуальної моделі для розпізнавання ключових слів у голосовому сигналі:

- обґрунтування вибору характеристик голосу (ознак для навчання), що використовуються для ідентифікації ключових слів;
- складність визначення оптимального виду моделі та її параметрів для точного розпізнавання ключових слів;
- необхідність удосконалення методів підготовки даних для навчання моделі, яка спрямована на визначення ключових слів.

Таким чином, у розробленій схемі передбачено можливість формування параметрів навчальних даних та їх валідації за допомогою експертної оцінки.

Аналіз структурних компонентів, [12] представлених на попередніх схемах, дає змогу зрозуміти, що ефективність процесу розпізнавання ключових слів у голосовому сигналі залежить від вибору базової моделі розпізнавання, серед яких можуть бути як нейронні мережі, так і інтеграція інших підходів, таких як статистичні методи та Марковські моделі.[9, 10, 11] Ефективність визначених операцій, що включають вибір характеристик голосового сигналу, формування та аналіз навчальних вибірок, а також тонке налаштування параметрів моделей, є вирішальними для точного розпізнавання мовлення. Ця систематизація дає змогу глибше зрозуміти взаємодію між різними компонентами моделі та їх вплив на загальну ефективність системи.

Також важливо уточнити основні операції, які впливають на ефективність розпізнавання емоцій.

1. Підготовка та навчання моделі:
 - а. Формування навчальних даних: створення репрезентативних наборів даних, які включають різні варіанти ключових слів, їхні контексти та варіації вимови.
 - б. Розробка навчальної бази даних: збір та систематизація навчальної бази, що охоплює різноманітність мовних та особливостей акцентів.
2. Обробка голосових даних:
 - а. Вибір ознак для навчання: визначення ключових акустичних та мовних ознак, таких як інтонація, темп, динаміка голосу.
 - б. Реєстрація та фільтрація даних: запис та відсіювання несуттєвих аспектів голосових сигналів для чіткого визначення ключових слів.
 - с. Аналіз даних за допомогою нейронних мереж або інших методів: застосування обраних моделей для виявлення та класифікації ключових слів.
3. Проектування та конфігурація моделі:

а. Вибір типу моделі: визначення підходящої архітектури (нейронна мережа, Марковські моделі, комбіновані підходи).

б. Налаштування параметрів моделі: точне налаштування гіперпараметрів для оптимізації точності та ефективності розпізнавання.

Згідно з фундаментальними принципами використання алгоритмів машинного навчання, ефективність систем розпізнавання ключових слів у голосовому сигналі значно зростає завдяки удосконаленню методик аналізу даних. Ці методики застосовуються для оптимізації обробки навчальних вибірок перед самим процесом навчання моделей, призначених для розпізнавання ключових слів.

Отже, формалізовано модель ефективності в контексті розпізнавання ключових слів може бути виражена за допомогою виразу виду:

$$E_{total} = f(E_{Dev}, E_{App}, E_{Proc}), (7)$$

де E_{total} – інтегральна ефективність системи розпізнавання; E_{Dev} , E_{App} – відповідно, ефективність розробки моделей та їхнього застосування; E_{Proc} – ефективність обробки голосових даних.

В свою чергу кожен зі складових можна деталізувати через залежність від інших складових:

$$E_{Dev} = f(w_1, w_2); (8)$$

$$E_{App} = f(w_3, w_4); (9)$$

$$E_{Proc} = f(w_5, w_6). (10)$$

де w_1 - формування навчальних даних: створення репрезентативних наборів даних, які включають різні варіанти ключових слів, їхні контексти та варіації вимови; w_2 - розробка навчальної бази даних: збір та систематизація навчальної бази, що охоплює різноманітність мовних та особливостей акцентів; w_3 - вибір ознак для навчання: визначення ключових акустичних та мовних ознак, таких як інтонація, темп, динаміка голосу; w_4 - реєстрація та фільтрація даних: запис та відсіювання несуттєвих аспектів голосових сигналів для чіткого визначення ключових слів; w_5 - проектування та конфігурація моделі: вибір типу моделі, визначення підходящої архітектури (нейронна мережа, Марковські моделі, комбіновані підходи); w_6 - налаштування параметрів моделі: точне налаштування гіперпараметрів для оптимізації точності та ефективності розпізнавання.

Кожна з цих складових сприяє загальній ефективності системи, забезпечуючи точність, адаптивність та ефективне використання ресурсів у процесі розпізнавання ключових слів.

Таким чином, вирази (7) - (10) складають аналітичне забезпечення концептуальної моделі в галузі визначення ефективності процесу розпізнавання ключових слів у голосовому сигналі. Вони деталізують інтегральну ефективність системи та її залежність від ефективності розробки моделей, їх застосування та обробки голосових даних. Кожна зі складових ефективності, у свою чергу, розкладається на конкретні операції, такі як формування навчальних даних, вибір ознак, проектування та налаштування моделей.

Висновки та перспективи подальшого дослідження.

У даній роботі було розроблено концептуальну модель підходу до розпізнавання ключових слів у голосовому сигналі. Ця модель забезпечує детальний та структурований опис процесу розробки засобів розпізнавання ключових слів, включаючи основні компоненти, їхню взаємодію та критичні фактори ефективності.

Запропонована концептуальна модель враховує низку викликів, таких як варіативність голосових сигналів, стійкість до шумів, обмеженість навчальних даних та вимоги до продуктивності в реальному часі. Модель інтегрує сучасні методи глибокого навчання, зокрема нейронні мережі, приховані марковські моделі та методи аугментації даних для підвищення точності розпізнавання.

Окремим внеском є розробка формалізованого підходу до оцінювання ефективності системи з використанням комплексних метрик, які охоплюють точність, швидкість реагування, надійність та масштабованість. Це дозволяє всебічно аналізувати та порівнювати різні методи та архітектури для задачі розпізнавання ключових слів.

Результати даного дослідження формують основу для подальшого вдосконалення систем розпізнавання мовлення, інтеграції з додатковими моделями, такими як розпізнавання мовця, а також адаптації до специфічних областей застосування.

Перспективні напрямки дослідження включають розвиток методів трансферного навчання, покращення якості та різноманітності навчальних даних, а також оптимізацію архітектур нейронних мереж для підвищення надійності та ефективності в динамічних умовах. Особливу увагу варто

звернути на застосування та вдосконалення методів динамічного програмування, зокрема алгоритму динамічного трансформування часу (DTW), що дозволяє ефективно справлятися з варіативністю та нестабільністю голосових сигналів. Ці методи забезпечують оптимальне вирівнювання часових послідовностей, що є критично важливим для точного розпізнавання мови в реальному часі. Розширення можливостей DTW та інших алгоритмів динамічного програмування сприятиме подоланню обмежень існуючих систем та підвищенню їхньої адаптивності до складних умов експлуатації.

Список бібліографічного опису

1. Umesh Dwivedia, T., Guptab, S., Upadhyayb, S. K., Shuklab, Y., & Ahujab, S. Automatic Speech Recognition System Using Hybrid Hidden Markov Model and Human Emotion Recognition System.
2. Rashmi, S., Hanumanthappa, M., & Reddy, M. V. (2018). Hidden Markov Model for speech recognition system—a pilot study and a naive approach for speech-to-text model. In *Speech and Language Processing for Human-Machine Communications: Proceedings of CSI 2015* (pp. 77-90). Springer Singapore.
3. Gunawan, A. (2010). English digits speech recognition system based on hidden Markov models. In *Proceedings of International Conference Computer*.
4. Deshmukh, A. M. (2020). Comparison of hidden markov model and recurrent neural network in automatic speech recognition. *European Journal of Engineering and Technology Research*, 5(8), 958-965.
5. Khurana, S., Laurent, A., Hsu, W. N., Chorowski, J., Lancucki, A., Marxer, R., & Glass, J. (2020). A convolutional deep markov model for unsupervised speech representation learning. *arXiv preprint arXiv:2006.02547*.
6. Chen, K. Y., Tsai, C. P., Liu, D. R., Lee, H. Y., & Lee, L. S. (2019). Completely unsupervised speech recognition by a generative adversarial network harmonized with iteratively refined hidden markov models. *arXiv preprint arXiv:1904.04100*.
7. Abdalla, M. I., & Ali, H. S. (2010). Wavelet-based mel-frequency cepstral coefficients for speaker identification using hidden markov models. *arXiv preprint arXiv:1003.5627*.
8. Belinkov, Y., & Glass, J. (2017). Analyzing hidden representations in end-to-end automatic speech recognition systems. *Advances in Neural Information Processing Systems*, 30.
9. Momeni, L., Afouras, T., Stafylakis, T., Albanie, S., & Zisserman, A. (2020). Seeing wake words: Audio-visual keyword spotting. *arXiv preprint arXiv:2009.01225*.
10. Berg, A., O'Connor, M., & Cruz, M. T. (2021). Keyword transformer: A self-attention model for keyword spotting. *arXiv preprint arXiv:2104.00769*.
11. Chen, G., Parada, C., & Sainath, T. N. (2015, April). Query-by-example keyword spotting using long short-term memory networks. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 5236-5240). IEEE.
12. Tereikovskiy, I., Hu, Z., Chernyshev, D., Tereikovska, L., Korystin, O., & Tereikovskiy, O. (2022). The method of semantic image segmentation using neural networks. *International Journal of Image, Graphics and Signal Processing*, 13(6), 1.
13. Toliupa, S., Tereikovskiy, I., Tereikovskiy, O., Tereikovska, L., Nakonechnyi, V., & Kulakov, Y. (2020, February). Keyboard dynamic analysis by Alexnet type neural network. In *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)* (pp. 416-420). IEEE.
14. Seilova, N., Tereikovskaya, L., & Nadgi, A. (2016). Conceptual model to ensure the efficiency of neural network recognition of phonemes in distance learning. *Seilova N, Tereikovskaya L, Nadgi A. Vestnik KazNRTU*, 114(2), 345-351.
15. Дичка, І., Терейковський, І., Самофалов, А., Терейковська, Л., & Романкевич, В. (2023). МНОЖИНА КРИТЕРІЇВ ЕФЕКТИВНОСТІ ФОРМУВАННЯ БАЗ ДАНИХ ЕМОЦІЙНО ЗАБАРВЛЕНИХ ГОЛОСОВИХ СИГНАЛІВ. *Електронне фахове наукове видання «Кібербезпека: освіта, наука, техніка»*, 1(21), 65-74.
16. I. A. Dychka, I. A. Tereikovskiy, O. S. Korovii, L. O. Tereikovska, and V. O. Romankevych, "Evaluation of the effectiveness of means for recognizing the emotional tonality of text fragments," *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, vol. 34 (73), no. 3, part 1, pp. 130-135, 2023.
17. Дичка, І.А., Терейковський, І.А., Дідус, А.В., Терейковська, Л.О., & Бояринова, Ю.С. (2023). Оцінка ефективності засобів розпізнавання ключових слів у голосовому сигналі. *Вчені записки*. <https://doi.org/10.32782/2663-5941/2023.3.1/19>

References

1. Umesh Dwivedia, T., Guptab, S., Upadhyayb, S. K., Shuklab, Y., & Ahujab, S. Automatic Speech Recognition System Using Hybrid Hidden Markov Model and Human Emotion Recognition System.
2. Rashmi, S., Hanumanthappa, M., & Reddy, M. V. (2018). Hidden Markov Model for speech recognition system—a pilot study and a naive approach for speech-to-text model. In *Speech and Language Processing for Human-Machine Communications: Proceedings of CSI 2015* (pp. 77-90). Springer Singapore.
3. Gunawan, A. (2010). English digits speech recognition system based on hidden Markov models. In *Proceedings of International Conference Computer*.
4. Deshmukh, A. M. (2020). Comparison of hidden markov model and recurrent neural network in automatic speech recognition. *European Journal of Engineering and Technology Research*, 5(8), 958-965.
5. Khurana, S., Laurent, A., Hsu, W. N., Chorowski, J., Lancucki, A., Marxer, R., & Glass, J. (2020). A convolutional deep markov model for unsupervised speech representation learning. *arXiv preprint arXiv:2006.02547*.

6. Chen, K. Y., Tsai, C. P., Liu, D. R., Lee, H. Y., & Lee, L. S. (2019). Completely unsupervised speech recognition by a generative adversarial network harmonized with iteratively refined hidden markov models. arXiv preprint arXiv:1904.04100.
7. Abdalla, M. I., & Ali, H. S. (2010). Wavelet-based mel-frequency cepstral coefficients for speaker identification using hidden markov models. arXiv preprint arXiv:1003.5627.
8. Belinkov, Y., & Glass, J. (2017). Analyzing hidden representations in end-to-end automatic speech recognition systems. *Advances in Neural Information Processing Systems*, 30.
9. Momeni, L., Afouras, T., Stafylakis, T., Albanie, S., & Zisserman, A. (2020). Seeing wake words: Audio-visual keyword spotting. arXiv preprint arXiv:2009.01225.
10. Berg, A., O'Connor, M., & Cruz, M. T. (2021). Keyword transformer: A self-attention model for keyword spotting. arXiv preprint arXiv:2104.00769.
11. Chen, G., Parada, C., & Sainath, T. N. (2015, April). Query-by-example keyword spotting using long short-term memory networks. In *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 5236-5240). IEEE.
12. Tereikovskiy, I., Hu, Z., Chernyshev, D., Tereikovska, L., Korystin, O., & Tereikovskiy, O. (2022). The method of semantic image segmentation using neural networks. *International Journal of Image, Graphics and Signal Processing*, 13(6), 1.
13. Toliupa, S., Tereikovskiy, I., Tereikovskiy, O., Tereikovska, L., Nakonechnyi, V., & Kulakov, Y. (2020, February). Keyboard dynamic analysis by Alexnet type neural network. In *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)* (pp. 416-420). IEEE.
14. Seilova, N., Tereikovskaya, L., & Nadgi, A. (2016). Conceptual model to ensure the efficiency of neural network recognition of phonemes in distance learning. Seilova N, Tereikovskaya L, Nadgi A. *Vestnik KazNRTU*, 114(2), 345-351.
15. Dychka, I., Tereikovskiy, I., Samofalov, A., Tereikovska, L., & Romankievych, V. (2023). A Set of Effectiveness Criteria for Creating Databases of Emotionally Colored Voice Signals. *Electronic Professional Scientific Publication "Cybersecurity: Education, Science, Technique"*, 1(21), 65-74.
16. I. A. Dychka, I. A. Tereikovskiy, O. S. Korovii, L. O. Tereikovska, and V. O. Romankevych, "Evaluation of the effectiveness of means for recognizing the emotional tonality of text fragments," *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, vol. 34 (73), no. 3, part 1, pp. 130-135, 2023.
17. Dychka, I. A., Tereikovskiy, I. A., Didus, A. V., Tereikovska, L. O., & Boyarynova, Y. Ye. (2023). Evaluation of the Effectiveness of Tools for Recognizing Keywords in Voice Signals. *Scholarly Notes*. <https://doi.org/10.32782/2663-5941/2023.3.1/19>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-10>

УДК 004.021

Дмитренко Тарас Васильович, аспірант

<https://orcid.org/0000-0002-2339-2903>

Національний авіаційний університет, м. Київ, Україна

ЗАСТОСУВАННЯ БПЛА З ТЕХНОЛОГІЄЮ ВЕЙВЛЕТ-ПЕРЕТВОРЕННЯ У ВІЙСЬКОВИХ ОПЕРАЦІЯХ: РОЗВІДКА, ВИЯВЛЕННЯ ЦІЛЕЙ ТА НАВІГАЦІЯ

Дмитренко Т.В. Застосування БПЛА з технологією вейвлет-перетворення у військових операціях: розвідка, виявлення цілей та навігація. У цій роботі основна увага зосереджена на використанні безпілотних літальних апаратів (БПЛА), робота яких ґрунтується на використанні вейвлет-перетворення як базової основи для забезпечення високої роздільної здатності зображень і реалізації великого оглядового району розвідки. Використання даної архітектури у військових умовах мотивується її вмінням здійснювати розвідку, виявлення цілей, навігацію та моніторинг територій з високою точністю, що дозволяє військовим здійснювати оперативне планування та приймати ефективні рішення на основі зібраної інформації. Проаналізовано функціональність вейвлет-перетворення у навігації БПЛА у просторі на прикладі імплементації даної технології у механізм знешумлювання у гіроскопі БПЛА. Описано дискретну вейвлет-функцію у дискретному вейвлет-перетворенні та наведено математичний аспект знаходження коефіцієнту дискретної вейвлет-трансляції. Наголошено, що метою вейвлет-перетворення у контексті знешумлення зображень є виділення корисного сигналу та усунення сигналу перешкод у вихідному сигналі. Охарактеризовано вейвлет-знешумлення та основну увагу приділено пороговому методу знешумлення з повним описанням його етапів. Зазначається, що вейвлет-перетворення являє собою математичний інструмент для обробки сигналів і зображень, пропонує унікальні переваги в розширенні можливостей БПЛА у сфері розвідки. Застосовуючи алгоритми вейвлет-перетворення до даних, отриманих бортовими датчиками, БПЛА можуть ефективно обробляти й аналізувати одержані зображення високої роздільної здатності та повну обізнаність про ситуацію на величезних географічних територіях. Загалом, у ході даного дослідження проводиться огляд інтеграції інноваційної технології вейвлет-перетворення в системи БПЛА, що революціонує підхід до ведення військових операцій в умовах сьогодення. Описаний функціонал допомагає військовим планувати маршрути та вибирати найбільш ефективні шляхи пересування з урахуванням поточних змін.

Ключові слова: вейвлет, функція, дистанційне зондування, обробка сигналів, збір розвідувальних даних, ідентифікація цілі.

Dmitrenko T. Application Of Uavs Based On The Wavelet Transform Technology In Military Operations: Scouting, Target Location And Navigation. This paper focuses on the use of unmanned aerial vehicles (UAVs), which work based on the use of wavelet transform as a basic framework to provide high-resolution images and realize a large reconnaissance observation area. The use of this architecture in military conditions is motivated by its ability to carry out reconnaissance, target detection, navigation and monitoring of territories with high accuracy, which allows the military to carry out operational planning and make effective decisions based on the collected information. The functionality of the wavelet transformation in the navigation of the UAV in space was analyzed using the example of the implementation of this technology in the denoising mechanism in the gyroscope of the UAV. The discrete wavelet function in the discrete wavelet transform is described and the mathematical aspect of finding the coefficient of the discrete wavelet translation is given. It is emphasized that the purpose of wavelet transformation in the context of image denoising is to extract the useful signal and eliminate the interference signal in the original signal. Wavelet denoising is characterized and the main attention is paid to the threshold method of denoising with a full description of its stages. It is noted that the wavelet transformation is a mathematical tool for processing signals and images, offering unique advantages in expanding the capabilities of UAVs in the field of intelligence. By applying wavelet transform algorithms to data received by onboard sensors, UAVs can efficiently process and analyze the resulting high-resolution images and complete situational awareness over vast geographic areas. In general, in the course of this study, a review of the integration of innovative wavelet-transformation technology into UAV systems is conducted, which revolutionizes the approach to conducting military operations in today's conditions. The described functionality helps the military to plan routes and choose the most effective ways of movement, taking into account current changes.

Key words: wavelet, function, remote sensing, signal processing, intelligence gathering, identification purposes.

Вступ та постановка проблеми. На даний час БПЛА широко використовуються у військових розвідувальних операціях завдяки їх здатності здійснювати зйомку з великої висоти і отримувати зображення з високою роздільною здатністю. Це дозволяє військовим отримувати детальну та точну інформацію про місцевість, ворожі позиції та інші важливі об'єкти на місцевості. Одночасно з цим, завдяки використанню вейвлет-перетворення, БПЛА можуть забезпечувати високу якість зображень навіть при поганих погодних умовах або в умовах обмеженої видимості.

Також однією з ключових функцій БПЛА з технологією вейвлет-перетворення є їх здатність до виявлення різних типів цілей на місцевості. Вони можуть виявляти ворожі технічні споруди, транспортні засоби, військову техніку та інші об'єкти, які мають стратегічне значення для проведення військових операцій. Завдяки використанню технології вейвлет-перетворення, БПЛА можуть забезпечувати високу точність виявлення цілей і детальний аналіз їх характеристик, що дозволяє військовим планувати ефективні операції з їх виявлення та знищення. У сучасних

військових операціях використання БПЛА з технологією вейвлет-перетворення відіграє ключову роль у забезпеченні розвідки, виявлення цілей та навігації. Ця технологія дозволяє здійснювати моніторинг територій з високою точністю та ефективністю.

Однією з ключових переваг використання БПЛА з технологією вейвлет-перетворення є їх можливість оперативної реакції на зміни в ситуації на місцевості. Швидкий доступ до розвідувальної інформації дозволяє військовим ефективно планувати дії та приймати стратегічні рішення. Більше того, БПЛА можуть здійснювати моніторинг руху ворожих військ та виявляти потенційні загрози для власних військ.

Аналіз останніх досліджень і публікацій. У дослідженні [1] наводиться аналіз методів розпізнавання зображень з метою автоматизації виявлення об'єктів і визначення їх стану. На основі аналізу розроблено покращений метод розпізнавання зображень об'єктів моніторингу згортковою нейронною мережею з використанням дискретного вейвлет-перетворення. В основі методу лежить задача автоматизації обробки зображень в БПЛА. Працездатність запропонованого методу було перевірено на прикладі обробки зображення (літака, танка, вертольота), отриманого оптичною системою безпілотного літального апарату. Дискретне вейвлет-перетворення було використано для створення бази даних вейвлет-зображень об'єктів і навчання згорткової нейронної мережі на їх основі. Це дозволило підвищити ефективність розпізнавання об'єктів моніторингу та автоматизувати заданий процес. Ефективність удосконаленого методу досягається попереднім розкладанням та апроксимацією цифрового зображення об'єкта спостереження дискретним вейвлет-перетворенням. Етапи даного методу включають побудову бази даних вейвлет-зображень заданих зображень і навчання згорткової нейронної мережі. Ефективність розпізнавання зображень об'єктів моніторингу вдосконаленим методом була перевірена на згортковій нейронній мережі, яка навчалася на зображеннях 300 об'єктів спостереження. При цьому час прийняття рішення на основі запропонованого методу зменшився в середньому з 0,7 до 0,84 с порівняно зі штучними нейронними мережами ResNet і ConvNets.

У роботі [2] було показано, що дискретний тип вейвлет-перетворення є одним із найкращих методів стиснення зображень. Він передбачає математичну нотацію для кодування інформації відповідно до необхідного рівня деталізації. Як базис DWT запропоновано вейвлет-функції Хаара. Надмірність коефіцієнтів деталізації DWT зменшується через порогові значення. Якість стиснутих зображень оцінювали за допомогою коефіцієнтів стиснення та PSNR. Експериментальні результати показують, що запропонована процедура забезпечує достатньо високий ступінь стиснення порівняно з іншими методами порогового стиснення. Виявлено, що недоліком даного методу є неможливість його використання для розпізнавання зображень об'єктів спостереження.

Крім того, варто зазначити праці наступних науковців: Дж. Сміт, Р. Джонсон [3], П. Сінгер [4], Ч. Чуй [5], Дж. Коллінз [6], П. Матер, М. Кох [7], С. Балакрішнан, А. Джамір, А. Рей [8], Т. Хуе, М. Жонг, Г. Лі [9], Х. Ян, Д. Лін, Ф. Чжан, Т. Сонг, Т. Цзян [10], Ф. Ван, Дж. Патрік [11], М. Проценко, Т. Курцеїтов, З. Бржезьська [12], Р. Крішнасвами, С. Нірмала Деві [13], Е. Аханону, М. Марцелін, А. Білгін [14], М. Павлунько, Д. Мороз, З. Бржезьська [15], С. Тепаде, Дж. Деван, С. Ерандол, Р. Джадхав [16], А. Пол, Т. Хан, П. Поддер, Р. Ахмед, М. Рахман, М. Хан [17] та інших.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з інтеграцією технології вейвлет-перетворення у архітектуру БПЛА задля їх використання у військових операціях все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Постановка завдання. Метою роботи є дослідження технології вейвлет-перетворення у БПЛА у контексті застосування їх у військових операціях.

Викладення основного матеріалу дослідження. Задля розуміння яким самим чином вейвлет-перетворення сприяє кращій роботі БПЛА в умовах бойових дій, необхідно визначитись з дефініцією даного терміну. Вейвлет-перетворення являє собою математичний метод, який використовується для аналізу сигналів і зображень у різних масштабах або роздільній здатності. Воно відрізняється від традиційного перетворення Фур'є тим, що в ньому використовуються невеликі локалізовані хвилі, які називаються вейвлетами, які масштабуються та зміщуються відповідно до різних характеристик у сигналі. Ця властивість локалізації має вирішальне значення, оскільки дозволяє вейвлет-перетворенню фіксувати перехідні характеристики та деталі ефективніше, ніж аналіз Фур'є.

Процес вейвлет-перетворення передбачає розкладання сигналу на ряд вейвлет-коефіцієнтів, кожен з яких представляє силу сигналу в різних масштабах і положеннях. Такого розкладання можна досягти за допомогою як безперервного вейвлет-перетворення (Continuous Wavelet Transform – CWT), так і дискретного вейвлет-перетворення (Discrete Wavelet Transform – DWT).

Головні відмінності безперервного та дискретного вейвлет-перетворення:

– безперервне вейвлет-перетворення передбачає згортання сигналу за допомогою безперервної групи вейвлетів, кожен з яких масштабується та зміщується, щоб охопити різні частотні діапазони та часові інтервали, в той час як дискретне вейвлет-перетворення розкладає сигнал на дискретні масштаби та позиції за допомогою серії вейвлет-фільтрів, зазвичай реалізованих за допомогою серії фільтрів високих і нижніх частот із подальшим зниженням дискретизації;

– отримані вейвлет-коефіцієнти забезпечують безперервне представлення частотного вмісту сигналу в різних масштабах і положеннях; проте у випадку з дискретним вейвлет-перетворенням, процес декомпозиції сигналів генерує представлення сигналу з різною роздільною здатністю, причому кожен рівень охоплює інший діапазон частот і рівень деталізації;

– з одного боку, безперервне вейвлет-перетворення є корисним для аналізу сигналів із безперервним частотним вмістом і для програм, що вимагають точної частотно-часової локалізації; з іншого боку, дискретний тип є обчислювально ефективним і широко використовується в практичних програмах, таких як стиснення зображень, усунення шумів і виділення ознак.

Найбільш значимі аспекти застосування БПЛА з інтегрованою технологією вейвлет-перетворення у військових операціях, що потребують детального аналізу та вирішення та вказують на потенційні можливості для покращення ефективності та функціональних можливостей, являють собою:

– вивчення підходів та шляхів збільшення автономності БПЛА у військових операціях, що включає в себе розробку та впровадження систем штучного інтелекту для автоматизації процесів розвідки та виявлення цілей;

– дослідження та впровадження нових систем передачі даних для забезпечення безперервного зв'язку між БПЛА та командним центром у будь-яких умовах;

– вивчення можливостей інтеграції БПЛА з вейвлет-перетворенням з іншими видами розвідувальних засобів та технологій для створення комплексних систем моніторингу та розвідки;

– дослідження можливостей швидкої адаптації БПЛА до змін в бойовій обстановці та розробка стратегій ефективного використання їх у реальному часі.

Дослідження застосування БПЛА у військових операціях підтверджує їх важливу роль у забезпеченні розвідки, виявлення цілей та навігації. Застосування таких БПЛА дозволяє отримувати детальну та високоякісну розвідувальну інформацію з великої висоти, що робить їх незамінними в сучасному військовому плануванні та стратегічних операціях, щодо планування воєнної операції.

Застосування БПЛА з технологією вейвлет-перетворення в розвідувальних операціях дозволяє військовим здійснювати оперативний моніторинг територій, виявлення ворожих сил та об'єктів, а також забезпечує необхідну інформацію для прийняття стратегічних рішень.

Також, БПЛА з технологією вейвлет-перетворення характеризуються високою точністю виявлення цілей на місцевості, що робить їх ефективними засобами для планування та здійснення військових операцій.

Основні напрямки покращення навігації БПЛА у просторі завдяки використанню вейвлет-перетворення:

– виділення основних деталей: вейвлет-перетворення має змогу виокремлювати важливі частини зображень, які захоплюються камерами БПЛА;

– стиснення та редукція даних, оскільки БПЛА часто наділені дуже обмеженими розрахунковими ресурсами та низькою пропускну здатністю;

– виявлення та розпізнавання об'єктів спостереження: завдяки декомпозиції зображень у рухливі частотні компоненти, вейвлет-перетворення має змогу виокремлювати специфічні об'єкти та патерни, спрощуючи БПЛА виявлення та розпізнавання перешкод, особливостей рельєфу тощо;

– реєстрація та вирівнювання зображень: завдяки даній операції, використовуючи зображення, зроблені в різний час або з різних ракурсів, БПЛА може створювати точні карти або відстежувати зміни в навколишньому середовищі з часом.

– зменшення шуму: зображення БПЛА можуть бути схильні до утворення шуму, особливо у складних умовах навколишнього середовища. Вейвлет-перетворення може знешумлювати зображення, відокремлюючи шум від сигналу в частотній області, покращуючи якість навігаційних даних, отриманих від датчиків БПЛА.

Проаналізувати функціональність вейвлет-перетворення у навігації БПЛА у просторі можливо на прикладі імплементації даної технології у механізм знешумлювання у гіроскопі БПЛА.

Вейвлет-перетворення є методом частотно-часового аналізу локалізації. Його розмір вікна є фіксованим, але форма піддається змінюванню. Крім того воно має високу частотну роздільну здатність і нижчу часову роздільну здатність у низькочастотній частині сигналу. А у високочастотній частині сигналу воно має високу роздільну здатність за часом і нижчу роздільну здатність за частотою.

Дискретна вейвлет-функція $\psi_{j,k}(t)$ у дискретному вейвлет-перетворенні може бути виражена як:

$$\psi_{j,k}(t) = s_0^{-j/2} \psi\left(\frac{t - ks_0^j \tau_0}{s_0^j}\right) = s_0^{-j/2} \psi(s_0^{-j} t - k\tau_0)$$

Коефіцієнт дискретної вейвлет-трансляції можна виразити як:

$$C_{j,k} = \int_{-\infty}^{\infty} f(t) \psi_{j,k}^*(t) dt = \langle f, \psi_{j,k} \rangle$$

Функцію реконструкції дискретного вейвлет-перетворення можна виразити як:

$$f(t) = C \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} C_{j,k} \psi_{j,k}(t)$$

Де s позначає масштабний коефіцієнт, τ позначає коефіцієнт трансляції; $f(t)$ представляє сигнальну функцію.

Метою вейвлет-перетворення у контексті знешумлення зображень є виділення корисного сигналу та усунення сигналу перешкод у вихідному сигналі. Іншими словами, корисний сигнал і шумовий сигнал розділені методом вейвлет-перетворення. Існує 4 поширених методи вейвлет-знешумлення:

– пороговий метод, який також має назву вейвлет-усадки. Основну ідею цього методу можна описати таким чином: вейвлет-коефіцієнти мають різні характери в окремих вейвлет-шкалах. Відповідно до цієї характеристики сигналу та шуму сигнали шуму перетворюються за допомогою вейвлет-перетворення в певних вейвлет-шкалах. Відповідно до певних стратегій обробки порогових значень для обробки вейвлет-коефіцієнтів, коефіцієнти, які перевищують порогове значення, зберігаються (метод жорсткого порогового значення) або скорочуються (метод м'якого порогового значення). Коефіцієнти, менші за порогове значення, вважаються шумом і зводяться до нульових значень. Потім на основі цих вейвлет-коефіцієнтів вихідний сигнал реконструюється за допомогою зворотного вейвлет-перетворення. І цей метод вимагає припущення, що сигнал шуму є білим шумом Гауса;

– метод вейвлет-декомпозиції та реконструкції, також метод Маллета. За допомогою нього сигнал із шумом у масштабі розкладається на різні смуги частот, та смуги з шумом зводяться до нульових значень, наприкінці сигнал реконструюється за допомогою вейвлет-методу. Цей метод видаляє шумові сигнали разом із корисними сигналами. Таким чином, реконструйований сигнал може виявитись спотвореним у кінцевому підсумку;

– метод максимуму модуля. У різних масштабах вейвлетів цей метод використовує особливості варіації максимальних значень модуля вейвлет-перетворення для зменшення шуму сигналу. Видаляються крайні точки, амплітуда яких зменшується зі збільшенням масштабу сигналу. Зберігаються крайні точки, амплітуда яких зростає зі збільшенням масштабу сигналу. При використанні методу змінної проєкції, вихідний сигнал реконструюється із знешумленої діаграми максимумів модуля, що у свою чергу призводить до знешумлення сигналу;

– інваріантний метод трансляції – являє собою вдосконалену версію методу порогового знешумлення. В ході використання цього методу, шумові сигнали приймаються за n разів циклічного зсуву. Перетворені сигнали прибираються за допомогою методу порогового усунення

шумів. Зрештою, знешумлений сигнал врівноважується. Цей метод має меншу середньоквадратичну похибку та покращує співвідношення сигнал/шум.

Увага у теперішньому аналізі буде зосереджена на пороговому методі знешумлення. Застосування даного методу складається з наступних етапів:

– обрання функції вейвлет. Тоді сигнал шуму $y_i, i = 0, 1, \dots, N - 1$ є дискретним при використанні вейвлет-перетворення. Отримання групи коефіцієнтів вейвлет-перетворення $d_{j,k}$; індекс j являє собою вейвлет-шкалу;

– порогове значення вейвлет-функції перетворює коефіцієнти $d_{j,k}$. Для роботи з коефіцієнтами можливо використовувати жорсткий поріг, м'який поріг або інший пороговий метод. Після обчислення отримують нові коефіцієнти вейвлет-перетворення $d_{j,k}$.

Оцінка жорсткого порогу визначається наступним співвідношенням:

$$\hat{d}_{j,k} = \begin{cases} d_{j,k}, & |d_{j,k}| \geq \lambda_j \\ 0, & |d_{j,k}| < \lambda_j \end{cases}$$

Оцінка м'якого порогу визначається таким чином:

$$\hat{d}_{j,k} = \begin{cases} \text{sgn}(d_{j,k})(|d_{j,k}| - \lambda_j), & |d_{j,k}| \geq \lambda_j \\ 0, & |d_{j,k}| < \lambda_j \end{cases}$$

λ_j виражає порогову постійну;

– вейвлет-реконструкція. Використовуючи формули, зворотні дискретному вейвлет-перетворенню, можна отримати сигнал із знешумленням \hat{y}_i .

Висновки. БПЛА з інтегрованою технологією вейвлет-перетворення забезпечують можливість дистанційного виявлення та ідентифікації цілей в умовах бойових дій. Така здатність дозволяє їм проникати в недоступні для людей місця, здійснюючи розвідку, не ризикуючи життям військовослужбовців. Таким чином, під час проведення військової операції відкривається можливість уникнення непотрібних втрат.

Одночасно БПЛА, оснащені технологією вейвлет-перетворення, можуть орієнтуватися в складних умовах, досліджуючи території та створюючи тривимірні моделі ландшафту. Цей функціонал допомагає військовим планувати маршрути та вибирати найбільш ефективні шляхи пересування з урахуванням поточних змін.

Підсумовуючи, використання безпілотних літальних апаратів із технологією вейвлет-перетворення у військових операціях є критично важливим компонентом сучасної військової стратегії. Інформація, отримана під час їхнього розгортання, надає військовим точні дані, необхідні для прийняття стратегічних рішень, одночасно знижуючи ризики для життя військовослужбовців та підвищуючи оперативну ефективність.

Список бібліографічного опису

1. Slyusar V., Protsenko M., Chernukha A., Gornostal S., Rudakov S., Shevchenko S., Chernikov O., Kolpachenko N., Timofeyev V., Artiukh R. Construction of an advanced method for recognizing monitored objects by a convolutional neural network using a discrete wavelet transform. *Eastern-European Journal of Enterprise Technologies*. 2021. №4. P. 65 - 77. DOI:10.15587/1729-4061.2021.238601.
2. Nashat A., Hussain Hassan, N. Image compression based upon Wavelet Transform and a statistical threshold. *2016 International Conference on Optoelectronics and Image Processing (ICOIP)*. 2016. DOI:10.1109/optip.2016.7528492
3. Smith J., Johnson R. *Military Robotics: Latest Trends and Spatial Wavelet Interpolation Techniques*. 2020.
4. Singer P. *Drone Warfare: The Development of Unmanned Aerial Conflict*. 2014.
5. Chui C. *Introduction to Wavelets and Wavelet Transforms: A Primer*. 2016.
6. Collins J. *Modern Military Strategy: An Introduction*. 2014.
7. Mather P., Koch M. *Remote Sensing and GIS for Military Applications*. 2019.
8. Balakrishnan S., Jamir A., Ray A. *Drone Technology and Applications*. 2021.
9. Xue T., Zhong M., Li G. Wavelet transform and parity space based actuator fault detection for unmanned aerial vehicle. 2016 № 33. P. 1193-1199. DOI:10.7641/CTA.2016.60034.
10. Yang X., Lin D., Zhang F., Song T., Jiang T. High Accuracy Active Stand-off Target Geolocation Using UAV Platform. *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*. 2019 DOI:10.1109/icsidp47821.2019.9172919
11. Van F., Patrick J. An introduction to digital images. *Discrete Wavelet Transformations: An Elementary Approach with Applications*. Wiley. 2019. P. 69–123. DOI:10.1002/9781119555414.ch3
12. Protsenko M., Kurtseitov T., Pavlunko M., Brzhevskaya Z. Wavelet transforms application for digital signal analysis. Use of packet wavelet transformation for radio signals processing. *Modern Information Security*. 2018. №3 (35). P. 11–15. DOI:10.31673/2409-7292.2018.031115
13. Krishnaswamy R., NirmalaDevi S. Efficient medical image compression based on integer wavelet transform. 2020

Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII). 2020. DOI:10.1109/icbsii49132.2020.9167597

14. Ahanonu E., Marcellin M., Bilgin A. Lossless Multi-component Image Compression Based on Integer Wavelet Coefficient Prediction using Convolutional Neural Networks. *2020 Data Compression Conference (DCC)*. 2020. DOI:10.1109/dcc47342.2020.00043

15. Protsenko M., Pavlun'ko M., Moroz D., Brzhevs'ka Z. Procedure of signal filtering based on wavelet transformation. *Modern Information Security*. 2019. № 1 (37). P. 64–69. DOI:10.31673/2409-7292.2019.016469

16. Thepade S., Dewan J., Erandole S., Jadhav S. Extended performance comparison of self mutated hybrid wavelet transforms in image compression with hybrid wavelet transforms & orthogonal transforms. *2015 Global Conference on Communication Technologies (GCCT)*. 2015. DOI: 10.1109/gcct.2015.7342675

17. Paul A., Khan T., Podder P., Ahmed R., Rahman M., Khan M. Iris image compression using wavelets transform coding. *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. 2015. DOI: <https://doi.org/10.1109/spin.2015.7095407>

References

1. Slyusar, V., Protsenko, M., Chernukha, A., Gornostal, S., Rudakov, S., Shevchenko, S., Chernikov, O., Kolpachenko, N., Timofeyev, V., & Artiukh, R. (2021). Construction of an advanced method for recognizing monitored objects by a convolutional neural network using a discrete wavelet transform. *Eastern-European Journal of Enterprise Technologies*, 4, 65-77. <https://doi.org/10.15587/1729-4061.2021.238601>

2. Nashat, A., & Hassan, N. H. (2016). Image compression based upon Wavelet Transform and a statistical threshold. *2016 International Conference on Optoelectronics and Image Processing (ICOIP)*. <https://doi.org/10.1109/OPTIP.2016.7528492>

3. Smith, J., & Johnson, R. (2020). Military Robotics: Latest Trends and Spatial Wavelet Interpolation Techniques.

4. Singer, P. (2014). *Drone Warfare: The Development of Unmanned Aerial Conflict*.

5. Chui, C. (2016). *Introduction to Wavelets and Wavelet Transforms: A Primer*.

6. Collins, J. (2014). *Modern Military Strategy: An Introduction*.

7. Mather, P., & Koch, M. (2019). *Remote Sensing and GIS for Military Applications*.

8. Balakrishnan, S., Jamir, A., & Ray, A. (2021). Drone Technology and Applications.

9. Xue, T., Zhong, M., & Li, G. (2016). Wavelet transform and parity space based actuator fault detection for unmanned aerial vehicle. *2016 № 33*, 1193-1199. <https://doi.org/10.7641/CTA.2016.60034>

10. Yang, X., Lin, D., Zhang, F., Song, T., & Jiang, T. (2019). High Accuracy Active Stand-off Target Geolocation Using UAV Platform. *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP)*. <https://doi.org/10.1109/ICSIDP47821.2019.9172919>

11. Van, F., & Patrick, J. (2019). An introduction to digital images. In *Discrete Wavelet Transformations: An Elementary Approach with Applications* (pp. 69–123). Wiley. <https://doi.org/10.1002/9781119555414.ch3>

12. Protsenko, M., Kurtseitov, T., Pavlun'ko, M., & Brzhevska, Z. (2018). Wavelet transforms application for digital signal analysis. Use of packet wavelet transformation for radio signals processing. *Modern Information Security*, 3(35), 11–15. <https://doi.org/10.31673/2409-7292.2018.031115>

13. Krishnaswamy, R., & NirmalaDevi, S. (2020). Efficient medical image compression based on integer wavelet transform. *2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII)*. <https://doi.org/10.1109/ICBSII49132.2020.9167597>

14. Ahanonu, E., Marcellin, M., & Bilgin, A. (2020). Lossless Multi-component Image Compression Based on Integer Wavelet Coefficient Prediction using Convolutional Neural Networks. *2020 Data Compression Conference (DCC)*. <https://doi.org/10.1109/DCC47342.2020.00043>

15. Protsenko, M., Pavlun'ko, M., Moroz, D., & Brzhevs'ka, Z. (2019). Procedure of signal filtering based on wavelet transformation. *Modern Information Security*, 1(37), 64–69. <https://doi.org/10.31673/2409-7292.2019.016469>

16. Thepade, S., Dewan, J., Erandole, S., & Jadhav, S. (2015). Extended performance comparison of self mutated hybrid wavelet transforms in image compression with hybrid wavelet transforms & orthogonal transforms. *2015 Global Conference on Communication Technologies (GCCT)*. <https://doi.org/10.1109/GCCT.2015.7342675>

17. Paul, A., Khan, T., Podder, P., Ahmed, R., Rahman, M., & Khan, M. (2015). Iris image compression using wavelets transform coding. *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*. <https://doi.org/10.1109/SPIN.2015.7095407>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-11>

УДК 519.816(004.056)

Добришин Юрій Євгенович, к.т.н., доцент

<https://orcid.org/0000-0003-2473-9507>

Національна академія Служби безпеки України, м. Київ

ФОРМАЛІЗАЦІЯ ПРОЦЕСУ ПРОЕКТУВАННЯ СИСТЕМ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ЩОДО ПІДВИЩЕННЯ РІВНЯ ЗАХИСТУ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

Добришин Ю.Є. Формалізація процесу проектування систем підтримки прийняття рішень щодо підвищення рівня захисту програмного забезпечення. В роботі представлені дослідження можливості формалізації процесу проектування систем підтримки прийняття рішень щодо відновлення пошкодженого програмного забезпечення внаслідок впливу кібератак. Проблема проектування вказаних систем пов'язана з тим, що внаслідок дії кібератак складно однозначно визначити їх тип, стан та характер пошкодженого програмного забезпечення автоматизованих систем та комплексів, суперечливість відомостей про дефекти, способи їх виявлення та усунення, а також призначити оптимальний перелік технологічних операцій з відновлення програмних компонентів, що експлуатуються. Рішення описаної задачі потребує розробки та застосування формалізованих методів щодо проектування компонентів систем підтримки прийняття рішень, призначених для відновлення пошкодженого програмного забезпечення внаслідок дії кібератак, та адаптування спеціального математичного апарату щодо класифікації інформації та технологій відновлення дефектів пошкодженого програмного забезпечення шляхом прийняття рішень відносно завдань, які складно формалізуються. Процеси проектування повинні включати аналіз проектної діяльності у галузі захисту інформації; оцінку інформації, що використовується протягом експлуатації різних автоматизованих систем; характеристику та можливості засобів захисту інформації, що здійснюють виявлення загроз та кібератак; аналіз дефектів програмного забезпечення; розробку алгоритмів та методичних матеріалів, що функціонують та використовуються у системах підтримки прийняття рішень.

Ключові слова: системи підтримки прийняття рішень, процес проектування, формалізація, кібератака, дефект програмного забезпечення, автоматизовані системи та комплекси, кіберзагроза, кібербезпека

Dobryshyn Yu. Formalization of the design process of decision-making support systems for increasing of the level of software protection. The work presents research opportunities formalization of the design process of decision-making support systems for the restoration of damaged software due to the impact of cyber attacks. The problem of designing the specified systems is related to the fact that due to cyberattacks it is difficult to unambiguously determine their type, state and nature of damaged software of automated systems and complexes, inconsistency of information about defects, methods of their detection and elimination, as well as to assign an optimal list of technological operations with recovery of software components that are in use. The solution to the described problem requires the development and application of formalized methods for the design of components of decision-making support systems intended for the restoration of damaged software as a result of cyber-attacks, and the adaptation of a special mathematical apparatus for the classification of information and technologies for repairing defects in damaged software by making decisions regarding tasks that are difficult are formalized. Design processes should include analysis of design activities in the field of information protection; evaluation of information used during operation of various automated systems; characteristics and capabilities of information protection tools that detect threats and cyberattacks; analysis of software defects; development of algorithms and methodical materials that function and are used in decision support systems.

Keywords: decision support systems, design process, formalization, cyber attack, software defect, automated systems and complexes, cyber threat, cyber security

Постановка наукової проблеми. Останнім часом серйозна увага приділяється питанням проектування та створення систем підтримки прийняття рішень, що застосовуються в підрозділах для виявлення, попередження та усунення вторгнень в комп'ютерні системи та мережі (далі-СППР). На думку фахівців, на теперішній час відсутній формалізований математичний опис багатьох процесів, що відбуваються під час та застосування сучасних кібератак, способів їх усунення, призначення оптимальних маршрутів відновлення програмного забезпечення автоматизованих систем та комплексів. Все це заважає здійснювати достовірно процес проектування компонентів та програмних модулів СППР щодо підвищення рівня захисту програмного забезпечення.

На теперішній час існує проблема, коли значна вартість та трудомісткість проектування систем підтримки прийняття рішень затримують впровадження компонентів, що належать до їх складу. Це призводить до погіршення використання засобів обчислювальної техніки та малої ефективності СППР.

Усунення зазначеної диспропорції можливе тільки за рахунок удосконалення технології проектування, яка має декілька суттєвих принципових відмінностей, внаслідок того, що СППР розробляється для автоматизованих систем та комплексів, які отримують пошкодження або збій у роботі програмного забезпечення, після дії різних кібератак, появи дефектів, для яких необхідно

прийняти рішення щодо характеру їх впливу на працездатність програмного забезпечення автоматизованих систем та комплексів, відповідно до чого призначити оптимальні способи відновлення їх працездатності.

Виникає задача, яка дозволяє здійснювати розробку СППР з використанням вже раніше існуючих типових частин програмного забезпечення, що достатньо якісно зарекомендували себе в роботі під час прийняття різних складних рішень, з додатковим доопрацюванням окремих пакетів прикладних програм СППР та з урахуванням захисту інформації в наслідок впливу кібератак. Не виключається можливість здійснювати розробку проектів СППР щодо захисту інформації з повторним використанням відомих проектів вже існуючих СППР.

Таким чином, за рахунок використання вказаних підходів, час розробки СППР зменшується, але збільшуються помилки проектування внаслідок неможливості вдосконалення технології проектування СППР, призначених щодо захисту інформації, враховуючи особливості технологій виявлення кібератак, оцінювання дефектів та визначення способів відновлення пошкодженого програмного забезпечення.

Виникає унікальна задача, яка передбачає проведення робіт з формалізації процесу проектування СППР, шляхом аналізу проектної діяльності у галузі захисту інформації, визначення відомостей, необхідних для проектування, розробки алгоритмів, що функціонують в СППР та створення необхідних методичних матеріалів.

Аналіз досліджень. Проблеми створення та застосування методології проектування систем підтримки прийняття рішень, останнім часом активно розглядаються у роботах багатьох вітчизняних та закордонних фахівців. Це пов'язано з тим, що для застосування політики безпеки під час обробки інформації та забезпечення безперервності бізнесу, компаніям необхідно впроваджувати системи управління ризиками в сфері інформаційної безпеки [1, 2].

На думку фахівців протистояти постійному збільшенню та складності кібератакам можна, зокрема, за допомогою програмних компонентів, що здійснюють розпізнавання кібератак та забезпечені модулями системи підтримки прийняття рішень, які поєднують знання та досвід прийняття рішень з питань відновлення пошкодженого програмного забезпечення.

Аналізуючи роботи багатьох авторів, необхідно зазначити, що у статтях окремим напрямком представлені дослідження у галузі проектування та застосування СППР, а також описуються методи розробки їх моделей та програмного забезпечення [3], включаючи розробку експертних систем [4, 5].

Суттєвим недоліком окремих статей [6, 7] є відсутність архітектурної реалізації СППР, адекватної до реального процесу обробки інформації. Це пов'язане з тим, що процеси, які виникають під час впливу кіберзагроз важно формалізувати. Тому, за висновком авторів, більшість СППР та експертних систем значну частину часу знаходяться у стадії тестування.

Інтерес представляє робота авторів [8], де розглядаються недоліки існуючих СППР та експертних систем, що застосовуються у галузі інформаційної безпеки. До таких недоліків належать необхідність присутності висококваліфікованих фахівців при складанні бази знань, труднощі алгоритмізації складних процесів відновлення програмного забезпечення тощо.

У науковій роботі представлена система підтримки прийняття рішень, заснована на аналітичному ієрархічному процесі та методах змішаного цілісного програмування для оптимального вибору дій щодо забезпечення інформаційної безпеки підприємства.

Запропонований підхід дозволяє максимізувати величину ризику при фіксованій сумі бюджету за рахунок визначення оптимального набору запобіжних заходів. СППР також допомагає особам, які ухвалюють рішення на підприємстві, визначити мінімальний бюджет інформаційної безпеки підприємства для заданого рівня ризику.

Продовження зазначеного підходу щодо проектування СППР досліджується у наукових працях [9 - 12], де автори розглядають теоретичні та практичні питання з технології підтримки прийняття рішень та пропонують ряд цікавих методів для розробки СППР, які базуються на аналізі та обробці інформації, шляхом використання експертних оцінок, певної послідовності управлінських операцій та процедур.

В зазначених наукових роботах питання розробки та застосування СППР вирішуються на підставі індивідуального аналізу політики безпеки суб'єкта господарювання з використанням індивідуальних методів та засобів проектування програмних компонентів СППР. Автори вважають,

що в основу розробки СППР повинен покладений системний підхід та моделювання процесами виявлення дефектів програмного забезпечення та їх усунення за рахунок взаємозв'язку між атаками та організаційними та технологічними заходами політики інформаційної безпеки. На думку фахівців такий підхід суттєво забезпечити рівень інформаційної безпеки на підприємстві.

Аналізуючи матеріали декілька робіт з питань розробки СППР [13 - 15], необхідно зазначити, що автори у своїх наукових працях пропонують шляхи щодо розробки моделі підтримки прийняття рішень у галузі кібербезпеки, яка базується на аналізі ризиків, враховуючи тільки фінансові фактори загроз. Такий підхід може бути ефективним для підтримки економічно-орієнтованих рішень, але не враховує структуру кібербезпеки підприємства.

Інша група авторів [16] пропонує підходи щодо проектування системи підтримки прийняття рішень, враховуючи аналіз ризиків та структури кібербезпеки, відповідно до існуючих стандартів. На думку авторів така модель забезпечує належний захист інформації та практично без помилок визначає загрози та кібератаки. У роботі, також надаються результати рекомендацій щодо проектування СППР у вигляді статистичної оцінки захисту інформації автоматизованої системи, які отримані шляхом практичного проведення атак на її програмне забезпечення.

Таким чином, огляд та систематизація матеріалу розглянутих наукових праць вказує на необхідність продовження робіт щодо формалізації процесів проектування СППР щодо підвищення рівня захисту програмного забезпечення. Такі дослідження повинні бути направлені на розробку методів проектування складних нетипових задач, які виникають під час виявлення та дії кібератак, враховуючи вимоги та обмеження, що визначають порядок обробки інформації на підприємстві під час її експлуатації, а також супроводження та адміністрування програмного забезпечення автоматизованих систем та комплексів.

Мета роботи. Метою статті є дослідження можливості формалізації процесу проектування систем підтримки прийняття рішень, призначених для відновлення пошкодженого програмного забезпечення внаслідок впливу кібератак та надання пропозицій щодо розробки структурно-логічної схеми вказаних систем.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Проектування компонентів СППР щодо відновлення пошкодженого програмного забезпечення внаслідок впливу кібератак, представляє собою складну задачу. Складність проектування визначається тим, що сам процес розробки компонентів вказаної системи ще недостатньо формалізований.

У процесі проектування вирішуються наступні проектні задачі B_j із застосування та допомогою різних формалізованих методик проектування

$$A = \{A_1, A_2, A_3, \dots, A_{11}\},$$

(1)

- де:
- A_1 – дослідження інформаційного об'єкту ;
 - A_2 – аналіз потоків інформації щодо появи кібератаки;
 - A_2 - аналіз характеристик кібератаки;
 - A_3 – попередня оцінка рівня загрози програмного забезпечення;
 - A_4 – надання пропозицій щодо можливих дефектів програмного забезпечення, які виникають у наслідок дії кібератаки;
 - A_5 – аналіз інформації, отриманої за результатами дефектації пошкодженого програмного забезпечення у наслідок дії кібератаки;
 - A_6 – оцінка дефектів пошкодженого програмного забезпечення після дії кібератаки;
 - A_7 – вибір переліку та способів відновлення пошкодженого програмного забезпечення у наслідок дії кібератаки;
 - A_8 – надання пропозицій щодо вибору необхідного переліку програмного та технічного забезпечення для відновлення пошкодженого програмного забезпечення;
 - A_9 – остаточний вибір способу усунення дефектів програмного забезпечення,
 - A_{10} розробка послідовності технологічних операцій щодо відновлення пошкодженого програмного забезпечення;
 - A_{11} – оцінка економічної ефективності щодо відновлення пошкодженого програмного забезпечення;

З метою здійснення робіт з проектування, технологія розробки СППР, передбачає використання двох основних етапів, а саме, по перше розробка технологічного проекту СППР, а по друге виконання його техніко-економічне обґрунтування. На теперішній час основним критерієм проектування СППР є його економічна ефективність. У зв'язку з чим, на початку проектування розробляється частина проектних матеріалів E_p , необхідних для розрахунку проектного коефіцієнту ефективності капітальних затрат E_{kf} щодо створення СППР, після чого здійснюється перевірка наступної умови:

$$E_p \gg E_{kf} \quad (2)$$

де, E_{kf} коефіцієнт ефективності капітальних затрат, який визначається під час створення СППР.

Аналізуючи стан проектування СППР, необхідно зазначити, що процес проектування має безліч особливостей. Особливість проектування пов'язана перш за все з неоднозначними явищами, що виникають під час впливу кібератак, фактами появи дефектів, способів їх виявлення та усунення, а також визначення оптимальних маршрутів з метою забезпечення відновлення програмного забезпечення автоматизованих систем та комплексів.

Визначено, що для розробки проекту СППР щодо відновлення пошкодженого програмного забезпечення внаслідок впливу кібератак, головним є аналіз та використання інформації, яка приймає участь у проектуванні, яка складається з керуючої, вихідний та термінальної, а саме:

-керівна інформація J_y регламентує технологічний процес проектування на підставі вимог керівних та методичних матеріалів з проектування СППР;

-вихідна J_o складається з нормативної інформації та інформації щодо об'єкту проектування, який є об'єктом впливу кібератак;

-у свою чергу термінальна інформація J_t є результатом безпосереднього проектування і містить відомості щодо текстових, табличних, графічних матеріалах, які, як правило, зберігаються носіях інформації.

Враховуючи зміст термінальної інформації, представимо її компоненти у вигляді певних складових, об'єднаних між собою:

$$J_t = \{J_1 \cup J_2 \cup J_3\} \quad (3)$$

де: J_1 – складові інформації, які раніше використовувалися для розробки СППР та застосовується для проектування без змін;

J_2 – складові інформації, отримані за результатами об'єднання елементів вихідної інформації;

J_3 – складові інформації, отримані за результатами попередньої обробки елементів вихідної інформації.

Процес проектування характеризується у функціональному відношенні як процес якісних та кількісних змін інформації, а у відношенні її структури – як сукупність взаємопов'язаних операцій. На підставі цього функцію (F_p) технологічного проектування СППР можна формалізувати у вигляді:

$$F_p = J_y \times J_o \rightarrow J_t \quad (4)$$

де, термінальна інформація J_t є відображенням вихідної J_o та керівної інформації J_y .

Структурні компоненти проектування $\{CK\}$ можна представити у вигляді наступного рівняння $CK = \{B_i\}$, де B_i є певної задачею, для якої виконується проектні процедури, наприклад, пошук інформації, попередня обробка інформації, класифікація кібератаки, оцінювання стану загроз, підготовка варіантів рішень щодо усунення кібератак.

Таким чином, проектування включає сукупність проектних процедур, які виконуються для кожної задачі. Функції кожної процедури f_i передбачають перетворення інформації з одного проміжного стану в інший.

У свою чергу структура кожної проектної задачі B_i визначається за результатами

застосування операції об'єднання графів - функцій процедур, а саме:

$$B_i = f_1 \cup f_2 \cup f_3 \dots \cup f_n \quad (5)$$

Остаточно структуру технологічного процесу проектування СППР можна описати графом, який представляє об'єднання графів-функцій проектних задач:

$$G(J, Q) = B_1 \cup B_2 \cup B_3 \dots \dots \cup B_n \dots \quad (6)$$

Зазначимо окремі особливості, що характеризують процес проектування СППР:

$$Q = \{Q_i\}, \quad i = \overline{1, k}, \quad (7)$$

- де: Q_1 – замкнутість процесу проектування;
 Q_2 – можливість удосконалення технології проектування;
 Q_3 – старіння результатів проектування.

Особистістю проектної діяльності є старіння проектних матеріалів. Істинність проекту під час розробки U_0 та під час реалізації U_p знаходиться у залежності

$$U_0 > U_p \quad (8)$$

Ця причина пояснюється тем, що об'єктивна реальність працездатності автоматизованих систем та комплексів постійно змінюється в наслідок імовірності появи та остаточного виявлення кібератак, дефектів програмного та апаратного забезпечення, способів їх відновлення, в наслідок чого змінюється адекватність програмного забезпечення для компонентів СППР, які раніше були розроблені.

Тобто можна зазначити, що інтервал часу T_i , протягом якого може бути отриманий економічний ефект від реалізації проекту, є величина, що функціонально залежить від багатьох факторів

$$T_i = f\{PR_1, PR_2 \dots \dots PR_n\} \quad (9)$$

- де: PR_1 – проектні рішення, що використовують нові технології та їх економічність;
 PR_2 – проектне рішення, яке враховує зовнішні впливи;
 PR_3 – інші проектні рішення.

Вибір рішення залежить від різних варіантів проектів $V_B = \{Vk\}$, $k = \overline{1, n}$, та передбачає вибір такого варіанту V_0 , де $V_0 \in V_B$, який відповідає завданню на проектування.

Загальний об'єм проектування суттєво зменшується, якщо вибір здійснюється не на множині варіантів проектів V_B , а множині варіантів проекту, які задовольняють умови

$$V_y = V_B \setminus V_n \quad (10)$$

де: V_n - множина незадовільних варіантів проекту.

За рахунок такого підходу, значно зменшується кількість варіантів проектів V_B , які потребують подальшої переробки, а сам процес проектування представляється у вигляді ієрархічної структури рівнів розробки, які характеризуються ступеню деталізації елементів.

На певному рівні здійснюється пошук умовного варіанту V_{yi} з деталізацією відомостей. На наступному рівні обрані варіанти переробляються більш детально на підставі певних критеріїв. Таким чином, у процесі проектування використовуються два напрямки – генерування різноманіття варіантів V_{bi} , їх аналіз та обмеження.

Послідовність проектування необхідно розподілити на дві паралельні задачі. Перша – це розробка системи збору інформації, її передавання та первинної обробки. Друга – розробка системи обробки даних. За часом вказані задачі можуть виконуватися паралельно або послідовно, в залежності від кількості осіб, які здійснюють розробку проекту.

Проектування передбачає ітераційний підхід щодо вирішення задач. Це пов'язано з тим, що на етапах розробки результати проектування необхідно порівнювати з нормативними показниками. Якщо терміни розробки не відповідають нормативним показникам, то здійснюється корегування

завдання і процес проектування продовжується. Необхідно зазначити що напрямком проектування є формалізація уніфікованих процедур розробки СППР, які відповідають сучасним технологія створення проекту. Послідовність проектування передбачає використання окремих процедур, пов'язаних між собою певними правилами.

Як було раніше зазначено, головними проектними процедурами під час проектування буде пошук інформації або її генерування, з подальшим вибором найкращого варіанту. У разі відсутності інформації, що необхідна для проектування, здійснюється корегування завдання на проектування. Якщо завдання неможливо скорегувати, то процес проектування або закінчується, або корегується для продовження проектування. Окрім цього виконується етап перевірки необхідності обробки інформації, яка аналізувалась раніше, якщо це необхідно, то для проектування застосовується процедура попередньої обробки інформації.

Важливим елементом є те, що процедури пошуку інформації та прийняття рішення виконуються на всіх рівнях проектування. Після чого застосовуються процедури, які здійснюють оцінку відповідності проектного рішення проектною задачі. У разі, якщо проектне рішення відповідає нормативними показникам, то оформлюється текстова та графічна документація проекту, якщо ні, то перевіряється можливість коректування завдання.

На проектування, впровадження та експлуатацію СППР впливають багато обмежень:

$$M = \{M_1, M_2, M_3, M_4\} \quad (11)$$

де: M_1 – грошові ресурси на придбання технічного обладнання та математичного забезпечення; M_2 – трудові ресурси проектувальників; M_3 – трудові ресурси кваліфікованих спеціалістів з експлуатації СППР; M_4 – капітальні вкладення тощо. Обмеження M також впливають на кількість підсистем та задач, відповідно і на множину алгоритмів. Тому склад першочергових задач СППР залежить від суб'єктивних факторів та не є оптимальним з економічної точки зору.

Дослідження розрахованих економічних показників ефективності СППР показують, що ступень впливу однакових факторів для різних проектів СППР суттєво відрізняється. Тому для кожної СППР включають тільки ті задачі, які забезпечують вплив факторів, що дають економію.

Висновки. Таким чином, для ефективності проектування СППР щодо підвищення рівня захисту програмного забезпечення, необхідно мати певний математичний апарат, за допомогою якого можна розробити оптимальну технологію проектування зазначених систем. Такий математичний апарат, повинен враховувати особливості процесів виявлення, оцінювання кібератак, а також обмежень на проектування підсистем та множину алгоритмів.

Розробка СППР щодо підвищення рівня захисту програмного забезпечення можлива тільки за рахунок удосконалення технології проектування, пов'язаної з формалізацією процесів, що приймають участь у технологічних операціях виявлення кібератак, оцінювання дефектів пошкодженого програмного забезпечення та призначення оптимального способу їх відновлення з метою забезпечення надійної працездатності автоматизованих систем та комплексів, для яких застосовуються зазначені системи.

Список бібліографічного опису

1. Delgado, F., Esenarro, D., Jubrez, R., Rebtegui, M. (2021). Methodology based on the NIST cybersecurity framework as a proposal for cybersecurity management in government organizations. *Cuadernos dedesarrollo aplicados a las TIC*, 10(2), 123–141.
2. Kure, H.I., Islam, S., Razzaque, MA. (2018). An integrated cyber security risk management approach for a cyber-physical system. *Appl. Sci. Switzerland*, 8(6), 1–29.
3. Fielder, A. (2016) Decision support approaches for cyber security investment. *Decision Support Systems*, 86, 13–23.
4. Atymtayeva, L. (2014) Building a Knowledge Base for Expert System in Information Security. *Chapter Soft Computing in Artificial Intelligence of the series Advances in Intelligent Systems and Computing*, 270, 57–76.
5. Gamal, M.M. (2011) Security Analysis Framework Powered by an Expert System. *International Journal of Computer Science and Security (IJCSS)*, 4(6), 505–527.
6. Lee, K., C. Wei, J., Mao, C.-H., Dai, J.-H., Kuang, Y.-T. (2016). Cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation. *Soft Computing*, 1–14.
7. Pan, S., Morris, T., Adhikari, U. (2015). Developing a Hybrid Intrusion Detection System Using Data Mining or Power. *Systems*, 6(6), 3104–3113.
8. Ferda, Ö., Banu, G. (2021). A Decision Support System for Optimal Selection of Enterprise Information Security Preventative Actions. *IEEE Transactions on Network and Service Management*, 18(3), 3260-3279

9. Бідюк, П.І., Кожухівський, А.Д., Кожухівська, О.А. (2013). Система підтримки прийняття рішень для аналізу і прогнозування стану підприємства. *Радіоелектроніка, інформатика, управління*, 1, 128-136.
10. Волошин, О.Ф. Машченко, С.О. (2010). *Моделі та методи прийняття рішень: навч. посіб.* Київ: Видавничополіграфічний центр "Київський університет".
11. Бурячок, В.Л., Толюпа, С.В., Аносов, А.О., Козачок, В.А., Лукова-Чуйко, Н.В. (2015). *Системний аналіз та прийняття рішень в інформаційній безпеці: підручник*. Київ: ДУТ.
12. Азарова, А.О., Дьогтева, І.О., Шиян, А.А. (2022). Система підтримки прийняття рішень щодо підвищення рівня інформаційної безпеки підприємства. *Інформаційні технології та комп'ютерна інженерія*, 1, 12-18.
13. Diesch, R., Pfaff, M., Krcmar, H. (2020). *A comprehensive model of information security factors for decision-makers. Comput. Secur.* Retrieved from <https://www.sciencedirect.com/science/article/pii/S0167404820300341>.
14. Roldán-Molina G., Almache-Cueva M., Silva-Rabadão C., Yevseyeva I., Basto-Fernandes V. (2017). A Decision Support System for Corporations Cybersecurity Management. In Proceedings of IEEE 12th Iberian Conference on Information Systems and Technologies, CISTI'2017, 21-24 June 2017, Lisbon, Portugal. IEEE.
15. Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., Smeraldi, F. (2016). *Decision support approaches for cyber security investment. Decis. Support.* Retrieved from: <https://www.sciencedirect.com/science/article/pii/S0167923616300239>.
16. Khairur, R., Benfano, S. (2022). Cybersecurity decision support model to designing information technology security system based on risk analysis and cybersecurity framework. *Indonesia gyptian Informatics Journal*, 23, 383–404.

References

1. Delgado, F., Esenarro, D., Juárez, R., Rebategui, M. (2021). Methodology based on the NIST cybersecurity framework as a proposal for cybersecurity management in government organizations. *Cuadernos de desarrollo aplicados a las TIC*, 10(2), 123–141.
2. Kure, H.I., Islam, S., Razzaque, M.A. (2018). An integrated cyber security risk management approach for a cyber-physical system. *Appl. Sci. Switzerland*, 8(6), 1–29.
3. Fielder, A. (2016) Decision support approaches for cyber security investment. *Decision Support Systems*, 86, 13–23.
4. Atymtayeva, L. (2014) Building a Knowledge Base for Expert System in Information Security. *Chapter Soft Computing in Artificial Intelligence of the series Advances in Intelligent Systems and Computing*, 270, 57–76.
5. Gamal, M.M. (2011) Security Analysis Framework Powered by an Expert System. *International Journal of Computer Science and Security (IJCSS)*, 4(6), 505–527.
6. Lee, K., C. Wei, J., Mao, C.-H., Dai, J.-H., Kuang, Y.-T. (2016). Cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation. *Soft Computing*, 1–14.
7. Pan, S., Morris, T., Adhikari, U. (2015). Developing a Hybrid Intrusion Detection System Using Data Mining or Power. *Systems*, 6(6), 3104–3113.
8. Ferda, Ö., Banu, G. (2021). A Decision Support System for Optimal Selection of Enterprise Information Security Preventative Actions. *IEEE Transactions on Network and Service Management*, 18(3), 3260-3279
9. Bidyuk, P.I., Kozhukhivskiy, A.D., Kozhukhivska, O.A. (2013). Decision support system for analyzing and forecasting the state of the enterprise. *Radio electronics, informatics, management*, 1, 128-136.
10. Voloshyn, O.F. Mashchenko, S.O. (2010). *Decision-making models and methods: teaching. manual* Kyiv: Kyiv University Publishing and Printing Center.
11. Buryachok, V.L., Tolyupa, S.V., Anosov, A.O., Kozachok, V.A., Lukova-Chuiko, N.V. (2015). *System analysis and decision-making in information security: a textbook*. Kyiv: DUT.
12. Azarova, A.O., Dyogteva, I.O., Shiyan, A.A. (2022). Decision support system for increasing the level of information security of the enterprise. *Information technologies and computer engineering*, 1, 12-18.
13. Diesch, R., Pfaff, M., Krcmar, H. (2020). *A comprehensive model of information security factors for decision-makers. Comput. Secur.* Retrieved from <https://www.sciencedirect.com/science/article/pii/S0167404820300341>.
14. Roldán-Molina G., Almache-Cueva M., Silva-Rabadão C., Yevseyeva I., Basto-Fernandes V. (2017). A Decision Support System for Corporations Cybersecurity Management. In Proceedings of IEEE 12th Iberian Conference on Information Systems and Technologies, CISTI'2017, 21-24 June 2017, Lisbon, Portugal. IEEE.
15. Fielder, A., Panaousis, E., Malacaria, P., Hankin, C., Smeraldi, F. (2016). *Decision support approaches for cyber security investment. Decis. Support.* Retrieved from: <https://www.sciencedirect.com/science/article/pii/S0167923616300239>.
16. Khairur, R., Benfano, S. (2022). Cybersecurity decision support model to designing information technology security system based on risk analysis and cybersecurity framework. *Indonesia gyptian Informatics Journal*, 23, 383–404.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-12>

УДК 004.056:004.725.5

Журавська Ірина Миколаївна, д.т.н., професор

<https://orcid.org/0000-0002-8102-9854>

Фісун Микола Тихонович, д.т.н., професор

<https://orcid.org/0000-0003-1297-6230>

Чорноморський національний університет імені Петра Могили, м. Миколаїв, Україна

РИЗИКИ ІНФОРМАЦІЙНИХ ВИТОКІВ З МОБІЛЬНИХ ПРИСТРОЇВ

Журавська І. М., Фісун М.Т. Ризики інформаційних витоків з мобільних пристроїв. У статті розглянуто загрози безпеці мобільних застосунків, уразливості мобільних операційних систем (ОС), канали витоку інформації, семантичну безпеку, безпеку, пов'язану з інфраструктурою, тощо. Зазначено, що порушення безпеки можна пом'якшити за допомогою покращення поведінки та ставлення до безпеки, а не лише за допомогою кращих технологій. Показано, що програми під керування як ОС Android, так і iOS за замовчуванням можуть отримувати доступ до датчиків, таких як акселерометр і гіроскоп, мікрофон і камера під час роботи у фоновому режимі без відома або згоди користувача. Проаналізовано ризики витоку інформації внаслідок надання прав системного доступу застосункам з root-правами, оновлення системи прошивками, що містять шкідливі програми або вразливості, які можуть бути використані зловмисниками. Проаналізовано методи перехоплення керування мобільними пристроями з використанням засобів пакету Smurf Suite, технологій фішингу та ін. Запропоновано практичні рекомендації щодо зменшення ризиків безпеки при використанні мобільних застосунків як для користувачів, так і для розробників. Наведено переваги використання, Apple Neural Engine і користувацьких моделей CreateML в мобільних пристроях з ОС iOS для обробки даних прямо на пристрої, щоб уникнути тривалих і потенційно ризикованих звертань до віддаленого сервера.

Ключові слова: загрози безпеці, виток інформації, мобільні пристрої, датчики, IoT-пристрої, конфіденційність, iOS, Android.

Zhuravska I., Fisun M. Risks of information leaks from mobile devices. The article covers mobile application security threats, mobile operating system (OS) vulnerabilities, information leakage channels, semantic security, infrastructure-related security, and more. It noted that security breaches can be mitigated through improved security behaviors and attitudes, not just better technology. By default, both Android and iOS apps have been shown to be able to access sensors such as accelerometer and gyroscope, microphone and camera while running in the background without the user's knowledge or consent. The risks of information leakage as a result of granting system access rights to applications with root rights, updating the system with firmware containing malicious programs or vulnerabilities that can be used by attackers have been analyzed. The methods of intercepting control of mobile devices using Smurf Suite tools, phishing technologies, etc. were analyzed. Practical recommendations for reducing security risks when using mobile applications are offered for both users and developers. Benefits of using Apple Neural Engine and custom CreateML models on iOS mobile devices to process data directly on the device to avoid lengthy and potentially risky calls to a remote server.

Keywords: security threats, leak of information, mobile devices, sensors, IoT devices, confidentiality, iOS, Android.

Постановка наукової проблеми. В сучасних умовах діджиталізації суспільства 78,3 % користувачів використовують свій телефон для банківських операцій, доступу до персональних даних (медичних та ін.), управління інвестиціями, торгівлі криптовалютами – по суті, для керування важливими даними та/або проведення конфіденційних транзакцій [1]. Але 44,5 % не використовують спеціальних програм або рішень для захисту власних мобільних пристроїв. В той же час, з поширенням інтелектуальних пристроїв Інтернету речей (IoT) на основі мобільних операційних систем (ОС), зростає кількість та функціонал шкідливих програм, націлених на пристрої IoT,

Зараз на ринку домінують дві ОС, а саме Android OS (74,69 % ринку) та iOS (22,34 % ринку). Відповідно, існує два окремих середовища розробки, включаючи мови програмування, SDK, API та різноманітні засоби створення компонентів, що відповідають двом ОС відповідно. Незважаючи на те, що розробники докладають величезних зусиль, щоб запобігти загрозам безпеці, реальність така, що жодне програмне забезпечення (ПЗ), у т. ч. ОС, не застраховано від злому, особливо під час цілеспрямованих атак [2]. Далі представлено більш детальне обговорення цих двох середовищ. Нещодавні дослідження показують, що порушення безпеки можна пом'якшити за допомогою покращення поведінки та ставлення до безпеки, а не лише за допомогою кращих технологій [3].

Аналіз досліджень. Проблеми з конфіденційністю та безпекою мобільних пристроїв активно вивчаються дослідниками [2; 3]. Але обмежена обчислювальна потужність та інтерфейс користувача мобільних пристроїв полегшують зловмисникам приховування своїх зловмисних дій. Основні теми, що обговорюються, включають загрози безпеці застосунків, уразливості мобільних




операційних систем (ОС), канали витоку інформації, семантичну безпеку, безпеку, пов'язану з інфраструктурою, тощо.

Захист мобільних застосунків від атак по побічним каналам можливий за допомогою вето на доступ до певних ресурсів [4], використання масштабованих систем виявлення зловмисного ПЗ мобільних ОС на основі аналізу шаблонів поведінки [5] та ін.

Сучасні смартфони зазвичай оснащені різними апаратними датчиками (наприклад, мікрофоном, GPS, датчиком освітлення, акселерометром тощо) для взаємодії з фізичним світом. Смартфони також мають доступ до особистої інформації користувача, такої як списки контактів, фотографії та паролі. Доступ сторонніх програм до цих датчиків/ресурсів та інформації користувача контролюється ОС за замовчуванням або під час встановлення програми, або під час її роботи.

Позитивним є те, що останні версії ОС забезпечують певний обмежений захист від цих атак із побічних каналів. Починаючи з Android 9.0, програми за замовчуванням більше не можуть отримувати доступ до датчиків, таких як акселерометр і гіроскоп, мікрофон і камера, під час роботи у фоновому режимі без запуску основної служби, видимої для користувачів як значок [6]. API AudioPlaybackCapture в Android 10, який дозволяє програмі записувати аудіо з іншої програми [7], також пропонує методи відмови, за допомогою яких програма може запобігти іншим програмам доступ до її аудіо. Як інший захист, Android надає доступ до камери та мікрофона лише одній програмі в кожен момент часу з певним обґрунтуванням включення, нп., мікрофона (табл. 1).

Таблиця 1. Обґрунтування включення мікрофона застосунком

Позитивне/ Негативне висловлювання	Приклад анонсу функції (призначення)	Обґрунтування
	Програма записує вночі, щоб виявити звуки хрюпіння.	Активне речення, яке чітко описує, як і чому програма збирає дані.
	Для кращої роботи потрібен доступ до мікрофона.	Пасивне речення, яке містить розпливчате, невизначене обґрунтування.
	Увімкніть доступ до мікрофона.	Наказове речення, яке не дає жодного обґрунтування.

Однак, обмеження, заявлені застосунками у діалогових вікнах на використання камери/мікрофона, не заважає зловмисній програмі використовувати ці ресурси/датчики для атаки побічного каналу, якщо програмі-жертві не потрібен доступ до цих ресурсів.

Виділення раніше невирішених питань. Нещодавні зміни в контролі доступу в Android та iOS спрямовані на забезпечення більшого контролю над ресурсами: дозвіл на фонове використання для місцезнаходження в Android 10 [8], iOS 13 [9] і одноразовий дозвіл на місцезнаходження, камеру та мікрофон в Android 11 [10]. Але навіть із цими вдосконаленнями користувачі все ще можуть дозволити програмі завжди отримувати доступ до цих ресурсів (особливо для програм, які довго працюють і активно використовуються); тобто до видалення такої програми можуть використовувати надані дозволи.

Однак ресурси, які вважаються незначними або взагалі не становлять ризику для безпеки/конфіденційності, як-от акселерометр або гіроскоп, можуть використовуватися будь-якою програмою без відома або згоди користувача. Було продемонстровано багато атак з використанням цих так званих небезпечних або звичайних ресурсів [11; 12], а також ресурсів, які вимагають явної згоди користувача [13; 14]. Окрім компрометації PIN-кодів і паролів, атака [13] показує, що, здавалося б, безпечний акселерометр також можна використовувати для прослуховування динаміка телефону. Поточні моделі дозволів на основі схвалення користувача не можуть протистояти цим прихованим, але високоефективним атакам.

Тому потрібні додаткові дослідження для зменшення ризиків безпеки як для користувачів, так і для розробників.

Формулювання мети і завдань дослідження. Метою статті є дослідження режимів роботи мобільних пристроїв для виконання ними необхідних функцій, не допускаючи витоку особистої інформації про пристрій або його користувача.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Конфіденційна інформація буде викрадена різними шляхами. Відмітимо основні з них:

- витоки інформації засобами операційної системи Андроїд;
- витоки інформації засобами фізичного доступу до телефону;
- витоки інформації засобами програмного забезпечення;
- витоки інформації шляхом фішингу та інших схожих методів.

Засобами ОС, здебільшого, дані можуть бути вкрадені, якщо не притримуватися оновлень системи, завантажувати і надавати права системного доступу застосункам з root-правами та перепрошивати систему, використовуючи чиїсь готові продукти (рис. 1).

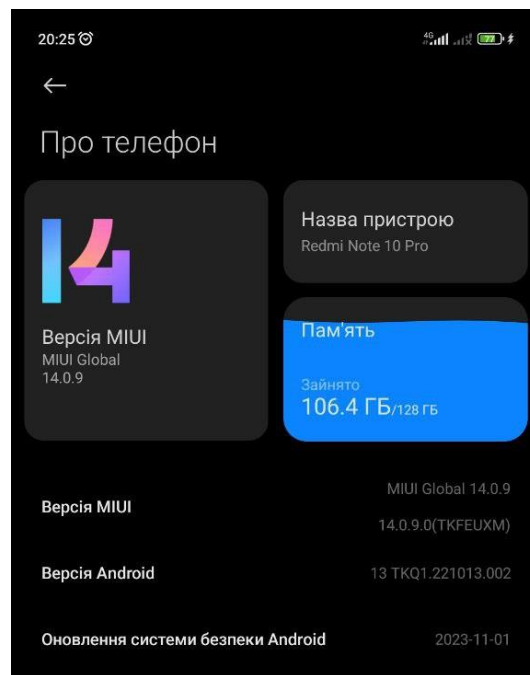


Рис. 1 – Відомості про телефон Xiaomi (версія телефону, ОС та вбудованого інтерфейсу)

Неактуальне ПЗ може бути однією з головних причин витоку інформації через операційну систему Android. Оновлення системи регулярно вирішує виявлені вразливості і запобігає можливим атакам.

Надання застосункам root-прав також може стати причиною витоку інформації, оскільки це дає їм розширений доступ до системи і даних користувача. Якщо застосунки з root-правами не контролюються або вони не мають достатнього рівня довіри, це може стати проблемою з безпекою.

Перепрошивка системи з використанням ненадійних або модифікованих прошивок також може призвести до витоку інформації, оскільки ці прошивки можуть містити шкідливі програми або вразливості, які можуть бути використані зловмисниками.

Витоки інформації засобами фізичного доступу до телефону можуть включати наступне:

Зловмисник, який має фізичний доступ до телефону, може отримати доступ до всієї збереженої інформації на пристрої, такої як фотографії, відео, повідомлення, контакти, електронні листи тощо. Зловмисник може встановити шпигунське ПЗ на телефон без відома користувача. Це ПЗ може стежити за активністю користувача, записувати клавіші, переглядати повідомлення та іншу конфіденційну інформацію.

Якщо телефон не захищений паролем або використовує прості паролі, зловмисник може отримати доступ до телефону і всіх збережених в ньому облікових записів, таких як соціальні мережі, електронна пошта, банківські акаунти тощо. Зловмисник може скопіювати SIM-карту і отримати доступ до всіх дзвінків і повідомлень, які надходять на номер телефону.

Якщо зловмисник отримав фізичний доступ до вашого телефону, він може спробувати використати методи атаки, такі як зламання пароля, використання експлойтів для отримання

доступу до системи або навіть фізичне пошкодження пристрою для отримання доступу до даних. Крім того, можуть бути використані методи перехоплення на кшталт «комплекту смурфів» (Smurf Suite) для Вкл./Викл. смартфона без відома Власника (Dreamy), відстеження Власника смартфона з більшою точністю за допомогою механізму геолокації (Tracker), управління мікрофоном смартфона для підслуховування того, що відбувається навколо смартфона (Curious) [15].

Для запобігання витокам інформації засобами фізичного доступу до телефону важливо встановлювати надійний пароль чи використовувати інші методи аутентифікації, такі як відбитки пальців або розпізнавання обличчя. Також рекомендується ніколи не залишати телефон без нагляду та використовувати функції шифрування даних, якщо вони доступні на пристрої.

Ні для кого не секрет, що компанії-велетні як Google чи Amazon збирають інформацію про користувачів своїх послуг, але роблять це «конфіденційно» для користувача, запевняючи того в цілісності даних, доступ до яких він сам надає [16]. Застосунки, такі як Instagram, TikTok та YouTube, пропонують рекомендації до відео на основі вподобань користувача, в незалежності від того, чи шукав він певну інформацію в застосунку, чи просто розмовляв з кимось, тримаючи телефон поряд.

Це відбувається через дозволи, які користувач сам надає компаніям, встановлюючи застосунки. При завантаженні на пристрій будь-якого застосунку він спитає у користувача надати йому доступ до певних функцій телефону користувача, тим самим відкриваючи можливість для застосунку на крадіжку інформації. Проте, з іншого боку, певні застосунки просто відмовляються виконувати свої функції без надання їм доступу. Саме через це треба уважно слідкувати, що і якому застосунку дозволяти зчитувати з свого телефону, але відмовлятися від усього теж не треба, адже на 100 % це не захистить користувача, а залишити без функціоналу застосунку може. В такому разі необхідно знайти пункт в налаштуваннях телефону «Захист конфіденційності», переглянути налаштування усіх пунктів і розібратися, які застосунки використовують певні доступи, котрі їм можна було б обмежити задля меншого витоку (рис. 2).

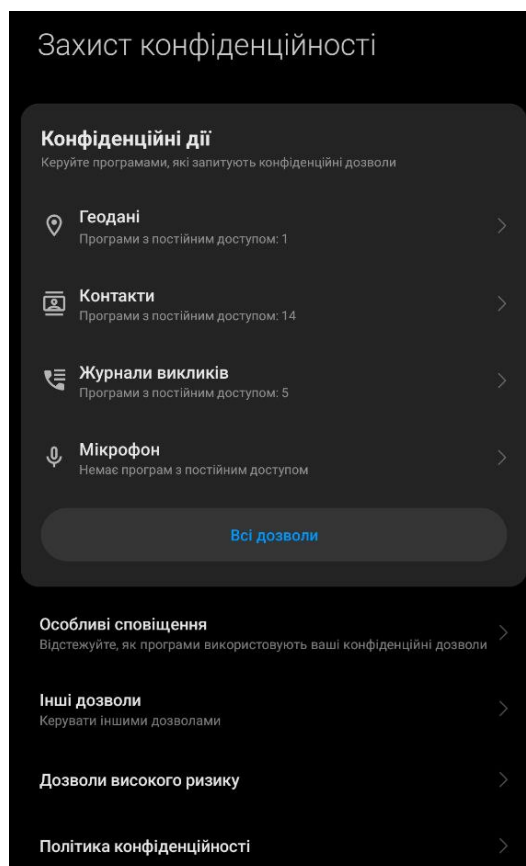


Рис. 2 – Налаштування «Захисту конфіденційності»

Як можна побачити на рис. 3, Instagram, Telegram, Viber та ще кілька застосунків мають доступ до списку контактів, а отже, користуючись цими застосунками, можливо бачити знайомих з

книги контактів одразу в них. Але тим самим власники цих застосунків вже можуть мати базу контактів власника у себе в офісі, а можливо навіть і в офісі у злодіїв, котрі скористалися нагодою вкрасти дані на етапі передачі з пристрою на сервер.

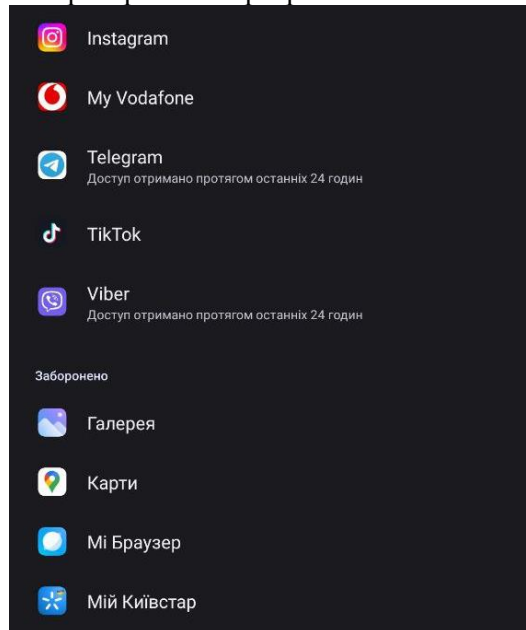


Рис. 3. Доступ до розділу «Контакти» на телефоні

Ну і останнє, але не найменше, – це витоки інформації шляхом фішингу. Все більше і більше людей стає популярними в наш час, і вони хочуть заохотити свою аудиторію роблячи розіграші і тому подібні конкурси. Тут на арену виступають «рибаки», котрі під прикриваючись вашим улюбленим блогером виманюють з вас данні карток, паролі і т. д. Ця схема працює не тільки в даному кейсі, а і під прикриттям банків, колекторів, розробників й т. ін.

Занотуймо основні способи:

Зловмисники можуть відправляти електронні листи, що виглядають як повідомлення від відомих компаній чи сервісів. У таких повідомленнях можуть міститися посилання на фальшиві веб-сайти, які виглядають як офіційні, але насправді призначені для крадіжки особистої інформації, такої як паролі або дані банківських карток.

Зловмисники також можуть використовувати соціальні мережі для поширення фішингових повідомлень, спробуючи отримати особисту інформацію від користувачів через маніпуляцію або використання підступних методів.

Зловмисники можуть відправляти фішингові SMS з повідомленнями про виграші, акції або інші привабливі пропозиції. Ці повідомлення можуть містити посилання на фальшиві веб-сайти або номери телефонів, що намагаються витягнути особисту інформацію від отримувачів.

Деякі шахраї можуть створювати фішингові застосунки або ігри, що виглядають як популярні або цікаві програми, але насправді призначені для збору особистої інформації з пристрою користувача.

Шахраї також можуть використовувати соціальну інженерію, щоб отримати особисту інформацію користувача, наприклад, видаючи себе під представників банків, компаній або інших організацій та запитуючи конфіденційні дані.

Висновки та перспективи подальшого дослідження. В результаті роботи було оглянуто ризики витоку інформації через мобільні платформи, представлено декілька прикладів різних видів витоку, розглянуто, чому і які дії можуть стати ключовими у втраті особистих даних через телефон, та розглянуто методи протидії певним з цих методів. Запропоновано практичні рекомендації щодо зменшення ризиків безпеки при використанні мобільних застосунків як для користувачів, так і для розробників, а саме:

а) обробляти дані на пристрої, де це можливо. В iOS, наприклад, можливо скористатися перевагами Apple Neural Engine і користувацьких моделей CreateML для обробки даних прямо на пристрої, щоб уникнути тривалих і потенційно ризикованих звертань до віддаленого сервера;

б) використовувати визначені системою засоби захисту конфіденційності та дотримуватися передових методів безпеки. Наприклад, в iOS 15 і новіших версіях можна покластися на CloudKit, щоб забезпечити шифрування та керування ключами для додаткових типів даних, як-от рядків, чисел і дат;

в) використовувати резервне копіювання важливих даних як на самому телефоні або інших зовнішніх носіях (нп., за допомогою iTunes для iPhone або Samsung Smart Switch для Samsung Galaxy), так і в хмарному сховищі, як-от Google Drive, iCloud, OneDrive і Dropbox;

г) запитувати дозвіл до:

– особистих даних, включаючи інформацію про місцезнаходження, стан здоров'я, фінанси, контакти та іншу особисту інформацію;

– контенту, створеного користувачами, як-от електронні листи, повідомлення, дані календаря, контакти, інформація про ігри, діяльність Apple Music, дані HomeKit, аудіо-, відео- та фотовміст;

– захищених ресурсів, як-от периферійні пристрої Bluetooth, функції домашньої автоматизації (IoT- пристрої та компоненти «розумного будинку»), з'єднання Wi-Fi та локальні мережі;

– такі можливостей пристрою, як камера, мікрофон тощо;

д) запитувати лише ті дані, які дійсно потрібні для роботи певного застосунку. Якщо запитувати більше даних, ніж потребує функція, користувачам буде важко довіряти такому застосунку.

Слід підкреслити, що будь-яке програмне забезпечення для мобільних платформ повинно мати документацію або навіть коротку анотацію, якій чітко описано, як програма використовує можливості, дані або ресурси, які запитує. Крім того, підтвердженням, що програма безпечна для використання, може бути підписання програми дійсним ідентифікатором розробника.

Список бібліографічного опису

1. Truță F. Top reasons why consumers shun mobile security apps. Publ. April 16, 2024. URL: <https://www.bitdefender.co.uk/blog/hotforsecurity/top-reasons-why-consumers-shun-mobile-security-apps/> (Last accessed: 03.05.2024).
2. Kadir A. F. A., Lashkari A. H., Firoozjaei M. D. Mobile application security. *In book: Understanding Cybersecurity on Smartphones*. Jan. 2024. P. 89–101. DOI: 10.1007/978-3-031-48865-8_6.
3. Balapour A., Nikkhah H. R., Sabherwal R. Mobile application security: Role of perceived privacy as the predictor of security perceptions. *International Journal of Information Management*. 2020. Vol. 52, No. 102063. DOI: 10.1016/j.ijinfomgt.2019.102063.
4. Osman T., Mannan M., Youssef A., Hengartner U., Youssef A. Securing applications against side-channel attacks through resource access veto. *Digital Threats: Research and Practice*. Dec. 2020. Vol. 1, No. 4. Article 22. 29 p. DOI: 10.1145/3416124.
5. Lei T., Qin Z., Zhibo W., Li Q., Dengpan Ye. EveDroid: Event-aware Android malware detection against model degrading for IoT devices. *IEEE Internet of Things Journal*. Aug. 2019. Vol. 6(4). P. 6668–6680. DOI: 10.1109/IIOT.2019.2909745.
6. Behavior changes: all apps. Android 9 (API level 28). Publ. Mar. 01, 2024. URL: <https://developer.android.com/about/versions/pie/android-9.0-changes-all> (Last accessed: 03.05.2024).
7. Turner D. Capturing Audio in Android Q. Publ. July 03, 2019. URL: <https://android-developers.googleblog.com/2019/07/capturing-audio-in-android-q.html> (Last accessed: 03.05.2024).
8. Privacy changes in Android 10. Publ. Apr. 30, 2024. URL: <https://developer.android.com/about/versions/10/privacy/changes#app-access-device-location> (Last accessed: 03.05.2024).
9. About privacy and Location Services in iOS, iPadOS, and watchOS. Publ. Mar. 27, 2024. URL: <https://support.apple.com/en-us/102515> (Last accessed: 03.05.2024).
10. Privacy in Android 11. URL: <https://developer.android.com/about/versions/11/privacy> (Last accessed: 03.05.2024).
11. Spreitzer R., Moonsamy V., Korak T., Mangard S. Systematic classification of side-channel attacks: A case study for mobile devices. *IEEE Commun. Surveys Tutor*. 2018. Vol. 20, Is. 1. P. 465–488.
12. Xu F., Diao W., Li Z., Chen J., Zhang K.. BadBluetooth: Breaking Android security mechanisms via malicious Bluetooth peripherals. *Proc. of the Network and Distributed System Security Symposium (NDSS'19)*, San Diego, CA, USA, 24–27 Feb. 2019. DOI: 10.14722/ndss.2019.23482.
13. Ba Z., Zheng T., Zhang X., Qin Z., Li B., Liu X., Ren K. Learning-based practical smartphone eavesdropping with built-in accelerometer. *Proc. of the Network and Distributed System Security Symposium (NDSS'20)*, San Diego, CA, USA, 23–26 Feb. 2020. DOI: 10.14722/ndss.2020.240762020.
14. Simon L., Anderson R. PIN Skimmer: Inferring PINs through the camera and microphone. *Proc. of the 3rd ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (SPSM'13)*. 08 Nov. 2013. P. 67–78. DOI: 10.1145/2516760.2516770.

15. UK spies can hack smartphones: Snowden. Publ. Oct. 05, 2015. URL: <https://phys.org/news/2015-10-uk-spies-hack-smartphones-snowden.html> (Last accessed: 03.05.2024).
 16. Відсота Д. Як глобальні медіа та IT-платформи збирають персональні дані користувачів. Опубл. 16 лютого, 2021. URL: <https://biz.nv.ua/ukr/experts/osobisti-dani-ta-socmerezhi-facebook-mesendzheri-ostanni-novini-50142213.html> (дата звернення: 03.05.2024).

References

1. Truță F. Top reasons why consumers shun mobile security apps. Publ. April 16, 2024. URL: <https://www.bitdefender.co.uk/blog/hotforsecurity/top-reasons-why-consumers-shun-mobile-security-apps/> (Last accessed: 03.05.2024).
2. Kadir A. F. A., Lashkari A. H., Firozjaei M. D. Mobile application security. In book: *Understanding Cybersecurity on Smartphones*. Jan. 2024. P. 89–101. DOI: 10.1007/978-3-031-48865-8_6.
3. Balapour A., Nikkha H. R., Sabherwal R. Mobile application security: Role of perceived privacy as the predictor of security perceptions. *International Journal of Information Management*. 2020. Vol. 52, No. 102063. DOI: 10.1016/j.ijinfomgt.2019.102063.
4. Osman T., Mannan M., Youssef A., Hengartner U., Youssef A. Securing applications against side-channel attacks through resource access veto. *Digital Threats: Research and Practice*. Dec. 2020. Vol. 1, No. 4. Article 22. 29 p. DOI: 10.1145/3416124.
5. Lei T., Qin Z., Zhibo W., Li Q., Dengpan Ye. EveDroid: Event-aware Android malware detection against model degrading for IoT devices. *IEEE Internet of Things Journal*. Aug. 2019. Vol. 6(4). P. 6668–6680. DOI: 10.1109/JIOT.2019.2909745.
6. Behavior changes: all apps. Android 9 (API level 28). Publ. Mar. 01, 2024. URL: <https://developer.android.com/about/versions/pie/android-9.0-changes-all> (Last accessed: 03.05.2024).
7. Turner D. Capturing Audio in Android Q. Publ. July 03, 2019. URL: <https://android-developers.googleblog.com/2019/07/capturing-audio-in-android-q.html> (Last accessed: 03.05.2024).
8. Privacy changes in Android 10. Publ. Apr. 30, 2024. URL: <https://developer.android.com/about/versions/10/privacy/changes#app-access-device-location> (Last accessed: 03.05.2024).
9. About privacy and Location Services in iOS, iPadOS, and watchOS. Publ. Mar. 27, 2024. URL: <https://support.apple.com/en-us/102515> (Last accessed: 03.05.2024).
10. Privacy in Android 11. URL: <https://developer.android.com/about/versions/11/privacy> (Last accessed: 03.05.2024).
11. Spreitzer R., Moonsamy V., Korak T., Mangard S. Systematic classification of side-channel attacks: A case study for mobile devices. *IEEE Commun. Surveys Tutor*. 2018. Vol. 20, Is. 1. P. 465–488.
12. Xu F., Diao W., Li Z., Chen J., Zhang K.. BadBluetooth: Breaking Android security mechanisms via malicious Bluetooth peripherals. *Proc. of the Network and Distributed System Security Symposium (NDSS'19)*, San Diego, CA, USA, 24–27 Feb. 2019. DOI: 10.14722/ndss.2019.23482.
13. Ba Z., Zheng T., Zhang X., Qin Z., Li B., Liu X., Ren K. Learning-based practical smartphone eavesdropping with built-in accelerometer. *Proc. of the Network and Distributed System Security Symposium (NDSS'20)*, San Diego, CA, USA, 23–26 Feb. 2020. DOI: 10.14722/ndss.2020.240762020.
14. Simon L., Anderson R. PIN Skimmer: Inferring PINs through the camera and microphone. *Proc. of the 3rd ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (SPSM'13)*. 08 Nov. 2013. P. 67–78. DOI: 10.1145/2516760.2516770.
15. UK spies can hack smartphones: Snowden. Publ. Oct. 05, 2015. URL: <https://phys.org/news/2015-10-uk-spies-hack-smartphones-snowden.html> (Last accessed: 03.05.2024).
16. Vidsota D. How global media and IT platforms collect users' personal data. Publ. Feb. 16, 2021. URL: <https://biz.nv.ua/ukr/experts/osobisti-dani-ta-socmerezhi-facebook-mesendzheri-ostanni-novini-50142213.html> (Last accessed: 03.05.2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-13>

УДК 004.7

Коляда Костянтин Вячеславович, к.т.н., старший викладач

<https://orcid.org/0000-0002-3962-5791>

Рожко Дмитро Вікторович, магістрант

<https://orcid.org/0009-0008-3444-8598>

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна.

ВИКОРИСТАННЯ МЕТАЕВРІСТИЧНИХ АЛГОРИТМІВ ДЛЯ РОЗВ'ЯЗКУ ЗАДАЧІ КОМІВОЯЖЕРА

Коляда К.В., Рожко Д.В. Використання метаевристичних алгоритмів для розв'язку задачі комівояжера. У статті розглядається та проводиться аналіз існуючих метаевристичних алгоритмів. Можливість їх використання при побудові та передачі даних всередині комп'ютерної системи. Також їх застосування у локальних чи глобальних мережах об'єднаних комп'ютерів. Згадано принципи, які лягли в основу формування поняття алгоритм у галузі інформаційних технологій. Описано актуальність використання метаевристичних алгоритмів в проектуванні мереж передачі даних. Встановлено зміст поняття алгоритм та його складність, використані у контексті даної статті. Визначено метаевристичний алгоритм та його відмінності від інших. Розглянуто особливості конкретних видів даних алгоритмів знаходження оптимального шляху. Визначено концепції метода рою часток, які використовують колективний інтелект для реалізації гравітаційного пошуку закладених в основу даного підходу. Розглянуто окремі види метаевристичних алгоритмів, зокрема мурашиний та бджолиний алгоритми. Описано метод гравітаційного пошуку, закладений в основу евристичних алгоритмів оптимізації, згаданих раніше. Встановлено визначення колективного інтелекту для даного контексту, що ліг в основу даного метода рішення завдання комівояжера. Описано актуальність використання конкретних, раніше згаданих, алгоритмів. Демонстрація систем, у яких можливе використання метаевристичних алгоритмів. Вказано можливість їх прикладного застосування у прикладних сферах діяльності людини та суспільства. Запропоновано методи ефективного застосування мурашиного та бджолиного алгоритмів. Виділено переваги та недоліки їх використання. Сформульовано можливості подальшої модернізації даних конкретних методів рішення завдання комівояжера. Описано переваги бджолиного та мурашиного алгоритмів над генетичними алгоритм. Виділення переваг даної належності для практичного застосування даних алгоритмів для рішення задач оптимізації.

Ключові слова: метаевристичний алгоритм, мурашиний алгоритм, бджолиний алгоритм, евристичний алгоритм, метод рою часток, алгоритм гравітаційного пошуку, колективний інтелект.

Kolyada K., Rozhko D. Using metaheuristic algorithms to solve the Traveling Salesman Problem. The article considers and analyzes the existing metaheuristic algorithms. The possibility of their use in the construction and transfer of data within the computer system. Also their application in local or global networks of united computers. The principles that formed the basis of the formation of the concept of algorithm in the field of information technologies are mentioned. The relevance of using metaheuristic algorithms in the design of data transmission networks is described. The content of the concept of algorithm and its complexity, used in the context of this article, are established. The metaheuristic algorithm and its differences from others are defined. The peculiarities of specific types of these algorithms for finding the optimal path are considered. Concepts of the particle swarm method, which use collective intelligence to implement gravitational search, are defined as the basis of this approach. Some types of metaheuristic algorithms are considered, in particular, ant and bee algorithms. The gravitational search method, which is the basis of the previously mentioned heuristic optimization algorithms, is described. The definition of collective intelligence for this context was established, which formed the basis of this method of solving the traveling salesman's task. The relevance of the use of specific, previously mentioned, algorithms is described. Demonstration of systems in which the use of metaheuristic algorithms is possible. The possibility of their applied application in applied spheres of human and social activity is indicated. Methods of effective application of ant and bee algorithms are proposed. Advantages and disadvantages of their use are highlighted. Possibilities of further modernization of these specific methods of solving the traveling salesman's task are formulated. Advantages of bee and ant algorithms over genetic algorithms are described. Highlighting the advantages of this property for the practical application of these algorithms for solving optimization problems.

Keywords: polynomial algorithm, ant colony optimization, artificial bee colony, heuristic algorithm, particle swarm optimization, graphiti-zation search algorithm, swarm intelligence.

Постановка наукової проблеми.

У наш час з розповсюдженням мережевих цифрових технологій особливої важливості набувають підходи взаємодії між елементами мережі. Величина та складність системи у багато разів виросла у порівнянні з першими поколіннями інтернету. За даними DataReportal зараз всього 5.35 мільярдів людей користується інтернетом, що становить 66.2% всього 8.08 мільярдного населення нашої планети [1]. Дана величина постійно зростає. Ще у 2022 році людей онлайн було 63% від загальної кількості населення, відповідно до даних із згаданого раніше аналітичного джерела. Вказана статистичної інформація нашттовхує на думку, що проблема пошуку оптимального шляху проходження повідомлень мережею є актуальною для ефективної передачі даних. Дана задача

комівояжера успішно вирішується існуючими метаевристичними алгоритмами. Проте їх удосконалення є актуальним для проходження повідомлень оптимальнішим шляхом, що дозволить не задіювати деякі вузли та ребра графа системи для передачі даних, що в свою чергу покращить ефективність передачі повідомлень мережею. Ефективне використання ресурсів мережевого апаратного забезпечення також є важливим при налаштуванні взаємодії між елементами комп'ютерної системи.

Аналіз досліджень.

Спочатку опису методів рішення поставленої задачі варто визначити з використанні термінами у контексті даного дослідження. Згадана ще в темі, задача комівояжера полягає у знаходженні найкоротшого маршруту, що проходить через всі вказані міста хоча б по одному разу. Існує багато різновидів цієї задачі, які вимагають знаходження оптимального маршруту з різними критеріями. Відомі дві групи методів розв'язання - точні і евристичні. Точні методи забезпечують оптимальні рішення за достатньо великий час, тоді як евристичні методи можуть знаходити ефективні шляхи, які не обов'язково є оптимальним рішенням за короткий час. На практиці частіше застосовуються більш складні методи, такі як алгоритм мурашиної колонії. Задача комівояжера є NP-повною, що означає, що для її розв'язання потрібно велика кількість обчислень [2]. Для вирішення проблеми комівояжера можливе її представлення у вигляді математичної моделі на графі. Зазвичай припускається, що кожна вершина графа відповідає місту, а ребро між вершинами показує зв'язок між містами. У рамках даного дослідження вершини являтимуть собою мережеві вузли, а ребра – з'єднання між ними. Кожному ребру можна призначити вагу, яка може відображати відстань, час або вартість подорожі між містами. Для спрощення задачі та забезпечення існування маршруту зазвичай припускають, що модельний граф задачі є повним, тобто між усіма парами вершин існує ребро. Якщо між деякими вузлами немає безпосереднього сполучення, то можна ввести ребра з максимальною вагою, щоб такі ребра не потрапляли до оптимального маршруту через їх велику довжину або вартість. Залежно від того, які ваги накладаються на ребра, існують різні варіанти задачі. Найважливіші серед них - симетрична та метрична задачі. Маршрут, який проходить по всіх вершинах графа лише один раз, який також називається гамільтоновим маршрутом. Метою задачі комівояжера є знайти найкоротший маршрут, який пройде через всі міста.

Метаевристичні методи рішення задачі комівояжера поєднують в собі пошук локальних та глобальних розв'язків для оптимізації задач. Багато з них спочатку знаходять початковий розв'язок і потім його покращують іншими методами. Для уникнення глухих кутів використовуються стратегії, такі як табу-пошук або симуляція відпалу. Деякі методи, наприклад мурашиний алгоритм або генетичні алгоритми, базуються на природніх процесах. Зазначені методи можуть знаходити як якісні, так і віддалені від оптимальних розв'язки. Результат та тривалість обчислення залежать від конкретних елементів, що використовуються. У даному випадку евристичний алгоритм може знайти краще рішення, яке не обов'язково є оптимальним. Часто евристичний алгоритм може бути найкращим рішенням для вирішення складних задач, навіть якщо він не завжди працює правильно [3]. Такі алгоритми широко використовуються на практиці, оскільки вони дозволяють отримувати швидкі та прийнятні результати в обмежені строки не залежно від оптимальності знайденого шляху. Гравітаційний пошук, закладений у метод оптимізації мурашиними та бджолиними колоніями, являє собою ефективний метод знаходження оптимального рішення завдання оптимізації, який базується на моделюванні гравітаційної взаємодії між рішеннями [4]. Дані алгоритми добре підходять для рішення задач з великою кількістю змінних та складними умовами завдань комбінаторної та багатокритерійної оптимізації. У реаліях мережевої чи комп'ютерної системи кількість ітерацій не завжди може бути константною для всіх випадків використання. Конфігурація мережі може змінюватись, тому для предметної області даної статі розглядатиметься поліноміальна версія метаевристичного мурашиного та бджолиного алгоритмів.

Поняття алгоритм за визначенням являє собою послідовність точно визначених кроків, які дозволяють вирішити певну задачу або проблему за скінченну кількість дій, закінчення яких обумовлюється логічними або апаратними обмеженнями електронно обчислювальної машини [5].

Популярним прикладом метаевристичних алгоритмів є мурашиний алгоритм, також відомий як алгоритм оптимізації мурашиної колонії. Він є ефективним методом для знаходження наближених розв'язків задачі комівояжера та схожих завдань на графах. Основна ідея полягає в тому, щоб мурахи маркували успішні шляхи великою кількістю феромону [6].

$$P_i = \frac{l_i^q \cdot f_i^p}{\sum_{k=0}^N l_k^q \cdot f_k^p}$$

де P_i – ймовірність проходження шляхом i , l_i – обернена до ваги i -того переходу величина, f_i – кількість феромону на i переході, q – жадібність алгоритму, p – стабільність алгоритму так, що $q + p = 1$.

Результат використання даного підходу може бути неточним з першого проходження або навіть гіршим, проте повторне застосування алгоритму може дати достатньо точний висновок через ймовірність правильного рішення.

У природі мурахи ходять довільно, але після знаходження їжі повертаються до колонії, залишаючи за собою феромонний слід. Інші мурахи, знаходячи цей шлях, починають за ним слідувати, підсилюючи його. Проте з часом феромони випаровуються, і довші шляхи втрачають привабливість. Як результат, короткі маршрути використовуються частіше, і феромони на них стають все більш концентрованими. Якби феромони не зникали, перші обрані мурахою шляхи були б пріоритетними для інших та розвідка нових шляхів була б обмежена. Даний принцип ігнорує вибір найбільш очевидного шляху і дозволяє мурахам знаходити найкращі шляхи до їжі. Мурашиний алгоритм полягає в наслідуванні цієї стратегії для пошуку оптимального рішення в задачах.

На рисунку 1 продемонстровано початкову концепцію, яка виникла під впливом спостережень за мурахами, які, не дивлячись на свою індивідуальну обмеженість, спроможні колективно знаходити найкоротший шлях до джерела їжі в гніздо.

Наступні кроки описують логіку прийняття рішення:

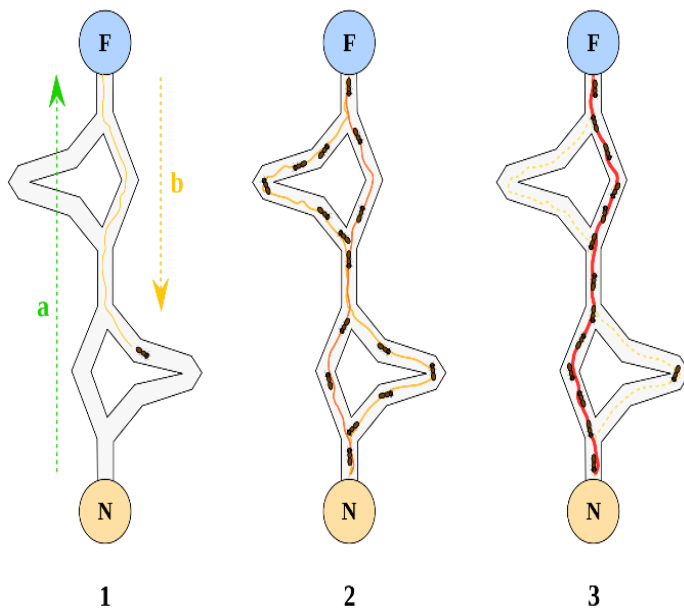


Рис. 1 – прийняття рішень у мурашиних алгоритмах.

1. Перша досліджувана мураха слідує з мурашника (F) в напрямку (b) до джерела їжі (N). Потім, шляхом (a) повертається до мурашника (F), залишаючи за собою феромоновий слід.

2. Мурахи випадково обирають всі шляхи руху між мурашником (F) та джерелом їжі (N).

3. Мурахи ідуть коротшим серед усіх раніше випробуваних шляхів через втрату щільності феромонових слідів на коротших відрізках шляху.

Для формування математичної моделі для використання мурашиним алгоритмом вважатимемо, що середовище для мурах – це неорієнтовний граф, де кожне ребро має вагу, що вказує на відстань між вершинами. Мурахи можуть проходити по графу у будь-якому напрямку, оскільки граф двох скерований. Для вибору шляху мурахи використовують феромони на ребрах: чим більше феромону на ребрі, тим вища ймовірність, що мураха

обере конкретний шлях. Маршрут, який мурахи виберуть, визначає кількість феромону, яку вони залишають на ребрах. Даний чинник забезпечує пошук найкоротшого шляху. При використанні тільки додатній зворотного зв'язку мурашки можуть застрягнути в локально-оптимальному маршруті. Утворення зациклень можна уникнути, використовуючи від'ємний зворотний зв'язок, який озчатиме випаровування феромону. Правильне встановлення часу випарування може допомогти колонії знайти глобально оптимальне рішення.

Рівняння знаходження кількості феромону [6]:

$$\Delta \tau_{ij}^k = \frac{Q}{L^k(t)}$$

Рівнянням можливо обрахувати кількість феромону, яка лишилась на кожному ребрі графа шляху мурахою k інтенсивності Q на шляху L . Як результат, отримуємо короткий шлях, що характеризується великою концентрацією феромону.

Довший шлях містить його значення нижче. У подальшому обрахований результат можна використати у рівнянні для збільшення кількості феромону ребер пройдених шляхів мурахою [6].

$$\tau_{ij}(t) = \Delta\tau_{ij}(t) + (\tau_{ij}^k \cdot \rho)$$

У даному рівнянні важливо, що кожне ребро шляху позначається феромоном, який пропорційний довжині шляху. Потрібно зачекати, доки мураха обійде весь шлях і тільки потім оновити рівні феромону, щоб точно визначити довжину маршруту. Константа ρ має значення від 0 до 1.

Спочатку шляху кожне ребро може обиратися. Для поступового видалення ребер довшого шляху використовується процес випаровування з ребра графу феромону [6]. З рівняння маємо наступну його модифікацію описану формулою:

$$\tau_{ij}(t) = \tau_{ij}(t) \cdot (1 - \rho)$$

Для коректного використання рівняння за випаровування береться обернений коефіцієнт обрахунку шляху.

Одною із варіацій рішення задачі комівояжера метаевристичним алгоритмом є штучна бджолина сім'я, основний принцип якої структурно схожий медоносним бджолам. Бджолиний алгоритм є відносно новим методом для пошуку глобальних екстремумів складних функцій у багатьох вимірах. Алгоритм поєднує сусідній та випадковий пошук і може також використовуватися для оптимізації як комбінаторних, так і функціональних завдань.

Бджолиний алгоритм базується на принципах пошуку оптимального рішення за допомогою колективного інтелекту. Бджоли в алгоритмі шукають найбільш придатні джерела для збирання нектару, використовуючи випадкові початкові координати і швидкості. Кожна бджола може запам'ятати місце, де вона знайшла багато квітів і порівнювати їх з іншими місцями, які відкрили інші бджоли. Коли бджола вибирає між поверненням до свого власного місця або дослідженням місця, визначеного іншими, вона вирішує, на якому місці шукати квіти, керуючись особистим досвідом або соціальними сигналами. Іноді бджола може виявити нове місце з більшою кількістю квітів, яке варто відвідати. Інші бджоли потім дізнаються про це і також направляються туди, спираючись на бджолиний танець. Процес, де бджоли досліджують поле, переходячи до місць з великою кількістю квітів і сповільнюючи рух, щоб перевірити їх. Вони шукають найбільшу концентрацію квітів, порівнюючи з раніше відвіданими місцями. Потім бджола рухається до місця з найбільшою кількістю квітів, і інші бджоли робітничі починають збирати нектар там [6].

У початковому етапі алгоритму методу бджолиної сім'ї встановлюють параметри для оптимізації і визначають допустимий інтервал для пошуку оптимальних значень. Потім бджоли розташовуються випадковим чином в допустимій області, а їх вектори і швидкості руху визначаються, при чому кожна частка повинна переміщатися через простір рішень, якщо вона була б бджолою в сім'ї. Далі проводиться обробка даних для кожної частинки, оцінюючи її придатність порівнюючи з персональною найкращою позицією та глобальною національною позицією. Якщо знайдено значення краще, ніж попереднє, то відповідні позиції замінюються. Коригування швидкості частинки є важливим елементом процесу оптимізації, де важливо правильно визначити швидкість залежно від розміщення позицій. Частинка стежить за найбільшою придатністю у напрямку цих позицій згідно з певним рівнянням:

$$v_n^{i+1} = w \cdot v_n^i + c_1 r(p_n - x_n) + c_2 r(g_n - x_n)$$

де v_n^i – швидкість бджоли а n вимірі попереднього кроку, x_n – координата бджоли n вимірі, p_n – найкраща персональна позиція, g_n – найкраща глобальна позиція, r – довільне число вірогідності від -1 до 1. Обчислення проводяться для кожного n кроку-виміру. Дане рівняння показує, що нова швидкість формується шляхом збільшення або зменшення початкової швидкості в w разів, а також додавання векторів швидкості глобальної і поточної глобальної позиції для конкретного напрямку. c_1 та c_2 являють собою пізнавальні та соціальні коефіцієнти визначення взаємного тяжіння до глобальної та персональної найкращої позиції. Вказані значення визначають вплив на бджолі і визначає її пам'ять про найкраще рішення вибору найкращої персональної позиції та інших членів сім'ї відповідно. Збільшення цих коефіцієнтів сприяє дослідженню простору рішень шляхом руху кожної частинки в напрямку своєї позиції та дослідження глобального максимуму. w вважається

інерціальною вагою від 0 до 1 та відображає міру бджоли, яка характеризує вірність своєму первинному курсові з ігноруванням впливу глобальних та персональних найкращих позицій.

Можливості врахування меж в задачах можуть бути різними. Один із способів – зробити їх поглинаючими. Даний підхід визначає, що бджола у випадку виходу за дозволену область буде повернена в заданий простір рішень. Інший варіант рішення полягає у відображенні швидкості частинки при наближенні до межі. Однак найефективнішим виявилось використання "невидимих стін", де бджола може вилітати за межі дозволеної області. Отримані значення у даному випадку не враховуються, поки вона не повернеться назад.

Метою роботи є аналіз існуючих метаевристичних алгоритмів з подальшим формулюванням нового підходу для рішення задачі комівояжера при відправленні повідомлень між елементами комп'ютерної системи чи мережі.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.

Результатом аналізу опису мурашиного алгоритму є те, даний метод рішення задачі комівояжера є метаевристичним методом оптимізації, інспірованим поведінкою мурах при пошуку шляху джерела їжі. Основна ідея алгоритму полягає в тому, що мурахи залишають феромони на своєму шляху, інші мурахи використовують цю інформацію для вибору найбільш оптимального шляху. Чим більше феромонів залишено на певному шляху, тим ймовірніше, що інші мурахи оберуть саме цей шлях. Переваги мурашиного алгоритму включають його здатність знаходити оптимальні рішення в складних просторах пошуку, а також його здатність до адаптації до змін в навколишньому середовищі. Крім того, цей алгоритм відмінно підходить для задач комбінаторної оптимізації. Однак у мурашиного алгоритму є недоліки. Наприклад, він може сходиться до локального оптимуму, якщо не вживати достатніх заходів для різноманітності шляхів пошуку. Також даний алгоритм може бути неефективним у випадках, коли мурахи не можуть комунікувати між собою або коли інформація про феромони швидко втрачається. У цілому нині, мурашиний алгоритм демонструє хороші результати у вирішенні складних завдань оптимізації, але його ефективного застосування необхідно враховувати специфіку завдання та особливості довкілля.

Результатом аналізу опису бджолиного алгоритму є те, даний метод рішення задачі комівояжера є метаевристичний алгоритм оптимізації, інспірований поведінкою бджіл під час пошуку їжі. Він заснований на ідеї множинного руху рою бджіл навколо квітів, збираючи їжу та обмінюючись інформацією про місцезнаходження квітів.

Принцип роботи бджолиного алгоритму полягає в наступному:

1. Ініціалізація популяції – початкове розподіл бджіл з різних рішень простору пошуку.
2. Визначення критерію якості – функція, яку потрібно оптимізувати.
3. Ітераційний процес – бджоли переміщуються по простору рішень, обмінюючись інформацією та оновлюючи своє положення відповідно до кращих рішень.
4. Конвергенція – алгоритм сходиться до раціонального рішення, враховуючи обмін інформацією та ітераційне оновлення популяції.

Перевагами бджолиного алгоритму являються простота реалізації та ефективність у вирішенні багатьох завдань оптимізації, багаторазове використання практичної інформації при ухваленні рішень про наступний крок, здатність знаходити оптимальне рішення за наявності кількох локальних екстремумів.

Недоліками алгоритму бджолиної колонії є нестійкість до вибору параметрів та функцій адаптації, висока обчислювальна складність зі збільшенням розмірності простору пошуку, необхідність ретельного налаштування параметрів для конкретного завдання.

Таким чином, бджолиний алгоритм є потужним інструментом для вирішення завдань оптимізації, який ефективний у вирішенні різних завдань, проте вимагає ретельного налаштування та вибору параметрів для досягнення оптимальних результатів.

Спільними рисами розглянутих алгоритмів є те, що і мурашині, і бджолині є евристичними ройовими алгоритмами гравітаційного пошуку з використанням колективного інтелекту. Вони мають випадкову поведінку, але був розроблений детермінований алгоритм на основі гравітаційної кінематики під назвою "центральна оптимізація сил" та починають пошук з однієї точки. Розвідка полягає в розширенні простору пошуку, а використання - у знаходженні оптимального розв'язку. Для ефективного пошуку потрібно знаходити баланс між цими двома аспектами. Щоб уникнути підтримання в локальному оптимумі, алгоритм повинен розглядати результати розвідки на

```
point:
| num | x | y |
+-----+-----+
| 0 | 12.6656 | 34.3967 |
| 1 | 58.3999 | 47.5795 |
| 2 | -50.2695 | 86.1596 |
| 3 | -78.4910 | 16.5287 |
| 4 | -24.0765 | 53.6886 |
| 5 | 40.6567 | -21.2777 |
| 6 | -10.8025 | -91.8239 |
| 7 | -83.2807 | -6.0268 |
| 8 | -39.3477 | -42.7264 |
| 9 | -32.1911 | 75.3474 |
```

Рис. 2 – задані точки з координатами для симуляції мурашиної колонії

```
best distance = 532.533
path:
| id from | id to | distance |
+-----+-----+
| 3 | 4 | 65.8923 |
| 4 | 2 | 41.7186 |
| 2 | 9 | 21.0650 |
| 9 | 0 | 60.7378 |
| 0 | 1 | 47.5963 |
| 1 | 5 | 71.1065 |
| 5 | 6 | 87.3202 |
| 6 | 8 | 56.7925 |
| 8 | 7 | 57.2448 |
| 7 | 3 | 23.0584 |
```

Рис. 4 – найкращий маршрут знайдений мурашиною колонією.

```
iter | min | max | best
+-----+-----+-----+
| 1 | 666.987 | 1052.95 | 666.987 |
| 2 | 661.275 | 971.391 | 661.275 |
| 3 | 532.533 | 1098.29 | 532.533 |
| 4 | 586.978 | 979.913 | 532.533 |
| 5 | 585.27 | 997.433 | 532.533 |
| 6 | 566.662 | 1038.35 | 532.533 |
| 7 | 532.533 | 978.047 | 532.533 |
| 8 | 621.485 | 1054.17 | 532.533 |
| 9 | 534.693 | 989.837 | 532.533 |
| 10 | 633.867 | 942.93 | 532.533 |
| 11 | 610.697 | 953.805 | 532.533 |
| 12 | 621.291 | 1005.77 | 532.533 |
| 13 | 586.978 | 1020.95 | 532.533 |
| 14 | 666.169 | 1017.43 | 532.533 |
| 15 | 601.881 | 995.576 | 532.533 |
| 16 | 594.415 | 1020.02 | 532.533 |
| 17 | 554.668 | 976.186 | 532.533 |
| 18 | 554.522 | 1031 | 532.533 |
| 19 | 620.638 | 1049.04 | 532.533 |
| 20 | 564.309 | 966.198 | 532.533 |
| 21 | 564.309 | 1019.38 | 532.533 |
| 22 | 619.242 | 999.078 | 532.533 |
| 23 | 659.674 | 950.696 | 532.533 |
```

Рис. 3 – обходи точок мурашиною колонією з знайденим найкращим результатом на кожному проході

```
iter | min | max | best
+-----+-----+-----+
| 1 | 522.257 | 1087.47 | 522.257 |
| 2 | 580.395 | 1081.39 | 522.257 |
| 3 | 557.754 | 1047.5 | 522.257 |
| 4 | 603.507 | 980.905 | 522.257 |
| 5 | 503.944 | 999.468 | 503.944 |
| 6 | 540.502 | 959.941 | 503.944 |
| 7 | 532.533 | 958.67 | 503.944 |
| 8 | 557.256 | 987.511 | 503.944 |
| 9 | 503.944 | 985.896 | 503.944 |
| 10 | 634.692 | 1041.03 | 503.944 |
| 11 | 569.059 | 971.638 | 503.944 |
| 12 | 634.949 | 999.326 | 503.944 |
| 13 | 576.659 | 1018.68 | 503.944 |
| 14 | 543.837 | 954.769 | 503.944 |
| 15 | 627.237 | 980.047 | 503.944 |
| 16 | 616.752 | 1041.28 | 503.944 |
| 17 | 566.662 | 972.093 | 503.944 |
| 18 | 566.662 | 971.183 | 503.944 |
| 19 | 522.257 | 952.629 | 503.944 |
| 20 | 572.834 | 1024.71 | 503.944 |
| 21 | 540.502 | 950.626 | 503.944 |
| 22 | 564.502 | 963.866 | 503.944 |
| 23 | 569.059 | 1025.25 | 503.944 |
| 24 | 623.793 | 1037.15 | 503.944 |
| 25 | 622.167 | 1020.78 | 503.944 |
```

Рис. 5 – обходи точок мурашиною колонією з обміном інформацією про знайдений найкращий результат на кожному проході.

```
best distance = 503.944
path:
| id from | id to | distance |
+-----+-----+
| 1 | 5 | 71.1065 |
| 5 | 6 | 87.3202 |
| 6 | 8 | 56.7925 |
| 8 | 7 | 57.2448 |
| 7 | 3 | 23.0584 |
| 3 | 2 | 75.1327 |
| 2 | 9 | 21.0650 |
| 9 | 4 | 23.1290 |
| 4 | 0 | 41.4989 |
| 0 | 1 | 47.5963 |
```

Рис. 6 – найкращий маршрут знайдений мурашиною колонією з обміном інформацією між мурахами.

початкових ітераціях. Вони надають хороші результати у цільових сферах використання, тому що немає одного універсального алгоритму, який вирішить всі оптимізаційні завдання. Мурашині та бджолині алгоритми симулюють активність методом рою часток. Даний метод оптимізації не вимагає знання градієнта оптимізованої функції. Він був створений для імітації соціальної поведінки, але виявився більш ефективним для оптимізації. За допомогою нього оптимізується функція, утримуючи популяцію часток, які переміщуються в просторі розв'язків за певною формулою [8]. Відповідно до даного методу, кожна частинка, являє собою потенційне рішення задачі, яке рухається в просторі з цілю знаходження оптимального рішення.

Для наочного аналізу мурашиного алгоритму створено симуляцію мурашиної колонії для знаходження оптимального шляху проходження через всі точки зазначені на рисунку 2. Позиції

точок задані величинами координатної площини x, y . Враховано також ситуації, коли координати є від'ємними величинами для коректності дослідження. Дані координати можуть бути також представлені у вигляді неорієнтовного графа з однаковими вагами на всіх ребрах. У рамках демонстрації та для практичного аналізу мурашиного алгоритму дані спрощення не є критичними. На рисунку 4 продемонстровано результати роботи оптимізації мурашиної колонії на всіх ітераціях проходження маршруту. Із отриманої таблиці можна спостерігати, що найкращий шлях було отримано ще при 3 проходженні через усі точки. Наступні 20 кроків отримання найкращих шляхів є рівними між собою. У випадку прикладного використання рішення завдання комівояжера для вибору короткого маршруту з'єднання згадані однакові шляхи можуть використовуватись при паралельній передачі даних, коли, наприклад, 20 комп'ютерів одночасно відправляють повідомлення в спільною мережею в одному отримувачу. Рисунок 4 демонструє конкретні вузли останнього знайденого короткого маршруту передачі даних від вузла 3 до вузла 7 та назад. Слід зазначити, що в отриманому результаті не з'являється жодного повторного проходження вузлами, тобто знайдено гамільтонів шлях.

Після аналізу опису бджолиного алгоритму з'являється гіпотеза для покращення продемонстрованої раніше мурашиної колонії. Особливістю бджолиної колонії є обмін інформацією між частинками рою, чого не роблять мурахи, виходячи з аналізу опису відповідного алгоритму. Як результат, існуючу мурашину колонію удосконалено обміном даними про знайдений шлях іншими частинками рою. Як результат отримано результат з рисунка 5, що демонструє дещо кращі показники величини найкоротшого шляху. Інші риси мурашиної колонії залишаються незмінними. У даному досліді також є 20 рівних між собою коротких шляхи для можливості відправки відповідної кількості повідомлень мережею альтернативними маршрутами. Відповідно до рисунка 6, в удосконаленій версії мурашиного алгоритму також не з'являється жодного повторювального вузла. У даному випадку також знайдено гамільтонів цикл.

Висновки та перспективи подальшого дослідження.

1. З описаних описів алгоритм оптимізації мурашиної колонії може використовуватися для розв'язання складних завдань оптимізації, які передбачають пошук найбільш відповідного рішення для оптимізації цільової функції. Це може бути пошук мінімуму або максимуму ціни, точності, часу, відстані та інших параметрів з дискретної множини можливих рішень.

Метод бджолиної сім'ї дозволяє ефективно розподілити процес на паралельні задачі, що значно прискорює його виконання. Порівняно з генетичним алгоритмом, де є кілька операторів, метод бджолиної сім'ї використовує лише один, який полягає у обчисленні швидкості, яка спрощує його використання. Даний підхід може мати схожу специфіку та сферу застосування як і метод оптимізації мурашиної колонії.

2. З аналізу алгоритму оптимізації за допомогою мурашиної колонії слідує те, що даний підхід може знаходити застосування у задачах календарного планування, маршрутизації транспорту, розподілу ресурсів та багатьох інших сферах. Дослідження показали, що метод мурашиної колонії може приносити результати, які перевершують результати генетичних мурашиних алгоритмів та нейронних мереж.

Методом бджолиної сім'ї легко визначається глобальний мінімум, на відміну від генетичних алгоритмах ця задача ускладнена. Ці методи базуються на різних природних процесах: метод бджолиної сім'ї використовує соціальну поведінку бджіл, а генетичний алгоритм імітує еволюцію та природний відбір. Є можливість об'єднати ці два методи для покращення результатів.

У випадку практичного застосування мурашиного та бджолиного алгоритмів для отримання оптимальних маршрутів для під'єднання у мережі «Інтернет» чи побудови оптимального логістичного шляху маємо важливим є також наявність набору оптимальних маршрутів, які формуються у процесі обрахунків. У випадку перевантаження мережевого чи логістичного шляху у залежності від поставленого завдання комівояжера є можливість використання менш оптимальних шляхів, які не є найвигіднішими серед оптимальних. Запропонований метод додаткового обміну даними між мурахами на кожній ітерації дозволяє деякою мірою покращити ефективність пошуку оптимального шляху. У подальшому даний підхід може застосовуватись для пошуку короткого маршруту відправки повідомлень мережею та покращуватись відповідно до вимог конкретної мережі.

Список бібліографічного опису:

1. S. Kemp, "Digital 2024: Global overview report," DataReportal – Global Digital Insights, Jan. 31, 2024. <https://datareportal.com/reports/digital-2024-global-overview-report>
2. E. L. Lawler, J. K. Lenstra, A H G Rinnooy Kan, and David Bernard Shmoys, The Traveling salesman problem : a guided tour of combinatorial optimization. Chichester West Sussex ; New York: Wiley, 1990.
3. M. J. Apter, The Computer Simulation of Behaviour. Routledge, 2018.
4. Sajad Ahmad Rather, Enhanced Gravitational Search Algorithms for Solving Global Optimization Problems. 2023.
5. S. S. Skiena, The algorithm design manual. London: Springer, 2012.
6. M. Dorigo and StützleT., Ant colony optimization / Ant colony optimization. Cambridge, Mass.: Mit Press, 2004.
7. Aboul Ella Hassanien and Eid Alamry, Swarm intelligence : principles, advances, and applications. Boca Raton, Florida ; London, England ; New York, New York: Crc Press, 2016.
8. R. C. Eberhart, Yuhui Shi, and J. F. Kennedy, Swarm intelligence c : y Russell C. Eberhart, Yuhui Shi, James Kennedy. Morgan Kaufmann, 2001.

References:

9. S. Kemp, "Digital 2024: Global overview report," DataReportal – Global Digital Insights, Jan. 31, 2024. <https://datareportal.com/reports/digital-2024-global-overview-report>
10. E. L. Lawler, J. K. Lenstra, A H G Rinnooy Kan, and David Bernard Shmoys, The Traveling salesman problem : a guided tour of combinatorial optimization. Chichester West Sussex ; New York: Wiley, 1990.
11. M. J. Apter, The Computer Simulation of Behaviour. Routledge, 2018.
12. Sajad Ahmad Rather, Enhanced Gravitational Search Algorithms for Solving Global Optimization Problems. 2023.
13. S. S. Skiena, The algorithm design manual. London: Springer, 2012.
14. M. Dorigo and StützleT., Ant colony optimization / Ant colony optimization. Cambridge, Mass.: Mit Press, 2004.
15. Aboul Ella Hassanien and Eid Alamry, Swarm intelligence : principles, advances, and applications. Boca Raton, Florida ; London, England ; New York, New York: Crc Press, 2016.
16. R. C. Eberhart, Yuhui Shi, and J. F. Kennedy, Swarm intelligence c : y Russell C. Eberhart, Yuhui Shi, James Kennedy. Morgan Kaufmann, 2001.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-14>

УДК 004.02

Коровій Олександр Сергійович, аспірант

<https://orcid.org/0009-0002-2835-9173>

Терейковський Ігор Анатолійович, д.т.н., професор

<https://orcid.org/0000-0003-4621-9668>

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

КОНЦЕПТУАЛЬНА МОДЕЛЬ ПРОЦЕСУ ВИЗНАЧЕННЯ ЕМОЦІЙНОГО ЗАБАРВЛЕННЯ ТА ТОНАЛЬНОСТІ ФРАГМЕНТІВ ТЕКСТУ

Коровій О.С., Терейковський І.А. Концептуальна модель процесу визначення емоційної тональності тексту. Задача розпізнавання емоційної тональності текстових фрагментів є важливою проблемою в галузі обробки природної мови з значним академічним і комерційним потенціалом. У роботі представлено концептуальну модель процесу визначення емоційної тональності тексту. Модель забезпечує формалізований та детальний опис задачі, сприяючи розробці відповідних засобів для розпізнавання емоційного забарвлення та тональності фрагментів тексту. Вона систематизує ключові етапи, включаючи формування навчальних даних, вибір архітектури нейронної мережі, її налаштування та навчання, а також безпосереднє використання для розпізнавання емоційного забарвлення та тональності. Концептуальна модель визначає залежності між різними факторами, що впливають на ефективність процесу розпізнавання емоційної тональності, такими як якість вхідних даних, параметри нейромережевої моделі, споживання ресурсів та похибки під час навчання і розпізнавання. Це дозволяє виявити критичні аспекти, які слід врахувати для підвищення загальної ефективності системи. Модель також передбачає залучення експертних знань на різних етапах для покращення якості вхідних даних та оптимізації параметрів моделі. Запропонована концептуальна модель може слугувати основою для розробки ефективних засобів розпізнавання емоційного забарвлення та тональності, що матиме практичне застосування в різних галузях, таких як аналіз соціальних мереж, обслуговування клієнтів та маркетинг.

Ключові слова: емоційна тональність, концептуальна модель, розпізнавання емоцій, обробка природної мови, нейронні мережі

Korovii O., Tereikovskiy I. Conceptual model of the process of determining the emotional tonality of the text. The task of recognizing the emotional tonality of text fragments is an important problem in the field of natural language processing with significant academic and commercial potential. This paper presents a conceptual model of the process of determining the emotional tonality of text. The model provides a formalized and detailed description of the task, facilitating the development of appropriate tools for recognizing the emotional coloring and tonality of text fragments. It systematizes the key stages, including the formation of training data, selection of the neural network architecture, its configuration and training, as well as direct use for recognizing emotional coloring and tonality. The conceptual model defines the dependencies between various factors that affect the efficiency of the emotional tonality recognition process, such as the quality of input data, parameters of the neural network model, resource consumption, and errors during training and recognition. This allows identifying critical aspects that should be considered to improve the overall efficiency of the system. The model also provides for the involvement of expert knowledge at various stages to improve the quality of input data and optimize model parameters. The proposed conceptual model can serve as a basis for the development of effective means of recognizing emotional coloring and tonality, which will have practical applications in various fields, such as social network analysis, customer service, and marketing.

Keywords: emotional tone, conceptual model, emotion recognition, natural language processing, neural networks

Постановка наукової проблеми.

Розв'язання проблеми ефективного розпізнавання емоційного забарвлення та тональності фрагментів тексту належить до актуальних завдань в галузі обробки природної мови, а також стало важливою темою через його значний академічний та комерційний потенціал. З розвитком технологій машинного навчання, це привертає все більше уваги тому, що задача розпізнавання емоційного забарвлення та тональності фрагментів тексту, розв'язанню якій і присвячена представлена наукова робота є актуальною на даний момент.

Аналіз досліджень. Важливість розпізнавання емоційного забарвлення та тональності (РЕЗТ) текстових фрагментів залишається незмінно великою сферою [1-3], особливо з огляду на масове використання соціальних мереж та інтернету загалом. Крім того, цей напрямок має значну наукову цінність, оскільки результати досліджень у цій сфері знаходять застосування в різноманітних сервісах і продемонстрували свою ефективність у комерційному секторі. Ефективність РЕЗТ фрагментів тексту визначається не лише точністю моделей, але і їхньою здатністю адаптуватися до динамічних умов зміни моделі світу, нових даних, тощо [3].

Існує багато різних підходів та методів, які використовуються для вирішення цього завдання, включаючи машинне навчання, глибинне навчання, обробку природної мови, лексикон-базовані методи та ін.

Більшість досліджень зосереджені на класифікації емоцій у текстах різних типах текстових документах, таких як відгуки про товари [6, 7], новинні статті [8], повідомлення в соціальних мережах [5, 9], електронні листи [4] тощо. Вибір конкретних типів емоцій для класифікації значно відрізняється в різних роботах. Найпоширенішим є розгляд базових емоцій - позитивних, негативних та нейтральних [6, 9], що відповідає завданню аналізу тональності. Однак, багато досліджень розглядають більш детальну класифікацію на конкретні емоції, такі як радість, сум, гнів, страх, здивування, огида [4, 7] та ін. Це дозволяє глибше зрозуміти емоційний стан автора тексту. Деякі роботи також враховують інтенсивність емоцій [5] або розглядають кілька емоційних міток одночасно для одного тексту [7].

Значна увага в літературі [1-11] приділяється використанню глибинних нейронних мереж та методів глибинного навчання для РЕЗТ у текстах [6, 10, 11]. Ці підходи показали високу ефективність порівняно з традиційними машинно-навчальними методами. Найбільш поширеними архітектурами є згорткові нейронні мережі (CNN) [6, 10], які дозволяють вилучати локальні ознаки з тексту, та рекурентні нейронні мережі (RNN, LSTM, GRU) [4, 10], здатні обробляти послідовності та враховувати контекст. Часто використовуються двонаправлені рекурентні нейронні мережі [6] або комбіновані варіанти цих архітектур, а також моделі які засновуються на механізмі уваги [6, 10], які дозволяють мережі зосереджуватися на найбільш значущих частинах вхідного тексту, це передові моделі трансформерів (BERT, GPT тощо) [11], які також успішно застосовуються для розпізнавання емоційної тональності. Крім того, досліджуються ансамблеві та гібридні підходи [7], що поєднують декілька архітектур для покращення продуктивності.

Проте, автори відзначають низку викликів та обмежень, серед яких:

- Нестача великих, якісних та розмічених емоційних даних.
- Розмиті межі між емоціями та їх суб'єктивність.
- Неповна емоційна інформація в текстах.
- Складність РЕЗТ у діалогах та контекстуальних повідомленнях.

Крім того, слід вказати на невизначеність процесу РЕЗТ тексту що являється важливою запорукою розробкою ефективних засобів розпізнавання. Як показують результати [4-11], для виправлення вказаного недоліку слід розробити концептуальну модель процесу в РЕЗТ.

Слід зазначити, що концептуальна модель може слугувати основою для розробки більш ефективних та універсальних алгоритмів РЕЗТ у текстах, що матиме практичне застосування в різних галузях, таких як аналіз соціальних мереж, клієнтський сервіс, маркетинг тощо. Крім того, формалізована концептуальна модель дозволить систематизувати існуючі підходи, виявити їхні переваги та недоліки, а також допоможе визначити напрямки для подальших наукових досліджень. Це сприятиме кращому розумінню природи емоцій у текстах, розробці нових методів представлення та вилучення емоційних ознак, удосконаленню архітектур нейронних мереж для обробки емоційної інформації. Концептуальна модель також може слугувати основою для створення уніфікованих емоційних корпусів та стандартизованих метрик оцінювання, що полегшить порівняння різних підходів та алгоритмів між собою.

Мета роботи. Метою є розробка концептуальної моделі процесу РЕЗТ текстових фрагментів, що забезпечує формалізований та детальний опис задачі розробки відповідних засобів розпізнавання.

Виклад основного матеріалу дослідження. У загальному випадку концептуальна модель розпізнавання емоційного забарвлення та тональності у текстах являє собою модель предметної області, що складається з переліку взаємопов'язаних концепцій та понять, які використовуються для опису цієї області. Вона включає властивості та характеристики емоційних проявів у текстах, класифікацію емоцій за типами, ситуаціями, ознаками, а також приховані патерни та принципи, що керують емоційними процесами у текстових даних. Концептуальна модель, в даному випадку, є відображенням концепції емоційності тексту, під якою розуміється спосіб трактування емоційних проявів, основна точка зору та провідні ідеї для їх систематичного аналізу та інтерпретації.

Для визначення ефективності процесу РЕЗТ тексту доцільно використовувати визначення та підходи з галузі комп'ютерної лінгвістики, обробки природної мови та машинного навчання, оскільки ці галузі лежать в основі методів та моделей для автоматизованого виявлення емоцій у текстових даних. Концептуальна модель має охоплювати різні аспекти емоційності текстів та *забезпечувати* систематичний підхід до їх аналізу та інтерпретації.

Згідно міжнародних стандартів, ефективність програмної системи визначається низкою атрибутів, що характеризують її продуктивність, споживання ресурсів та дотримання встановлених норм і вимог. Ключовими показниками ефективності є швидкість виконання операцій, обсяг спожитих обчислювальних ресурсів та відповідність системи галузевим стандартам. Зокрема, до основних атрибутів ефективності належать: швидкість (час відгуку та обробки даних), ресурсоємність (споживання апаратних та програмних ресурсів) і узгодженість (дотримання правил, регламентів та специфікацій).

На першому етапі створення концептуальної моделі для розпізнавання емоційної тональності у текстах було проведено гармонізацію термінології, що використовується в цій галузі. Гармонізація виконувалась з метою відобразити сучасний стан наукових досліджень та практичних застосувань, а також забезпечити підтримку завдань, пов'язаних з розробкою систем РЕЗТ текстів. В результаті були визначені такі ключові терміни:

- Емоції в тексті - це суб'єктивне емоційне забарвлення, яке виражається автором через слова, фрази, стиль написання та передає певні почуття, настрої або ставлення до описуваних явищ, ситуацій чи об'єктів.
- Базові емоції - фундаментальні емоційні стани, до яких належать гнів, огида, смуток, зневага, страх, подив, радість.
- Нейронна мережа (НМ) - обчислювальна система, що імітує роботу біологічних нейронних мереж і складається з взаємопов'язаних простих елементів (штучних нейронів) для паралельної обробки інформації.
- Глибинне навчання - це підрозділ машинного навчання, який використовує багатопланові нейронні мережі для автоматичного вивчення ієрархічних представлень даних. Воно дозволяє моделям самостійно вилучати та трансформувати релевантні ознаки з необроблених даних.
- Нейромережева модель (НММ) - модель машинного навчання, побудована на основі нейронної мережі, яка здатна вивчати та узагальнювати приховані закономірності в даних.
- Глибинна нейронна мережа (або, Глибока нейронна мережа) - це тип нейронної мережі, що містить кілька прихованих шарів між вхідним та вихідним шарами. Глибокі архітектури можуть ефективно вивчати складні взаємозв'язки та абстракції в даних завдяки великій кількості нелінійних перетворень.
- Фрагмент тексту - обмежена послідовність слів, словосполучень або речень, що формують змістовну цілісність.
- Емоційні ознаки - лінгвістичні, стилістичні або контекстуальні маркери в тексті, що вказують на певні емоції.
- Тональність - загальна емоційна оцінка фрагменту тексту, або всього тексту, що класифікується як позитивна, негативна або нейтральна, вимірюється значення від -1 до 1, де -1 – негативна, 0 – нейтральна, 1 – позитивна.
- Похибка першого роду - це помилка, яка виникає, коли система класифікує негативний приклад, як позитивний. У контексті розпізнавання емоцій це означає, що система невірно визначила наявність емоції в тексті, де її немає.
- Похибка другого роду - це помилка, коли система класифікує позитивний приклад як негативний. У контексті розпізнавання емоцій це означає, що система пропустила наявну емоцію в тексті.
- Емоційне забарвлення - це міра присутності базових емоцій у текстовому фрагменті. Воно може бути представлене, як набір значень від 0 до 1 для кожної базової емоції, де 0 означає повну відсутність цієї емоції, а 1 - максимальну присутність.

Ці визначення узгоджені з сучасними поглядами на природу емоцій та їх прояв у текстових даних. Вони забезпечують концептуальну основу для формального опису та моделювання процесу виявлення емоційного забарвлення фрагментів тексту за допомогою комп'ютерних алгоритмів.

Окрім того, в рамках дослідження цієї статті було встановлено, що концептуальна модель є засобом для систематизації причинно-наслідкових відносин, які є характерними для процесу ідентифікації емоцій. Це обумовлено потребою в оптимізації ефективності вказаного процесу.

У концептуальній моделі для РЕЗТ у текстах необхідно врахувати кілька ключових аспектів. По-перше, умови функціонування самої системи розпізнавання та її взаємодії з іншими компонентами. По-друге, потрібно забезпечити ефективне застосування нейромережевої моделі для

РЕЗТ. Крім того, модель повинна передбачати можливість управління системою та визначення її керованих параметрів.

Згідно із загальноприйнятою технологією використання НММ, процес розпізнавання емоційної тональності у текстах можна декомпозувати на кілька ключових етапів:

1. Формування бази даних фрагментів текстів: На цьому етапі створюється набір текстових даних, з розміткою відповідних емоцій, які будуть використовуватися для навчання та тестування нейронної мережі.

2. Формування параметрів навчальних та тестових вибірок: На цьому кроці визначаються та налаштовуються параметри для формування навчальної та тестової вибірок з наявних текстових даних.

3. Формування навчальної та тестової вибірки: Цей крок передбачає підготовку даних та поділ їх на навчальну та тестову вибірки для навчання та оцінки якості нейронної мережі відповідно.

4. Визначення типу НММ: Залежно від типу даних і вимог, які покладаються на НМ, приймається рішення щодо вибору найбільш відповідного типу нейронної мережі для розв'язання поставленої задачі.

5. Навчання НММ: Обрана нейромережева модель навчається на навчальній вибірці даних.

6. Верифікація НММ та оцінка якості: На цьому етапі відбувається верифікація та тестування НММ на навчальній та тестовій вибірках для оцінки похибок 1-го та 2-го роду при РЕЗТ та перевірка моделі відповідності вимогам до точності та ресурсоемкості.

7. Використання НММ для розпізнавання: Тут відбувається власне застосування навченої нейромережевої моделі для РЕЗТ на нових текстових даних.

Ця послідовність кроків зображена на рисунку 1.

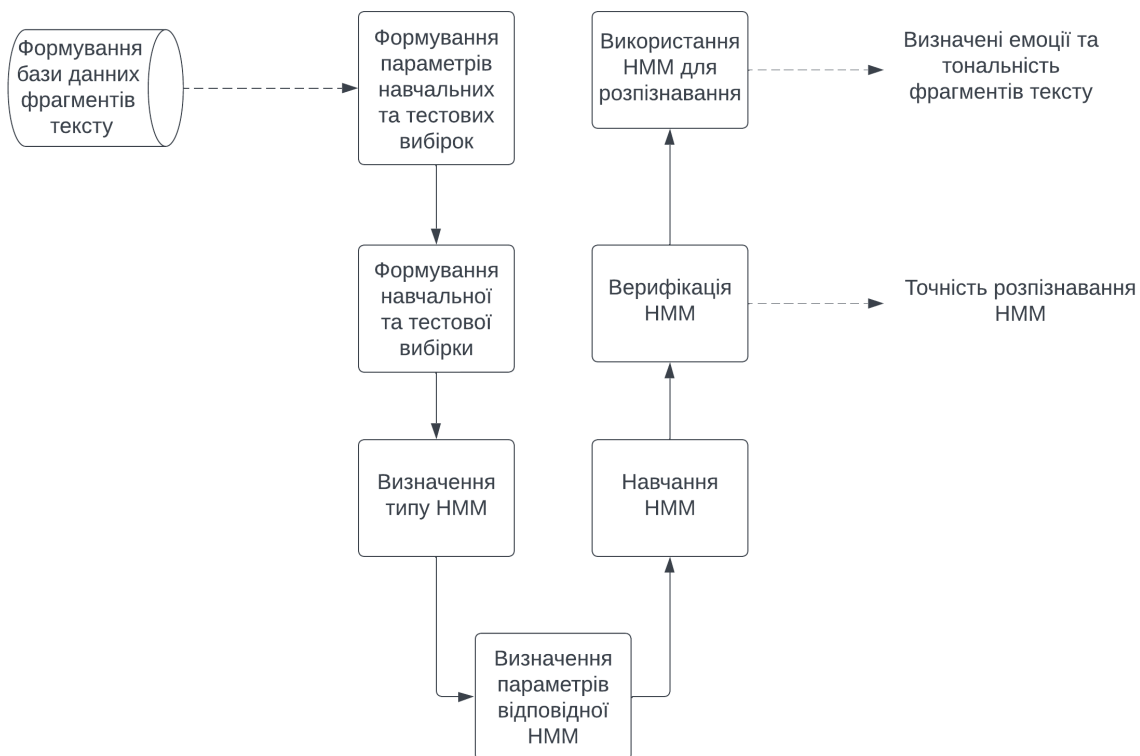


Рис. 1. Схема застосування нейромережевих моделей для розпізнавання.

При розробці концептуальної схеми РЕЗТ у текстових фрагментах на основі нейронних мереж необхідно врахувати кілька ключових особливостей. По-перше, важливим етапом є обґрунтований вибір баз даних фрагментів текстів, які будуть використовуватися для навчання та тестування системи. Від якості та репрезентативності цих текстових даних залежить здатність системи адекватно розпізнавати емоції в різноманітних текстах. По-друге, визначення оптимального виду та параметрів нейромережевої моделі для розпізнавання емоційної тональності

є складним завданням, що потребує ретельного налаштування та експериментів. Архітектура, глибина, функція оцінки похибки та інші параметри моделі суттєво впливають на її продуктивність.

Також слід зазначити недосконалість існуючих способів формування навчальних прикладів для неймережових моделей, призначених для РЕЗТ у текстах. Це стосується як вибору відповідних текстових фрагментів, так і встановлення релевантних емоційних категорій для кожного прикладу. Через ці особливості концептуальна модель передбачає залучення експертів на ключових етапах - формування баз даних текстових фрагментів, створення навчальних прикладів та визначення параметрів неймережової моделі. Експертний досвід та знання дозволяють підвищити якість вхідних даних та оптимізувати роботу системи розпізнавання емоційної тональності. Тому це було враховано при створенні схеми взаємодії компонентів НММ РЕЗТ у фрагментах текстів зображених на рисунку 2.

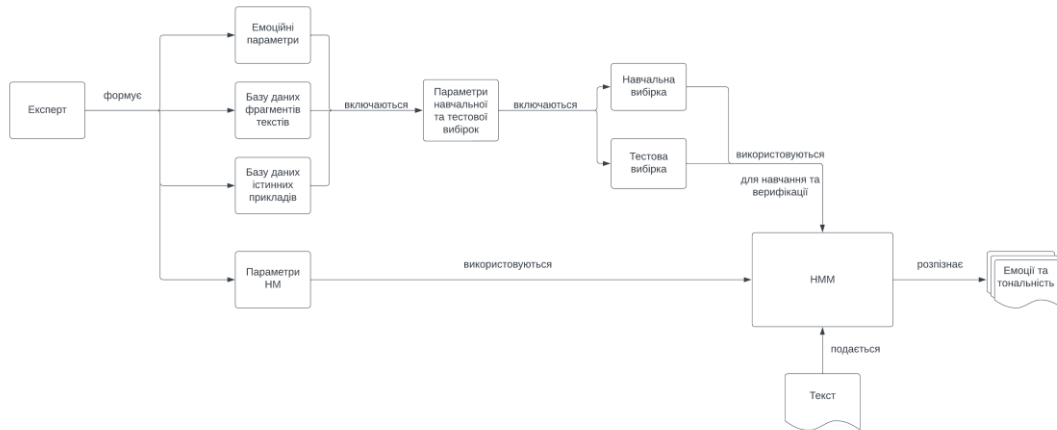


Рис. 2. Схема взаємодії компонентів НММ РЕЗТ у фрагментах текстів

Виходячи з аналізу схем, зображених на рис. 1 та рис. 2, було проаналізовано залежності, які впливають на ефективність НММ для РЕЗТ.

На рис. 3. представлено схему взаємопов'язаних процесів та чинників, які визначають ефективність НММ у завданні РЕЗТ у текстових фрагментах.

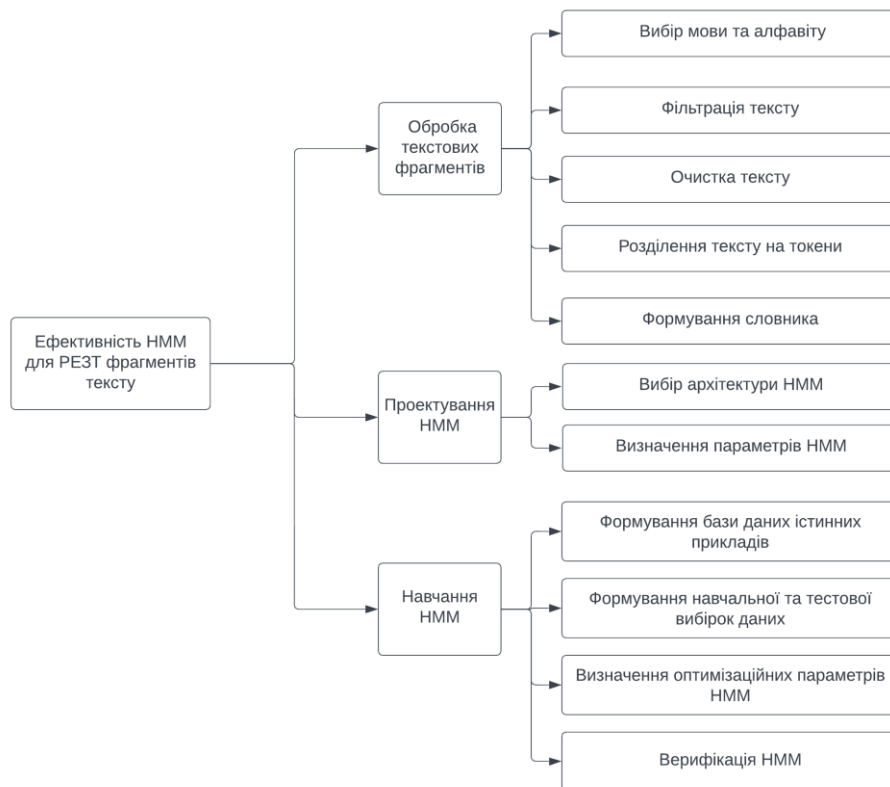


Рис. 3. Схема взаємопов'язаних процесів та чинників, які визначають ефективність НММ, для розпізнавання емоційної тональності фрагментів тексту.

Ключовими залежностями, що впливають на ефективність НММ, є:

- належна обробка вхідних текстових даних є критично важливою для коректного навчання та роботи НММ. Якість та репрезентативність навчальних даних визначають здатність НММ узагальнювати та розпізнавати емоції та тональність в різноманітних текстах.
- правильний вибір архітектури та налаштування параметрів нейромережі забезпечують її здатність ефективно вивчати закономірності в даних та розпізнавати емоції.
- правильний підбір оптимізаційних параметрів під час навчання дозволяє досягти максимальної ефективності НММ у завданні РЕЗТ.

Також, в рамках концептуальної моделі потрібно навести формули для метрик оцінки якості РЕЗТ:

1. Вірність (англ. Accuracy) - вимірює частку всіх вірно класифікованих результатів (і позитивних, і негативних) відносно всіх результатів. Вона включає істинно позитивні та істинно негативні результати та є мірою загальної правильності прогнозів моделі.:

$$\begin{aligned} \text{Вірність} &= \frac{\text{Істинно Позитивні (TP)} + \text{Істинно Негативні (TN)}}{\text{Всього прикладів}} = \\ &= \frac{\text{ІП (TP)} + \text{ІН (TN)}}{\text{ІП (TP)} + \text{ХП (FP)} + \text{ІН (TN)} + \text{ХН (FN)}} \end{aligned} \quad (1)$$

де TP - істинно позитивні (англ. true positives) класифіковані приклади, TN - істинно негативні (англ. true negatives), FP - хибно позитивні (англ. false positives), FN - хибно негативні (англ. false negatives).

2. Точність (англ. Precision) - відображає частку істинно позитивних результатів відносно всіх результатів, які модель класифікувала як позитивні (сума істинно позитивних та хибно позитивних результатів). Це міра якості позитивних прогнозів моделі:

$$\text{Точність} = \frac{\text{Істинно Позитивні (TP)}}{\text{Істинно Позитивні (TP)} + \text{Хибно Позитивні (FP)}} \quad (2)$$

3. Повнота (англ. Recall) - вимірює здатність моделі ідентифікувати всі релевантні випадки в даних. Вона показує, яку частку із всіх реально позитивних випадків модель змогла правильно класифікувати як позитивні.:

$$\text{Повнота} = \frac{\text{Істинно Позитивні (TP)}}{\text{Істинно Позитивні (TP)} + \text{Хибно Негативні (FN)}} \quad (3)$$

Вимірювання похибки є критично важливим етапом у процесі навчання нейронних мереж для завдань РЕЗТ у текстових даних. Функція втрат кількісно оцінює розбіжність між передбаченими моделлю імовірностями та фактичними мітками емоцій у навчальному наборі даних. Її значення використовується, як цільова функція для оптимізації вагових коефіцієнтів нейронної мережі з метою мінімізації похибки під час навчання.

У випадку завдання визначення присутності базових емоцій у текстах, яке розглядається в рамках розпізнавання емоційного забарвлення тексту (РЕЗТ), широко застосовується бінарна перехресна ентропія (Binary Cross-Entropy, BCE) як функція втрат. BCE є спеціалізованою метрикою для задач бінарної класифікації, де необхідно передбачити імовірність належності прикладу до позитивного класу (наявність емоції) або негативного класу (відсутність емоції).

Також вона Формула:

$$\text{BCE} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)] \quad (4)$$

де:

- N — кількість спостережень у наборі даних,
- y_i — фактична мітка класу для i -го спостереження, яка може приймати значення 0 або 1,
- \hat{y}_i — передбачена моделлю ймовірність того, що i -те спостереження належить класу 1,
- \log — натуральний логарифм.

Під час навчання нейромережева модель намагається мінімізувати значення бінарної перехресної ентропії шляхом налаштування своїх вагових коефіцієнтів (параметрів) таким чином, щоб передбачені імовірності якомога точніше відповідали фактичним міткам емоцій у навчальних даних.

Зазвичай під час навчання нейронних мереж метрики вірності, точності, повноти та похибки розраховуються на валідаційній та тестовій вибірці для оцінки якості моделі та вибору оптимальних параметрів.

Для обчислення ресурсозатратності нейронної мережі, тобто кількості операцій, необхідних для її обчислення, можна використати наступну формулу:

$$R = k_p \times P \quad (5)$$

де:

- R – ресурсозатратність нейронної мережі (кількість операцій обчислення)
- P - кількість параметрів (вагових коефіцієнтів) нейронної мережі
- k_p - коефіцієнт, що враховує тип нейронної мережі.

Для повнозв'язаних, рекурентних та заснованих на основі механізму уваги нейронних мереж коефіцієнт $k_t \sim 1$ оскільки всі параметри моделі задіяні в обчислювальному процесі, де кожен нейрон бере участь у поширенні та трансформації інформаційного потоку через повний масив зв'язків між нейронами сусідніх шарів.

Для згорткових нейронних мереж (англ. Convolutional Neural Networks, CNN) коефіцієнт $k_t > 1$, оскільки згорткові шари мають інший тип зв'язків між нейронами, ніж повнозв'язані, завдяки використанню ядер згортки та потребують значно більше операцій обчислення. Значення k_t залежить від конкретної архітектури мережі та розмірів ядер згортки.

Формула дозволяє приблизно оцінити ресурсозатратність нейронної мережі, що є важливим фактором при виборі архітектури та налаштуванні параметрів моделі для забезпечення її ефективної роботи на доступних обчислювальних ресурсах.

Таким чином, сумарна ефективність РЕЗТ у тексті ($E_{ст}$) залежить від ефективності розробки системи (E_p) та ефективності самого процесу розпізнавання за допомогою навченої нейромережевої моделі ($E_{НММ}$). Цю залежність можна виразити формулою:

$$E_{ст} = f(E_p, E_{НММ}) \quad (6)$$

де f - деяка функція, що визначає вплив кожного чинника на загальну ефективність.

Ефективність розробки (E_p) залежить від термінів, обсягу та ресурсозатратності формування бази даних істинних прикладів, а також від похибок на тестовій вибірці під час налаштування моделі.

Ефективність розпізнавання НММ ($E_{НММ}$) визначається ресурсозатратністю та швидкістю під час навчання та розпізнавання текстових фрагментів, а також похибками 1-го та 2-го роду під час розпізнавання. Для підвищення сумарної ефективності $E_{ст}$ необхідно враховувати розвиток нейромережевих технологій та покращувати відповідні показники E_p та $E_{НММ}$. Це може включати використання більш ефективних архітектур нейронних мереж, оптимізацію алгоритмів навчання, збільшення обсягів та якості навчальних даних тощо. Крім того, слід враховувати матеріальні та часові ресурси, необхідні для розробки та впровадження системи РЕЗТ. Ці чинники можуть накладати обмеження на можливості підвищення ефективності та вимагати компромісів між якістю розпізнавання та вартістю реалізації проекту.

Тому, сумарна ефективність $E_{ст}$ залежить від комплексної взаємодії багатьох факторів, пов'язаних з розробкою та використанням нейромережевих моделей для РЕЗТ у текстах, а також від доступних ресурсів та врахування новітніх технологічних досягнень у цій галузі. Виходячи з цього було сформовано схему сумарної ефективності зображену на рисунку 4.

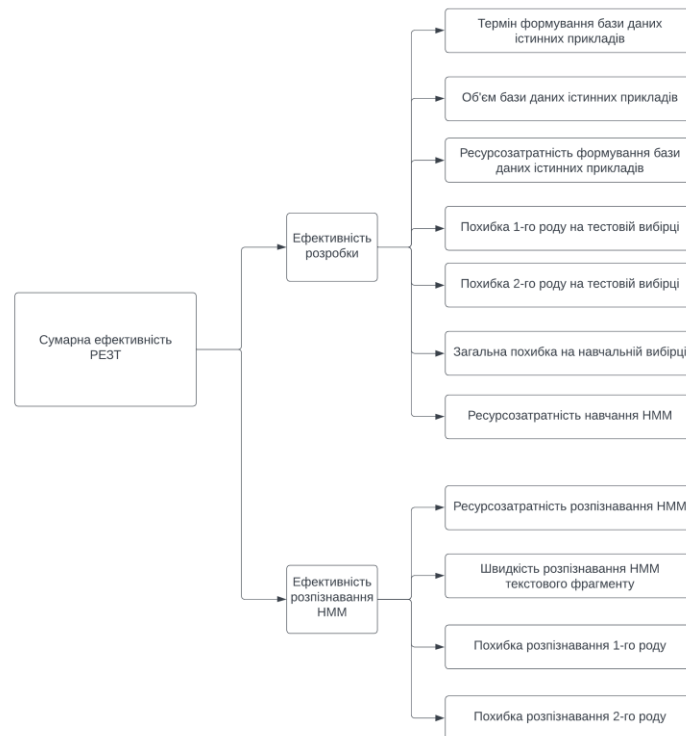


Рисунок 4. Схема сумарної ефективності розпізнавання емоційної тональності в тексті.

Висновки. РЕЗТ текстових фрагментів є важливою задачею в галузі обробки природної мови [1-3] та має значний практичний потенціал у різних сферах застосування. Формалізація процесу побудови систем РЕЗТ є необхідною для систематичного вивчення та розробки ефективних засобів автоматизованого виявлення емоційного забарвлення та тональності у текстових даних. У представленій роботі запропоновано концептуальну модель процесу РЕЗТ тексту, яка формалізує та систематизує підходи до розв'язання цієї задачі. Модель охоплює ключові етапи, такі як формування навчальних даних, вибір архітектури нейронної мережі, її налаштування та навчання, а також безпосереднє використання для РЕЗТ. Концептуальна модель визначає залежності між різними факторами, що впливають на ефективність процесу РЕЗТ, зокрема якістю вхідних даних, параметрами нейромережевої моделі, ресурсозатратністю та похибками під час навчання і розпізнавання. Це дозволяє виявити критичні аспекти, на які слід звернути увагу для підвищення загальної ефективності системи. Модель також передбачає залучення експертних знань на різних етапах для покращення якості вхідних даних та оптимізації параметрів моделі.

Список бібліографічного опису

1. I. Tereikovskiy, Z. Hu, D. Chernyshev, L. Tereikovska, O. Korystyn, and O. Tereikovskiy, "The Method of Semantic Image Segmentation Using Neural Networks," *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, vol. 14, no. 6, pp. 1-14, 2022, doi: 10.5815/ijigsp.2022.06.01.
2. S. Toliupa, Y. Kulakov, I. Tereikovskiy, O. Tereikovskiy, L. Tereikovska, and V. Nakonechniy, "Keyboard Dynamic Analysis by Alexnet Type Neural Network," in *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 2020, pp. 416-420, doi: 10.1109/TCSET49122.2020.235466.
3. I. A. Dychka, I. A. Tereikovskiy, O. S. Korovii, L. O. Tereikovska, and V. O. Romankevych, "Evaluation of the effectiveness of means for recognizing the emotional tonality of text fragments," *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, vol. 34 (73), no. 3, part 1, pp. 130-135, 2023.
4. Z. Halim, M. Waqar, and M. Tahir, "A machine learning-based investigation utilizing the in-text features for the identification of dominant emotion in an email," *Knowledge-Based Systems*, vol. 208, pp. 1-17, 2020.
5. A. Adikari, G. Gamage, D. de Silva, N. Mills, S.-M. J. Wong, and D. Alahakoon, "A self structuring artificial intelligence framework for deep emotions modelling and analysis on the social web," *Future Generation Computer Systems*, vol. 116, pp. 302-315, 2021.
6. M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279-294, 2021.
7. A. M. G. Almeida, R. Cerri, E. C. Paraiso, R. G. Mantovani, and S. Barbon Junior, "Applying multi-label techniques in emotion identification of short texts," *Neurocomputing*, vol. 320, pp. 35-46, 2018.

8. A. Bandhakavi, N. Wiratunga, D. Padmanabhan, and S. Massie, "Lexicon based feature extraction for emotion text classification," *Pattern Recognition Letters*, vol. 93, pp. 133–142, 2017, doi: 10.1016/j.patrec.2016.12.009.
9. K. Sailunaz and R. Alhajj, "Emotion and sentiment analysis from Twitter text," *Journal of Computational Science*, vol. 36, pp. 1–18, 2019.
10. M. Usama, B. Ahmad, E. Song, M. S. Hossain, M. Alrashoud, and G. Muhammad, "Attention-based sentiment analysis using convolutional and recurrent neural network," *Future Generation Computer Systems*, vol. 113, pp. 571–578, 2020.
11. J. Deng and F. Ren, "A Survey of Textual Emotion Recognition and Its Challenges," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 49–67, Jan.–Mar. 2023.

References:

1. I. Tereikovskiy, Z. Hu, D. Chernyshev, L. Tereikovska, O. Korystin, and O. Tereikovskiy, "The Method of Semantic Image Segmentation Using Neural Networks," *International Journal of Image, Graphics and Signal Processing (IJIGSP)*, vol. 14, no. 6, pp. 1–14, 2022, doi: 10.5815/ijigsp.2022.06.01.
2. S. Toliupa, Y. Kulakov, I. Tereikovskiy, O. Tereikovskiy, L. Tereikovska, and V. Nakonechnyi, "Keyboard Dynamic Analysis by Alexnet Type Neural Network," in *2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)*, Lviv-Slavske, Ukraine, 2020, pp. 416–420, doi: 10.1109/TCSET49122.2020.235466.
3. I. A. Dychka, I. A. Tereikovskiy, O. S. Korovii, L. O. Tereikovska, and V. O. Romankevych, "Evaluation of the effectiveness of means for recognizing the emotional tonality of text fragments," *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, vol. 34 (73), no. 3, part 1, pp. 130–135, 2023.
4. Z. Halim, M. Waqar, and M. Tahir, "A machine learning-based investigation utilizing the in-text features for the identification of dominant emotion in an email," *Knowledge-Based Systems*, vol. 208, pp. 1–17, 2020.
5. A. Adikari, G. Gamage, D. de Silva, N. Mills, S.-M. J. Wong, and D. Alahakoon, "A self structuring artificial intelligence framework for deep emotions modelling and analysis on the social web," *Future Generation Computer Systems*, vol. 116, pp. 302–315, 2021.
6. M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: an attention-based bidirectional CNN-RNN deep model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, 2021.
7. A. M. G. Almeida, R. Cerri, E. C. Paraiso, R. G. Mantovani, and S. Barbon Junior, "Applying multi-label techniques in emotion identification of short texts," *Neurocomputing*, vol. 320, pp. 35–46, 2018.
8. A. Bandhakavi, N. Wiratunga, D. Padmanabhan, and S. Massie, "Lexicon based feature extraction for emotion text classification," *Pattern Recognition Letters*, vol. 93, pp. 133–142, 2017, doi: 10.1016/j.patrec.2016.12.009.
9. K. Sailunaz and R. Alhajj, "Emotion and sentiment analysis from Twitter text," *Journal of Computational Science*, vol. 36, pp. 1–18, 2019.
10. M. Usama, B. Ahmad, E. Song, M. S. Hossain, M. Alrashoud, and G. Muhammad, "Attention-based sentiment analysis using convolutional and recurrent neural network," *Future Generation Computer Systems*, vol. 113, pp. 571–578, 2020.
11. J. Deng and F. Ren, "A Survey of Textual Emotion Recognition and Its Challenges," *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 49–67, Jan.–Mar. 2023.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-15>

УДК 004.347

Лукашук Микола Миколайович, аспірант

<http://orcid.org/0009-0009-5448-4315>

Київський національний університет культури і мистецтв, м. Київ, Україна

СТАНДАРТИЗАЦІЯ ВЕБ-ДОСТУПНОСТІ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ ЯК СПОСІБ ЗАЛУЧЕННЯ БІЛЬШОГО ЧИСЛА КОРИСТУВАЧІВ

Лукашук М.М. Стандартизація веб-доступності інформаційно-комунікаційних технологій як спосіб залучення більшого числа користувачів. У роботі досліджено стандартизацію веб-доступності інформаційно-комунікаційних технологій як способу залучення більшого числа користувачів. Зазначається, що під веб-доступністю розуміють інклюзивну практику, за якої веб-сайти, онлайн-інструменти та веб-технології розробляються таким чином, щоб їх могли використовувати люди з особливими потребами, у тому числі й особливо особи з обмеженими можливостями здоров'я. Наголошується, що під час організації доступного веб-контенту необхідно враховувати індивідуальні особливості, які можуть перешкоджати доступу користувача до Інтернету. Наголошується, що дизайн веб-сторінки доступний, якщо користувач може вільно оперувати веб-контентом без сторонньої допомоги, а саме сприймати матеріал і розуміти його, орієнтуватися в межах сторінки або групи зв'язаних сторінок, переміщатися за гіперпосиланнями та вносити свої дані до інтерактивних форм. Зазначається, що головним органом стандартизації веб-технологій є Міжнародний консорціум весвітнього павутиння, який ґрунтується на принципі, що доступність Інтернету для осіб з обмеженими можливостями здоров'я може бути досягнута, якщо доступні та перебувають у взаємодії такі компоненти: веб-контент (змістова частина та розмітка); користувальницькі агенти (браузери та інші веб-додатки); асистивні технології (допоміжне програмне та апаратне забезпечення); користувачі з обмеженими можливостями здоров'я, які знають адаптивні стратегії оперування веб-сторінками; веб-розробники, які мають компетенції до розробки доступних веб-сторінок; програмне забезпечення веб-розробника; інструменти оцінки доступності веб-ресурсів. Описано детально всі версії WCAG, розвиток яких продовжується безперервно. Наголошено, що кожна нова версія ініціюється W3C з метою покращення рекомендацій доступності та побудована на принципі зворотної сумісності з версією-батьком. Важливість веб-доступності додатково підкреслюється необхідністю цілісного підходу, який залучає всіх зацікавлених сторін, включаючи розробників, користувачів з обмеженими можливостями та постачальників технологій. Стандартизуючи веб-доступність, мета полягає в тому, щоб створити інклюзивне цифрове середовище, де всі користувачі, незалежно від їхніх фізичних чи когнітивних здібностей, зможуть отримати доступ і скористатися інформацією та послугами, доступними в мережі Інтернет. Це не тільки збільшує потенційну базу користувачів, але й сприяє рівним можливостям і соціальній інтеграції в епоху цифрових технологій.

Ключові слова: веб-доступність, інклюзивні технології, стандартизація ІКТ, WCAG, інтерактивна доступність.

Lukashuk M. Standardization Of Web Accessibility Of Information And Communication Technologies As A Way To Attract More Users. The paper examines the standardization of web accessibility of information and communication technologies as a way to attract a larger number of users. It is noted that web accessibility is understood as an inclusive practice in which websites, online tools and web technologies are developed in such a way that they can be used by people with special needs, including and especially people with disabilities. It is emphasized that when organizing accessible web content, it is necessary to take into account individual characteristics that may hinder the user's access to the Internet. It is emphasized that the design of a web page is accessible if the user can freely operate the web content without external assistance, namely, perceive and understand the material, navigate within a page or group of linked pages, navigate by hyperlinks, and enter data into interactive forms. It is noted that the main standardizing body for web technologies is the International World Wide Web Consortium, which is based on the principle that the accessibility of the Internet for people with disabilities can be achieved if the following components are available and interact: Web content (content part and marking); user agents (browsers and other web applications); assistive technologies (supporting software and hardware); users with disabilities who know adaptive strategies for operating web pages; web developers who have the competence to develop accessible web pages; web developer software; tools for assessing the availability of web resources. All versions of WCAG are described in detail, the development of which continues continuously. It is emphasized that each new version is initiated by the W3C in order to improve accessibility recommendations and is built on the principle of backward compatibility with the parent version. The importance of web accessibility is further underscored by the need for a holistic approach that involves all stakeholders, including developers, users with disabilities, and technology providers. By standardizing web accessibility, the goal is to create an inclusive digital environment where all users, regardless of their physical or cognitive abilities, can access and benefit from the wealth of information and services available on the internet. This not only increases the potential user base but also promotes equal opportunities and social inclusion in the digital age.

Key words: web accessibility, inclusive technologies, ICT standardization, WCAG, interactive accessibility.

Вступ. Зростаюча кількість користувачів щодня взаємодіє з веб-сайтами та мобільними додатками, різноманітність та кількість останніх зростає з геометричною прогресією, і користувачі цінують зручність використання та користувацький досвід (UX), які вони пропонують. Зростаюча потреба користувачів у мобільних програмах призвела до відкриття гібридних програм, які поєднують веб-технології та власні функції. Однак і гібридні, і нативні програми мають деякі обмеження, такі як зобов'язання щодо вільного місця на пристроях і складніші мультиплатформенні оновлення. У 2015 році Google запропонувала вирішення вищезгаданих обмежень, представивши

нову технологію під назвою прогресивні веб-програми (PWA). PWA реалізують функції, які користувач очікує від нативних програм, у мобільному браузері [3]. Вони створені та вдосконалені за допомогою сучасних інтерфейсів прикладного програмування (API), щоб забезпечити розширені можливості, надійність і можливість встановлення, охоплюючи будь-кого, будь-де та на будь-якому пристрої за допомогою єдиної кодової бази [1]. У той час, коли веб-технології стрімко розвиваються, інтерес розробників до веб-доступності згасає. Потреба людей з обмеженими можливостями в рівному доступі до нових технологій зростає. Навіть академічні та державні веб-сайти не пропонують реальної доступності, ускладнюючи таким чином життя людей з обмеженими можливостями.

Аналіз попередніх досліджень. У сучасній науковій площині з'являються роботи присвячені дослідженням інклюзивності та веб-доступності, їх впровадження у інформаційні і комунікаційні технології для підвищення якості та рівня використання.

Н. С. Тюркеджи [4] розглядає проблему створення доступного інформаційного середовища та розвитку безбар'єрного простору для людей з інвалідністю. Авторкою розкрито принципи та послідовність розробки та реалізації концептуальної моделі впровадження асистивних технологій у документноінформаційних структурах на фізичному рівні доступності та в реалізації веб-доступності цифрових ресурсів.

У [2] запропоновані і досліджуються в загальних рисах складові інформаційного середовища, зокрема, технологічного базису інформаційно-комунікаційних систем та основні процеси його вдосконалення для побудови інформаційного суспільства. Відстежується взаємодія користувач – Web-сервіс (інформаційні послуги, які надаються користувачам за допомогою Web-технологій).

Із зарубіжних авторів варто відмітити роботи таких науковців як: Крус К., Флор Е. [5], Естрада-Мартінес Ф., Хілера Х., Отон С., Агуадо-Дельгадо Х. [6], Хемблі А., Есілада Ю., Віго М., Харпер С. [7], Хортісуела Р. [8], Хусніта Л., Рахаюні А., Фусфітасарі Й., Сісванто Е., Рінтанінгрим Р. [9], Джорданоскі З., Мейерхофф Нільсен М. [10] та інших. Однак незважаючи на масштабність наукових досліджень питання актуальності даної роботи не викликає сумнівів.

Мета. Метою роботи є дослідження стандартизації веб-доступності інформаційно-комунікаційних технологій як способу залучення більшого числа користувачів.

Результати та обговорення. Веб-доступність дозволяє кожному, включаючи людей з обмеженими можливостями, сприймати, розуміти, орієнтуватися в Інтернеті та взаємодіяти з додатками. Те саме стосується мобільних програм і PWA. Веб-доступність можна визначити як спосіб створення веб-сайту, який можна навігувати та просувати різним категоріям користувачів, особливо користувачам з обмеженими можливостями, які зазвичай стикаються з перешкодами та обмеженнями під час взаємодії з Інтернетом через електронні пристрої. Веб-доступність пов'язана з практикою створення веб-сторінок, до яких мають доступ люди з будь-якими здібностями та вадами. Незважаючи на технологічний розвиток і міжнародні правила, більшість веб-сайтів і додатків залишаються недоступними для певних груп людей. Зростаюча кількість національних і міжнародних стандартів, що стосуються доступності інформаційно-комунікаційних технологій (ІКТ), включаючи Інтернет, призвела до появи багатьох різних підходів на практиці. Першим законом про доступність, прийнятим у США в 1990 році, є Закон про американців з обмеженими можливостями (ADA). Відповідно до закону веб-сайти та програми вважаються місцями розміщення інформації та мають бути доступними; ті, до яких немає доступу, вважаються такими, що дискримінують людей з обмеженими можливостями.

Існує широкий спектр людей з обмеженими можливостями, включаючи зорові, слухові, фізичні, мовленнєві, когнітивні, мовні, навчальні та неврологічні порушення, а також їх поєднання. Незважаючи на те, що було розроблено численні вказівки та інструменти, щоб допомогти покращити доступ до вмісту веб-сайтів і зрозуміти його, найбільш актуальними є W3C WCAG та ISO.

Керівництво щодо інтерфейсів користувача всесвітньої павутини містить вказівки щодо орієнтованого на людину дизайну веб-інтерфейсів користувача програмного забезпечення. ISO 9241-151 зосереджується на аспектах проектування та надає вказівки та рекомендації щодо проектування в чотирьох основних сферах:

- високорівневі дизайнерські рішення та стратегія дизайну;
- наповнення та функціональність;

- навігація, пошук і взаємодія;
- медіа дизайн та презентація.

Консорціум Всесвітньої павутини (W3C) розробив Рекомендації щодо доступності веб-контенту (WCAG), щоб зробити Інтернет доступним для людей з обмеженими можливостями. Дотримання цих вказівок робить контент доступнішим для більшого кола людей з обмеженими можливостями. Реалізація Інтернету доступним приносить користь окремим особам, підприємствам і суспільству, і метою WCAG є надання єдиного спільного стандарту для доступності веб-вмісту, який відповідає потребам окремих осіб, організацій та урядів у всьому світі.

Існує чотири різні версії WCAG: WCAG 1.0, WCAG 2.0, WCAG 2.1 і WCAG 2.2. Кожна версія розширює свою попередницю та враховує її вимоги. WCAG 1.0, створений W3C 5 травня 1999 року, є основою правил WCAG. Ця перша версія складається з 14 вказівок або загальних принципів доступного дизайну. Кожна настанова має унікальне положення, номер настанови та контрольні точки. Відповідно, контрольні точки мають унікальний оператор, номер контрольної точки та посилання на розділ документу про методи, де обговорюються реалізації та приклади контрольної точки. Щоб класифікувати контрольні пункти залежно від їх важливості, кожен контрольний пункт має рівень пріоритету, призначений робочою групою на основі впливу контрольного пункту на доступність. У цьому процесі виникло три рівні відповідності:

- Рівень відповідності «А»: усі контрольні точки Пріоритету 1 задоволені;
- Рівень відповідності «Double-A»: усі контрольні точки пріоритету 1 і 2 задоволені;
- Рівень відповідності «Triple-A»: усі контрольні точки пріоритету 1, 2 і 3 задоволені.

Незважаючи на те, що рівень відповідності було створено для визначення пріоритету кожної контрольної точки, водночас із підвищенням рівня веб-розробникам важко досягти відповідних вимог, тобто контрольні точки рівня А легше досягти, ніж потрібний рівень.

WCAG 2.0 вийшов як оновлення до WCAG 1.0 у 2008 році. Вони покращили структуру та узгодженість попередніх рекомендацій і, водночас, додали додаткові критерії успіху. Контрольно-пропускні пункти були перейменовані на критерії успіху, а пріоритет було асимільовано з Рівня А до ААА. WCAG 2.0 організовано навколо чотирьох принципів дизайну, які забезпечують основу для веб-доступності (відчутний, працездатний, зрозумілий і надійний):

Рівень А – це мінімальний рівень, який визначає основи веб-доступності. Він складається з 12 настанов, включаючи 25 критеріїв успіху.

Рівень АА включає всі вимоги рівня А та АА. Крім того, він додає 13 нових критеріїв успіху до існуючих шести настанов. Ці критерії успіху визначають основні цілі, яких мають досягти автори, щоб створити більш доступний вміст для користувачів із різними рівнями інвалідності.

Рівень ААА включає всі вимоги до рівнів А, АА та ААА. Крім того, він додає 23 нових критеріїв успіху до існуючих дев'яти керівних принципів. Ці критерії успіху дозволяють оцінити вимоги та потреби, такі як специфікації проекту, закупівлі, регулювання або договірні угоди.

Оновлена версія WCAG 2.1 вийшла в червні 2018 року. WCAG 2.1 став розширенням Версія 2.0, додавши до існуючих інструкцій п'ять критеріїв успіху А, сім АА та п'ять ААА. Найновішу версію WCAG 2.2 було представлено як робочу чернетку в серпні 2020 року та оновлено в травні 2021 року. WCAG 2.2 розширює вказівки свого попередника, додаючи такі дев'ять критеріїв успіху:

Доступна автентифікація (3.3.7; рівень А).

Якщо процес автентифікації покладається на тест когнітивних функцій, тоді має бути доступний принаймні один інший метод, який не покладається на тест когнітивних функцій. Когнітивні дефіцити зазвичай виникають у літніх людей, яким важко запам'ятовувати паролі, робити розрахунки чи навіть розгадувати головоломку. Оскільки цей критерій успіху необхідний для відповідності рівню А, повинен існувати альтернативний метод у випадку, якщо веб-сайт використовує функцію когнітивної автентифікації.

Надлишковий запис (3.3.8; рівень А).

Надлишковий запис є критерієм успіху рівня А, який вимагає автоматичного заповнення або вибору користувачем інформації в полях, які раніше вводив користувач. Однак існує виняток, якщо здійснюється повторне введення важливої інформації.

Послідовна допомога (3.2.6; рівень А).

Зіштовхнувшись із проблемами під час виконання завдання на веб-сайті, люди з певними типами обмежень можуть не впоратися з проблемою без додаткової допомоги. Проблеми можуть

включати труднощі із заповненням форми або пошуком документа чи сторінки, які надають інформацію, необхідну для виконання завдання. Тому потрібні актуальні розділи «Поширені запитання» та контактні дані людини або механізм обміну повідомленнями.

Навігація розривом сторінки (2.4.13; рівень А).

Мета цього критерію успіху полягає в тому, щоб дозволити людям, які використовують допоміжні технології або програми зчитування з екрана, знаходити посилання на вміст на основі локаторів розривів сторінок, які можна знайти в режимі перегляду за замовчуванням або в друкованій версії публікації.

Рухи перетягування (2.5.7; рівень AA).

Усі функції, які використовують рух перетягування, можуть працювати за допомогою одного вказівника без перетягування, якщо перетягування не є необхідним. Рухи перетягування спостерігаються в таких службах, як Google Maps на мобільних телефонах, де для перетягування карти потрібно більше вказівників.

Цільовий розмір (2.5.8; мінімум; рівень AA).

Цей критерій успішності належить до рівня відповідності AA та спрямований на те, щоб допомогти користувачам із тремором рук і тим, хто має труднощі з дрібною моторикою, точно активувати інтерактивні області, такі як вбудовані посилання та спливаючий вміст.

Видимі елементи керування (3.2.7; рівень AA).

Елементи керування, необхідні для просування або завершення процесу, видимі в той час, коли вони потрібні, не вимагаючи наведення вказівника або фокусування клавіатури, або доступний механізм, який робить їх видимими постійно. Такі завдання, як взаємодія з користувачем при наведенні миші, можуть ускладнити це людям із порушеннями пам'яті та іншими когнітивними розладами та вадами навчання.

Фокусний зовнішній вигляд (2.4.11; мінімум; рівень AA).

Мета цього критерію успіху полягає в тому, щоб індикатор фокусування клавіатури був чітко видимим і помітним після відповідної мінімальної площі та контрасту.

Фокусний зовнішній вигляд (2.4.12; покращений; рівень AAA).

Цей критерій успіху є розширенням попереднього, розширюючи мінімальну площу, збільшуючи контраст і виключаючи затемнені елементи.

Успіх веб-додатків залежить від того, наскільки добре вони сприймаються кінцевими користувачами. Надання доступу до веб-сайтів і програм є обов'язковим. Так, наприклад, у кожній новій технології, яку представляє Google, наголошується на необхідності створення орієнтованих на людину веб-додатків. Крім того, веб-розробники, мотивовані найкращим рейтингом пошукових систем, дотримуються деяких вказівок щодо доступності. Однак більшість створених веб-додатків не є повністю доступними, що створює проблеми з доступністю для людей з обмеженими можливостями. Хоча деякі закони зобов'язують власників веб-сайтів як у державному, так і в приватному секторах створювати веб-сайти, доступні для всіх, ці закони не застосовуються, якщо особа з обмеженими можливостями не поскаржиться на недотримання.

Попередні дослідження показали, що, незважаючи на те, що рекомендації щодо доступності закладають основу для більш доступного Інтернету, вони також мають деякі обмеження. Докази свідчать про те, що дотримання стандартів доступності не завжди гарантує задовільний UX в Інтернеті. Крім того, керівні принципи важко оцінити, навіть якщо використовується комбінація людського та машинного аудиту. У дослідженні за участю учасників з інтелектуальними вадами було виявлено, що навіть якщо веб-сайт відповідає W3C WCAG, задоволеність користувачів не є 100%.

З точки зору розробника, робота зі спеціальними можливостями є складним завданням, яке потребує багато часу для дотримання кожної з інструкцій і критеріїв успіху. Щоб реалізувати реальну доступність, кожне агентство веб-розробників має найняти спеціальну команду, яка перевірить доступність програми як за допомогою онлайн-інструментів, так і вручну. Крім того, існує проблема, яка пов'язана з тим, що відсутність обізнаності, освіти та мотивації змушує розробників створювати недоступні веб-сайти.

Онлайн-інструменти (машинний аудит) і плагіни були розроблені, щоб допомогти власникам веб-сайтів і веб-розробникам створювати більш доступні веб-сайти. За допомогою онлайн-інструментів оцінки, таких як WAVE (від WebAim), валідатор W3 (від W3C) і Lighthouse (від Google), кожен веб-сайт можна перевірити на відповідність WCAG і критеріям успіху. Ці

інструменти сканують вихідний код веб-сайту та виділяють виправлення, які необхідно внести, щоб зробити веб-сайт більш доступним. Деякі з них використовують візуальні композитори, щоб виділити області, де виникають помилки, тоді як інші звертаються до відповідного критерію успіху, щоб зрозуміти, що саме потрібно виправити. Хоча інструменти машинної оцінки пропонують корисні поради, кожен із них представляє різні помилки веб-сайту порівняно з іншими. Крім того, деякі плагіни для популярних платформ з відкритим вихідним кодом обіцяють повну доступність одним клацанням миші. Ці плагіни вносять деякі зміни у вихідний код веб-сайту, не обов'язково охоплюючи будь-які вказівки чи критерії успіху.

Таким чином, на сьогодні, потреба в доступності зростає, однак все більше веб-сайтів – навіть державних, онлайн-бібліотек і університетських веб-сайтів – не дотримуються цих інструкцій. З одного боку, хоча інструкції щодо доступності не можуть охопити кожне питання інвалідності, вони вдосконалюються з кожним днем. З іншого боку, веб-розробники мають моральний обов'язок створити веб-сайт, який є максимально доступним. Щоб досягти цього, необхідно проводити як ручний, так і машинний аудит.

В якості прикладу доступного дизайну у даному дослідженні наведено веб-ресурс який спеціалізується на продажу цифрових та аудіокниг. Веб-сайт відповідає різноманітним вимогам доступності для користувачів з вадами зору. Він використовує висококонтрастну кольорну схему з темним текстом на світлому фоні, що забезпечує читабельність для користувачів із слабким зором або дальтонізмом. Користувачі мають можливість налаштувати параметри кольору за допомогою панелі «Параметри», дозволяючи їм вибирати з кількох кольорних схем або регулювати рівні контрастності відповідно до своїх уподобань. Усі зображення продукту містять описовий альтернативний текст, надаючи користувачам програми зчитування з екрана або текстові браузері з інформацією про обкладинки книг, авторів і жанри. Елементи заголовків (<h1><h6>), списки (,) і орієнтири (<nav>, <main>, <footer>) належним чином використовуються для передачі ієрархії та організації вмісту, полегшення навігації для користувачів. Розмір тексту можна змінювати, що дозволяє користувачам регулювати розмір шрифту за допомогою налаштувань браузера без шкоди для читабельності чи цілісності макета.

Висновки. Метою веб-доступності є надання доступної веб-інформації та послуг якомога більшій кількості людей. Зростає кількість технологій, які характеризуються як доступні без фактичного дотримання вказівок щодо доступності. Існує суттєва потреба адаптувати доступність до Всесвітньої павутини, охоплюючи кожну платформу та кожну технологію. У цій статті запропоновано поглиблений огляд стандартів веб-доступності, як технологій для залучення більшої кількості користувачів.

Обізнаність веб-розробників є ключовою перевагою для досягнення оновлення Всесвітньої павутини до місця, де кожен користувач, незалежно від його/її здібностей, може мати рівний і безпроблемний доступ.

Майбутні дослідження за темою полягають у створенні інструментів, які не лише перевірятимуть веб-сайти на основі списку вказівок щодо веб-доступності, але й включатимуть пропозиції від осіб, які стикаються з проблемою виділення частини веб-сайту, яку їм важко прочитати, зрозуміти або отримати доступ.

Список бібліографічного опису

1. Болюбаш Н., Олійник М. Методи підвищення продуктивності прогресивного вебзастосування бібліотеки на основі моделі RAIL. *Інформаційні технології та суспільство*. 2023. Вип. 1 № 7. С. 13–20. DOI: 10.32689/maup.it.2023.1.2
2. Міхалевський В.Ц., Міхалевська Г. І. Розвиток інформаційного середовища для формування інтелектуальної надбудови інформаційного суспільства. *Вісник Хмельницького національного університету*. 2020. №1 С. 163–170.
3. Ткачук В. PWA, як перспективний напрямок об'єднання мобільних технологій. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. Вип. 46. 2022. С. 83–87. DOI: 10.36910/6775-2524-0560-2022-46-12
4. Тюркеджи Н. С. Асистивні технології в документно-інформаційних структурах: сервісний підхід. – Кваліфікаційна наукова праця на правах рукопису. Дисертація на здобуття наукового ступеня доктора філософії зі спеціальності 029 Інформаційна, бібліотечна та архівна справа. – *Міністерство культури та інформаційної політики України, Харківська державна академія культури*. 2023. 224 с.
5. Cruz C., Flor E. Vidal Caicedo Maria. User-centered web accessibility : recommendations for ensuring access to government information for older adults. *Ingenieria Solidaria*. 2023. № 19. P. 1–22. DOI:10.16925/2357-6014.2023.03.02.

6. Estrada-Martínez F., Hilera J., Otón S., Aguado-Delgado J. Semantic web technologies applied to software accessibility evaluation: a systematic literature review. *Universal Access in the Information Society*. 2022. №21. P.1-25. DOI:10.1007/s10209-020-00759-y.
7. Hambley A., Yesilada Y., Vigo M., Harper S. Web Structure Derived Clustering for Optimised Web Accessibility Evaluation. 2023. P. 1345-1354. DOI:10.1145/3543507.3583508.
8. Hortizuela R. Towards Web Equality: Efforts on Web Accessibility for Persons with Cognitive Disability. *International Journal of Research In Science & Engineering*. 2022. P.1-16. DOI:10.55529/ijrise.23.1.16.
9. Husnita L., Rahayuni A., Fusfitasari Y., Siswanto E., Rintaningrum R. The Role of Mobile Technology in Improving Accessibility and Quality of Learning. *al-fikrah. Jurnal Manajemen Pendidikan*. 2023. №11. P. 259. DOI:10.31958/jaf.v11i2.10548.
10. Jordanoski Z., Meyerhoff Nielsen M. The challenge of web accessibility: an evaluation of selected government websites and service portals of high, middle and low-income countries. 2023. P. 101-110. DOI:10.1145/3614321.3614343.

References

1. Boliubash N., Oliinyk M. (2023). Metody pidvyshchennia produktyvnosti prohresyvnoho vebzastosunku biblioteki na osnovi modeli RAIL. [Methods for increasing the performance of a progressive web application library based on the RAIL model]. *Informatsiini tekhnologii ta suspilstvo*. 1(7), 13-20. DOI: 10.32689/maup.it.2023.1.2 [in Ukrainian].
2. Mikhalevskiy V.Ts., Mikhalevska H.I. (2020). Rozvytok informatsiinoho seredovyscha dlia formuvannia intelektualnoi nadbudovy informatsiinoho suspilstva. [Development of the information environment for the formation of the intellectual superstructure of the information society]. *Visnyk Khmelnytskoho natsionalnogo universytetu*, 1, 163-170. [in Ukrainian].
3. Tkachuk V. (2022). PWA, yak perspektyvnyi napriamok obiednannia mobilnykh tekhnologii. [PWA as a promising direction for combining mobile technologies]. *Kompiuterno-intehrovani tekhnologii: osvita, nauka, vyrobnytstvo*, 46, 83-87. DOI: 10.36910/6775-2524-0560-2022-46-12 [in Ukrainian].
4. Tiurkedzhi N.S. (2023). Asystyvni tekhnologii v dokumentno-informatsiinykh strukturakh: servisnyi pidkhid. [Assistive technologies in document and information structures: a service approach]. PhD thesis in Information, Library, and Archival Science. Ministry of Culture and Information Policy of Ukraine, Kharkiv State Academy of Culture. 224 p. [in Ukrainian].
5. Cruz, C., Flor, E., & Vidal Caicedo, M. (2023). User-centered web accessibility: recommendations for ensuring access to government information for older adults. *Ingenieria Solidaria*, 19, 1–22. <https://doi.org/10.16925/2357-6014.2023.03.02>
6. Estrada-Martínez, F., Hilera, J., Otón, S., & Aguado-Delgado, J. (2022). Semantic web technologies applied to software accessibility evaluation: A systematic literature review. *Universal Access in the Information Society*, 21, 1–25. <https://doi.org/10.1007/s10209-020-00759-y>
7. Hambley, A., Yesilada, Y., Vigo, M., & Harper, S. (2023). Web structure derived clustering for optimised web accessibility evaluation. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1345-1354. <https://doi.org/10.1145/3543507.3583508>
8. Hortizuela, R. (2022). Towards web equality: efforts on web accessibility for persons with cognitive disability. *International Journal of Research In Science & Engineering*, 1–16. <https://doi.org/10.55529/ijrise.23.1.16>
9. Husnita, L., Rahayuni, A., Fusfitasari, Y., Siswanto, E., & Rintaningrum, R. (2023). The role of mobile technology in improving accessibility and quality of learning. *al-fikrah: Jurnal Manajemen Pendidikan*, 11(2), 259. <https://doi.org/10.31958/jaf.v11i2.10548>
10. Jordanoski, Z., & Meyerhoff Nielsen, M. (2023). The challenge of web accessibility: An evaluation of selected government websites and service portals of high, middle and low-income countries. *Proceedings of the International Conference on Human-Computer Interaction*, 101-110. <https://doi.org/10.1145/3614321.3614343>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-16>

УДК: 004.7

Маруняк Станіслав Тарасович, аспірант

<http://orcid.org/0009-0006-0635-512X>

Національний університет «Львівська Політехніка», м. Львів, Україна

ВИЯВЛЕННЯ ТА ПОМ'ЯКШЕННЯ ВРАЗЛИВОСТЕЙ БЕЗПЕКИ В ПРОТОКОЛАХ ДИНАМІЧНОЇ МАРШРУТИЗАЦІЇ: ПОТОЧНІ ВИКЛИКИ ТА РІШЕННЯ

Маруняк С.Т. виявлення та пом'якшення вразливостей безпеки в протоколах динамічної маршрутизації: поточні виклики та рішення. У сучасному світі, де цифрова інфраструктура стає все більш комплексною та взаємопов'язаною, динамічна маршрутизація відіграє ключову роль у підтримці ефективності та гнучкості мережевих систем. Метою статті є визначення актуального стану безпекових вразливостей у протоколах динамічної маршрутизації, що застосовуються в мережевих системах, та запропонувати комплексну стратегію зниження ризиків з використанням можливостей штучного інтелекту. Стаття розглядає різні протоколи динамічної маршрутизації, такі як OSPF, IS-IS, BGP, та EIGRP, їх особливості та сфери застосування в залежності від конкретних потреб та вимог до мережевого середовища. Особлива увага приділяється вибору між цими протоколами з урахуванням факторів, таких як масштабування, гнучкість управління та наявність специфічного обладнання. Акцентовано, що безпека динамічної маршрутизації залишається вразливою до різноманітних загроз через недоліки в аутентифікації, масштабуванні, розподілі навантаження та можливості фальсифікації маршрутизаційної інформації. Наголошено на значенні комплексного підходу до забезпечення безпеки, включаючи застосування сучасних методів аутентифікації, оновлення програмного забезпечення, моніторинг мережі та використання шифрування. Особливо висвітлено роль штучного інтелекту в ідентифікації та пом'якшенні загроз, що відкриває нові перспективи для підвищення безпеки мережевих інфраструктур. Підкреслено необхідність неперервного розвитку та інвестицій в кібербезпеку, а також потребу в стійкій увазі до безпеки на рівні маршрутизаційних протоколів для ефективного протистояння кіберзагрозам. Запропоноване дослідження може бути корисним фаховим дослідникам, аналітикам, бізнесу. Подальші дослідження фокусуватиметься на детальному аналізі впливу штучного інтелекту на виявлення та пом'якшення вразливостей безпеки в протоколах динамічної маршрутизації.

Ключові слова: протоколи динамічної маршрутизації; безпека мережі; вразливості мережі; заходи безпеки; штучний інтелект у кібербезпеці.

Maruniak S. Detecting and mitigating security vulnerabilities in dynamic routing protocols: current challenges and solutions. In the modern world, where digital infrastructure is becoming increasingly complex and interconnected, dynamic routing plays a key role in maintaining the efficiency and flexibility of network systems. The purpose of the article is to determine the current state of security vulnerabilities in dynamic routing protocols used in network systems and to propose a comprehensive risk mitigation strategy using artificial intelligence capabilities. The article discusses various dynamic routing protocols, such as OSPF, IS-IS, BGP, and EIGRP, their features and applications depending on specific needs and requirements for the network environment. Particular attention is paid to the choice between these protocols, taking into account factors such as scalability, management flexibility, and the availability of specific equipment. It is emphasized that the security of dynamic routing remains vulnerable to various threats due to shortcomings in authentication, scaling, load balancing, and the possibility of falsifying routing information. The importance of an integrated approach to security is emphasized, including the use of modern authentication methods, software updates, network monitoring, and encryption. The role of artificial intelligence in identifying and mitigating threats is highlighted, which opens up new prospects for improving the security of network infrastructures. The need for continuous development and investment in cybersecurity, as well as the need for sustained attention to security at the level of routing protocols to effectively counter cyber threats, is emphasized. The proposed research can be useful for professional researchers, analysts, and businesses. Further research will focus on a detailed analysis of the impact of artificial intelligence on the detection and mitigation of security vulnerabilities in dynamic routing protocols.

Keywords: dynamic routing protocols; network security; network vulnerabilities; security measures; artificial intelligence in cybersecurity.

Вступ. В останні роки, зі зростанням глобальної інформаційної мережі, важливість надійних і безпечних протоколів динамічної маршрутизації стає все більш виразною. Ці протоколи, які є ключовими компонентами мережевої інфраструктури, забезпечують визначення та оптимізацію шляхів передачі даних між вузлами в складних і динамічно змінюваних мережевих середовищах. Однак, разом з технологічним прогресом зростає і кількість потенційних загроз для цих протоколів, що робить питання їх безпеки особливо актуальним. Відповідно постає нагальна потреба аналізу сучасних викликів, пов'язаних з вразливістю безпеки в протоколах динамічної маршрутизації, і вивчення ефективних методів їх виявлення та пом'якшення. Важливо розглянути різні аспекти безпеки динамічної маршрутизації, включаючи ідентифікацію потенційних загроз, аналіз ризиків, а також розробку і впровадження механізмів захисту, що можуть ефективно протистояти виявленим вразливостям. Особлива увага приділяється новітнім дослідженням і розробкам у цій галузі, з акцентом на інноваційні підходи та технології, які обіцяють значне покращення захисту протоколів динамічної маршрутизації. Важливо проаналізувати існуючі рішення, їх переваги та недоліки, а

також висвітлити перспективні напрямки досліджень, спрямовані на забезпечення більшої стійкості мережевої інфраструктури до зовнішніх і внутрішніх загроз. У контексті стрімкого розвитку цифрових технологій, питання безпеки динамічної маршрутизації набуває особливої актуальності. Від ефективного виявлення та пом'якшення вразливостей безпеки залежить не лише стабільність роботи окремих мереж, але й безпека цілих систем, від корпоративних мереж до глобального Інтернету. Відповідно розробка та впровадження новітніх методів захисту є ключовою для забезпечення цілісності, доступності та конфіденційності інформації в сучасному цифровому світі. В силу цього важливо поглибити розуміння поточного стану безпеки протоколів динамічної маршрутизації, визначити основні виклики та вказати на можливі шляхи їх вирішення. Відповідно це допоможе ефективніше протидіяти загрозам і підвищити рівень захищеності мережевої інфраструктури.

Огляд попередніх досліджень. Проблематиці виявлення та зниження ризиків безпеки в динамічних протоколах маршрутизації присвячено ряд робіт українських і іноземних дослідників. В роботі Sakthivel T., Chandrasekaran R. [1] пропонують комбіноване рішення, що інтегрує методи шифрування та аутентифікації для захисту мережі від атак, важливою особливістю є використання фіктивних пакетів для протидії атакуючим вузлам, що дозволяє підвищити надійність передачі даних у мережі. Chen I. та ін. [2] зосереджуються на розробці системи DTN, яка може бути застосована для безпечної маршрутизації. Автори представляють модель, яка адаптується до змінних умов мережі та використовує різноманітні параметри, такі як історія взаємодії та поведінка вузлів, для визначення надійності маршрутів. Цей підхід допомагає покращити безпеку та ефективність маршрутизації в DTN, зменшуючи ризики, пов'язані з маршрутизацією через потенційно ненадійні вузли. Korir F., Cheruiyot W. [3] здійснюють аналіз безпекових викликів протоколів маршрутизації в мережах MANET. Автори аналізують різні типи атак, з якими можуть зіткнутися протоколи маршрутизації MANET, та надають огляд існуючих рішень для захисту цих мереж. Стаття важлива тим, що надає систематизований огляд викликів безпеки та способів їхнього подолання, що може слугувати основою для розробки нових, більш ефективних рішень для захисту MANET. Mohanapriya M., Krishnamurthi I. [4] зосереджуються на протоколі DSR, модифікуючи його з використанням механізмів оцінки, щоб ефективно виявляти та ізолювати зловмисні вузли. Цей підхід покликаний підвищити безпеку та надійність маршрутизації в бездротових мережах. Anupam W., Ghosh U. [5] розглядають методи забезпечення безпеки при маршрутизації та передачі даних в мобільних мережах.

Автори пропонують комплексний підхід до захисту мережі, який включає алгоритми аутентифікації, шифрування та інтеграції безпеки на рівні маршрутизації. Цей підхід дозволяє забезпечити конфіденційність, цілісність та доступність даних у мережах без центрального керування. Younes O. [6] пропонує модифіковану версію протоколу DHCP, яка націлена на протидію атакам в локальних мережах LAN. Автор аналізує поширені види атак на DHCP, такі як DoS, та пропонує методи їх нейтралізації через вдосконалення процедур аутентифікації та верифікації. Song W. та ін. [7] аналізують вразливості в дизайні маршрутизаторів процесорних інтерконектів, які можуть бути використані для проведення атак на архітектуру обчислювальних систем. Автори вказують на специфічні вразливі місця в дизайні маршрутизаторів і пропонують рекомендації щодо їх усунення або мінімізації ризиків.

Поміж українських досліджень в даній царині виділяється ряд робіт. Снігуров А., Чакрян В. [8] розглядають специфіку маршрутизації в спеціалізованих безпроводних телекомунікаційних мережах, що працюють в умовах активної інформаційної протидії. Автори пропонують підхід до управління маршрутизацією, який забезпечує підвищену надійність та безпеку передачі даних за рахунок адаптивної зміни маршрутів в залежності від поточної обстановки в мережі. Єременко О., Андрушко Д. [9] вивчають моделі маршрутизації для мереж, яка дозволяє підвищити ефективність та надійність передачі даних за рахунок використання шляхів, що перетинаються не лише в межах окремих вузлів, але й на різних рівнях мережевої архітектури. Кулаков Ю. та ін. [10] аналізують спосіб організації маршрутизації в MPLS мережах, який забезпечує підвищену безпеку передачі даних. Автори розглядають механізми багато-шляхової маршрутизації та їх застосування для захисту від потенційних загроз у безпроводних мережах. Бабенко Т. [11] зосереджується на аналізі мережевого трафіка з метою виявлення DoS атак. Автор пропонує використання ентропії трафіка як одного з індикаторів, що дозволяє вчасно виявити та відреагувати на спроби DoS атак, тим самим підвищуючи загальний рівень безпеки мережі.

Аналіз попередніх досліджень показав різноманітність підходів та стратегій, що використовуються для підвищення безпеки в бездротових мережах. З одного боку, інноваційні методи, такі як використання фіктивних пакетів, динамічне управління довірою, адаптивна маршрутизація в умовах інформаційних протидій, механізми оцінки довіри, спрямовані на ідентифікацію та ізоляцію зловмисних вузлів, демонструють широкий розвиток безпекових викликів та важливість адаптивності у захисті сучасних мережевих інфраструктур. З іншого боку, розробка та впровадження вдосконалених протоколів, які здатні протистояти специфічним атакам вказує на необхідність постійного оновлення безпекових механізмів та методів аналізу мережевого трафіку. Сукупність досліджень підкреслює значення комплексного підходу до забезпечення безпеки, що об'єднує криптографічні методи, системи довіри, аутентифікацію та алгоритми адаптивної маршрутизації. Це не лише сприяє захисту від існуючих загроз, але й створює основу для розробки стійких до майбутніх викликів мережевих систем. Такий інтегрований підхід важливий для розробки надійних та безпечних мережевих інфраструктур, здатних адаптуватися до змінних умов та витримувати нові форми кібератак. Однак, питання комплексного підходу до ефективних методів їх пом'якшення, враховуючи можливості штучного інтелекту недостатньо досліджена та потребує додаткового вивчення.

Мета статті – проаналізувати сучасний стан вразливостей безпеки в протоколах динамічної маршрутизації, які використовуються в комп'ютерних мережах і запропонувати комплексну стратегію пом'якшення з урахуванням потенціалу штучного інтелекту.

Виклад основного матеріалу. Кібербезпека перебуває в стадії стрімкого розвитку, що обумовлено постійно зростаючими загрозами в інтернет-просторі та збільшенням кількості цифрових даних. За даними Statista [12], обсяг світового ринку рішень в царині кібербезпеки, який станом на 2023 р. становив 166,2 млрд дол США, продовжуватиме зростати високими темпами й досягне 273,5 млрд дол США до 2028 р.. Сервісам кібербезпеки представляють найбільший сегмент на даному ринку з обсягом близько 88 млрд дол США на 2023 р., що складає приблизно 53% світових ринку. Найбільший дохід на ринку кібербезпеки у 2023 р. – у США з 72 млрд дол США. В даному ключі провідні позиції також займають Китай і Великобританія із показниками у 14 млрд дол США і 10 млрд дол США відповідно. Ключовими гравцями, представленими на ринку є Broadcom, Cisco, IBM, Microsoft, Palo Alto Network.

Аналіз існуючих протоколів динамічної маршрутизації відіграє ключову роль у розумінні процесів передачі даних у сучасних комп'ютерних мережах. Динамічна маршрутизація, яка дозволяє мережам автоматично пристосовуватися до змін, використовує різноманітні протоколи, кожен з яких має свої особливості та сфери застосування. Розглянемо ключові з них. Одним з найпопулярніших протоколів є OSPF, який являє собою протокол стану каналу з відкритим стандартом. Цей протокол забезпечує ефективну маршрутизацію на основі стану каналів і підтримує розділення мереж на зони для кращого масштабування. Втім, OSPF вимагає значних ресурсів обчислення та пам'яті, що може стати проблемою в великих мережах. Протокол IS-IS працює за принципом маршрутизації стану каналу, тобто найкоротший шлях до сусіднього вузла обчислюється на основі топологічної карти мережі, яку будує кожен маршрутизатор цієї мережі. Пакети IS-IS не вразливі до атак типу IP spoofing і DDoS, оскільки це протокол каналного рівня моделі OSI, а не IP-орієнтований протокол. Протокол BGP дозволяє обмін маршрутною інформацією між різними автономними системами. BGP забезпечує масштабування та гнучкість, але його налаштування та оптимізація вимагають високого рівня знань та можуть бути ускладнені ризиком неправильної конфігурації. Протокол EIGRP використовує алгоритми для ефективного визначення найкращого шляху і підтримує навантажувальне балансування. Хоча EIGRP став відкритим стандартом у 2013 році, даний протокол переважно асоціюється з обладнанням Cisco і може не бути оптимальним вибором для великих мереж. Кожен з цих протоколів сприяє розвитку динамічної маршрутизації, маючи свої специфічні переваги та обмеження. Вибір протоколу маршрутизації залежить від специфічних вимог до мережі, її розміру, а також необхідності масштабування та гнучкості. Важливо оцінити всі ці фактори в рамках забезпечення надійної та ефективної передачі даних. Зведені результати аналізу за основними характеристиками найбільш використовуваних протоколів динамічної маршрутизації подамо в таблиці 1.

Звернімо увагу, що кожен з цих протоколів має унікальні особливості та відповідає на певні вимоги мережі. Вибір між ними залежить від специфіки мережевого середовища, вимог до масштабування, гнучкості управління та наявності обладнання.

В рамках аналізу вразливості протоколів динамічної маршрутизації, ми стикаємося з низкою потенційних ризиків, що можуть відкрити мережеві системи загрозам безпеки. Ці протоколи, які лежать в основі автоматизації маршрутизації даних в комп'ютерних мережах, виявляються вразливими для різноманітних атак.

Перш за все, недостатня аутентифікація учасників у процесі маршрутизації створює можливість для несанкціонованого доступу до маршрутизаційної інформації, дозволяючи атакуючим модифікувати або перехоплювати дані. Це стає особливо небезпечним, коли атакуючі надсилають підроблені маршрутизаційні оновлення, ефективно перенаправляючи трафік через мережі під своїм контролем або створюючи маршрутизаційні петлі, що може спричинити збої у роботі мережі.

Таблиця 1. Основні характеристики найбільш використовуваних протоколів динамічної маршрутизації [13]

Характеристики	OSPF	IS-IS	BGP	EIGRP
Розробник	DEC/IETF	DEC/IETF	IETF	Cisco Systems
Тип	Відкритий стандарт	Відкритий стандарт	Відкритий стандарт	Відкритий стандарт
Метрики	Вартість шляху	Вартість шляху	Політики, префікси, атрибути	Затримка, пропускна здатність, надійність, навантаження, MTU
Масштабування	Добре	Добре	Відмінне	Добре
Основне застосування	Великі корпоративні мережі	Великі провайдерські мережі	Міждоменна маршрутизація в Інтернеті	Внутрішньої та зовнішньої маршрутизації в мережах Cisco

Проблеми з масштабуванням і розподілом навантаження також мають бути враховані в даному контексті враховувати. Велика кількість оновлень маршрутизації або навмисно створені маршрутизаційні петлі можуть перевантажити обладнання, викликаючи затримки або втрату пакетів. Недоліки в реалізації протоколів, такі як переповнення буферу в програмному забезпеченні маршрутизаторів, відкривають шлях для атак, спрямованих на виконання довільного коду або отримання несанкціонованого доступу до систем. Механізми експлуатації цих вразливостей різноманітні та включають атаки типу MitM, де атакуючі можуть перехоплювати, переглядати або змінювати передані дані. Атаки типу DoS, засновані на надсиланні великої кількості фальсифікованих маршрутизаційних повідомлень, можуть вивести з ладу маршрутизатори або знизити продуктивність мережі. Крім того, модифікація маршрутизаційної інформації дозволяє атакуючим перенаправляти чутливі дані через контрольовані ними мережі, підвищуючи ризики їх перехоплення та аналізу. Захист від таких вразливостей вимагає багаторівневого підходу до безпеки, що включає застосування сучасних методів аутентифікації, регулярне оновлення програмного забезпечення маршрутизаторів, моніторинг мережі для виявлення аномального трафіку та використання шифрування для захисту переданих даних. Такий підхід допомагає зміцнити стійкість мережі до атак і забезпечити надійну передачу даних.

Динамічна маршрутизація є ключовою для ефективності та гнучкості сучасних мережевих інфраструктур, проте безпека цих протоколів зазнає постійних випробувань через різноманітні вразливості. Розглядаючи конкретні вразливості та їх наслідки, стає можливим розробити стратегії для підвищення безпеки мереж. Виділимо основні вразливості та наслідки. По-перше, це *фальсифікація маршрутизаційної інформації*. Однією з найсерйозніших вразливостей є можливість фальсифікації маршрутизаційної інформації. Атакуючі можуть анонсувати неправильні маршрути, внаслідок чого трафік буде неналежним чином перенаправлений через атакуючого, що дозволяє здійснювати атаки MitM. Наслідки можуть включати втрату конфіденційності, цілісності даних та доступності сервісу. По-друге, це *ризик успішних DoS атак*. Механізми маршрутизації можуть бути

використані для виклику відмов в обслуговуванні, коли система або мережа стає перевантаженою або повністю недоступною для легітимних користувачів через масові фальсифіковані оновлення маршрутів. По-третє, це *використання недоліків у реалізації*. Слабкі місця у реалізації конкретних маршрутизаторів або протоколів можуть дозволити виконання довільного коду або злам обладнання, що призводить до втрати управління над мережевими компонентами.

Відповідно дані вразливості та їх наслідки потребують інструментарію для їх пом'якшення. Виділимо основні стратегії пом'якшення. Дані стратегії подано на рисунку 1.

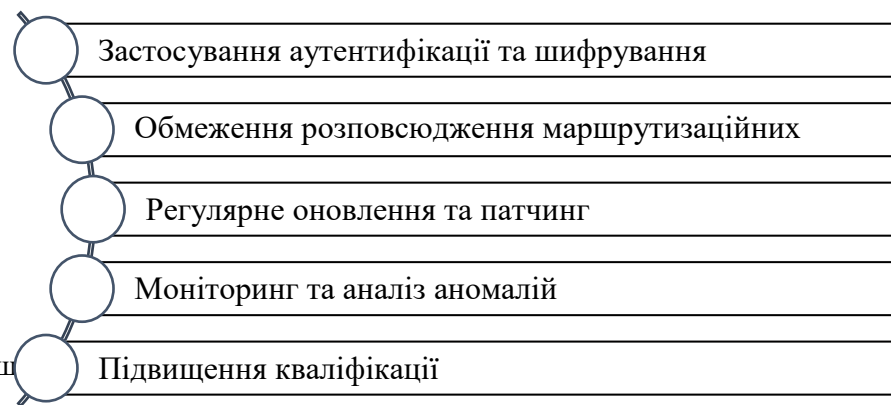


Рис. 1 – Основні стратегії пом'якшення маршрутизації [14-16]

Розглянемо в більших деталях виділені стратегії пом'якшення:

Застосування аутентифікації та шифрування: використання технологій аутентифікації та шифрування для маршрутизаційних оновлень забезпечує захист від несанкціонованого доступу та модифікації. Наприклад, протоколи BGP і OSPF з захищеними з'єднаннями, можуть значно знизити ризик фальсифікації маршрутизаційної інформації;

Обмеження розповсюдження маршрутизаційних оновлень: встановлення політик, які обмежують розповсюдження маршрутизаційних оновлень від ненадійних джерел або до критично важливих частин мережі, допомагає запобігти несанкціонованому перенаправленню трафіку;

Регулярне оновлення та патчинг: оскільки багато вразливостей виникають через недоліки в програмному забезпеченні, регулярне оновлення маршрутизаторів та іншого мережевого обладнання є критично важливим. Виробники часто випускають патчі та оновлення для виправлення виявлених вразливостей;

Моніторинг та аналіз аномалій: активний моніторинг мережевого трафіку та аналіз аномалій можуть допомогти виявити незвичайну активність, яка може вказувати на спробу експлуатації вразливостей маршрутизації. Використання систем SIEM та IDS може бути корисним у цьому контексті.

Підвищення кваліфікації: ріст обізнаності та тренінг персоналу щодо потенційних вразливостей та загроз також є важливою складовою стратегії безпеки. Знання про те, як вразливості можуть бути експлуатовані та як їх можна запобігти, допомагає зміцнити загальну безпеку мережі.

Застосування цих стратегій дозволяє зміцнити безпеку мережі, знизити ризики, пов'язані з вразливістю динамічної маршрутизації, гарантувати надійність та доступність мережесервісів. Зауважимо, що найбільшу ефективність дані стратегії матимуть за їх комплексного застосування.

В рамках даної комплексної стратегії важливо використати потенціал штучного інтелекту. Виявлення та пом'якшення вразливостей безпеки в протоколах динамічної маршрутизації за допомогою штучного інтелекту є суттєвим для забезпечення надійності та ефективності мережевої інфраструктури. Використання штучного інтелекту у цій сфері відкриває нові можливості для аналізу поведінки мережі в реальному часі, виявляючи незвичайну активність, яка може свідчити про атаки або внутрішні збої. Штучний інтелект допомагає прогнозувати потенційні вразливості, засновуючись на аналізі історичних даних, що дозволяє вжити профілактичних заходів до того, як ці слабкі місця стануть критичними. Завдяки алгоритмам машинного навчання можливо розробити системи адаптивної маршрутизації, які здатні самостійно реагувати на атаки, змінюючи маршрути трафіку для забезпечення безперервності обслуговування. Ці системи також можуть

вдосконалювати політики безпеки, аналізуючи їхню ефективність та рекомендуючи зміни, щоб адаптуватися до постійно змінюваного ландшафту кіберзагроз. Штучний інтелект також відіграє ключову роль у створенні симуляційних моделей мережі, які дозволяють проводити навчання та тестування оборонних стратегій в контрольованих умовах. Це сприяє глибшому розумінню потенційних загроз і ефективності різних підходів до захисту мережі. Впровадження штучного інтелекту у захист протоколів динамічної маршрутизації вимагає від організацій забезпечити високу точність обробки даних, адаптивність до нових загроз і гладку інтеграцію з наявними мережевими рішеннями. Важливим аспектом є знаходження балансу між автоматизацією, що її пропонує штучний інтелект, та контролем з боку людини, що покликано гарантувати, що системи діятимуть в межах визначених політик безпеки та етичних норм.

Висновки. В підсумку, динамічна маршрутизація, яка дозволяє мережам автоматично пристосовуватися до змін, використовуючи різноманітні протоколи – такі як OSPF, IS-IS, BGP, EIGRP – є ключовою для забезпечення ефективності та гнучкості сучасних мережових інфраструктур. Кожен з цих протоколів має свої особливості та сфери застосування, вибір між якими залежить від специфіки мережевого середовища, вимог до масштабування, гнучкості управління та наявності обладнання.

Однак, безпека цих протоколів зазнає постійних випробувань через різноманітні вразливості, які можуть бути експлуатовані атаками, спрямованими на недоліки в аутентифікації, масштабуванні, розподілі навантаження, а також через можливість фальсифікації маршрутизаційної інформації. Відтак, застосування комплексного підходу до безпеки, який включає застосування сучасних методів аутентифікації, регулярне оновлення програмного забезпечення, моніторинг мережі, використання шифрування та підвищення кваліфікації персоналу, стає критично важливим для забезпечення надійності та доступності мережових сервісів.

Отримані результати підкреслюють не лише важливість неперервного розвитку та інвестицій у кібербезпеку на глобальному рівні, але й необхідність постійної уваги до безпеки на рівні протоколів маршрутизації та мережових інфраструктур, щоб адекватно протистояти зростаючим кіберзагрозам. Окремим блоком даної комплексної стратегії має стати застосування потенціалу штучного інтелекту для виявлення загроз і їх пом'якшення. Подальші перспективи досліджень передбачають детальне вивчення впливу штучного інтелекту на виявлення вразливостей безпеки в протоколах динамічної маршрутизації та їх пом'якшення.

Список бібліографічного опису

1. Sakthivel T., Chandrasekaran R. A dummy Packet-Based hybrid security framework for mitigating routing misbehavior in Multi-Hop wireless networks. *Wireless Personal Communications*. 2018. №101(3). P. 1581–1618.
2. Chen I., Bao F., Chang M., Cho J. Dynamic trust management for delay tolerant networks and its application to secure routing. *IEEE Transactions on Parallel and Distributed Systems*. 2014. №25(5). P. 1200–1210.
3. Korir F., Cheruiyot W. A survey on security challenges in the current MANET routing protocols. *Global Journal of Engineering and Technology Advances*. 2022. №12(1). P. 78–91.
4. Mohanapriya M., Krishnamurthi I. Trust based DSR routing protocol for mitigating cooperative black hole attacks in ad hoc networks. *Arabian Journal for Science and Engineering*. 2013. №39(3). P. 1825–1833.
5. Alnumay W., Ghosh U. Secure routing and data transmission in mobile ad hoc networks. *International Journal of Computer Networks and Communications*. 2014. №6(1). P. 111–127.
6. Younes O. A secure DHCP protocol to mitigate LAN attacks. *Journal of Computer and Communications*. 2016. №4(1). P. 39–50.
7. Song W., Kim J., Lee J., Abts D. Security Vulnerability in Processor-Interconnect Router Design. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 2014. P. 1–11.
8. Снігуров А., Чакрян В. Підхід до управління маршрутизацією в безпроводових телекомунікаційних мережах спеціального призначення, функціонуючих в умовах інформаційної протидії. *Захист інформації і безпека інформаційних систем : II міжнародна наук.-техн. конф. : Збірник тез доповідей*. Львів, 2013. С. 16–17.
9. Єременко О., Андрушко Д. Модель маршрутизації в телекомунікаційній мережі з використанням шляхів, що перетинаються за вузлами. *Вісник Національного університету «Львівська політехніка» серія: «Радіоелектроніка та телекомунікації»*. 2015. №818. С. 181–188.
10. Кулаков Ю., Лукашенко В., Левчук А. Спосіб організації безпечної багатопляхової маршрутизації в безпроводовій мережі MPLS. *Вісник Національного Авіаційного Університету*. 2012. Т. 50. № 1. С. 101–105.
11. Бабенко Т. Дослідження ентропії мережевого трафіка як індикатора DDOS-атак. *Науковий вісник Національного гірничого університету*. 2013. №2. С. 86–89.
12. Statista. Cybersecurity: Market Data & Analysis 2023. URL: <https://www.statista.com/study/124902/cybersecurity-report/> (дата доступу: 08.04.2024).

13. Cisco. Dynamic Routing Protocols: OSPF, EIGRP, RIPv2, IS-IS, BGP. URL: <https://community.cisco.com/t5/networking-knowledge-base/dynamic-routing-protocols-ospf-eigrp-ripv2-is-is-bgp/ta-p/4511577> (дата доступу: 08.04.2024).
14. Australian Government. Australian Signals Directorate. Strategies to Mitigate Cyber Security Incidents. URL: <https://www.cyber.gov.au/sites/default/files/2023-03/PROTECT%20-%20Strategies%20to%20Mitigate%20Cyber%20Security%20Incidents%20-%20Mitigation%20Details%20%28February%202017%29.pdf> (Дата доступу: 08.04.2024).
15. Cisco. General Design Considerations for Secure Networks. URL: <https://www.ciscopress.com/articles/article.asp?p=174313&seqNum=5> (Дата доступу: 08.04.2024).
16. CSRIC. Network Reliability and Security Risk Reduction. URL: <https://www.fcc.gov/file/13925/download> (дата доступу: 08.04.2024)

References:

17. Sakthivel T., Chandrasekaran R. A dummy Packet-Based hybrid security framework for mitigating routing misbehavior in Multi-Hop wireless networks. *Wireless Personal Communications*. 2018. №101(3). P. 1581–1618.
18. Chen I., Bao F., Chang M., Cho J. Dynamic trust management for delay tolerant networks and its application to secure routing. *IEEE Transactions on Parallel and Distributed Systems*. 2014. №25(5). P. 1200–1210.
19. Korir F., Cheruiyot W. A survey on security challenges in the current MANET routing protocols. *Global Journal of Engineering and Technology Advances*. 2022. №12(1). P. 78–91.
20. Mohanapriya M., Krishnamurthi I. Trust based DSR routing protocol for mitigating cooperative black hole attacks in ad hoc networks. *Arabian Journal for Science and Engineering*. 2013. №39(3). P. 1825–1833.
21. Alnumay W., Ghosh U. Secure routing and data transmission in mobile ad hoc networks. *International Journal of Computer Networks and Communications*. 2014. №6(1). P. 111–127.
22. Younes O. A secure DHCP protocol to mitigate LAN attacks. *Journal of Computer and Communications*. 2016. №4(1). P. 39–50.
23. Song W., Kim J., Lee J., Abts D. Security Vulnerability in Processor-Interconnect Router Design. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*. 2014. P. 1–11.
24. Снігуров А., Чакрян В. Підхід до управління маршрутизацією в безпроводових телекомунікаційних мережах спеціального призначення, функціонуючих в умовах інформаційної протидії. *Захист інформації і безпека інформаційних систем : II міжнародна наук.-техн. конф. : Збірник тез доповідей*. Львів, 2013. С. 16–17.
25. Єременко О., Андрушко Д. Модель маршрутизації в телекомунікаційній мережі з використанням шляхів, що перетинаються за вузлами. *Вісник Національного університету «Львівська політехніка» серія: «Радіоелектроніка та телекомунікації»*. 2015. №818. С. 181–188.
26. Кулаков Ю., Лукашенко В., Левчук А. Спосіб організації безпечної багатошляхової маршрутизації в безпроводовій мережі MPLS. *Вісник Національного Авіаційного Університету*. 2012. Т. 50. № 1. С. 101–105.
27. Бабенко Т. Дослідження ентропії мережевого трафіка як індикатора DDOS-атак. *Науковий вісник Національного гірничого університету*. 2013. №2. С. 86–89.
28. Statista. Cybersecurity: Market Data & Analysis 2023. URL: <https://www.statista.com/study/124902/cybersecurity-report/> (дата доступу: 08.04.2024).
29. Cisco. Dynamic Routing Protocols: OSPF, EIGRP, RIPv2, IS-IS, BGP. URL: <https://community.cisco.com/t5/networking-knowledge-base/dynamic-routing-protocols-ospf-eigrp-ripv2-is-is-bgp/ta-p/4511577> (дата доступу: 08.04.2024).
30. Australian Government. Australian Signals Directorate. Strategies to Mitigate Cyber Security Incidents. URL: <https://www.cyber.gov.au/sites/default/files/2023-03/PROTECT%20-%20Strategies%20to%20Mitigate%20Cyber%20Security%20Incidents%20-%20Mitigation%20Details%20%28February%202017%29.pdf> (Дата доступу: 08.04.2024).
31. Cisco. General Design Considerations for Secure Networks. URL: <https://www.ciscopress.com/articles/article.asp?p=174313&seqNum=5> (Дата доступу: 08.04.2024).
32. CSRIC. Network Reliability and Security Risk Reduction. URL: <https://www.fcc.gov/file/13925/download> (дата доступу: 08.04.2024)

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-17>

УДК 004.75; 004.02

Марченко Олексій Олександрович, асистент

<https://orcid.org/0000-0002-5080-4811>

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

МЕТОД РОЗПАРАЛЕЛЕННЯ ПОШУКУ ПО ДЕРЕВУ МЕТОДОМ MCTS

Марченко О.О. Метод розпаралелення пошуку по дереву методом MCTS. Метод, який пропонується в цій статті, ґрунтується на двох напрямках подальшого розвитку пошуку по дереву методом Монте-Карло (MCTS): напрямку вдосконалення теоретичних підходів цього методу і напрямку розпаралелення процесу пошуку цим методом. Розпаралелення пошуку MCTS виконується на основі раніше запропонованих авторському варіанті MCTS-TSC (Monte-Carlo Tree Search with Tree Shape Control) з контролем форми дерева пошуку, узагальненій графово-структурній гетерогенній моделі динамічного розпаралелення пошуку по дереву методом Монте-Карло, а також моделі ресурсів гетерогенної розподіленої комп'ютерної системи з локальними зв'язками та її графу. Запропонований метод розпаралелення враховує всі три відомі методи апаратно-незалежного розпаралелення MCTS: кореневе, деревне і листкове розпаралелення і складається із семи етапів. Принциповою відмінністю запропонованого метода від методів інших дослідників є більш інтелектуальне прийняття рішення про вид розпаралелення та степінь розпаралелення у вершинах дерева пошуку на етапах 1-3, яке приймається як на основі наявних апаратних ресурсів, так і на основі поточної форми дерева пошуку, щоб спрямувати подальшу побудову дерева у бажаному напрямку (в глибину, чи в ширину), в той час, як іншими дослідниками, як правило, виконується тільки достатньо жорстко прив'язана до конкретної конфігурації обчислювальної системи реалізація кореневого розпаралелення між декількома комп'ютерами, реалізація деревного розпаралелення чи реалізація листового розпаралелення на платах GPU. Запропонований метод розпаралелення пошуку методом MCTS забезпечує більш ефективне завантаження наявних апаратних ресурсів і дозволяє як прискорити процес пошуку в цілому, так і підвищити ефективність пошуку методом MCTS.

Ключові слова: пошук в дереві методом Монте-Карло, MCTS, методи розпаралелення MCTS, паралельні та розподілені комп'ютерні системи, модель ресурсів комп'ютерної системи.

Marchenko O. Method of the MCTS Tree Search Method Parallelization. The method proposed in this article is based on two directions of further development of Monte Carlo tree search method (MCTS): the direction of improving the theoretical approaches of this method and the direction of parallelization of the search process by this method. Parallelization of MCTS search is performed on the basis of previously proposed author's variant MCTS-TSC (Monte-Carlo Tree Search with Tree Shape Control) with search tree shape control, a generalized graph-structural heterogeneous model of dynamic parallelization of Monte Carlo tree search, as well as a resource model of a heterogeneous distributed computer system with local connections and its graph. The proposed parallelization method takes into account all three known MCTS hardware-independent parallelization methods: root, tree, and leaf parallelization and consists of seven stages. The fundamental difference between the proposed method and the methods of other researchers is a more intelligent decision-making about the type of parallelization and the degree of parallelization at the vertices of the search tree at stages 1-3, which is made both on the basis of available hardware resources and on the basis of the current shape of the search tree in order to direct further building a tree in the desired direction (in depth or width), while other researchers, as a rule, perform only the implementation of root parallelization between several computers, which is strictly tied to a specific configuration of the computer system, the implementation of tree parallelization or implementation of leaf parallelization on GPU boards. The proposed method of MCTS search parallelization ensures more efficient loading of available hardware resources and allows both speeding up the search process as a whole and increasing the efficiency of MCTS search.

Keywords: Monte-Carlo tree search method, MCTS, MCTS parallelization methods, parallel and distributed computer systems, resource model for computer system.

Постановка наукової проблеми.

Задачі штучного інтелекту оперують дуже великими обсягами даних, обробка яких потребує значного часу, що часто перевищує ліміти часу, який є доцільним для обробки. Тому, напрями досліджень, які пропонують ефективні розв'язки таких задач є актуальними і залишаються такими надалі, оскільки обсяги інформації у світі, які потрібно оброблювати, постійно зростають.

Для розв'язку зазначених задач наразі широко використовується пошук по дереву інформації методом Монте-Карло (Monte Carlo Tree Search – MCTS) [1]. Цей метод показав себе з найкращою сторони при розв'язку багатьох надскладних задач штучного інтелекту, однією з яких є найскладніша з існуючих ігор – гра Го. Але, незважаючи на його теоретичну потужність, сам по собі цей метод не забезпечує достатньої швидкості обробки інформації, обсяг якої часто є надзвичайно великим. Тому, подальшому розвитку цього методу було присвячено багато досліджень, причому ці дослідження виконувалися як з точки зору вдосконалення його теоретичних засад, так і з точки зору розробки його варіантів, орієнтованих на певні класи задач, так і з точки зору розпаралелення роботи пошуку MCTS для прискорення процесу пошуку. По всіх цих напрямках вже були досягнуті вагомі успіхи, але й, як зазначалося вище, обсяги інформації, яку потрібно оброблювати, також зростають,

що вимагає ще більшої швидкості та ефективності обробки інформації. Тому продовження досліджень по всіх зазначених напрямках залишається актуальним, зокрема й по напрямку розпаралелення MCTS, чому й присвячена ця стаття.

Аналіз досліджень.

Метод, який пропонується в цій статті ґрунтується на двох напрямках подальшого розвитку методу пошуку по дереву MCTS: напрямку вдосконалення теоретичних підходів цього методу і напрямку розпаралелення процесу пошуку цим методом. Відомий, можна сказати вже класичний, огляд існуючих варіантів реалізації MCTS та підходів до його подальшого розвитку, як ефективних модифікацій основної формули UCS1 методу MCTS, так методів розпаралелення, зроблених в роботі [1]. Власний погляд на структуру та взаємозв'язки між підходами до підвищення ефективності пошуку MCTS був запропонований у вигляді класифікації [2]. В роботі [3] була запропонована авторська ідея контролю форми дерева пошуку під час процесу пошуку на основі критеріїв типу «глибина-ширина» DWC (Depth-Width Criteria), а також запропонований подальший розвиток методу MCTS, що базується на цій ідеї і який був названий пошуком по дереву методом Монте-Карло з контролем форми дерева (Monte-Carlo Tree Search with Tree Shape Control – MCTS-TSC) [3].

Запропонований в цій статті метод розпаралелення ґрунтується саме на варіанті MCTS-TSC, а також на раніше запропонованих узагальненій графово-структурній гетерогенній моделі динамічного розпаралелення пошуку по дереву методом Монте-Карло [4] та моделі ресурсів гетерогенної розподіленої комп'ютерної системи з локальними зв'язками (ГРКСЛЗ) та її графу [5]. Крім авторських досліджень, для розробки запропонованого методу розпаралелення є важливими дослідження, виконані в роботах [1, 6-12].

Як відомо, для апаратно-незалежного розпаралелення MCTS були запропоновані три основних методи: кореневе розпаралелення, деревне розпаралелення і листкове розпаралелення, що можуть бути застосовані на різних етапах процесу пошуку MCTS, який складається з чотирьох етапів [1]: 1) етапу вибору; 2) етапу розширення; 3) етапу моделювання; 4) етапу переобчислення.

Кореневе розпаралелення [6, 7] передбачає побудову декількох дерев пошуку MCTS від поточного кореня паралельно. Реалізація такого розпаралелення є доцільною або на окремих комп'ютерах обчислювального кластера, або на дуже потужних комп'ютерах з великою кількістю ядер.

Деревне розпаралелення полягає у паралельному виконанні кількох ітерацій процесу пошуку загальної схеми MCTS на одному й тому ж дереві пошуку [8, 9, 10]. Головною проблемою при деревному розпаралеленні є можливість виникнення конфліктів одночасного доступу до вершин дерева пошуку з різних потоків. Ця проблема розв'язується встановленням або глобального м'ютекса на всьому дереві, або локальних м'ютексів на кожній вершині дерева пошуку.

Листкове розпаралелення [9] полягає у паралельному виконанні множини ітерацій етапу моделювання загальної схеми MCTS, які виконуються від нової доданої листкової вершини дерева пошуку. Потрібно зазначити, що листкове розпаралелення може бути реалізоване ефективно як на центральних багатоядерних процесорах CPU, так і на графічних платах GPU.

У роботах [11, 12] пропонуються цікаві варіанти розпаралелення MCTS, в яких для планування розподілу окремих дій процесу пошуку MCTS між апаратними ресурсами обчислювальної системи використовується різного виду таблиці.

Зазначимо, що дослідники розпаралелення методу пошуку MCTS [6-12] спрямовували свої дослідження та реалізації на окремі аспекти розпаралелення: на реалізацію кореневого, деревного чи листкового розпаралелення на певній комп'ютерній системі, на залежності прискорення пошуку від кількості процесорів та ядер, на варіанти реалізації листкового розпаралелення на GPU тощо. Але комплексний підхід на основі теоретичних моделей в цих дослідженнях відсутній.

Метою даної роботи розробка нового методу розпаралелення MCTS, який ґрунтується на раніше запропонованих моделі динамічного розпаралелення пошуку по дереву методом Монте-Карло [4], моделі ресурсів ГРКСЛЗ та її графу [5] і подальшого розвитку методу пошуку по дереву MCTS на основі контролю форми дерева пошуку, базуючись на критеріях типу «глибина-ширина» [3], що забезпечить більш ефективний розподіл наявних апаратних ресурсів для розпаралелення пошуку методом MCTS на різних етапах його роботи та прискорить процес пошуку в цілому.

Метод розпаралелення пошуку по дереву методом MCTS

В класифікації [2], зазначено, що серед методів розпаралелення MCTS можна виділити дві

групи методів: апаратно-незалежні методи та апаратно-орієнтовані методи. Запропонований метод є гібридним, оскільки об'єднує головні підходи апаратно-незалежного розпаралелення MCTS з орієнтацією на наявні для розпаралелення в поточний момент апаратні ресурси. Зі сторони апаратно-незалежних методів розпаралелення він охоплює всі відомі методи: кореневе розпаралелення, деревне розпаралелення і листкове розпаралелення. Апаратно-залежна орієнтація цього методу полягає у тому, що він ґрунтується на узагальненій графово-структурній гетерогенній моделі динамічного розпаралелення DPM(t), яка вперше була запропонована в [4], але пізніше була вдосконалена і розширена.

Модель DPM(t) [4] визначається як

$$DPM(t) = (STM(t), RM(t), TP(t), EP(t), SP(t), BPP(t), PP(t), t),$$

де STM(t) (Search Tree Model) – модель дерева пошуку MCTS в момент часу t;

RM(t) (Resource Model) [5] – модель ресурсів ГРКСЛЗ в момент часу t;

TP(t) (Tree Policy) – політика дерева (політика вибору вершини дерева для розширення) в момент часу t;

EP(t) (Expansion Policy) – політика кроку розширення в момент часу t;

SP(t) (Simulation Policy) – політика кроку моделювання в момент часу t;

BPP(t) (BackPropogation Policy) – політика зворотного обчислення параметрів передбачення (фактично є певною формулою переобчислення) в момент часу t;

PP(t) (Parallelization Policy) – політика розпаралелення в момент часу t.

Підхід запропонованого методу полягає у співставленні моделі дерева STM(t) дерева пошуку MCTS та політик виконання кожного етапу пошуку MCTS TP(t), EP(t), SP(t), BPP(t) в момент часу t узагальненої моделі DPM(t) до графу моделі ресурсів ГРКСЛЗ RM(t) в момент часу t за допомогою політики розпаралелення PP(t) в момент часу t.

Базовий варіант застосування запропонованого методу розпаралелення показаний на рис. 1. На цьому рисунку показані перші 4 етапи методу, на яких відбувається власне процес розпаралелення.

Розглянемо запропонований метод детально.

Початковий стан. Процес пошуку MCTS в момент часу t знаходиться перед черговою ітерацією цього процесу у вершині дерева пошуку V, яка відповідає певній позиції гри (стану прикладної задачі) і від якої потрібно знайти найкращий наступний хід (найкращу подальшу дію).

Етап 1. Для кореневого вузла Node[0] (Host) 0-го рівня моделі ресурсів RM(t) в момент часу t на основі моделі піддерева STM(t) вершини V, критеріїв DWC(t), кількості та типів наявних ресурсів 1-го рівня моделі ресурсів RM(t) визначається степінь кореневого розпаралелення RPD(t) і, в залежності від типів вузлів Node[i], можливо деревного розпаралелення TPD(t) політики розпаралелення PP(t).

Після цього відбувається створення відповідної кількості паралельних потоків, які передають необхідні структури даних вузлам моделі ресурсів RM(t) наступного рівня для паралельного виконання етапів вибору та розширення процесу пошуку MCTS.

Етап 2. Для кожного вузла Node[i] 1-го рівня моделі ресурсів RM(t) в момент часу t на основі моделі піддерева STM(t) вершини V, критеріїв DWC(t), кількості та типів наявних ресурсів 2-го рівня моделі ресурсів RM(t) визначається степінь деревного розпаралелення TPD(t) і, можливо, листкового розпаралелення LPD(t) політики розпаралелення PP(t).

Далі відбувається створення відповідної кількості паралельних потоків, які на ядрах Core_i виконують різні варіанти етапів вибору та розширення дерева пошуку MCTS (тобто деревне розпаралелення) і визначення множини нових доданих вершин для виконання етапу моделювання процесу пошуку MCTS. Після цього ці потоки передають необхідні структури даних вузлам моделі ресурсів RM(t) типу GBoards для паралельного виконання етапу моделювання від нових доданих на етапі розширення вершин дерева пошуку MCTS. Крім того, частина ядер Core_i також може бути розподілена для виконання етапу моделювання.

Етап 3. Для кожного вузла типу GBoards 2-го рівня моделі ресурсів RM(t) в момент часу t на основі моделі піддерева STM(t) вершини V, критеріїв DWC(t), кількості та типів наявних ресурсів 3-го рівня моделі ресурсів RM(t) визначається степінь листкового розпаралелення LPD(t) політики розпаралелення PP(t).

Після цього для кожного ресурсу 3-го рівня типу GPU моделі ресурсів RM(t) відбувається створення відповідної кількості паралельних потоків для виконання етапу моделювання від нових

доданих на етапі розширення вершин дерева пошуку MCTS.

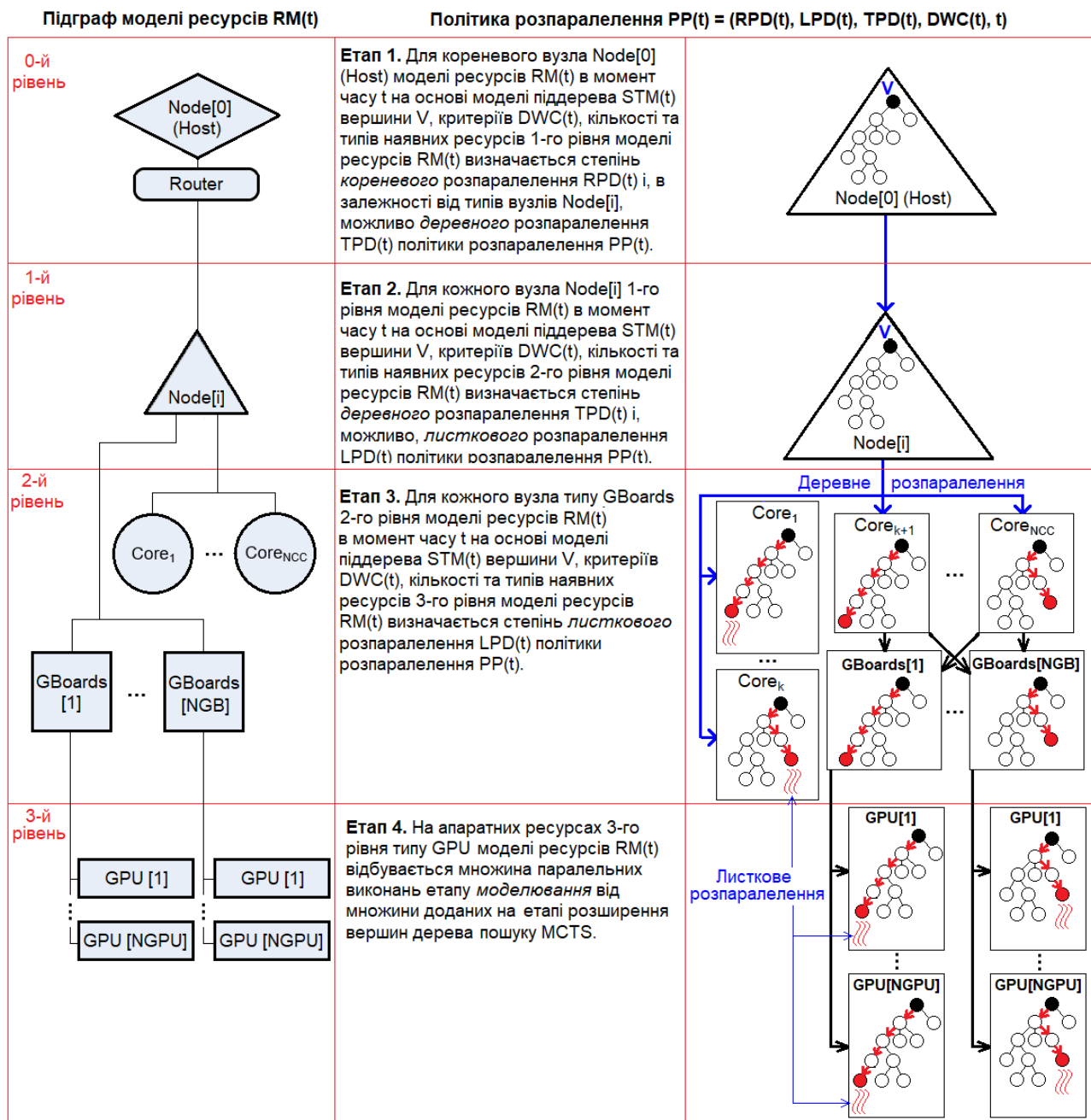


Рис.1. Метод підвищення ефективності пошуку по дереву методом MCTS розпаралеленням на основі моделі DPM(t).

Етап 4. На апаратних ресурсах 3-го рівня типу GPU моделі ресурсів RM(t) відбувається множина паралельного виконання етапу моделювання від множини доданих на етапі розширення вершин дерева пошуку MCTS.

Етап 5. На 5-му етапі методу відбувається повернення результатів виконаних моделювань з ресурсів 3-го рівня моделі ресурсів RM(t) на 2-й рівень, на якому кожен із потоків виконує етап переобчислення пошуку методом MCTS на спільному дереві пошуку так, щоб не виникало колізій з виконанням свого переобчислення іншими потоками.

Етап 6. На 6-му етапі методу відбувається повернення результуючого дерева пошуку з 2-го рівня моделі ресурсів RM(t) на 1-й, на якому це дерево об'єднується з результатами моделювання та переобчислення, виконаних на цьому рівні (це випадок, якщо на 1-му рівні було виконане як *деревне*, так і *листяне* розпаралелення). кожен із потоків виконує етап переобчислення пошуку методом

MCTS на спільному дереві пошуку так, щоб не виникало колізій з виконанням свого переобчислення іншими потоками.

Етап 7. На 7-му етапі методу відбувається повернення остаточно сформованого у вузлі Node[i] результуючого дерева пошуку з 1-го рівня моделі ресурсів $RM(t)$ на 0-й, тобто на вузол Node[0] (Host), на якому згідно з прийнятим способом оцінювання результатів пошуку в різних деревах визначається найкращий наступний хід гри (наступна дія прикладної програми) для поточної позиції (стану прикладної програми), якій відповідає вершина V дерева пошуку MCTS.

Принциповою відмінністю запропонованого метода розпаралелення MCTS від методів інших дослідників є більш інтелектуальне прийняття рішення про вид розпаралелення та степінь розпаралелення у вершинах дерева пошуку на етапах 1-3, яке приймається як на основі наявних апаратних ресурсів, так і на основі поточної форми дерева пошуку, щоб спрямувати подальшу побудову дерева у бажаному напрямку (в глибину, чи в ширину), в той час, як іншими дослідниками, як правило, виконується тільки достатньо жорстко прив'язана до конкретної конфігурації обчислювальної системи реалізація кореневого розпаралелення між декількома комп'ютерами, реалізація деревного розпаралелення чи реалізація листкового розпаралелення на платах GPU.

Висновки. На основі узагальненої графово-структурної гетерогенної моделі динамічного розпаралелення DPM(t) вперше запропоновано метод підвищення ефективності пошуку по дереву методом MCTS, який полягає у розпаралеленні процесу такого пошуку з врахуванням наявних апаратних ресурсів гетерогенної розподіленої комп'ютерної системи в кожен момент часу та використанням критеріїв виду «глибина-ширина» для прийняття рішення про вибір наступної вершини дерева пошуку.

Використання запропонованого методу динамічного розпаралелення пошуку MCTS забезпечує більш ефективне завантаження наявних апаратних ресурсів і дозволяє як досягти більш високого рівня розпаралелення та прискорити процес пошуку в цілому, так і підвищити ефективність пошуку методом MCTS.

Список бібліографічного опису

1. Cameron Browne. A Survey of Monte Carlo Tree Search Methods / Cameron Browne, Edward Powley, Daniel Whitehouse, and others // IEEE Trans. on Computational Intelligence and AI in Games. – vol. 4. – no. 1. – March 2012. – P. 1-49.
2. Марченко О.І. Класифікація способів реалізації та покращення пошуку по дереву методом Монте-Карло / Марченко О.І., Марченко О.О., Орлова М.М. // Штучний інтелект. – 2016. – №2(72). – С. 59-69.
3. Oleksandr I. Marchenko. Monte-Carlo Tree Search with Tree Shape Control. / Oleksandr I. Marchenko, Oleksii O. Marchenko // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). Conference Proceedings. May 29 – June 2, 2017., Kyiv, Ukraine. – 2017. – P. 812-8173.
4. Марченко О.О. Модель динамічного розпаралелення пошуку в дереві методом Монте-Карло для grid-систем. / Марченко О.О., Марченко О.І. // Системний аналіз та інформаційні технології: матеріали 19-ї Міжнародної науково-технічної конференції SAIT 2017, Київ, 22 – 25 травня 2017 р. / ННК "ІПСА" НТУУ "КПІ ім. Ігоря Сікорського". – К.: ННК "ІПСА" НТУУ "КПІ ім. Ігоря Сікорського", 2017., с.213-214. – Текст: укр.
5. Марченко О.О. Модель ресурсів неоднорідної розподіленої комп'ютерної системи з локальними зв'язками та її граф. / Марченко О.О., Марченко О.І. // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – 2020. – № 39. – с.83-88.
6. T. Cazenave. On the Parallelization of UCT / T. Cazenave and N. Jouandeau // Proc. Comput. Games Workshop, Amsterdam, Netherlands. – 2007. – P. 93–101.
7. Y. Soejima. Evaluating Root Parallelization in Go / Y. Soejima, A. Kishimoto, and O. Watanabe // IEEE Trans. Comp. Intell. AI Games. – vol. 2. – no. 4. – 2010. – P. 278–287.
8. E. Steinmetz. More Trees or Larger Trees: Parallelizing Monte Carlo Tree Search / E. Steinmetz, M. Gini // IEEE Transactions on Games, vol. 13, No. 3. – 2021. – P. 315–320.
9. G. M. J.-B. Chaslot. Parallel Monte-Carlo Tree Search / G. M. J.-B. Chaslot, M. H. M. Winands, and H.J. van den Herik // Proc. Comput. And Games, LNCS 5131, Beijing, China. – 2008. P.60–71.
10. A. Bourki. Scalability and Parallelization of Monte-Carlo Tree Search / A. Bourki, G. M. J.-B. Chaslot, M. Coulm, V. Danjean, H. Doghmen, J.-B. Hoock, T. Herrault, A. Rimmel, F. Teytaud, O. Teytaud, P. Vayssie`re, and Z. Yu // Proc. Int. Conf. Comput. and Games, LNCS 6515, Kanazawa, Japan. – 2010. – P. 48–58.
11. Xiufeng Yang. Practical Massively Parallel Monte-Carlo Tree Search Applied to Molecular Design / Xiufeng Yang, Tanuj Kr Aasawat, Kazuki Yoshizoe // ICLR 2021 – The Ninth International Conference on Learning Representations, May 3-7. – Access: <https://iclr.cc/virtual/2021/poster/2528>
12. L. Schaefers. Distributed Monte Carlo Tree Search: A Novel Technique and its Application to Computer Go / L. Schaefers, M. Platzner // IEEE Transactions on Computational Intelligence and AI in Games, vol. 7, No. 4. – 2015. – P. 361–374.

References

1. Cameron Browne. A Survey of Monte Carlo Tree Search Methods / Cameron Browne, Edward Powley, Daniel

- Whitehouse, and others // IEEE Trans. on Computational Intelligence and AI in Games. – vol. 4. – no. 1. – March 2012. – P. 1-49.
2. Oleksandr I. Marchenko Classification of Monte-Carlo tree search enhancement techniques oriented to specifics of the method / Oleksandr I. Marchenko, Oleksii O. Marchenko, Mariia M. Orlova // Artificial Intelligence. – 2016. – №2 (72). – P. 59-69.
 3. Oleksandr I. Marchenko. Monte-Carlo Tree Search with Tree Shape Control / Oleksandr I. Marchenko, Oleksii O. Marchenko // 2017 IEEE First Ukraine Conference on Electrical and Computer Engineering (UKRCON). Conference Proceedings. May 29 – June 2, 2017., Kyiv, Ukraine. – 2017. – P. 812-8173.
 4. Oleksii O. Marchenko. Dynamic Parallelizing Model of the Monte-Carlo Tree Search Method for grid-systems / Oleksii O. Marchenko, Oleksandr I. Marchenko // System analysis and information technology: 19-th International conference SAIT 2017, Kyiv, Ukraine, May 22 – 25, 2017. Proceedings. – ESC “IASA” NTUU “Igor Sikorsky Kyiv Polytechnic Institute”, 2017. – P. 213-214.
 5. Oleksii O. Marchenko. Resource model for heterogenous distributed computer system with local connections and its graph. / Oleksii O. Marchenko, Oleksandr I. Marchenko // Computer-Integrated Technologies: Education, Science, Production. – 2020. – № 39. – P. 83-88.
 6. T. Cazenave. On the Parallelization of UCT / T. Cazenave and N. Jouandeau // Proc. Comput. Games Workshop, Amsterdam, Netherlands. – 2007. – P. 93–101.
 7. Y. Soejima. Evaluating Root Parallelization in Go / Y. Soejima, A. Kishimoto, and O. Watanabe // IEEE Trans. Comp. Intell. AI Games. – vol. 2. – no. 4. – 2010. – P. 278–287.
 8. E. Steinmetz. More Trees or Larger Trees: Parallelizing Monte Carlo Tree Search / E. Steinmetz, M. Gini // IEEE Transactions on Games, vol. 13, No. 3. – 2021. – P. 315–320.
 9. G. M. J.-B. Chaslot. Parallel Monte-Carlo Tree Search / G. M. J.-B. Chaslot, M. H. M. Winands, and H.J. van den Herik // Proc. Comput. And Games, LNCS 5131, Beijing, China. – 2008. P.60–71.
 10. A. Bourki. Scalability and Parallelization of Monte-Carlo Tree Search / A. Bourki, G. M. J.-B. Chaslot, M. Coulm, V. Danjean, H. Doghmen, J.-B. Hoock, T. Herrault, A. Rimmel, F. Teytaud, O. Teytaud, P. Vayssiè`re, and Z. Yu // Proc. Int. Conf. Comput. and Games, LNCS 6515, Kanazawa, Japan. – 2010. – P. 48–58.
 11. Xiufeng Yang. Practical Massively Parallel Monte-Carlo Tree Search Applied to Molecular Design / Xiufeng Yang, Tanuj Kr Aasawat, Kazuki Yoshizoe // ICLR 2021 – The Ninth International Conference on Learning Representations, May 3-7. – Access: <https://iclr.cc/virtual/2021/poster/2528>
 12. L. Schaefers. Distributed Monte Carlo Tree Search: A Novel Technique and its Application to Computer Go / L. Schaefers, M. Platzner // IEEE Transactions on Computational Intelligence and AI in Games, vol. 7, No. 4. – 2015. – P. 361–374.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-18>

UDC: 004.451

Melnyk Vasyl¹, Ph.D., Associate Professor

<http://orcid.org/0000-0001-8282-6639>

Bahniuk Nataliia², Ph.D., Associate Professor

<https://orcid.org/0000-0002-7120-5455>

Roiko Oleksandr¹, Ph.D.

<https://orcid.org/0000-0001-8642-7707>

Bortnyk Kateryna², Ph.D., Associate Professor

<https://orcid.org/0000-0001-5282-099X>

Kizym Svitlana¹,

<https://orcid.org/0009-0003-1205-2086>

¹Volyn Professional College of the National University of Food Technologies, Lutsk, Ukraine

²Lutsk National Technical University, Lutsk, Ukraine

SOCKET PERFORMANCE INFLUENCE ON DATA PROCESSING INTENSITY IN A VIRTUAL MACHINE CLUSTER WITH HETEROGENEOUS CONDITIONS

Melnyk V., Bahniuk N., Roiko O., Bortnyk K., Kizym S. **Socket performance influence on data processing intensity in a virtual machine cluster with heterogeneous conditions.** The paper presents the dependences of data processing in heterogeneous conditions taking into account the impact of socket performance. As high-performance, sockets are proposed on the virtual interface architecture to use a component basis in order to support the application performance for intensive data processing. High-performance applications, using the traditional Linux-based TCP/IP interface, require guarantees of performance and its scalability to adapt for heterogeneous networks. In case of using highly effective protocols for their operation, some additional methods should be used at the user level, including for sockets on the virtual interface architecture. The limitation on the high-performance socket layers is tied with the application developing completion to full performance salvation during the TCP/IP data transfer process. There also revealed that small-block load balancing has been found to increase heterogeneous network adaptation.

Keywords: virtual interface architecture, sockets, network performance, heterogeneous conditions.

Мельник В.М., Багнюк Н.В., Ройко О.Ю., Бортник К.Я., Кізим С.О. **Вплив продуктивності сокетів на інтенсивність обробки даних на кластері віртуальних машин в гетерогенних умовах.** У роботі наведено залежності обробки даних в неоднорідних умовах з урахуванням впливу продуктивності сокетів. Як високопродуктивні сокети пропонується на архітектурі віртуального інтерфейсу використовувати компонентну основу для підтримки продуктивності програми для інтенсивної обробки даних. Високопродуктивні програми, які використовують традиційний інтерфейс TCP/IP на базі Linux, вимагають гарантій продуктивності та масштабованості для адаптації до гетерогенних мереж. У разі використання високоєфективних протоколів для їх роботи необхідно використовувати деякі додаткові методи на рівні користувача, в тому числі для сокетів на архітектурі віртуального інтерфейсу. Обмеження на рівні високопродуктивних сокетів пов'язане із завершенням розробки програми для повного збереження продуктивності під час процесу передачі даних TCP/IP. Також було виявлено, що балансування навантаження малих блоків збільшує адаптацію гетерогенної мережі.

Ключові слова: архітектура віртуального інтерфейсу, сокети, продуктивність мережі, гетерогенні умови.

Introduction. Nowadays, the research that is using high-performance data processing relies on analytical methods and their implementation for the network applications development. High-performance applications are used for clusters in technical, scientific, medical fields, in cosmonautics and others [1, 2]. Sensor technologies, for example, allow high-resolution support for multidimensional data captured by microscopes. On this basis, there appears an interest in creating the high-performance network applications for data processing, which could interactively perform basic researches with the possibility of big data generalization and analysis [3]. In their architecture, computing clusters use conventional hardware, becoming the basis of supercomputers for the operation of various applications with intensive processing of large-scale data in the form of arrays, which require some powerful support from system platforms and resources. All these features are required for efficient movement of large data between the memory cells of the computing integrated processor. To achieve the cluster high performance, all operations in data exchange and processing must perform with high precision by managing means and their internal dynamics. High-performance applications also require a high level of guarantees for productive work with the ability to adapt to the heterogeneity of work environments and variations in the support system resources.

Network applications and their performance are related to data organization and basic structures that have a component structure [4-6] and are aimed at supporting their flexible operation on distributed computing platforms. Because of the basic structures, the network application integrates a set of interaction software components arranged according to computing resources. This placement architecture directly

affects the flexibility of network applications and their productive optimization. In the parallel data processing process, there is also multicopying of data [4] around the storage cluster and computing nodes.

Another implementation method with increased productivity is a pipelining application in the computing cluster with high-speed data processing with redistribution of the data entire volume into user-defined blocks. Such processing can perform in a pipeline, taking into account their new defragmentation volume. Data delivery and processing are CPU-dependent and may overlap during operation, and the computing performance of the cluster will depend on the data blocks or thier transferred volume. In [7] was proved that small data forwarding blocks lead to an increase in the load balance and perfection of the pipeline. Due to this fact, a large number of messages should be transfered in blocks of small sizes after the defragmentation process, even smaller than the characteristic size of the payload message. However, the data delivery in the form of bulk blocks reduces the number of forwarding messages in the network, thereby increasing the bandwidth of the transmission channel with the possible manifestation of load imbalance and reduced pipeline.

Analysis of the latest research. The application on high-speed systems with intercomponent interaction in the process of exchanging messages for software applications was studied in [8,9] with the aim of creating new high-performance flowlines at the user level for exchange systems, which are aimed at reducing the parameters of the delay time and improving the throughput. However, some studies [10] aimed at the high-performance protocols standardization, which also includes the virtual interface architecture. The authors of works [11,12], where high-speed applications were focused on basic TCP or UDP protocols, used traditional socket interface. In order to support applications, there were socket connection approaches implemented at the user level with the high-performance protocols involvement. Applications written for the fundamental TCP/IP protocols for Linux kernel used normal communication. However, high-speed means own other supporting characteristics that affect the critical area of data exchange and its performance [13]. The influential parameter changes for the network application, such as the size of the defragmented blocks, provide an opportunity to obtain advantages in the productivity of working tools, increasing the speed of block processing and adaptability in a cluster system. Such actions and their application in the computing process also apply to sockets based on the virtual interface architecture in order to improve data exchange, increase the component system performance and its operation flexibility. Therefore, the data defragmentation mechanism of large blocks into sub-blocks with a system-specific volume, taking into account its component approach, leads to adaptation ensuring in the proces of data reception and transmission support [4] for applications with intensive processing integrated in the system.

The average delay time of data blocks with updating guarantee per every second under homogeneity conditions is studied in [14]. The first group of studies there concerns the improvement of frame-by-frame data processing level depending on the number of newly generated images or full updates per each second. During the request execution, the average delay time is determined for a partial image update per second, where "depending on the specified frame rate for a request with new images, the TCP socket puts a certain requirement to set the appropriate size of the data block to achieve the intended throughput". Research data obtined claim that for a fixed size of data fragmentation and the use of sockets under the virtual interface architecture there is ensured higher system performance, but from the side of sockets is a requirement to reduce the defragmentation sub-block size.

Our experimental results also confirm the performance advantage in the case of using the same sockets without computational costs presence on the processing nodes. Thus, for the case of using the TCP sockets, a higher update limit of 3.25 per processing second is not reached yet. The sockets based on the virtual interface architecture with data defragmentation into sub-blocks achieve the specified above rate without performance reducing. Received experimental results in [14] withot sub-block redistribution reveal an increase in performance ~3.5 times, and for availability of data block redistribution it goes up more than 10 times. Hance, in case of use the sockets using under the virtual interface architecture and the resizing of subblocks, the application shows significant performance improvements, scalability, and compliance with performance guarantees. It is also submitted that the application performance with data processing costs on the nodes depending from the sub-block size is linear. The values of the processing time for the partial visualization of taken image with using a digital microscopy application [15] and data defragmentation availability are on average ~18/byte of processing. The experiment data in the conditions of heterogeneity for sockets based on the virtual interface architecture with data defragmentation exceeded 3.25 updates/second due to the limitation of bandwidth. However, the experiment shows an overall performance

improvement of 3.9-4.1 times for the case of non-defragmented data and ~12 times more with partitioning into sub-blocks.

The second part of the experiment in [14] concerns the update per second with delay time guarantees where the increase in the number of full updates for each image processing second is stated at a constant value of the delay time, which is the maximum for performing a partial update request. The results of application performance growth in homogeneous conditions for a constant value of delay time and TCP protocols reveal the performance dependence on the processing data block size, while in the same conditions, the performance improvement for sockets with a virtual interface architecture is achieved. The data refragmentation at fixed bandwidth and latency results significantly higher performance for these sockets. In the absence of processing costs and a delay time of ~100 μ s, the processing performance for TCP sockets drops significantly, and for sockets with a virtual interface architecture it continues to increase up to peak values, achieving the performance improvement in almost 6 times without the participation of data defragmentation. The performance improvement of more than 8 times is observed for sockets with a virtual interface architecture during data block redistribution under homogeneity conditions. Due to the processing nodes overheads, the socket performance under the virtual interface architecture and TCP with high execution guarantee and constant latency is similar. For a processing cost value of ~18 ns/byte, the level of data processing for sockets based on the virtual interface architecture reaches some constant values, but for TCP sockets it is achieved faster. For the sockets proposed in this work, the framing level and the latency decrease are changing slightly, but the data processing performance increases by four times.

As described in [16], a high-performance socket used in a network interactive application to guarantee the productive operation of such a system in relation to the user due to the adaptability of the executable programs for productive processing in heterogeneity conditions. The results of the given data indicate that with the help of computing system reorganization with high-speed data processing, it is possible to achieve high performance of applications. Productivity growth should conduct high-speed computing guarantees and favorable stimulation of high scalability. At the same time, the involvement of an approach to data fragmentation and the structuring involvement to achieve a load balance in order to adapt additives to the environment heterogeneity, which as a result reduces resource variability, has a great impact.

Computing technology improvements, its power and memory have led to an increase to the potential for the network applications development and using in combination with the ability to store data in large volumes. To analyze the preservation of big data, the use of visualization codes and interactive visualization implemented the possibility of taking over a specific system from the inside. At the same time, the analysis and visual representation of data blocks can play a decisive role in many scientific studies. This mostly applies to programs implemented to work in an interactive high-speed way of executing requests and analyzing data blocks. The most frequently used high-performance application is the analysis of data taken from a microscope and their restoration with a high resolution [13] and digitization. Such an application uses relevant characteristics in its work, such as an instance of a video-captured study and a motivating script for data processing.

Powerful network applications should always work in batch mode with the ability to generate and process large-scale data arrays, the processing of which depends on external and internal high-speed parameters, which, for example, were studied in [13] on a cluster of virtual machines for high-speed processing. The cluster must exhibit binary compatibility for network applications with a traditional interface. With the use of the simplified communication mechanism in the system based on Xen 3.2 and the Linux kernel, an improvement of the external and internal speed parameters in the virtual machines communication was found, similar to the organization of UNIX DOMAIN communication. Bandwidth at the same time between machines using simplified communication became 2.12% greater than for traditional TCP/IP protocols, and the speed of message exchange improves gone up near by ~7.8%.

The operation of application support for searching, processing and saving digitized slides of studies made with a microscope, taking into account interactivity in the operating interval process of the microscope and behavioral displays are presented in [17]. The main difficulty that occurs in the big data processing of high-resolution digital images is proportional to the size of their volume. Such images can vary from hundreds of Mbytes to several hundreds of Gbytes, and the system and software must clearly display the current data of the working microscope with high accuracy, taking into account the continuous manipulation of its volume and frame-by-frame acquisition. The processing of such customer requests requires high-resolution projection of data on a selected resolution area, taking into account the pixel

distribution at each resolution point. In this case, the visualization server of digital microscopy should be ready to receive and store typical requests in conditions of heterogeneity that will come from the client side.

In practice, the most frequent are requests for a complete update of the aggregate image, which after processing should present a completely new image. Requests for partial typical updates to display the image in panned or zoomed views are also common. The server must handle both types of requests. The digital processing of big data by query applications that modify a set of scientific data is quite often similar to an acyclic coarse flow in the form of defined blocks from one or more sources to processing nodes of clients. It can be one or more data sets with a corresponding distribution, corresponding to the system of their storage. The interest data for such a request is removed from the defined sets and is processed in the appropriate sequence of scheduled operations on the nodes. Therefore, the data of the digital microscopy frame processing application is combined during the addition of digital sub-blocks and the corresponding viewing [4,18], and next it is sent to the client.

Stored data sub-blocks contain partial indexed images or their fragments necessary to obtain a block, where a complete image can consist of multiple blocks. A partial update request may require some defined portion of blocks, and to execute them, the byte size and length of the block affect the amount of data that is important to exchange between processing nodes.

High-performance application support. To improve the high-speed system performance, one can take into account various effective influence directions, the first of which is the intermediate refragmentation introduction of data volumes in order to establish parallelism during the execution of I/O operations and obtaining data for the request execution. Due to the block pre-calculated matching after declustering, a query started for execution can occupy the maximum number of disks involved. On the other hand, the manifestation of the effective processing power of data blocks in the system affects the working application, developed with parallelism in mind. A third effect of improving the performance of applications working interactively with defragmented data sets is their pipelining, where the total time spent on data processing can be correspondingly reduced by defragmenting the data and pipelining it for processing, which in turn provides a final mechanism, forming the final version of their integration. Therefore, it is not necessary to wait for the completion of the request, but to keep track of the block results that are partially processed. This does not reduce the full time of processing the request, but it is effective in the work of interactive applications, taking into account the shift of the playback horizon. The parameter of the sub-blocks size should take into account the network bandwidth and the delay time in the process of their network exchange. The delay time also takes into account the transmission of the full processing message using the protocols used by the current system along the full transmission path from the sender to the receivers. The number of messages to transmit the same data volume decreases with the increase in its size due to defragmentation [14], and the system bandwidth or performance plays a more important role. As the size of the sub-block increases, the processing time for each of them increases, and as the number of separate or consecutive updates of the aggregate image decreases, the system loses sensitivity, and if the size of the sub-block decreases, then the number of them increases for the transmission of the same data amount over the network. Then the message waiting time can be dominant in the efficient operation of the network application. However, smaller subblocks result in better application load balancing, and performance may suffer slightly due to connection overhead.

According to previous studies [12,14] was found that increase in the size of the transmitted data blocks leads to an improvement in the parameter of the system response time for the complete request update, which is proportional to its bandwidth. In the implementation of a partial update request, there is an increase in costs for processing sub-blocks accompanied by an increase in the data amount. Conversely, as transmission block sizes decrease, a partial update request will receive less redundant data. However, a full update request in this case will suffer losses due to reduced system bandwidth. An increase in bandwidth and a decrease in system delay time with a decrease in the size of the message causes also increase in network bandwidth [16]. For example, when working with TCP/IP kernel sockets, the reduction in system latency during the transmission of messages with fixed volume is quite noticeable and leads to an increase in the performance of its work.

The load inhomogeneity of the due to heterogeneity can be argued by the hardware environment, the components of which can be machines with different power and memory capacity, as well as the resources distribution can overlap between other working applications. In such cases, CPU and memory resources can undergo dynamic changes, and the working application must be adapted to the heterogeneity of the environment during operation, optimized for the use of shared resources and for their changing availability and accessing. Here it is required to use its own adaptation mechanisms to balance the inter-node workload according to the computing capabilities of each involved work node. For this purpose, the distribution planning of predetermined output data and calculations should be implemented on the application processing nodes. The data fragmentation may be done

so that processing and inter-node communication can be performed in pipeline mode. Data fragmentation can be performed taking into account the demand pattern, in which fast nodes could receive a larger volume of data. If such a high-speed node slows down in the data processing process due to the actions of other applications or other reasons, then the load balancing mechanism must quickly detect changes in the resources availability and/or their accessibility.

Application developers need to take into account that in the course of work, component-based tools [5, 6] have the ability to ensure the efficiency of the working environment. Such tools can be located on different resources, but the execution of the task and the flow of parallel data processing can be significantly improved due to the pipeline of their multi-copies. In our work, we also use a component infrastructure for dividing data into blocks [4] in order to support the operation of high-intensity applications in distributed environments. For this purpose, an implemented highly efficient socket interface is developed according to the virtual interface architecture for TCP/IP applications, to provide an advantage in the implementation of increased data processing performance.

The data defragmentation method into blocks. Let's describe the means structure to divide the data into blocks [14] in order to implement a software filter-flow model for applications with intensive data processing and their development. It need to submit that according to this model, the application structure is implemented with components, where each of them is a filter exchanging data using the separation of their starting flow. The filter interface includes three functions: an initialization function for allocating and initializing memory for data structures and required resources; user-defined functions for processing operations on data elements; and initialization functions for completing and freeing resources. The filters connected by unidirectional data streams from the sender of mailings to the recipient perform the functions of reading them from the incoming streams and writing them to its output streams. A data buffer, represented by an array of elements, also works for transferring them between filters. A logical flow implementation divides data into fixed-size buffers and feeds them over a TCP connection between interaction points.

Therefore, the application structure is realized with groups of interconnected filters and logical flows. At a moment, any filter group is started to process a request, before executing it the operating system makes a socket connection between the filters on the remote hosts. Filters on a single machine represent separate threads, and the request of the running application is processing as a single module through the intended filter group. Data processing by one working module is carried out in the pipeline mode, in which different filters simultaneously work with different data. Processing by a single worker module occurs when the filtering process requests the function to initialize the filter, which is together with all the necessary resources [14]. On the next, the function of processing and reading data from input streams, the function of working with data to place them in buffers, and the function of writing data to output streams come into operation. After processing finish the last data buffer, a system marker transmits to submit the end of a distinct work module. After the processing end of the current work module, the finalization function is triggering that frees the involved resources and the workspace. To process the next single working module, the mentioned interface functions should be activated again.

The program model completes several types of distribution to facilitate performance optimization by performing distribution of logical input and output streams and a transparent filter copy, which can be implemented without affecting the filtered group semantics. In this case, the vision of the work block is unchanged regarding the transparent copies and their number. The transparent filter copies also provide the ability to implement data parallelism for a single query, and multiple filter groups allow the parallelism implementation between multiple queries.

To organize communication between sender and receiver logical filters at runtime, the filter system maintains a single point-to-point logical flow that is responsible for buffering the data flow between filter transparent copies. That is, if copy P1 performs a buffer write to the logical flow for the connection from filter P to filter Q, then, for example, the data buffer will be sent in copies to Host3 or Host4. During the distribution between transparent copies, the working system supports a cyclic migration mechanism and a demand management mechanism based on the level of buffer consumption. The demand management mechanism sends buffers to the designated filter for the fastest their processing. At the beginning of the data buffer processing received from the sending filter, the consumer filter informs its message that this buffer is being processed. For sending processing buffer data, the sender filter also selects a usage filter with the minimum number of unacknowledged buffers, achieving improved load balancing.

The research purpose and tasks. The study of the data processing intensity dependence from the socket performance with taking into account the environment heterogeneity in the processing cluster aimed to use a component infrastructure with a combination of secondary sub-block division for output data streams by the

processing system to increase the performance of TCP applications in distributed environments together with a combination of a highly efficient socket interface created according to the virtual interface architecture. To achieve this goal, it is necessary to complete the following tasks:

–To implement a cluster system on personal computers using TCP-sockets and sockets based on the virtual interface architecture using a mechanism for dividing data into sub-blocks to study the heterogeneity parameters of message processing.

–Provide the performance indicators assessment for the message processing with using TCP- and virtual interface architecture sockets and combination of splitting data into sub-blocks in the research process by the mini-tests method for delay time and bandwidth parameters during direct and indirect influences on the application operation.

–to investigate the delay time average value for the processing of the data original block in heterogeneous conditions with a updating per second guarantee.

Theoretical foundations of virtual interface architecture sockets. Along with the application development based on TCP and UDP protocols, nowadays there is a rapid development of protocols with high throughput or low latency at the user level. The socket interface based on the traditional TCP and UDP protocols is widely used in network applications, but the cLAN network is a clear example of the hardware implementation for the virtual interface architecture [14]. The socket implementation in the cLAN network is of two types, one of which is based on maintaining the imitation socket, immutability of TCP/UDP and IP levels. Also, an additional layer is introduced to connect the IP layer and the kernel layer of the virtual interface. There is introduced application of a LAN emulator [8] at the socket level for use from the IP level to the virtual interface level. During the final actions of system calls including kernel context switch, TLB processing, cache filling, actions of sublevel device handlers, etc., and also introduced multicopies in this implementation, the applications with using LANE did not achieve a clear advantage in performance increasing for the basic network [14].

To provide the socket interface in the cLAN network, there is another type based to use on the user level libraries that practice primitives based on the virtual interface architecture, and takes place in our work. According to the virtual interface architecture, there is a reference to the specified socket level. This implementation for sockets based on the architecture of the virtual interface is not the main one, and the micro-test results for them are given in [14, 19]. The paper will conduct mini-test studies of performance growth depending on the delay time and bandwidth using sockets based on the virtual interface architecture. To evaluate these parameters, the direct and indirect influences on the performance of high-production socket applications in the presence of splitting the output data into blocks will also be considered. The research was conducted on a personal computers cluster (420 nodes) connected by Gigaset cLAN and Fast Ethernet networks using cLAN 1000 guest adapters and cLAN5300 cluster switches. Each node is subordinated to two Pentium III – 1 GHz processors on the Intel 840 chipset with four 32-bit 33 MHz PCI connectors, which are equipped with 512 MB SDRAM with 256 K L2 cache memory on the Linux-2.2.17 kernel.

Fig. 1 presents the theoretical and experimental dependences of the delay time for high-performance sockets, sockets based on the virtual interface architecture, and the dependence of the same parameter for traditional TCP sockets at their top on the message size [14]. To compare the received data, it can submit that the practically implemented socket level shows a low delay time value of 9.5 s., which is quite close to the theoretically calculated one for sockets based on the virtual interface architecture. From the dependencies shown in Fig. 1 is well observed that the delay time parameter increased by almost five times for traditional TCP/IP sockets to compare with high-performance ones. However, when the size of the message blocks increases from ~128 bytes with their further growth, the delay time monotonically increases sharply for all types of sockets, and it can be seen on the behavior of the dependencies shown in this figure.

Fig. 2 shows the theoretically calculated and experimentally obtained dependencies of the bandwidth achieved for high-performance sockets based on the virtual interface architecture and traditional cLAN-based one [14]. As you can see from the given comparative dependence curves, the throughput for high-performance sockets based on the virtual interface architecture reaches a peak behavior at a value of ~763 Mbytes/s, while for their theoretically calculated throughput curve, the peak occurs at a value of ~795 Mbytes/s, and for the traditional implementation of TCP – at a value of 510 Mbytes/s. Therefore, comparing the peak throughput for different sockets, it can be seen that it is improved by almost 50% for sockets with a virtual interface architecture compared to traditionally implemented TCP sockets.

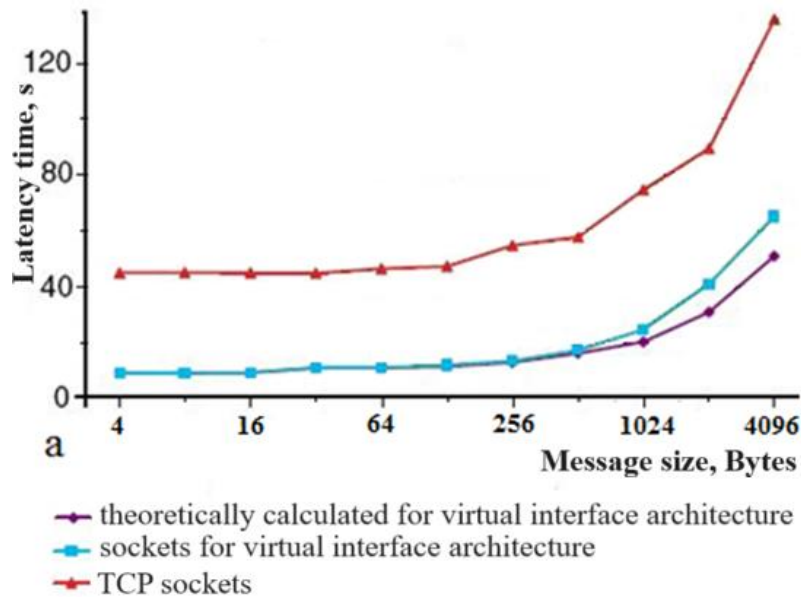


Fig. 1. Marks to compare the delay time dependence on the message size [14]

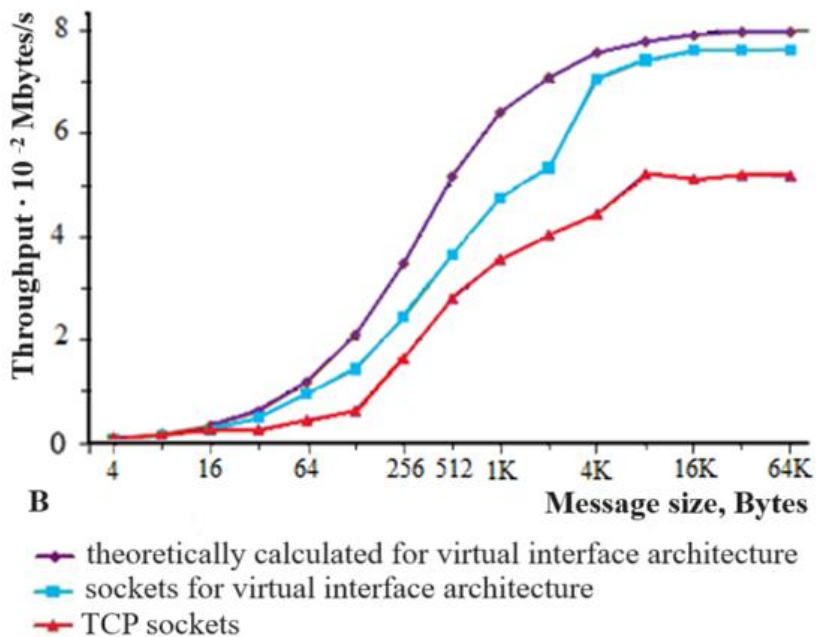


Fig. 2. Marks to compare bandwidth dependence on message size [14]

Experimental equipment. Two types of applications were used in the performance indicators study. The first type is intended for emulating the visualization server and with its help a four-stage conveyor mechanism is implemented programmatically, at the final stage of which the visualization filter is included. In order to improve throughput at the final stage of the pipeline, three copies of each filter were performed for each study. Visually, the layout of the filters is shown in Fig. 3 [13]. The image was reproduced by the user in the corresponding visualization node with the help of a filter, after which the necessary data was selected from the storage and transferred to other filters for the next work stages. Each of the following filters was placed in the system on another node included in the specified pipeline.

The image for consideration was 16 Mbytes of data for its acquisition and processing. Such predefined distribution block data will contain parts of the image, but the complete image for each distribution block will consist of several sub-blocks. To fulfill a request for a partial or full update of an integrated image, the system must deliver all the necessary blocks for their inclusion in it. It should be taken into account that each block must be delivered intact, which may require additional data initialization and their costs, including their transmission over the network to the place of unification into a single object.

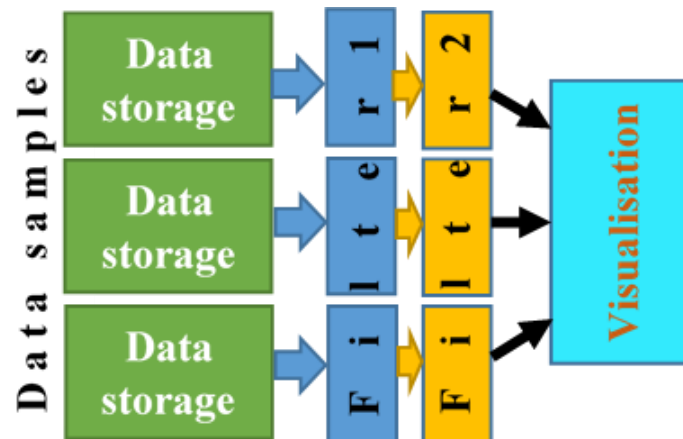


Fig. 3 – Experimental scheme [13]

During the research, two types of requests were implemented. The first one required a complete image renewing or a complete merging of all blocks involved in the composite image. This type of update from the side of practical implementation encourages the use with large block sizes and is bandwidth sensitive. Therefore, for the full update request with taking into account the corresponding speed, it is necessary to determine in advance the size of the blocks for storage and transmission, which would fulfill the user's conditions.

The completion of the second request or a partial image update takes place in the case of some image displacement rendering window or a change in its scale. For this request and its productive execution it is necessary to add only additional blocks to the existing blocks if possible in a small number compared to the full image reproduction. Productive execution of this request becomes dependent already on the delay time of additional blocks transmission. For partial recovery mapping fragments are usually extracted or narrowed, and the use of small data blocks is preferred in this case.

If the size of the blocks is relatively large in volume, then the partial update of the image will take a longer exposure time due to the modification of each block, even when a part of it is involved in the composite image. Using small block sizes during a full update will take longer due to changing a large number of blocks. If the application is configured to handle two types of mentioned requests, then it is necessary to correlate this agreement in the process of their execution. In works [13, 14] the improvement in the application performance using sockets based on the virtual interface architecture and guarantees of the requests execution for two types of updates was found, the results of which were compared with the results of the same studies on traditional TCP sockets.

Another implementation of the productive execution for the mentioned requests includes a balancing mechanism for distribution the load on the applications work due to the means of data fragmentation into blocks of appropriate sizes. For data processing by several nodes, the pipeline improvement will be achieved under the condition that the balanced load time goes to the time of data unit forwarding to the processing node and is equal to the time spent processing it on this node. Thus, taking into account the time intervals for processing and communication, a characteristic block size establishes for the application productive operation and achieving a better balance in the pipeline. In [14] there was taken into account the invariance of the computing nodes power during the application's operation time. However, due to the heterogeneity of the mentioned parameters the implementation of this approach in a dynamic heterogeneous environment is not possible.

According to the data obtained from the mentioned researches and their comparison, a high consistency for TCP/IP sockets during the pipeline implementation is achieved in case of breaking the initial data into sub-blocks of 16 Kbytes. For sockets based on the virtual interface architecture, such consistency is achieved during the refragmentation of the initial data into sub-blocks in size of 2 Kbytes. From such results follows that for TCP/IP sockets the characteristic block size is larger when for sockets with a virtual interface architecture it is smaller. During the information transmission through heterogeneous networks, the sizes of the blocks can be heterogeneous and quite large. In this case, anomalies and deviations from normal load balancing during the transfer of data blocks to a node, in which processing is slowed down, can lead to an increase in costs. Comparing the data presentation scheme below for the case of a heterogeneous cluster (Fig. 4), which was taken to analyze the impact of block heterogeneity on system performance, it is possible to predict cases of such growth.

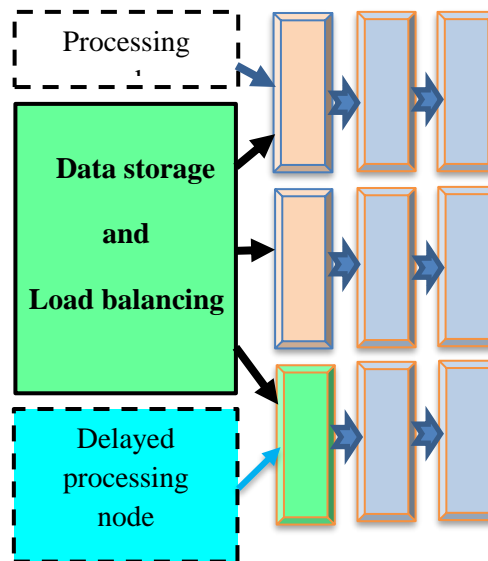


Fig. 4. Scheme of data management in heterogeneity conditions

Results and their discussion. In our conducted studies, the influence of virtual interface architecture sockets on the performance indicators of cluster computing nodes with a heterogeneous data set is observed. To gain this a slowed nodes emulation is made in the network, which, accordingly, force the data processing to be performed on other nodes more than once. For TCP protocols, the slowdown in data processing speed may take place due to a decrease in the speed of their exchange and an increase in their processing time. However, the following conditions are observed in the experiments: the communication time remains constant, and the data processing time may change in magnitude.

Also, the experiment investigates the impact of the cyclically migrating mechanism embedded in the buffer to perform data defragmentation on the performance for their processing. To compare the obtained results, the research was conducted using traditional TCP sockets and sockets based on the virtual interface architecture. The time of data transfer to the processing node under the compliance conditions with high pipeline should be the same as the time of their processing on each working node in particular. For this purpose during the experiment is monitored the balance of the load distribution on the active filters of the visualization application, or on the first nodes of the conveyor line (see Fig. 3). The data processing time on each of the filters depends linearly on the message volume and is approximately equal to 18 ns for each of its bytes. According to the obtained experimental data for traditional TCP sockets, the ideal pipeline was achieved for message sub-blocks with a size of ~16 Kbytes. In the case of the experiment for sockets based on the virtual interface architecture, the high co-verification rate was observed for message blocks with a volume of ~2 Kbytes. This proved that in the course of balancing the load and achieving high conveyance, it is necessary to analyze the influence parameters on them in more detail.

The dependence of the time amount for load balancing on nodes in conditions of heterogeneity with the growth of the network heterogeneity factor is shown in fig. 5. It can be seen from the given dependence that the value of this factor is affected by the processing speeds of the slowest and fastest processors. For the experiment with TCP, the message size is relatively large, which is ~16Kbytes, and an error occurs in the load balancing process. A block of data, which is transferring to process on the slower node leads to an increase in time costs. In general, this situation leads to an increase in load balancing time, which is an important factor for the system productive implementation. In studies with virtual interface architecture sockets, the characteristic value of the block size is smaller and the load balancing time required for its processing in a slow node, taking into account the error occurrence, is smaller to compare with TCP sockets. At the same time, the response time parameter for load balancing also decreases. Therefore, the obtained experiment results with sockets based on the virtual interface architecture show a decrease in the response time for load balancing compared to TCP sockets by approximately 8 times.

There was implemented cluster system in the experiment, which embodies the planned request management scheme (fig. 4). Based on such a system, the effect of heterogeneity on data processing speed, i.e., on the performance of the request control mechanism, is studied by performing the task of buffer planning with the division of data into blocks for TCP sockets and virtual interface architecture sockets. The result of high pipelining according to the cyclic scheduling scheme for sockets according to the virtual interface architecture was achieved by dividing the data into blocks of 2 Kbytes and for traditional TCP sockets the block size was 16 Kbytes. Fig. 6

shows dependence graphs for application execution time on the probability indicator of its slowdown. Note that the probability of slowing down is proportional to the heterogeneity parameter at the nodes. In the experiment, it was assumed that the dynamic value of the processing node should decrease over time. On the abscissa axis, as can be seen from the above graphs, was plotted the probable percentage of the slowdown. Analyzing the slowdown probability indicator equal to 20%, it can be understood that the data processing of 20% received by it is performed at a certain time with a slowdown while at the same time, the other received data processing of 80% takes place in its initial dynamics.

Figure 6 shows the dependencies for the two types of sockets described in the work with a variable heterogeneity index. From the analysis of the given dependencies it can be seen that the performance indicator of applications with traditional TCP protocols and the same coefficient of heterogeneity is quite close to applications with sockets based on the virtual interface architecture. Convergence in the magnitude of these indicators proves that, in accordance with the mechanism for managing requests to consumers and the data blocks sizes, processing nodes on less loaded processors are given the opportunity to perform a larger volume of work. At the same time, observing data pipelining ensures that it is forwarded and timed between the communication and the processing process.

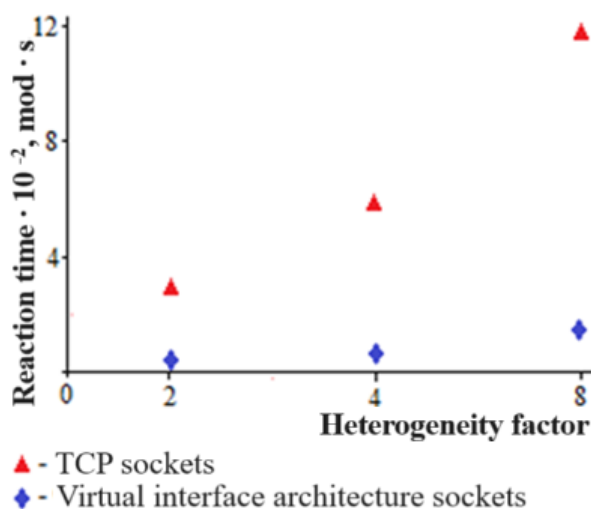


Fig. 5. The processing speed dependence during load balancing according to a cyclic planning scheme in conditions of heterogeneity

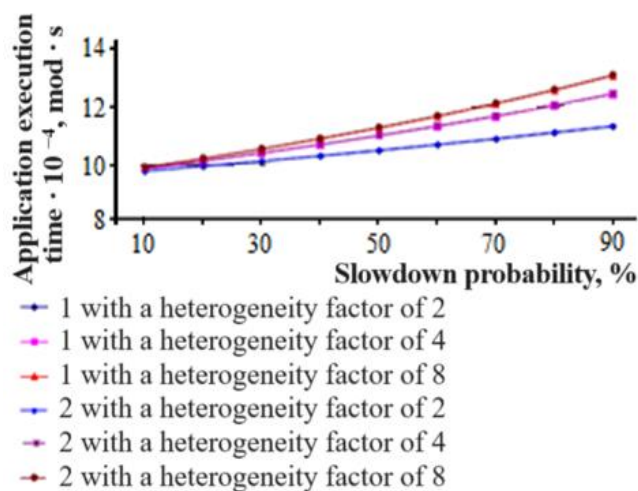


Fig. 6. The data processing speed dependence during load balancing according to the scheme of planned demand management: 1 – socket application for virtual interface architecture; 2 – application for TCP sockets

The obtained experimental results also confirm that in the case of using traditional sockets on the hardware operating configuration, the applications should be structured in such a way that it is possible to implement data

pipeline for them in accordance with their dynamic planning. Also, in order to ensure the performance of applications and the guarantees for the waiting time given by them, the use of high-performance sockets is essential.

Conclusions and research perspectives. Based on personal computers the cluster system is implemented with the involvement of traditional TCP and virtual interface architecture sockets, and the data redistribution mechanism for heterogeneous processing. According to the results of the mini-tests, an approximate estimate of the maximum data processing performance for both types of sockets is given, based on the experimentally obtained values of the delay time and bandwidth parameters and taking into account the direct and indirect effects on the applications. Without taking into account the processing costs on the nodes for sockets based on the virtual interface architecture, there is a performance improvement of more than 3.5 times, and for the same sockets with available data block redistribution it takes more than 10 times. More than in 8 times the performance gains are seen when using block redistribution actions with latency-per-second guarantees.

Experimental analysis proves that intensive applications using the interface of traditional TCP/IP sockets based on Linux require a guarantee of performance, scalability and adaptation to heterogeneous networks. For the same applications, the use of high-performance protocols, which include sockets based on the virtual interface architecture, requires the involvement of a number of methods at the user level. Highly efficient socket layers are generally limited by applications written for performance-preserving TCP/IP connections. Due to the componentization and block fragmentation to ensure the operation of applications with intensive data processing, the obtained experiment results prove the increase in productivity conditioned by reorganization of their individual components. The restructuring approach stimulates the application scalability growth with performance guarantees. An increase in the adaptation of applications to heterogeneous networks is also observed in the course of small-block load balancing.

The obtained experimental results also prove that the parametric characteristics of the sockets according to the virtual interface architecture and the use of data fragmentation on the initial processing nodes lead to an improvement in performance, and in some cases by an order of magnitude. In this case, high-performance sockets with significantly lower data computing costs allow data processing applications to achieve significantly higher quality parameters. This fact can be openly applied on modern high-performance processing clusters and in the design of intensive data processing applications on nodes.

References

1. S. Iannucci, V. Cardellini, O. D. Barba, I. Banicescu. (2020) A hybrid model-free approach for the near-optimal intrusion response control of non-stationary systems. // *Future Generation Computer Systems*. 2020, V 109, p. 111-124.
2. Nagarjuna, Setti Vidya Sagar Appaj. A real-time communication strategy for wireless sensor Networks with differentiated services. / *Industrial Engineering Journal*. 2022. Vol51, Issue 03. P.36-52. ISSN: 0970-2555.
3. Howard F. What Is Quality of Service (QoS) in Networking? 2023. E-Resource. Available at: <https://community.fs.com/article/what-is-quality-of-service-qos-in-networking.html>.
4. E.R. Rhythm, R. A. Shuvo, H. K. Mehedi, S. Hossain, A. A. Rasel. Distributed Computing for Big Data Analytics: Challenges and Opportunities. / *Computer Science and Engineering. Distributed Computing*. Preprint. 2022. 5 p. DOI:10.13140/RG.2.2.12511.53927/1. Available at: https://www.researchgate.net/publication/366466213_Distributed_Computing_for_Big_Data_Analytics_Challenges_and_Opportunities
5. F. Isaila, J. Carretero, R. Ross. CLARISSE: a middleware for data-staging coordination and control on large-scale HPC platforms. *CCGRID '16: Proceedings of the 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing*. May 2016. P. 346-355. Available at: <https://doi.org/10.1109/CCGrid.2016.24>.
6. Morgan Fluhler. Effectively Handling Large Datasets. 2023. Dataiku. E-Resource. Available at: <https://blog.dataiku.com/effectively-handling-large-datasets>.
7. V. Melnyk, K. Melnyk, S. Lavrenchuk, I. Burchak, O. Kaganiuk. Influence of the message direct search mechanism based on the TCP protocols to the exchange process. *East-European journal of Enterprise Technologies*. Kharkov (Scopus DOI:10.15587/1729-4061.2019.167995). 2019. №3(2)99. p. 36-42.
8. Durgesh Pandey. Inter Process Communication (IPC). 2023. E-Resource. Available at: <https://www.geeksforgeeks.org/inter-process-communication-ipc/>.
9. Tom Tobul. Powering Advanced Creative Workloads with Entry-Level Workstations. 2021. Dell Blog. Available at: <https://www.dell.com/en-us/blog/powering-advanced-creative-workloads-with-entry-level-workstations/>
10. Virtual Interface Architecture and Support for SAN. 2021. E-Resource. Available at: <https://learn.microsoft.com/en-us/windows-hardware/drivers/network/virtual-interface-architecture-and-support-for-san>.
11. C. Shim, R. Shinde, M. Choi. Compatibility Enhancement and Performance Measurement for Socket Interface with PCIe Interconnections. / *Hum. Cent. Comput. Inf. Science*. 2019. Doi: <https://doi.org/10.1186/s13673-019-0170-0>.
12. Shah H.V., Pu C., Madukkarumukumana R.S. High Performance Sockets and RPC over Virtual Interface (VI) Architecture. In the Proceedings of CANPC workshop (held in conjunction with HPCA Conference), p. 91-107, 1999.
13. Melnyk V., Melnyk K., Kuzmych O., Bahniuk N., Kraviez O. Internal and external communication performance improving research on a cluster of communicated virtual machines. / *Scientific journal "CIT: education, science, production"*. 2020, N.39. P. 162-174.

14. Melnyk V., Kaganiuk O., Kozlenko M. et al. Data processing intensity dependence in the cluster on socket performance avoiding heterogeneity. / Scientific journal "CIT: education, science, production" 2020, № 40. P. 128-139.
15. M. Hortsch, N.K. Koney, A.M. Oommen, D.G. Yohannan, Y. Li, A.C. Rocha, V.C. Girao-Carmona. Virtual Microscopy Goes Global: The Images are Virtual and the Problems are Real. Advances in Experimental Medicine and Biology. / Part in the book series Biomedical Visualisation. 2023. Vol. 1421. P. 79-124.
16. Melnyk V., Bahniuk N., Melnyk K. Influence of High Performance Sockets on Data Processing Intensity. // Scientific Journal "ScienceRise". Kharkiv. 2015. V.6, № 2(11). P. 38-48.
17. R.C.N. Melo, M.W.D. Raas, C. Palazzi, V.H. Neves, K.K. Malta, T.P.Silva. Whole Slide Imaging and Its Applications to Histopathological Studies of Liver Disorders. / Frontiers in Medicine. 2020. 13 P. doi: 10.3389/fmed.2019.00310. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6960181/pdf/fmed-06-00310.pdf>.
18. M. Abughazala, H. Muccini, M. Sharaf. Architecture Description Framework For Data-Intensive Applications. / Fourth International Conference on Intelligent Data Science Technologies and Applications (Conference Paper, 8 pages, October 24-26, 2023). ResearchGate. DOI:10.1109/IDSTA58916.2023.10317869. Available at: <https://ieeexplore.ieee.org/document/10317869/metrics#metrics>.
19. J. Rexford, I. Stoica. Socksdirect: datacenter sockets can be fast and compatible. / SIGCOMM '19: Proceedings of the ACM Special Interest Group on Data Communication (August 2019). P. 90–103. Available at: <https://doi.org/10.1145/3341302.3342071>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-19>

УДК 004.42

Мороз Борис Іванович, д.т.н., професор

<https://orcid.org/0000-0002-5625-0864>

Круглик Андрій Сергійович, аспірант

Національний технічний університет «Дніпровська політехніка», м. Дніпро, Україна

КОНЦЕПТУАЛЬНА МОДЕЛЬ ОБРОБКИ ІНФОРМАЦІЙНИХ ПОТОКІВ В СИСТЕМІ БЕЗПЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

Мороз Б.І., Круглик А.С. Концептуальна модель обробки інформаційних потоків в системі безпілотних літальних апаратів. У межах статті описано концептуальну модель організації обробки інформаційних потоків у системах безпілотних (літальних) апаратів що можуть використовуватись у межах трансформації ЗСУ. Концептуальні модель не є вичерпною і може бути доповнена чи змінена з появою нових технічних рішень. Підкреслюється, що впровадження безпілотних систем у збройні сили є одним із пріоритетних задач трансформації в умовах ведення довгострокових військових дій для забезпечення мінімізації втрат особового складу. Представлено базові сценарії можливого використання і комбінування безпілотних засобів для досягнення мети – зменшення втрат особового складу. Оглядом досліджено існуючі технічні рішення для використання у безпілотних літальних апаратах, наземних роботизованих системах, морських дронах і т. д. Продемонстровано зміну інтенсивності повідомлень окремого типу а також зміну цінності конкретного повідомлення в залежності від оперативної ситуації. Увага у статті приділяється не лише використанню існуючих рішень, а, в першу чергу, розробку вітчизняних програмних і технічних рішень для можливості їх безперервної модифікації і вдосконалення за рахунок появи нових рішень і оптимізації існуючих.

Ключові слова: клієнт-серверна архітектура, інтенсивність повідомлень, цінність повідомлення, БПЛА, мікроконтролер, мікрокомп'ютер.

Moroz B., Kruhlyk A. Conceptual model of information flow processing in the system of unmanned aerial vehicles.

The article describes a conceptual model of the organization of information flow processing in unmanned (aircraft) systems that can be used within transformation process of the Armed Forces. The conceptual model is not exhaustive and may be supplemented or changed when new technical solutions appear. It is emphasized that implementation of unmanned systems in the armed forces is one of the priority tasks of transformation in the conditions of conducting long-term military operations to ensure the minimization of personnel losses. Basic scenarios of the possible use and combination of unmanned vehicles to achieve the goal of reducing personnel losses are presented. The existing technical solutions for use in unmanned aerial vehicles, ground robotic systems, marine drones, etc. have been reviewed. The change in the intensity of messages of a particular type, as well as the change in the value of a specific message depending on the operational situation, has been demonstrated. The article pays attention not only to the use of existing solutions, but, first of all, to the development of domestic software and technical solutions for the possibility of their continuous modification and improvement due to the appearance of new solutions and optimization of existing ones.

Keywords: client-server architecture, message intensity, message value, UAV, microcontroller, microcomputer.

Постановка проблеми. Два роки російсько-української війни показали що питання впровадження і використання безпілотних дронів на сьогоднішній день є дуже актуальним і, можливо, навіть змінює правила ведення бою. У роботі [1] автори вказали, що провідні країни світу ведуть активну розробку і впровадження безпілотних (літальних) апаратів у військові формування. Ведеться розробка не лише літальних апаратів, але й також наземних роботизованих систем. Як було сказано раніше, провідні країни світу вже мають дорожні карти інтегрування безпілотних літальних апаратів у свої збройні сили.

Літальні апарати змінили умови та правила ведення бою, і це стало очевидним не лише для України, а й для більшості розвинутих країн світу. Україно-російська війна стала стимулом для розвитку безпілотних літальних систем у всьому світі, іншими словами, пройшла революція в середовищі розробки таких систем. Автори вважають що це тільки перші кроки у сфері розробки безпілотних літальних апаратів (БПЛА), і на жаль, усі наступні війни будуть вестися із масовим використанням безпілотних наземних, морських, літальних систем. Саме тому автори цілком закономірно впевнені у доцільності описати концептуальну модель використання безпілотних літальних а також наземних роботизованих систем для впровадження їх у військову сферу.

Концептуально автори бачать таку систему як набагато ширшу аніж використання дронів спостереження чи frv дронів. Частково концепцію було показано в роботі [1]. Автори пропонують розглянути таку модель, де буде зберігатись чітка взаємодія між безпілотними системами і окремими літальними апаратами, наземними системами та ін., що забезпечить, з поміж іншого, можливість злагодженої взаємодії між підрозділами, частинами, з'єднаннями і т. д.

Аналіз останніх досліджень і публікацій. Розвиток літальних та безпілотних літальних апаратів в Україні поки що є хаотичним. Задача сьогодення - створення дешевих first person view (fpv) дронів або дронів спостереження для вирішення нагальних проблем. Це цілком зрозуміло і логічно, тому що в умовах ведення активних бойових дій неможливо планування концепції на роки. Але автори закономірно припускають, що всі майбутні війни будуть вестися літальними безпілотними або наземними системами. Тому Україна повинна розробляти концепцію безпілотних систем і впроваджувати її в збройні сили.

Як можна побачити з джерел [2-3] США вже запустили в стратосферу дрон спостереження, який, за інформацією розробника, має можливість вести цілодобове спостереження, забезпечувати зв'язок, охороняти кордони, вести розвідку. Крім того, вказаний вище дрон є не чутливим до ворожих РЕБ і не вразливим до ППО. З поміж іншого, як заявляють розробники, такий дрон може бути безперервно у повітрі до 2-років, а отже є повністю автономним.

Іншим вагомим прикладом хотілося б вказати розробку США [4], де морськими піхотинцями під час навчань було випробувано безпілотну систему, що стріляє ракетами по кораблям. Ці навчання пройшли в рамках модернізації морської піхоти США, так як в країні вважають, що в гіпотетичному військовому конфлікті будуть використовуватись саме безпілотні технології. Це ще раз підкреслює важливість впровадження безпілотних систем у межах трансформації ЗСУ. На сьогоднішній день автори не знайшли робіт, де було б описано концепції розробок і впровадження безпілотних систем у ЗСУ.

Постановка завдання. Метою роботи є опис концептуальної моделі способів взаємодії між центрами управління і прийняття рішень та окремими безпілотними апаратами чи системами.

Викладення основного матеріалу дослідження. З появою інноваційних рішень у сфері БПЛА цілком можливо змінити умови і методи проведення оборонних та наступальних операцій. Разом із цим потрібно розуміти, середовище театру воєнних дій є дуже динамічним, а отже, вектор проведення певної операції може бути швидко змінений. Іншими словами, можлива стрімка зміна інтенсивності інформаційних потоків певного типу а значить, система повинна бути готовою до таких змін і мати змогу обробляти інформаційні потоки в часових межах, визначених для такого типу повідомлень.

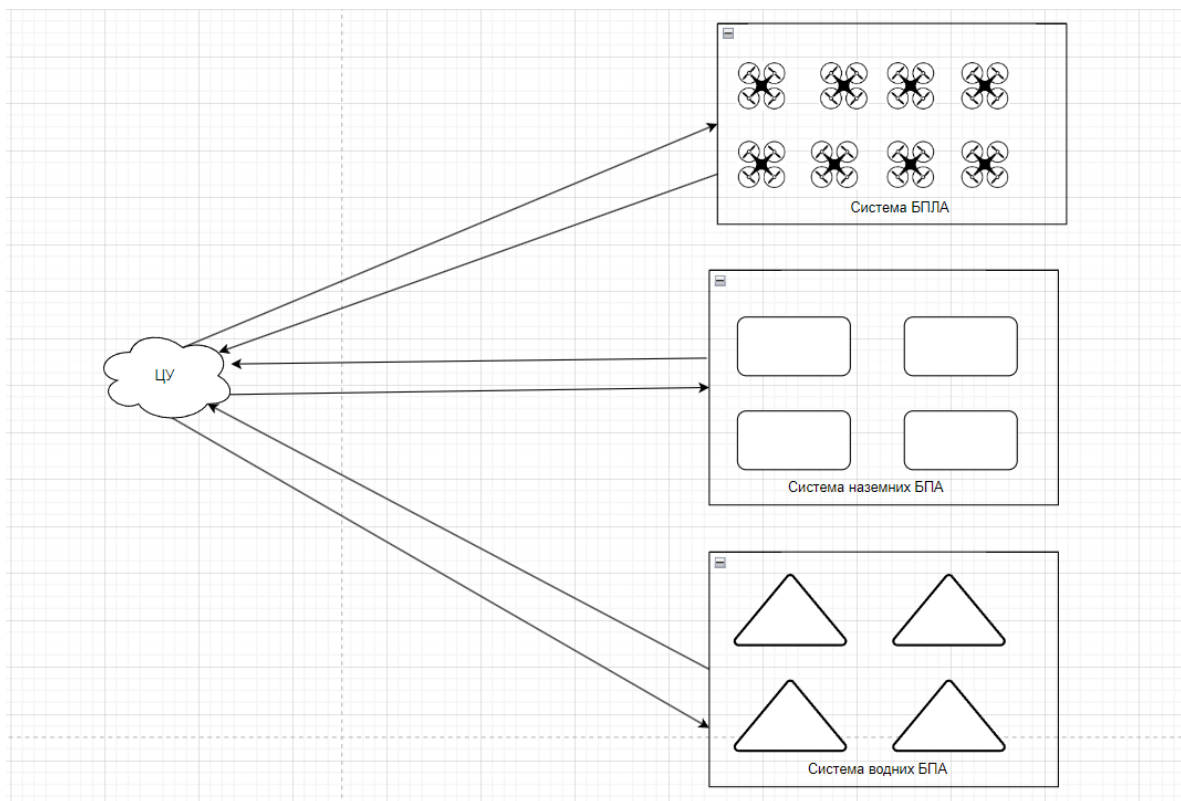


Рис. 1. Схематичне зображення взаємодії між ЦУ та системами безпілотних (літальних) апаратів.

В загальному вигляді система управління безпілотними (літальними) апаратами може бути представлена наступним чином (див. рис. 1.) З рисунку видно, окремі безпілотні апарати можуть відноситися до певної групи – літальні апарати, наземні роботизовані системи, морські дрони і т. ін. В свою чергу кожна з груп може поділятися на підгрупи. Кожна підгрупа може бути представлена як окрема система зі своїм центром управління. З розвитком безпілотних технологій підгрупи можуть бути знову поділені чи об'єднані. Головним елементом будь-якої системи є центр управління, який у свою чергу поділяється на декілька окремих взаємопов'язаних модулів. По-перше це неформальні оператори, які аналізують потоки повідомлень і приймають безпосередні рішення. Іншим важливим елементом центру управління є набір програм, модулів, функцій, критичних секцій і т. ін. для забезпечення керування як окремими безпілотними апаратами, так і комплексами та окремими системами є модуль керування. Цей модуль повинен забезпечувати максимальну автономність дій. Тобто, за можливості за допомогою заданих алгоритмів приймати певні рішення без необхідності втручання з боку оператора. Обробляти помилки, аналізувати, сповіщати операторів і т. ін. Саме на цей модуль лягає обов'язок попереднього прийому, аналізу, очистки вхідних інформаційних потоків з боку зовнішніх систем.

Сфера задач, яку безпосередньо можуть виконувати безпілотні (літальні) апарати, наземні комплекси, морські і т. д. є майже невичерпною з розвитком і вдосконаленням технологій. Це і спостереження, нанесення уражень, керування іншими системами, мінування та розмінування, транспортування поранених з важко прохідних місць, доставка боєкомплекту і т. д. Як бачимо цей список не є вичерпним і показує лише найближчу перспективу розвитку. Також хотілося б додати, що один і той же безпілотний засіб може виконувати декілька видів операцій. Це обов'язково потрібно враховувати при розробці модулів управління. Наприклад, один літальний апарат, що призначений для транспортування боєкомплекту міг би при необхідності, транспортувати пораненого з поля бою.

Керування подібною концептуальною системою здається цілком логічним і правильним за допомогою клієнт-серверної архітектури, де клієнтом є окремий літальний апарат або наземний комплекс, або система літальних апаратів. Сервером у даному випадку виступає блок управління. Який приймає вхідні заявки, аналізує, оброблює їх і керує роботою того чи іншого безпілотного апарату.

Розглянемо концептуальну модель використання дронів спостереження у парі з роботою наземних роботизованих комплексів типу безпілотник-кулемет [5], див. рис. 2. З рисунку видно, що ЛА знаходяться в русі і спостерігають за оперативною обстановкою у зоні відповідальності. Для виконання завдання спостереження і контролю за обстановкою можливо використання навіть звичайного дрону невійськового призначення. В даній ситуації ЦУ аналізує інформацію з БПЛА і приймає рішення на ураження у виникаючій ситуації. Для керування наземним роботом-кулеметом (РК), ЦУ віддає йому команди, при необхідності аналізує результати враження через БПЛА, корегує напрямок пострілу і віддає повторну команду на враження. Звичайно, деякі команди і функції управління лягають безпосередньо на операторів. Але виконання деяких команд можливо налаштувати на автоматичне виконання, за заданою умовою. При цьому вивільняється час оператора для виконання іншого завдання. Наприклад, при зменшенні боєкомплекту до граничного рівня, РК може повідомити про це ЦУ – сервер, який проаналізує вхідну інформацію, і по заданому алгоритму прийме рішення про повернення даного наземного комплексу у місце постійної дислокації.

При цьому РК з самого початку не володіє інформацією стосовно точки повернення, щоб у випадку захоплення ці дані не були використані ворогом. Іншими словами, ЦУ видає набір координат – детальний маршрут на карті, або частина маршруту з точкою, при досягненні якої наземному роботу буде видано наступні координати руху. Іншим аналогічним прикладом може слугувати сигнал від БПЛА про низький заряд акумуляторної батареї (АКБ), при якому ЦУ може проаналізувати і направити такий апарат до найближчої або безпечної точки заміни/заряджання АКБ. Як видно з наведених прикладів, ЦУ повинен вміти самостійно обробляти частину інформації.

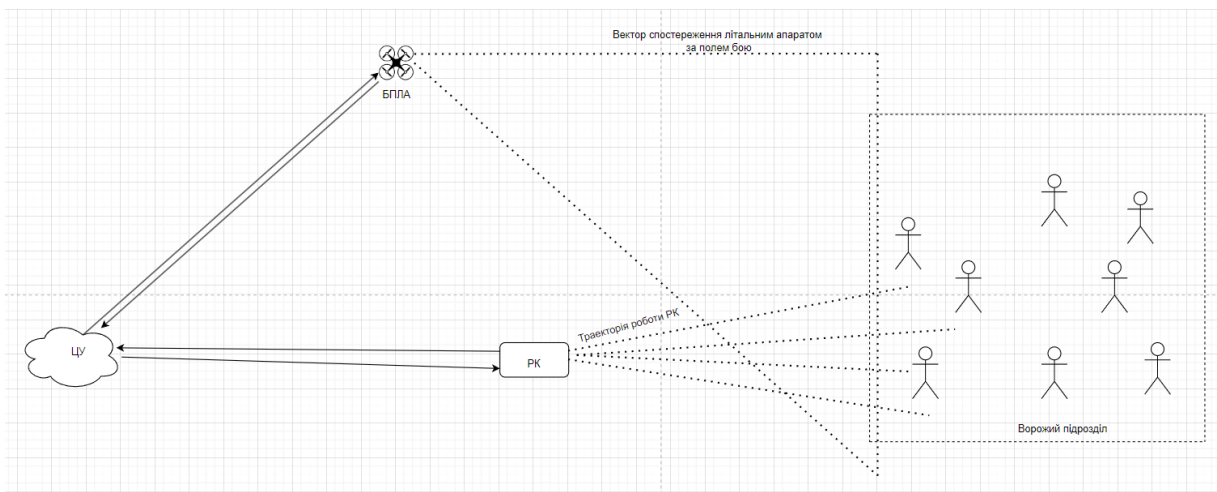


Рис. 2. Схематичне зображення взаємодії між ЦУ – БПЛА, ЦУ – РК.

Розглянемо інший приклад. Задача як і попередньому прикладі – відбиття ворожого штурму наявними засобами: ЦУ, БПЛА, РК. Припустимо, між ЦУ і РК знаходиться густе лісонасадження, яке перешкоджає стійкій роботі засобів зв'язку (див. рис. 3). У такому випадку напряму керування РК з боку ЦУ стає неможливим. Для вирішення задачі коригування вогню логічно використати БПЛА як посередника між ЦУ і РК. У такому випадку БПЛА, що є клієнтом по відношенню до ЦУ, стає сервером по відношенню до РК. Звідси з'являються вимоги, що можуть мати вищу цінність по відношенню до інших. Наприклад, БПЛА спостереження, при зміні вектору атаки ворожих сил, повинен фокусуватися на певних цілях, отримувати команди з боку ЦУ та керувати РК. У цей час основна задача даного конкретного БПЛА – спостереження за усією ділянкою відповідальності, стає менш пріоритетною задачею. Але в той же час автори хочуть підкреслити, що зміна вектору інтенсивності обробки повідомлень з боку БПЛА – РК не означає, що такий БПЛА перестає виконувати свою задачу спостереження, лише акцентують, що квантування часу тепер буде перерозподілено, що дозволить виконувати більш цінні задачі інтенсивніше, а отже кількість виконаних менш цінних задач буде зменшена.

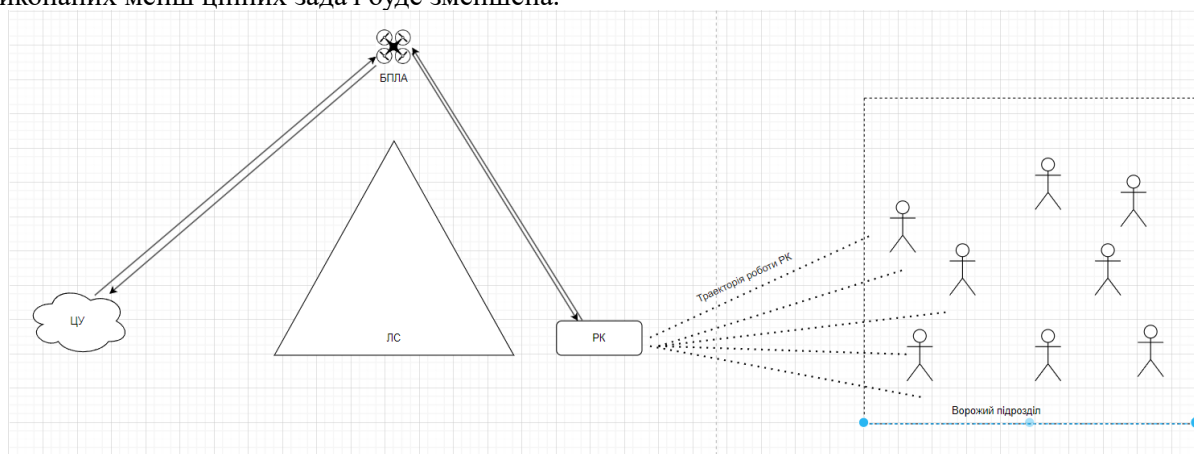


Рис. 3. Схематичне зображення взаємодії між ЦУ – БПЛА, БПЛА – РК.

Де ЛС – лісонасадження, лісосмуга чи інша перешкода, що унеможливує пряму взаємодію між ЦУ та РК.

Для забезпечення вирішення описаної вище задачі розробки управління системою літальних апаратів на думку авторів необхідно виконати подальші наступні завдання:

- Формалізований опис системи безпілотних (літальних) апаратів і зв'язків поміж її окремими елементами;
- Постановку і розробку окремих функціональних задач які повинна виконувати система безпілотних (літальних) апаратів;
- Формалізований опис інформаційних потоків, які функціонують поміж окремими елементами системи літальних апаратів;

- Класифікація окремих повідомлень, що функціонують в межах інформаційних потоків поміж елементами системи літальних апаратів;
- Визначення і дослідження якісно-кількісних характеристик (зокрема цінності і старіння) для врахування при організації їх обробки;
- Формулювання задач оптимальної (раціональної) організації обробки інформаційних потоків для управління системою безпілотних (літальних) апаратів для забезпечення вирішення функціональних задач;
- Розробка методів вирішення задач оптимальної (раціональної) організації обробки інформаційних потоків в системі літальних апаратів.

Читач міг простежити думку, що в останньому наведеному прикладі використання цивільного БПЛА ускладнюється або унеможлиблюється тим, що такий ЛА повинен містити додатковий модуль керування чи трансляції. В дронах невійськового призначення це додає накладні витрати на злітну масу такого апарату а отже, і час його безперервної роботи. Тому автори фокусуються на тому, що для розвитку вітчизняного комплексу безпілотних систем потрібно робити акцент на розробку вітчизняного програмного забезпечення (ПЗ) для керування подібними БПЛА у сумісності з вітчизняними зразками техніки.

Було проведено порівняльне дослідження існуючих технічних рішень, які б могли бути використані для побудови клієнт-серверної архітектури між вказаними системами. Це мікроконтролер STM32, мікроконтролер ESP32, мікрокомп'ютер Raspberry Pi. В дослідженні приймалося до уваги декілька важливих параметрів:

- наявність достатньої потужності для виконання поточних і перспективних задач;
- вбудовані можливості для розширення;
- перспективи можливостей розширення за допомогою допоміжних плат чи інших мікроконтролерів;
- наявність стійкого, достатнього ком'юніті для обговорення і вирішення питань, що будуть виникати в процесі розробки чи тестування ПЗ на даному виді технічного пристрою;
- наявність засобів безпеки для шифрування вихідних даних.

На думку авторів прийнятними є технічні рішення на базі ESP32 та Raspberry Pi. На сьогоднішній день STM32 мікроконтролер також розвивається, але вбудовані можливості для розширення є недостатніми на думку авторів. Частково це можна компенсувати завдяки пристроям розширення. ESP32 в свою чергу має достатню потужність, вбудовані можливості для розширення, широкий набір периферичних інтерфейсів, які дозволять іще розширити можливості пристрою. Окрім того, мікроконтролер ESP32 має криптографічне апаратне прискорення для забезпечення усіх функцій безпеки [10]. Цей мікроконтролер має низьке живлення, що також характеризує його як потенційного кандидата для БПЛА. Мінусом даного мікроконтролера може бути низька потужність у випадках, коли його треба використовувати у БПЛА спостереження, і при цьому стискати відео для економії трафіку. У цьому випадку автори розглядають беззаперечного лідера – мікрокомп'ютер Raspberry Pi. Цей технічний пристрій містить в собі повноцінну операційну систему (ОС), зазвичай на базі ядра Linux, але не обмежений лише однією ОС. Наявність повноцінної ОС дає можливість розробникам абстрагуватись від реалізації тієї чи іншої функціональності. Це у свою чергу дозволяє пришвидшити процес розробки програмного забезпечення шляхом розпаралелювання процесу розробки на модулі, абстрагуючись від роботи один одного. Єдиним недоліком Raspberry Pi можна вважати збільшене споживання електроенергії, у порівнянні з вище названими мікроконтролерами. Але цей недолік поступово усувається з розвитком більш ємнісних акумуляторних батарей.

Висновки. У межах даної роботи описано концептуальну модель взаємодії між центрами управління та окремими безпілотними (літальними) апаратами та системами. Запропоновано технічний підхід, при якому забезпечиться керування окремими безпілотними елементами без прямого контакту між центром управління та безпосередньо пристроєм. Зроблено порівняльне дослідження існуючих технічних засобів, що забезпечать розробку і використання програмного забезпечення вітчизняного походження.

Перспективи подальшого дослідження. У даній роботі клієнт-серверне рішення описано лише поверхнево. У подальшому необхідно проаналізувати існуючі програмні засоби, такі як Apache HTTP Server, nginx та можливо інші на предмет використання у описаній вище концептуальній моделі. Окрім того, в даній роботі не було описано використання шифрування повідомлень, що є обов'язковим в умовах ведення бойових дій.

Список бібліографічного опису

1. Мороз Б.І., Круглик А.С., Мороз Д.М., Мартиненко А.А. Математична модель і загальний алгоритм вирішення задачі обробки повідомлень з урахуванням їх цінності і старіння в системах літальних апаратів. *Системні технології. Регіональний міжвузівський збірник наукових праць*. 2024. № 5(154). С. 3 – 18. DOI 10.34185/1562-9945-5-154-2024-01
2. Зіграє вирішальну роль у війні з Росією: США запустили в стратосферу дрон-шпигун (відео) / Фокус: веб-сайт. URL: <https://focus.ua/uk/digital/580632-zigraye-virishalnu-rol-u-vijni-z-rosiyeyu-ssha-zapustili-v-stratosferu-dron-shpigun-video>. (дата звернення 02.02.2024).
3. В США успішно протестували дрон-шпигун, якого не помічає ППО / РБК-Україна: веб-сайт. URL: <https://www.rbc.ua/rus/news/ssha-uspishno-protestuvali-dron-shpigun-kogo-1690039622.html>. (дата звернення 28.05.2024).
4. Армія США випробувала безпілотну систему, що стріляє ракетами по кораблях / Фокус: веб-сайт. URL: <https://focus.ua/uk/digital/580545-armiya-ssha-viprobuvala-bezpiilotnu-sistemu-sho-strilyaye-raketami-po-korablyah>. (дата звернення 01.05.2024).
5. В Україні створили безпілотник-кулемет (Відео) / Військовий кур'єр: веб-сайт. URL: <https://mil.co.ua/v-ukrayini-stvoryly-bezpiilotnyk-kulemet-video/> (дата звернення 20.04.2024).
6. Беляков Р. О. Проблема інтеграції повітряної мережі класу FANET в мобільну комунікаційну мережу спеціального призначення. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2023. №53. С. 263-276. DOI: <https://doi.org/10.36910/6775-2524-0560-2023-53-40>.
7. Дичка І.А., Радченко К.О., Терейковський І.А., Терейковська Л.О. Концептуальна модель процесу прогнозування навантаження на вебсервер. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. №54. С. 74-83. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-09>.
8. Беляков Р.О. Ієрархічна модель інтелектуального управління наземно-повітряної комунікаційної мережі спеціального призначення. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. №54. С. 225-235. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-28>.
9. Озерчук І.М. Архітектура платформи програмно-визначеного радіо на основі процесора загального призначення. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. №54. С. 165-170. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-20>.
10. STM32. Електронний ресурс. URL: <https://ru.wikipedia.org/wiki/STM32>. (дата звернення 01.06. 2024).
11. ESP32. Електронний ресурс. URL: <https://ru.wikipedia.org/wiki/ESP32>. (дата звернення 01.06.2024).
12. Raspberry Pi. Електронний ресурс. URL: https://ru.wikipedia.org/wiki/Raspberry_Pi. (дата звернення 01.06.2024).

References

1. B. Moroz, A. Kruhlyk. Mathematical model and general algorithm for solving the problem of processing messages taking into account their value and aging in aircraft systems. *System technologies*. 2024. № 5(154). P. 3 – 18. DOI 10.34185/1562-9945-5-154-2024-01
2. It will play a decisive role in the war with Russia: the USA launched a spy drone into the stratosphere (video) / Focus: website. URL: <https://focus.ua/uk/digital/580632-zigraye-virishalnu-rol-u-vijni-z-rosiyeyu-ssha-zapustili-v-stratosferu-dron-shpigun-video>. (date of app. 02.02.2024).
3. The USA successfully tested a spy drone that is not noticed by air defense / RBC-Ukraine. website. URL: <https://www.rbc.ua/rus/news/ssha-uspishno-protestuvali-dron-shpigun-kogo-1690039622.html>. (date of app. 28.05.2024).
4. The US Army has tested an unmanned system that fires missiles at ships / Focus: website. URL: <https://focus.ua/uk/digital/580545-armiya-ssha-viprobuvala-bezpiilotnu-sistemu-sho-strilyaye-raketami-po-korablyah>. (date of app. 01.05.2024).
5. An unmanned machine gun was created in Ukraine (Video) / Military courier: website. URL: <https://mil.co.ua/v-ukrayini-stvoryly-bezpiilotnyk-kulemet-video/>. (date of app. 20.04.2024).
6. Bieliakov R. The problem of integrating the FANET class air network into a special purpose mobile communication network. *Computer-integrated technologies: education, science, production*. 2023. №53. P. 263-276. DOI: <https://doi.org/10.36910/6775-2524-0560-2023-53-40>
7. Dychka I., Radchenko K., Tereikovskiy I. Tereikovska L. Conceptual model of the forecasting process on web server load. *Computer-integrated technologies: education, science, production*. 2024. №54. P. 74-83. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-09>.
8. Bieliakov R. Hierarchical model of intelligent management of special purpose ground-air communication network. *Computer-integrated technologies: education, science, production*. 2024. №54. P. 225-235. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-28>.
9. Ozerchuk I. A Software-Defined Radio Platform Architecture Based on a General-Purpose Processor. *Computer-integrated technologies: education, science, production*. 2024. №54. P. 165-170. DOI: <https://doi.org/10.36910/6775-2524-0560-2024-54-20>.
10. STM32. Electronic resource. URL: <https://ru.wikipedia.org/wiki/STM32>. (date of ap. 01.06. 2024).
11. ESP32. Electronic resource. URL: <https://ru.wikipedia.org/wiki/ESP32>. (date of ap. 01.06. 2024).
12. Raspberry Pi. Electronic resource. URL: https://ru.wikipedia.org/wiki/Raspberry_Pi. (date of ap. 01.06. 2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-20>

УДК 004.651.4:629.5

Морохович Василь Степанович, к.ф.-м.н., доцент

<https://orcid.org/0000-0002-4939-6566>

Лях Ігор Михайлович, к.т.н., доцент

<https://orcid.org/0000-0001-5417-9403>

Хом'як Максим Ігорович, студент

<https://orcid.org/0009-0003-6769-7009>

Морохович Богдан Васильович, магістр

<https://orcid.org/0000-0002-3498-6547>

Ужгородський національний університет, м. Ужгород, Україна

ПРОГНОЗУВАННЯ ПАСАЖИРІВ, ЯКІ ВИЖИЛИ ПІД ЧАС КАТАСТРОФИ «ТИТАНІКА», ЗА ДОПОМОГОЮ ДЕРЕВА ПРИЙНЯТТЯ РІШЕНЬ

Морохович В.С., Лях І.М., Хом'як М.І., Морохович Б.В. Прогнозування пасажирів, які вижили під час катастрофи «Титаніка», за допомогою дерева прийняття рішень. У статті запропоновано застосування методу дерева прийняття рішень для прогнозування пасажирів, які вижили в результаті катастрофи лайнера «Титанік». У якості вхідних даних використано набори даних «Titanic – Machine Learning from Disaster», що наявні у відкритому доступі на платформі Kaggle. Дерева прийняття рішень добре підходять для вирішення задач класифікації, а їхня простота інтерпретації робить цей метод кращим вибором серед інших алгоритмів машинного навчання. У навчальному наборі даних було виконано модифікацію, щоб заповнити відсутні значення. Оцінку розподілу якісних і кількісних ознак даних та пошуку в них закономірностей здійснено за допомогою візуального аналізу даних, що дало змогу виділити ознаки пасажирів, які корелюють з їхнім виживанням найбільше, і відповідно покращити набір даних. Дерево прийняття рішень для кінцевого набору даних побудовано за допомогою бібліотеки scikit-learn (sklearn), яка надає потужні інструменти для машинного навчання в Python. Точність побудованого дерева рішень становить 77% на відкладеній вибірці. Подальше дослідження застосування дерева прийняття рішень для даного набору даних може бути проведено шляхом використання методу налаштування гіперпараметрів дерева (hyperparameter tuning), який допоможе покращити точність побудованого дерева рішень.

Ключові слова: машинне навчання, аналіз даних, дерева прийняття рішень, прогнозування, платформа Kaggle.

Morokhovych V., Liakh I., Khomyak M., Morokhovych B. Predicting passengers who survived the Titanic disaster using a decision tree. The article proposes the use of the decision tree method for predicting the passengers who survived the Titanic liner disaster. The Titanic - Machine Learning from Disaster dataset, which is publicly available on the Kaggle platform, is used as input. Decision trees are well suited for solving classification and forecasting problems, and their ease of interpretation makes this method the best choice among other machine learning algorithms. On training data set, modification has been performed to fill the missing values. The distribution of qualitative and quantitative data features and the search for patterns in the data were evaluated using visual data analysis, which allowed us to identify the passenger features that correlate with their survival the most and improve the data set accordingly. The decision tree for the final dataset was built using the scikit-learn library (sklearn), which provides powerful tools for machine learning Python. The accuracy of the built decision tree is 77% of the deferred sample. Further study of the application of the decision tree for this dataset can be done by using the hyperparameter tuning method, which will help to improve the accuracy of the constructed decision tree.

Keywords: machine learning, data analysis, decision trees, forecasting, Kaggle platform.

Постановка проблеми. Катастрофа «Титаніка» сталася більше 100 років тому, але все ще приваблює дослідників зрозуміти та вивчити, як одні пасажирів вижили, а інші загинули. Лайнер «Титанік» був найбільшим судном свого часу. Його трюм складався з шістнадцяти частин, навіть повне затоплення чотирьох із яких не могло призвести до його затоплення [7]. Не дивлячись на високі стандарти безпеки, які вкладали в цей корабель його конструктори, однак несприятливі події призвели до зіткнення «Титаніка» з айсбергом і подальшої жахливої катастрофи, яка забрала життя багатьох людей. Не зважаючи на давнину подій, дослідження даних про катастрофу продовжується і понині, в тому числі із використанням методів штучного інтелекту та машинного навчання. Різні атрибути пасажирів, такі як стать, вік, категорія, до якої вони належать, і їхній соціальний клас тощо, дозволяють побудувати прогностичні моделі, забезпечуючи якісну базу даних для аналізу. Дані дослідження надають можливість передбачити виживання пасажирів на «Титаніку» за допомогою різних методів на основі даних платформи Kaggle «Titanic – Machine Learning from Disaster».

Аналіз останніх досліджень і публікацій. Дослідження прогнозування розподілу виживання пасажирів під час катастрофи «Титаніка» із використанням набору даних Kaggle проводились науковцями, використовуючи різні алгоритми машинного навчання. Приміром, для прогнозування рівня виживання пасажирів лайнера дослідники застосовували різні алгоритми,

включаючи логістичну регресію, k -найближчих сусідів, метод опорних векторів, «випадкового лісу» (random forest), штучних нейромереж та дерева рішень [4, 5]. У роботі [3] була проведена класифікація з двома класами (пасажирів, які вижили і не вижили) за допомогою дерева рішень, де виживання аналізувалося на кожному рівні. Кластеризація виконувалась за допомогою алгоритму машинного навчання KMeans, а його реалізацію виконано за допомогою програмування Python. Науковцями встановлено, що пасажирів, які подорожували з невеликою сім'єю, яка складала від 2 до 4 осіб, мали більше шансів на виживання. У статті [2] розглянуто створення систем прогнозування на основі методу опорних векторів (SVM). Дослідження полягало в тому, щоб побудувати серію моделей машинного навчання з точністю f -вимірювання понад 80% на заданому наборі демографічної інформації на Kaggle. Автором було досягнуто найкращого результату, тобто 82,82% правильних прогнозів.

Сумнозвісний інцидент затоплення найбільшого круїзного лайнера змушує дослідників і надалі заглиблюватися у масиви даних та проводити дослідницький аналіз даних, щоб зрозуміти вплив ключових параметрів на виживання людей на його борту.

Формулювання мети дослідження. Метою статті є аналіз даних про пасажирів круїзного лайнера «Титанік» та побудова дерева рішень з метою прогнозування пасажирів, які вижили під час його катастрофи, використовуючи набори даних, що наявні у відкритому доступі на платформі Kaggle.

Виклад основного матеріалу. Одним із методів розв'язання задач класифікації та прогнозування є дерева прийняття рішень, що представляє собою ієрархічну структуру наборів правил, які послідовно дають відповіді «так» або «ні». Структура дерева починається з кореневого вузла, який не має жодних вхідних гілок. Гілки, що виходять з кореневого вузла, потрапляють до внутрішніх вузлів, відомих як вузли прийняття рішень. На основі існуючих ознак обидва типи вузлів виконують оцінки і формують однорідні підмножини, які позначаються листовими вузлами. Саме вони являють собою всі можливі результати набору даних.

Метод дерева прийняття рішень має такі переваги над іншими алгоритмами машинного навчання: їх легше інтерпретувати, може працювати з якісними та кількісними ознаками, не потребує попередньої нормалізації даних.

Kaggle є відомою онлайн-платформою для спільної роботи у сфері аналізу даних, машинного навчання і штучного інтелекту [1]. Вона надає користувачам можливість обмінюватися знаннями, виконувати проекти та приймати участь у них, які пов'язані з розробкою моделей машинного навчання. Однією із поставлених задач є «Titanic – Machine Learning from Disaster», яка є навчальною.

Набір даних на платформі Kaggle складається з двох груп: дані для навчання моделі та для її тестування. Структура цих даних однакова, за виключенням відсутності ознаки виживання пасажирів в тестових даних (рис. 1).

PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

Рис. 1. Набір даних пасажирів «Титаніка» на платформі Kaggle

Загалом навчальний набір даних містить інформацію про 814 пасажирів та 12 ознак. Також із цих даних за необхідності можна одержати інформацію про середні значення, стандартні відхилення та інші статистичні характеристики ознак.

Для подальшого використання даних необхідно провести їх очищення, тобто позбутися неінформативних ознак, розібратися з відсутніми значеннями в даних і модифікувати їх [6].

Суттєвою проблемою є наявність в наборі даних відсутніх значень, а саме – інформації про вік деяких пасажирів. Цей недолік було усунуто шляхом заповнення відсутніх даних середнім значенням віку пасажирів. На рис. 2 представлено набір даних після їх очищення.

	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Embarked
0	0	3	1	22.0	1	0	7.2500	0
1	1	1	0	38.0	1	0	71.2833	1
2	1	3	0	26.0	0	0	7.9250	0
3	1	1	0	35.0	1	0	53.1000	0
4	0	3	1	35.0	0	0	8.0500	0

Рис. 2. Набір даних після їх очищення

Перед застосуванням методів машинного навчання є корисним застосування методу візуального аналізу даних для досліджуваного набору даних. Цей метод полягає в аналізі вигляду розподілів ознак даних та пошуку в них закономірностей. Це дозволить виділити ознаки пасажирів, які корелюють з їхнім виживанням найбільше, і відповідно покращити набір даних, що призведе до покращення результатів передбачення дерева рішень.

Для оцінки розподілу якісних та кількісних ознак побудовано їх графіки (рис. 3, 4).

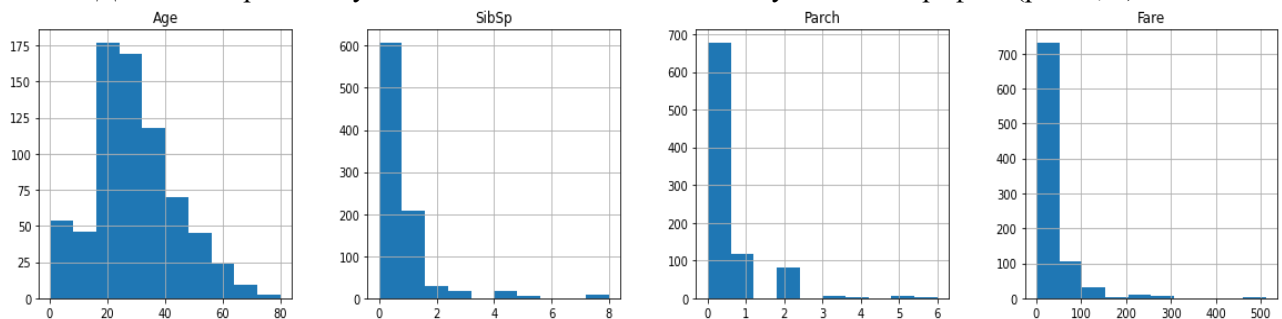


Рис. 3. Гістограми кількісних ознак даних

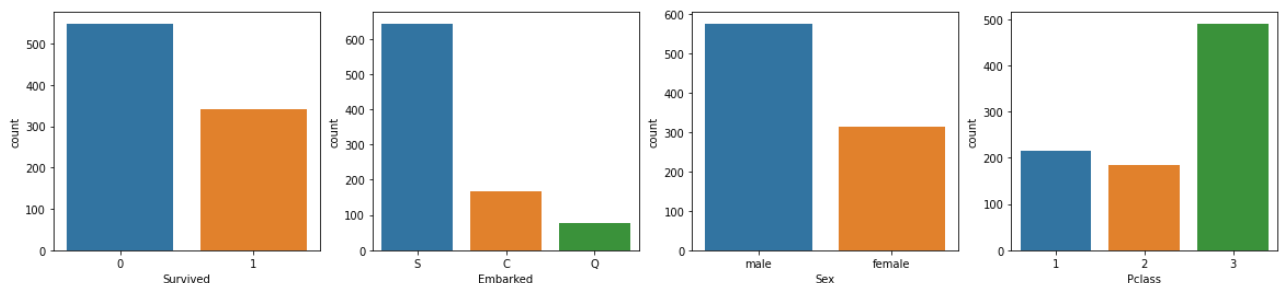


Рис. 4. Гістограми якісних ознак даних

Аналізуючи графіки, можна зазначити, що тільки вік пасажирів (Age) розподілений нормально на відміну від інших ознак, таких як: стать (Sex), кількість братів і сестер/подружжя (SibSp), порт посадки (Embarked), вартість квитка (Fare), клас квитка (Pclass), кількість батьків/дітей на борту (Parch), кількість пасажирів, які вижили (Survived).

Гістограми вказують, що в ознаках статі та кількості пасажирів, які вижили, майже є баланс, проте в ознаці «порти посадки» домінує порт Cherbourg, а серед класу квитків пасажирів – найбільше 3-го класу. Такий дисбаланс у двох ознаках міг би негативно вплинути на навчання моделі ШІ для застосування її на великих наборах даних.

Отже, в наборі даних є всього вісім дійсно корисних для аналізу ознак, чотири із яких – кількісні, а інші чотири – якісні, причому однією із якісних є цільова ознака пасажирів, які вижили. Саме для цієї ознаки необхідно визначити фактори, які найбільше на неї впливають. Порівнювати між собою всі ознаки одночасно не можливо, тому є необхідність побудувати графіки для різних типів ознак. Нижче побудовано графік типу «boxplot», який дозволяє виявляти зв'язки між якісною ознакою пасажирів, які вижили, та кількісними ознаками (Age, SibSp, Parch, Fare) (рис. 5).

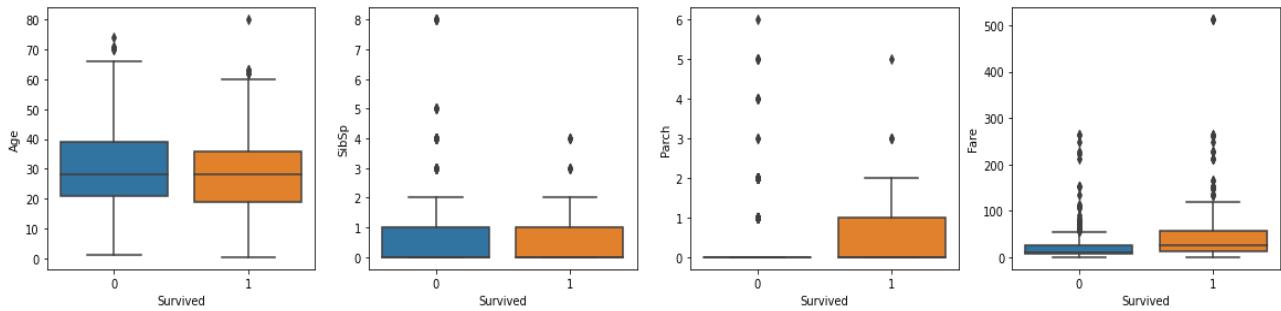


Рис. 5. Графіки «boxplot» порівняння кількісних ознак з ознакою пасажирів, які вижили

На основі отриманих графіків, можна зробити висновок, що вік пасажирів незначно впливає на їх виживання – у молодших пасажирів трохи більші шанси вижити. Ознака «кількість братів і сестер/подружжя» не впливає на виживання пасажирів. Наявність у пасажиря батьків/дітей на борту сильно корелюється з ознакою пасажирів, які вижили, а також, вартість квитка пасажиря корелюється з його виживанням.

Для аналізу якісної ознаки виживання пасажирів з іншими якісними ознаками побудовано наступні гістограми (рис. 6).

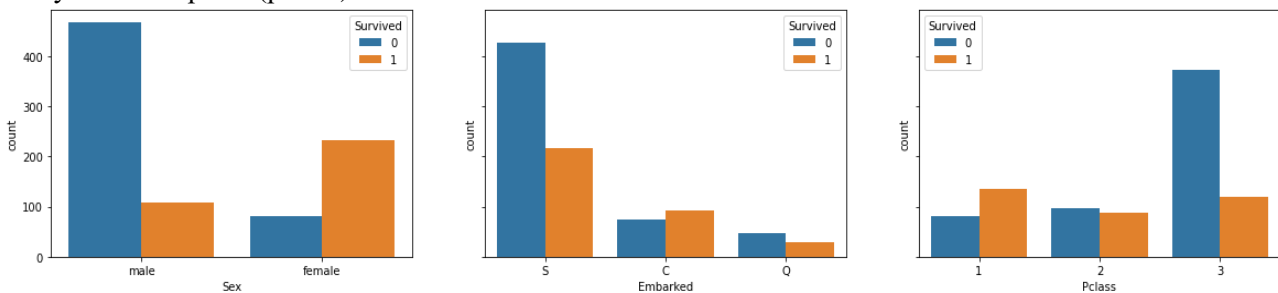


Рис. 6. Гістограми порівняння якісних ознак з ознакою виживання пасажирів (0 – пасажиря, які загинули; 1 – пасажиря, які вижили)

Проаналізувавши гістограми, можна відмітити, що стать пасажиря сильно впливає на його виживання – серед жінок набагато більше змогли врятуватися. Найбільше загинуло людей серед пасажирів третього класу. Цікавим є графік з портами посадки пасажирів – серед пасажирів, які сіли в порту «Southampto», найбільше загинув, тобто ознака порту посадки пасажирів також виявилась важливою. Тому доцільно побудувати гістограми для порівняння порту посадки пасажирів та їх класу квитків і статі (рис. 7).

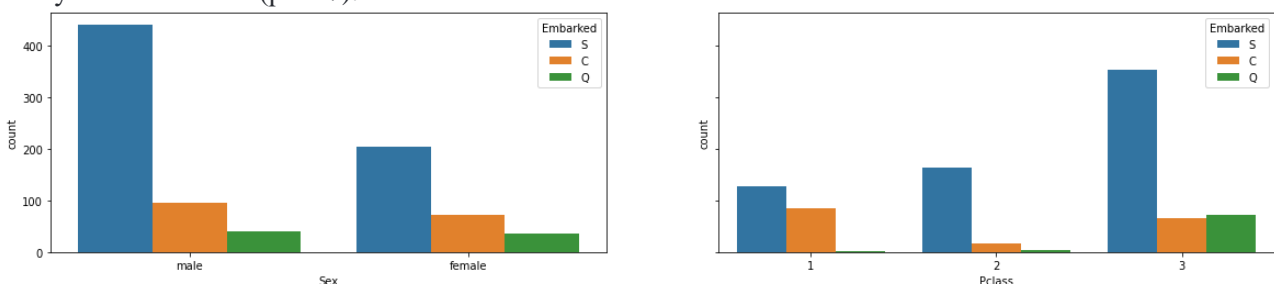


Рис. 7. Порівняння портів посадки пасажирів з їх класом квитків та статтю

З рис. 7 видно, що в порту «Southampto» на лайнер «Титанік» сіли, в основному, чоловіки та пасажиря з квитками третього класу. Тому буде доцільно вилучити ознаку порту посадки з набору даних, оскільки вона є зайвою. Також було вилучено ознаку класу квитка, оскільки вона прямо корелює з його вартістю.

Переконавшись, що дані переведено в числову форму і в них немає відсутніх значень, можна перейти до побудови дерева рішень для прогнозування цільової ознаки, а саме пасажирів, які вижили.

Максимальна глибина побудованого дерева дорівнює трьом, а його ефективність вимірювалась за допомогою метрики точності та становить 77%. Графічне представлення дерева рішень наведено на рис. 8.

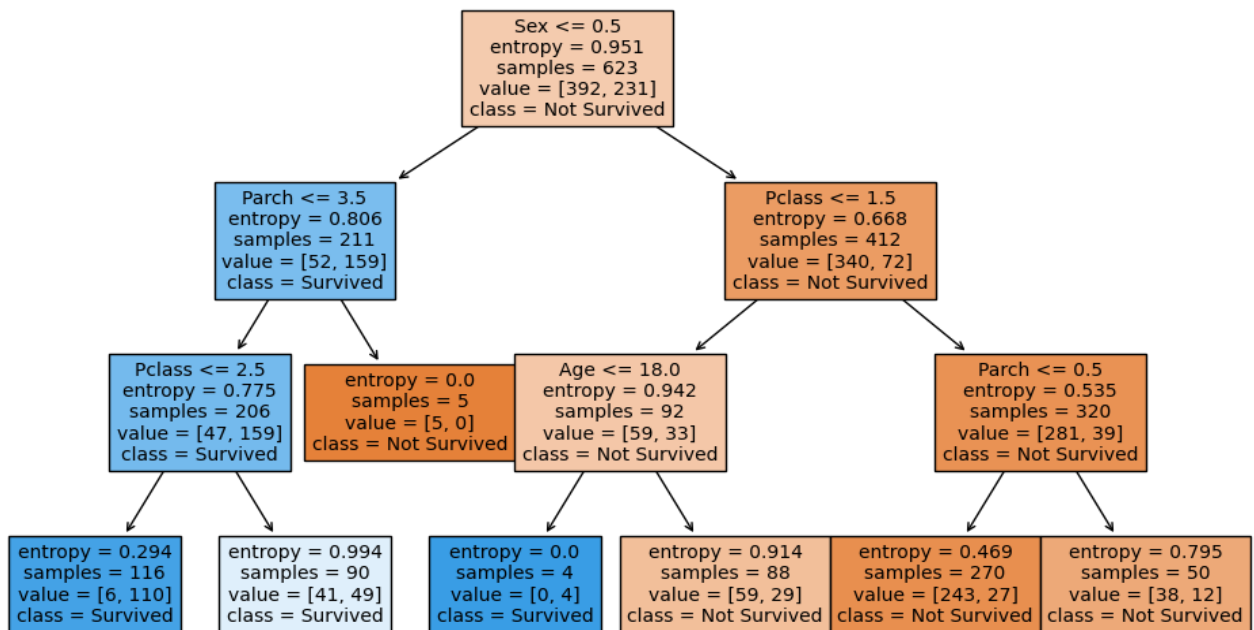


Рис. 8. Дерево прийняття рішень для набору даних

Дерево прийняття рішень для кінцевого набору даних було побудовано за допомогою бібліотеки scikit-learn (sklearn), яка надає потужні інструменти для машинного навчання в Python. У бібліотеці sklearn існує клас DecisionTreeClassifier, який дозволяє побудувати дерево рішень для класифікації, тобто передбачення категоріальної цільової ознаки.

Висновки та перспективи подальшого дослідження. У результаті проведеного дослідження було проаналізовано дані про пасажирів круїзного лайнеру «Титанік», які перебували на ньому під час його катастрофи 14 квітня 1912 року. Дані про пасажирів було очищено та використано для побудови дерева рішень з метою прогнозування пасажирів, які вижили. Точність побудованого дерева рішень становить 77% на відкладеній вибірці. Таким чином можна зробити висновок, що дерева прийняття рішень добре підходять для розв'язування задач класифікації, а легкість їх інтерпретації робить цей метод пріоритетним вибором серед інших алгоритмів машинного навчання, коли необхідне чітке розуміння, як саме приймається певне рішення.

Подальше дослідження застосування дерева рішень на даному наборі даних може бути проведено шляхом покращення точності вже побудованого дерева рішень хоча б до значення 80% на відкладеній вибірці. Метод, який допоможе покращити точність побудованого дерева рішень, є налаштування гіперпараметрів дерева (hyperparameter tuning). Наприклад, при побудові дерева рішень можна вказати такі гіперпараметри, як максимальна глибина дерева та максимальна кількість ознак набору даних, які враховуються алгоритмом при побудові чергових вузлів дерева.

Список бібліографічного опису

1. Лук'янюк В. (2017). Цей день в історії: Загибель «Титаніка». URL: <https://www.jnsm.com.ua/h/0414M/>.
2. Ai Yu. (2023). Predicting Titanic Survivors by Using Machine Learning, *Highlights in Science, Engineering and Technology*, 34, 360-367, <https://doi.org/10.54097/hset.v34i.5494>.
3. Haque MA, Shivaprasad G. & Guruprasad G. (2021). Passenger data analysis of Titanic using machine learning approach in the context of chances of surviving the disaster, *IOP Conference Series: Materials Science and Engineering*, 1065(1), <https://doi.org/10.1088/1757-899X/1065/1/012042>.
4. Singh A., Saraswat S. & Faujdar N. (2017). Analyzing Titanic disaster using machine learning algorithms, *International Conference on Computing, Communication and Automation (ICCCA)*, Greater Noida, India, 406-411, <https://doi.org/10.1109/CCAA.2017.8229835>.
5. Singh K., Nagpal R. & Sehgal R. (2020). Exploratory Data Analysis and Machine Learning on Titanic Disaster Dataset, *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 320-326, <https://doi.org/10.1109/Confluence47617.2020.9057955>.

6. Song YY, Lu Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*, Vol. 27(2), 130-135, <https://doi.org/10.11919/j.issn.1002-0829.215044>.
7. Titanic – Machine Learning from Disaster. URL: <https://www.kaggle.com/c/titanic/data>.

References

1. Lukyaniuk, V. (2017). This day in history: The sinking of the Titanic. URL: <https://www.jnsm.com.ua/h/0414M/>.
2. Ai, Yu. (2023). Predicting Titanic Survivors by Using Machine Learning, *Highlights in Science, Engineering and Technology*, 34, 360-367, <https://doi.org/10.54097/hset.v34i.5494>.
3. Haque, MA, Shivaprasad, G. & Guruprasad, G. (2021). Passenger data analysis of Titanic using machine learning approach in the context of chances of surviving the disaster, *IOP Conference Series: Materials Science and Engineering*, 1065(1), <https://doi.org/10.1088/1757-899X/1065/1/012042>.
4. Singh, A., Saraswat, S. & Faujdar, N. (2017). Analyzing Titanic disaster using machine learning algorithms, *International Conference on Computing, Communication and Automation (ICCCA)*, Greater Noida, India, 406-411, <https://doi.org/10.1109/CCAA.2017.8229835>.
5. Singh, K., Nagpal, R. & Sehgal, R. (2020). Exploratory Data Analysis and Machine Learning on Titanic Disaster Dataset, *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, Noida, India, 320-326, <https://doi.org/10.1109/Confluence47617.2020.9057955>.
6. Song, YY, Lu, Y. (2015). Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*, Vol. 27(2), 130-135, <https://doi.org/10.11919/j.issn.1002-0829.215044>.
7. Titanic – Machine Learning from Disaster. URL: <https://www.kaggle.com/c/titanic/data>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-21>

УДК 621.383

Нижник Олександр Олександрович, бакалавр

Ящинський Леонід Васильович, к.ф.-м.н., доцент

<https://orcid.org/0000-0003-3018-3904>

Федосов Сергій Анатолійович, д.ф.-м.н., професор

<https://orcid.org/0000-0003-3457-8911>

Луцький національний технічний університет, м. Луцьк, Україна

ПРИНЦИП РОБОТИ І ПЕРСПЕКТИВИ ВИКОРИСТАННЯ ГОЛОГРАФІЧНИХ ДИСПЛЕЇВ

Нижник О.О., Ящинський Л.В., Федосов С.А. Принцип роботи і перспективи використання голографічних дисплеїв. У цій роботі ми досліджуємо принцип роботи голографічних дисплеїв (моніторів), їхні технічні характеристики та потенційні переваги порівняно з традиційними дисплеями. Також розглядаються перспективи використання голографічних дисплеїв у різних галузях, включаючи медицину, освіту, розваги, телекомунікації та виробництво. Розуміння принципів функціонування цих пристроїв та їхніх потенційних можливостей може привести до нових можливостей для візуалізації даних та створення іммерсивних візуальних досвідів.

Ключові слова: голографічний дисплей, монітор, 3D-зображення, інтерференція.

Nyzhnyk O., Yashchynskyy L., Fedosov S. The principle of operation and prospects for the use of holographic displays. In this work, we investigate the principle of operation of holographic displays (monitors), their technical characteristics and effective advantages replaced by traditional displays. Prospects for the use of holographic displays in various industries, including medicine, education, entertainment, telecommunications, and manufacturing, are also considered. Understanding how these devices work and their capabilities can lead to new possibilities for data visualization and creating immersive visual experiences.

Keywords: holographic display, monitor, 3D image, interference.

Постановка наукової проблеми.

У сучасному світі швидкого розвитку технологій візуальне відображення інформації стає все більш важливим і вимагає нових підходів до його забезпечення. У цьому контексті голографічні монітори (дисплеї) представляють собою перспективний напрямок, що набуває все більшого значення в галузі відтворення зображень і візуалізації даних (рис. 1). Голографічні дисплеї працюють на основі принципів голографії, яка дозволяє створювати тривимірні зображення без необхідності в спеціальних окулярах або інших пристроях.

Аналіз досліджень.

Було запропоновано ряд технік для перегляду сцен шляхом повторного відображення попередньо відрендерених або оцифрованих зображень. Також були запропоновані методи інтерполяції між видами шляхом деформації вхідних зображень, використання інформації про глибину або відповідності між кількома зображеннями. У праці [1] описано простий і надійний метод для генерації нових видів з довільних положень камери без інформації про глибину або відповідності функцій, просто об'єднуючи та повторюючи вибірку доступних зображень. Ключ до цієї техніки полягає в інтерпретації вхідних зображень як 2D-зрізів 4D-функції – світлового поля. Ця функція повністю характеризує потік світла через безперешкодний простір у статичній сцені з фіксованим освітленням. Авторами описано вибіркоче представлення для світлових полів, яке дозволяє як ефективно створювати та відображати зовнішніх і внутрішніх виглядів. Створено світлові поля з великих масивів відтворених і оцифрованих зображень. Останні знімаються за допомогою відеокамери, встановленої на порталі, керованому комп'ютером. Після створення світлового поля можна створювати нові види в режимі реального часу шляхом виділення зрізів у відповідних напрямках. Оскільки успіх методу залежить від високої частоти дискретизації, описано систему стиснення, яка здатна стискати згенеровані світлові поля більше ніж у 100:1 із дуже невеликою втратою точності. Авторами також розглянуто проблеми згладжування під час створення та повторної дискретизації під час вилучення фрагментів.

Нещодавно було показано, що поверхні, вкриті ультратонкими плазмонними структурами – так звані метаповерхні – здатні повністю контролювати фазу світла, що представляє нову парадигму для розробки інноваційних оптичних елементів, таких як ультратонкі плоскі лінзи, спрямовані зв'язки для поверхневих плазмонних поляритонів. і генерація вихрового променя хвильової пластини. Серед різних типів метаповерхень геометричні метаповерхні, які складаються з масиву

плазмонних нанострижнів із просторово змінною орієнтацією, показали чудовий контроль фази завдяки геометричній природі їхнього фазового профілю. Метаповерхні нещодавно використовувалися для створення голограм, згенерованих комп'ютером, але ефективність голограм залишалася надто низькою у видимих довжинах хвиль для практичних цілей. Авторами [2] повідомлено про дизайн і реалізацію геометричної метаповерхневої голограми, яка досягає ефективності дифракції 80 % на 825 нм і має широку смугу від 630 нм до 1050 нм. Продемонстрована 16-рівнева фазова голограма, згенерована комп'ютером, поєднує в собі переваги геометричної метаповерхні для чудового контролю профілю фази та рефлекторних масивів для досягнення високої ефективності перетворення поляризації. Зокрема, дизайн голограми об'єднує шліфовану металеву площину з геометричною метаповерхнею, що підвищує ефективність перетворення між двома станами кругової поляризації, що забезпечує високу ефективність дифракції без ускладнення процесу виготовлення. Завдяки цим перевагам дана стратегія може бути життєздатною для різних практичних голографічних застосувань.

Метаповерхні дозволили реалізувати безліч нових функцій у надтонкому вимірі, прокладаючи шлях до плоских та високоінтегрованих фотонних пристроїв. Незважаючи на швидкий прогрес у цій галузі, одночасна реалізація реконфігурації, високої ефективності та повного контролю над фазою та амплітудою розсіяного світла становить велику проблему. У [3] автори намагалися розв'язати цю задачу, представивши концепцію перепрограмованої голограми на основі 1-бітових метаповерхень кодування. Стан кожної елементарної комірки кодуєвої метаповерхні можна перемикає між «1» і «0» шляхом електричного керування навантаженими діодами. Експерименти з підтвердження концепції показують, що кілька бажаних голографічних зображень можна реалізувати в режимі реального часу лише за допомогою однієї метаповерхні кодування. Запропонована перепрограмована голограма може стати ключем до створення майбутніх інтелектуальних пристроїв із реконфігурованими та програмованими функціями, які можуть призвести до прогресу в різноманітних додатках, таких як мікроскопія, дисплей, безпека, зберігання даних та обробка інформації.

Методи цифрової голографії удосконалюються для отримання високоякісних зображень кількісної фазово-контрастної мікроскопії. Зокрема, видно, що метод кутового спектру розрахунку голографічного оптичного поля має значні переваги, включаючи жорсткий контроль паразитних шумових компонентів. Голографічні фазові зображення отримують із боковою роздільною здатністю 0,5 мкм, обмеженою дифракцією, і значною мірою захищені від когерентного шуму, звичайного в інших голографічних методах. Фазовий профіль має точність приблизно до 30 нм оптичної товщини [4].

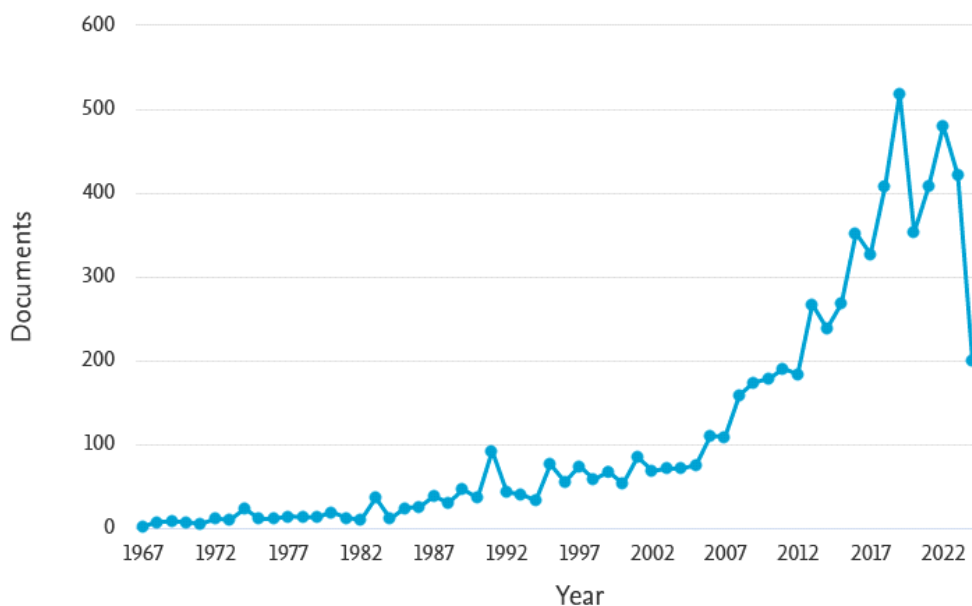


Рис. 1 – Кількість публікацій за роками, пов'язаних з голографічними дисплеями

Рідкокристалічні лазери – це розвиваюча область у галузі фотоніки м'якої матерії, яка може стати провідником нової ери ультратонких, надзвичайно універсальних лазерних джерел. Такі

лазери мають безліч чудових особливостей, включаючи широкосмугову перестроюваність, велику площу когерентності та, у деяких випадках, багатоспрямоване випромінювання. Вони мають потенціал для поєднання великої вихідної потужності з мініатюрними розмірами порожнини – дві властивості, які традиційно були несумісними. Їх потенційні можливості застосування різноманітні, починаючи від мініатюрних медичних діагностичних інструментів і закінчуючи голографічними лазерними дисплеями великої площі. У роботі [5] обговорено наукові витoki цієї технології та дано короткий огляд передових досліджень, які зараз проводяться в усьому світі.

Постановка завдань. Метою дослідження є вивчення основних принципів роботи голографічних моніторів, їхні технічні характеристики і переваги над традиційними дисплеями. Розглянути перспективи використання голографічних моніторів у різних галузях, включаючи медицину, освіту, розваги, телекомунікації та виробництво.

Викладення основного матеріалу й обґрунтування отриманих результатів дослідження.

Загальні відомості. Голографічний дисплей – це тип дисплея, в якому для створення віртуального тривимірного зображення використовується дифракція світла. Голографічні екрани відрізняються від інших способів отримання 3D-зображень тим, що вони не потребують будь-яких спеціальних окулярів або зовнішнього обладнання для перегляду зображення.

Голографія – набір технологій для точного запису, відтворення і переформатування хвильових полів. Це спосіб одержання об'ємних зображень предметів на фотопластинці (голограми) за допомогою когерентного випромінювання лазера. Голограма фіксує не саме зображення предмета, а структуру відбитої від нього світлової хвилі (її амплітуду та фазу). Для отримання голограми необхідно, щоб на фотографічну пластинку одночасно потрапили два когерентних світлових пучки: предметний, відбитий від об'єкта та опорний, що проходить безпосередньо від лазера. Світло обох пучків інтерферує, створюючи на пластинці чергування дуже вузьких темних і світлих смуг – інтерференційну картину.

Голографічні монітори мають ряд переваг перед традиційними дисплеями. Вони більш реалістичні, інтерактивні та можуть відображати більше інформації. Завдяки цьому вони мають потенціал революціонізувати спосіб, яким ми взаємодіємо з комп'ютерами та інформацією.

Принцип роботи. Розглянемо принцип роботи голографічного монітора. Голографічні дисплеї на основі інтерференції:

цей тип дисплея використовує два пучки світла: опорний пучок і предметний пучок;

опорний пучок – це еталонний пучок світла, який не змінюється;

предметний пучок – це пучок світла, який містить інформацію про зображення, яке потрібно створити;

при перетині цих двох пучків світла утворюється інтерференційна картина;

ця інтерференційна картина потім перетворюється на тривимірне зображення спеціальним матеріалом.

Для того щоб побачити голографічне зображення необхідно його записати та в подальшому реконструювати з запису. Процес запису зображення зображено на рисунку 2.

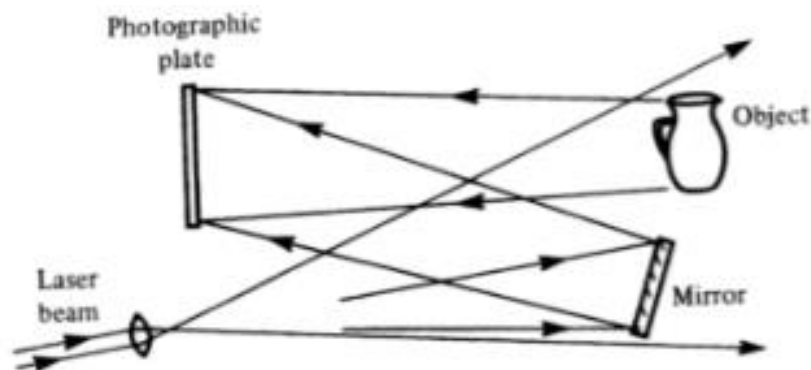


Рис. 2 – Запис голограми

На рисунку 2 фотопластинка записує інтерференційну картину, створювану світловими хвилями, розсіяними від об'єкта, і опорною хвилею, відбитою від нього дзеркалом.

Після того, як голограму було записано, її необхідно відтворити, цей процес представлено на рисунку 3.

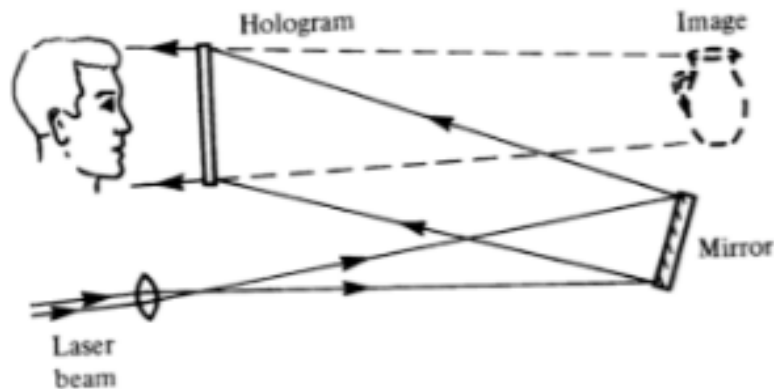


Рис. 3 – Реконструкція зображення

На рисунку 4 представлена інша схема, яка демонструє принцип роботи голографічного монітора.

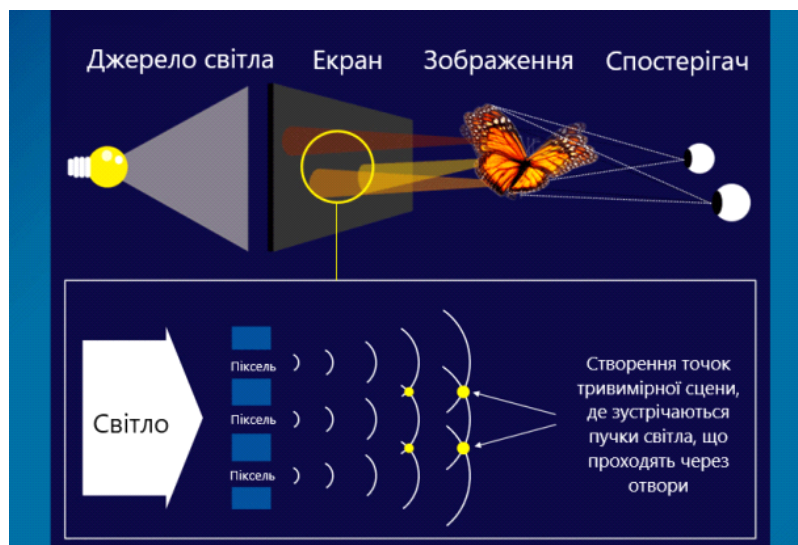


Рис. 4 – Принцип роботи голографічного монітора

Голографія використовує лазерні джерела світла для реконструкції тривимірних зображень. Світло від лазерів зустрічається, щоб створити точку, і безліч таких точок об'єднується, щоб утворити загальну 3D-сцену.

Лазер із когерентним світлом використовується як джерело підсвічування, тоді як пікселі панелі дисплея калібруються і контролюються, щоб світло зустрічалось в певній точці простору для створення точок тривимірної сцени на основі принципу конструктивної інтерференції.

У загальному можна виокремити 3 основні фізичні принципи які використовуються в роботі голографічних моніторів:

інтерференція – це явище, при якому два хвилі світла складаються, щоб створити нову хвилю;

дифракція – це явище, при якому хвиля світла розсіюється, коли вона зустрічає перешкоду;

голографія – це процес запису та відтворення тривимірних зображень.

Перспективи використання. Голографічні монітори пропонують ряд перспективних переваг порівняно з традиційними дисплеями, що робить їх потенційно революційними в багатьох сферах. Розглянемо основні переваги:

1. **Реалістичність.** Голографічні зображення значно реалістичніші, ніж зображення на традиційних дисплеях. Їх тривимірність створює відчуття глибини та об'єму, що робить їх більш захоплюючими та інформативними.

2. Інтерактивність. Голографічні монітори можуть бути інтерактивними, що дозволяє користувачам взаємодіяти з зображеннями та маніпулювати ними. Це може бути корисно для таких завдань, як віртуальне навчання, хірургія та ігри.

3. Відображення інформації. Голографічні монітори можуть відображати значно більше інформації, ніж традиційні дисплеї. Це пов'язано з тим, що вони можуть проєктувати зображення в трьох вимірах, а не лише на двовимірній поверхні.

4. Нові можливості. Голографічні монітори відкривають нові можливості для спілкування, розваг та освіти. Вони можуть використовуватися для створення віртуальних середовищ, які неможливо відтворити за допомогою традиційних дисплеїв.

Враховуючи основні переваги голографічних моніторів можна виокремити декілька основних сфер їх використання.

Освіта. Голографічні монітори можуть використовуватися для створення більш захоплюючого та інтерактивного досвіду навчання. Вони можуть дозволити учням візуалізувати складні концепції та взаємодіяти з ними, що може призвести до кращого засвоєння матеріалу.

Медицина. Голографічні монітори можуть використовуватися для візуалізації анатомії та проведення хірургічних операцій. Це може допомогти хірургам краще планувати та виконувати операції, що може призвести до кращих результатів для пацієнтів.

Розваги. Голографічні монітори можуть використовуватися для створення більш реалістичного досвіду ігор і фільмів. Вони також можуть використовуватися для створення віртуальних концертів та інших заходів.

Реклама. Голографічні монітори можуть використовуватися для створення більш привабливих рекламних дисплеїв. Вони можуть дозволити рекламодавцям розповідати історії та демонструвати продукти більш захоплюючим способом.

Висновки та перспективи подальшого дослідження. Голографічні монітори є перспективним напрямком в сучасних технологіях візуалізації і відображення інформації. Їхній принцип роботи базується на використанні світлової інтерференції та дифракції для створення тривимірних зображень без необхідності у спеціальних окулярах чи гарнітурах. Основні фізичні властивості, такі як дифракція світла й інтерференція, лежать в основі їхньої роботи.

Перспективи використання голографічних моніторів дуже широкі. Вони можуть знайти застосування в медицині для відображення тривимірних моделей органів і тканин, у навчанні для візуалізації складних концепцій, у розважальній індустрії для створення іммерсивних віртуальних середовищ, у телекомунікаціях для покращення спілкування та багато іншого.

Голографічні монітори можуть змінити спосіб, яким ми сприймаємо і взаємодіємо з інформацією, роблячи візуалізацію більш іммерсивною, ефективною та інтерактивною. Хоча вони ще не стали повсякденними в нашому житті, їхні потенційні можливості надають підстави для оптимізму щодо майбутнього розвитку цієї технології.

Список бібліографічного опису

1. Левоу, М., Ханрахан, П. (1996). Візуалізація світлового поля. *Матеріали 23-ї щорічної конференції з комп'ютерної графіки та інтерактивних технологій, SIGGRAPH 1996*, 31–42. DOI: [10.1145/237170.237199](https://doi.org/10.1145/237170.237199)
2. Чжен, Г., Мюленбернд, Х., Кенні, М., Зентграф, Т., Чжан, С. (2015). Метаповерхневі голограми досягають 80% ефективності. *Природні нанотехнології*, 10(4), 308–312. DOI: [10.1038/nnano.2015.2](https://doi.org/10.1038/nnano.2015.2)
3. Лі, Л., Джун Цуй, Т., Джі, В.,...Кю, К.-В., Чжан, С. (2017). Електромагнітні перепрограмовані голограми кодування метаповерхні. *Комунікації природи*, 8(1), 197. DOI: [10.1038/s41467-017-00164-9](https://doi.org/10.1038/s41467-017-00164-9)
4. Манн, К.Дж., Ю, Л., Ло, Ч.-М., Кім, М.К. (2005). Кількісна фазово-контрастна мікроскопія високої роздільної здатності методом цифрової голографії. *Оптика Експрес*, 13(22), 8693–8698. DOI: [10.1364/opex.13.008693](https://doi.org/10.1364/opex.13.008693)
5. Коулз, Х., Морріс, С. (2010). Рідкокристалічні лазери. *Фотоніка природи*, 4(10), 676–685. DOI: [10.1038/nphoton.2010.184](https://doi.org/10.1038/nphoton.2010.184)

References

1. Levoy, M., Hanrahan, P. (1996). Light field rendering. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1996*, 31–42. DOI: [10.1145/237170.237199](https://doi.org/10.1145/237170.237199)
2. Zheng, G., Mühlenbernd, H., Kenney, M.,...Zentgraf, T., Zhang, S. (2015). Metasurface holograms reaching 80% efficiency. *Nature Nanotechnology*, 10(4), 308–312. DOI: [10.1038/nnano.2015.2](https://doi.org/10.1038/nnano.2015.2)
3. Li, L., Jun Cui, T., Ji, W., ...Qiu, C.-W., Zhang, S. (2017). Electromagnetic reprogrammable coding-metasurface holograms. *Nature Communications*, 8(1), 197. DOI: [10.1038/s41467-017-00164-9](https://doi.org/10.1038/s41467-017-00164-9)
4. Mann, C.J., Yu, L., Lo, C.-M., Kim, M.K. (2005). High-resolution quantitative phase-contrast microscopy by digital holography. *Optics Express*, 13(22), 8693–8698. DOI: [10.1364/opex.13.008693](https://doi.org/10.1364/opex.13.008693)
5. Coles, H., Morris, S. (2010). Liquid-crystal lasers. *Nature Photonics*, 4(10), 676–685. DOI: [10.1038/nphoton.2010.184](https://doi.org/10.1038/nphoton.2010.184)

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-22>

УДК 004.75

Пригода Андрій Ярославович, аспірант, асистент

<https://orcid.org/0000-0003-3774-4583>

Державний торговельно-економічний університет Київ, Україна

ОЦІНКА ЕФЕКТИВНОСТІ ПРОЄКТУ РОЗРОБКИ ТА ВПРОВАДЖЕННЯ CRM-СИСТЕМИ НА ОСНОВІ МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ

Пригода А.Я. Оцінка ефективності проєкту розробки та впровадження сrm-системи на основі мікросервісної архітектури. У статті проведено оцінку ефективності проєкту розробки та впровадження CRM-системи на основі мікросервісної архітектури. Описано принципи та сутність CRM-системи. Детально розглянуто види мікросервісів, описано технологію переходу між будівельними блоками архітектури. Наголошено, що застосування мікросервісів у проєктуванні та впровадженні систем має багато переваг, таких як автономність, стабільний зв'язок, можливість компонування, масштабованість, стійкість до відмов. Описано принципи деталізації MSA, які сприяють узгодженню функціональності мікросервісу з однією окремою можливістю, яку сервіс надає загальній архітектурі програмного забезпечення. Підкреслено, що мікросервісна архітектура дозволяє уникнути більшості труднощів, які є в монолітній архітектурі пропонуючи концепцію часткового розгортання, коли виробничий сервер втрачає лише частину функціоналу. Зазначається, що додавання нового сервісу також не впливає на працездатність всієї хмари. Окреслено основні переваги мікросервісної архітектури, які включають в себе: простоту, зосередження на бізнес-функціоналі, покращену продуктивність і швидкість, масштабованість і гнучкість, автономні, багатопрофільні команди. Визначено недоліки мікросервісної архітектури такі як, збільшена складність, витрати, більш високий попит ресурсу, безпека, надійність, неузгодженість. Наголошено, що мікросервісна архітектура передбачає розподіл відповідальності між командами розробки за цілком природними межами: за межами самих сервісів. Так, у віданні кількох команд розробки знаходяться непересічні Набори мікросервісів, що забезпечує високу ізоляцію роботи. Додатковою приємною перевагою архітектури мікросервісу є можливість вибору різних технологій у різних мікросервісах. Загалом, підкреслено, що мікросервісна архітектура CRM-системи дозволяє підвищити надійність і стабільність функціонування за рахунок відносної автономності її окремих мікросервісних компонентів. Архітектура мікросервісу CRM-системи має високу масштабованість, у тому числі за рахунок використання хмарних технологій.

Ключові слова. CRM-система, мікросервіс, архітектура, проєкт, оцінка, ефективність, розробка.

Pryhoda A. Evaluation Of The Effectiveness Of The Crm System Development And Implementation Project Based On Microservice Architecture. The article evaluates the effectiveness of the CRM system development and implementation project based on microservice architecture. The principles and essence of the CRM system are described. The types of microservices are considered in detail, the technology of transition between building blocks of architecture is described. It is emphasized that the use of microservices in the design and implementation of systems has many advantages, such as autonomy, stable communication, the possibility of composition, scalability, resistance to failures. MSA's granularity principles are described, which help to align the functionality of a microservice with a single capability that the service provides to the overall software architecture. It is emphasized that the microservice architecture avoids most of the difficulties that exist in the monolithic architecture by offering the concept of partial deployment, when the production server loses only a part of the functionality. It is noted that adding a new service also does not affect the performance of the entire cloud. The main advantages of microservice architecture are outlined, which include: simplicity, focus on business functionality, improved performance and speed, scalability and flexibility, autonomous, multidisciplinary teams. Disadvantages of microservice architecture are identified, such as increased complexity, costs, higher resource demand, security, reliability, inconsistency. It is emphasized that the microservice architecture provides for the distribution of responsibility between development teams along completely natural boundaries: beyond the boundaries of the services themselves. So, several development teams have unique sets of microservices, which ensures high isolation of work. An additional nice benefit of microservice architecture is the ability to choose different technologies in different microservices. In general, it is emphasized that the microservice architecture of the CRM system allows to increase the reliability and stability of functioning due to the relative autonomy of its individual microservice components. The microservice architecture of the CRM system is highly scalable, including due to the use of cloud technologies.

Keywords. CRM system, microservice, architecture, project, evaluation, efficiency, development.

Вступ та постановка проблеми дослідження. Архітектура програмного забезпечення системи зображує організацію або структуру системи та надає пояснення того, як вона веде себе. Система являє собою сукупність компонентів, які виконують певну функцію або набір функцій. Іншими словами, архітектура програмного забезпечення забезпечує міцну основу, на якій можна будувати програмне забезпечення [1]. Серія архітектурних рішень і компромісів впливають на якість, продуктивність, ремонтпридатність і загальний успіх інформаційної системи. Неврахування поширених проблем і довгострокових наслідків може поставити систему під загрозу. Існує безліч моделей і принципів архітектури високого рівня, які зазвичай використовуються в сучасних інформаційних системах. Їх часто називають архітектурними стилями. Архітектура програмної системи рідко обмежується одним архітектурним стилем. Натомість комбінація стилів часто складає повну систему [2].

В наш час конкуренція на ринку досить висока та при ситуації взаємодії з клієнтами кожна компанія замислюється про впровадження CRM систем, а саме використання хмарних CRM, які доступні не тільки великим компаніям та й малим підприємцям. Концепція CRM (Customer Relationship Management – система управління взаємовідносинами з клієнтами) реалізується шляхом застосування програмного забезпечення для досягнення наступних цілей: – автоматизації стратегій взаємодії із клієнтами; – оптимізації надання послуг; – отримання інформації про клієнтів та історію взаємовідносин з ними; – удосконалення процедур керування процесами спрямованих на надання клієнту повного пакету послуг відповідно до умов контракту; – забезпечення контролю результатів роботи.

Таким чином, система управління взаємовідносинами з клієнтами або CRM-система характеризується як інструмент підвищення ефективності взаємодії підприємства з клієнтами. Сьогодні CRM-системи є невід'ємним інструментом діяльності багатьох світових підприємств, що дозволяє їм набагато ефективніше встановлювати та розвивати відношення з клієнтами.

В умовах сьогодення мікросервісна архітектура відома як успішна та перспективна архітектура для CRM-систем. Застосування мікросервісів у проектуванні та впровадженні систем має багато переваг, таких як автономність, стабільний зв'язок, можливість компонування, масштабованість, стійкість до відмов. Однак складність викликів між мікросервісами призводить до проблем із безпекою, доступністю та керуванням даними під час виконання систем. Щоб вирішити ці проблеми, в останні роки різні дослідники та розробники зосередилися на використанні мікросервісних шаблонів у реалізації систем на основі мікросервісів. Патерни мікросервісів є результатом успішного досвіду розробників у вирішенні типових проблем у системах на основі мікросервісів. Однак досі не було надано жодних вказівок для поглибленого розуміння шаблонів мікросервісів і того, як їх застосувати до реальних CRM-систем.

Аналіз останніх досліджень і публікацій. Наукові діячі сьогодення внесли значний вклад у розробку та впровадження CRM-системи на основі мікросервісної архітектури. Був проведений ряд досліджень для вирішення даної проблеми.

У роботі [3] доведено, що архітектура мікросервісу забезпечує масштабованість, надійність і орієнтований на користувача підхід, який узгоджується з динамічною природою клінічних робочих процесів. Дослідження базуються на прикладній методології дослідження, одночасно використовуючи підхід дизайнерського мислення для ітеративної розробки масштабованої базової системи, яка може врахувати особливості суб'єктивних даних пацієнтів і полегшити впровадження систем CBR у CDSS.

Дослідження [4] описує практичне дослідження переходу моноліту до архітектури мікросервісу, де модульна монолітна архітектура використовується як проміжний крок. Вплив на зусилля з міграції та продуктивність вимірюється для обох кроків. Поточний стан техніки аналізує міграцію монолітних систем до архітектури мікросервісів, зазначається, що зусилля з міграції та проблеми з продуктивністю вже значні при переході до модульного моноліту. Таким чином, для кожного з кроків встановлено чітке розмежування, яке може інформувати архітекторів програмного забезпечення про планування міграції монолітних систем. Зокрема, розглянуто компроміси виконання всього процесу міграції або просто переходу на модульний моноліт.

Тим не менш, окрім зазначеної вище документації, варто зазначити праці ще наступних науковців: Міліч Мілош, Ніколіч Драгана [5], Лі Сяочжоу, Калефато Фабіо, Ленардуцці Валентина, Тайбі Давіде [6], Чжен Лін, Вей Бо [7], Хамза Мухаммад [8], Чаплія Олег, Клім Галина [9], Темоор Музаффар [10], Маркес Гастон, Тарамаско Карла, Астудільо Ернан, Зальк Вінсент, Істрате Дан [11], Лі Вей [12], Існаені Рискулла, Утама Нур, Суаканто Сінунг [13], Странд Рубен, Крістенсен Ларс, Петручі Лауре [14], Браубах Ларс, Джандер Кай, Покар Олександр [15] та інших.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з розробкою та впровадженням CRM-системи на основі мікросервісної архітектури, все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Мета дослідження. Метою даного дослідження є оцінка ефективності проєкту розробки та впровадження CRM-системи на основі мікросервісної архітектури.

Викладення основного матеріалу дослідження. Архітектура мікросервісу (MSA) – це новий підхід до реалізації сервісної архітектури програмного забезпечення (SBSA – Service-based Software Architecture) [16]. Загалом, термін «Мікросервісна архітектура» вперше використаний на конференції Cloud Computing Expo в 2005-му році Пітером Роджерсом, а почав набирати популярність у практиці та дослідженнях у 2014 та 2015 роках відповідно. Архітектура є сучасним

представленням SOA (Service-Oriented Architecture) і являє собою де-факто сучасний стандарт написання сучасних розподілених обчислювальних систем. Починаючи з 2012-го року термін набув широкого поширення, і саме з цього моменту часу багато команд підтримки веб-додатків стали розбивати свої моноліти на мікросервіси [17]. Отже, SOA та MSA мають спільне те, що вони використовують сервіси як функціональні будівельні блоки для реалізації архітектур розподіленого програмного забезпечення. Таким чином, вони стикаються з проблемами, властивими поняттю сервісу та пов'язаними, наприклад, з налаштуванням можливостей, композицією та координацією для виконання завдань і розгортанням середовища виконання. Однак мікросервіс MSA демонструє характеристики, які відрізняють його від сервісу SOA.

Деталізація MSA сприяє узгодженню функціональності мікросервісу з однією окремою можливістю, яку сервіс надає загальній архітектурі програмного забезпечення. SOA явно не передбачає певного ступеня деталізації сервісу. Натомість може реалізовувати детальну функціональність, специфічну для конкретної програми, але також враховувати реалізацію бізнес-процесів для всієї системи.

Крім того, SOA спрямований на максимізацію повторного використання послуг. MSA та покладається на адаптацію сервісів і прагне зберегти кількість мікросервісів. Тобто зменшити зв'язок між мікросервісами та підвищити автономію та ізоляцію мікросервісів з метою сприяння масштабованості та модифікованості архітектури [18]. Спільні бібліотеки – це підхід до зменшення повторного використання мікросервісів шляхом переміщення незалежних від сервісів функціональних можливостей до повторно використовуваних артефактів [19].

Окрім доступних операцій сервісу, інтерфейси також визначають структуру та формати обмінюваних даних. SOA зазвичай використовує Enterprise Service Bus (ESB) як платформу інтеграції послуг [20]. ESB може перетворювати дані, передані однією службою, у структури та формати, які можна інтерпретувати іншими несумісними службами. MSA, з іншого боку, не використовує складні засоби інтеграції для абстрагування від структур даних або форматів, передбачених інтерфейсами. Натомість мікросервіси самі відповідають за структуру даних і перетворення формату за необхідності. Знову ж таки, обґрунтуванням цієї характеристики MSA є зменшення зв'язку та підвищення автономності та ізоляції обслуговування.

Окрім перетворення структури даних і формату, ESB також можуть трансформувати протоколи технічного зв'язку, наприклад, MQTT [21] у SOAP і навпаки. Таким чином, SOA може забезпечити взаємодію, навіть якщо служби використовують різні технічні протоколи зв'язку. Навпаки, MSA делегує відповідальність за такі перетворення мікросервісам. Щоб уникнути або принаймні пом'якшити складність, спричинену перетвореннями протоколів, архітектури мікросервісів зазвичай покладаються щонайбільше на два протоколи – один для взаємодії мікросервісів «один-до-одного» (зазвичай синхронний), а інший – для мікросервісу «один-до-багатьох» (зазвичай асинхронний).

Для взаємодії між сервісами однієї архітектури програмного забезпечення (взаємодія між сервісами) SOA підтримує оркестровку та хореографію як основні моделі взаємодії [22]. У сценаріях оркестровки центральний посередник, наприклад, ESB, визначає взаємодію служб. У сценаріях хореографії служби самостійно приймають рішення про свою взаємодію. MSA зазвичай використовує шаблон хореографії для взаємодії між службами, щоб збільшити автономію мікросервісів.

Для взаємодії із зовнішніми компонентами архітектури (взаємодія поза послугами) SOA зазвичай покладається на ESB для маршрутизації та, можливо, перетворення обмінюваних даних. В архітектурах мікросервісів шлюзи API можуть використовуватися для посередництва між зовнішніми компонентами архітектури та внутрішніми мікросервісами архітектури. Однак, порівняно з ESB, шлюзи API забезпечують досить прості фасади, які агрегують операції інтерфейсу вибраних мікросервісів і надають їх зовнішнім компонентам [23].

Сфера застосування SOA призначена для реалізації корпоративних і міжкорпоративних SBSA, які вимагають різноманітних гетерогенних зовнішніх систем для участі в складних бізнес-процесах. Отже, SOA особливо кваліфікується як засіб інтеграції. MSA, однак, сприяє розробці, наприклад, хмарних додатків (CNA) або SBSA з чітко визначеними потоками обробки. Крім того, MSA часто використовується для декомпозиції додатків зі зниженою масштабованістю та зручністю обслуговування. Крім того, на відміну від SOA, MSA зазвичай зосереджується на розробці або міграції окремих програмних систем.

Порівняно з SOA передбачається, що MSA краще відповідає практичним потребам практиків, які займаються розробкою SBSA, наприклад, архітекторів програмного забезпечення та розробників. Це припущення викликано тим фактом, що, окрім SOA, MSA виникла виключно з практики застосування. Поки академічні кола не почали досліджувати це в 2015 році, MSA керувалося виключно потребами галузі.

Крім того, вважається, що MSA демонструє нижчий рівень складності, ніж SOA, це сприйняття значною мірою нав'язано такими характеристиками MSA:

MSA зводить до мінімуму спільне використання функцій між службами. Хоча ця характеристика підвищує масштабованість і придатність до обслуговування, дозволяючи службам розвиватися незалежно, вона також сприяє гнучкості.

MSA опускає поняття ESB. Як результат, архітектури мікросервісів зазвичай не інтегрують комбіновані централізовані засоби для перетворення формату даних, структури та протоколу, а також оркестровки сервісу. Однак мікросервіси мають збільшений набір обов'язків порівняно з сервісами SOA.

MSA зменшує таксономію послуг. У SOA існує декілька типів сервісів, наприклад, бізнес-сервіси, корпоративні сервіси, сервіси додатків та сервіси інтеграції. Вони відрізняються деталізацією, сферою застосування та типом можливостей. З іншого боку, MSA зазвичай складається лише з двох типів послуг, які відрізняються лише типом можливостей. Функціональні служби реалізують бізнес-логіку, а служби інфраструктури забезпечують технічні можливості, наприклад, для виявлення служби, моніторингу або автентифікації.

MSA покладається на «полегшені» технології. Наприклад, SOA часто ототожнюють із застосуванням SOAP, тоді як очікується, що MSA використовуватиме Representational State Transfer (REST) [24] для синхронної взаємодії сервісів. Враховуючи підтримку SOAP різноманітних стандартів, наприклад мови опису веб-сервісів (WSDL) і сімейства стандартів WS-*, він зазвичай сприймається як більш складний ніж REST.

Право власності на послуги в SOA зазвичай залежить від типу служби. Наприклад, розробники несуть відповідальність за реалізацію служб додатків, а архітектори програмного забезпечення компонує послуги додатків у виконуваних бізнес-послугах для всього підприємства. Тому SOA зазвичай збирає однорідні команди на основі професійного досвіду їхніх членів.

MSA, з іншого боку, сприяє передачі права власності на мікросервіс виділеній команді, яка несе повну відповідальність за всі проблеми, пов'язані з проектуванням, впровадженням і роботою мікросервісу.

Зосередженість MSA на неоднорідному складі команди та невеликому розмірі команди зрештою сприяє прийняттю DevOps [25]. DevOps – це набір практик, спрямованих на скорочення часу між фіксацією та розгортанням зміни, одночасно забезпечуючи високу якість програмної системи, що розробляється. Практики DevOps включають, наприклад, постійну інтеграцію та розгортання, а також проактивний моніторинг розглянутої системи програмного забезпечення. Хоча SOA може не перешкоджати прийняттю DevOps, вона все ж виходить за межі архітектурного стилю SOA.

Підсумовуючи, мікросервіс – це сервіс SBSA з наступними характеристиками:

- надає окремі можливості для інших компонентів архітектури. Тобто всі функціональні можливості мікросервісу вирішують одну проблему. Ця проблема має функціональний або інфраструктурний характер;

- максимально незалежний від інших компонентів з точки зору його впровадження, управління даними, тестування, розгортання та роботи;

- повністю відповідає за всі аспекти, пов'язані із взаємодією з іншими компонентами. Такі аспекти включають, наприклад, рішення щодо взаємозв'язків взаємодії, визначення протоколу зв'язку, структуру даних і перетворення формату та обробку відмов. Без серйозних технічних причин мікросервіс підтримує щонайбільше два протоколи зв'язку – один для синхронної взаємодії «один-до-одного», а інший – для асинхронної взаємодії «один-до-багатьох»;

- належить лише одній команді. Команда несе повну відповідальність за всі аспекти, пов'язані з розробкою, реалізацією та роботою мікросервісу.

Архітектура мікросервісу – це SBSA, функціональні компоненти якого, тобто компоненти, які реалізують його бізнес-логіку, є мікросервісами.

Мікросервісна архітектура дозволяє уникнути більшості труднощів, які є в монолітній архітектурі пропонуючи концепцію часткового розгортання, коли виробничий сервер втрачає лише частину функціоналу. Додавання нового сервісу також не впливає на працездатність всієї хмари. Ці переваги досягаються за рахунок того, що додаток зберігає більшу частину свого функціоналу, в той час як один мікросервіс знаходиться в стані перезавантаження або запуску.

Основними перевагами мікросервісної архітектури є наступне [26].

Простота. Програми стає легше побудувати і підтримати, коли вони розділені на ряд менших, складових фрагментів.

Зосередження на бізнес-функціоналі. Мікросервіси дозволяють будувати продукт замість проєктів.

Покращена продуктивність і швидкість. Мікросервісна архітектура вирішує проблему продуктивності і швидкості, розкладаючи програми на керовані сервіси, які швидко розвиваються.

Масштабованість і гнучкість. Кожний мікросервіс може бути написаним з використанням різних технологій.

Автономні, багатопрофільні команди. Мікросервіси дуже зручні для розподілених команд. Розвиток великої системи моноліту може бути складним, якщо співробітники працюють з підрозділами у всьому світі [27].

Найбільший недолік мікросервісної архітектури – збільшена складність. Складність мікросервісного застосування безпосередньо корелює з кількістю залучених сервісів. У цього типу архітектури є набагато більше рухомих частин, ніж у традиційних програмних систем, що вимагає додаткового зусилля, планування, і автоматизації, щоб контролювати міжсервісне спілкування, моніторинг, тестування та розгортання. Ініціатива створення програмної системи мікросервісного типу вимагатиме структурної зміни в команді. Вона вимагає зрілої та гнучкої культури DevOps. Із заснованою на мікросервісах програмою командам потрібно вміти управляти всім життєвим циклом усіх сервісів. Це часто вимагає мігруючих компетенцій та прийняття різних рішень від менеджерів і архітекторів окремих команд. Ця зміна в ієрархії може бути важкою для компанії [28].

Крім того, зв'язок між людьми і командами стає набагато більш складним, оскільки у команд може не бути видимості усієї картини і того, як окремі послуги повинні працювати одна з одною, щоб створити повноцінне програмне забезпечення. Організація повинна також визначити, чи володіють їх люди навичками і мають необхідний досвід, щоб розробляти засновану на мікросервісах програмну систему. Оскільки команда може бути відповідальною за один сервіс, то розробники повинні бути добре обізнані про розвиток, розгортання, тестування та контроль застосування. Останньою вимогою буде мати хоча б одного DevOps розробника з відповідними навичками на команду.

Серед інших недоліків мікросервісів – витрати. Сервіси повинні будуть спілкуватися один з одним, що призведе до великої кількості запитів. Ці віддалені запити призводять до більш високих цін, пов'язаних з мережевим часом очікування і обробки, порівняно з традиційною архітектурою. Розробники повинні прикласти максимальних зусиль, щоб скоротити кількість запитів [29].

Інший фактор збільшеної вартості – більш високий попит ресурсу, оскільки кожен сервіс вимагатиме свого власного середовища і центрального процесора. Це – вимога, щоб зберегти кожен сервіс ізольованим. Крім того, через кожний сервіс, що використовує власну мову і технології, розробка додатку та неоднорідність архітектури можуть збільшити ресурси, які організація витрачає на управління і обслуговування [30].

Мікросервіси ставлять під сумнів рівень проблеми безпеки. Мікросервіси можуть поставити величезні проблеми безпеки через збільшення кількості комунікації міжсервісного спілкування по мережі. Всі ці взаємодії створюють можливість зовнішнім елементам отримати доступ до системи. Мікросервіси можуть стати причиною низької продуктивності. Наприклад, якщо сервіс викликає п'ять віддалених сервісів, кожен з яких викликає ще п'ять віддалених сервісів, то загальний час затримки може стати надмірно великим. Для вирішення цієї проблеми було придумано кілька способів. Можна робити менше віддалених викликів, збільшивши їх деталізацію, але це ускладнює модель програмування, так як тепер потрібно думати про те, як об'єднувати взаємодії між сервісами. Ще одним варіантом вирішення може буде асинхронність. Якщо зробити кілька асинхронних викликів одночасно, то загальний час очікування буде визначатися самим повільним викликом, а не сумою всіх. Це може дати великий вигравш в продуктивності, але тягне за собою іншу проблему – асинхронність складно програмувати і ще важче налагоджувати.

Ще одна проблема при будівництві програмної системи мікросервісного типу – надійність. Очікується, що внутрішні виклики завжди будуть працювати, але віддалений виклик може відмовити в будь-який момент. З ростом кількості мікросервісів потенційних точок відмови стає більше. Розробники намагаються проектувати систему з урахуванням відмов. Існують техніки, які потрібні для реалізації асинхронної взаємодії, вони добре підходять для роботи з відмовами, підвищуючи відмовостійкість. Однак, це не повна компенсація, тому що як і раніше залишаються додаткові складності з виявлення наслідків відмови для кожного зовнішнього виклику.

Мікросервіси схильні до проблеми неузгодженості. У монолітній системі можуть бути оновленими великою кількістю об'єктів в одній транзакції. З мікросервісами потрібно кілька ресурсів для оновлення та розподілення транзакції.

Як результат, розробники повинні пам'ятати про можливість неузгодженості і думати про те, як визначати моменти розсинхронізації, перш ніж програмувати щось, що потрібно буде покращувати з часом. Один з недоліків – складний онбордінг (процес адаптації користувача до продукту). Існує точка зору, що оскільки кожен сервіс невеликий, його простіше зрозуміти, але небезпека полягає в тому, що складність не зникає, вона просто зміщується на взаємозв'язку між сервісами. Це може маскувати збільшення складності експлуатації, наприклад, труднощі при налагодженні функцій, які включають кілька сервісів. Якісний вибір меж сервісів зменшує цю проблему, інший – робить її набагато складніше.

У сукупності з підходом безперервної інтеграції, що прийшов зі світу DevOps, мікросервісна архітектура дозволяє вводити у виробничу фазу повністю протестований (пройшов модульні тести, інтеграційні тести і end-to-end тести) функціонал. Написання якісних масштабних тестів для монолітного додатку часто виявляється на порядок більш важким завданням в силу змішування функціоналу в одному великому додатку [31].

Мікросервіси за своєю суттю дотримуються філософії UNIX: «робіть одне, але добре». Кожен сервіс має вузьку зону відповідальності, який якісно вирішує малий обсяг покладених на нього завдань. Нові розробники легше вливаються в робочий процес, оскільки вивчення і підтримка вихідного коду об'ємом в 2–3 тисячі рядків є куди більш простим завданням, ніж підтримка 30–40 тисяч рядків коду. Коли мова заходить про вихідний код, не можна не згадати філософію слабкої зв'язності мікросервісів. Використання повідомлень для взаємодії сервісів між собою дозволяє розробнику зосередитися на вирішенні поточного завдання, нехтуючи випадковим пошкодженням вже існуючий функціонал, яку часто відчують розробники монолітних додатків.

Мікросервісна архітектура передбачає розподіл відповідальності між командами розробки за цілком природними межами: за межами самих сервісів. Так, у віданні кількох команд розробки знаходяться непересічні Набори мікросервісів, що забезпечує високу ізоляцію роботи. По суті, кожна команда вільна робити в своїх сервісах будь-які бажані зміни, якщо ці зміни не порушують наступних умов [32]:

Сервіс зберігає працездатність і виконує покладені на нього функціональні обов'язки, обумовлені проектними керівниками.

Дотримано виконання контрактів споживачів. Ці контракти наказують, яку інформацію даний сервіс повинен надавати іншим сервісам за допомогою REST API або асинхронних повідомлень, що передаються в рамках повсюдно впровадженої event-driven архітектури.

З вище сказаного можна зробити висновок, що кожен мікросервіс зобов'язаний бути незалежним компонентом. Так як кожен мікросервіс працює в своєму процесі, він має чітко винести інтерфейс взаємодії з ним (API). Інші структурні елементи системи можуть взаємодіяти з сервісом лише через API, тому мінімізація зв'язків – один з найважливіших процесів в плануванні такої архітектури.

Додатковою приємною перевагою архітектури мікросервісу є можливість вибору різних технологій у різних мікросервісах. Наприклад, один сервіс може використовувати базу даних Neo4j, що працює з графами даних, прекрасно підходить для зберігання соціальних даних, а інший – класика PostgreSQL. І при розумному розбитті розв'язуваного веб-додатком завдання на контексти (області відповідальності, що характеризуються сильним зчепленням за змістом, наприклад, контекст користувачів, контекст електронних листів, контекст замовлень) простота підтримки програми і його продуктивність можуть відчутно зрости. Ця ж перевага дозволяє командам експериментувати з новими технологіями, які будуть впроваджені і в інші сервіси, якщо запропоновані рішення виявляться вдалими.

Висновки та перспективи подальших досліджень. Мікросервісна архітектура CRM-системи дозволяє підвищити надійність і стабільність функціонування за рахунок відносної автономності її окремих мікросервісних компонентів. Архітектура мікросервісу CRM-системи має високу масштабованість, у тому числі за рахунок використання хмарних технологій. Складність таких систем полягає в підтримці стану, синхронізації та узгодженості даних але вона є максимально масштабованою та володіє підвищеною здатністю ефективно виконувати всі функції покладені на неї.

Список бібліографічного опису

1. Аль Равашдех Лейт Ахмед Мустафа. Оцінка похибок динамічних нейронних мереж для вимірювальних систем. *Метрологія та прилади*. 2018. № 3. С. 33–35.
2. Бойчук В. О. Сучасні штучні нейронні мережі та підходи до їх моделювання. *Вимірювальна та обчислювальна техніка в технологічних процесах*. 2014. № 4. С. 216–219.
3. Jaiswal Amar, Meisingset Ingebrigt. A Microservice-Based Architecture for Clinical Decision Support System for Addressing Non-Specific Musculoskeletal Disorders: The SupportPrim Project. 2024. DOI: 10.21203/rs.3.rs-4122773/v1.
4. Faustino Diogo, Gonçalves Nuno, Portela Manuel, Silva António. Stepwise migration of a monolith to a microservice architecture: Performance and migration effort evaluation. *Performance Evaluation*. 2024. № 164. P. 102-411. DOI: 10.1016/j.peva.2024.102411.
5. Milić Miloš, Nikolić Dragana. Development of a Quality-Based Model for Software Architecture Optimization: A Case Study of Monolith and Microservice Architectures. *Symmetry*. 2022. № 14. 1824 p. DOI: 10.3390/sym14091824.
6. Li Xiaozhou, Calefato Fabio, Lenarduzzi Valentina, Taibi Davide. Toward Collaboration Optimization in Microservice Projects based on Developer Personalities. 2024.
7. Zheng Ling, Wei Bo. Application of microservice architecture in cloud environment project development. *MATEC Web of Conferences*. 2018. 189 p. 03023. DOI: 10.1051/mateconf/201818903023.
8. Hamza Muhammad. Transforming Monolithic Systems to a Microservices Architecture. *ACM SIGSOFT Software Engineering Notes*. 2023. №48. P. 67-69. DOI: 10.1145/3573074.3573091.
9. Chaplia Oleh, Klym Halyna. NODE.JS PROJECT ARCHITECTURE WITH SHARED DEPENDENCIES FOR MICROSERVICES. *Measuring Equipment and Metrology*. 2023. №84. P. 53-58. DOI: 10.23939/istcmtm2023.03.053.
10. Temoor Muzaffar. Architecture for Microservice Based System. *A Report*. 2020. DOI: 10.13140/RG.2.2.17340.16004/1.
11. Márquez Gastón, Taramasco Carla, Astudillo Hernán, Zalc Vincent, Istrate Dan. Involving Stakeholders in the Implementation of Microservice-Based Systems: A Case Study in an Ambient-Assisted Living System. *IEEE Access*. 2021. PP. 1-1. DOI: 10.1109/ACCESS.2021.3049444.
12. Li Wei. Design and Implementation of Flight Inquiry and Booking System Based on Microservice Architecture. *Highlights in Science, Engineering and Technology*. 2023. № 56. P. 697-703. DOI: 10.54097/hset.v56i.10836.
13. Isnaeni Rizqullah, Utama Nur, Suakanto Sinung. Backend Development of a Microservice-Based Website Application for Public Issue Reporting: Case Study in People Representative Council. *Journal La Multiapp*. 2024. № 5. P. 63-77. DOI: 10.37899/journallamultiapp.v5i2.1148.
14. Strand Ruben, Kristensen Lars, Petrucci Laure. Development and Verification of a Microservice Architecture for a Fire Risk Notification System. 2023. DOI: 10.1007/978-3-662-68191-6_2.
15. Braubach Lars, Jander Kai, Pokahr Alexander. (2022). Macro Architecture for Microservices: Improving Internal Quality of Microservice Systems. 2022. DOI: 10.1007/978-3-030-96627-0_10.
16. Jo Yong, Cho Se, Choi Byoungwook. Towards a ROS2-based software architecture for service robots. *Bulletin of Electrical Engineering and Informatics*. 2023. №12. P.3027-3038. DOI: 10.11591/eei.v12i5.5590.
17. Fajar Ahmad, Limonthy Stanley, Handopo Josua, Purnawan Fandy, Kesuma Adidharma. (2023). System Architecture for IT Talent Ecosystem Using Service Oriented Approach. *HighTech and Innovation Journal*. 2023. № 4. P. 739-748. DOI: 10.28991/HIJ-2023-04-04-03.
18. Graves Tom. Basics – Service-Oriented Architecture. 2023. DOI: 10.1007/978-1-4842-9189-4_2.
19. Gottardi Thiago, Braga Rosana. Run-time Adaptable Service Oriented Architecture in the Context of Repository Systems. *Anais do Workshop em Modelagem e Simulação de Sistemas Intensivos em Software (MSSiS)*. 2023. P.21-30. DOI: 10.5753/mssis.2023.235434.
20. Bais Aditya, Mishra Varun. Analysis of Data Functionality in Enterprise Service Bus. 2017. DOI: 10.4018/978-1-5225-2157-0.ch005.
21. Shah Pujan. MQTT Systems: A Survey. *International Journal for Research in Applied Science and Engineering Technology*. 2024. № 12. P. 1063-1067. DOI: 10.22214/ijraset.2024.59000.
22. Karande Aarti, Karande Milind, Meshram Bhushan. Choreography and Orchestration using Business Process Execution Language for SOA with Web Services. *International Journal of Computer Science Issues*. 2011. № 8.
23. Roshen Waseem. Service oriented architecture enterprise service bus with universal ports. 2014.
24. Arsana I Nyoman, Purnawati Niwayan. Analysis of the implementation of Restful (Representational State Transfer) web services for data integration between information systems. 2023. № 7. DOI: 10.35335/mantik.v7i2.4016.
25. Tatineni Sumanth. Applying DevOps Practices for Quality and Reliability Improvement in Cloud-Based Systems. 2023. DOI: 10.13140/RG.2.2.25688.88326.
26. Kamil Figura. Why you should choose the microservices architecture? Режим доступу до ресурсу: <https://pretius.com/blog/benefits-of-microservices/> Дата звернення: 14.05.2024

27. Antipattern-Based Problem Injection for Assessing Performance and Reliability Evaluation Techniques. *IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. 2016. С. 64–70.
28. Тарарака В. Д. Архітектура комп'ютерних систем: навчальний посібник. Житомир: ЖДТУ, 2018. 383 с.
29. Rajesh RV, Spring Microservices. Packt Publishing, 2016. 436 с.
30. AlQershhi N., Mokhtar S.S.M. Innovative CRM and Performance of SMEs: The Moderating Role of Relational Capital. *Journal of Open Innovation*. 2020, № 6, Pp. 155. <https://doi.org/10.3390/joitmc6040155>.
31. Anna K., Lazaros P. Exploring E-CRM Implementation in Sport Tourism Hotels in Peloponnese. In: *Springer Proceedings in Business and Economics*. 2020. Pp. 611–627. https://doi.org/10.1007/978-3-030-36342-0_47.
32. Wilson H., Daniel E., McDonald M. Factors for success in customer relationship management (CRM) systems. *Journal of marketing management*, 2002, № 18(1-2), Pp. 193-219.

References:

1. Al Rawashdeh Leit Ahmed Mustafa. Fehlerschätzung dynamischer neuronaler Netze für Messsysteme. *Messtechnik und Geräte*. 2018. Nr. 3. S. 33
2. Boychuk V. O. Moderne künstliche neuronale Netze und Ansätze zu ihrer Modellierung. Mess- und Rechengerte in technologischen Prozessen. 2014. Nr. 4. S. 216–219.
3. Jaiswal Amar, Meisingset Ingebrigt. A Microservice-Based Architecture for Clinical Decision Support System for Addressing Non-Specific Musculoskeletal Disorders: The SupportPrim Project. 2024. DOI: 10.21203/rs.3.rs-4122773/v1.
4. Faustino Diogo, Gonçalves Nuno, Portela Manuel, Silva António. Stepwise migration of a monolith to a microservice architecture: Performance and migration effort evaluation. *Performance Evaluation*. 2024. № 164. P. 102-411. DOI: 10.1016/j.peva.2024.102411.
5. Milić Miloš, Nikolić Dragana. Development of a Quality-Based Model for Software Architecture Optimization: A Case Study of Monolith and Microservice Architectures. *Symmetry*. 2022. № 14. 1824 p. DOI: 10.3390/sym14091824.
6. Li Xiaozhou, Calefato Fabio, Lenarduzzi Valentina, Taibi Davide. Toward Collaboration Optimization in Microservice Projects based on Developer Personalities. 2024.
7. Zheng Ling, Wei Bo. Application of microservice architecture in cloud environment project development. *MATEC Web of Conferences*. 2018. 189 p. 03023. DOI: 10.1051/mateconf/201818903023.
8. Hamza Muhammad. Transforming Monolithic Systems to a Microservices Architecture. *ACM SIGSOFT Software Engineering Notes*. 2023. №48. P. 67-69. DOI: 10.1145/3573074.3573091.
9. Chaplia Oleh, Klym Halyna. NODE.JS PROJECT ARCHITECTURE WITH SHARED DEPENDENCIES FOR MICROSERVICES. *Measuring Equipment and Metrology*. 2023. №84. P. 53-58. DOI: 10.23939/istcmtm2023.03.053.
10. Temoor Muzaffar. Architecture for Microservice Based System. *A Report*. 2020. DOI: 10.13140/RG.2.2.17340.16004/1.
11. Márquez Gastón, Taramasco Carla, Astudillo Hernán, Zalc Vincent, Istrate Dan. Involving Stakeholders in the Implementation of Microservice-Based Systems: A Case Study in an Ambient-Assisted Living System. *IEEE Access*. 2021. PP. 1-1. DOI: 10.1109/ACCESS.2021.3049444.
12. Li Wei. Design and Implementation of Flight Inquiry and Booking System Based on Microservice Architecture. *Highlights in Science, Engineering and Technology*. 2023. № 56. P. 697-703. DOI: 10.54097/hset.v56i.10836.
13. Isnaeni Rizqullah, Utama Nur, Suakanto Sinung. Backend Development of a Microservice-Based Website Application for Public Issue Reporting: Case Studyn in People Representative Council. *Journal La Multiapp*. 2024. № 5. P. 63-77. DOI: 10.37899/journallamultiapp.v5i2.1148.
14. Strand Ruben, Kristensen Lars, Petrucci Laure. Development and Verification of a Microservice Architecture for a Fire Risk Notification System. 2023. DOI: 10.1007/978-3-662-68191-6_2.
15. Braubach Lars, Jander Kai, Pokahr Alexander. (2022). Macro Architecture for Microservices: Improving Internal Quality of Microservice Systems. 2022. DOI: 10.1007/978-3-030-96627-0_10.
16. Jo Yong, Cho Se, Choi Byoungwook. Towards a ROS2-based software architecture for service robots. *Bulletin of Electrical Engineering and Informatics*. 2023. №12. P.3027-3038. DOI: 10.11591/eei.v12i5.5590.
17. Fajar Ahmad, Limonthy Stanley, Handopo Josua, Purnawan Fandy, Kesuma Adidharma. (2023). System Architecture for IT Talent Ecosystem Using Service Oriented Approach. *HighTech and Innovation Journal*. 2023. № 4. P. 739-748. DOI: 10.28991/HIJ-2023-04-04-03.
18. Graves Tom. Basics – Service-Oriented Architecture. 2023. DOI: 10.1007/978-1-4842-9189-4_2.
19. Gottardi Thiago, Braga Rosana. Run-time Adaptable Service Oriented Architecture in the Context of Repository Systems. *Anais do Workshop em Modelagem e Simulação de Sistemas Intensivos em Software (MSSiS)*. 2023. P.21-30. DOI: 10.5753/mssis.2023.235434.
20. Bais Aditya, Mishra Varun. Analysis of Data Functionality in Enterprise Service Bus. 2017. DOI: 10.4018/978-1-5225-2157-0.ch005.
21. Shah Pujan. MQTT Systems: A Survey. *International Journal for Research in Applied Science and Engineering Technology*. 2024. № 12. P. 1063-1067. DOI: 10.22214/ijraset.2024.59000.
22. Karande Aarti, Karande Milind, Meshram Bhushan. Choreography and Orchestration using Business Process Execution Language for SOA with Web Services. *International Journal of Computer Science Issues*. 2011. № 8.
23. Roshen Waseem. Service oriented architecture enterprise service bus with universal ports. 2014.
24. Arsana I Nyoman, Purnawati Niwayan. Analysis of the implementation of Restful (Representational State Transfer) web services for data integration between information systems. 2023. № 7. DOI: 10.35335/mantik.v7i2.4016.
25. Tatineni Sumanth. Applying DevOps Practices for Quality and Reliability Improvement in Cloud-Based Systems. 2023. DOI: 10.13140/RG.2.2.25688.88326.

26. Kamil Figura. Why you should choose the microservices architecture? Режим доступу до ресурсу: <https://pretius.com/blog/benefits-of-microservices/> Дата звернення: 14.05.2024
 27. Antipattern-Based Problem Injection for Assessing Performance and Reliability Evaluation Techniques. *IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW)*. 2016. С. 64–70.
 28. Тарарака В. Д. Архітектура комп'ютерних систем: навчальний посібник. Житомир: ЖДТУ, 2018. 383 с.
 29. Rajesh RV, Spring Microservices. Packt Publishing, 2016. 436 с.
 30. AlQershi N., Mokhtar S.S.M. Innovative CRM and Performance of SMEs: The Moderating Role of Relational Capital. *Journal of Open Innovation*. 2020, № 6, Pp. 155. <https://doi.org/10.3390/joitmc6040155>.
 31. Anna K., Lazaros P. Exploring E-CRM Implementation in Sport Tourism Hotels in Peloponnese. In: *Springer Proceedings in Business and Economics*. 2020. Pp. 611–627. https://doi.org/10.1007/978-3-030-36342-0_47.
- Wilson H., Daniel E., McDonald M. Factors for success in customer relationship management (CRM) systems. *Journal of marketing management*, 2002, № 18(1-2), Pp. 193-219

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-23>

УДК 004.942

Решетняк Валентин Валентинович, аспірант

<https://orcid.org/0000-0002-2865-8678>

Фауре Еміль Віталійович, д.т.н., професор

<https://orcid.org/0000-0002-2046-481X>

Черкаський державний технологічний університет, м. Черкаси, Україна

EYE-TRACKING AS A TOOL FOR RESEARCHING USER BEHAVIOR

Reshetniak V., Faure E. Eye-tracking as a tool for researching user behavior. This article discusses eye-tracking technology, one of the critical technologies in analyzing user behavior when interacting with computer systems, and has a wide range of practical applications. The current capabilities and limitations of existing methods and equipment for eye tracking were reviewed and analyzed, and problems and promising areas for further research were identified. The central hardware systems were studied, such as desktop trackers, mobile glasses for eye-tracking, systems for immersive environments, and systems based on embedded cameras. This research also includes analyzing and classifying eye localization methods: eye shape-based, feature-based, and appearance-based. Also were investigated modern gaze-tracking methods, which are crucial to analyzing eye movement and determining the user's point of view. Feature-based methods provide accuracy but have limitations in strong lighting conditions. Model-based methods require high image resolution and calibration to measure gaze direction using facial characteristics. Cross-ratio-based methods do not require calibration but require additional lighting. Appearance-based methods are effective with low-quality images but have limitations concerning head position. The study indicates significant potential for improving gaze-tracking technologies to enhance user interfaces and increase accessibility. It emphasizes the need to develop adaptive methods that function in different environmental conditions and hardware limitations.

Keywords: human oculomotor system, eye-tracking, eye-tracker, user experience, eye gaze estimation, pattern recognition.

Решетняк В.В., Фауре Е.В. Відслідковування погляду як інструмент дослідження поведінки користувача. В цій статті розглядається технологія відслідковування погляду, яка є однією з ключових в аналізі користувацької поведінки при взаємодії з комп'ютерними системами й має широке практичне застосування. Було розглянуто і проаналізовано сучасні можливості та обмеження існуючих методів і обладнання для айтрекінгу, а також на визначено проблеми та перспективні напрямки для подальших досліджень. Були розглянуті основні апаратні комплекси, такі як настільні трекери, мобільні окуляри для айтрекінгу, системи для імерсивних середовищ, а також системи засновані на вбудованих камерах. Це дослідження також включає аналіз та класифікацію методів локалізації очей: на основі форми очей, на основі особливостей форми та на основі оточення області очей. Також було досліджено сучасні методи гейз-трекінгу, які є ключовими для аналізу переміщення очей і визначення точки погляду користувача. Методи на основі особливостей форми забезпечують точність, але мають обмеження при сильному освітленні. Методи на основі моделей вимагають високої роздільної здатності зображень і калібрування, але дозволяють визначати напрямок погляду за допомогою простих характеристик обличчя. Методи на основі перехресного співвідношення не потребують калібрування, але потребують додаткового освітлення. Методи на основі оточення області очей ефективні з низькоякісними зображеннями, але мають обмеження щодо положення голови. Дослідження вказує на значний потенціал вдосконалення технологій відслідковування погляду для покращення користувацьких інтерфейсів та підвищення доступності. Підкреслюється необхідність розробки адаптивних методів, які можуть функціонувати в різних умовах середовища та обмеженнях обладнання.

Ключові слова: окуло-моторна система людини, відслідковування погляду, айтрекінг, айтрекер, досвід користувача, розпізнавання образів.

Introduction.

Relevance of the study.

Eye tracking technology allows the discovery of user behavior when interacting with computer systems. According to [1], information is perceived during the moments of eye fixation between rapid eye movements (saccades). Thus, the duration of the fixation time on a stimulus can indicate the efficiency of the information-reading process. In this study, we observed the hardware and software complexes (eye trackers) to explore this efficiency. These complexes collect temporal and spatial information about gaze movement and fixation.

Since eye tracking technology allows the analysis of human behavior, it has been widely used in research in many areas. In [2], it provides an analysis of the use of eye tracking in medical research. For example, it is used to identify and research autism spectrum disorders [3], schizophrenia, depression, anxiety, Parkinson's disease, and Alzheimer's disease. [4] and [5] describes the use of eye tracking in design research. It is used to analyze the design of websites, mobile applications, and games. It is also widely used in marketing research [6]. [12] provides an overview of the application of eye tracking technology in virtual reality, [13] and [14] describe the application of this technology for augmented reality. In [15], the use of eye tracking is suggested to access mediation. It has also been proposed that this technology be used in the

military to analyze soldier fatigue [16]. There are studies of the application of eye trackers in the automotive industry, where this technology can be used to track a driver's concentration and field of view [17]. Also, recent studies have used eye tracking not only to analyze human behavior but also that of animals, for example, during their training [18].

Despite a long history of development and widespread use of the technology, there are still challenges in implementing it in some areas. An example is the design research of mobile applications using eye tracking, which requires either expensive equipment or the fixation of the smartphone's position.

The purpose of the study

This study aims to explore the current capabilities of eye tracking technology and analyze the limitations of existing methods and equipment. We observed existing hardware systems and video processing methods for localizing eyes and matching eye movements with the gaze direction. Particular emphasis is placed on identifying problematic aspects of using these technologies and identifying promising research areas. This article seeks to identify the reasons for the limitations of existing systems and propose research objectives for further improvement of eye tracking technology.

Existing hardware systems for eye tracking

General overview of hardware systems

The study of eye movement began in the mid-twentieth century, but as can be seen in the graph of mentions of the term "eye tracking" in printed sources (Fig.1), eye tracking technology has been most developed since 1990 when personal computers became widespread. In recent years, the need arose to study the users' interaction in human-computer systems in different conditions and for various tasks. As is shown on the graph, eye tracking technology has been developing rapidly in recent years due to the increased use of computer systems, including smartphones and tablets, as well as the expansion of the scope of this technology.

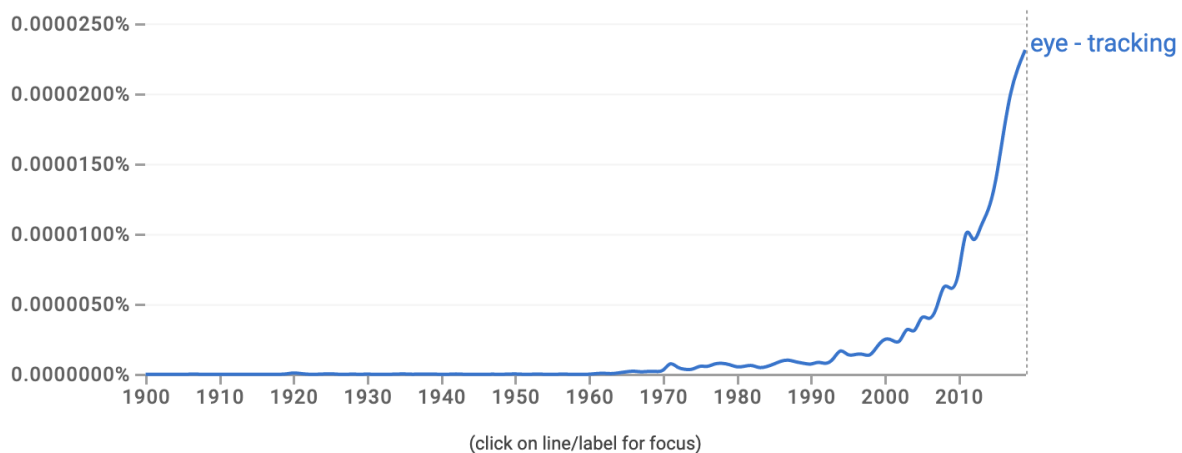


Fig. 1. Mention of the term "eye tracking" in printed publications from 1900 to 2019. Data from Google Books Ngram Viewer [19]

Software and hardware systems are used to study eye movements, which are classified according to various parameters, such as mobility, type of equipment used, areas of use, and others. In this section, the following main types of software and hardware systems will be considered, and their features, advantages, and disadvantages will be analyzed:

- Desktop eye tracking complexes;
- Mobile eye tracking glasses;
- Virtual reality eye tracking complexes.

It should also be noted that eye tracking can be performed without third-party devices, using computers and smartphones' built-in cameras (or external webcams). In this case, eye tracking is based on visible light, without using infrared light, which is used in external devices.

Desktop eye tracking complexes

Systems used with a stationary computer screen are called Desktop Eye Trackers. Such devices are the most common because they have an extended development history. [20] They consist of an infrared

light emitter (or several) and one or more digital cameras that capture images in the near-infrared spectrum directed at the user's eyes. These complexes are mounted under or above the monitor, which allows real-time reading of gaze movement and transmitting the coordinates of the focus point movement on the screen. These devices include Tobii Pro Spectrum, Tobii Pro Spark, TOBII PRO NANO, iMotions Gazepoint GP3 HD, and iMotions Smart Eye AI-X.

This type of devices also includes stationary eye tracking devices on mobile devices (smartphones and tablets). Such devices have the same principle of operation as those used with computer screens. An example of such a device is the Tobii Pro Nano, which can be used with both computers and mobile devices (Fig. 2).



Fig. 2. Using Tobii Pro Nano with a smartphone [21]

Such devices provide accurate gaze tracking results, but they limit the user's head movements, require calibration every time they are used and are not mobile in their movement, which imposes restrictions on interaction with smartphones and tablets.

Mobile eye tracking glasses

Glasses equipped with sensors for tracking eye movements have replaced massive and uncomfortable head-mounted devices, such as the ASL Headmounted Eye Tracker. [20] Such glasses use the same technology as screen-based systems, but the key feature is an additional camera directed in the direction of gaze. This feature, combined with the high mobility of such devices, allows researchers to study not only human interaction with computer screens but also interaction with the environment in real-world settings such as shops, workplaces, public transportation, and so on, which is relevant, for example, in advertising. Such devices include Pupil Invisible, Tobii Pro Glasses 3 (Fig. 3), and Viewpointssystem VPS 19.



Fig. 3. Structure of Tobii Pro Glasses 3 [22]

Devices of this type have the advantage of mobility, do not require additional calibration, and do not restrict head movement. Such devices have a more comprehensive range of practical applications, as they do not need a monitor to be looked at.

Virtual reality eye tracking complexes

With the growing popularity of immersive environments, the need for eye tracking has arisen. Such devices are usually integrated into virtual reality helmets and augmented reality glasses. Still, add-on devices are also made separately and can be connected to helmets and glasses. An example of an add-on device is the Binocular Add-on from Pupil Labs for HTC virtual reality helmets. Such systems also use the technology of reading the infrared reflection of light from the eye, which allows for calculating the direction of gaze.

Compared to other devices used primarily for user behavior research, the devices used in virtual reality helmets and augmented reality glasses have more practical applications. For example, augmented reality glasses, such as Microsoft HoloLens 2, use gaze tracking to achieve a hologram effect - the image on the glasses' screen changes according to the angle of the gaze on the screen. Another application of eye-tracking technology is Foveated Rendering, which is used in virtual reality helmets:

The foveated rendering system presents the foveal vision with full-resolution rendering and the peripheral vision with low-resolution rendering. [23]

In immersive environments, gaze tracking improves user interaction. It makes it more natural, but due to the peculiarities of such environments, additional challenges arise, such as a limited workspace or constant shifting of the user's head position.

Mobile and webcam eye tracking

The previously mentioned systems require complex and expensive equipment, which significantly limits the use of the technology. At the same time, there are developing areas of eye tracking that use built-in cameras of laptops, tablets, smartphones, and regular webcams. These areas significantly reduce the entry threshold for using the technology, although they are somewhat less accurate compared to professional eye trackers.

Several software products are currently available for tracking eyes using a computer's built-in or external webcam. For example, GazeRecorder has several software products based on webcam tracking for creating heat maps of observing websites, controlling the mouse cursor, typing, and scrolling through pages [24]. A similar example is the free myEye program, which allows users to control the cursor with their eyes. It was developed to improve usability for people with amyotrophic lateral sclerosis, who can only interact with their eyes [25]. The open-source JavaScript library WebGazer.js also uses webcam data to determine where the user is looking on the page. This library can be integrated into websites to analyze user behavior [26]. One of the leaders in the tracking market, Tobii, also has a software product that supports webcam eye tracking. The Tobii Sticky software allows the researcher to conduct eye tracking research on advertising, video, and packaging using a standard webcam [27]. The company iMotions also offers its software for webcam tracking - iMotions Online [28].

Compared to professional equipment, webcam-based eye tracking has advantages primarily due to its affordability and ease of installation. This approach significantly reduces the cost of implementing the technology, as it does not require additional equipment: most laptops are already equipped with built-in cameras, and desktop computers can easily be supplemented with a regular webcam. Software solutions for webcam tracking, such as the WebGazer.js library, can be easily integrated with existing websites. Moreover, some of these software products can be used simultaneously with other programs, which allows for the research of these programs.

Despite the advantages of webcam tracking, there are also disadvantages to this approach. Using standard built-in or regular webcams can reduce the accuracy of measurements, especially in low-light conditions, compared to specialized eye tracking equipment. There is also a fairly wide variation in the quality of cameras, which makes it impossible to predict the result of the study. This method highly depends on the distance between the camera and the user. As this distance increases, the tracking quality is getting worse, similar to changes in the position of the user's head.

As for mobile eye-tracking, several test programs are available in the Google Play and Apple AppStore that allow eye tracking but do not have any practical applications. However, the Hawkeye [29] and SeeSo [30] apps can make eye-tracking research on websites.

Like webcam eye tracking, the main advantage of mobile eye tracking is its accessibility: almost all smartphones and tablets are equipped with a front-facing camera used for research. Another significant advantage is mobility: since there is no need for stationarity, it is possible to capture the natural interaction of users with mobile interfaces, which provides more accurate behavioral data.

Mobile eye tracking has the same disadvantages as webcam eye tracking: a wide variety of cameras and lower accuracy than professional equipment. However, this approach also has disadvantages related to its use on mobile devices. The small size of smartphone screens limits the amplitude of gaze movements, which makes it difficult to track the gaze accurately. The head position during usage of mobile devices can be changed more frequently and extensively than when interacting with a laptop or monitor screen. Lighting conditions may change regularly and abruptly while using a mobile device. There are also technical limitations: not all smartphone models can provide efficient real-time video processing due to high computational loads, and active use of the camera and data processing can quickly drain a mobile device's battery. Another problem is the use of mobile tracking technology for mobile application research. Unlike websites, for which it is possible to embed the tracking function in the browser, mobile applications must embed it for each application.

Webcam and mobile eye tracking are actively developing, but several challenges must be resolved. First of all, there is a need to improve tracking accuracy, especially in low-light conditions and when using low-quality cameras. This can be achieved by further improving calibration. It is necessary to find ways to minimize the impact of external factors, such as lighting and changes in head position, which can distort the tracking data. Another challenge for mobile tracking is to develop accurate algorithms that require less computing power. This will ensure high performance even on devices with limited resources. Thus, both areas need further research and development to overcome the existing limitations.

Parameters that affect tracking quality

Most devices use infrared light reflection analysis captured by one or more cameras. Therefore, the main parameters that can affect the quality of tracking are:

- Camera resolution. Higher resolution provides more accurate eye movement data.
- Number of cameras. Since the cameras are placed at different angles to the eye, it gives the ability to recognize the direction of gaze more accurately.
- Frame rate. Higher frame rates allow for smoother tracking of gaze movements and shorter changes in gaze direction.
- Accuracy. It is the offset between the actual gaze position and what the eye tracker recorded as the gaze position [31].
- Precision. It is a measure of variation in the recorded data. It is defined as the ability of an eye tracker to reliably reproduce the same gaze point measurement from one sample to the next. Precision is calculated via the Root Mean Square (RMS) of the sampled points [32].
- Synchronization latency. It is defined as the time between the time the eye image is captured and the time the data is available in the program on the host computer. This time includes image exposure time, image acquisition and transmission time, processing time, and data transfer time to the main computer [33]. Low latency is essential for virtual and augmented reality systems devices to improve the user experience.
- Gaze recovery time. This is the time it takes for gaze tracking to resume after a temporary loss of gaze (e.g. when the participant turns their back on the eye tracker). If the device cannot detect the eyes at the last location within a few hundred milliseconds, it will start searching for eyes across the entire tracking area [34].
- Battery capacity. For autonomous devices, battery capacity is an important indicator, which can impose restrictions on the tracking process.
- Ergonomic features. Systems that restrict user movement and uncomfortable devices can lead to user discomfort, which can affect data accuracy.

Researching and improving these parameters is essential for developing more efficient and accurate eye tracking systems. Equally important is a comprehensive assessment of these parameters and their impact on each other. Attention to these factors can increase tracking accuracy, improve the user's experience with the system, and increase user comfort.

Summary of the review of existing eye tracking devices

There are three main types of eye tracking devices: desktop eye tracking complexes, mobile eye tracking glasses, and virtual reality eye tracking complexes. Web-camera and mobile eye tracking stand apart because they use built-in standard webcams.

Each of these types has its advantages and disadvantages, and the choice of a complex for a particular study depends on the goals set and many requirements such as accuracy, synchronization latency, mobility, etc. Currently, hardware systems are more focused on scientific research and, accordingly, less accurate and cheaper systems for commercial, more massive use.

In general, to successfully conduct eye tracking research, it is essential to carefully analyze all aspects related to the use of different hardware systems, environmental conditions, and the goals of this research.

In the further development of eye tracking technologies, we can expect to see improvements in the characteristics of hardware systems to obtain more accurate and informative research results, the use of built-in cameras in laptops and smartphones for eye tracking, and the expansion of the technology to even more scientific and applied areas.

Methods of eye tracking

An overview of eye tracking methods

Since gaze tracking technology began, several areas have been invented: electro-oculography, electromagnetic oculography, and video oculography (eye tracking). In electro-oculography methods, sensors are placed around the eyes to measure the electric field that changes when the eyes turn, and the eye's position is determined by changes in slight potential differences around the eye [35, 36]. In electromagnetic oculography, special contact lenses with an integrated inductor are used, and the eye's position is determined by the orientation of the coil in the magnetic field. Commercial devices commonly use video oculography or eye tracking, which allows for detecting gaze direction by analyzing a video stream with a face [37].

Modern eye tracking solves two tasks: eye localization and matching the movement of the gaze with the corresponding objects of observation (gaze tracking). There are 3 main approaches to eye localization:

- Shape-based Methods;
- Feature-based Shape Methods;
- Appearance-based Methods.

Gaze tracking makes it possible to correlate eye movements with objects and areas to which the gaze is directed and thus track the movement of the gaze. Depending on the equipment, gaze-tracking methods can be characterized by the number of cameras (one or more), the presence of infrared lighting (active and passive), and the number of additional infrared lighting sources (one or more). Like eye localization methods, gaze tracking methods have the following main approaches:

- Feature-based Methods;
- Model-based Methods;
- Cross Ratio based Methods;
- Appearance-based Methods.

It should be noted that most eye localization and gaze tracking methods use additional infrared lighting and cameras capable of perceiving it. It imposes some limitations on the use of such methods. For example, using an eye tracker in bright sunlight makes it difficult. Also, since most modern smartphones do not have a front-facing infrared camera (or do not allow third-party developers to use it), it is impossible to use the following methods in mobile eye tracking.

Methods of eye localization

The primary task in image tracking when analyzing a video is to detect the eyes in it. The main advantage is that the appearance of the eyes has many similar features among people, regardless of their age, gender, or ethnicity. However, some challenges complicate such analysis: wearing glasses, adapting to different head positions, blink rate, eye closure, individual differences in eye size, and other factors. In computer vision, eye localization is often performed parallel to face detection. To effectively solve this problem, it is possible to use methods that first detect the face, which significantly simplifies the localization of eyes in the area.

Shape-based methods propose the construction of a geometric model of the eye. Among them, there are methods that use a simple elliptical eye model, where the eye model is based on the elliptical shape of the iris contour, and methods where the eye is represented by a more complex shape, such as two parabolas for the eyelids and an ellipse for the iris [39]. An essential feature of these methods is their general ability to handle shape, scale, and rotation changes [40].

In Feature-based Shape methods, the characteristics of the eye are examined to identify a set of characteristics around the eyes. The limbus, pupil, and corneal reflectance are standard features used for eye localization [40].

Appearance-based Methods detect and track eyes directly based on the photometric appearance, which is characterized by the eye's color distribution or filter response and surroundings. One of the main benefits of detection is to reduce the impact of light changes by preserving sub-bands that are less sensitive to light and removing sub-bands that are sensitive to light changes. Appearance-based methods can be based on image templates, where spatial and intensity information of each pixel is stored, or on a holistic approach, where the intensity distribution is characterized by ignoring spatial information. Appearance-based methods typically require collecting a large amount of training data representing the eyes of different subjects under different face orientations and lighting conditions. Still, the models underlying them are essentially independent of the object class [37].

Gaze-tracking methods

The main task of gaze tracking or gaze prediction is to compare eye movements with the Point of Regard (PoR), the position in the displayed content the user is supposed to view. By analyzing the movement of the PoR, it is possible to detect gaze fixations and saccades (fast gaze movements) and to detect, according to [1], the level of information absorption.

Near-infrared eye illumination is widely used in gaze-tracking methods, which allows cameras to identify eye movements. There are two methods of such lighting: the "light pupil" method and the "dark pupil" method.

The "light pupil" method (Fig. 4) utilizes the contrast between the pupil and iris created by infrared rays parallel to the camera's optical axis. It effectively reduces noise caused by eyelashes or glare and is suitable for various lighting conditions. The pupil in the image appears bright due to the reflection of light from the retina [33].

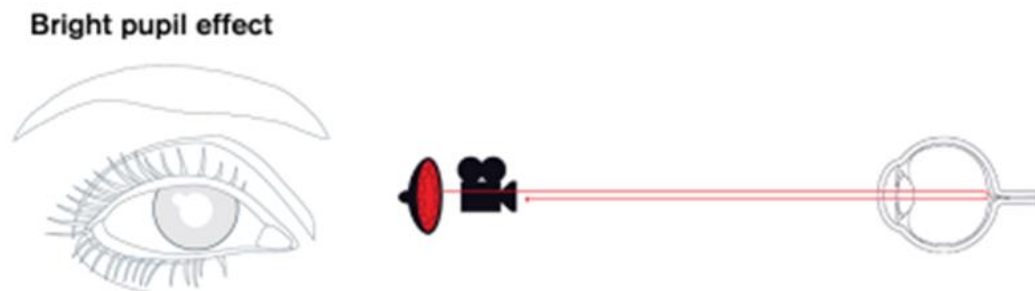


Fig. 4. The "light and light" method [38]

The "dark pupil" method (Fig. 5) also uses infrared light, but the illumination is placed at an angle to the camera's optical axis. It makes it possible to create stationary systems independent of the camera. However, the pupil becomes more sensitive to ambient light, which limits the use of the method in very bright environments [38].

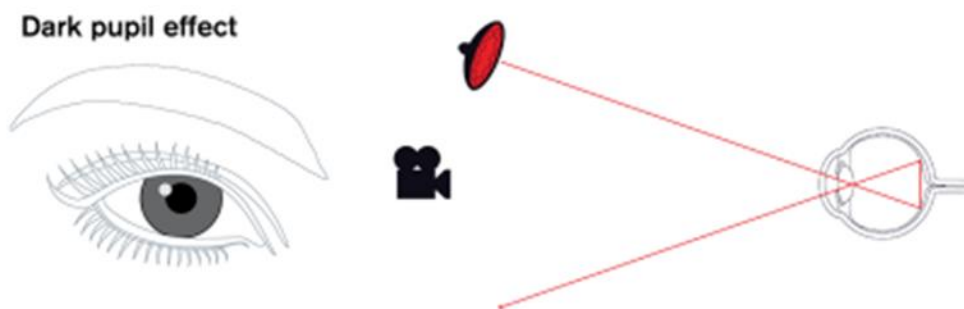


Fig. 5. Dark Pupil Method [38]

In general, the gaze focus point is determined by comparing the position of the near-infrared light (reflected by the eye) to the position of the pupil. By combining this information with head position data, the point of focus can be determined.

Using gaze tracking methods based on the visible light spectrum poses a significant challenge due to the low contrast between the iris and pupil. Compared to methods that use infrared light, this makes it difficult to determine the exact position of the pupil, which is why the position of the center of the iris is often used as an alternative in the analysis. This approach can lead to some inaccuracy in measurements, as it does not provide effective mechanisms for filtering out noise. It is necessary to implement video processing algorithms that can better distinguish between the pupil and iris to improve accuracy, as well as the use of machine learning methods to improve the accuracy of gaze detection based on data analysis. However, the use of visible light spectrum methods is unavoidable in cases where it is not possible to provide infrared light and when the camera does not support infrared perception. Thus, such methods are advisable in mobile eye tracking and webcam eye tracking.

Let's look at the main approaches of gaze tracking.

Feature-based Methods use eye features (pupil border and contrast, illumination distribution, etc.) to determine geometric vectors that reflect the shape of the eyes and head position to estimate the direction of gaze. The advantage of such methods is that the pupil and glare are easy to find and that these features can be formally related to gaze. These methods have performance issues in high light [41]. [40] provides a broad overview of feature-based methods.

Model-based Methods determine the gaze direction by fitting face and eye models to the image. These methods first build the eye's optical axis, then the gaze axis, and as a result, the point of view is determined by intersecting the gaze axis with the scene's geometry. The gaze axis is constructed by finding the pupil's cornea and center. The disadvantage of this approach is the need to use high-resolution images of the eyes and manually calibrate the system for each user to determine the eyes' location on the face accurately. The advantage of model-based methods is that they allow gaze direction to be determined using simple facial characteristics, minimizing the need for volumetric data for training for a particular person [41].

Cross Ratio based methods can determine the gaze using several sources of IR illumination and detect their glare on the cornea. [42] describes a method based on four infrared illuminators and [43] - based on five illuminants, which makes it possible to avoid additional calibration. The advantage of these methods is that there is no need for pre-calibration. Still, using additional illuminants may impose limitations due to the inability to use them with some devices.

Appearance-based methods use eye images as input to find the point of view using machine learning. These methods can work effectively with low-quality images but require significant data to train the models. Although they offer a direct image-to-gaze point conversion, they face limitations related to head pose, as changes in pose and gaze direction are not always reflected in changes in the appearance of the eye region [40].

Summary of eye tracking algorithms review

Thanks to decades of research into eye tracking methods, this technology has significantly progressed in developing various approaches and their features. The first task of eye tracking, eye localization, includes methods based on the eye area's shape, shape features, and appearance. Each approach has advantages, including accuracy in different lighting conditions and head positions. Gaze tracking methods, particularly those that use infrared light, provide high accuracy in determining the position of the

pupil and the direction of gaze, although they also have limitations. Using machine learning in appearance-based methods expands eye tracking possibilities, allowing it to work effectively with low-quality images. Modern eye tracking methods have significantly improved accuracy and user-friendliness, opening up various applications, from medical research to interface design. However, existing challenges, such as wearing glasses, lighting variability, and the need for calibration, point to further study and improvement of these methods.

Conclusions.

The study discusses in detail the use of eye tracking technologies, their types, features, and the algorithms on which their operation is based. It is noted that such systems can accurately determine the user's point of view both when interacting with computer systems and with objects in real and immersive environments.

The first part discusses various hardware systems for gaze tracking: systems based on screen interaction, systems using glasses, and systems for virtual and augmented reality. Attention was also paid to mobile and webcam eye tracking, which uses built-in cameras of devices and household webcams to track eyes.

The second part of the article examines the algorithmic side of eye tracking. The main tasks solved by the eye tracking algorithms are outlined: localization of eyes in a video series and comparison of eye movement with the point of view. The main approaches to solving these problems are reviewed, and the challenges and limitations of these technologies are discussed, including the need to overcome gaze point errors and improve measurement accuracy.

Potential directions for future research are outlined, including creating more affordable gaze tracking systems that do not depend on expensive equipment, which could be widely used in various fields of research.

References

1. Irwin, David E., Laura A. Carlson-Radvansky, and Rachel V. Andrews. 'Information Processing during Saccadic Eye Movements'. *Acta Psychologica, Discrete and Continuous Information Processing*, 90, no. 1 (1 November 1995): 261–73. [https://doi.org/10.1016/0001-6918\(95\)00024-0](https://doi.org/10.1016/0001-6918(95)00024-0).
2. Zammarchi, Gianpaolo, and Claudio Conversano. 'Application of Eye Tracking Technology in Medicine: A Bibliometric Analysis'. *Vision* 5, no. 4 (11 November 2021): 56. <https://doi.org/10.3390/vision5040056>.
3. Wei, QiuHong, Huiling Cao, Yuan Shi, Ximing Xu, and Tingyu Li. 'Machine Learning Based on Eye-Tracking Data to Identify Autism Spectrum Disorder: A Systematic Review and Meta-Analysis'. *Journal of Biomedical Informatics* 137 (January 2023): 104254. <https://doi.org/10.1016/j.jbi.2022.104254>.
4. Schall, Andrew, and Jennifer Romano Bergstrom. '1 - Introduction to Eye Tracking'. In *Eye Tracking in User Experience Design*, edited by Jennifer Romano Bergstrom and Andrew Jonathan Schall, 3–26. Boston: Morgan Kaufmann, 2014. <https://doi.org/10.1016/B978-0-12-408138-3.00001-7>.
5. Bojko, Agnieszka. *Eye Tracking the User Experience: A Practical Guide to Research*, 2013.
6. Białowas, Sylwester, and Adrianna Szyszka. 'Eye-Tracking in Marketing Research'. In *Managing Economic Innovations – Methods and Instruments*, edited by Robert Romanowski, 91–104. Bogucki Irwin, David E., Laura A. Carlson-Radvansky, and Rachel V. Andrews. 'Information Processing during Saccadic Eye Movements'. *Acta Psychologica, Discrete and Continuous Information Processing*, 90, no. 1 (1 November 1995): 261–73. [https://doi.org/10.1016/0001-6918\(95\)00024-0](https://doi.org/10.1016/0001-6918(95)00024-0).
7. Zammarchi, Gianpaolo, and Claudio Conversano. 'Application of Eye Tracking Technology in Medicine: A Bibliometric Analysis'. *Vision* 5, no. 4 (11 November 2021): 56. <https://doi.org/10.3390/vision5040056>.
8. Wei, QiuHong, Huiling Cao, Yuan Shi, Ximing Xu, and Tingyu Li. 'Machine Learning Based on Eye-Tracking Data to Identify Autism Spectrum Disorder: A Systematic Review and Meta-Analysis'. *Journal of Biomedical Informatics* 137 (January 2023): 104254. <https://doi.org/10.1016/j.jbi.2022.104254>.
9. Schall, Andrew, and Jennifer Romano Bergstrom. '1 - Introduction to Eye Tracking'. In *Eye Tracking in User Experience Design*, edited by Jennifer Romano Bergstrom and Andrew Jonathan Schall, 3–26. Boston: Morgan Kaufmann, 2014. <https://doi.org/10.1016/B978-0-12-408138-3.00001-7>.
10. Bojko, Agnieszka. *Eye Tracking the User Experience: A Practical Guide to Research*, 2013.
11. Wydawnictwo Naukowe, 2019. <https://doi.org/10.12657/9788379862771-6>.
12. Adhanom, Isayas Berhe, Paul MacNeilage, and Eelke Folmer. 'Eye Tracking in Virtual Reality: A Broad Review of Applications and Challenges'. *Virtual Reality* 27, no. 2 (1 June 2023): 1481–1505. <https://doi.org/10.1007/s10055-022-00738-z>.
13. Oney, Seyda, Nils Rodrigues, Michael Becher, Thomas Ertl, Guido Reina, Michael Sedlmair, and Daniel Weiskopf. 'Evaluation of Gaze Depth Estimation from Eye Tracking in Augmented Reality'. In *ACM Symposium on Eye Tracking Research and Applications*, 1–5. ETRA '20 Short Papers. New York, NY, USA: Association for Computing Machinery, 2020. <https://doi.org/10.1145/3379156.3391835>.
14. Lu, Shang, Yerly Paola Sanchez Perdomo, Xianta Jiang, and Bin Zheng. 'Integrating Eye-Tracking to Augmented Reality System for Surgical Training'. *Journal of Medical Systems* 44, no. 11 (29 September 2020): 192. <https://doi.org/10.1007/s10916-020-01656-w>.

15. Pavlenko, Vitaliy, Tetiana Shamanina, and Vladislav Chori. 'Biometric Method of Personality Authentication Based on the Eye Tracking Data'. *Bulletin of the National Technical University KhPI A Series of Information and Modeling*, 25 October 2021, 142–52. <https://doi.org/10.20998/2411-0558.2021.01.11>.
16. Schweizer, Theresa, Thomas Wyss, and Rahel Gilgen-Ammann. 'Detecting Soldiers' Fatigue Using Eye-Tracking Glasses: Practical Field Applications and Research Opportunities'. *Military Medicine* 187 (14 December 2021). <https://doi.org/10.1093/milmed/usab509>.
17. Rosner, Agnes, Thomas Franke, Frederik Platten, and Christiane Attig. 'Eye Movements in Vehicle Control'. In *Eye Movement Research: An Introduction to Its Scientific Foundations and Applications*, edited by Christoph Klein and Ulrich Ettinger, 929–69. Studies in Neuroscience, Psychology and Behavioral Economics. Cham: Springer International Publishing, 2019. https://doi.org/10.1007/978-3-030-20085-5_22.
18. Pelgrim, Madeline H., Julia Espinosa, and Daphna Buchsbaum. 'Head-Mounted Mobile Eye-Tracking in the Domestic Dog: A New Method'. *Behavior Research Methods* 55, no. 4 (1 June 2023): 1924–41. <https://doi.org/10.3758/s13428-022-01907-3>.
19. 'Google Books Ngram Viewer'. Accessed 18 March 2023. https://books.google.com/ngrams/graph?content=eye-tracking&year_start=1900&year_end=2019&corpus=en-2019&smoothing=0.
20. Smith, John D., Roel Vertegaal, and Changuk Sohn. 'ViewPointer: Lightweight Calibration-Free Eye Tracking for Ubiquitous Handsfree Deixis'. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*, 53–61. Seattle WA USA: ACM, 2005. <https://doi.org/10.1145/1095034.1095043>.
21. 'Perfect for Usability Tests - Mobile Testing Accessory'. Accessed 18 March 2023. <https://www.tobii.com/products/accessories/mobile-testing-accessory>.
22. 'Tobii Pro Glasses 3 | Latest in Wearable Eye Tracking'. Accessed 20 March 2023. <https://www.tobii.com/products/eye-trackers/wearables/tobii-pro-glasses-3>.
23. Meng, Xiaoxu, Ruofei Du, and Amitabh Varshney. 'Eye-Dominance-Guided Foveated Rendering'. *IEEE Transactions on Visualization and Computer Graphics* 26, no. 5 (May 2020): 1972–80. <https://doi.org/10.1109/TVCG.2020.2973442>.
24. GazeRecorder. 'Online Eye Tracking | Webcam Eye-Tracking Software'. Accessed 18 May 2024. <https://gazerecorder.com/>.
25. myEye Project. 'Home'. Accessed 18 May 2024. <http://myeye.jimdofree.com/>.
26. 'WebGazer.js: Democratizing Webcam Eye Tracking on the Browser'. Accessed 18 May 2024. <https://webgazer.cs.brown.edu/>.
27. 'Online Eye Tracking Software | Free Trial - Sticky - Tobii'. Accessed 18 May 2024. <https://www.tobii.com/products/software/remote-testing-software/sticky>.
28. 'iMotions Online - iMotions'. Accessed 18 May 2024. <https://imotions.com/products/imotions-online/>.
29. 'Hawkeye | Learn Where People Look in Your Products'. Accessed 18 May 2024. <https://www.usehawkeye.com/>.
30. VisualCamp. 'Home-Edite - VisualCamp', 2 February 2024. <https://visual.camp/>.
31. 'Tobii Customer Portal'. Accessed 18 May 2024. https://connect.tobii.com/s/article/eye-tracker-accuracy-and-precision?language=en_US.
32. 'Tobii Customer Portal'. Accessed 18 May 2024. https://connect.tobii.com/s/article/What-is-the-latency-of-my-eye-tracker?language=en_US.
33. 'Tobii Customer Portal'. Accessed 18 May 2024. https://connect.tobii.com/s/article/tobii-eye-tracker-glossary?language=en_US.
34. Global, Tobii Dynavox. 'What Factors in Your Environment Can Affect Gaze Interaction?' Tobii Dynavox Global. Accessed 1 April 2023. <https://www.tobiidynavox.com/blogs/support-articles/what-factors-in-your-environment-can-affect-gaze-interaction>.
35. Jia, Yingxin, and Christopher W. Tyler. 'Measurement of Saccadic Eye Movements by Electrooculography for Simultaneous EEG Recording'. *Behavior Research Methods* 51, no. 5 (1 October 2019): 2139–51. <https://doi.org/10.3758/s13428-019-01280-8>.
36. Bharadwaj, Dr, and Bandna Kumari. 'Electrooculography: Analysis On Device Control By Signal Processing'. *International Journal of Advanced Computer Research* 8 (11 September 2018): 787–90. <https://doi.org/10.26483/ijarcs.v8i3.3098>.
37. Chennamma, Hr, and Xiaohui Yuan. 'A Survey on Eye-Gaze Tracking Techniques'. *Indian Journal of Computer Science and Engineering* 4 (22 December 2013).
38. 'Tobii Customer Portal'. Accessed 18 May 2024. https://connect.tobii.com/s/article/What-is-dark-and-bright-pupil-tracking?language=en_US.
39. Yuille, A.L., D.S. Cohen, and P.W. Hallinan. 'Feature Extraction from Faces Using Deformable Templates'. In *Proceedings CVPR '89: IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 104–9. San Diego, CA, USA: IEEE Comput. Soc. Press, 1989. <https://doi.org/10.1109/CVPR.1989.37836>.
40. Hansen, Dan Witzner, and Qiang Ji. 'In the Eye of the Beholder: A Survey of Models for Eyes and Gaze'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32, no. 3 (March 2010): 478–500. <https://doi.org/10.1109/TPAMI.2009.30>.
41. Kottwani, Aveena, and Ayush Kumar. 'Eye Gaze Estimation Model Analysis', 28 July 2022. <https://doi.org/10.13140/RG.2.2.22546.99522>.
42. Yoo, Donghyun, and Myung Chung. 'A Novel Non-Intrusive Eye Gaze Estimation Using Cross-Ratio under Large Head Motion'. *Computer Vision and Image Understanding* 98 (1 April 2005): 25–51. <https://doi.org/10.1016/j.cviu.2004.07.011>.
43. Yoo, Donghyun, Jae Kim, Bang Lee, and Myoung Chung. *Non-Contact Eye Gaze Tracking System by Mapping of Corneal Reflections*, 2002. <https://doi.org/10.1109/AFGR.2002.1004139>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-24>

УДК 681.3.05:004.056

Розломій Інна Олександрівна¹, к.т.н., доцент

<https://orcid.org/0000-0001-5065-9004>

Симонюк Володимир Павлович², к.т.н., доцент

<https://orcid.org/0000-0002-7624-4760>

Науменко Сергій Васильович¹, аспірант

<https://orcid.org/0000-0002-6337-1605>

Михайловський Павло Васильович¹, аспірант

<https://orcid.org/0009-0008-4324-1724>

¹Черкаський національний університет імені Богдана Хмельницького, м. Черкаси, Україна

²Луцький національний технічний університет, м. Луцьк, Україна

МОДЕЛЬ БЕЗПЕКИ ВЗАЄМОПОВ'ЯЗАНИХ ОБЧИСЛЮВАЛЬНИХ ПРИСТРОЇВ НА ОСНОВІ ПОЛЕГШЕНОЇ СХЕМИ ШИФРУВАННЯ ДЛЯ ІОТ

Розломій І.О., Симонюк В.П., Науменко С.В., Михайловський П.В. Модель безпеки взаємопов'язаних обчислювальних пристроїв на основі полегшеної схеми шифрування для ІоТ. У сучасному світі число пристроїв, підключених до інтернету речей (ІоТ), стрімко зростає, що породжує нові виклики у сфері забезпечення їхньої безпеки. Стаття присвячена розробці моделі безпеки для ІоТ пристроїв, яка враховує обмежені обчислювальні ресурси і мінімальне енергоспоживання, необхідні для ефективного функціонування криптографічних алгоритмів. Проблема полягає у відсутності універсальних полегшених криптографічних рішень, здатних забезпечити надійний захист в умовах ІоТ без значного збільшення витрат енергії чи обчислювальних ресурсів. У статті представлено аналіз сучасних досліджень та публікацій у цій сфері, зокрема, розглянуто новітні розробки в галузі полегшеної криптографії, що оптимізовані для ІоТ пристроїв. Особлива увага приділена легким криптографічним алгоритмам, які можуть забезпечити високу безпеку при мінімальному використанні ресурсів, таких як енергія та обчислювальна потужність. Запропонована модель включає як апаратні, так і програмні рішення для шифрування та управління ключами, з метою мінімізації ризиків компрометації даних. Досліджуються методи оцінки ефективності засобів захисту та моделі врахування обмежень ресурсів. У статті також розглядаються перспективи використання квантово-стійких криптографічних протоколів, які можуть стати основою для майбутніх стандартів безпеки ІоТ. Висновки підкреслюють важливість постійного вдосконалення захисних механізмів для підвищення рівня безпеки ІоТ технологій, що сприятиме зростанню довіри користувачів до цих систем і забезпечить надійний захист даних у глобальних мережах. Розробка та імплементація таких полегшених криптографічних рішень є критично важливими для захисту даних і забезпечення стабільної роботи ІоТ екосистем в умовах сучасних кіберзагроз.

Ключові слова: Інтернет речей, полегшена криптографія, шифрування, апаратні шифратори, криптографічні бібліотеки, керування ключами, обмеження ресурсів.

Rozlomi I., Symonyuk V., Naumenko S., Mykhailovskiy P. The security model of interconnected computing devices based on a lightweight encryption scheme for IoT. Today world, the number of devices connected to the Internet of Things (IoT) is growing rapidly, which creates new challenges in the field of ensuring their security. The article is devoted to the development of a security model for IoT devices, which takes into account the limited computing resources and minimal energy consumption necessary for the effective functioning of cryptographic algorithms. The problem is the lack of universal lightweight cryptographic solutions capable of providing reliable protection in the IoT environment without a significant increase in energy consumption or computing resources. The article presents an analysis of modern research and publications in this field, in particular, the latest developments in the field of lightweight cryptography optimized for IoT devices are considered. Special attention is paid to lightweight cryptographic algorithms that can provide high security with minimal use of resources such as energy and computing power. The proposed model includes both hardware and software solutions for encryption and key management in order to minimize the risks of data compromise. The methods of assessing the effectiveness of protective measures and the models for taking into account resource limitations are being studied. The article also discusses the prospects for using quantum-resistant cryptographic protocols that could become the basis for future IoT security standards. The conclusions emphasize the importance of continuous improvement of protective mechanisms to increase the level of security of IoT technologies, which will contribute to the growth of user trust in these systems and ensure reliable protection of data in global networks. The development and implementation of such lightweight cryptographic solutions are critical for protecting data and ensuring the stable operation of IoT ecosystems in the face of modern cyber threats.

Key words: Internet of Things, lightweight cryptography, encryption, hardware ciphers, cryptographic libraries, key management, resource limitations.

Постановка проблеми та її зв'язок із важливими науковими чи практичними завданнями. З розвитком інтернету речей (ІоТ) постійно збільшується кількість пристроїв, які взаємодіють між собою через глобальні мережі. В кінці 2023 року кількість таких пристроїв перевищила 30 мільярдів, що в два рази більше, ніж п'ять років тому. Прогнози показують, що ця кількість може досягти 50 мільярдів до 2025 року, рисунок 1 [1]. Однак, стрімкий розвиток ІоТ

також призводить до збільшення кількості загроз безпеці цих пристроїв. До таких загроз належать витоки даних, несанкціоноване втручання та атаки на конфіденційність даних користувачів [2].

В умовах, коли кіберзлочинці вдосконалюють свої методи атак на системи IoT, наукові та практичні спільноти стикаються з нагальною потребою розробки ефективних, але при цьому ресурсно ощадних методів шифрування [3]. Багато існуючих рішень у сфері криптографії не враховують специфіку пристроїв IoT, які характеризуються обмеженими обчислювальними потужностями, малим розміром пам'яті та низьким енергоспоживанням.

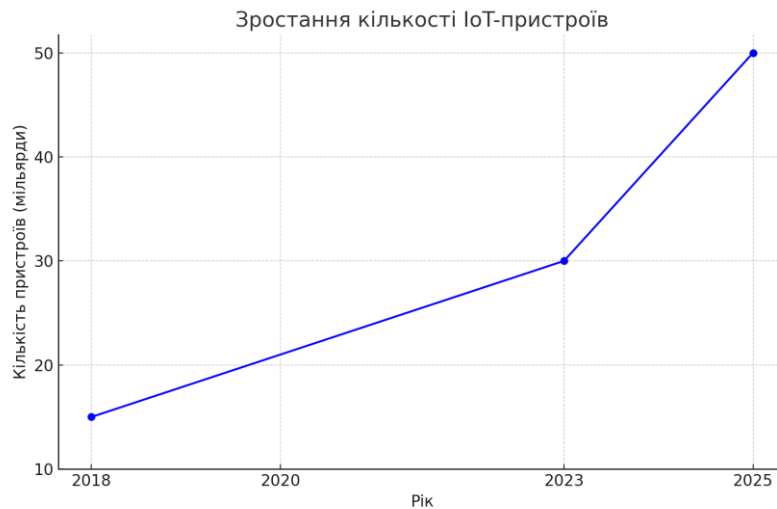


Рис. 1. Графік зростання кількості IoT-пристроїв

Проблема полягає у відсутності універсальних легких криптографічних рішень, які здатні забезпечити надійний захист в обмежених умовах IoT без значного збільшення витрат енергії чи обчислювальних ресурсів. Це вимагає розробки нових алгоритмів шифрування та методів управління ключами, оптимізованих для масштабів IoT. Науковий та практичний інтерес до цієї проблеми обумовлений глобальною потребою в стандартизації та імплементації таких рішень на законодавчому рівні, що включає забезпечення відповідності міжнародним нормам та регуляціям у сфері захисту даних та конфіденційності. Вирішення цієї проблеми не лише посилить захист приватності та даних кінцевих користувачів, але й сприятиме подальшому розвитку та інтеграції IoT технологій у повсякденне життя. Отже, пошук рішень для цієї проблеми має важливе значення як з наукової, так і з практичної точки зору.

Актуальність дослідження зумовлена кількома ключовими факторами. Зростання кількості IoT-пристроїв, які є частиною критично важливих систем, таких як системи охорони здоров'я, інтелектуальні транспортні системи та промислове обладнання, вимагає високого рівня безпеки через їхній вплив на безпеку та добробут людей. Законодавчі та нормативні вимоги щодо захисту даних, встановлені регуляторними органами, такими як Європейський Союз із своїм Загальним регламентом про захист даних (GDPR), а також іншими міжнародними організаціями, посилюють необхідність розробки відповідних криптографічних технологій [4]. Збільшення різноманітності та складності кібератак змушує зловмисників шукати нові способи обходу традиційних захисних механізмів, що вимагає більш вдосконалених рішень для шифрування та аутентифікації.

Наукове співтовариство та індустрія також виявляють зростаючий інтерес до розвитку стійких до квантових обчислень криптографічних рішень, оскільки квантові технології можуть згодом зламати багато сучасних систем шифрування [5]. Адаптація легких криптографічних методів, що забезпечують захист від потенційних квантових загроз, стає критично важливою.

Всі ці фактори разом з нагальною потребою в ефективному, але ресурсно невимогливному шифруванні для забезпечення безпеки величезного масиву IoT-пристроїв роблять дане дослідження надзвичайно актуальним та важливим як з наукової, так і з практичної точки зору. Розробка та імплементація нових легких криптографічних рішень може суттєво зміцнити захист в IoT екосистемах, підвищити довіру користувачів і стимулювати подальший розвиток інтелектуальних технологій.

Аналіз останніх досліджень та публікацій. Сфера криптографічного захисту IoT набуває все

більшої актуальності, а зростаючий інтерес дослідників та розробників призводить до появи значної кількості наукових робіт, спрямованих на вдосконалення методів шифрування, які би задовольняли вимоги до пристроїв IoT [6]. Основна увага в сучасних дослідженнях приділяється розробці так званих «легких» криптографічних алгоритмів, що оптимізовані для обмеженого обчислювального потенціалу та мінімального енергоспоживання [7].

Одним з найважливіших напрямків у цьому контексті є робота [8], де представлено новий алгоритм шифрування, базований на структурі Feistel Network. Алгоритм показав високу стійкість до атак з відкритим текстом і при цьому демонструє значно нижчі вимоги до обчислювальних ресурсів порівняно з традиційними методами.

Варто також виділити дослідження Mahmood та співавторів, в якому представлено метод дворівневого шифрування, що інтегрує асиметричне шифрування для управління ключами і симетричне шифрування для передачі даних [9]. Цей підхід не тільки забезпечує високу безпеку, але й оптимізує процес використання енергії пристроєм.

Значну увагу привертає дослідження в області квантово-стійкого шифрування, яке є відповіддю на потенційні майбутні загрози квантових обчислень [10]. Робота демонструє новітній підхід до створення квантово-стійких криптографічних протоколів, які можуть бути ефективно імплементовані на IoT пристроях.

З огляду на зростаючі вимоги до безпеки даних в IoT, вчені також акцентують увагу на розробці універсальних рішень для управління ключами. Важливим напрямком є дослідження, які спрямовані на інтеграцію блокчейн технологій для децентралізованого управління криптографічними ключами в мережах IoT [11].

Ці дослідження є важливими кроками в розробці криптографічних рішень, які можуть ефективно функціонувати в умовах обмежених ресурсів, що є характерним для IoT пристроїв, і при цьому забезпечувати високий рівень безпеки і захисту даних.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Пристрої IoT широко використовуються в різних сферах, від побутових приладів до медичних пристроїв та інфраструктурних систем. Однак, через обмежені ресурси таких пристроїв, вони часто мають слабкі місця в системах захисту, що робить їх вразливими до різноманітних кібератак. Важливість забезпечення надійного захисту даних та безперебійної роботи цих пристроїв не можна переоцінити, оскільки порушення їх безпеки може мати серйозні наслідки для користувачів та організацій.

У взаємопов'язаних обчислювальних пристроях, зокрема в IoT, існують загрози безпеці, які можуть негативно вплинути на функціонування та цілісність системи, табл. 1.

Таблиця 1. Загрози безпеці пристроїв IoT

Загроза	Опис	Наслідки
Неавторизований доступ	Несанкціоноване проникнення у систему або пристрої	Крадіжка даних, маніпуляції пристроями
Атаки типу DoS	Спроба зробити систему або пристрій недоступними шляхом перенавантаження ресурсів	Недоступність системи, зупинка роботи
Атаки на конфіденційність	Перехоплення та використання чутливих даних	Витік конфіденційної інформації
Атаки на цілісність даних	Модифікація даних, що передаються між пристроями	Неправильні рішення або дії системи
Атаки на доступність	Спрямовані на порушення безперервної роботи пристроїв	Порушення функціонування системи
Зловмисне програмне забезпечення (Malware)	Впровадження шкідливих програм з метою крадіжки даних або порушення роботи пристроїв	Використання у ботнет-атаках, крадіжка інформації

У зв'язку з численними загрозами, які постають перед пристроями IoT, важливо забезпечити

надійний захист даних і безперебійне функціонування цих пристроїв. Основна увага приділяється розробці ефективних засобів захисту, які можуть протистояти кібератакам, зокрема неавторизованому доступу, атакам типу DoS, та іншим видам зловмисних дій [12]. Одним з таких засобів є полегшена криптографія, яка спеціально розроблена для пристроїв з обмеженими ресурсами. Обмеження обчислювальних можливостей пристроїв IoT вимагає пошуку ресурсоощадних засобів захисту інформації. Одним з таких є полегшена криптографія. Вона характеризується низьким споживанням енергії, невеликою пам'яттю та обчислювальними вимогами, що дозволяє застосовувати її в умовах обмежених апаратних можливостей.

Полегшена криптографія є важливим напрямом у галузі інформаційної безпеки, спрямованим на забезпечення захисту даних в умовах обмежених обчислювальних ресурсів, таких як енергія, пам'ять і обчислювальна потужність. Традиційні криптографічні алгоритми, такі як AES, потребують значних ресурсів, що робить їх непридатними для використання в багатьох пристроях IoT.

Полегшена криптографія базується на таких принципах:

1. Ефективність. Алгоритми повинні бути достатньо швидкими і не споживати багато енергії. Це дозволяє пристроям IoT працювати довше без заміни батарей або зарядки.
2. Мінімальне використання пам'яті. Алгоритми повинні вимагати мінімум оперативної пам'яті (RAM) для обробки даних, що особливо важливо для пристроїв з обмеженим об'ємом RAM.
3. Безпека. Незважаючи на свою «легкість», алгоритми повинні забезпечувати достатній рівень захисту даних від різних видів атак.

Захист пов'язаних пристроїв IoT на рівні мікроконтролерів можна забезпечити шляхом використання криптографічних бібліотек, які інтегровані у програмне забезпечення мікроконтролерів, або за допомогою апаратних шифраторів, якщо мікроконтролер має вбудований апаратний модуль для шифрування. Розглянемо детальніше захист інформації за допомогою використання бібліотек та апаратних шифраторів.

Криптографічні бібліотеки забезпечують програмну реалізацію полегшених криптографічних алгоритмів. Вони можуть бути інтегровані у програмне забезпечення мікроконтролерів для забезпечення шифрування та розшифрування даних. Це дозволяє пристроям IoT використовувати ефективні методи захисту інформації, не перевантажуючи обчислювальні ресурси. До популярних бібліотек належать [13]:

1. TinyCrypt – легка криптографічна бібліотека, розроблена для IoT пристроїв, що забезпечує базові криптографічні операції з мінімальним використанням ресурсів. Вона підтримує такі функції, як блокове шифрування, хешування та генерування ключів.
2. WolfSSL відома своєю високою продуктивністю та низькими вимогами до пам'яті. Ця бібліотека підтримує різні полегшені шифри та протоколи, що робить її ідеальною для IoT пристроїв. Вона включає такі функції, як SSL/TLS протоколи, а також підтримку алгоритмів, таких як RSA, ECC (еліптичні криві) та AES.
3. mbed TLS – легка криптографічна бібліотека з відкритим кодом, яка забезпечує широкий набір криптографічних примітивів і протоколів для IoT пристроїв. Вона підтримує різні алгоритми шифрування, аутентифікації та хешування.

Інтеграція криптографічної бібліотеки у програмне забезпечення мікроконтролера здійснюється на рівні прикладної програми та проміжного програмного шару (middleware). Спочатку бібліотека додається до проекту та ініціалізується у коді програми. Потім вона використовується для виконання криптографічних операцій, таких як шифрування, розшифрування та хешування даних. Це забезпечує захист інформації на рівні програмного забезпечення мікроконтролера. У контексті забезпечення безпеки IoT пристроїв важливо розглядати різні рівні, на яких можуть бути реалізовані захисні механізми. На рисунку 2 представлена схема основних шарів, починаючи від фізичного рівня IoT пристрою до апаратних і програмних рішень, які забезпечують криптографічний захист.

IoT Device	Фізичний пристрій, який використовує мікроконтролер для взаємодії з навколишнім середовищем та користувачами.
Application	Прикладна програма, що виконує основні функції пристрою IoT, такі як збір даних з сенсорів або контроль виконавчих механізмів.
Middleware Layer	Проміжний програмний шар, що забезпечує взаємодію між прикладною програмою та криптографічним API.
Cryptographic API	Спеціалізоване програмне забезпечення, яке надає інтерфейс для взаємодії з апаратним шифратором.
Hardware Encryptor	Апаратний модуль у складі мікроконтролера, який виконує криптографічні операції на апаратному рівні.
Hardware Layer	Апаратний шар мікроконтролера, що включає процесорні ядра, пам'ять та інші фізичні компоненти, необхідні для роботи пристрою.

Рис. 2. Схема основних шарів IoT-системи

Більшість сучасних мікроконтролерів мають вбудовані апаратні шифратори, що забезпечують ефективне виконання криптографічних операцій. Ці апаратні модулі дозволяють швидко і з низьким енергоспоживанням виконувати шифрування, розшифрування та інші криптографічні завдання, не навантажуючи основний процесор мікроконтролера. Використання апаратних шифраторів значно підвищує продуктивність і безпеку IoT пристроїв. Апаратні шифратори інтегровані безпосередньо у мікроконтролер і можуть бути використані через спеціальні API, що надаються виробниками мікроконтролерів. Це забезпечує доступ до апаратних криптографічних модулів без необхідності реалізовувати складні криптографічні алгоритми на програмному рівні.

Мікроконтролери з вбудованими апаратними шифраторами, такі як STM32F429, ESP32, NXP LPC55S69, Nordic nRF52840 та Renesas RA6M3, підтримують різні алгоритми шифрування, включаючи AES, SHA-256, RSA та ECC. Ці мікроконтролери інтегрують спеціалізовані криптографічні блоки, які виконують шифрування та розшифрування даних на апаратному рівні, що значно знижує навантаження на основний процесор [14].

Криптографічні бібліотеки забезпечують програмну реалізацію криптографічних алгоритмів і можуть бути інтегровані у програмне забезпечення мікроконтролерів. Апаратні шифратори вбудовані безпосередньо у мікроконтролери і виконують криптографічні операції на апаратному рівні [15]. Це дозволяє значно підвищити продуктивність та знизити енергоспоживання.

Апаратне шифрування на мікроконтролерах має суттєві переваги перед програмним шифруванням за допомогою криптографічних бібліотек, особливо в контексті IoT пристроїв, де важливими є продуктивність, енергоефективність та безпека. Хоча обидва підходи мають свої переваги та недоліки, апаратні шифратори часто забезпечують кращу продуктивність і безпеку завдяки спеціалізованим модулям, які оптимізовані для виконання криптографічних операцій. В таблиці 2 представлено порівняння шифрування з використанням криптографічних бібліотек та апаратних шифраторів.

Вибір між криптографічними бібліотеками та апаратними шифраторами залежить від конкретних вимог до продуктивності, енергоефективності, гнучкості та безпеки IoT пристрою. Криптографічні бібліотеки пропонують велику гнучкість та низьку вартість впровадження, що робить їх привабливими для менш вимогливих застосувань. Водночас апаратні шифратори забезпечують високу продуктивність, енергоефективність та безпеку, що є критичним для багатьох застосувань IoT, особливо там, де потрібна висока швидкість обробки даних і низьке енергоспоживання.

Таблиця 2. Порівняння

Параметр	Криптографічні бібліотеки	Апаратні шифратори
Продуктивність	Помірна, залежить від потужності мікроконтролера	Висока, завдяки спеціалізованому обладнанню
Енергоефективність	Відносно високе споживання енергії	Низьке споживання енергії
Гнучкість	Висока, легко оновлюється та змінюється	Низька, важко змінювати алгоритми
Вартість	Низька, не потребує додаткового обладнання	Вища початкова вартість обладнання
Впровадження	Швидке, за рахунок програмного інтегрування	Складніше, потребує спеціалізованого апаратного забезпечення
Безпека	Залежить від реалізації та безпеки ПЗ	Висока, ключі та дані зберігаються у спеціалізованих модулях

Для побудови математичної моделі безпеки взаємопов'язаних обчислювальних пристроїв на основі полегшеної схеми шифрування для IoT, розглянемо основні компоненти та операції, які включаються до такої моделі. Модель безпеки враховує дані, що передаються між пристроями, можливі загрози, а також засоби захисту даних як на апаратному рівні (апаратні шифратори), так і на програмному рівні (криптографічні бібліотеки). Представимо процес шифрування та дешифрування даних між пристроями IoT виразами (1) і (2) відповідно.

$$D'_i = E(D_i, K_i) \quad (1)$$

$$D_i = D(D'_i, K_i) \quad (2)$$

де D_i – дані, які передаються між пристроями; K_i – криптографічні ключі, використовувані для шифрування даних; E – функція шифрування (апаратна або програмна); D – функція дешифрування (апаратна або програмна).

IoT-пристрої мають обмежені обчислювальні ресурси, тому важливо враховувати ці обмеження при реалізації заходів безпеки. Для врахування обмежень ресурсів IoT пристроїв при розробці криптографічних алгоритмів, необхідно детально проаналізувати їх вплив на ефективність і безпеку шифрування. Важливо, щоб обрані алгоритми були оптимізовані для мінімального використання енергії та обчислювальних потужностей, зберігаючи при цьому високий рівень безпеки. Розглянемо математичний вираз, який враховує ці обмеження (3).

$$\sum_{i=1}^n C_i \leq C_{total}$$

де C_i – обчислювальні ресурси, необхідні для виконання певної криптографічної операції на пристрої i ; C_{total} – загальні доступні обчислювальні ресурси. Такий підхід дозволяє формалізувати та проаналізувати захист інформації в умовах обмежених ресурсів, характерних для IoT.

Для обчислення ризику компрометації даних залежно від ефективності засобів захисту та наявних загроз пропонується використання математичної моделі (4).

$$R = \sum_{j=1}^m \sum_{k=1}^n T_{jk} \cdot (1 - P_{jk})$$

де R – ризик компрометації даних; T_{jk} – ймовірність загрози j для засобу захисту k ; P_{jk} – ймовірність протидії загрози j засобом захисту k . Ця модель дозволяє кількісно оцінити рівень ризику, враховуючи різні фактори загроз і ефективність захисних заходів. Використання такого підходу допомагає ідентифікувати найуразливіші точки в системі та оптимізувати стратегії захисту для IoT

пристроїв.

Ефективність засобів захисту визначається як сукупна оцінка, що враховує вагу кожного засобу захисту і його здатність протидіяти загрозам (5).

$$E(P) = \sum_{k=1}^n \alpha_k \cdot P_k$$

де $E(P)$ – загальна ефективність засобів захисту; α_k – ваговий коефіцієнт для засобу захисту k . Ця модель дозволяє оцінити загальну ефективність захисних заходів, враховуючи відносну важливість кожного з них. Таким чином, можна визначити, які засоби захисту мають найбільший вплив на зменшення ризику компрометації даних в системі IoT.

Запропонована модель безпеки для пристроїв IoT забезпечує комплексний підхід до захисту даних, враховуючи як апаратні, так і програмні засоби. Вона дозволяє мінімізувати ризики, пов'язані із внутрішніми та зовнішніми загрозами, шляхом оптимального вибору засобів захисту при обмежених обчислювальних ресурсах.

Висновки та перспективи подальшого дослідження. У даній статті було розглянуто проблему забезпечення безпеки взаємопов'язаних обчислювальних пристроїв на основі полегшеної схеми шифрування для IoT. Було визначено, що сучасні методи шифрування часто не враховують обмежені обчислювальні ресурси та енергоспоживання IoT пристроїв, що створює необхідність у розробці спеціалізованих легких криптографічних рішень.

Запропонована модель безпеки, яка враховує як апаратні, так і програмні засоби шифрування та управління ключами, продемонструвала можливість забезпечення надійного захисту даних при мінімальних витратах ресурсів. Аналіз сучасних досліджень та публікацій виявив перспективні напрямки розвитку полегшеної криптографії, зокрема, використання квантово-стійких протоколів та інтеграції блокчейн технологій для децентралізованого управління криптографічними ключами.

Подальші дослідження можуть бути спрямовані на розробку нових легких криптографічних алгоритмів та методів управління ключами, які ще більше знизять енергоспоживання та покращать безпеку IoT пристроїв. Також перспективним є дослідження інтеграції штучного інтелекту для автоматизації процесів виявлення та протидії кібератакам на IoT системи.

Важливим напрямком є проведення експериментальних досліджень та тестування запропонованих рішень в реальних умовах, що дозволить оцінити їхню ефективність та надійність. Поглиблене вивчення можливостей апаратних шифраторів та оптимізація їх взаємодії з програмними засобами захисту може суттєво підвищити загальний рівень безпеки IoT екосистем.

Список бібліографічного опису

1. Marton, A., 2023. State of IoT – spring 2023, report by IoT analytics, safepay systems. Available from: <https://iotac.eu/state-of-iot-spring-2023-by-iot-analytics/>.
2. Hassan, W. H. (2019). Current research on Internet of Things (IoT) security: A survey. *Computer networks*, 148, 283-294.
3. Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., & Sikdar, B. (2019). A survey on IoT security: application areas, security threats, and solution architectures. *IEEE Access*, 7, 82721-82743.
4. Zaem, R. N., & Barber, K. S. (2020). The effect of the GDPR on privacy policies: Recent progress and future promise. *ACM Transactions on Management Information Systems (TMIS)*, 12(1), 1-20.
5. Easttom, C. (2022). Quantum computing and cryptography. In *Modern Cryptography: Applied Mathematics for Encryption and Information Security* (pp. 397-407). Cham: Springer International Publishing.
6. Rozlomi, I., Yarmilko, A., & Naumenko, S. (2024, April). Data security of IoT devices with limited resources: challenges and potential solutions. In *Proceedings of the 4th Edge Computing Workshop (doors 2024)*, Zhytomyr, Ukraine (pp. 85-96).
7. Rozlomi, I., Yarmilko, A., Naumenko, S., & Mykhailovskyi, P. (2023). IoT Smart Implants: Information Security and the Implementation of Lightweight Cryptography. In *Proceedings of the 6th International Conference on Informatics & Data-Driven Medicine (IDDM'2023)* (pp. 145-146). Bratislava.
8. Abd Ali, S. M., & Hasan, H. F. (2019). Novel encryption algorithm for securing sensitive information based on feistel cipher. *Test Engineering Management*, 19(80), 10-16.
9. Mahmood, Z., Ning, H., & Ghafoor, A. (2016, December). Lightweight two-level session key management for end user authentication in Internet of Things. In *2016 IEEE international conference on internet of things (iThings) and IEEE Green computing and communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 323-327). IEEE.
10. Fernández-Caramés, T. M. (2019). From pre-quantum to post-quantum IoT security: A survey on quantum-resistant cryptosystems for the Internet of Things. *IEEE Internet of Things Journal*, 7(7), 6457-6480.
11. Panda, S. S., Jena, D., Mohanta, B. K., Ramasubbareddy, S., Daneshmand, M., & Gandomi, A. H. (2021). Authentication and key management in distributed iot using blockchain technology. *IEEE Internet of Things Journal*, 8(16), 12947-12954.

12. Yugha, R., & Chithra, S. (2020). A survey on technologies and security protocols: Reference for future generation IoT. *Journal of Network and Computer Applications*, 169, 102763.
13. Restuccia, G., Tschofenig, H., & Baccelli, E. (2020, December). Low-power IoT communication security: On the performance of DTLS and TLS 1.3. In *2020 9th IFIP International Conference on Performance Evaluation and Modeling in Wireless Networks (PEMWN)* (pp. 1-6). IEEE.
14. Pop, A. A. (2022). Incremental encoder speed acquisition using an STM32 microcontroller and NI ELVIS. *Sensors*, 22(14), 5127.
15. Розломий І.О., Косенюк Г.В., Науменко С.В., Михайловський П.В. (2023) Моделювання системи датчиків на базі мікроконтролера в ігровій симуляції «Смарт-будинок» з використанням шифрування. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. (53), 292-299. <https://doi.org/10.36910/6775-2524-0560-2023-53-43>

References

1. Marton, A., 2023. State of IoT – spring 2023, report by IoT analytics, safepay systems. Available from: <https://iotac.eu/state-of-iot-spring-2023-by-iot-analytics/>.
2. Hassan, W. H. (2019). Current research on Internet of Things (IoT) security: A survey. *Computer networks*, 148, 283-294.
3. Hassija, V., Chamola, V., Saxena, V., Jain, D., Goyal, P., & Sikdar, B. (2019). A survey on IoT security: application areas, security threats, and solution architectures. *IEEE Access*, 7, 82721-82743.
4. Zaeem, R. N., & Barber, K. S. (2020). The effect of the GDPR on privacy policies: Recent progress and future promise. *ACM Transactions on Management Information Systems (TMIS)*, 12(1), 1-20.
5. Easttom, C. (2022). Quantum computing and cryptography. In *Modern Cryptography: Applied Mathematics for Encryption and Information Security* (pp. 397-407). Cham: Springer International Publishing.
6. Rozlomii, I., Yarmilko, A., & Naumenko, S. (2024, April). Data security of IoT devices with limited resources: challenges and potential solutions. In *Proceedings of the 4th Edge Computing Workshop (doors 2024)*, Zhytomyr, Ukraine (pp. 85-96).
7. Rozlomii, I., Yarmilko, A., Naumenko, S., & Mykhailovskyi, P. (2023). IoT Smart Implants: Information Security and the Implementation of Lightweight Cryptography. In *Proceedings of the 6th International Conference on Informatics & Data-Driven Medicine (IDDM'2023)* (pp. 145-146). Bratislava.
8. Abd Ali, S. M., & Hasan, H. F. (2019). Novel encryption algorithm for securing sensitive information based on feistel cipher. *Test Engineering Management*, 19(80), 10-16.
9. Mahmood, Z., Ning, H., & Ghafoor, A. (2016, December). Lightweight two-level session key management for end user authentication in Internet of Things. In *2016 IEEE international conference on internet of things (iThings) and IEEE Green computing and communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)* (pp. 323-327). IEEE.
10. Fernández-Caramés, T. M. (2019). From pre-quantum to post-quantum IoT security: A survey on quantum-resistant cryptosystems for the Internet of Things. *IEEE Internet of Things Journal*, 7(7), 6457-6480.
11. Panda, S. S., Jena, D., Mohanta, B. K., Ramasubbareddy, S., Daneshmand, M., & Gandomi, A. H. (2021). Authentication and key management in distributed iot using blockchain technology. *IEEE Internet of Things Journal*, 8(16), 12947-12954.
12. Yugha, R., & Chithra, S. (2020). A survey on technologies and security protocols: Reference for future generation IoT. *Journal of Network and Computer Applications*, 169, 102763.
13. Restuccia, G., Tschofenig, H., & Baccelli, E. (2020, December). Low-power IoT communication security: On the performance of DTLS and TLS 1.3. In *2020 9th IFIP International Conference on Performance Evaluation and Modeling in Wireless Networks (PEMWN)* (pp. 1-6). IEEE.
14. Pop, A. A. (2022). Incremental encoder speed acquisition using an STM32 microcontroller and NI ELVIS. *Sensors*, 22(14), 5127.
15. Rozlomii I.O., Kosenyuk G.V., Naumenko S.V., Mykhaylovskiy P.V. (2023) Modeling a Microcontroller-Based Sensor System in a Smart Home Game Simulation Using Encryption. *Computer-Integrated technologies: education, science, production*. (53), 292-299. <https://doi.org/10.36910/6775-2524-0560-2023-53-43>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-25>

УДК 004.932.2

Седінкін Олександр Анатолійович¹, студент

<https://orcid.org/0009-0000-6500-4887>

Деркач Марина Володимирівна^{1,2}, к.т.н., доцент

<https://orcid.org/0000-0001-8977-2776>

Скарга-Бандурова Інна Сергіївна², д.т.н., професор

<https://orcid.org/0000-0003-3458-8730>

Матюк Данило Сергійович¹, студент

<https://orcid.org/0000-0001-5851-8433>

¹Східноукраїнський національний університет імені Володимира Даля, м. Київ, Україна

²Тернопільський національний технічний університет імені Івана Пулюя, м. Тернопіль, Україна

СИСТЕМА ДЛЯ ВІДСТЕЖЕННЯ РУХУ ОЧЕЙ НА ОСНОВІ МАШИННОГО НАВЧАННЯ

Седінкін О.А., Деркач М.В., Скарга-Бандурова І.С., Матюк Д.С. Система для відстеження руху очей на основі машинного навчання. У статті розглянута актуальна технологія визначення точки фіксації погляду, що надає можливість інтерактивної взаємодії у різних сферах, зокрема у вдосконаленні комп'ютерних інтерфейсів. Розроблена система для відстеження руху очей ґрунтується на методи HOG та лінійному алгоритмі машинного навчання SVM й реалізована на мові програмування Python з підключенням бібліотек: OpenCV, Dlib, numpy. Представлена система виділяє зіниці на зображенні обличчя людини, визначає координати зіниць лівого і правого ока, визначає напрямок погляду людини.

Ключові слова: система, розпізнавання обличчя, машинне навчання, технологія, зіниці, око.

Sedinkin O., Derkach M., Skarga-Bandurova I., Matiuk D. Eye tracking system based on machine learning. The article discusses current technology for determining the gaze fixation point, which provides possibility of interactive interaction in various areas in computer interfaces. The developed eye tracking system is based on the HOG method and linear SVM and is implemented in Python with libraries: OpenCV, Dlib, numpy. The presented system selects the pupils in the image of a person's face, determines the coordinates of the pupils of the left and right eye, and determines the direction of the person's gaze.

Keywords: system, face recognition, machine learning, technology, pupils, eye.

Постановка проблеми. Вивчення траєкторії руху очного яблука стає ключем до розуміння процесів розпізнавання зорових образів та когнітивних механізмів, що охоплюють всю глибину внутрішнього світу людського розуму [1]. Крім того, відстеження руху очей відкриває нові перспективи для розробки новаторських методів взаємодії між людиною та технічними системами, надаючи можливість впливу на них через найпростіший інструмент — погляд. Саме тому, актуальність технологій визначення точки фіксації погляду стає основною науково-технічною дилемою, яка стимулює багаторічні дослідження та інноваційні розробки. Ця технологія надає можливість інтерактивної взаємодії у різних сферах, зокрема в ігровій індустрії, розробці інтерфейсів користувача [2], віртуальній та доповненій реальності, аналізі користувацьких пристроїв та додатків, а також в галузі кібербезпеки та ідентифікації [3]. Технологія відстеження руху очей на основі машинного навчання відіграє ключову роль у вдосконаленні інтерфейсів та взаємодії з комп'ютером.

Аналіз останніх досліджень і публікацій. Останнім часом застосування технології відстеження руху очей значно поширюється на різноманітні області та відкриває нові перспективи для розвитку цифрового світу. Відомі підходи для вимірювання та аналізу рухів очей включають електроокулографію (ЕОГ), електромагнітні методи, методи відеоокулографії, тощо.

ЕОГ є біомедичними сигналами, зазвичай які отримують від ЕЕГ-електродів і використовують для виявлення руху очей і моргання [4]. Так, наприклад, дослідження [5] зосереджено на зменшенні кількості електродів за допомогою методів вибору ознак без будь-яких наслідків для статистичних показників ефективності ЕЕГ-пристроїв для прогнозування стану очей, досягнута точність склала 96,3% за допомогою класифікатора KStar (K*). А у дослідженні [6] було навчено дев'ять алгоритмів машинного навчання, таких як штучні нейронні мережі, лінійний дискримінантний аналіз, дерево рішень, метод k-найближчих сусідів, наївний баєсів класифікатор і метод опорних векторів, тощо, на наборі даних, створеному шляхом виділення ознак ЕЕГ-сигналів.

В той час, як відеоокулографія заснована на записі положення очей за допомогою відеокамер і може виконуватися кількома техніками: відеоокулографія межі лімба/райдужки і склери, зінична відеоокулографія й відеоокулографія відображень зіниці та рогівки. Для прикладу у статті [7]

автори досліджують як можна передбачити стан очей (відкриті чи закриті) шляхом вимірювання мозкових хвиль за допомогою камери під час вимірювання ЕЕГ, протестовано 42 різні алгоритми машинного навчання. А у дослідженні [8] запропонували нову бінокулярну систему виявлення зіниці та погляду, яка використовує дистанційну камеру з роздільною здатністю Full HD і використовує алгоритми обробки зображень LabVIEW.

Метою статті є розроблення системи відстеження руху очей на основі машинного навчання для визначення положення зіниць та напрямку погляду людини на зображенні.

Вирішення проблеми. У центрі сприйняття світу людською свідомістю знаходиться заплутана мережа зорової системи, що складається з інтегрованого комплексу ока, нервових шляхів та зорового аналізатора головного мозку. Існує 52 поля, що визначив німецький невропатолог Корбініан Бродман завдяки поділу кори великих півкуль на цитоархітектонічні зони, згруповані в 11 гістологічних регіонів, так як ці ділянки мають різну структуру та виконують різні функції. Результат зорової активності проходить через бічні колінчасті ядра таламусу, а потім досягає зорової кори, тобто частини кори головного мозку, що відіграє важливу роль в обробці візуальної інформації. Обидві півкулі мозку містять зорову кору. Зорова кора лівої півкулі отримує сигнали від правого зорового поля, а зорова кора в правій півкулі отримує сигнал від лівого поля зору. Зорова кора розташована в потиличній частці, в задній частині черепа. Частина зорової кори, яка отримує сенсорні входи від таламуса, називається первинною зоровою корою, або першою візуальною зоною (відповідає полю Бродмана 17). Частина екстрастріарної кори, яка складається з зорових областей другої, третьої, четвертої і п'ятої, відповідає полям Бродмана 18 і 19. Стійка активність у зоровій корі передається з потилично-скроневої ділянки задньої частини мозку назад у префронтальну кору. І навіть після зникнення візуальної активності, мозкова активність значно знижується, але зберігається. Так працює мозок людини, що дає можливість миттєво розпізнати будь-який об'єкт.

Що ж стосується технічного зору, то комп'ютери на сьогодні не здатні до узагальнень високого рівня і для задач розпізнавання [9], в тому числі задачі відстеження руху очей, потрібно побудувати систему, що складається з кількох етапів розпізнавання окремих об'єктів і послідовно поєднує декілька алгоритмів машинного навчання.

Звісно існують сервіси розпізнавання й порівняння обличчя на зображенні/відео, такі як, наприклад, Microsoft Azure Face API, Google Cloud Vision API та Amazon Rekognition Image, але це комерційні проєкти.

Одним з головних критеріїв роботи подібних сервісів, є час, за який система розпізнає обличчя на зображенні/відео, особливо для інтерфейсів, де в режимі реального часу будується людино-машинна взаємодія. Результат порівняння за часом опрацювання фото згаданих вище сервісів і розробленої системи показано в таблиці 1.

Таблиця 1. Порівняння часу, витраченого для пошуку обличчя

Назва сервісу	Face ID	Microsoft Azure Face API	Google Cloud Vision API	Amazon Recognition Image	Розроблена система
Час опрацювання, с	0,77	3,08	6	1,77	1,67

Отримавши прийнятний результат розпізнавання обличчя на першому етапі, для наступного етапу реалізації розробки побудовано систему для визначення положення зіниць та напрямку погляду людини.

Тепер безпосередньо представимо розроблену систему для відстеження руху очей, що реалізовано на мові програмування Python з підключенням бібліотек: OpenCV, Dlib, numpy, й ґрунтується на методі HOG (Histogram of Oriented Gradients, гістограма спрямованих градієнтів) та лінійному алгоритмі машинного навчання SVM (Support Vector Machines, метод опорних векторів):

1. Основна ідея HOG полягає в наступних кроках:

Крок 1. Поділити зображення на маленькі з'єднані клітинки.

Крок 2. Обчислити гістограму для кожної клітинки.

Крок 3. Об'єднати всі гістограми разом, щоб сформувати одну унікальну для кожного обличчя гістограму.

2. SVM є алгоритмом класифікації, що визначає гіперплощину, яка робить поділ між класами. Функцію прийняття рішень SVM можна виразити як:

$$f(x) = \text{sign} \left(\sum_{i=1}^N a_i y_i K(x_i, x) + b \right),$$

де N — розмір навчальних даних, K — функція ядра, яка вимірює подібність між x_i (опорний вектор) і x (значення ознак), a_i — множник Лагранжа, y_i представляє клас належності кожного даного (± 1), а b — числова константа.

Лінійний SVM: функція ядра виражається як внутрішній добуток опорного вектора та значень ознак:

$$K(x_i, x) = x_i^T \cdot x.$$

Фактично, HoG застосовується для обчислення дескрипторів, що використовуються для навчання лінійної SVM, яка виявляє обличчя.

Алгоритм функціонування системи зображено на рис. 1. Спочатку створюється відеопотік, у циклі зчитуються кадри відеопотоку один за одним, після чого кожен кадр перетворюється в чорно-біле зображення. Потім відбувається пошук та виявлення обличчя на кадрі завдяки методу HOG та лінійному SVM.

Наступний крок - аналіз кожного кадру для розпізнавання обличчя та ініціалізація очей на зображенні. Спочатку знаходимо обличчя, після чого визначаємо ключові точки (орієнтири) обличчя за допомогою навченої моделі. Ці орієнтири включають точки навколо очей, що використовується для визначення положень очей та зіниць, й містять порогове значення для бінаризації зображення очей для визначення райдужної оболонки ока. Під час знаходження орієнтирів також калібруємо алгоритм, для покращення результатів. Далі обробляємо основний кадр з виділеними зіницями і визначаємо напрям погляду користувача.

Для перевірки положення зіниць використовується метод, завдяки якому можна отримати цілі значення координат x та y , що вказують на координати центра для зіниць лівого і правого ока:

Лістинг 1. Метод визначення положення зіниць

```
def pupils_located(self):
    try:
        int(self.eye_left.pupil.x)
        int(self.eye_left.pupil.y)
        int(self.eye_right.pupil.x)
        int(self.eye_right.pupil.y)
        return True
    except Exception:
        return False
```

Потім візуалізуємо кожну зіницю на кадрі, щоб перевірити, що зіниці були знайдені та правильно проаналізовані. Система призначена для розпізнавання руху очей однієї людини в кадрі.

Результати досліджень. Розроблена система протестована за допомогою камери мобільного телефону Redmi 9 Pro, що розташована перед людиною у фіксованому положенні навпроти її голови.

Характеристики камери: матриця 64 мегапікселі; роздільна здатність: 9248 x 6920; сенсор: Samsung Bright S5KGW1; діафрагма: $f/ 1.89$; розмір пікселю: 0.80 μm ; розмір датчику: 1/1.72"; зчитування кадру: до 60 кадрів при записуванні відео 1080 роздільності.

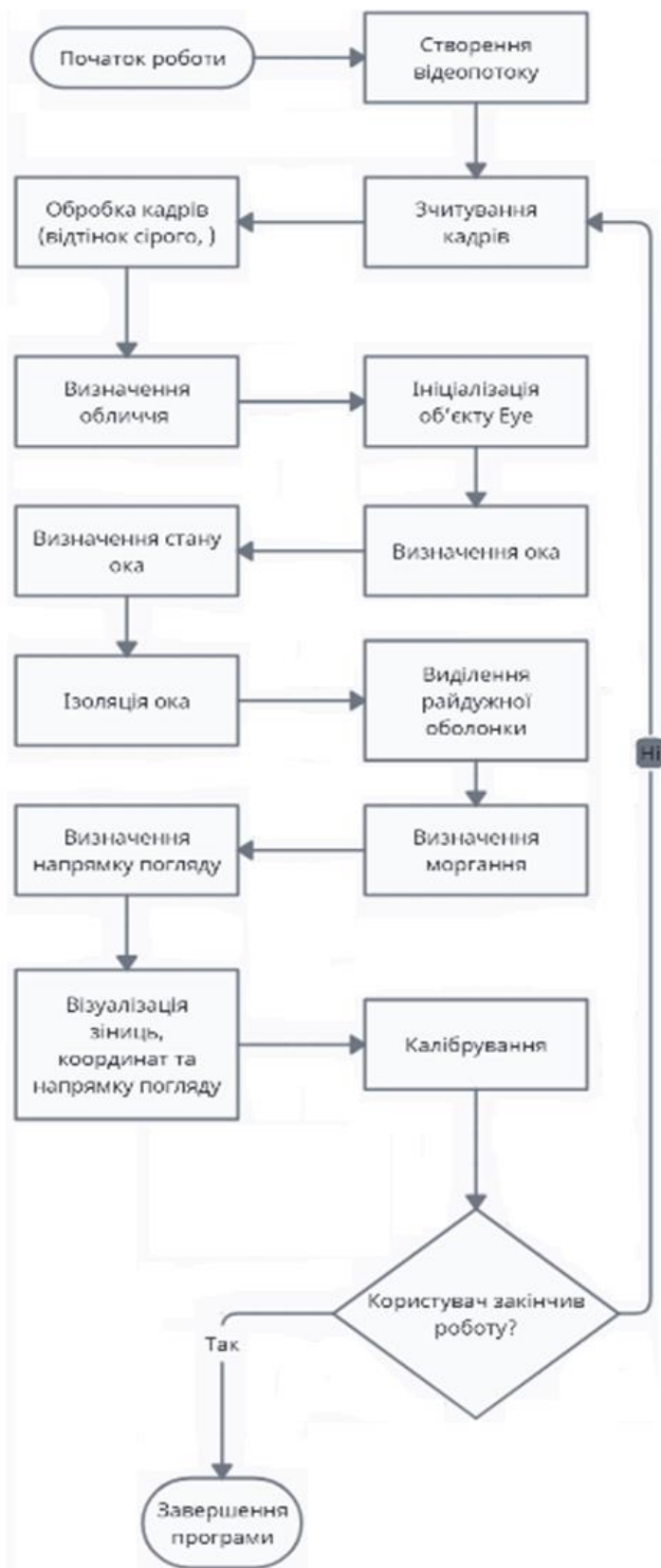


Рис. 1. Алгоритм системи для відстеження руху очей на основі машинного навчання
Результати тестування для визначення координат центрів зіниць та напрямку погляду
вліво/вправо представлені на рисунку 2.

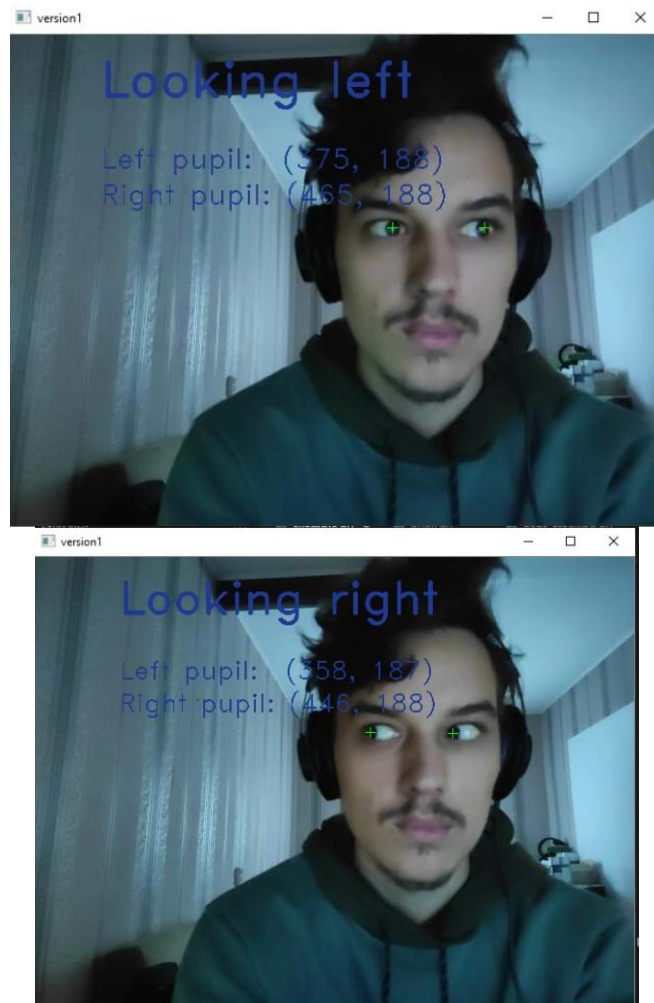


Рис. 2 – Визначення напрямку погляду

Проведено тестування системи в стані заплющених очей та в умовах слабкої освітленості приміщення (рис. 3).

Результати тестування підтверджують, що система для відстеження руху очей:

- виділяє зіниці на зображенні обличчя людини;
- визначає координати зіниць лівого і правого ока;
- визначає напрямок погляду людини.

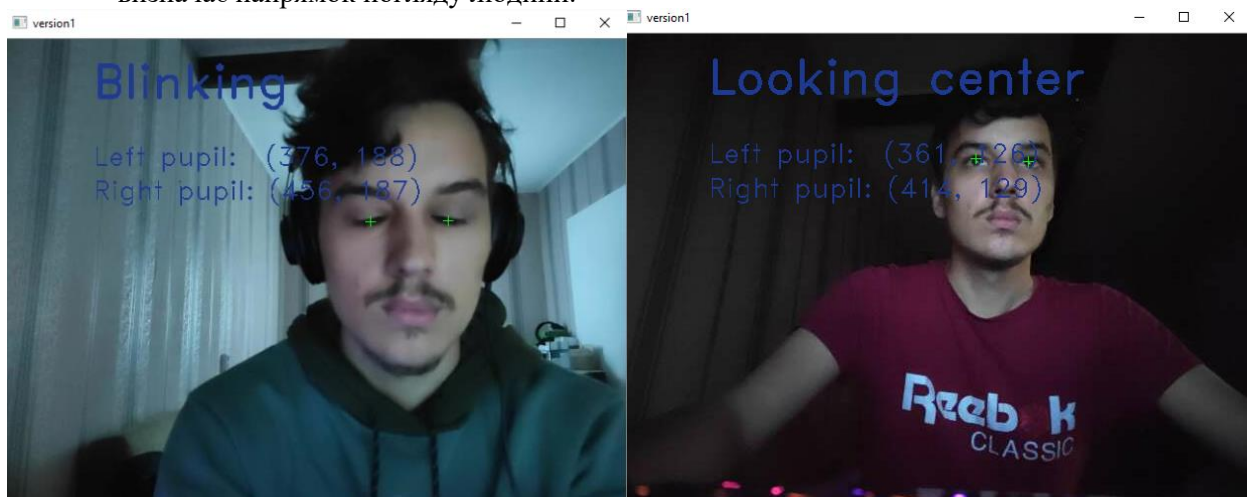


Рис. 3 – Результати тестування системи

Крім того, визначено, що система розпізнає стан заплющених очей, та однаково добре працює при помірній та слабкій освітленості приміщення. При цьому виявлено недоліки системи: при надмірній освітленості приміщення система може не виявити очі, а також при швидкому русі голови система не встигає обробляти результати і перестає визначати напрям погляду та координати зіниць. В подальшому планується вирішення цих проблем та покращення точності розпізнавання шляхом збільшення частоти кадрів камери та збільшення набору даних для навчання моделі.

Висновки. Розроблено систему для відстеження руху очей на основі машинного навчання. Система здатна виділяти зіниці на зображенні обличчя людини, визначати координати зіниць лівого і правого ока, визначати напрям погляду людини. Для виявлених недоліків запропоновано методи їх вирішення. З перспективних напрямів розвитку, можна очікувати застосування систем відстеження руху очей для особистісної ідентифікації, оскільки характер рухів очей має індивідуальну природу, схожу на почерк, але неможливий для точної імітації. Ще один значний напрям застосування подібних систем полягає у комп'ютерних інтерфейсах, наприклад, можливість керувати курсором при роботі з ПК без допомоги рук, що допоможе людям з обмеженими можливостями після травм чи захворювань продовжувати жити повноцінним життям.

Список бібліографічного опису

1. Mai, T.D.T., Phung, TN. (2023). Evaluating the Performance of Some Deep Learning Model for the Problem of Emotion Recognition Based on EEG Signal. In: Nghia, P.T., Thai, V.D., Thuy, N.T., Son, L.H., Huynh, VN. (eds) *Advances in Information and Communication Technology. ICTA 2023. Lecture Notes in Networks and Systems*, vol 847. Springer, Cham. https://doi.org/10.1007/978-3-031-49529-8_19.
2. М. В. Деркач, Д. С. Матюк. Дослідження технології ЕЕГ на прикладі реалізованого нейрокомп'ютерного інтерфейсу. Сучасні інформаційні технології та системи в управлінні [Електронний ресурс]: Зб. матеріалів IV Міжнар. наук.-практ. конф. молодих вчених, аспірантів і студентів; 20–21 квітня 2023 р. Київ: КНЕУ, 2023. С. 81 – 83.
3. Jayarathne I, Cohen M, Amarakeerthi S. (2020) Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio. *PLoS ONE* 15(9): e0238872. <https://doi.org/10.1371/journal.pone.0238872>.
4. Д. С. Матюк, М. В. Деркач. Оцінка спектральної щільності потужності ЕЕГ сигналу. Актуальні задачі сучасних технологій : зб. тез доповідей XII міжнар. наук.-практ. конф. Молодих учених та студентів, (Тернопіль, 6-7 грудня 2023). М-во освіти і науки України, Терн. націон. техн. ун-т ім. І. Пулюя [та ін.]. Тернопіль: ФОП Паляниця В. А., 2023. С. 404-405.
5. Uwisengeyimana, J. de D., Khalid Al_Salihi, N., & Ibriki, T. (2016). Statistical Performance Effect of Feature Selection Techniques on Eye State Prediction Using EEG. *International Journal of Statistics in Medical Research*, 5(3), 224–230. <https://doi.org/10.6000/1929-6029.2016.05.03.9>.
6. Ramirez-Arias, F.J.; Garcia-Guerrero, E.E.; Tlelo-Cuautle, E.; Colores-Vargas, J.M.; Garcia-Canseco, E.; López-Bonilla, O.R.; Galindo-Aldana, G.M.; Inzunza-González, E. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies* 2022, 10, 79. <https://doi.org/10.3390/technologies10040079>.
7. Rösler, Oliver and David Suendermann. "A First Step towards Eye State Prediction Using EEG." (2013).
8. Durna Y, Ari F. Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images. *Applied Sciences*. 2017; 7(5):498. <https://doi.org/10.3390/app7050498>.
9. Y. Boltov, I. Skarga-Bandurova and M. Derkach, "A Comparative Analysis of Deep Learning-Based Object Detectors for Embedded Systems," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1156-1160, doi: 10.1109/IDAACS58523.2023.10348642.

References

1. Mai, T.D.T., Phung, TN. (2023). Evaluating the Performance of Some Deep Learning Model for the Problem of Emotion Recognition Based on EEG Signal. In: Nghia, P.T., Thai, V.D., Thuy, N.T., Son, L.H., Huynh, VN. (eds) *Advances in Information and Communication Technology. ICTA 2023. Lecture Notes in Networks and Systems*, vol 847. Springer, Cham. https://doi.org/10.1007/978-3-031-49529-8_19.
2. M. V. Derkach, D. S. Matiuk. Research of EEG technology on the example of a implemented neurocomputer interface. *Modern information technologies and management systems [Electronic resource]: Collection. materials IV International Scientific and Practical Conference of Young Scientists, Postgraduate Students and Students; April 20–21, 2023. Kyiv: KNEU, 2023. P. 81–83.*
3. Jayarathne I, Cohen M, Amarakeerthi S (2020) Person identification from EEG using various machine learning techniques with inter-hemispheric amplitude ratio. *PLoS ONE* 15(9): e0238872. <https://doi.org/10.1371/journal.pone.0238872>.
4. D. S. Matiuk, M. V. Derkach. Estimation power spectral density of EEG signal. *Actual problems of modern technologies: book of abstracts of the XII International scientific and practical conference of young researchers and*

- students, (Ternopil, December, 6th-7th, 2023). Ministry of Education and Science of Ukraine, Ternopil Ivan Puluj National Technical University [and other.]. Ternopil: PE Palianytsia V.A., 2023. P. 404-405.
5. Uwisengeyimana, J. de D., Khalid Al_Salihy, N., & Ibrikci, T. (2016). Statistical Performance Effect of Feature Selection Techniques on Eye State Prediction Using EEG. *International Journal of Statistics in Medical Research*, 5(3), 224–230. <https://doi.org/10.6000/1929-6029.2016.05.03.9>.
 6. Ramírez-Arias, F.J.; García-Guerrero, E.E.; Tlelo-Cuautle, E.; Colores-Vargas, J.M.; García-Canseco, E.; López-Bonilla, O.R.; Galindo-Aldana, G.M.; Inzunza-González, E. Evaluation of Machine Learning Algorithms for Classification of EEG Signals. *Technologies* 2022, 10, 79. <https://doi.org/10.3390/technologies10040079>.
 7. Rösler, Oliver and David Suendermann. "A First Step towards Eye State Prediction Using EEG." (2013).
 8. Durna Y, Ari F. Design of a Binocular Pupil and Gaze Point Detection System Utilizing High Definition Images. *Applied Sciences*. 2017; 7(5):498. <https://doi.org/10.3390/app7050498>.
 9. Y. Boltov, I. Skarga-Bandurova and M. Derkach, "A Comparative Analysis of Deep Learning-Based Object Detectors for Embedded Systems," 2023 IEEE 12th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Dortmund, Germany, 2023, pp. 1156-1160, doi: 10.1109/IDAACS58523.2023.10348642.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-26>

УДК 004.4

Філь Наталія Юріївна, к.т.н., доцент

<https://orcid.org/0000-0003-2081-7176>

Харківський національний автомобільно-дорожній університет, м. Харків, Україна

МОДЕЛЬ ВИБОРУ ДАТЧИКА ТЕМПЕРАТУРИ ТА ВОЛОГОСТІ ДЛЯ МОДУЛЯ СИСТЕМИ ЕКОЛОГІЧНОГО МОНІТОРИНГУ

Філь Н.Ю. Моделі вибору датчика температури та вологості для модуля системи екологічного моніторингу.

Система екологічного моніторингу в Україні є застарілою та не дозволяє збирати дані в режимі реального часу та накопичувати їх в єдиному цифровому форматі. Сьогодні ситуація особливо загострилася, бо деякі метеорологічні станції зруйновані, тому заміри концентрацій забруднюючих речовин проводяться нерегулярно, що впливає на отримані результати. Проблема реінжинірингу системи екологічного моніторингу є актуальною. В роботі розроблена модель вибору датчика температури та вологості повітря для модуля системи екологічного моніторингу. Наведено приклад вибору датчику за обраними критеріями.

Ключові слова: екологічний моніторинг, датчик, критерій, альтернатива, функція корисності.

Fil N. Model for Selecting a Temperature and Humidity Sensor for a Module of an Environmental Monitoring System. The system of environmental monitoring in Ukraine is outdated, it does not allow accumulating data in a single digital format and viewing them in real time. Today, the situation has become particularly acute, as some meteorological stations have been destroyed, so measurements of pollutant concentrations are not carried out regularly, which affects the results obtained. The problem of re-engineering the environmental monitoring system is urgent. The paper develops models for selecting a temperature and humidity sensor for a module of an environmental monitoring system. An example of sensor selection according to the selected criteria is given.

Keywords: environmental monitoring, sensor, criterion, alternative, utility function.

Постановка наукової проблеми

Екологічна катастрофа довкілля, яку спричинила війна, не знає кордонів, адже циркуляція повітря, води поширюють негативний вплив війни по всьому світу. Необхідно вже зараз забезпечити ефективне подолання екологічних наслідків війни, яке неможливо зробити без сучасної системи екологічного моніторингу [1].

В Україні існує державна системи екологічного моніторингу (СЕМ). Проби атмосферного повітря відбираються з періодичністю 3-4 рази на добу 6 днів на тиждень на 129 державних стаціонарних постах. Під обстріли попадають промислові об'єкти, склади нафтопродуктів, тому необхідно проводити постійно моніторингу повітря в онлайн режимі. Але на більшості постів використовується старе апаратне забезпечення, тому потрібно проводити реінжиніринг системи екологічного моніторингу. Всі ці недоліки не дозволяють приймати ефективні рішення щодо зменшення негативного впливу на довкілля [2].

В прийнятій «Концепції Державної цільової екологічної програми моніторингу довкілля» зазначено, що екологічна ситуація в Україні є вкрай складною, навантаження на навколишнє природне середовище зростає. У більшості міст за окремими показниками спостерігається високий рівень забруднення атмосферного повітря. Така незадовільна екологічна ситуація зумовлена рядом факторів, зокрема неефективним функціонуванням державної системи моніторингу навколишнього природного середовища [3].

Через воєнний стан та економічні складнощі державна система екологічного моніторингу не відповідає викликам та вимогам сьогодення, потребує реінжинірингу апаратного забезпечення на основі сучасних технологій.

Аналіз досліджень. Проблеми екологічного моніторингу довкілля присвячено багато досліджень вітчизняних та закордонних авторів.

Захист повітря від забруднення є важливим елементом європейської політики захисту довкілля. В державах Європейського Союзу міжнародний стандарт з управління навколишнім середовищем ведеться з початку 1990-х рр [4].

Забруднення повітря є першою у списку найбільших екологічних проблем у Європі. На якість повітря в державах ЄС впливає рівень забруднення повітря, а концентрація твердих частинок значно перевищує останні рекомендації Всесвітньої організації охорони здоров'я.

Повітря є одним з елементів природного середовища, захист якого належить до пріоритетних напрямків державної політики в Польщі. Станції моніторингу якості атмосферного

повітря встановлюються за рахунок Воєводського фонду охорони довкілля і водного господарства, а результати щогодинного автоматичного вимірювання забруднення повітря доступні на порталі «Якість повітря» в модулі «Поточні дані вимірювань» та в мобільних програмах «Якість повітря в Польщі» [4].

Україна їде по шляху вдосконалення системи оцінки школи довкіллю. Нещодавно, підписана угода про співпрацю з Фінляндією для розбудови національної системи моніторингу води та біорізноманіття [5]. Проект технічної допомоги планують реалізувати у 2024-2027 роках, що дозволить покращити екологічний води і біорізноманіття в Україні.

В роботі [6] автори пропонують використання технологія Інтернету речей (IoT), що дозволить збирати, обробляти дані в реальному режимі. Така система повинна бути розроблена з використанням датчиків, мікроконтролерів на основі IoT. Це дозволить проводити ефективний екологічний моніторинг та прогнозувати зміни навколишнього середовища.

В роботі [7] зазначено, що використання сучасних інформаційних технологій з використанням геопросторових даних екологічних викликів скоротить час прийняття управлінських рішень, а також. Необхідно проводити постійний моніторинг атмосферного повітря, у частині впливу постійно тліючих торф'яних полів навколо Києва, попередження пожеж у лісах, вибухів на об'єктах нафтопереробної промисловості.

В роботах [8-9] розглядається застосування бездротових датчиків на базі Інтернету речей для контролю за довкіллям. Використовуються дві методології, перша з яких – зв'язок Wi-Fi на основі протоколу користувацьких дейтаграм (UDP), а інші – через Wi-Fi, а також протокол передачі машиночитаемого тексту (HTTP). Обидві системи дозволяють спостерігати за географічно широкими територіями.

В роботі [10] автори розглянули основні вимоги для СЕМ. Запропоновано структурна схема багато сенсорної системи екологічного моніторингу. Автори розглянули взаємозв'язок різних елементів СЕМ.

Таким чином, через воєнний стан та економічні складнощі державна система екологічного моніторингу не відповідає викликам та вимогам сьогодення та потребує реінжинірингу.

Мета та задачі дослідження

Метою дослідження є розробка моделей датчика температури та вологості для модуля системи екологічного моніторингу (МСЕМ) в умовах невизначеності вхідної інформації, що дозволить автоматизувати процес розробки модуля системи екологічного моніторингу та підвищити швидкість збору і точність даних, які спостерігаються.

Для досягнення поставленої мети необхідно виконати наступні завдання:

- провести аналіз проблеми вибору датчика температури та вологості для МСЕМ;
- розробити модель вибору датчику температури та для МСЕМ в умовах невизначеності вхідної інформації;
- навести приклад використання розробленої моделі.

Модель вибору датчика температури та вологості повітря для МСЕМ

Відомо множина датчиків для вимірювання температури та вологості повітря сумісні з мікроконтролером, який буде використовуватися для МСЕМ $STV = \{STV_j\}$, $(j = \overline{1, j^*})$, які можливо використовувати для модуля системи екологічного моніторингу, де j^* – кількість датчиків для вимірювання температури та вологості повітря.

Кожен датчик для вимірювання температури та вологості повітря характеризується наступними параметрами:

- діапазон вимірювання температури – $STV^{DT} = \{STV_j^{DT}\}$, $(j = \overline{1, j^*})$;
- діапазон вимірювання вологості – $STV^B = \{STV_j^B\}$, $(j = \overline{1, j^*})$;
- похибка вимірювання температури – $STV^{IT} = \{STV_j^{IT}\}$, $(j = \overline{1, j^*})$;
- похибка вимірювання вологості – $STV^{IB} = \{STV_j^{IB}\}$, $(j = \overline{1, j^*})$;
- максимальне значення струму, що споживається – $STV^C = \{STV_j^C\}$, $(j = \overline{1, j^*})$;

– час відгуку – $STV^{CB} = \{STV_j^{CB}\}, (j = \overline{1, j^*})$

– вартістю – $STV^{BAP} = \{STV_j^{BAP}\}, (j = \overline{1, j^*})$.

Введемо змінну $x_j^{STV} = 1$, якщо обрано j - датчик для вимірювання температури та вологості повітря, та $x_j^{STV} = 0$ в протилежному випадку.

Основними вимогами для вибору датчика для вимірювання температури та вологості повітря системи екологічного моніторингу є:

– мінімальна вартість датчика для вимірювання температури та вологості повітря для модуля системи екологічного моніторингу

$$STV^{BAP} = \sum_{j=1}^{j^*} x_j^{STV} STV_j^{BAP} \rightarrow \min ; \quad (1)$$

– максимальний діапазон вимірювання температури

$$STV^{DT} = \sum_{j=1}^{j^*} x_j^{STV} STV_j^{DT} \rightarrow \max ; \quad (2)$$

– максимальний діапазон вимірювання вологості

$$STV^B = \sum_{j=1}^{j^*} x_j^{STV} STV_j^B \rightarrow \max ; \quad (3)$$

– мінімальна похибка вимірювання температури

$$STV^{ITT} = \sum_{j=1}^{j^*} x_j^{STV} STV_j^{ITT} \rightarrow \min ; \quad (4)$$

– мінімальна похибка вимірювання вологості

$$STV^{IB} = \sum_{j=1}^{j^*} x_j^{STV} STV_j^{IB} \rightarrow \min ; \quad (5)$$

– мінімальний час відгуку

$$STV^{CB} = \sum_{j=1}^{j^*} x_j^{STV} STV_j^{CB} \rightarrow \min . \quad (6)$$

Область допустимих рішень визначається обмеженнями:

– витрати на датчика для вимірювання температури та вологості повітря системи екологічного моніторингу не повинні бути більше заданих $STV_{зад}^{BAP}$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^{BAP} \leq STV_{зад}^{BAP} ; \quad (7)$$

– діапазон вимірювання температури повинен бути не менше заданого $STV_{зад}^{DT}$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^{DT} \geq STV_{зад}^{DT} ; \quad (8)$$

– діапазон вимірювання вологості повинен бути не менше заданого $STV_{зад}^B$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^B \geq STV_{зад}^B ; \quad (9)$$

– похибка вимірювання температури повинна бути не більше заданої $STV_{зад}^{ITT}$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^{ITT} \leq STV_{зад}^{ITT} ; \quad (10)$$

– похибка вимірювання вологості повинна бути не більше заданої $STV_{зад}^{IIB}$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^{IIB} \leq STV_{зад}^{IIB} ; \quad (11)$$

– час відгуку повинен бути не більше заданого $STV_{зад}^{ЧВ}$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^{ЧВ} \leq STV_{зад}^{ЧВ} ; \quad (12)$$

– максимальний струм, що споживається повинен бути не більше заданого $STV_{зад}^C$

$$\sum_{j=1}^{j^*} x_j^{STV} STV_j^C \leq STV_{зад}^C ; \quad (13)$$

– необхідно обрати тільки один датчик для вимірювання температури та вологості повітря

$$\sum_{j=1}^{j^*} x_j^{STV} = 1. \quad (14)$$

Таким чином, розроблена модель вибору датчика температури та вологості повітря для МСЕМ, що дозволяє вибрати ефективний датчик температури та вологості повітря за заданими критеріями й обмеженням в умовах нечіткої вхідної інформації.

Розроблена модель (1) – (14) відноситься до задач багатокритеріального дискретного програмування з булевими змінними.

Найбільш поширений підхід до розв'язання багатокритеріальної задачі – її зведення до однокритеріальної [11]. Основою такого підходу є теорія корисності.

Для вирішення задачі використовується функція корисності часткових критеріїв $R_f(k_f)$, $f = 1, F$, яка є універсальною та відображає особливості конкретних альтернатив і критеріїв [7]

$$R_f(k_f) = \frac{k_f - k_f^{HG}}{k_f^{HK} - k_f^{HG}}, f = \overline{1, F}, \quad (15)$$

де k_f, k_f^{HG}, k_f^{HK} – поточне, найгірше (гранично допустиме) і найкраще значення f -го часткового критерію, що відповідають наблизеній області компромісів.

Функція корисності часткових критеріїв (15) має наступні властивості:

- має єдиний інтервал змін $[0, 1]$;
- є безрозмірною;
- інваріантною до виду екстремуму часткового критерію (min и max), тобто, найкращому значенню відповідає 1, а найгіршому – 0.

Інформація про значущість часткових критеріїв визначається безрозмірними коефіцієнтами, що враховують важливість критеріїв, тобто:

$$\sum_{f=1}^F \lambda_f = 1, \text{ де } \lambda_f \in [0,1] \text{ (} j = \overline{1, F} \text{)}. \quad (16)$$

Велика група схем компромісу заснована на принципі максимальної адитивної корисності часткових критеріїв. Якщо відомі значення вагових коефіцієнтів λ_f ($j = \overline{1, F}$) часткових критеріїв та їх функцій корисності $R_f(x)$ оцінка проектних рішень $x \in X$ та вибір найкращого x^0 проводиться за узагальненим критерієм виду [7-8]:

$$W(x^0) = \max_{x \in X} \sum_{f=1}^F \lambda_f R_f(x) \quad \left| \quad \sum_{f=1}^F \lambda_f = 1. \quad (17)$$

Метод багатокритеріальної оптимізації має певні переваги. Цей метод є більш стійким в порівнянні з методом аналізу ієрархій, та має меншу трудомісткість [11-12].

Як перевагу можна зазначити, що метод багатокритеріальної оптимізації дозволяє враховувати приховану інформацію. Експерти визначають значущість критеріїв за якими проводиться вибір альтернатив. В цьому випадку особа, що приймає рішення покладається не тільки на свій досвід прийняття рішень.

Розглянемо приклад використання розробленої моделі.

Приклад використання розробленої моделі

В якості альтернатив розглянемо датчики температури та вологості повітря, які є в наявності на ринку України. Технічні характеристики альтернатив – датчиків температури та вологості повітря, представлені в таблиці 1.

Таблиця 1. Технічні характеристики датчиків температури та вологості повітря

	Діапазон виміру температури, °C	Похибка виміру температури, °C	Діапазон виміру вологості, %	Похибка виміру вологості, %	Час відгуку, с	Максимальний струм, що споживається, мА	Вартість, грн
DHT-11	0-50	2	20-90	5	1	2	59
DHT-22	-40-80	0,5	0-100	2	2	1,5	145
GYSHT31-D	-40-125	0,3	0-100	2	1	2	277
GY-21 HTU21	-10-85	0,3	0-80	3	2	0,15	98
Sensiron SHT31-D	0-90	0,2	0-100	2	8	2	177
BME280 5B I2C	-40-85	0,5	0-100	3	2	0,35	235

В середовищі Microsoft Excel були виконані розрахунки функцій корисності за кожним критерієм для всіх альтернатив за формулою (15), а також узагальненого критерію – за формулою (17) (Табл. 2). Вагові коефіцієнти значущості критеріїв були визначені експертами та відповідають умові (16).

Таблиця 2. Значення функції корисності для альтернатив

	Похибка виміру температури, °C	Похибка виміру вологості, %	Час відгуку, с	Максимальний струм, що споживається, мА	Вартість, грн	Загальний пріоритет
DHT-11	0,000	0,000	1,000	0,000	1,000	0,450
DHT-22	0,833	1,000	0,857	0,270	0,606	0,576
GYSHT31-D	0,944	1,000	1,000	0,000	0,000	0,294
GY-21 HTU21	0,944	0,667	0,857	1,000	0,821	0,884
Sensiron SHT31-D	1,000	1,000	0,000	0,000	0,459	0,361
BME280 5B I2C	0,833	0,667	0,857	0,892	0,193	0,615

найгірше	2,000	5	8	2	277	
найкраще	0,200	2	1	0,15	59	
Важливість	0,10	0,10	0,10	0,35	0,35	1,00

Функція корисності за критеріями діапазони виміру температури та вологості не розраховувались, адже всі альтернативи мають необхідні допустимі діапазони виміру температури та вологості.

За розрахунками кращою альтернативою є датчик GY-21 HTU21, який має максимальне значення загального пріоритету.

Висновки та перспективи подальшого дослідження.

В роботі розроблена модель вибору датчика температури та вологості повітря для модуля системи екологічного моніторингу в умовах нечіткої вхідної інформації, яка дозволяє вибирати ефективний датчик за багатьма функціональними та вартісними критеріями та обмеженнями. Модель вибору датчика температури та вологості повітря для модуля системи екологічного моніторингу відноситься до задач багатокритеріального дискретного програмування з булевими змінними. Для розв'язання поставленого завдання запропоновано використовувати метод багатокритеріальної оптимізації, який є ефективним при розробці систем підтримки прийняття рішень для слабоформалізованих багатокритеріальних задач. Метод багатокритеріальної оптимізації дозволяє здійснити вибір датчика температури та вологості повітря для модуля системи екологічного моніторингу шляхом розрахунку функції корисності.

В роботі наведено приклад використання розробленої моделі вибору комутатору для системи цифрового відеоспостереження.

Подальші дослідження будуть спрямовані на розробку системи підтримки прийняття рішень, щодо вибору сучасного апаратного забезпечення для модуля системи екологічного забезпечення.

Список бібліографічного опису

1. Вишницька А. Повітря під час війни. Чому важливо моніторити забруднення та розповісти про це. <https://ua.boell.org/uk/2022/11/16/povitrya-pid-chas-viyny-chomu-vazhlyvo-monitoryty-zabrudnennya-ta-rozpovidaty-pro-tse>
2. Черноцова М. Роль повітря війни: доповідь у Верховній Раді про моніторинг якості атмосферного повітря та радіації. <https://cleanair.org.ua/7758/parlament>
3. Про схвалення Концепції Державної цільової екологічної програми моніторингу довкілля. Розпорядження Кабінету Міністрів України; Концепція від 07.07.2023 № 610-р. <https://zakon.rada.gov.ua/laws/show/610-2023-%D1%80#n9>
4. Моніторинг якості повітря у Республіці Польща. <https://epl.org.ua/environment/monitoring-atmosfernogo-povitrya-v-polshhi/>.
5. Україна посилює співпрацю з Фінляндією для розбудови національної системи моніторингу води та біорізноманіття. <https://www.kmu.gov.ua/news/ukraina-posyliuie-spivpratsiu-z-finliandiieiu-dlia-rozbudovy-natsionalnoi-systemy-monitorynhu-vody-ta-bioriznomanittia>
6. Hulwan P. D. B. Environmental Monitoring System. International Journal for Research in Applied Science and Engineering Technology. 2023. Vol. 11, no. 5. P. 3590–3595. URL: <https://doi.org/10.22214/ijraset.2023.51826>.
7. Бондар О.І., Фінін Г.С., Шевченко Р.Ю. (2022) Екологічні виклики воєнного часу: оцінка впливу на довкілля космічними системами дистанційного зондування та GPS-навігації. *Екологічні науки*. 4 (43), 40-49. <https://doi.org/10.32846/2306-9716/2022.eco.4-43.7>.
8. Mois, George, Silviu Folea, and Teodora Sanislav (2017). Analysis of three IoT-based wireless sensors for environmental monitoring. *IEEE Transactions on Instrumentation and Measurement* (66.8) 2056-2064.
9. Santhosh G., Dhanne B., Upender G., (2020). Design and Implementation of IoT-Based Wireless Sensors for Ecological Monitoring System. In: Solanki, V., Hoang, M., Lu, Z., Pattnaik, P. (eds) *Intelligent Computing in Engineering. Advances in Intelligent Systems and Computing*, Springer, Singapore (Vol 1125) https://doi.org/10.1007/978-981-15-2780-7_39.
10. Nefedov L, Fil, N. The model of the regional environmental monitoring system organization // 2023 13th International Conference on Dependable Systems, Services and Technologies (DESSERT), Athens 13 October 2023 до 15 October 2023. 1-6.
11. Петров Э. Г., Брынза Н. А., Колесник Л. В., Пискалова О.А. Методы и модели принятия решений в условиях многокритериальности и неопределенности. Херсон, 2014. 192 с.
12. Безкоровайний В. В., Петришин Л. Б., Шевченко О. Ю. (2020) Виділення підмножин ефективних варіантів в технологіях прийняття проектних рішень. *Прикладні аспекти інформаційних технологій*. (V 3. №. 1) 443–455. <https://doi.org/10.15276/aait.01.2020.6>.
13. Датчики температури. <https://arduino.ua/ru/cat38-emperatyravlajnost>

References

1. Vyshnytska A. Air during the war. Why it is important to monitor pollution and talk about it. <https://ua.boell.org/uk/2022/11/16/povitrya-pid-chas-viyny-chomu-vazhlyvo-monitoryty-zabrudnennya-ta-rozpovidaty-pro-tse>
2. Chernotsova M. The Role of the Air of War: A Report to the Verkhovna Rada on Monitoring Air Quality and Radiation. <https://cleanair.org.ua/7758/parlament>
3. On approval of the Concept of the State Target Environmental Monitoring Program. Order of the Cabinet of Ministers of Ukraine; Concept of 07.07.2023 № 610-p. <https://zakon.rada.gov.ua/laws/show/610-2023-%D1%80#n9>
4. Air quality monitoring in the Republic of Poland. <https://epl.org.ua/environment/monitoryng-atmosfernogo-povitrya-v-polshhi/>
5. Ukraine strengthens cooperation with Finland to develop a national water and biodiversity monitoring system. <https://www.kmu.gov.ua/news/ukraina-posyluie-spivpratsiu-z-finliandiieiu-dlia-rozbudovy-natsionalnoi-systemy-monitorynhu-vody-ta-bioriznomanittia>
6. Hykavchuk M.S., Petrovsky S.S., Skripnik T.K. (2019) Information technology analysis competitiveness of web portal. *Herald of Khmelnytskyi national university*, Issue 6, (279). P. 120-124. <https://doi.org/10.31891/2307-5732-2019-279-6-120-124>.
7. Bondar O., Finin G., Shevchenko R. (2022). Environmental challenges of wartime: environmental impact assessment by Space Remote Sensing Systems and GPSnavigation. *Ecological Sciences*. 4 (43), 40-49. <https://doi.org/10.32846/2306-9716/2022.eco.4-43.7>
8. Mois, George, Silviu Folea, and Teodora Sanislav (2017). Analysis of three IoT-based wireless sensors for environmental monitoring. *IEEE Transactions on Instrumentation and Measurement* (66.8) 2056-2064.
9. Santhosh G., Dhanne B., Upender G., (2020). Design and Implementation of IoT-Based Wireless Sensors for Ecological Monitoring System. In: Solanki, V., Hoang, M., Lu, Z., Pattnaik, P. (eds) *Intelligent Computing in Engineering. Advances in Intelligent Systems and Computing*, Springer, Singapore (Vol 1125) https://doi.org/10.1007/978-981-15-2780-7_39.
10. Nefedov L., Fil, N. The model of the regional environmental monitoring system organization // 2023 13th International Conference on Dependable Systems, Services and Technologies (DESSERT), Athens 13 October 2023 до 15 October 2023. 1-6.
11. Petrov E. G., Brynza N. A., Kolesnik L. V., Pisklakova O.A. *Metody i modeli priniattia reshenii v usloviakh mnogokriterial'nosti i neopredelennosti [Methods and models of decision-making in conditions of multicriteria and uncertainty]*. Kherson, 2014. 192 p.
12. Beskorovainyi V., Petryshyn L., Shevchenko O (2022). Models for the selection of car service equipment. *Applied Aspects of Information Technology*, (V3. 1), 443–455. <https://doi.org/10.15276/aait.01.2020.6>.
13. Temperature sensors. <https://arduino.ua/ru/cat38-emperatyravljajnost>

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-27>

УДК 004.42

Хамбір Владислав Русланович, магістр, головний інженер-програміст

<http://orcid.org/0009-0008-6812-2422>

Кепітал Ван, США

АВТОМАТИЗАЦІЯ ПРОЦЕСІВ ТЕСТУВАННЯ МОБІЛЬНИХ ЗАСТОСУНКІВ

Хамбір В.Р. Автоматизація процесів тестування мобільних застосунків. У статті розглянуто актуальну проблему автоматизації процесів тестування мобільних застосунків. Доведено, що автоматизована концепція має на меті полегшити й покращити процес тестування мобільних застосунків. Висвітлено сучасні тенденції та виклики, пов'язані із зазначеною темою. Розглянуто основоположні підходи та методи до автоматизації тестування мобільних застосунків, а також розглянуто новітні технології й інструменти, які можна використати для покращення процесів тестування. Розроблено схеми переваг і рівнів автоматизації тестування мобільних застосунків. Розглянуто причини автоматизації процесів тестування мобільних застосунків. Досліджено такі важливі аспекти автоматизації як тестування функціональності, інтерфейсу, продуктивності та безпеки мобільних застосунків. Акцентовано увагу на вирішенні проблем, пов'язаних з різними платформами та версіями операційних систем. На основі аналізу літературних джерел, практичних досліджень і власного досвіду автора надано рекомендації щодо ефективної автоматизації процесів тестування мобільних застосунків. Попри наявні проблеми й обмеження стосовно автоматизації, як-от початкове налаштування та крива навчання, тестування складних сценаріїв і вимоги до обслуговування, переваг значно більше, ніж недоліків. Отже, організації можуть приймати обґрунтовані рішення щодо впровадження процесів автоматизованого тестування для підвищення якості та надійності своїх мобільних застосунків, ретельно оцінюючи аргументи «за» і «проти» автоматизації.

Ключові слова: інформаційні технології, автоматизоване тестування, мобільний застосунок, тестування, тестування мобільних застосунків.

Khambir V. Automation of mobile application testing processes. The article considers the trendy problem of automating mobile application testing processes. It is proved that the automated concept aims to facilitate and improve the testing process of mobile applications. Current trends and challenges related to this topic are highlighted. The basic approaches and methods to automate mobile application testing are described, and the latest technologies and tools that can be used to improve testing processes are considered. Schemes of advantages and levels of automation of mobile application testing are developed. A scheme for classifying approaches to automating mobile application testing is presented. Ways to automate the testing processes of mobile applications are considered and demonstrated. The stages of automation of mobile application testing processes are clarified, which is shown in the table. The reasons for automating mobile application testing processes are considered. It explores such important aspects of automation as testing mobile application functionality, interface, performance, and security. Attention is focused on solving problems related to different platforms and versions of operating systems. Based on the analysis of literature sources, practical research, and the author's own experience, recommendations for effective automation of mobile application testing processes are provided. While there are challenges and limitations to automation, such as initial setup and learning curve, testing complex scenarios, and maintenance requirements, the benefits often outweigh the drawbacks. Consequently, organizations can make informed decisions about implementing automated testing processes to improve the quality and reliability of their mobile apps by carefully evaluating the pros and cons of automation.

Keywords: information technology, automated testing, mobile application, testing, mobile application testing.

Постановка наукової проблеми. Стрімке зростання ролі тестування призвело до формування процесу автоматизації, яке безпосередньо впливає на якість кінцевого продукту. Об'єктивно те, що приріст популярності мобільних застосунків у сучасному світі спричинив збільшення вимог до їх якості й надійності. Варто переконатися під час випуску нових мобільних застосунків, що вони належним чином працюють, не містять помилок і задовольняють потреби користувачів. Однак, при цьому процес тестування мобільних застосунків може бути кропітким і витратним. Необхідно враховувати швидкість технологічних змін і різноманіття операційних систем і пристроїв. Тобто йдеться про появу концепції автоматизації, яка має на меті полегшити й покращити процес тестування мобільних застосунків. Слід зауважити, що автоматизація процесів тестування або автоматизація тестування зі свого боку означає використання програмного забезпечення й інструментів для автоматизації виконання тестів. Цей метод дає змогу прискорити та підвищити ефективність процесів тестування. Наразі це невід'ємна частина процесів розробки програмного забезпечення та забезпечення якості. Це принципово дає змогу організаціям швидше створювати високоякісні продукти. Як правило, автоматизувавши повторювані й трудомісткі завдання тестування, команди можуть зосередитися на більш стратегічних аспектах тестування, таких як розробка й аналіз тестів. Тож вона є ключовою складовою успішної роботи будь-якого розробника програмного забезпечення. При цьому автоматизація процесів тестування мобільних застосунків стає загальнозживаною завдяки потенційним перевагам стосовно ефективності, економічності, масштабованості й гнучкості.

По-перше, це дає змогу знизити трудомісткість, скоротити час тестування та підвищити якість застосунків. По-друге, це дає змогу виконувати досить ефективно рутинні та повторювані завдання, забезпечувати прискорений зворотній зв'язок стосовно помилок і дефектів, а також забезпечувати неперервне тестування під час розробки застосунків. По-третє, це зменшить час на розробку продукту, підвищить якість і забезпечить надійність функціонування програмного забезпечення на всіх платформах. Однак, існують певні проблеми й обмеження, пов'язані з автоматизацією, які необхідно враховувати. У зв'язку із цим загострилися питання щодо розгляду такого аспекту розробки програмного забезпечення як тестування, оскільки саме цей процес забезпечує належну роботу застосунків на різних пристроях і платформах.

Аналіз досліджень. Різні аспекти процесів тестування мобільних застосунків вивчали Р. О. Багрій, В. О. Браун, К. П. Газдюк, В. М. Джулій, Л. І. Д'яченко, Т. О. Жирова, А. В. Ілін, Н. О. Котенко, М. Б. Кулеба, К. А. Кизима, С. С. Петровський, Д. О. Селюков, Л. В. Солодєєва, О. Ю. Тарновецька, В. О. Ткачук, А. В. Шестакова, Л. М. Шумиляк, N.G. Verihun, C. Dongmo, Van der J.A. Poll та інші. Р. О. Багрій, С. С. Петровський дослідили та визначили проблеми автоматизованого тестування вебзастосунків на сучасних платформах і запропонували методи й підходи для їх вирішення. У праці розглянуто автоматизацію тестів на основі піраміди Майкла Кона [1]. Вчені Г. М. Кодола, Н. С. Волинець, І. В. Сербулова [2] наголошують: «Тестування проводять на трьох рівнях: серверному, клієнтському та на рівні бази даних. На серверному рівні тестують стійкість сервера та програмного забезпечення до навантажень, на клієнтському – функціональність і сумісність застосунку, а на рівні бази даних – цілісність і стійкість до змін даних». У дослідженні Н. О. Котенко, Т. О. Жирова, М. Б. Кулеба [3] визначено особливості тестування мобільних застосунків, як орієнтир у складанні тест-плану для фахівця-початківця із забезпечення якості програмного забезпечення. Л. І. Д'яченко, А. В. Ілін, Л. М. Шумиляк, К. П. Газдюк, О. Ю. Тарновецька [4] розкрили методи автоматизованого тестування програмного забезпечення, а також дослідили прийоми й інструменти для створення автоматизованих тестів. В. О. Ткачук, А. В. Шестакова [5] розкривають технічні переваги використання мобільних застосунків з особливостями для ресторанного бізнесу. В. О. Браун, В. М. Джулій, К. А. Кизима, Д. О. Селюков, Л. В. Солодєєва [6] виявили особливості та методи тестування застосунків для мобільних пристроїв. N.G. Verihun, C. Dongmo, J.A. Van der Poll зазначили: «Тестування мобільної автоматизації – це один з етапів, який виконують тестувальники, щоб перевірити, чи задовольняють застосунки необхідну функціональність. Мобільні тестувальники наразі адаптують автоматизацію для досягнення якості та продуктивності. Однак ера мобільного автоматизованого тестування має свої виклики. Деякі з підводних каменів мобільного автоматизованого тестування містять незнання або визначення того, які тести можуть і повинні бути автоматизовані, автоматизацію невідповідних тестових кейсів розробниками, а також автоматизацію тестів на неправильному рівні і в неправильний час» [7]. Однак, сучасні дослідження у сфері автоматизації процесів тестування мобільних застосунків потребують комплексного підходу.

Метою дослідження є аналіз актуальної проблеми автоматизації процесів тестування мобільних застосунків. Слід виділити такі завдання дослідження: 1) визначити сучасні тенденції та виклики, пов'язані з автоматизацією тестування мобільних застосунків; 2) розглянути переваги та недоліки, основні методи, новітні технології й інструменти, рівні та підходи до автоматизації тестування мобільних застосунків. По-третє, також одним із завдань є дослідження таких суттєво важливих аспектів автоматизації як тестування функціональності, інтерфейсу, продуктивності та безпеки мобільних застосунків.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Автоматизація процесів тестування – це важливий аспект розробки програмного забезпечення, який допомагає забезпечити якість і надійність застосунків. Автоматизувавши процес тестування, організації можуть заощадити час і ресурси, одночасно покращуючи покриття і точність тестів.

Автоматизація ручних тестів: визначення вимог; розробка шаблонів; ручна розробка текстової функціональної специфікації; ручна розробка тестових сценаріїв у текстовому вигляді; автоматизація тестових сценаріїв; автоматизоване тестування. Метод автоматичної розробки тестів (РА-метод) вимагає текстового опису тестових сценаріїв. По суті створюються тестові сценарії в текстовому вигляді, як і у випадку РР-методу, а на основі даних тестових сценаріїв розробляються автоматизовані тести. Такий підхід вимагає більше часу на

розробку тестів, але значно скорочує час проведення тестування. Отже, такий підхід ефективний під час збільшення кількості циклів розробки [7].

На думку А. О. Мельник, Ю. Є. Кинаш: «Автоматизація надання клієнтських послуг містить автоматизацію процесів, пов'язаних із наданням підтримки клієнтам у використанні, виявленні, оптимізації й усуненні недоліків щодо послуги чи продукту [8]».

О. І. Міскевич, Д. Ю. Омельчук зазначають: «Автоматизоване тестування має свої переваги. Наприклад, автоматизовані тести можна виконувати швидко і велику кількість разів, що прискорює процес тестування. Ще автоматизоване тестування забезпечує послідовне виконання тестів, у цьому виді тестування неможливі помилки людського фактору та варіантів у тестуванні. Під час життєвого циклу розробки постійно вносяться нові зміни або оновлення. За допомогою автоматизованих тестів можна перевіряти новий код щодо відповідності наявним тестовим випадкам. Автоматизоване тестування може симулювати велику навантаженість для оцінки продуктивності під час стресового тестування. В автоматизованого тестування також є і недоліки. Створення автоматизованих тестів може бути витратним завданням, особливо на початковому етапі розробки. Автоматизовані тести потребують постійного обслуговування й оновлення під час змін у коді програми. Автоматизовані тести не можуть виявити помилки, непередбачені в сценаріях» [9].

На рисунку 1 наведено різноманітні підходи щодо автоматизації тестування мобільних застосунків.



Рис. 1 – Класифікація підходів до автоматизації тестування мобільних застосунків

Тобто можна бачити, що за інструментами автоматизації виокремлюють інструменти на основі SDK, зображень, гібридні інструменти. За типом тестування виділяють функціональне тестування, тестування не функціональних вимог, універсальне тестування. Наразі автоматизація процесів тестування мобільних застосунків передбачає написання скриптів або використання спеціалізованих інструментів для автоматизації різних видів діяльності з тестування, таких як виконання тестових кейсів, генерація тестових даних й аналіз результатів. Автоматизувавши ці завдання, тестувальники можуть заощадити час і зусилля, а також отримати більш точні та послідовні результати. За рівнем автоматизації буває ручне частково автоматизоване та повне автоматизування. За середовищем виконання підходів до автоматизації тестування мобільних застосунків бувають локальне віддалене хмарне емулятори та симулятори. При цьому вибір середовища виконання для автоматизації тестування мобільних застосунків може суттєво вплинути на ефективність і результати тестування. Існує декілька способів створення тестів автоматизації тестування мобільних застосунків. Зокрема, запис і відтворення на основі ключових слів, моделі об'єктів, поведінки [10].

Крім того, автоматизація тестування сприяє повторному використанню тестових скриптів, дозволяючи командам запускати тести багаторазово без необхідності ручного втручання. Таке багаторазове використання не лише економить час, але й забезпечує послідовне виконання тестів, що призводить до більш надійних результатів тестування.

У контексті мобільних застосунків автоматизація процесів тестування відіграє вирішальну роль у забезпеченні якості та надійності програмного забезпечення. Із підвищенням складності та різноманітності мобільних пристроїв й операційних систем, ручного тестування недостатньо, щоб охопити всі можливі сценарії. Автоматизація тестування дає змогу здійснювати ретельне тестування на різних пристроях і платформах, гарантуючи, що застосунок працює належним чином для всіх користувачів. Щоб ефективно впровадити автоматизацію процесів тестування, тестувальники повинні використовувати правильні інструменти та дотримуватися встановлених процесів.

Крім того, автоматизація тестування мобільних застосунків дає змогу скоротити час виходу на ринок, забезпечуючи швидкі та безперервні цикли тестування. Така гнучкість має важливе значення у швидкозмінній індустрії розробки мобільних застосунків, де швидкі випуски й оновлення є ключовим фактором для збереження конкурентоспроможності на ринку.

Розглянемо детальніше переваги автоматизації тестування мобільних застосунків (рис. 2).

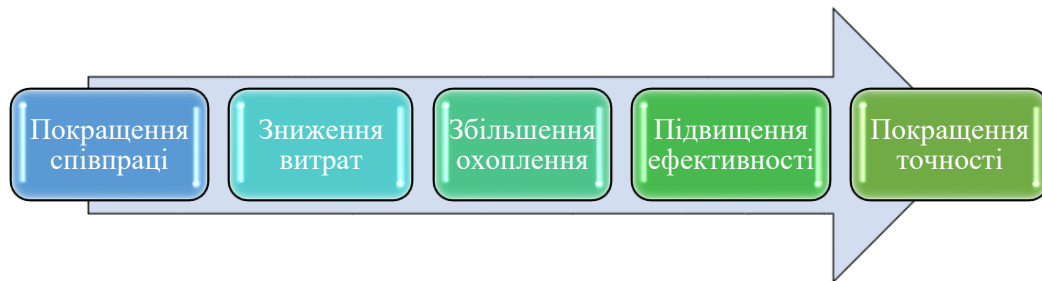


Рис. 2. – Переваги автоматизації тестування мобільних застосунків

Як правило, автоматизація процесів тестування мобільних застосунків забезпечує підвищену ефективність і точність. По-перше, автоматизація може виконувати повторювані тести набагато швидше, ніж тестування вручну. Це значно підвищує загальну ефективність процесу тестування. Крім того, автоматизовані тести виконуються точніше, що зі свого боку зменшує ймовірність людських помилок. При цьому автоматизація дає змогу проводити тести одночасно на кількох пристроях і платформах, заощаджуючи час і зусилля. Виявляючи на ранніх стадіях дефекти процесу розробки, автоматизація дає змогу оперативно розв'язувати проблеми, що сприяє створенню мобільних застосунків вищої якості.

Причинами автоматизації тестування мобільних застосунків є:

1. Зростання обсягів тестового навантаження. Коли розробляється більше застосунків для різноманітних платформ і пристроїв, кількість випадків тестування показово зростає. Це призводить до значного збільшення часу та коштів на тестування, якщо використовувати традиційні методи. Автоматизація дає змогу оптимізувати процес тестування, зменшивши час і ресурси, потрібні для його виконання.

2. Збільшення складності застосунків. Сучасні застосунки характеризуються складнішою структурою та більшою кількістю функцій, що мають бути протестованими. Це може призвести до появи значної кількості випадків тестування і складніших сценаріїв, які важко оцінити вручну. Автоматизація дає змогу обійти цю проблему, забезпечуючи тестування всіх аспектів застосунку на кожній стадії розробки.

3. Потреба в ресурсах. Для тестування традиційними методами необхідно мати достатньо ресурсів, таких як персонал, обладнання і час. Це може бути проблемою для малих і середніх розробників програмного забезпечення, які можуть не мати таких ресурсів. Автоматизація дає змогу використовувати наявні ресурси більш ефективно та зменшує вартість тестування.

Тестувальники переважно можуть зосередитися на більш важливих аспектах тестування, автоматизуючи повторювані й трудомісткі тести. Автоматизація тестування також зменшує ймовірність людської помилки, що призводить до більш точних і надійних результатів тестування.

Ручне тестування може бути трудомістким і дорогим процесом, особливо коли мова йде про тестування мобільних застосунків на різних пристроях і платформах. Натомість слід зазначити, що автоматизація тестування дає змогу тестувальникам виконувати тести швидко й ефективно, заощаджуючи час і витрати, пов'язані з ручним тестуванням.

До того ж автоматизація процесів тестування дає змогу створювати комплексні набори тестів, які можна повторно використовувати для регресійного тестування. Це зі свого боку означає, що одного разу створений набір тестів можна запускати багато разів без додаткових зусиль, тим самим забезпечивши повне тестове покриття різних сценаріїв і функціональних можливостей.

Одночасно завдяки автоматизації тестування тестувальники можуть легко охопити широкий спектр тестових кейсів, які можна виконувати вручну. Таке розширене покриття тестів допомагає виявляти дефекти на ранніх стадіях циклу розробки, що призводить до підвищення якості програмного забезпечення та покращення користувацького досвіду [11].

Зверніть увагу, що автоматизація може значно скоротити час і ресурси, необхідні для тестування мобільних застосунків. Автоматизовані тести можна виконувати швидше й частіше, ніж ручні тести, що дасть змогу швидко виявляти помилки. Вони менш схильні до людської помилки, що може призвести до більш точного тестування. Їх можна використовувати для тестування ширшого спектра сценаріїв, ніж ручні тести. Також їх можна легко поширювати й інтегрувати із системами безперервної інтеграції та безперервної доставки (CI/CD), що покращить співпрацю між командами розробників і тестування.

З іншого боку, автоматизація процесів тестування мобільних застосунків пов'язана з такими проблемами, як початкове налаштування та навчання. Упровадження інструментів автоматизації потребує часу та ресурсів для налаштування та навчання, що може стати перешкодою для команд, які переходять від ручного до автоматизованого тестування. Крім того, автоматизація може вимагати спеціальних навичок, якими можуть не володіти наявні ручні тестувальники.

Саме тому автоматизація процесів тестування мобільних застосунків виявилася рентабельною. Зменшуючи потребу в ручних тестерах, автоматизація дає змогу заощадити на оплаті праці. Можливість проводити регресійне тестування без додаткових витрат є ще однією економічною перевагою автоматизації. Крім того, автоматизація сприяє швидкому виходу мобільних застосунків на ринок, що призводить до потенційної економії коштів і конкурентних переваг на ринку.

Автоматизація процесів тестування мобільних застосунків може мати обмеження під час тестування складних сценаріїв попри економічну ефективність. Автоматизовані тести можуть мати проблеми зі сценаріями, які вимагають людської інтуїції або творчих здібностей, наприклад дослідницьке тестування. Крім того, певні візуальні проблеми чи проблеми з користуванням неможливо ефективно виявити за допомогою автоматизованих тестів. Нездатність автоматизації адаптуватися до вимог спеціального тестування так само ефективно, як ручні тестери, створює проблему для забезпечення повного охоплення тестуванням.

Масштабованість і гнучкість, які пропонує автоматизація процесів тестування мобільних застосунків, є безумовно значними перевагами. Автоматизація може легко масштабуватися відповідно до тестувальних вимог для різних типів і конфігурацій пристроїв, забезпечуючи повне охоплення тестами. Багаторазове використання тестових сценаріїв у різних тестувальних сценаріях підвищує ефективність і зменшує дублювання зусиль. Крім того, інтеграція з конвеєрами CI/CD забезпечує безперебійне тестування протягом життєвого циклу розробки, сприяючи швидкій доставці високоякісних мобільних застосунків [12].

Однак, технічне обслуговування й оновлення є ключовими факторами автоматизації процесів тестування мобільних застосунків. Автоматизовані сценарії тестування потребують регулярного обслуговування, щоб забезпечити їх релевантність і точність виявлення дефектів. Зміни в програмі можуть вимагати оновлення сценаріїв автоматизованого тестування, що потребує додаткових зусиль і ресурсів. Крім того, оновлення інструментів або платформ тестування можуть порушити структуру автоматизації, вимагаючи коригування та потенційно вплинути на часові рамки тестування. Розглянемо детальніше недоліки автоматизації тестування мобільних застосунків (рис. 3).

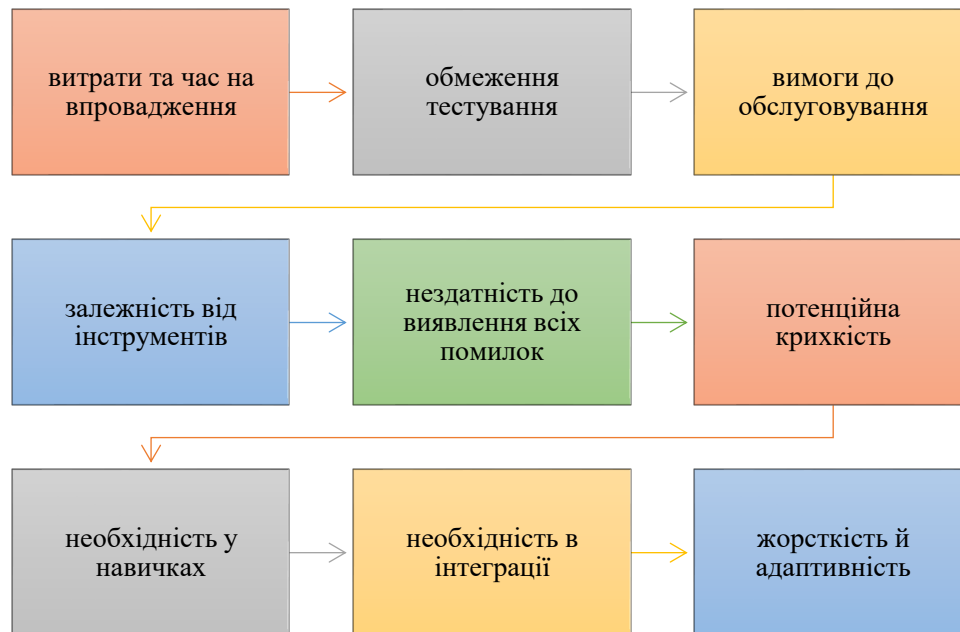


Рис. 3 – Недоліки автоматизації тестування мобільних застосунків

Незважаючи на те, що автоматизація процесів тестування пропонує низку переваг, існують певні виклики, з якими організації можуть зіткнутися під час упровадження. Автоматизація тестування вимагає технічних знань у написанні сценаріїв і використанні інструментів автоматизації. Організаціям необхідно буде інвестувати в навчання своїх тестувальників або найняти спеціалізовані ресурси для подолання цих викликів. Крім того, проблеми сумісності з різними операційними системами, браузерами й пристроями можуть створювати технічні труднощі під час автоматизації тестування. Забезпечення безперебійної роботи автоматизованих тестів на різних платформах вимагає ретельного тестування та налагодження, що ускладнює процес упровадження.

Інтеграція автоматизації тестування в наявні та робочі процеси може бути складним завданням. Для забезпечення безперебійного виконання автоматизованих тестів потрібна координація між різними командами, наприклад, розробниками й тестувальниками. Крім того, підтримка тестових скриптів та їхнє оновлення у зв'язку із частими змінами в програмному забезпеченні може бути також складним завданням. Будь-які зміни в застосунку, що тестується, можуть вимагати відповідних оновлень сценаріїв автоматизованого тестування, що збільшує зусилля з обслуговування і потенційно може спричинити затримки у виконанні тестів. Розробка й налаштування автоматизованих тестових сценаріїв може потребувати значно більше часу та ресурсів, ніж ручне тестування. Це може бути особливо актуально для складних мобільних застосунків або для тестів, які потребують частої адаптації до змін в UI або функціональності. Автоматизовані тести не завжди можуть повністю відтворити дії та поведінку людського користувача. Це може призвести до того, що деякі помилки або дефекти залишаться невиявленими. Крім того, автоматизовані тести можуть бути не такими ефективними під час тестування творчого контенту або складних взаємодій з користувачем.

Розробка й обслуговування автоматизованих тестових сценаріїв зазвичай потребує певних технічних навичок, знань програмування та розуміння платформ автоматизації. Це може бути бар'єром для команд тестувальників, які не володіють необхідним досвідом. Автоматизовані тести можуть бути жорсткими й складними для адаптації до змін у мобільному застосунку. Це може призвести до того, що тести стануть застарілими та потребуватимуть частого оновлення, що може бути витратним за часом і ресурсами. Автоматизовані тести не завжди можуть виявити всі типи помилок, особливо ті, що пов'язані з UI, продуктивністю або складними взаємодіями з користувачем. Для виявлення таких помилок все ще може бути необхідне ручне тестування. Автоматизація тестування залежить від надійності та доступності використовуваних інструментів. Технічні проблеми або збої в роботі інструментів можуть призвести до перерв у тестуванні та затримок. Автоматизовані тести можуть потребувати інтеграції з іншими системами, такими як

інструменти CI/CD або системи управління тестами. Це може потребувати додаткових зусиль та налаштування [13].

Тепер розглянемо етапи та рівні автоматизації процесів тестування мобільних застосунків (Табл. 1 та рис. 4).

Таблиця 1. Етапи автоматизації процесів тестування мобільних застосунків

Етап	Опис
Планування й аналіз	Визначте цілі, обсяг, інструменти й план тестування
Розробка тестів	Створіть тестові сценарії, дані та скрипти.
Виконання тестів	Налаштуйте середовище та запустіть скрипти
Аналіз результатів	Перевірте результати, знайдіть помилки й оновіть тести
Звітність	Створіть звіти та поділіться ними із зацікавленими сторонами

Для автоматизації процесів тестування мобільних застосунків необхідно врахувати декілька етапів:

1. Визначення потреби в автоматизації тестування. Дослідіть й оцініть процеси тестування в вашій організації, визначте, які саме завдання можуть бути автоматизовані та які інструменти краще використовувати для цього.

2. Вибір інструментів для автоматизації. Розгляньте різні інструменти для автоматизації тестування мобільних застосунків, такі як Appium, Selenium, TestComplete тощо, й оберіть той, який найкраще відповідає вашим потребам.

3. Розробка автоматизованих тестів. Створіть скрипти для автоматизованого тестування мобільних застосунків з урахуванням різних сценаріїв використання та вимог до застосунків.

4. Виконання тестових сценаріїв. Запустіть автоматизовані тести на різних пристроях і платформах, щоб переконатися в їх працездатності та здатності виявляти помилки.

5. Аналіз результатів. Оцініть результати автоматизованих тестів, виявлені помилки та пропонуйте відповідні виправлення розробникам застосунку.

6. Підтримка автоматизованих тестів. Регулярно оновлюйте та підтримуйте автоматизовані тести, додавайте нові сценарії та вдосконалюйте їх для покращення ефективності тестування.

Дотримуючись цих етапів і вимог, ви зможете успішно автоматизувати процеси тестування мобільних застосунків у вашій організації.

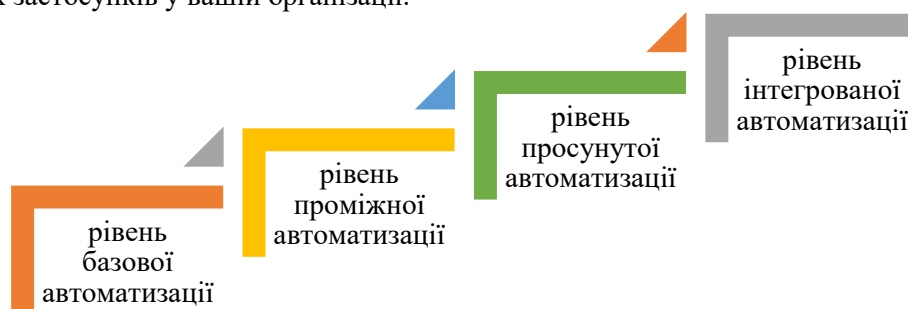


Рис. 4 – Рівні автоматизації тестування мобільних застосунків

Зверніть увагу на те, що автоматизація тестування мобільних застосунків може бути організована за кількома рівнями, кожен з яких описує певний ступінь охоплення і складності тестових сценаріїв.

На рівні базової автоматизації іде фокусування на автоматизації простих тестових сценаріїв, таких як перевірка входу в систему, навігація за основними функціями й тестування базової функціональності. Зазвичай використовуються прості інструменти, такі як Appium або XCTest, з базовими можливостями запису та відтворення. До переваг належить швидке впровадження,

економія часу на повторюваних тестах, до недоліків – обмежене охоплення складних сценаріїв, низька гнучкість для тестування нестандартної логіки.

На рівні проміжної автоматизації відбувається розширення базової автоматизацію, включаючи тестування API, тестування продуктивності та деяких аспектів тестування UI. Використовуються більш просунуті інструменти, такі як Robot Framework або Calabash, які пропонують кращу структуру та можливості тестування. До переваг належить збільшене охоплення функціональності, можливість тестування складніших сценаріїв, до недоліків – більші навички розробки й тестування, може потребувати більше часу на налаштування.

На рівні просунутої автоматизації відбувається автоматизація всебічного тестування мобільного додатку, включаючи інтеграційне тестування, тестування навантаження та тестування безпеки. Використовуються комплексні фреймворки, такі як Appium Studio або EarlGrey, які пропонують розширені можливості тестування та інтеграції. Переваги – максимальне охоплення тестування, покращення якості та надійності мобільного застосунку, недоліками – складність упровадження й обслуговування, потребує значних ресурсів і досвіду [14].

На рівні інтегрованої автоматизації відбувається об'єднання автоматизації тестування мобільних застосунків з іншими процесами DevOps, такими як безперервна інтеграція та безперервна доставка (CI/CD). Використовуються платформи CI/CD, такі як Jenkins або GitLab CI, разом з інструментами автоматизації тестування. Перевагами є вбудовування в процес розробки, забезпечуючи швидке виявлення та виправлення помилок. Недоліками є потреба значної інтеграції та налаштування, може бути складним для команд, які не мають досвіду роботи з DevOps.

Слід зазначити, що вибір рівня автоматизації тестування мобільних застосунків залежить від декількох факторів. Як правило, складніші додатки потребують вищого рівня автоматизації. Просунуті рівні автоматизації потребують більших інвестицій у час і кошти.

Водночас команда повинна володіти необхідними навичками для роботи з обраними інструментами та платформами. Необхідно чітко визначити, які аспекти мобільного застосунку потребують автоматизованого тестування. Важливо ретельно оцінити ці фактори, щоб вибрати оптимальний рівень автоматизації, який збалансує охоплення тестування, економію ресурсів і досяжність для команди. Упровадження автоматизації процесів тестування вимагає системного підходу.

Процес включає планування тестів, розробку тестових кейсів, розробку сценаріїв, виконання тестів й аналіз результатів. Тестувальники повинні розуміти додаток, що тестується, і визначити критичні сценарії, які необхідно автоматизувати. Визначаючи пріоритетність тестових кейсів на основі ризику та впливу, тестери можуть оптимізувати зусилля з автоматизації та зосередитися на областях, які є найбільш важливими для функціональності програми. Крім того, установлення чітких каналів комунікації між розробниками, тестувальниками й іншими зацікавленими сторонами має важливе значення для успішного автоматизованого тестування. Співпраця та зворотній зв'язок дають змогу виявляти проблеми на ранніх стадіях циклу розробки, що призводить до швидшого їх вирішення та покращення якості програмного забезпечення.

Зосередимо увагу на способах автоматизації тестування мобільних застосунків (рис. 5).



Рис. 5 – Способи автоматизації тестування мобільних застосунків

Один з найпопулярніших інструментів з відкритим кодом для автоматизації тестування мобільних застосунків є Appium, який підтримує iOS, Android і Windows Phone. Він використовує WebDriver інтерфейс для взаємодії з мобільними елементами та дає змогу писати тести на різних мовах програмування, таких як Java, Python, Ruby та C#. Ще один популярний інструмент з відкритим кодом – це Selenium, який використовується переважно для тестування вебзастосунків, але також може бути адаптований для тестування мобільних вебзастосунків. Він пропонує широкий спектр команд для взаємодії з вебелементами та підтримує різні мови програмування.

TestComplete – це комплексний інструмент для автоматизації тестування, який пропонує широкий спектр функцій для тестування мобільних застосунків, включаючи функціональне тестування, тестування UI, тестування продуктивності та тестування навантаження. Він підтримує різні мобільні платформи та пропонує інтеграцію з іншими інструментами DevOps.

UFT Mobile – інструмент для автоматизації тестування мобільних застосунків від HP, який пропонує схожі функції, як і TestComplete. UFT Mobile підтримує iOS, Android і Windows Phone та пропонує інтеграцію з іншими продуктами HP, такими як ALM і QC [13].

Robot Framework – це платформа для тестування з відкритим кодом, яка використовує простий ключовий підхід до написання тестів. Вона добре підходить для автоматизації простих і складних тестових сценаріїв, а також може бути інтегрована з іншими інструментами тестування.

Cucumber – це ще одна популярна платформа для тестування з відкритим кодом, яка використовує Gherkin, мову, подібну до англійської, для опису тестових сценаріїв. Вона підходить для написання читабельних і зрозумілих тестів, а також може бути інтегрована з різними інструментами автоматизації.

Appium Studio – це хмарна платформа тестування мобільних застосунків, яка пропонує просту графічну оболонку для написання та виконання тестів Appium. Appium Studio також пропонує функції звітності й аналітики, що дають змогу відстежувати хід тестування.

А також Experitest – ще одна хмарна платформа тестування мобільних застосунків, яка пропонує широкий спектр функцій, включаючи підтримку iOS, Android і Windows Phone, тестування на реальних пристроях й інтеграцію із CI/CD інструментами.

Тож на ринку доступні різні інструменти для автоматизації тестування мобільних застосунків. Ці інструменти надають такі функції, як запис і відтворення скриптів, управління тестовими кейсами й аналіз результатів. Деякі популярні інструменти містять Appium, Selenium та TestComplete. Ці інструменти пропонують гнучкість і масштабованість, що дає змогу тестувальникам створювати надійні автоматизовані сценарії тестування для різних типів застосунків.

Крім того, деякі інструменти пропонують інтеграцію з конвеєрами безперервної інтеграції/безперервного розгортання (CI/CD), що забезпечує безперебійну автоматизацію протягом життєвого циклу розробки програмного забезпечення. Ця інтеграція гарантує, що автоматизовані тести запускаються автоматично під час унесення нових змін до коду, забезпечуючи негайний зворотній зв'язок з розробниками [15].

Основними технічними рішеннями автоматизації тестування мобільних застосунків є застосування фреймворків тестування та скриптоване тестування. Існують різні фреймворки тестування мобільних застосунків, такі як: Appium, Espresso, XCUITest тощо. Вони забезпечують інтегровану платформу для автоматизації процесів тестування на всіх основних платформах і пристроях. Основним інструментом автоматизації тестування є скриптоване тестування, яке дає змогу використовувати різні мови програмування, такі як: Java, Python, JavaScript тощо, для написання тестових сценаріїв.

Оскільки технології постійно розвиваються, автоматизація процесів тестування також розвивається.

Нові технології, такі як роботизована автоматизація процесів (RPA) і штучний інтелект (AI), інтегруються в автоматизацію тестування. Ці технології можуть розширити можливості автоматизованого тестування, дозволяючи створювати більш складні тестові сценарії та покращувати покриття тестів.

Штучний інтелект і машинне навчання революціонізують сферу автоматизації тестування. Ці технології можуть аналізувати величезні обсяги тестових даних, виявляти закономірності й

аномалії, а також надавати ідеї для покращення тестів. Вони також дають змогу створювати інтелектуальні тестові кейси та самовідновлювані тестові сценарії.

Незважаючи на те, що штучний інтелект і машинне навчання пропонують величезний потенціал для покращення автоматизації тестування, існують також проблеми, які необхідно подолати. Однією з таких проблем є потреба у кваліфікованих фахівцях, які можуть розробляти та підтримувати платформи для тестування на основі ШІ. Організації повинні інвестувати в навчання своєї робочої сили, щоб ефективно використовувати ці технології.

Крім того, інтеграція ШІ в наявні процеси тестування вимагає ретельного планування і виконання, щоб забезпечити плавний перехід. Важливо визначити правильні варіанти використання ШІ в автоматизації тестування і поступово впроваджувати рішення для ШІ, щоб звести до мінімуму збої в поточній діяльності з тестування.

Висновки та перспективи подальшого дослідження. Отже, автоматизація процесів тестування мобільних застосунків життєво важливою для забезпечення якості та надійності мобільних застосунків. Розуміючи її концепцію, використовуючи правильні інструменти та долаючи труднощі впровадження, організації можуть отримати вигоду від підвищення ефективності, точності й економії коштів. Автоматизація процесів тестування мобільних застосунків може допомогти покращити ефективність, надійність і якість мобільних застосунків. Існує багато інструментів для автоматизації тестування мобільних застосунків, доступних на ринку, тому важливо вибрати інструмент, який відповідає потребам. Крім того, відстеження майбутніх тенденцій, таких як нові технології та штучний інтелект, допоможе організаціям залишатися попереду у сфері автоматизації тестування, що постійно розвивається.

Отже, автоматизація процесів тестування мобільних застосунків пропонує переконливі переваги з точки зору ефективності, економічності, масштабності та гнучкості. Хоча існують проблеми й обмеження, пов'язані з автоматизацією, як-от початкове налаштування та крива навчання, тестування складних сценаріїв і вимоги до обслуговування, переваг значно більше, ніж недоліків. Ретельно оцінюючи аргументи «за» і «проти» автоматизації, організації можуть приймати обґрунтовані рішення щодо впровадження процесів автоматизованого тестування для підвищення якості та надійності своїх мобільних застосунків. Перспективою подальших досліджень є розробка автоматизованих тестових сценаріїв, які динамічно адаптуються до змін в UI або функціональності мобільного застосунку.

Список бібліографічного опису

1. Багрій Р.О., Петровський С.С. Особливості сучасного тестування веб-додатків. *Вісник Хмельницького національного університету. Технічні науки*. 2022. № 3. С. 70-74. DOI 10.31891/2307-5732-2022-309-3-70-74 URL: <https://elar.khmnu.edu.ua/handle/123456789/12457> (дата звернення: 07.06.2024).
2. Кодола Г.М., Волинець Н.С., Сербулова І.В. Автоматизоване тестування веб-додатків з різномірною архітектурою. *Вісник НТУ "ХПИ": Серія «Нові рішення в сучасних технологіях»*. 2019. № 5 (1330). С. 91-100. DOI 10.20998/2413-4295.2019.05.12 URL: <https://core.ac.uk/download/211066507.pdf> (дата звернення: 07.06.2024).
3. Котенко Н.О., Жирова Т.О., Кулеба М.Б. Дослідження особливостей тестування мобільних додатків. *Управління розвитком складних систем*. 2020. № 41. С. 55-60. DOI <https://doi.org/10.32347/2412-9933.2020.41.55-60> URL: <http://mdcs.knuba.edu.ua/article/view/203613> (дата звернення: 07.06.2024).
4. Д'яченко Л.І., Ілін А.В., Шумиляк Л.М., Газдюк К.П., Тарновецька О.Ю. Огляд програмних засобів та методологій для реалізації систем автоматизованого тестування. *Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки*. 2021. № 2. Том 32 (71). Ч. 1. С. 122-129. DOI <https://doi.org/10.32838/2663-5941/2021.2-1/20> URL: https://tech.vernadskyjournals.in.ua/journals/2021/2_2021/part_1/22.pdf (дата звернення: 07.06.2024).
5. Ткачук В.О., Шестакова А.В. Використання мобільних додатків для цифровізації бізнес-процесів у ресторанному бізнесі та їх оптимізація на основі методу А/В тестування. 2022. № 4 (102). С. 27-34. DOI [https://doi.org/10.26642/jen-2022-4\(102\)-27-34](https://doi.org/10.26642/jen-2022-4(102)-27-34) URL: <https://eztuir.ztu.edu.ua/jspui/bitstream/123456789/8097/1/27.pdf> (дата звернення: 07.06.2024).
6. Браун В.О., Джулії В.М., Кизима К.А., Селюков Д.О., Солодєва Л.В. Інтегральний критерій оцінки методів тестування додатків для мобільних пристроїв. *Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка*. Київ : ВІКНУ, 2018. Вип. 61. С. 40-49. URL: <https://miljournals.knu.ua/index.php/zbirnik/article/download/332/326/> (дата звернення: 07.06.2024).
7. Berihun NG, Dongmo C, Van der Poll JA. The Applicability of Automated Testing Frameworks for Mobile Application Testing: A Systematic Literature Review. *Computers*. 2023; 12(5): 97. DOI <https://doi.org/10.3390/computers12050097> URL: https://www.mdpi.com/2073-431X/12/5/97?type=check_update&version=2 (date of access: 07.06.2024).
8. Мельник А.О., Кинаш Ю.Є. Розробка підсистеми автоматизованого управління відносинами з клієнтами малого бізнесу. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2024. Вип. № 54. Луцьк, С.

- (54). C. 27-35. DOI <https://doi.org/10.36910/6775-2524-0560-2024-54-03> URL: <http://www.cit-journal.com.ua/index.php/cit/article/view/521> (дата звернення: 07.06.2024).
9. Міскевич О.І., Омельчук Д.Ю. Ручне та автоматизоване тестування блокчейну за допомогою Postman. *Комп'ютерно-інтегровані технології: освіта, наука, виробництво*. 2023. № 53. С. 23-30. DOI <https://doi.org/10.36910/6775-2524-0560-2023-53-04> URL: <http://www.cit-journal.com.ua/index.php/cit/article/view/476> (дата звернення: 07.06.2024).
10. Bannikov, Valentyn. (2022). Foresight as a Strategic Management Tool in the ICT Industry. *Business Inform.* DOI 10.225-230. 10.32983/2222-4459-2022-10-225-230. (date of access: 07.06.2024).
11. Hunkoi I. (2023). Software testing in 2023: new trends and challenges. *Herald of Kyiv Institute of Business and Technology*, 49(1-2), 25-36. DOI <https://doi.org/10.37203/kibit.2023.49.03> (date of access: 07.06.2024).
12. Tramontana, P.; Amalfitano, D.; Amatucci, N. Automated functional testing of mobile applications: A systematic mapping study. *Softw. Qual. J.* 2019, 149–201. URL: <https://link.springer.com/article/10.1007/s11219-018-9418-6> (date of access: 07.06.2024).
13. Kong, P.; Li, L.; Gao, J.; Liu, K.; Bissyande, T.F.; Klein, J. Automated Testing of Android Apps: A Systematic Literature Review. *IEEE Trans. Reliab.* 2019, 68, 45–66. URL: <https://ieeexplore.ieee.org/document/8453877> (date of access: 07.06.2024).
14. Tirodkar, A.A.; Khandpur, S.S. EarlGrey: iOS UI Automation Testing Framework. In *Proceedings of the 2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*, Montreal, QC, Canada, 25–25 May 2019; pp. 12–15. URL: <https://ieeexplore.ieee.org/document/8816871> (date of access: 07.06.2024).
15. Salam, M.A., Taha, S., Hamed, M.G. Advanced Framework for Automated Testing of Mobile Applications. In *Proceedings of the 2022 4th Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, Giza, Egypt, 22–24 October 2022; pp. 233–238. URL: <https://ieeexplore.ieee.org/document/9942374> (date of access: 07.06.2024).

References

1. Bahrii R.O., Petrovskiy S.S. (2022). Osoblyvosti suchasnoho testuvannia veb-dodatkov. *Visnyk Khmelnytskoho natsionalnoho universytetu [Features of modern web application testing. Bulletin of Khmelnytsky National University]. Tekhnichni nauky - Technical sciences.* 3, 70-74. DOI 10.31891/2307-5732-2022-309-3-70-74 Retrieved from <https://elar.khmnu.edu.ua/handle/123456789/12457> [in Ukrainian].
2. Kodola H.M., Volynets N.S., Serbulova I.V. (2019). Avtomatyzovane testuvannia veb-dodatkov z riznorivnevoiu arkhitekturoiu [Automated testing of web applications with multilevel architecture]. *Visnyk NTU "KhPI": Seriya «Novi rishennia v suchasnykh tekhnolohiiakh» - Bulletin of NTU "KhPI": Series "New solutions in modern technologies".* 5 (1330), 91-100. DOI 10.20998/2413-4295.2019.05.12 Retrieved from <https://core.ac.uk/download/211066507.pdf> [in Ukrainian].
3. Kotenko N.O., Zhyrova T.O., Kuleba M.B. (2020). Doslidzhennia osoblyvostei testuvannia mobilnykh dodatkov [Research of mobile application testing features]. *Upravlinnia rozvytkom skladnykh system - Managing the development of complex systems.* DOI: <https://doi.org/10.32347/2412-9933.2020.41.55-60> Retrieved from <http://mdcs.knuba.edu.ua/article/view/203613> [in Ukrainian].
4. Diachenko L.I., Ilin A.V., Shumylyak L.M., Hazdiuk K.P., Tarnovetska O.Iu. (2021). Ohliad prohramnykh zasobiv ta metodolohii dlia realizatsiii system avtomatyzovanoho testuvannia [Review of software tools and methodologies for the implementation of automated testing systems]. *Vcheni zapysky TNU imeni V.I. Vernadskoho. Seriya: Tekhnichni nauky - Scientific notes of Vernadsky National University. Series: Technical sciences.* 2, 32 (71), 1, 122-129. DOI <https://doi.org/10.32838/2663-5941/2021.2-1/20> Retrieved from https://tech.vernadskyjournals.in.ua/journals/2021/2_2021/part_1/22.pdf [in Ukrainian].
5. Tkachuk, V. O., & Shestakova, A. V. (2022). Using mobile applications for digitalization of business processes in the restaurant business and their optimization based on the A/B testing method. *Journal of Electronic Scientific and Information Technologies*, 4(102), 27-34. [https://doi.org/10.26642/jen-2022-4\(102\)-27-34](https://doi.org/10.26642/jen-2022-4(102)-27-34)
6. Braun, V. O., Dzhuili, V. M., Kyzyma, K. A., Selyukov, D. O., & Solodiieva, L. V. (2018). Integral criterion for evaluating testing methods for mobile applications. *Collection of Scientific Works of the Military Institute of the Taras Shevchenko National University of Kyiv*, 61, 40-49. <https://miljournals.knu.ua/index.php/zbirnik/article/download/332/326/>
7. Melnyk, A. O., & Kinash, Yu. Ye. (2024). Development of a subsystem for automated customer relationship management for small businesses. *Computer-Integrated Technologies: Education, Science, Production*, 54, 27-35. <https://doi.org/10.36910/6775-2524-0560-2024-54-03>
8. Miskevich, O. I., & Omelchuk, D. Yu. (2023). Manual and automated testing of the blockchain using Postman. *Computer-Integrated Technologies: Education, Science, Production*, 53, 23-30. <https://doi.org/10.36910/6775-2524-0560-2023-53-04>
9. Berihun, N. G., Dongmo, C., & Van der Poll, J. A. (2023). The Applicability of Automated Testing Frameworks for Mobile Application Testing: A Systematic Literature Review. *Computers*, 12(5), 97. <https://doi.org/10.3390/computers12050097>
10. Bannikov, V. (2022). Foresight as a Strategic Management Tool in the ICT Industry. *Business Inform*, 10(1), 225-230. DOI 10.225-230.
11. Hunkoi, I. (2023). Software testing in 2023: new trends and challenges. *Herald of Kyiv Institute of Business and Technology*, 49(1-2), 25-36. <https://doi.org/10.37203/kibit.2023.49.03>
12. Tramontana, P., Amalfitano, D., & Amatucci, N. (2019). Automated functional testing of mobile applications: A systematic mapping study. *Software Quality Journal*, 149–201. <https://link.springer.com/article/10.1007/s11219-018-9418-6>

13. Kong, P., Li, L., Gao, J., Liu, K., Bissyande, T.F., & Klein, J. (2019). Automated Testing of Android Apps: A Systematic Literature Review. *IEEE Transactions on Reliability*, 68, 45–66. <https://ieeexplore.ieee.org/document/8453877>
14. Tirodkar, A.A., & Khandpur, S.S. (2019). EarlGrey: iOS UI Automation Testing Framework. In Proceedings of the 2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft), Montreal, QC, Canada, 25–25 May.
15. Salam, M.A., Taha, S., Hamed, M.G. (2022). Advanced Framework for Automated Testing of Mobile Applications. In Proceedings of the 2022 4th Novel Intelligent and Leading Emerging Sciences Conference (NILES), Giza, Egypt, 22–24 October; 233–238. URL: <https://ieeexplore.ieee.org/document/9942374>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-28>

УДК 004.75

Христинець Наталія Анатоліївна, к.т.н., доцент

<https://orcid.org/0000-0002-4836-7632>

Вікторова Яна Андріївна, студентка

Жуковська Софія Юрїївна, студентка

Радіон Юлія Андріївна, студентка

Луцький національний технічний університет, м. Луцьк, Україна

РОЗРОБКА ФАЙЛОВОГО МЕНЕДЖЕРА ОПЕРАЦІЙНОЇ СИСТЕМИ

Христинець Н.А., Вікторова Я.А., Жуковська С.Ю., Радіон Ю.А. Розробка файлового менеджера ОС. В цій роботі описується розробка та реалізація простого файлового менеджера для операційної системи. Файловий менеджер - це програма, яка надає користувачеві інтерфейс для взаємодії з файловою системою комп'ютера. Він дозволяє створювати, видаляти, копіювати, переміщати та шукати файли та папки, а також переглядати їх властивості. Файлові менеджери є невід'ємною частиною будь-якої операційної системи. Функціонал цього файлового менеджера включає можливість працювати з різними типами файлів та забезпечує зручний спосіб їх упорядкування. Інтерфейс програми був розроблений таким чином, щоб бути інтуїтивно зрозумілим для користувачів з різним рівнем технічної підготовки. Важливою особливістю є підтримка операцій з великим обсягом даних, що підвищує ефективність роботи користувачів. У процесі розробки було враховано вимоги безпеки для захисту даних від несанкціонованого доступу. Крім того, файловий менеджер підтримує інтеграцію з іншими додатками, що розширює його функціональні можливості.

Ключові слова: файловий менеджер, файлова система, операції над файлами, Python, командний рядок.

Khrystynets N., Viktorova Y., Zhukovska S., Radion Yu. OS file manager development. This paper describes the development and implementation of a simple file manager for the operating system. A file manager is a program that provides the user with an interface to interact with the computer's file system. It allows you to create, delete, copy, move, and search for files and folders, as well as view their properties. File managers are an essential part of any operating system. The function of the file manager is to enable the reader to follow the file's instructions and thus avoid a serious injury. Interpreters are proving to be a great help to Chinesees, but they are incredibly dangerous for the brains of young people. The important thing is to bring about a reduction in the number of animals, which in turn leads to the effective breeding of robots. In the process, the pink ball turned out to be a nightmare for the security forces of the unsanctioned government. Indeed, Filovy Medvedev is attempting to integrate into other dimensions, he is expanding yoga's functional capabilities.

Keywords: file manager, file system, file operations, Python, command line.

Постановка наукової проблеми. У наш час практично неможливо уявити будинок людини без комп'ютера. Ви не знайдете компанії, офісу, фабрики чи банку, які не надають такі пристрої. Люди використовують персональні комп'ютери для зберігання необхідної інформації, даних і файлів. Але якими б корисними не були ПК, технологія з кожним днем стає все більш необхідною та популярною, щоб уникнути значного дискомфорту, спричиненого використанням комп'ютера. Великі системні блоки, потреба у великому обсязі пам'яті та спеціальному робочому місці, значна кількість шнурів, випадки поломок - це проблеми, які необхідно вирішувати якнайшвидше для продовження технічного прогресу людства. В сучасному цифровому світі, де ми постійно генеруємо та зберігаємо великі обсяги даних, ефективне керування файлами стає життєво необхідною потребою. Саме тут на допомогу приходять файлові менеджери – програмні інструменти, які надають користувачам зручний інтерфейс для взаємодії з файловою системою комп'ютера.

Аналіз досліджень. Існує багато різних файлових менеджерів, доступних для різних операційних систем. Деякі популярні файлові менеджери для Windows включають Провідник Windows, Total Commander та FreeCommander. Деякі популярні файлові менеджери для Linux включають Nautilus, Dolphin та Krusader. Файлові менеджери відрізняються за функціональністю, можливостями та інтерфейсом користувача.

Мета роботи: Метою цієї роботи є розробка простого файлового менеджера для операційної системи, який буде підтримувати основні операції над файлами та папками. Файловий менеджер написаний засобами мови програмування Python.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. В сучасному цифровому світі, де ми постійно генеруємо та зберігаємо великі обсяги даних, ефективне керування файлами стає життєво необхідною потребою. Саме тут на допомогу приходять файлові менеджери – програмні інструменти, які надають користувачам зручний інтерфейс для взаємодії з файловою системою комп'ютера. Файлові менеджери надають графічний інтерфейс користувача (GUI), який полегшує навігацію по файловій системі та виконання операцій над файлами та папками. Існує багато різних файлових менеджерів, доступних для різних операційних систем,

кожен з власним набором функцій та можливостей. Файлові менеджери є невід'ємним інструментом для керування файлами та папками на комп'ютері. Вони надають графічний інтерфейс користувача, який полегшує навігацію по файловій системі та виконання операцій над файлами та папками. Файлові менеджери можуть використовуватися для створення, видалення, копіювання, переміщення та пошуку файлів та папок, а також для перегляду їх властивостей. Вибір відповідного файлового менеджера залежить від індивідуальних потреб та вподобань користувача. Деякі користувачі можуть віддавати перевагу файловим менеджерам з багатим набором функцій, тоді як інші можуть віддавати перевагу простішим та зручнішим у використанні.

Менеджер файлів — це комп'ютерна програма, яка забезпечує інтерфейс користувача, необхідний для роботи з каталогами та файлами. Менеджер файлів дозволяє виконувати основні операції з файлами - створювати, відкривати і перемішувати, переглядати, редагувати, видаляти, перейменовувати, шукати файли, копіювати, змінювати властивості і властивості, призначати дозволи. Менеджери файлів забезпечують більш зручний та інтуїтивно зрозумілий спосіб спілкування з вашим ПК, ніж операційна система (ОС). Однією з найвідоміших ранніх програмних оболонок був Norton Commander. Його розробив американський програміст Пітер Нортон. Цей файловий менеджер чітко відображає на моніторі всю файлову структуру пристрою: диски, каталоги та файли. Використання такої програми позбавляє від необхідності вводити складні команди MS-DOS у командному рядку.

В даний час в операційній системі Windows є свій інструмент для візуальної роботи з файловою системою - вбудована програма Провідник. Однак, незважаючи на це, файлові менеджери все ще користуються великою популярністю. Таким чином, файловий менеджер - це програма, спеціально розроблена для полегшення спілкування користувача з командами операційної системи. Ці програми працюють під керуванням операційної системи і займають проміжне положення між операційною системою та додатками. Крім основних функцій, багато додаткових функцій: мережеві операції, можливість налаштування інтерфейсу, можливість створення макросів і інших зовнішніх функцій пристроїв. Розрізняють основні типи файлових менеджерів: навігаційний і пробіл - іноді можуть підтримувати перемикання між цими режимами; дві панелі - зазвичай оснащені двома однаковими панелями для списку файлів і дерева каталогів. Представником навігаційного (космічного) файлового менеджера є звичайний Windows Explorer. Це не найкраща, але корисна програма. Він не завжди здатний виконувати операції з великими файлами, наприклад, коли вам потрібно перемістити або перейменувати сотню фотографій. При роботі з великою кількістю файлів Провідник може значно уповільнити роботу системи. Але найсуттєвішим недоліком є відсутність другої панелі для копіювання або переміщення файлів. Інші відомі навігаційні (просторові) файлові менеджери включають: Pocket Total Commander, Directory Opus Magellan, Software Standard File Manager, Resco Explorer, Workbench — включено в AmigaOS, Konqueror — включено в KDE, Dolphin — включено в KDE 4, Nautilus (файловий менеджер) – поставляється з GNOME, Thunar – поставляється з Xfce, Bynarys Smart Explorer, ROX-Filer – поставляється з ROX Desktop. У другому типі менеджера (двопанельному), який використовується для багатьох операцій, вікно програми розділене на дві частини. При цьому робота з файлами стає зручнішою та швидшою. А якщо врахувати, що більшість файлових менеджерів дозволяють контролювати всі дії з клавіатури, то швидкість роботи зростає в кілька разів. Найвідоміші ортодоксальні (двопанельні) файлові менеджери: Norton Commander, Total Commander, DOS Navigator, Volkov Commander, PIE Commander, Microsoft Windows, FreeCommander, Nomad.NET, Unreal Commander, File Navigator, GNOME Commander.

Програма Windows Explorer — це спеціальний плагін типу файлового менеджера, вбудований в операційну систему Windows. Програма призначена для навігації по файловій структурі Windows і її обслуговування (копіювання, перейменування, видалення і переміщення папок і файлів). Запуск програми здійснюється через команду Пуск/Програми/Провідник. Провідник Windows одночасно відображає вкладену структуру папок на комп'ютері (її ієрархію) і відображає внутрішню частину вибраної папки. Це дуже зручно під час копіювання та переміщення даних: просто відкрийте папку з потрібними файлами та перетягніть їх у папку призначення. Коли вам потрібно відкрити, скопіювати, перемістити, видалити, перейменувати чи іншим чином змінити папки чи файли, ви можете використовувати Провідник файлів або працювати з папками та файлами безпосередньо на робочому столі. Основною перевагою, на відміну від звичайних вікон папок на Робочому столі, є те, що програма Провідник надає можливість одночасно працювати як з вмістом правої панелі вікна, так і з усією архітектурою файлової системи комп'ютера користувача, тобто з

лівою панеллю. Це, наприклад, значно спрощує копіювання файлів із правої панелі в каталог або логічний диск, який зараз відображається на лівій панелі. При цьому немає необхідності відкривати велику кількість вікон, як це довелося б робити на Робочому столі. Робота з утилітою Провідник мало чим відрізняється від роботи зі звичайними вікнами Windows, за винятком того, що користувач може одночасно переглядати ієрархічну структуру своєї файлової системи. Він може клацнути правою кнопкою миші будь-яку папку та переглянути її вміст або клацнути будь-яку піктограму, щоб перетягнути її в інше місце.

Total Commander, колишній Windows Commander, є найпопулярнішим файловим менеджером для ОС Windows, крім Windows Explorer, і розроблений для зручної роботи з файлами та каталогами користувача. Подібно до Провідника, він дозволяє виконувати всі основні операції з файлами та каталогами, такі як копіювання, переміщення, видалення та зміни. Однак Total Commander використовує інший, більш зручний підхід до організації інтерфейсу: він має дві постійні панелі, які спрощують навігацію користувача та полегшують пошук необхідного місця для копіювання файлу. В усьому цьому допомагають постійно видимі каталог джерела файлу та каталог приймача для зручності користувача. Frigate - це дуже потужний файловий менеджер, який дозволяє швидко, інтуїтивно зрозуміло і комфортно працювати з комп'ютером. На сьогоднішній день Frigate є найбільш функціональним та потужним засобом для роботи з файлами.

Основні характеристики Frigate: Цей файловий менеджер включає в себе підтримку багатозадачності при виконанні файлових операцій, що дозволяє користувачу виконувати декілька операцій одночасно. Він також забезпечує підтримку усіх відомих архівів, що полегшує роботу з архівами. Навігація в менеджері є зручною завдяки "швидким" папкам та історією переміщення по папках. Така система навігації полегшує пошук та доступ до потрібних файлів. Крім цього, у менеджері доступні корисні утиліти, такі як калькулятор, системна інформація, телефонна книга, зберігач паролів та інші, що розширює його функціональність і корисність для користувача. Також, вбудований перегляд (F3) дозволяє швидко переглядати RTF-текст і файли DBF, а вбудований редактор (F4) надає можливість редагувати текст з підсвічуванням синтаксису і RTF-текстові файли. Менеджер також забезпечує відновлення файлів після обриву копіювання, що є важливою функцією для збереження даних. Крім цього, він підтримує перегляд і редагування MP3-тегів, вдосконалений пошук та файловий фільтр для швидкого відбору типів файлів. Користувач також може працювати з картинками, оскільки менеджер підтримує всі популярні графічні формати. Інші можливості Frigate включають широку функціональність, гнучкість у налаштуваннях та інтеграцію з іншими додатками.

У нашій роботі ми використали мову програмування Python, оскільки основними архітектурними особливостями її є динамічна типізація, механізм обробки винятків, автоматичне керування пам'яттю, підтримка багатопоточних обчислень і повний самоаналіз. Динамічна семантика та прив'язки, а також високорівневі структури даних роблять цю мову привабливою для прискореної розробки додатків. Python підтримує велику кількість парадигм програмування, включаючи об'єктно-орієнтовану, процедурну, імперативну, функціональну та аспектно-орієнтовану. Крім того, Python є кросплатформною і не вимагає додаткових зусиль для перенесення з однієї платформи на іншу.

Інтерпретатор написаний на C, а вихідний код доступний для будь-яких маніпуляцій. Якщо необхідно, можна вставити його у свою програму та використовувати, як вбудовану оболонку. Або, написавши свої доповнення на C, отримати «просунутий» інтерпретатор з новими можливостями. Корисним є функціонал розширення для програм, які потребують подальшого налагодження. Підтримка модулів та пакетів модулів значно збільшує повторне використання коду, що корисно під час розробки великих програм. Інтерпретатор Python і стандартна бібліотека доступні у скомпільованій та вихідній формі на всіх основних платформах. Для виконання програми потрібно використання бібліотеки send2trash. Вона використовується для перенесення файлів до смітника. Її можна встановити за командою `pip install send2trash`. Обговоримо основні функції, які ми використали у цій роботі.

Імпорт необхідних модулів:

```

1  #! python3
2
3  import sys
4  import os
5  import shutil
6  import send2trash

```

Рис. 1. Початок програми, підключення бібліотек

sys: Цей модуль використовується для доступу до вбудованих функцій та змінних Python, таких як exit(). os: Цей модуль використовується для взаємодії з операційною системою, наприклад, для отримання доступу до файлів та каталогів, переміщення каталогів та видалення файлів. shutil: Цей модуль використовується для копіювання, переміщення та видалення файлів та каталогів. Send2trash: Цей модуль використовується для надсилання файлів та каталогів до кошика.

Змінна drives для зберігання списку доступних дисків:

```

10 # Зберігає кожний підключений до ПК диск у списку.
11 drives = [chr(x) + ':' for x in range(65, 90) if os.path.exists(chr(x) + ':')]

```

Рис. 2. Функція збереження списку доступних дисків

Цей код використовує цикл for, щоб перебрати символи з A по Z (65 - 90 в ASCII). Для кожного символу перевіряється, чи існує диск з такою літерою, використовуючи os.path.exists(). Якщо диск існує, його додається до списку drives.

Функція listDirectories() для виведення списку файлів та каталогів:

```

13 # Виводить список кожної теки та файлу, що знаходяться в поточній робочій директорії.
14 10 usages
15 def listDirectories():
16     listdir = os.listdir(os.getcwd())
17     for x in listdir:
18         print(x)

```

Рис. 3. Функція для виведення списку файлів та каталогів

Ця функція використовується для виведення списку файлів та каталогів у поточній робочій директорії. Вона використовує os.listdir() для отримання списку елементів у поточній директорії. Потім вона циклічно перебирає цей список і друкує назву кожного елемента.

Головний цикл while True для роботи менеджера файлів:

```

19 while True:
20     print("1. Відкрити файл/теку \n2. Переіменувати \n3. Перемістити та вставити \n4. Скопіювати та вставити \n5. Видалити\n")
21     result = input("Оберіть одну з наступних опцій: ")
22
23     if result == '1':
24         # Головний екран
25         print("\nВидки доступу:\n1. Документи\n2. Відео\n3. Зображення\n4. Завантаження\n")
26
27         print("Диски: ")
28         for x in range(len(drives)):
29             print(str(5 + x) + '. ' + drives[x])
30
31         while True:
32             inp = input("\nВведіть ваш вибір: ")
33
34             if inp == '1':
35                 path = 'C:\\Users\\$USERNAME\\Documents'
36                 os.chdir(os.path.expandvars(path))
37                 break
38
39             elif inp == '2':
40                 path = 'C:\\Users\\$USERNAME\\Videos'
41                 os.chdir(os.path.expandvars(path))
42                 break
43
44             elif inp == '3':
45                 path = 'C:\\Users\\$USERNAME\\Pictures'
46                 os.chdir(os.path.expandvars(path))

```

Рис. 4. Головний цикл while True. Частина 1

Цикл while True забезпечує безперервну роботу менеджера файлів, поки користувач не вирішить вийти. У середині циклу: Відображається меню з 5 опціями. Користувач вводить свій вибір. Вибір обробляється за допомогою вкладених операторів if elif else. Якщо введено неправильний вибір, друкується повідомлення про помилку.

Обробка вибору користувача відбувається через відкриття файлу/папки. Фрагмент коду нижче використовується для надання швидкого доступу до документів, відео, зображень та завантажень. Користувач може вибрати один із цих місць або перейти на диск.

```

44         elif inp == '3':
45             path = 'C:\\Users\\$USERNAME\\Pictures'
46             os.chdir(os.path.expandvars(path))
47             break
48
49         elif inp == '4':
50             path = 'C:\\Users\\$USERNAME\\Downloads'
51             os.chdir(os.path.expandvars(path))
52             break
53
54         elif inp in drives:
55             os.chdir(inp + '\\')
56             break
57
58         else:
59             print('Помилка\nВведіть правильний ввід або назву диска.\n')
60

```

Рис. 5. Головний цикл while True. Частина 2

Після вибору диска або місця код переходить до каталогу та відображає список файлів та каталогів у ньому. Користувач може вибрати файл, щоб відкрити його, або каталог, щоб перейти до нього. Користувач також може ввести exitManager, щоб вийти з менеджера файлів, або backManager, щоб перейти на один каталог вище. **Перейменувати:** Користувач може вибрати диск, щоб перейти до нього. Після вибору диска код відображає список файлів та каталогів у ньому. Користувач може вибрати файл для перейменування. Користувач буде запитуватися на введення нового імені, а потім старе ім'я буде змінено на нове за допомогою функції os.rename. Користувач також може ввести exitManager, щоб вийти з менеджера файлів, або backManager, щоб перейти на один каталог вище.

```

85         if result == '2':
86             print("Ви вибрали перейменувати")
87             print('Диски: ')
88             for x in range(len(drives)):
89                 print(str(1 + x) + '. ' + drives[x])
90
91             while True:
92                 inp = input("\nВведіть ваш вибір: ")
93
94                 if inp in drives:
95                     os.chdir(inp + '\\')
96                     break
97                 else:
98                     print('Помилка\nВведіть правильну назву диска.\n')
99

```

Рис. 6. Функція для виконання опції «Перейменувати»

Перемістити та вставити: Цей код реалізує файловий менеджер з можливістю перегляду та переміщення файлів та каталогів. Після запуску користувачу відображається список дисків. Він може обрати диск, щоб перейти до нього. Після вибору диска відображається список файлів та каталогів у ньому. Користувач може обрати файл або каталог для переміщення. Потім користувачу буде запропоновано ввести місце призначення, і обраний файл або каталог буде переміщено до нового розташування за допомогою функцій shutil.move(). Крім того, користувач може ввести команду exitManager, щоб вийти з менеджера файлів, або backManager, щоб перейти на один каталог вище.

Скопіювати та вставити: Цей код створює файловий менеджер з можливістю перегляду та копіювання файлів та каталогів. При запуску програми користувачу відображається список дисків, з якими він може працювати. Він може обрати диск для подальшої роботи. Після вибору диска відображається список файлів та каталогів, що містяться на цьому диску. Користувач може обрати файл або каталог, який він бажає скопіювати. Після цього він буде запитаний про нове місце розташування, куди слід скопіювати вибраний файл або каталог. Копіювання виконується за допомогою функцій shutil.copytree() для каталогів та shutil.copy2() для файлів. Крім того, користувач може ввести команду exitManager, щоб вийти з менеджера файлів, або backManager, щоб перейти на один каталог вище.

Видалити: Цей код включає опцію для користувача вибрати, чи він хоче видалити файл або каталог назавжди, чи перемістити його до кошика. Якщо користувач обере видалення назавжди, він отримає попередження про можливі наслідки цієї дії та буде запрошений підтвердити видалення. Якщо видалення підтверджено, відповідний файл або каталог буде видалено назавжди за допомогою функцій `os.remove()` для файлів та `shutil.rmtree()` для каталогів.

У випадку, якщо користувач вирішить перемістити файл або каталог до кошика, вони будуть переміщені до кошика за допомогою модуля `send2trash`. Це надає можливість відновлення файлів або каталогів у разі необхідності, адже вони не видаляються остаточно, а лише переміщуються до системного кошика.

Обробка помилок вводу користувача в файловому менеджері включає перевірку правильності введених користувачем опцій. Коли користувач вводить опцію, код перевіряє, чи вона є дійсною та чи належить до діапазону доступних опцій. Це реалізується за допомогою операторів `if`, `elif` і `else`. Якщо введена опція неправильна, код виводить повідомлення про помилку та просить користувача ввести опцію знову, поки не буде введено правильну.

Також в файловому менеджері передбачена обробка помилок вводу назви диска, якщо користувач виконує дії, такі як перейменування, переміщення, копіювання або видалення файлів або каталогів. Код перевіряє правильність введеної назви диска, наприклад, чи існує такий диск та чи є введена назва диском. Якщо введена назва диска неправильна, виводиться повідомлення про помилку, і користувачу пропонується ввести правильну назву.

Додаткові функції, такі як `backManager` та `exitManager`, реалізовані для полегшення навігації користувача в менеджері файлів. Функція `backManager` дозволяє користувачеві перейти на один рівень вище в поточній робочій директорії, щоб швидше навігуватися по системі. Функція `exitManager` дає можливість користувачеві вийти з менеджера файлів у будь-який момент. Обидві ці функції спрощують використання менеджера файлів та зроблюють його більш зручним для користувача.

Вихід: Користувач може вийти з менеджера файлів, ввівши `exit` або `exitManager` у головному меню. У виводі цієї програми можна отримати такі результати:

```
C:\Users\Asus\PycharmProjects\pythonProject1\.venv\Scripts\python.exe C:\Users\Asus\PycharmProjects\pythonProject1\main.py
Ласкаво просимо до файлового менеджера на Python!

1. Відкрити файли/теки
2. Перейменувати
3. Перемістити та вставити
4. Скопіювати та вставити
5. Видалити

Оберіть одну з наступних опцій:
```

Рис. 7. Початок виводу програми

Наприклад, якщо ви оберете опцію 1, вивід буде таким:

```
Оберіть одну з наступних опцій: 1

Швидкий доступ:
1. Документи
2. Відео
3. Зображення
4. Завантаження

Диски:
5. C:
6. D:
7. E:

Введіть ваш вибір:
```

Рис. 8. Вивід опції 1

```
Введіть ваш вибір: 1
cc_20220319_000048.reg
desktop.ini
IISExpress
iZotope
LabVIEW Data
MATLAB
My Music
My Pictures
My Videos
My Web Sites
National Instruments
ViberDownloads
Visual Studio 2022
Zoom
Без назви-1.dxf
Без назви-1_1.png
Без назви-1_2.png
Без назви-1_3.png
Настраиваемые шаблоны Office
Настроенные шаблоны Office
Пользовательские шаблоны Office
```

Рис. 9. Вивід тек для опції 1 на диску 1. Частина 1

```
Напишіть "exitManager", щоб вийти з менеджера файлів.
Напишіть "backManager", щоб перейти вище на один каталог.

Оберіть файл/теку:
```

Рис. 10. Вивід тек для опції 1 на диску 1. Частина 2

Якщо обрати іншу опцію зі списку, вивід буде побудований по тому ж принципу. Першою дією буде вибір опції, наступним кроком потрібно буде обрати диск з яким постає необхідність взаємодіяти. Далі обираються папки для роботи.

Висновки.

У сучасному світі, де комп'ютери та цифрові технології стали невід'ємною частиною нашого життя, виникає гостра потреба у покращенні та спрощенні роботи з ними. Існує ряд проблем, з якими стикаються користувачі персональних комп'ютерів. Для вирішення цих проблем необхідні інноваційні рішення, спрямовані на покращення ергономіки, простоти використання та надійності комп'ютерних систем. Розробка простого файлового менеджера є одним із актуальних кроків у цьому напрямку. Отож, був розроблений простий файловий менеджер для операційної системи, який відповідає поставленим цілям. Файловий менеджер підтримує основні операції над файлами та папками та написаний чітким та зрозумілим кодом.

Список бібліографічного опису

1. DSpace. ELAKPI: Репозитарій КПІ ім. Ігоря Сікорського. URL: <https://ela.kpi.ua/server/api/core/bitstreams/98d6a1b6-b74a-4c04-9e26-6ce460015be6/content> (дата звернення: 10.05.2024).
2. Міністерство освіти і науки України. Інститут інформатики та радіоелектроніки, Факультет комп'ютерних наук і технологій. Кафедра програмних засобів. РОЗРОБКА ПРОГРАМИ ФАЙЛОВОГО МЕНЕДЖЕРА ДЛЯ ОПЕРАЦІЙНОЇ СИСТЕМИ ANDROID. FILE MANAGER DEVELOPMENT FOR ANDROID OS. URL: https://eir.zntu.edu.ua/bitstream/123456789/10723/1/BR_Burdakova.pdf (дата звернення: 10.05.2024).

References

- 1.DSpace. ELAKPI: KPI Repository. Igor Sikorsky. URL: <https://ela.kpi.ua/server/api/core/bitstreams/98d6a1b6-b74a-4c04-9e26-6ce460015be6/content> (Last accessed: 10.05.2024).
- 2.Ministry of Education and Science of Ukraine. Institute of Informatics and Radioelectronics, Faculty of Computer Science and Technology. Department of Software. DEVELOPMENT OF A FILE MANAGER APPLICATION FOR THE ANDROID OPERATING SYSTEM. FILE MANAGER DEVELOPMENT FOR ANDROID OS. URL: https://eir.zntu.edu.ua/bitstream/123456789/10723/1/BR_Burdakova.pdf (Last accessed: 10.05.2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-29>

УДК 614.8:004.942:616-083

Чибіряк Яна Іванівна¹, к.т.н., доцент,

<https://orcid.org/0000-0002-0634-7609>

Захарова Анна Миколаївна¹, бакалавр,

<https://orcid.org/0000-0001-5242-4065>

Лавров Євгеній Анатолійович¹, д.т.н., професор,

<https://orcid.org/0000-0001-9117-5727>

Сірик Ольга Євгенівна², к.ф.-м.н.,

<https://orcid.org/0000-0001-9360-4388>

Баранова Ірина Володимирівна¹, к.т.н., доцент,

<https://orcid.org/0000-0002-3767-8099>

Татарчук Руслан Миколайович¹, магістр

¹Сумський державний університет, м. Суми, Україна

²Київський національний університет імені Тараса Шевченка, м. Київ, Україна

КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ДЛЯ МЕДИЦИНИ КАТАСТРОФ. ПІДХІД ДО АВТОМАТИЗОВАНОГО АНАЛІЗУ РІШЕНЬ З ОРГАНІЗАЦІЇ МОБІЛЬНИХ МЕДИЧНИХ ШПИТАЛІВ

Чибіряк Я.І., Захарова А.М., Лавров Є.А., Сірик О.Є., Баранова І.В., Татарчук Р.М. Комп'ютерне моделювання для медицини катастроф. Підхід до автоматизованого аналізу рішень з організації мобільних медичних шпиталів. Засоби імітаційного моделювання все частіше використовуються спеціалістами органів управління для оцінювання ефективності роботи медичних організацій. У роботі запропоновано структуру та розроблено імітаційну модель мобільного медичного шпиталю, яка дозволяє оцінити робоче навантаження на окремі відділення шпиталю у залежності від інтенсивності вхідного потоку поранених. Побудовано імітаційну систему аналізу та візуалізації даних, що може бути використана для прийняття рішень щодо розподілу, мобілізації сил та засобів медичних служб в умовах бойових дій. Дані для побудови моделі були отримані з довідкової медичної літератури, також надавалися пропозиції від окремих медичних організацій міста Суми та від бойових медиків з прикордоння. Імітаційне моделювання здійснюється у програмі FlexSim. Дискретно-подійна модель побудована на базі математичного апарату, який використовується для оцінювання часових показників проведення лікувальних заходів під час обстрілів, пожеж та інших надзвичайних ситуацій. До них відносяться такі параметри, як тривалість надання медичної допомоги, середній час очікування поранених в чергах та ін. параметри. Імітаційна модель надає можливість виводити інформацію про завантаженість окремих відділень, це дозволяє швидко приймати рішення про зосередження сил і засобів у потрібних підрозділах шпиталю. Впровадження імітаційної моделі медичного шпиталю в діяльність ВСУ дозволить у режимі реального часу управляти ефективністю надання медичної допомоги пораненим у зонах воєнних конфліктів.

Ключові слова: імітаційна модель, мобільний медичний шпиталь, час очікування, показники ефективності, інтенсивність надходження, довжина черги, робоча завантаженість, структура системи, пропускна здатність.

Chybiriak Ya. I., Zakharova A. M., Lavrov E. A., Siryk O. E., Baranova I. V., Tatarchuk R. M. Computer modeling for disaster medicine. An approach to automated analysis of solutions for organizing mobile medical hospitals. Experts of authority bodies increasingly use simulation tools to evaluate the effectiveness of medical units. This paper suggests a structure and develops a simulation model of a mobile medical hospital that allows one to estimate the workload of each hospital department depending on the intensity of the incoming flow of wounded. A simulation system for analyzing and visualizing data is built. It can be used to make decisions on the distribution and mobilization of forces and means of medical services in combat conditions. Data for building the model were obtained from the reference medical literature. Separate medical organizations in the city of Sumy and combat medics working at the border area also provided their suggestions. The simulation was performed in the FlexSim program. The discrete-event model is based on a mathematical apparatus used to estimate the timing of medical treatment during shelling, fires, and other emergencies. These include such parameters as the duration of medical care, the average waiting time for the wounded in queues, etc. The simulation model provides an opportunity to display information about the workload of each department, which allows one to quickly make decisions about the allocation of forces and resources in the necessary departments of the hospital. The usage of a simulation model of a medical hospital by the Armed Forces of Ukraine will make it possible to manage in real time the effectiveness of providing medical care to the wounded in war zones.

Keywords: simulation model, mobile medical hospital, waiting time, performance indicators, admission rate, queue length, workload, system structure, throughput.

Вступ. У регіонах України, де тривають бойові дії, медична служба стикається із задачею створення та підвищення ефективності і якості роботи спеціалізованих мобільних медичних підрозділів, призначених для надання своєчасної медичної допомоги населенню під час масованих обстрілів, вибухів, пожеж та інших надзвичайних ситуацій, викликаних воєнними діями.

Зокрема, до таких медичних формувань відносяться:

- медичні мобільні шпиталі;

– медичні мобільні загони.

Діяльність даних формувань регламентується "Положенням про Державну службу медицини катастроф", затвердженим Постановою Кабінету Міністрів України від 11.07.2001 р. № 827 [1].

Зокрема, у сфері управління МНС України створено мобільний медичний госпіталь, який брав участь у організації ліквідації наслідків надзвичайних ситуацій, у забезпеченні медичних масових заходів, медичних навчань в Україні та за кордоном [2].

Мобільні бригади лікарів надавали медичну допомогу у населених пунктах Київщини після деокупації Бучі, Ірпеню, Бородянки, та в інших постраждалих регіонах [3].

У своїй більшості медики працюють на волонтерських засадах або у рамках спеціальних проєктів. Так, на Дніпропетровщині діють мобільні медичні бригади, що створені за підтримки Товариства Червоного Хреста України [4].

У жовтні 2023 року Україна отримала перший польовий мобільний госпіталь від німецької компанії "Rheinmetall ZMS" загальною вартістю 9 млн. євро [5]. У систему входять відділення невідкладної допомоги, операційні, інтенсивної терапії, аптеки. Присутні намети для розміщення медичного персоналу, адміністративної частини, зони очікування (рис.1).



Рис. 1. Польовий госпіталь від німецької компанії "Rheinmetall ZMS"[[5]]

У квітні 2024 року на Сумщині у місті Охтирка відкрито модульну станцію, призначену для надання екстреної допомоги та у селі Хухра Охтирського району встановлено модульну медичну амбулаторію [6] (рис. 2).

Необхідність у розгортанні таких польових медичних шпиталів виникає ще й у зв'язку з тим, що на окремих територіях руйнуються заклади охорони здоров'я, техніка, медичні засоби.



Рис. 2. Амбулаторія модульного типу в м. Охтирка [[6]]

Особливої уваги потребує ситуація, що складається на прикордонних територіях на півночі та сході нашої країни. Існуюча організаційно-технічна система при збільшенні навантаження не здатна забезпечити своєчасне надання медичної допомоги пораненим.

Тому задача даного дослідження полягає у комп'ютерному моделюванні процесу функціонування мобільного медичного шпиталю, призначеного для вчасного надання медичної допомоги пораненим в умовах бойових дій.

Аналіз існуючих досліджень та публікацій. У дослідженнях, проведених методами комп'ютерного моделювання в Українській військовій медичній академії, приділялася увага вирішенню наступних задач:

- оцінювання ефективності роботи медичних закладів [7];
- удосконалення структур медичних формувань [8];
- раціональний розподіл медичного персоналу [9].

Для дослідження систем масового обслуговування використовуються два основних методи дослідження ефективності систем – аналітичний та імітаційний, що передбачає дослідження систем на комп'ютерних моделях.

У наукових публікаціях описано використання методів та підходів, що застосовуються для дослідження медичних систем. Так, у [10] вирішується задача скорочення середнього часу очікування методом спеціальної організації черг у медичній системі з дефіцитом ресурсів, при цьому робочі процеси представляються Марківськими процесами, використовується чисельне моделювання у середовищі програмного забезпечення ProModel.

У статті [11] для сервісних медичних систем типу М/М/1, аналітичними методами досліджується тривалість очікування пацієнтів в чергах у залежності від різної частоти прибуття.

В [12] мовою програмування GPSS розроблено імітаційну модель медичного закладу. Медична система, що відображає процес функціонування медичного центру, представлена мережею масового обслуговування відкритого типу, що складається з 5 окремих вузлів, які є системами масового обслуговування. Модель оптимізована за кількістю відмов та за довжинами черг на прийом до лікаря.

Робота [13] присвячена вирішенню задачі раціонального розташування каналів обслуговування у медичній системі з дефіцитом персоналу, медичного обладнання та інших ресурсів. Імітаційне моделювання виконано у програмному середовищі Arena.

Таким чином, у теорії масового обслуговування використовується два основних підходи по дослідженню ефективності медичних систем: аналітичний та імітаційний.

Аналітичний підхід для моделювання складних стохастичних систем є непридатним, оскільки призводить до [14], [15], [16]:

- обчислювальної складності;
- низької точності результатів;
- значних часових витрат.

На противагу аналітичним, імітаційні моделі надають можливість [17], [18], [19]:

– швидко, у режимі реального часу, проводити дослідження та на їх основі приймати ефективні рішення;

- моделювати складні стохастичні процеси;
- отримувати результати моделювання із заданою точністю.

Тому, у даній роботі для моделювання та дослідження процесів пересувного медичного шпиталю використано метод імітаційного моделювання [20]. У якості програмного середовища моделювання обрано програму FlexSim [16].

Постановка задачі дослідження. Метою роботи є створення імітаційної моделі пересувного медичного шпиталю для дослідження ефективності роботи при різних значеннях інтенсивності вхідного потоку поранених.

Для досягнення поставленої мети необхідно виконати наступні задачі:

– розробити структуру медичного шпиталю, виконати опис робочих процесів системи шпиталю;

- формалізувати робочі процеси шпиталю у вигляді мережі масового обслуговування;
- виконати програмну реалізацію моделі у середовищі імітаційного моделювання FlexSim;
- виконати дослідження роботи шпиталю протягом робочого тижня.

У ході моделювання планується виявити «вузькі місця» та оцінити роботу медичної системи за наступними показниками ефективності:

- середні та максимальні довжини черг;
- середній та максимальний час очікування в чергах;
- середній час перебування постраждалих у системі;
- завантаженість робочих місць та людських ресурсів;
- пропускна здатність шпиталю.

Виклад основного матеріалу та обґрунтування отриманих результатів дослідження.

Структура пересувного медичного шпиталю спроектована по блочно-модульному принципу і є подібною до класичного варіанту організації медичної допомоги з використанням функціональних підрозділів в наметах. Структурно-логічна схема пересувного медичного шпиталю представлена на рис. 3.

Вхідний потік поранених надходить до системи за законом розподілу Пуассона і потрапляє до приймального відділення, де визначається ступінь важкості та вид поранення.

За ступенем важкості поранення потерпілі поділяються на чотири категорії:

- легкі поранення – 19 %;
- поранення середнього ступеню важкості – 27%;
- важкопоранені – 42%;
- дуже важкі поранення – 12%.

За видом поранення потерпілі поділяються на дев'ять груп:

- поранення і травми кінцівок – 20%;
- поранення і травми грудної клітини – 17.5%;
- поранення і травми голови – 16%;
- поранення і травми щелепно-лицьової ділянки – 9.7%;
- поранення і травми органів зору – 9%;
- поранення і травми черевної частини – 8.7%;
- опіки – 8.6%;
- поранення і травми хребта – 7.5%;
- психогенно-реактивні стани, контузії – 3%.

Після приймального відділення поранені надходять у відповідні відділення у залежності від важкості та від типу поранення.

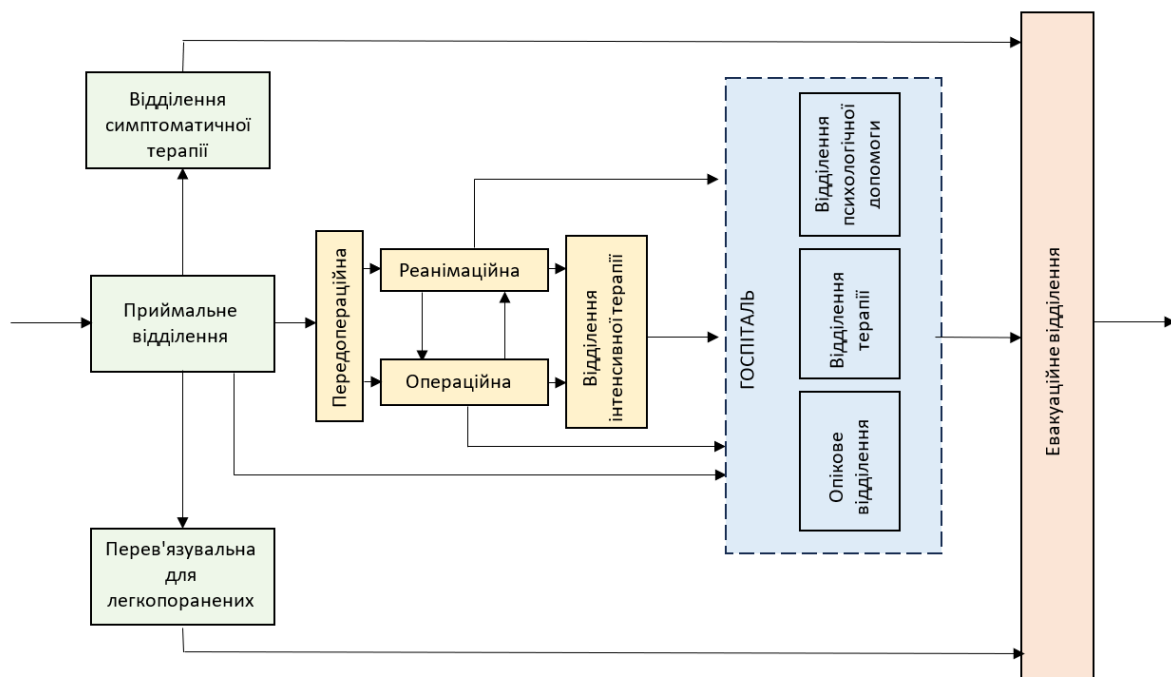


Рис 3. Структурно-логічна схема пересувного медичного шпиталю

Час надання медичної допомоги у підрозділах шпиталю визначається за трикутним законом розподілу, параметри якого залежать від ступеню важкості поранення.

Медичні системи відносяться до типу сервісних систем, їх робочі процеси можна представити мережею масового обслуговування, до складу якої входять обслуговуючі канали –

робочі ресурси (ліжкомісця, медичне обладнання, персонал та ін.) та черги, де відбувається очікування надання медичної допомоги.

У теорії черг найбільш відомими і поширеними серед спеціалістів засобом формалізації процесів функціонування дискретних систем є схеми у вигляді мереж масового обслуговування, що відображають склад, структуру системи, а також порядок надходження і правила обслуговування вимог (заявок).

Формалізація процесів системи мобільного медичного шпиталю у вигляді мережі масового обслуговування показана на рис. 4.

На рис. 4 використано умовні позначення:

- $\boxed{ПВ_i}$ – робочі місця приймального відділення ($i = \overline{1, n}$);
- $\boxed{ПЛ_i}$ – робочі місця у перев'язувальній для легкопоранених ($i = \overline{1, k}$);
- $\boxed{СТ_i}$ – робочі місця у відділенні симптоматичної терапії ($i = \overline{1, m}$);
- $\boxed{ПО_i}$ – робочі місця у передопераційній ($i = \overline{1, s}$);
- $\boxed{Р_i}$ – робочі місця у відділенні реанімації ($i = \overline{1, r}$);
- $\boxed{ОП_i}$ – робочі місця в операційній ($i = \overline{1, p}$);
- $\boxed{ІТ_i}$ – робочі місця у відділенні інтенсивної терапії ($i = \overline{1, t}$);
- $\boxed{ЕВ_i}$ – робочі місця у відділенні евакуації ($i = \overline{1, h}$);
- \diamond – умови переходу з одного відділення у інший ($i = \overline{1, z}$).

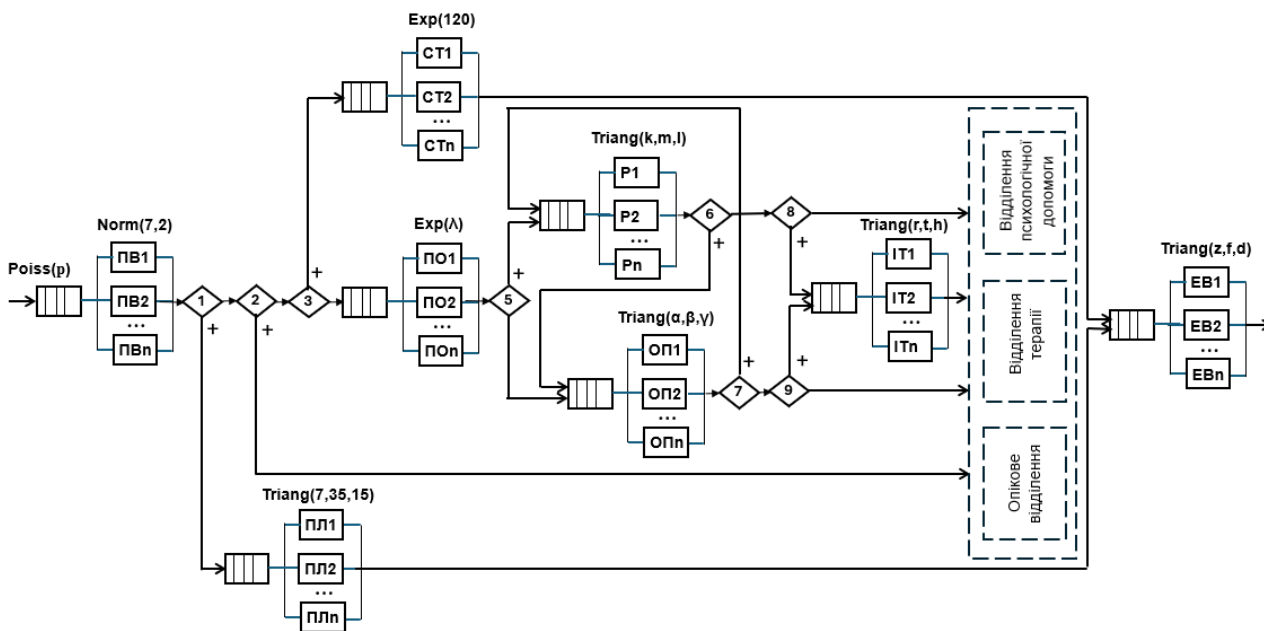


Рис 4. Мережа масового обслуговування системи мобільного медичного шпиталю

Програмна реалізація імітаційної моделі відбувається в кілька етапів:

- відтворення структури шпиталю;
- налаштування функціоналу об'єктів моделі;
- виведення показників ефективності у вигляді графіків і діаграм;
- налаштування візуалізації процесів.

Структура шпиталю, побудована з використанням бібліотеки фіксованих та мобільних об'єктів FlexSim, має вигляд, показаний на рис. 5.

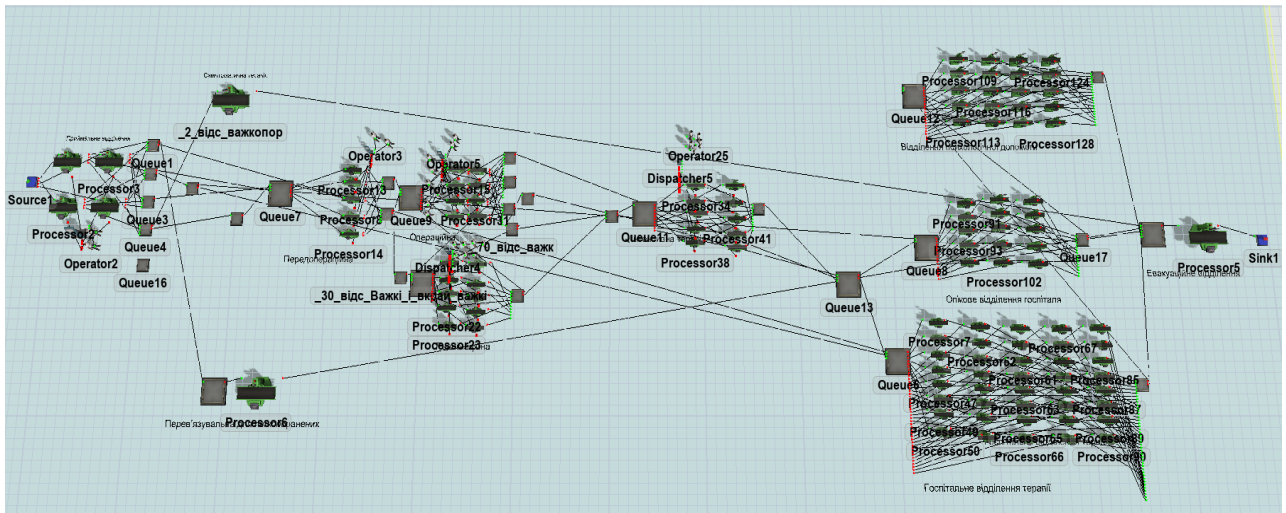


Рис. 5. Структура мобільного шпиталю у середовищі FlexSim

Ліжкомісця у відділеннях моделюються об'єктом Processor, який є фіксованим ресурсом і позначає канал обслуговування в системі шпиталю. Черги (Queues) використовуються для очікування елементів потоку, якщо ліжкомісця зайняті і відділення не можуть приймати заявки на обслуговування. Функції персоналу у моделі виконують оператори (Operators). У якості елементів потоку (заявок) виступають поранені, які надходять до відділень та затримуються на час надання медичної допомоги. За замовчуванням черги працюють за правилом «першим прийшов — першим обслужений» (FIFO). На програмному рівні було реалізовано пріоритетне обслуговування у залежності від важкості поранення.

Час надання допомоги у підрозділах шпиталю задається трикутним законом розподілу та залежить від типу і виду поранення.

У моделі автоматично розраховуються показники ефективності, які відображено у графічному вигляді на рис. 6:

- середній, мінімальний та максимальний час перебування поранених у шпиталі;
- пропускна здатність шпиталю;
- робоча завантаженість підрозділів шпиталю;
- довжини черг у відділеннях;
- час очікування поранених в чергах.

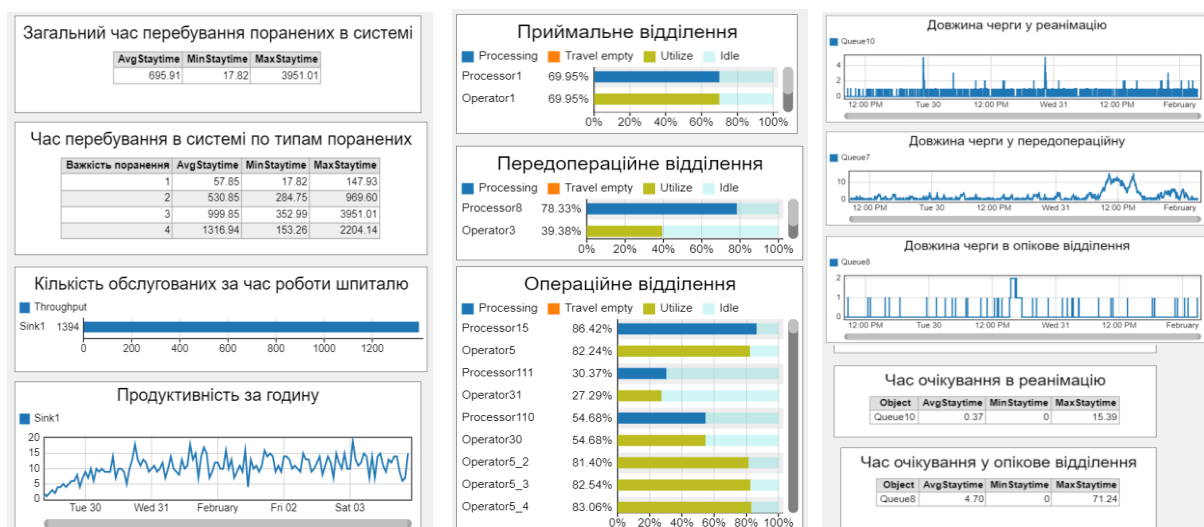


Рис. 6. Показники ефективності роботи системи шпиталю

На рис. 7 представлено фрагмент моделі мобільного медичного шпиталю, в якій налаштовано візуалізацію робочих процесів.

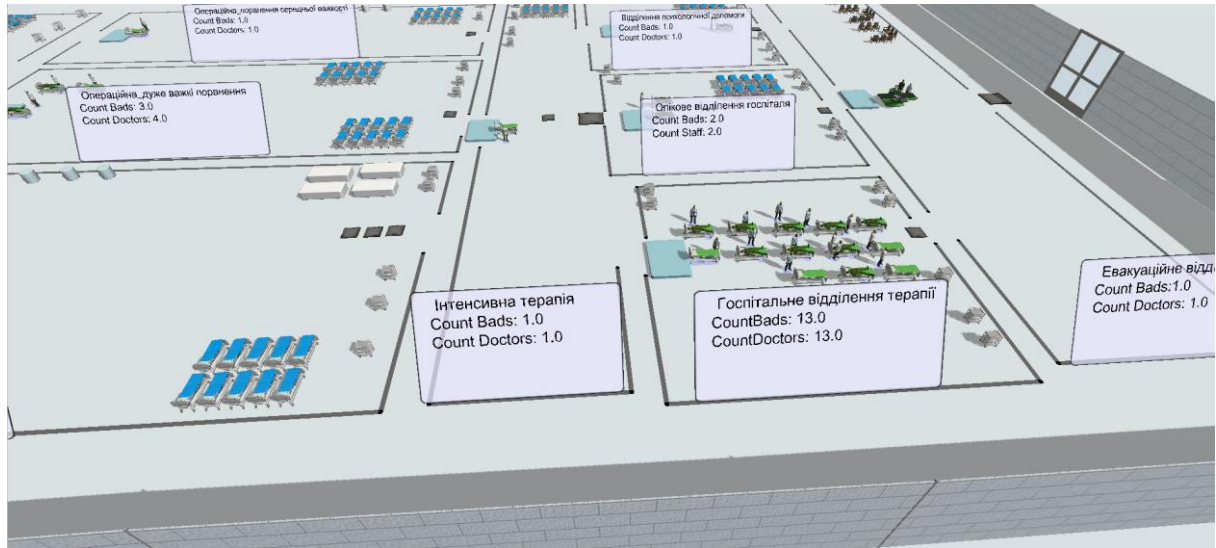


Рис. 7. Імітаційна модель мобільного медичного шпиталю (фрагмент)

На етапі дослідження проведено моделювання роботи шпиталю для базової конфігурації із заданою кількістю ліжок/місць та персоналу у відділеннях. Результати показали наявність «вузького місця» в операційній (табл. 1).

При надходженні 40 поранених/добу маємо такі показники:

- середнє значення довжин черг у відділеннях не перевищує 3-5 осіб;
- найбільша завантаженість 85% у операційному відділенні;
- найвищий середній час очікування 41 хв., також спостерігається в операційній.

Табл. 1. Результати імітаційного моделювання для базової конфігурації мобільного шпиталю

Структурний підрозділ	Кількість ліжок/місць	Кількість персоналу	Середня довжина черги, осіб	Середній час очікування, хв.	Завантаженість, %	Середній час перебування, хв.	Пропускна здатність, осіб.
Надходження 40 поранених/добу							
Приймальне відділення	1	1	0	0	22	-	-
Передопераційна	4	3	0	0	33	-	-
Реанімація	6	4	3	4	80	-	-
Операційна	12	12	5	41	85	-	-
Інтенсивна терапія	8	6	1	7	73	-	-
Психологічної допомоги	7	2	0	0	25	-	-
Опікове	2	2	2	5	80	-	-
Терапія	13	7	2	3	78	-	-
Евакуаційна	1	1	0	0	40	-	-
Увесь шпиталь	54	38	-	-	-	1300	120
Надходження 60 поранених/добу							
Приймальне відділення	1	1	0	0	33	-	-
Передопераційна	4	3	0	0	50	-	-

Реанімація	6	4	12	96	89	-	-
Операційна	12	12	21	720	92	-	-
Інтенсивна терапія	8	6	2	98	81	-	-
Психологічної допомоги	7	2	0	0	38	-	-
Опікове	2	2	9	27	92	-	-
Терапія	13	7	6	14	87	-	-
Евакуаційна	1	1	0	0	60	-	-
Увесь шпиталь	54	38	-	-	-	1950	148
Надходження 80 поранених/добу							
Приймальне відділення	1	1	0	0	44	-	-
Передопераційна	4	3	0	0	76	-	-
Реанімація	6	4	24	120	92	-	-
Операційна	12	12	40	1160	98	-	-
Інтенсивна терапія	8	6	4	140	97	-	-
Психологічної допомоги	7	2	0	0	50	-	-
Опікове	2	2	18	100	98	-	-
Терапія	13	7	12	116	90	-	-
Евакуаційна	1	1	0	0	75	-	-
Увесь шпиталь	54	38	-	-	-	2600	166

При збільшенні вхідного потоку до 60 поранених/добу:

- середня кількість осіб, що чекають в операційну зростає до 21;
- час очікування на операцію збільшується до 12 год.

Подальше збільшення інтенсивності надходження підтверджує наявність «вузького місця» в операційному відділенні. При надходженні 80 поранених/добу:

- середня довжина черга в операційну зростає до 40;
- середній час очікування досягає 48 год.
- додатково з'являються вузькі місця у відділеннях реанімації, опіковому та в інтенсивній терапії.

Для усунення «вузьких місць» рекомендовано збільшити кількість хірургічних бригад до 17, збільшити кількість ліжкомісць у реанімації до 10, у опіковому відділенні – до 12 та в інтенсивній терапії – до 16. Результати моделювання зі змінами, внесеними до базового варіанту моделі, показані в табл. 2.

Табл. 2. Результати імітаційного моделювання для базової конфігурації мобільного шпиталю зі збільшеною кількістю хірургічних бригад та ліжкомісць у відділеннях

Структурний підрозділ	Кількість ліжкомісць	Кількість персоналу	Довжина черги, осіб	Час очікування, хв.	Завантаженість, %	Час перебування, хв.	Пропускна здатність, осіб
Надходження 40 поранених/добу							
Приймальне відділення	1	1	0	0	22	-	-
Передопераційна	4	3	0	0	33	-	-
Реанімація	10	4	0	0	27	-	-
Операційна	17	12	1	2	45	-	-

Інтенсивна терапія	16	6	1	0	23	-	-
Психологічної допомоги	7	2	0	0	25	-	-
Опікове	12	2	0	0	47	-	-
Терапія	13	7	0	0	68	-	-
Евакуаційна	1	1	0	0	40	-	-
Увесь шпиталь	54	38	-	-	-	1120	146
Надходження 60 поранених/добу							
Приймальне відділення	1	1	0	0	33	-	-
Передопераційна	4	3	0	0	45	-	-
Реанімація	10	4	0	0	74	-	-
Операційна	17	12	1	2	75	-	-
Інтенсивна терапія	16	6	1	0	74	-	-
Психологічної допомоги	7	2	0	0	38	-	-
Опікове	12	2	0	0	72	-	-
Терапія	13	7	0	0	67	-	-
Евакуаційна	1	1	0	0	60	-	-
Увесь шпиталь	54	38	-	-	-	976	162
Надходження 80 поранених/добу							
Приймальне відділення	1	1	0	0	44	-	-
Передопераційна	4	3	0	0	76	-	-
Реанімація	10	4	0	0	75	-	-
Операційна	17	12	2	12	78	-	-
Інтенсивна терапія	16	6	1	0	70	-	-
Психологічної допомоги	7	2	0	0	50	-	-
Опікове	12	2	0	0	78	-	-
Терапія	13	7	0	0	72	-	-
Евакуаційна	1	1	0	0	75	-	-
Увесь шпиталь	54	38	-	-	-	828	246

Спостерігається зменшення довжини черги та часу очікування в операційному відділенні. При цьому, внаслідок ліквідації «вузьких місць» у підрозділах, збільшилась пропускна здатність шпиталю.

Таким чином, імітаційна модель мобільного медичного шпиталю дозволяє оцінити його можливості по об'єму робіт, що очікуються, визначити потреби у ресурсах і засобах та підвищити ефективність роботи під час надання медичної допомоги пораненим.

Висновки. На територіях України, де тривають бойові дії, медичні установи стикаються з додатковим навантаженням. Це призводить до нестачі ресурсів та зниження ефективності медичної допомоги. За таких умов необхідно впроваджувати допоміжні мобільні медичні точки для своєчасного надання медичної допомоги. Тому в роботі було розроблено і реалізовано комп'ютерну модель мобільного медичного шпиталю. З комп'ютерною моделлю були проведені дослідження для різної інтенсивності вхідного потоку поранених. Дослідження роботи шпиталю виконано протягом одного робочого тижня. Часові характеристики процесів задаються в одиницях модельного часу – у хвилинах.

Наукова новизна роботи полягає в тому, що отримано нову імітаційну модель, придатну для оцінювання ефективності шпиталю за множиною показників: пропускна здатність, час очікування в чергах, довжини черг у відділення, час перебування поранених у шпиталі.

Використання імітаційної моделі на практиці дозволить визначати потрібну кількість ресурсів, готуватися до можливих навантажень, ефективно використовувати мобільні шпиталі за потребами та приймати обґрунтовані рішення у режимі реального часу.

Подальші дослідження системи шпиталю будуть спрямовані на:

- проведення експериментів типу «що буде, якщо?...»;
- автоматичний розрахунок рекомендованої кількості ліжокмісць та персоналу у відділеннях шпиталю;
- вирішення багатокритеріальної задачі оптимізації за показниками ефективності.

Список бібліографічного опису

1. Г. Г. Рошін Георгій і О. Mazurenko, «Мобільні медичні формування служби медицини катастроф», Мобільні медичні формування служби медицини катастроф, вип. 11, Чер 2010.
2. П. Б. Волянський, «Мобільні медичні формування як механізм медичного захисту за умов надзвичайних ситуацій природного і техногенного характеру», Інвестиції: практика та досвід, вип. 18, с. 104–106, 2011.
3. «Медична допомога після окупації - як працюють медичні мобільні бригади в Україні - ZN.ua». Дата звернення: 06, Квітень 2024. [Online]. Доступний у: <https://zn.ua/ukr/HEALTH/deokupovani-teritoriji-medichni-mobilni-brigadi.html>
4. «За три місяці роботи понад 190 виїздів: мобільні медичні бригади допомагають пацієнтам у найвіддаленіших куточках області | Дніпропетровська обласна державна адміністрація». Дата звернення: 06, Квітень 2024. [Online]. Доступний у: <https://adm.dp.gov.ua/news/za-tri-misyaci-roboti-ponad-190-viyizdiv-mobilni-medichni-brigadi-dopomagayut-paciyentam-u-najviddalenishih-kutochkah-oblasti>
5. «Мобільний госпіталь від 'Rheinmetall'. Що це таке і чому важливо. - портал новин LB.ua». Дата звернення: 06, Червень 2024. [Online]. Доступний у: https://lb.ua/society/2023/10/02/577438_mobilnyi_gospital_vid.html
6. «В Охтирці відкрили модульну екстрену меддопомогу — Суспільне Суми». Дата звернення: 06, Квітень 2024. [Online]. Доступний у: <https://suspilne.media/sumy/726373-v-ohirci-na-sumsini-vidkrili-modulne-ukritta-ekstrenoi-meddopomogi/>
7. A. V. Verba, T. M. Ostashchenko, I. B. Plis, S. S. Tuz, i I. V. Sobko, «Досвід використання зведених медичних загонів у складі передової госпітальної бази в умовах збройного конфлікту високої інтенсивності», Ukrainian Journal of Military Medicine, вип. 4, вип. 1, с. 13–20, Бер 2023, doi: 10.46847/UJMM.2023.1(4)-013.
8. O. P. Yavorovsky et al., «Система забезпечення безпеки персоналу і пацієнтів у лікарняних закладах: стратегія удосконалення», Ukrainian Journal of Military Medicine, вип. 2, вип. 2, с. 23–33, Лип 2021, doi: 10.46847/UJMM.2021.2(2)-023.
9. Y. O. Zui, «Методичний підхід до кількісного оцінювання спроможностей медичних підрозділів Збройних Сил України», Ukrainian Journal of Military Medicine, вип. 2, вип. 2, с. 14–22, Лип 2021, doi: 10.46847/UJMM.2021.2(2)-014.
10. J. Li i Y. Lu, «A dynamic overflow triage model in an outpatient queuing system», IMA Journal of Management Mathematics, вип. 35, вип. 2, с. 301–320, Квіт 2024, doi: 10.1093/imaman/dpad027.
11. P. Sarla et al., «Study on patients arrival at hospital using queuing model with self-similar characteristics», 2022, с. 060005. doi: 10.1063/5.0081946.
12. T. Grober, O. Grober, i O. Savchenko, «An innovative approach to the study of the model of a medical institution», E3S Web of Conferences, вип. 210, с. 11002, Чер 2020, doi: 10.1051/e3sconf/202021011002.
13. F. Mallor, C. Azcárate, i J. Barado, «Control problems and management policies in health systems: application to intensive care units», Flex Serv Manuf J, вип. 28, вип. 1–2, с. 62–89, Чер 2016, doi: 10.1007/s10696-014-9209-8.
14. E. Lavrov, O. Siryk, I. Kirichenko, N. Barchenko, i Y. Chybiriak, «The Methodology of Managed Functional Networks for Organizing Effective and Adaptive Human-Machine Dialogue in Automated Systems», в CEUR Workshop Proceedings, 2021.
15. E. Lavrov, O. Siryk, Y. Chybiriak, L. Danilova, V. Nahomyi, i S. Vakal, «A Model for the Organization of Adaptive Dialogue Interaction 'Man-Computer' Taking into Account the Requirements of Reliability and Efficiency», в 2021 IEEE 4th International Conference on Advanced Information and Communication Technologies, AICT 2021 - Proceedings, 2021. doi: 10.1109/AICT52120.2021.9628939.
16. E. Lavrov, Y. Chybiriak, O. Siryk, V. Logvinenko, i A. Zakharova, «Training of Specialists for Adaptive management. Techniques for Teaching Computer Analysis of Automated Production Systems in the FlexSim Environment», в CEUR Workshop Proceedings, 2022.
17. E. A. Lavrov, O. E. Siryk, Y. I. Chybiriak, A. L. Zolkin, i N. A. Sedova, «Human-centered management in polyergatic information systems. Multi-criteria distribution of functions between operators», в IOP Conference Series: Earth and Environmental Science, 2022. doi: 10.1088/1755-1315/1049/1/012020.
18. E. Lavrov et al., «Human Factors in Automated Systems. Methodology for the Computer System 'International Center for the Support of Ergonomic Design and Research'», в 5th IEEE International Conference on Advanced Information and Communication Technologies, AICT 2023 - Proceedings, 2023. doi: 10.1109/AICT61584.2023.10452692.
19. E. A. Lavrov, V. G. Logvinenko, V. V. Osadchyi, O. Ye. Siryk, i Y. I. Chybiriak, «Adaptive learning system based on cognitive independence», в CEUR Workshop Proceedings, 2023.
- Я. Чибіряк, І. Баранова, і К. Ніколаєнко, «Метод наскрізного навчання студентів ІТ-спеціальностей імітаційному моделюванню у середовищі FlexSim для пошуку резервів підвищення ефективності автоматизованих систем.», Computer-Integrated Technologies: Education, Science, Production, вип. 42, с. 119–129, Бер 2021, doi: 10.36910/6775-2524-0560-2021-42-17.

References

1. G. G. Roschin Georgiy and O. Mazurenko, "Mobile medical formations of the disaster medicine service", *Mobile medical formations of the disaster medicine service*, vol. 11, June 2010.
2. P. B. Volyanskiy, "Mobile medical formations as a mechanism of medical protection under conditions of emergency situations of a natural and man-made nature", *Investments: practice and experience*, vol. 18, p. 104–106, 2011.
3. "Medical care after the occupation - how mobile medical teams work in Ukraine - ZN.ua". Access date: June 6, 2024. [Online]. Available at: <https://zn.ua/ukr/HEALTH/deokupovani-teritoriji-medichni-mobilni-brihadi.html>
4. "Over 190 visits in three months of work: mobile medical teams help patients in the most remote corners of the region | Dnipropetrovsk Regional State Administration". Access date: June 6, 2024. [Online]. Available at: <https://adm.dp.gov.ua/news/za-tri-misyaci-roboti-ponad-190-viyizdiv-mobilni-medichni-brigadi-dopomagayut-paciyentam-u-najviddalenishih-kutochkah-oblasti>
5. "Mobile hospital from 'Rheinmetall'. What is it and why is it important? - news portal LB.ua". Access date: June 6, 2024. [Online]. Available at: https://lb.ua/society/2023/10/02/577438_mobilniy_gospital_vid.html
6. "A modular emergency medical center was opened in Okhtyrka - Sumy Community". Access date: June 6, 2024. [Online]. Available at: <https://suspilne.media/sumy/726373-v-ohirci-na-sumsini-vidkrili-modulne-ukritta-ekstrenoi-meddopomogi/>
7. V. Verba, T. M. Ostashchenko, I. B. Plis, S. S. Tuz, and I. V. Sobko, "Experience of using consolidated medical units as part of a forward hospital base in the conditions of a high-intensity armed conflict", *Ukrainian Journal of Military Medicine*, vol. 4, issue 1, p. 13–20, May 2023, doi: 10.46847/UJMM.2023.1(4)-013.
8. O. P. Yavorovsky et al., "The system for ensuring the safety of personnel and patients in hospitals: a strategy for improvement", *Ukrainian Journal of Military Medicine*, vol. 2, issue 2, p. 23–33, July 2021, doi: 10.46847/UJMM.2021.2(2)-023.
9. Y. O. Zui, "Methodical approach to quantitative assessment of the capabilities of medical units of the Armed Forces of Ukraine", *Ukrainian Journal of Military Medicine*, vol. 2, issue 2, p. 14–22, July 2021, doi: 10.46847/UJMM.2021.2(2)-014.
10. J. Li and Y. Lu, "A dynamic overflow triage model in an outpatient queuing system," *IMA Journal of Management Mathematics*, vol. 35, issue 2, p. 301–320, Apr 2024, doi: 10.1093/imaman/dpad027.
11. P. Sarla et al., "Study on patients arrival at hospital using queuing model with self-similar characteristics", 2022, p. 060005. doi: 10.1063/5.0081946.
12. T. Grober, O. Grober, and O. Savchenko, "An innovative approach to the study of the model of a medical institution", *E3S Web of Conferences*, vol. 210, p. 11002, Jun 2020, doi: 10.1051/e3sconf/202021011002.
13. F. Mallor, C. Azcárate, and J. Barado, "Control problems and management policies in health systems: application to intensive care units," *Flex Serv Manuf J*, vol. 28, issue 1–2, p. 62–89, Jun 2016, doi: 10.1007/s10696-014-9209-8.
14. E. Lavrov, O. Siryk, I. Kirichenko, N. Barchenko, and Y. Chybiriak, "The Methodology of Managed Functional Networks for Organizing Effective and Adaptive Human-Machine Dialogue in Automated Systems," in *CEUR Workshop Proceedings*, 2021.
15. E. Lavrov, O. Siryk, Y. Chybiriak, L. Danilova, V. Nahorny, and S. Vakal, "A Model for the Organization of Adaptive Dialogue Interaction 'Man-Computer' Taking into Account the Requirements of Reliability and Efficiency", in *2021 IEEE 4th International Conference on Advanced Information and Communication Technologies, AICT 2021 - Proceedings*, 2021. doi: 10.1109/AICT52120.2021.9628939.
16. E. Lavrov, Y. Chybiriak, O. Siryk, V. Logvinenko, and A. Zakharova, "Training of Specialists for Adaptive management. Techniques for Teaching Computer Analysis of Automated Production Systems in the FlexSim Environment", in *CEUR Workshop Proceedings*, 2022.
17. E. A. Lavrov, O. E. Siryk, Y. I. Chybiriak, A. L. Zolkin, and N. A. Sedova, "Human-centered management in polyergatic information systems. Multi-criteria distribution of functions between operators", in *IOP Conference Series: Earth and Environmental Science*, 2022. doi: 10.1088/1755-1315/1049/1/012020.
18. E. Lavrov et al., "Human Factors in Automated Systems. Methodology for the Computer System 'International Center for the Support of Ergonomic Design and Research'", in *5th IEEE International Conference on Advanced Information and Communication Technologies, AICT 2023 - Proceedings*, 2023. doi: 10.1109/AICT61584.2023.10452692.
19. E. A. Lavrov, V. G. Logvinenko, V. V. Osadchyi, O. Ye. Siryk, and Y. I. Chybiriak, "Adaptive learning system based on cognitive independence," in *CEUR Workshop Proceedings*, 2023.
20. Ya. Chibyryak, I. Baranova, and K. Nikolayenko, "A method of end-to-end training of students of IT specialties in simulation modeling in the FlexSim environment for finding reserves for increasing the efficiency of automated systems.", *Computer-Integrated Technologies: Education, Science, Production*, vol. 42, p. 119–129, May 2021, doi: 10.36910/6775-2524-0560-2021-42-17.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-30>

УДК: 004.4

Шваліковський Дмитро Миколайович, інженер

<https://orcid.org/0009-0009-8860-0329>

Волинський національний університет імені Лесі Українки, м. Луцьк, Україна

ДВІ ЗАДАЧІ ТЕПЛОПРОВІДНОСТІ У ДВОКОМПОНЕНТНИХ СИСТЕМАХ

Шваліковський Д.М. Дві задачі теплопровідності у двокомпонентних системах. Здійснено числовий розрахунок розподілу температурного поля у одновимірній системі із двох компонент з різними тепловими характеристиками. Розрахунок провадився способом розв'язання дискретизованого рівняння теплопровідності за методом скінченних різниць. Використано прямий двокроковий метод для вирішення відповідної задачі Коші у ДРЧП другого порядку. Розглянуто випадок охолодження такої системи із початкового стану з заданим температурним розподілом, та випадок нагрівання системи за наявності зовнішнього джерела тепла.

Ключові слова: рівняння теплопровідності, числовий розрахунок, задача Діріхле, задача Робіна, CAS Maxima.

Shvalikovskiy D. Two problems of heat equation in two-component systems. We have investigated a numerical calculation of the temperature field distribution in a one-dimensional system of two components with different thermal characteristic. The calculation involved solving the discretized heat conduction equation utilizing the finite difference method. We use a direct two-step method to solve the corresponding Cauchy problem in the second-order PDE. The case of system cooling from the initial state with a given temperature distribution, and the case of heating the system in the presence of an external heat source are considered.

Keywords: heat equation, numerical calculation, Dirichlet problem, Robin problem, CAS Maxima.

Вступні зауваги.

Загальне рівняння, котре описує процеси теплопровідності і дифузії, має вигляд

$$\frac{\partial u(\vec{r}, t)}{\partial t} = \text{div}(a(\vec{r}) \cdot \text{grad } u(\vec{r}, t)) + F(\vec{r}, t).$$

Для задачі теплопровідності $u(\vec{r}, t)$ – значення температури у даній точці тіла в заданий момент часу, $a(\vec{r})$ – величина, що має назву «тепературопровідність» (thermal diffusivity). В такому випадку $a = k/(C\rho)$, де C – питома теплоємність, k – коефіцієнт теплопровідності, ρ – густина тіла. Загалом a може залежати від точки простору (за умови неоднорідності структури матеріалу), однак фактично для практичних задач можна вважати її сталою для встановленого матеріалу. Ми розглянемо одновимірну задачу температурного розподілу та поширення тепла у тонкому стрижні довжини L , тому початкове рівняння спрощується та зводиться до вигляду [1]

$$u_t = a u_{xx}.$$

Для його однозначного вирішення необхідне задання початкових та крайових умов. Зазвичай розглядають два способи задання:

1) початковий розподіл температури та закон її зміни на краях відрізка (задача Діріхле) [2]:

$$u(x, 0) = A(x); \quad u(0, t) = B_1(t); \quad u(L, t) = B_2(t);$$

2) початковий розподіл температури, один кінець міняє свою температуру за певним законом, через інший кінець відбувається теплообмін із зовнішнім джерелом тепла (задача Робіна):

$$u(x, 0) = A(x); \quad u(0, t) = B(t); \quad u_x(L, t) = C(t).$$

При числовому розв'язанні цього рівняння будемо користуватись методом скінченних різниць, тому для його розгляду необхідно провести процедуру дискретизації [3]. За цим методом просторово-часові проміжки розгляду розбиваються на дискретну множину вузловими точками. Розіб'ємо довжину стрижня L на m частин, а проміжок часу S на n частин. Тоді вузлових точок буде відповідно $m + 1$ та $n + 1$, а довжина кожного відрізка $h = L/m$ та $s = S/n$. Нехай функція $u(x, t)$ приймає значення у відповідних вузлах $u(x_i, t_k)$, тоді вона в цих точках замінюється елементами дискретного набору $u_{i,k}$, а частинні похідні основного рівняння замінюються на різницеві відношення

$$u_t(x_i, t_k) = \frac{1}{s}(u_{i,k+1} - u_{i,k}); \quad u_{xx}(x_i, t_k) = \frac{1}{h^2}(u_{i+1,k} - 2u_{i,k} + u_{i-1,k}).$$

Як видно, при підстановці цих співвідношень у основне рівняння можливо виразити значення функції у наступний момент часу $k + 1$ через її значення у попередній момент k в різних просторових вузлових точках. Ця обставина дозволяє застосувати прямий метод розрахунку відповідної задачі Коші, задавши початкові величини в граничних вузлах.

Постановка проблеми

Нехай маємо стрижень загальної довжини L , що складається з двох частин L_1 та L_2 з різними характеристиками. Це може бути система з двох спаяних металічних провідників, або деталей з різнорідних матеріалів, які щільно прилягають одна до одної. Кожна з частин має свої властивості, які загалом у нашій моделі можна звести до трьох величин: довжина (L_1 та L_2), температуропровідність (a_1 та a_2), коефіцієнт переносу тепла (c_1 та c_2).

Розіб'ємо першу частину на m_1 , а другу частину на m_2 проміжків, загальна їх кількість буде $m = m_1 + m_2$. Тоді довжина кожного проміжку $h_1 = L_1/m_1$, $h_2 = L_2/m_2$. Дискретизована просторова система першої частини нумерується послідовністю $0, 1, \dots, m_1$, другої частини послідовністю $m_1, m_1 + 1, \dots, m$. Точка m_1 вважається спільною для обох відрізків, вона міститься у місці спайки двох матеріалів. Часові проміжки розіб'ємо на відрізки $s = S/n$, тоді дискретизована часова система нумерується послідовністю $0, 1, \dots, n$.

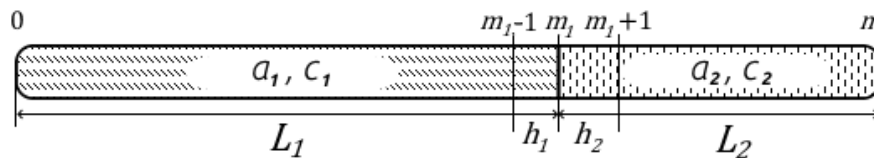


Рис. 1. Стрижень із двох різнорідних матеріалів.

Рівняння теплопровідності, записане для внутрішніх вузлів, приймає вигляд

$$\frac{1}{s}(u_{i,k+1} - u_{i,k}) = \frac{a}{h^2}(u_{i+1,k} - 2u_{i,k} + u_{i-1,k}).$$

Виразивши звідси величину $u_{i,k+1}$, маємо вираз для послідовного обрахунку температурної функції:

$$u_{i,k+1} = \frac{as}{h^2}(u_{i+1,k} - 2u_{i,k} + u_{i-1,k}) - u_{i,k}.$$

Отримана рівність повинна доповнюватись відповідними значеннями функції у початковий момент часу та умовами на краях системи. Ці додаткові умови накладаються на граничні вузли, і для кожної конкретної задачі вони будуть мати свій вигляд.

Множник біля дужок повинен задовольняти умову $\frac{as}{h^2} < 0.5$ для забезпечення збіжності розрахунку, в іншому випадку обраховані значення u швидко йдуть до нескінченності, тому за виконанням цієї умови потрібно уважно слідкувати при розбитті на вузлові точки.

Виклад основного матеріалу й обґрунтування отриманих результатів

Ми будемо провадити розрахунки у відкритій системі комп'ютерної алгебри **Maxima**, оскільки вона має зручний інтерактивний інтерфейс та дозволяє отримувати негайний відгук і відразу спостерігати обчислений результат [4].

Задача 1. Розрахувати еволюцію температури на стрижні довжиною $L = 1$ м, що складається з двох частин довжинами $L_1 = 0.5$ м та $L_2 = 0.5$ м, температуропровідностями $a_1 = 1$, $a_2 = 0.5$, на якому заданий початковий температурний розподіл $A(x) = 1000x(L - x)$. Кінці стрижня мають фіксовану температуру $B_1(t) = 0, B_2(t) = 0$.

Це стандартна задача Діріхле для ДРЧП, тобто задача із закріпленими кінцями. Сталість температури на кінцях стрижня може забезпечуватись поміщенням їх у теплові резервуари із нескінченною теплопровідністю та нескінченною теплоємністю, що звісно ж є модельною абстракцією. Ці резервуари мають свою власну температуру, котра змінюється у часі за законами $B_1(t)$ та $B_2(t)$. При вирішенні задачі Діріхле на дискретизовану функцію накладаються наступні початкові та крайові умови:

$$u_{i,0} = A_i = A(x_i); \quad u_{0,k} = B_{1;k} = B_1(t_k); \quad u_{m,k} = B_{2;k} = B_2(t_k).$$

Врахуємо, що швидкість зміни температури у початковий момент часу є нульовою, звідки

$$\left. \frac{\partial u}{\partial t} \right|_{t=0} = \frac{1}{s}(u_{i,1} - u_{i,0}) = 0; \quad u_{i,1} = u_{i,0}.$$

Різномірність структури стрижня забезпечимо заданням кускових функцій $a(x)$ та $h(x)$, які будуть приймати сталі значення на різних просторових відрізках L_1 та L_2 .

Переведемо задачу на мову **Maxima**, обравши часовий відрізок розгляду $S = 0.3$ с. Кількість точок розбиття виберемо $n = 3000, m_1 = 35, m_2 = 35$.

Вносимо початкові дані у систему.

```
(%i6) L:1$ L1:0.5$ L2:L-L1$ S:0.3$ a1:1$ a2:0.5$
```

```
(%i14) n:3000$ m1:35$ m2:35$ h1:L1/m1$ h2:L2/m2$ s:S/n$ m:m1+m2$ A0:1000$
```

Записуємо кускові функції для $a(x)$ та $h(x)$.

```
(%i15) a(x):= if x<=L1 then a1 else a2 $
```

```
(%i16) h(x):= if x<=L1 then h1 else h2 $
```

Перевіряємо умову стійкості.

```
(%i18) a1*s/h1^2, numer; a2*s/h2^2, numer;
```

```
(%o17) 0.49 (%o18) 0.245
```

На обох частинах умова виконується. Як видно, ми могли б обрати більшу кількість вузлових точок на правій частині (що пов'язано з меншим значення коефіцієнта a_2), однак пізніше це б призвело до нерівномірності графічного представлення початкового розподілу.

Тепер записуємо вирази для початкових та крайових умов. Застосуємо до функцій команду `float`, щоб їх значення одразу перетворювались на числа з плаваючою точкоюю.

```
(%i21) A(x):=float(A0*x*(L-x))$ B1(t):=float(0)$ B2(t):=float(0)$
```

Створюємо масиви X та T , які заповнюємо просторовими та часовими вузловими точками.

```
(%i22) for i:1 thru m1 do
(arraymake(X, [i]), X[0]:0,
X[i]:X[i-1]+h1);
```

```
(%o22) done
```

```
(%i23) for i:m1+1 thru m do
X[i]:X[i-1]+h2;
```

```
(%o23) done
```

```
(%i24) for k:1 thru n do
(arraymake(T, [k]), T[0]:0,
T[k]:T[k-1]+s);
```

```
(%o24) done
```

Тепер у шуканий масив u (він буде двоіндексним, перший індекс i відповідає за просторову змінну, другий k за часову) вносимо початкові та крайові умови.

```
(%i25) for i:0 thru m do
(u[i,0]:A(X[i]), u[i,1]:A(X[i]));
```

```
(%o25) done
```

```
(%i26) for k:0 thru n do
(u[0,k]:B1(T[k]), u[m,k]:B2(T[k]));
```

```
(%o26) done
```

Записуємо вираз для базового рівняння, створивши цикл для його обрахування.

```
(%i27) for k:1 thru n do
for i:1 thru m-1 do
(u[i,k+1]:((a(X[i]))^2*s/(h(X[i]))^2)*(u[i+1,k]-2*u[i,k]+u[i-1,k])+u[i,k]));
```

```
(%o27) done
```

Розрахунок розподілу температури здійснений. Він знаходиться у вигляді масиву значень чисел $u[i, k]$, котрі залежать від координат дискретної просторово-часової сітки. Ці координати є цілими числами, тому, щоб побудувати температурну поверхню, запровадимо функцію, яка залежить від округлених значень x і t до найближчого цілого, та побудуємо її тривимірний графік (Рис. 2).

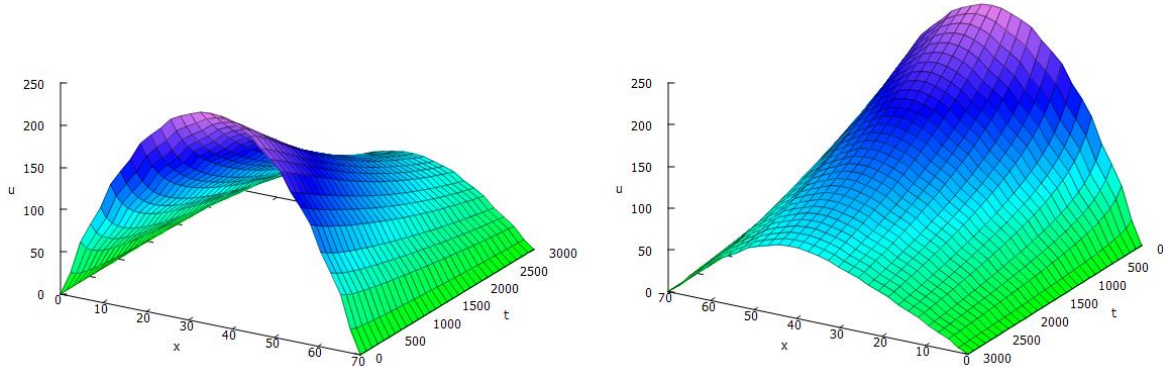
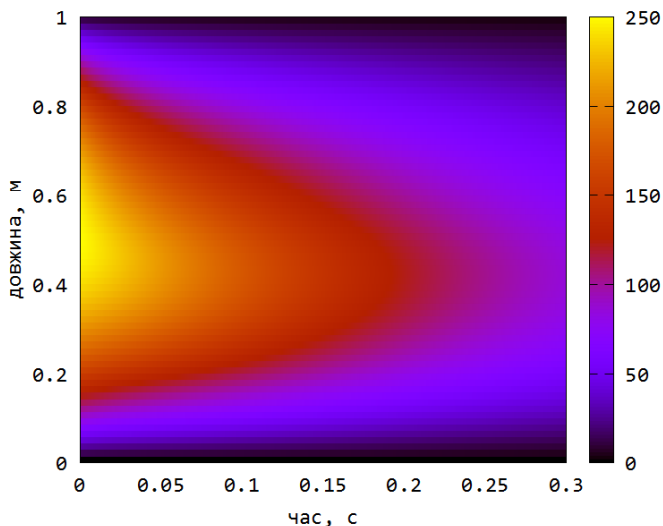


Рис. 2. Обрахована температурно-часова поверхня двокомпонентного стрижня.

```
(%i28) H(x,t):=float(u[round(x),round(t)])$
(%i29) plot3d(H(x,t),[x,0,m],[t,0,n]);
(%t29) (Graphics)
```

Побудований графік можна оглядати, обертаючи мишкою. Із поверхні чітко видно, що часова еволюція розв'язку викривлює на пізніших етапах початковий симетричний розподіл. Більший температурний «горб» буде знаходитись на правій частині поверхні, оскільки друга компонента стрижня має меншу температуропровідність, а отже повільніше втрачає тепло.

Для більшої наочності скористаємось можливістю розфарбовування елементів матриці згідно з її числовим значенням у комірці. Таке представлення здійснюється командою `image` із пакету `draw`. Для цього створимо нульову матрицю розмірності $m \times n$ та заповнюємо її елементами масиву $u[i, k]$. Компоненти матриці запишемо таким чином, щоб їх значення узгоджувались із геометрією задачі.



```
(%i30) load(draw)$
(%i31) U:zeromatrix(m,n)$
(%i32) for i:1 thru m do
for k:1 thru n do
U[i,k]:u[i,k];
(%o32) done
(%i33) wxdraw2d(image(U,0,0,0.3,1),
font = "Consolas",
font_size = 18,
xlabel = "час, с",
ylabel = "довжина, м");
(%t33) (Graphics)
```

Рис. 3. Температурно-часова діаграма двокомпонентного стрижня.

Бачимо, що розподіл температури з часом зміщується вправо, і тут знову ж таки проявляється залежність від величини a .

Наостанок, зробимо два зауваження. По-перше, реальні значення коефіцієнта теплопровідності a становлять $\sim 10^{-5}$ м²/с. Обрані нами модельні величини 1 та 0.5 є надзвичайно великими, і це пояснює, чому температура так швидко спадає фактично до нуля за 0.3 с. По-друге, запропоновану схему розв'язку можна легко масштабувати на більшу кількість частин стрижня: для цього потрібно лише модифікувати кількість та величину вузлових точок у масиві X та запровадити нові кускові функції для виразів $a(x)$ та $h(x)$, які і будуть описувати ці нові додаткові складові.

Задача 2. Розрахувати еволюцію температури на стрижні довжиною $L = 1$ м, що складається з двох частин довжинами $L_1 = 0.5$ м та $L_2 = 0.5$ м, теплопровідностями $a_1 = 1$, $a_2 = 0.7$, коефіцієнтами теплового переносу $c_1 = 0.95$, $c_2 = 0.85$, та на якому заданий початковий температурний розподіл виду $A(x) = \sin \pi x$. Кінці стрижня мають фіксовану температуру $B_1(t) = 0$, $B_2(t) = 0$. У місці з'єднання на стрижень діє джерело тепла з потужністю $g(t) = 8t$.

Це різновид задачі Робіна для ДРЧП, тобто задачі із фіксованою умовою на функцію на одному краю відрізка, та заданим значенням похідної на другому краю. Загалом для широкого класу явищ потік тепла між двома з'єднаними тілами при різних температурах є прямо пропорційним до різниці цих температур (закон Фур'є). Нехай до кінця стрижня довжини L приєднане зовнішнє джерело з температурою, яка міняється за законом $g(t)$. Тоді можна записати

$$u_x(x, t)|_{x=L} = c(g(t) - u(L, t))$$

де u_x – зміна температури на кінці стрижня, c – коефіцієнт теплового переносу (визначає частку енергії, що йде від нагрівника на нагрівання). Запишемо це рівняння в дискретизованому вигляді:

$$\frac{1}{h}(u_{m,k} - u_{m-1,k}) = c(g_k - u_{m,k}),$$

звідки можна визначити температуру на кінці стрижня

$$u_{m,k} = \frac{u_{m-1,k} + chg_k}{1 + ch}.$$

У нашому випадку «кінцем» для обох компонент буде точка з'єднання m_1 . Інші початкові та крайові умови будуть аналогічними до попередньої задачі:

$$u_{i,0} = A_i = A(x_i); \quad u_{0,k} = B_{1,k} = B_1(t_k); \quad u_{m,k} = B_{2,k} = B_2(t_k).$$

Ми будемо вирішувати поставлене завдання як дві задачі Робіна для двох компонент стрижня, вважаючи початком першої частини точку 0, її кінцем точку m_1 , а початком другої частини останню точку m , її кінцем точку m_1 , тобто для другої компоненти розрахунок буде вестись у зворотному просторовому порядку. Це необхідна обставина, оскільки цикл повинен розпочинатись із заданого значення функції. Умова для крайової точки m_1 буде дописуватись до основного рівняння в обох випадках як окремий вираз, що пов'язує її з попередньою.

Запрограмуємо задачу на мові **Maxima**, обравши часовий відрізок розгляду $S = 0.3$ с. Кількість точок розбиття виберемо $n = 3000$, $m_1 = 35$, $m_2 = 35$.

Вносимо початкові характеристики стрижня.

```
(%i6) L:1$ L1:0.5$ L2:L-L1$ S:0.3$ a1:1$ a2:0.7$
```

```
(%i16) n:3000$ m1:35$ m2:35$ h1:L1/m1$ h2:L2/m2$ s:S/n$ m:m1+m2$ c1:0.95$ c2:0.85$
```

Тепер запроваджуємо кускові функції, які описують неоднорідність системи.

```
(%i19) a(x):= if x<=L1 then a1 else a2$
```

```
h(x):= if x<=L1 then h1 else h2$
```

```
c(x):= if x<=L1 then c1 else c2$
```

Перевіряємо умову стійкості.

```
(%i21) a1*s/h1^2, numer; a2*s/h2^2, numer;
```

```
(%o20) 0.49
```

```
(%o21) 0.343
```

Записуємо функціональні залежності початкових, крайових умов, функцію джерела.

```
(%i25) A(x):=float(sin(%pi*x))$ B1(t):=float(0)$ B2(t):=float(0)$ g(t):=float(8*t);
```

Створюємо масиви вузлових точок та заповнюємо їх.

```
(%i26) for i:1 thru m1 do
(arraymake(X, [i]), X[0]:0,
X[i]:X[i-1]+h1);
```

```
(%o26) done
```

```
(%i27) for i:m1+1 thru m do
(X[m1]:L1,
X[i]:X[i-1]+h2);
```

```
(%o27) done
```

```
(%i28) for k:1 thru n do
(arraymake(T, [k]), T[0]:0,
T[k]:T[k-1]+s);
```

```
(%o28) done
```

Вносимо у шуканий масив u початкові та крайові значення.

```
(%i29) for i:0 thru m do
(u[i,1]:A(X[i]), u[i,0]:A(X[i]));
```

```
(%o29) done
```

```
(%i30) for k:0 thru n do
(u[0,k]:B1(T[k]), u[m,k]:B2(T[k]));
```

```
(%o30) done
```

Записуємо цикл обрахунку основного рівняння для точок від 1 до $m_1 - 1$, та окреме рівняння для точки m_1 . Таким чином буде знайдена температурна поверхня лівої компоненти.

```
(%i31) for k:1 thru n do
for i:1 thru m1-1 do
(u[i,k+1]:(a(X[i])^2*s/h(X[i])^2)*(u[i+1,k]-2*u[i,k]+u[i-1,k])+u[i,k],
u[m1,k]:(u[m1-1,k]+c(X[i])*h(X[i])*g(T[k]))/(c(X[i])*h(X[i])+1));
```

```
(%o31) done
```

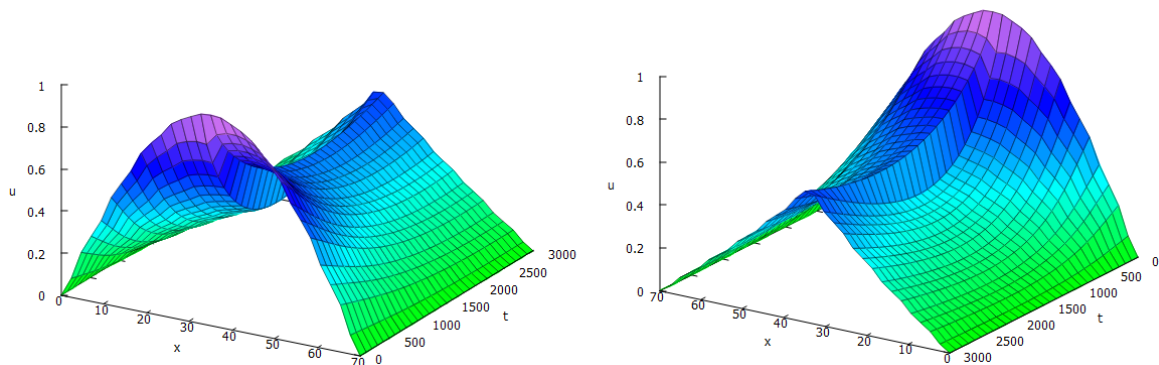


Рис. 4. Обрахована температурно-часова поверхня двокомпонентного стрижня за наявності джерела тепла.

Записуємо аналогічний цикл обрахунку у зворотному порядку базового рівняння для точок від m до $m_1 + 1$, та окреме рівняння для точки m_1 , знайшовши температурну поверхню правої компоненти.

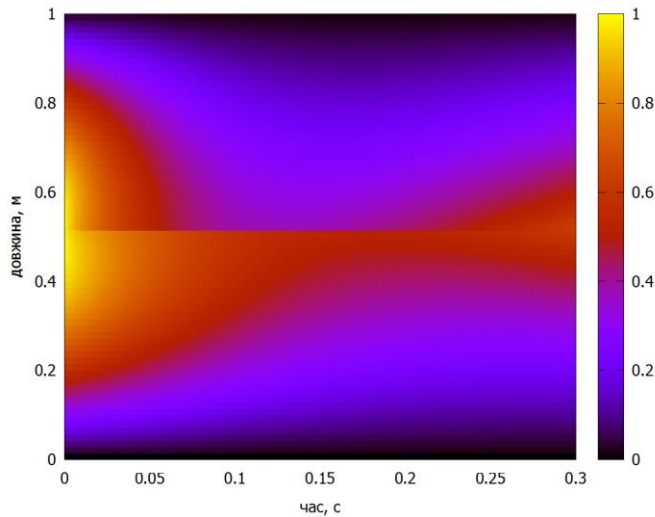
```
(%i32) for k:1 thru n do
for i:m-1 thru m1+1 step -1 do
(u[i,k+1]:(a(X[i])^2*s/h(X[i])^2)*(u[i+1,k]-2*u[i,k]+u[i-1,k])+u[i,k],
u[m1,k]:(u[m1+1,k]+c(X[i])*h(X[i])*g(T[k]))/(c(X[i])*h(X[i])+1));
```

```
(%o32) done
```

Обчислення завершені. Запроваджуємо функцію $H(x, t)$ округленням змінних, графік якої буде температурною поверхнею (Рис. 4).


```
(%i33) H(x,t):=float(u[round(x),round(t)])$
(%i34) wxplot3d(H(x,t),[x,0,m],[t,0,n]);
(%t34) (Graphics)
```

Тепер скористаємось побудовою кольорової матриці для отримання температурно-часової діаграми (Рис. 5).

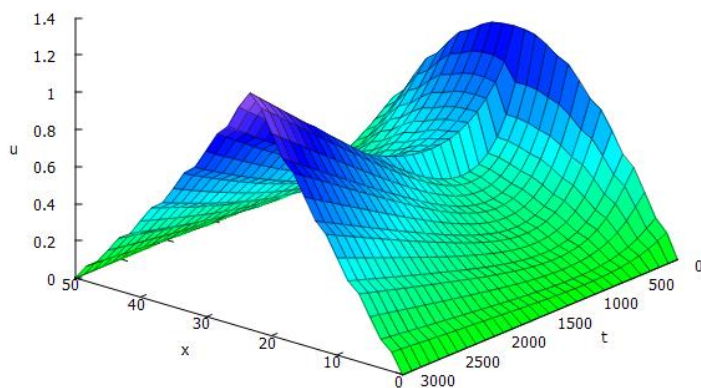


```
(%i35) load(draw)$
(%i36) U:zeromatrix(m,n)$
(%i37) for i:1 thru m do
for k:1 thru n do
U[i,k]:u[i,k];
(%o37) done
(%i38) wxdraw2d(image(U,0,0,0.3,1),
font = "Consolas",
font_size = 18,
xlabel = "час, с",
ylabel = "довжина, м");
(%t38) (Graphics)
```

Рис. 5. Температурно-часова діаграма двокомпонентного стрижня з джерелом тепла.

Обговоримо одержані результати. Бачимо, що присутність джерела тепла створює розрив між двома функціями температурної поверхні. У цьому характерна відмінність від першої задачі: там зберігає свою неперервність як функція, так і її похідна; тоді як у другій задачі присутній розрив і функції, і її похідної (відмінність теплових потоків спричиняється різницею у коефіцієнтах теплопереносу c_1 та c_2). На початку часового проміжку температури обох компонент зменшуються внаслідок охолодження до крайових значень, причому ліва компонента швидше, оскільки вона має більшу температуропровідність. Але внаслідок постійного теплового притоку обидві поверхні підвищуються і утворюють трикутний пік.

Хоча б здавалося, що після точки з'єднання дві поверхні повинні знову розійтись через різницю у коефіцієнтах a та c , насправді цей пік буде завжди присутній. Різниця між температурними поверхнями, яка обумовлена температуропровідністю та коефіцієнтом теплопереносу, буде набагато меншою за градієнт температур, створений безперервним притоком тепла від джерела (потужність котрого весь час збільшується).



Для прикладу, обчислена поверхня задачі у випадку розгляду часового проміжку 0.6 с зображена на Рис. 6. Поведінка системи повністю відповідає загальній теоремі: для задачі Робіна при достатньо великому часі спостереження температурна функція вже не залежить від початкового розподілу і є лінійною функцією зі змінним кутовим коефіцієнтом [5]:

$$u(x,t) = \frac{g(t) - T_1}{L} x.$$

Рис. 6. Обрахована поверхня для часу 0.6 с.

На рисунку в кінцевій частині справді ми бачимо майже трикутну форму профіля, де температурні функції є фактично прямими відрізками із ледь помітними кривинами.

Виникає природне питання про масштабованість запропонованої схеми. Розглянемо стрижень, який складається вже з трьох компонент: ліва, середня і права. Нехай у місцях з'єднань діють теплові джерела з різними потужностями. Для лівої та правої частин виникає задача Робіна, у якій відома функція на одному краї та її похідна на іншому краї в кожен момент часу (а це крайова задача третього роду), тому ми можемо застосувати запропоновану схему та обрахувати відповідну поверхню. Для середньої частини будуть відомі лише теплові потоки на краях, тобто похідні температурної функції (крайова задача другого роду, або задача Неймана). Способом її вирішення є лише розв'язання системи з m рівнянь, де m – кількість просторових вузлів сітки, для прикладу, методом Кранка-Ніколсона [6]. Причому для спостереження часової еволюції таких систем має бути n – кількість часових вузлів. У нашому випадку це мало б бути 3000 систем із 35 рівнянь. Такі завдання вимагають вже вищих обчислювальних потужностей.

Висновки та перспективи подальшого дослідження. У роботі представлений алгоритм аналізу та числового розрахунку температурного розподілу для двокомпонентної системи; розв'язані задачі охолодження такої системи та її нагрівання за наявності джерела тепла. Запропоновану схему можна вільно розширити на подібні випадки для більшого числа компонент, або для складових частин із неоднорідних матеріалів (зі змінними тепловими характеристиками). Подальшим природним узагальненням є перехід до двовимірного середовища: розгляд 2d-плити із кількох частин або зі змінною температуропровідністю та за наявності нагрівальних елементів.

У дослідженні ми використовували вільне програмне забезпечення CAS Maxima, яка є досить потужним та розробленим застосунком для розв'язання широкого кола математичних та фізичних задач моделювання.

Список бібліографічного опису

1. Nakhle H. Asmar. Partial Differential Equations with Fourier Series and Boundary Value Problems: Third Edition. – Dover Publications, 2016. – 816 p.
2. S.R.K. Iyengar, R.K. Jain. Numerical Methods. – New Age International Limited Publishers, 2009. – 326 p.
3. W.Y. Yang, W. Cao, T.S. Chung, J. Morris. Applied Numerical Methods using MATLAB. – A John Wiley & Sons, Inc., 2005. – 512 p.
4. Шваліковський Д. М. Моделювання процесів та систем у середовищі CAS Maxima. – Луцьк: ВНУ імені Лесі Українки, 2024. – 252 с.
5. Ryan C. Daileda. Partial Differential Equations: Lecture Handouts & Slides. The 1-D Heat Equation, Part 2. – Trinity University, 2023. – 24 p.
6. A. Kharab, R.B. Guenther. An Introduction to Numerical Methods. – Taylor & Francis Group, 2019. – 615 p.

References

1. Nakhle H. Asmar. Partial Differential Equations with Fourier Series and Boundary Value Problems: Third Edition. – Dover Publications, 2016. – 816 p.
2. S.R.K. Iyengar, R.K. Jain. Numerical Methods. – New Age International Limited Publishers, 2009. – 326 p.
3. W.Y. Yang, W. Cao, T.S. Chung, J. Morris. Applied Numerical Methods using MATLAB. – A John Wiley & Sons, Inc., 2005. – 512 p.
4. D.M. Shvalikovskiy. Simulation of Processes and Systems in the CAS Maxima Environment. – Lutsk: Lesya Ukrainka Volyn National University, 2024. – 252 p.
5. Ryan C. Daileda. Partial Differential Equations: Lecture Handouts & Slides. The 1-D Heat Equation, Part 2. – Trinity University, 2023. – 24 p.
6. A. Kharab, R.B. Guenther. An Introduction to Numerical Methods. – Taylor & Francis Group, 2019. – 615 p.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-31>

УДК 519.854.2+004.021

Шенгелія Володимир Олександрович, магістр

<https://orcid.org/0009-0001-6356-962X>

Жданова Олена Григорівна, к.т.н., доцент

<https://orcid.org/0000-0002-8787-846X>

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

МЕТОДИ СКЛАДАННЯ МАРШРУТІВ ПОЛЬОТУ ДЛЯ БПЛА З ПЕРЕСУВНИМ ЦЕНТРОМ ОБСЛУГОВУВАННЯ

Шенгелія В.О., Жданова О.Г. Методи складання маршрутів польоту для БПЛА з пересувним центром обслуговування. У статті розглядаються методи складання маршрутів польоту для безпілотних літальних апаратів (БПЛА) з пересувним центром обслуговування (ПЦО). Досліджувана задача передбачає наявність одного БПЛА та одного ПЦО, який має множину місць дислокації. Метою роботи є підвищення якості маршрутів польоту. Для розв'язання задачі розроблено жадібний алгоритм, алгоритм штучної бджолоїної колонії та табу-пошук. Запропонований бджолоїний алгоритм налічує три способи організації роботи бджіл-розвідників і п'ять способів організації роботи бджіл-фуражирів, комбінації яких утворюють 15 різноманітних алгоритмів. Для алгоритму табу-пошуку розроблено чотири способи роботи локальної оптимізації. Експерименти показали, що найкращий із алгоритмів табу-пошуку працював значно швидше за найкращий з бджолоїних алгоритмів, але в середньому трохи поступився в якості побудованих маршрутів.

Ключові слова: безпілотні літальні апарати (БПЛА), пересувний центр обслуговування (ПЦО), теорія маршрутизації, комбінаторна оптимізація, жадібний алгоритм, алгоритм штучної бджолоїної колонії, алгоритм табу-пошуку.

Shenheliia V., Zhdanova O. Path planning for UAVs with a moving maintenance center. The article examines methods for path planning of unmanned aerial vehicles (UAVs) with a mobile maintenance center (MMC). The task under study involves a single UAV and a single MMC, which has multiple deployment locations. The aim of the work is to improve the quality of flight paths. To solve the problem, a greedy algorithm, an artificial bee colony algorithm, and a tabu search algorithm were developed. The proposed bee algorithm includes three methods for organizing the work of scout bees and five methods for organizing the work of forager bees, with combinations forming 15 different algorithms. For the tabu search algorithm, four methods of local optimization were developed. Experiments showed that the best of the tabu search algorithms operated significantly faster than the best of the bee algorithms, but on average, it slightly lagged in the quality of the constructed flight paths.

Keywords: unmanned aerial vehicles (UAV), mobile maintenance center (MMC), combinatorial optimization, routing theory, greedy algorithm, artificial bee colony algorithm, tabu search algorithm.

Постановка проблеми. Безпілотні літальні апарати (БПЛА) стрімко починають займати важливе місце в різних сферах діяльності людини, від військових до цивільних застосувань. Вже активно використовуються рятувальними службами для доставки медикаментів та їжі у віддалені та важкодоступні місця. Військові застосовують БПЛА під час ведення розвідки для виявлення та атаки ворожих цілей, що допомагає проводити операції без прямої участі пілотів. Протягом останніх років виникло багато інтересу до їх використання у комерційних потребах, це трапилось завдяки розвитку технологій і зменшенню витрат на виробництво даних апаратів. Великі логістичні компанії, такі як Amazon, United Parcel Service і DHL, шукають альтернативні підходи зменшення витрат на останній кілометр доставки замовлень клієнтам. Amazon в 2021 році оголосив, що вже почав розробку власних БПЛА й всієї необхідної під них інфраструктури для вирішення цієї проблеми. Однак існують численні виклики, пов'язані з використанням БПЛА. Основним обмеженням є тривалість польоту, цю проблему можна вирішити розробивши різні типи акумуляторів, використовуючи двигуни внутрішнього згорання або гібридні системи. Проте, більш перспективним рішенням проблеми визнали пересувні центри обслуговування (ПЦО), які здатні перезаряджати чи замінювати акумулятори, зберігати БПЛА та взаємодіяти з ними [1]. ПЦО є інноваційним концептом, який передбачає використання мобільної платформи (наприклад, автомобіля, корабля або наземної станції), яка служить як точка запуску та посадки для БПЛА. Застосування ПЦО дозволяє вирішити проблему тривалості польоту та рухатися далі у напрямку автоматизації систем, через це основною областю досліджень стало створення енергоефективних маршрутів. Таким чином, розробка ефективних методів складання маршрутів польоту БПЛА з ПЦО може стати ключовим елементом у розширенні можливостей використання цих апаратів в різних галузях.

Аналіз останніх досліджень і публікацій. Наявна велика кількість типів задач маршрутизації БПЛА, їх класифікують в залежності від цільової функції, обмежень та змінних

факторів. Більшість задач складання маршрутів польоту для БПЛА ставить за мету оптимізувати пройдено відстань та витрачений час. Обмеження ресурсу заряду акумулятору, наявність заборонених для польоту зон або перешкод, вантажопідйомність є найбільш поширеними. Змінні фактори представляють собою кількість БПЛА та їх характеристики, наявність динамічного середовища або ПЦО, а також багато інших. Ці особливості грають значну роль під час створення математичної моделі та вибору методів розв'язання поставленої задачі.

Методи складання маршрутів польоту для БПЛА поділяються на традиційні та сучасні інтелектуальні. Традиційні включають методи найшвидшого спуску, оптимального керування та кореляції похідних. Вони мають деякі суттєві недоліки, такі як погана адаптивність та низька стабільність у відтворенні отриманих результатів, з легкістю можуть зацикловатись в локальній оптимізації, що не допустимо для маршрутизації БПЛА [2]. Серед сучасних інтелектуальних алгоритмів можна зустріти еволюційні, метаевристики основані на ройовому інтелекті, методи локального пошуку та багато інших [3]. Дана група методів вже довела свою ефективність та використовується в більшості робіт, пов'язаних зі складанням маршрутів польоту для БПЛА. Далі окремо розглянуто переваги та недоліки найпопулярніших з них.

Генетичні алгоритми. Незважаючи на переваги, такі як глобальна оптимізація, можливість паралельних обчислень та гнучкість (застосування до широкого діапазону задач), вони мають деякі недоліки, а саме низьку швидкість збіжності, проблеми з масштабованістю (збільшення розмірності може значно знизити якість), потребують великої кількості обчислювальних ресурсів [4]. Алгоритм все ж таки є ефективним в розв'язанні задач маршрутизації БПЛА. Проте, складність визначення принципів роботи операторів, часто призводять до того, що розроблений алгоритм є неефективним [5] та потребує модифікацій чи повної переробки [6].

Метод рою частинок. Завдяки своїй ефективності з часом став набирати популярність для розв'язання задач маршрутизації. Не зважаючи на просту структуру і малу кількість параметрів, даний метод не дуже підходить для розв'язання дискретних задач, адже через наявну схему роботи виникає потреба у введенні форм «рух» і «швидкість», доволі часто це відбувається неефективно [7]. Проте, особливої уваги вартує основний недолік, від якого страждає якість його побудови, а саме часте впадання в локальні оптимуми [8].

Мурашиний алгоритм. Даний метод оптимізації є більш популярним та ефективним за якістю складання маршрутів польоту БПЛА ніж попередньо розглянуті. Має низьку вагомих відомих переваг, а саме наявний глобальний пошук, децентралізовану обробку та підвищення якості знайдених розв'язків з часом [9]. Проте, алгоритм страждає від повільної збіжності, що призводить до його низької швидкості роботи [10].

Бджолиний алгоритм. Ефективний для задач складання маршрутів польоту БПЛА, адже передбачає локальне вдосконалення знайдених розв'язків, простий у реалізації, з гнучкою параметризацією як кількості бджіл-розвідників та бджіл-фуражирів, так і умови завершення роботи. Під час пошуку може вимагати великої кількості обчислень через наявну постійну локальну оптимізацію [11]. В порівнянні з мурашиним алгоритмом показує значно кращі результати як в якості отриманих маршрутів, так і за часом виконання [12].

Методи локального пошуку. Група методів може бути майже такою ж ефективною як ройові, проте кращою за часом роботи, перевірка цієї гіпотези заслуговує на увагу. Серед наявних алгоритмів найбільшої уваги заслуговує алгоритм табу-пошуку, адже надає можливість уникнення локальних мінімумів, здатний працювати зі складними обмеженнями, простий в реалізації та доволі ефективний на практиці.

Для подальшої розробки та аналізу ефективності серед ройових обрано алгоритм штучної бджолиної колонії, адже попередньо він виявився кращим за якістю отриманих маршрутів. Для оцінки ефективності побудови методів локального пошуку обрано алгоритм табу-пошуку.

Постановка досліджуваної задачі. Нехай маємо один БПЛА, множину об'єктів розвідки та ПЦО, що також має множину (більше трьох) місць дислокації. Ресурс польоту БПЛА є обмеженим, приземлившись на місці дислокації ПЦО, він проходить обслуговування (заміну акумулятору або дозаправку паливом), після чого повністю відновлює ресурс польоту. Об'єкти розвідки мають різну вагу (пріоритет для відвідання), при цьому кожен з них можна відвідати лише один раз протягом виконання усього маршруту польоту.

БПЛА починає рух з першого місця дислокації і облітає певні об'єкти розвідки, до закінчення ресурсу польоту має здійснити посадку на друге місце дислокації, де проходить обслуговування. Послідовність дій, описана попередньо, аналогічна для наступних підмаршрутів. Завдання полягає

у визначенні підмаршрутів, що дозволять обстежити максимальну зважену кількість об'єктів розвідки під час даного вильоту.

Дано:

$J = \{ 1, 2, \dots, n \}$ — множина об'єктів розвідки, кожен має вагу $1 \leq w_j \leq 10, j = 1, \dots, n$;

$\{ A_1, A_2, \dots, A_m \}$ — множина місць дислокації ПЦО, де $3 < m \leq 100$;

t — ресурс польоту БПЛА в часовому вимірі (без обслуговування);

τ — час відновлення ресурсу польоту (обслуговування);

v — швидкість польоту БПЛА.

Проміжні дані:

c_{pq} ($p, q \in \{ 1, 2, \dots, n, n+1, n+2, \dots, n+m \}$) — матриця часів перельоту.

Значення під індексами $n+1, n+2, \dots, n+m$ визначають час перельоту до місця дислокації ПЦО.

Необхідно знайти такі підмаршрути обльоту об'єктів розвідки:

$$J_1 = \{ A_1, j_1^1, j_2^1, \dots, j_{n_1}^1, A_2 \},$$

$$\dots$$

$$J_{m-1} = \{ A_{m-1}, j_1^{m-1}, j_2^{m-1}, \dots, j_{n_{m-1}}^{m-1}, A_m \},$$

де n_1, \dots, n_{m-1} вказують на кількість обстежених об'єктів розвідки кожного з підмаршрутів, ці величини також відносяться до шуканих.

До знайдених підмаршрутів накладаються наступні обмеження:

$$\{ j_1^1, j_2^1, \dots, j_{n_1}^1 \} \cup \dots \cup \{ j_1^{m-1}, j_2^{m-1}, \dots, j_{n_{m-1}}^{m-1} \} \subseteq J, \quad (1.1)$$

$$\{ j_1^p, j_2^p, \dots, j_{n_p}^p \} \cap \{ j_1^q, j_2^q, \dots, j_{n_q}^q \} = \emptyset, \quad p, q = \overline{1, m-1}, p \neq q. \quad (1.2)$$

Формула 1.1 вказує, що усі відвідані об'єкти розвідки включені або є еквівалентними заданій множині об'єктів розвідки. Формула 1.2 вводить обмеження, що кожен об'єкт розвідки може бути відвіданий лише раз. Сумарний час польоту кожного підмаршруту не має перевищувати ресурс польоту БПЛА у часовому вимірі:

$$C_1 = c_{A_1 j_1^1} + \sum_{q=1}^{n_1-1} c_{j_q^1 j_{q+1}^1} + c_{j_{n_1}^1 A_2} \leq t,$$

$$\dots$$

$$C_{m-1} = c_{A_{m-1} j_1^{m-1}} + \sum_{q=1}^{n_{m-1}-1} c_{j_q^{m-1} j_{q+1}^{m-1}} + c_{j_{n_{m-1}}^{m-1} A_m} \leq t.$$

Цільовою функцією є максимізація сукупної ваги обстежених об'єктів розвідки:

$$\sum_{i=1}^{m-1} \sum_{l=j_1^i}^{j_{n_i}^i} w_l \rightarrow \max.$$

Розроблені алгоритми. Вводиться умова використання двовимірного простору для спрощення процесу розробки алгоритмів. Пошук розв'язків кожним з них відбувається послідовно за підмаршрутами. Очікується, що під час ініціалізації задач наявна матриця, що містить відомості про час переміщення та співвідношення ваги об'єкту розвідки до часу перельоту для всіх наявних

точок. Жадібний алгоритм введено як референтну точку відліку для оцінки якості результатів, які були отримані за допомогою розроблених алгоритмів.

Жадібний алгоритм (Greedy). Натхнений класичним алгоритмом найближчого сусіда, що зазвичай найчастіше обирається людиною для пошуку розв'язків подібних задач. У базовій постановці на кожному кроці вибору підлягає найближча за відстанню точка, проте ця модифікація передбачає пошук найбільшого співвідношення ваги об'єкта розвідки до відстані. Пошук відбувається без урахування відвіданих об'єктів розвідки, проте враховує наявність ресурсу польоту для приземлення на кінцевий ПЦО.

Алгоритм має описаний нижче вигляд.

ЦИКЛ ЗА ПІДМАРШРУТАМИ:

1. ВКЛЮЧИТИ початковий ПЦО в підмаршрут.
2. ВИКОНУВАТИ (формування підмаршруту):
 - 2.1. ВИЗНАЧЕННЯ найкращого об'єкта розвідки для візиту з поточної точки.
 - 2.2. ЯКЩО наступний для відвідання об'єкт розвідки не знайдено, ТО:
 - 2.2.1. ВКЛЮЧИТИ кінцевий ПЦО в підмаршрут.
 - 2.2.2. ВИЙТИ З ЦИКЛУ.

ІНАКШЕ:

- 2.2.1. ВКЛЮЧИТИ знайдений об'єкт розвідки в підмаршрут.
- 2.2.2. ПОЗНАЧИТИ знайдений об'єкт розвідки як відвіданий.
- 2.2.3. ПРИЗНАЧИТИ знайдений об'єкт розвідки поточним.
- 2.2.4. ПЕРЕРАХУВАТИ залишок ресурсу польоту.

Бджолиний алгоритм (ABC). Побудований на основі алгоритму штучної бджолоїної колонії, який є натхненим поведінкою рою бджіл в процесі пошуку нектару. Бджоли-розвідники вилітають з вулика у різних випадкових напрямках та шукають місця знаходження нектару, формуючи початкові розв'язки. З часом повертаються у вулик, передаючи інформацію про розташування нектару. Бджоли-фуражири летять у місця, де його очікується найбільше, та проводять процедуру пошуку в околах знаходження нектару. Даний процес повторюється доти, доки не виконається певна кількість ітерацій без покращення отриманого розв'язку.

Алгоритм має схему, подану нижче.

ЦИКЛ за підмаршрутами:

ВИКОНУВАТИ:

1. ЦИКЛ за кількістю бджіл-розвідників:
 - 1.1. ВИПУСТИТИ бджолу-розвідника.
2. ОБРАТИ найкращих бджіл-розвідників.
3. ЦИКЛ за найкращими бджолами-розвідниками:
 - 3.1. ПОКИ результат покращується:
 - 3.1.1. ЦИКЛ за бджолами-фуражирами:
 - 3.1.1.1. ВИПУСТИТИ бджолу-фуражира.
 4. ЯКЩО результат покращився, ТО СКИНУТИ В НУЛЬ кількість ітерацій без покращення.

ІНАКШЕ ІНКРЕМЕНТУВАТИ кількість ітерацій без покращення.

ПОКИ кількість ітерацій без покращення < максимальної кількості ітерацій без покращення.

ЗБЕРЕГТИ отриманий підмаршрут.

ВИКЛЮЧИТИ відвідані об'єкти розвідки з огляду.

Запропоновано три способи організації роботи бджіл-розвідників. Розглянемо їх детально.

Випадковий (Random — R). Під час побудови підмаршруту визначається список об'єктів розвідки, які можна відвідати з поточного місця дислокації ПЦО. Випадковим чином генерується кількість тих, які має відвідати БПЛА в процесі побудови, об'єкти розвідки довільно обираються зі створеного раніше списку. Після чого підмаршрут перевіряється на допустимість, в разі успішності перевірки додається до вже знайдених.

Ймовірнісний (Probable — P). На кожному кроці пошуку об'єктів розвідки створюється список тих, що можна відвідати з поточної точки. Будується одинична пряма з нанесеними відрізками співвідношень ваги до часу перельоту для обраних об'єктів розвідки. Генерується раціональне число від нуля до одиниці, що вказує на точку, яка належить відріzkу певного об'єкта розвідки. Таким чином, ймовірність вибору об'єкта розвідки прямо пропорційна вазі та обернено пропорційна часу перельоту для кожного з об'єктів розвідки.

Ймовірнісний в околі (Probable in the neighborhood — PN). Додатково враховується, що шуканий об'єкт розвідки має знаходитись в околі радіусом одна четверта від максимального ресурсу польоту БПЛА.

Локальна оптимізація початкових підмаршрутів є ключовою у знаходженні розв'язків. Запропоновано п'ять способів організації її роботи, кожен з яких розглянуто нижче.

Додати найближчий об'єкт розвідки (Add nearest — AN). Випадково обирається один з об'єктів розвідки підмаршруту, в залежності від ефективності перебудованого підмаршруту найближчий об'єкт розвідки до даного додається для відвідання перед або після нього.

Перебудувати підмаршрут ймовірно (Rebuild probable — RP). З підмаршруту випадковим чином обирається об'єкт розвідки, після чого наступні за ним об'єкти розвідки виключаються, а сам підмаршрут добудовується за способом організації роботи ймовірнісних бджіл-розвідників.

Перебудувати підмаршрут ймовірно в околі (Rebuild probable in the neighborhood — RPN). Подібний до попереднього, але при перебудові підмаршруту точки обираються з околу.

Перебудувати підмаршрут ймовірно та додати найближчий об'єкт розвідки (Rebuild probable and add nearest — RPAAN). Представляє собою комбінацію способів організації роботи RP та AN, спочатку підмаршрут підлягає перебудові, а потім покращується додаванням найближчого об'єкта розвідки.

Перебудувати підмаршрут ймовірно в околі та додати найближчий об'єкт розвідки (Rebuild probable in the neighborhood and add nearest — RPNAAN). Аналогічний попередньому, але перебудова підмаршрутів відбувається з урахування пошуку об'єктів розвідки в околі.

Для подальшого дослідження розроблено три способи організації роботи бджіл-розвідників і п'ять способів організації роботи бджіл-фуражирів, комбінації яких утворюють 15 різноманітних алгоритмів.

Алгоритм табу-пошуку (Tabu). Перш за все, відбувається пошук допустимого розв'язку. За допомогою локальних модифікацій поточний підмаршрут ітеративно змінюється, намагаючись перейти до кращого розв'язку. При цьому отриманий розв'язок може погіршуватися, що дозволяє вийти з «пастки» локального мінімуму. Усі розв'язки, отримані в процесі пошуку, додаються до списку заборонених (табу), отже вони не будуть оцінюватися як потенційні кандидати для подальшого відвідування.

Нижче наведено схему алгоритму.

ЦИКЛ за підмаршрутами:

1. ФОРМУВАННЯ допустимого розв'язку.
2. ІНІЦІАЛІЗУВАТИ табу-списку.
3. ПРИЗНАЧИТИ поточний розв'язок рекордним.
4. ВИКОНУВАТИ (локальний пошук):
 - 4.1. ПОШУК допустимих розв'язків в околі.
 - 4.2. ЯКЩО один із знайдених розв'язків краще рекордного, ТО:
 - 4.2.1. ПРИЗНАЧИТИ даний розв'язок рекордним.
 - 4.2.2. ОБНУЛИТИ кількість ітерацій без покращення.
ІНАКШЕ ІНКРЕМЕНТУВАТИ кількість ітерацій без покращення.
 - 4.3. ВКЛЮЧИТИ знайдені розв'язки в табу-список.
ПОКИ кількість ітерацій без покращення < максимальної кількості ітерацій без покращення.
5. ЗБЕРЕГТИ отриманий підмаршрут.
6. ВИКЛЮЧИТИ відвідані об'єкти розвідки з огляду.

У розробленій системі алгоритмів для ініціалізації початкового допустимого розв'язку використано жадібний алгоритм. Для локальної оптимізації застосовано всі способи організації роботи бджіл-фуражирів, окрім AN, оскільки постійні спроби додавання найближчого об'єкта розвідки до єдиного підмаршруту не дозволяють отримати бажану якість оптимізації початкового розв'язку.

Експериментальне дослідження. Експеримент спрямовано на визначення найкращого з розроблених алгоритмів для досліджуваної задачі. Кожним із запропонованих алгоритмів розв'язується 100 індивідуальних задач у 10 прогонах, що дозволяє оперувати даними 1000 розв'язаних під час порівняння результатів роботи. Даний процес має місце для кожної з розглянутих розмірностей.

Кількість місць дислокації ПЦО (розмірність) варіюється від 4 до 10. Кількість об'єктів розвідки в задачі визначається як розмірність помножена на коефіцієнт в діапазоні від 6 до 15. Отже,

для кожної з розглянутих розмірностей буде наявно 10 груп задач, які мають по 10 індивідуальних та відрізняються за кількістю об'єктів розвідки. Наприклад, для розмірності 10 кількість об'єктів розвідки індивідуальних задач буде варіюватись від 60 до 150 одиниць.

Експеримент проводився за умови, що жодна інша програма не запущена на комп'ютері з процесором Intel Core i7-1185G7 з частотою 3.00 – 4.80 ГГц, а також наявною оперативною пам'яттю DDR4 на 32 ГБ з постійною частотою 3200 МГц.

Таблиця 1 відображає результати роботи алгоритмів для розмірності 10. Зображені там характеристики є ключовими для визначення ефективності розроблених алгоритмів.

Таблиця 1. Результати експерименту для розмірності 10

Алгоритм	Відхилення значення ЦФ від ЦФ жадібного (%)			Час роботи (мс.)		
	Найкр.	Серед.	Найгір.	Макс.	Серед.	Мін.
Greedy	0	0	0	8.75	2.91	1.75
ABC-R-AN	77.71	40.83	-33.67	747.40	536.01	378.09
ABC-P-AN	79.18	48.86	-31.31	779.53	570.45	471.25
ABC-PN-AN	81.72	49.64	-27.88	731.15	509.68	488.90
ABC-R-RP	75.47	44.45	-25.59	1331.06	1082.33	731.00
ABC-P-RP	77.57	50.75	-24.40	1539.47	1119.88	940.62
ABC-PN-RP	78.63	51.67	-23.41	2206.13	1065.68	971.76
ABC-R-RPN	72.20	47.12	-26.37	1506.19	1032.25	624.71
ABC-P-RPN	75.17	51.35	-24.17	1497.63	1066.04	845.61
ABC-PN-RPN	82.33	51.84	-23.69	1670.43	1122.77	876.43
ABC-R-RPAAAN	86.14	51.56	-14.97	1684.92	1416.36	952.03
ABC-P-RPAAAN	89.10	52.34	-13.42	2014.57	1408.14	1182.65
ABC-PN-RPAAAN	91.47	52.99	-12.63	2051.16	1575.52	1224.42
ABC-R-RPNAAAN	87.07	52.84	-13.14	1440.79	1255.88	892.45
ABC-P-RPNAAAN	92.09	53.83	-12.63	1595.07	1385.08	1092.57
ABC-PN-RPNAAAN	95.27	54.36	-12.04	1910.80	1308.39	1106.50
Tabu-RP	71.78	40.77	-16.38	551.98	250.38	113.76
Tabu-RPN	77.72	41.16	-15.96	514.23	233.64	104.44
Tabu-RPAAAN	84.54	47.22	-14.25	648.15	272.92	113.45
Tabu-RPNAAAN	80.88	46.88	-14.85	647.62	253.93	104.24

Метрикою якості маршрутів польоту, що використовується в даному експерименті, є середнє відхилення значення цільової функції (ЦФ) алгоритмів від ЦФ жадібного. Результати експерименту для всіх розмірностей проілюстровано на одному графіку, який зображено на рисунку 1.

З графіку, зображеного на рисунку 1, добре видно, що зі збільшенням кількості місць дислокації ПЦО значення середнього відхилення ЦФ розроблених алгоритмів від ЦФ жадібного зростає. Отже, зі збільшенням розмірності розроблені алгоритми стають більш ефективними в порівнянні з жадібним. Значення зростають без суттєвих коливань, що свідчить про сталість запропонованих алгоритмів.

Результати роботи бджолиних алгоритмів вказують на те, що використання способу організації роботи бджіл-розвідників R призводить до найгірших результатів, спосіб P в середньому поступається PN на 1%. Розглядаючи локальну оптимізацію, можна прийти до висновку, що спосіб організації роботи AN виявився найгіршим. Майже однакові значення показали RP та RPN, відхилення варіюється в діапазоні до 1%. RPAAAN та RPNAAAN є кращими за попередні в середньому більш ніж на 2%, а другий ефективніше за перший майже на 1.5%. Жорсткіші обмеження за околону способів організації роботи призводять до кращих результатів.

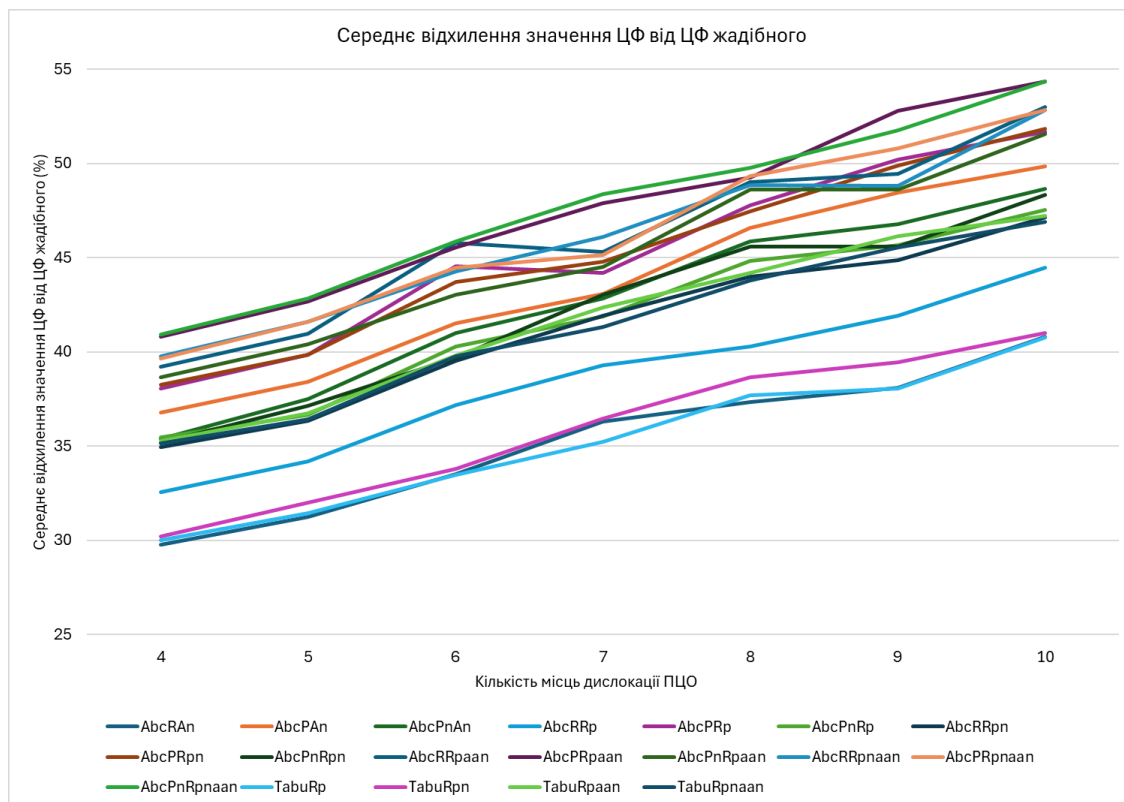


Рис.1. Середнє відхилення значення ЦФ алгоритмів від ЦФ жадібного

Для алгоритмів табу-пошуку попередні твердження, відносно локальної оптимізації, є схожими, проте RPAAN є навпаки кращим за RPNAAN приблизно на 0.5%. Відхилення в значеннях між кращим бджолиним і табу-пошуку складає приблизно 7%, а для розмірності 4 є лише біля 2.5%. Ці результати були очікуваними під час висування гіпотези для даного дослідження, проте для повного її підтвердження треба проаналізувати швидкодію. Графік середнього часу роботи алгоритмів проілюстровано на рисунку 2.

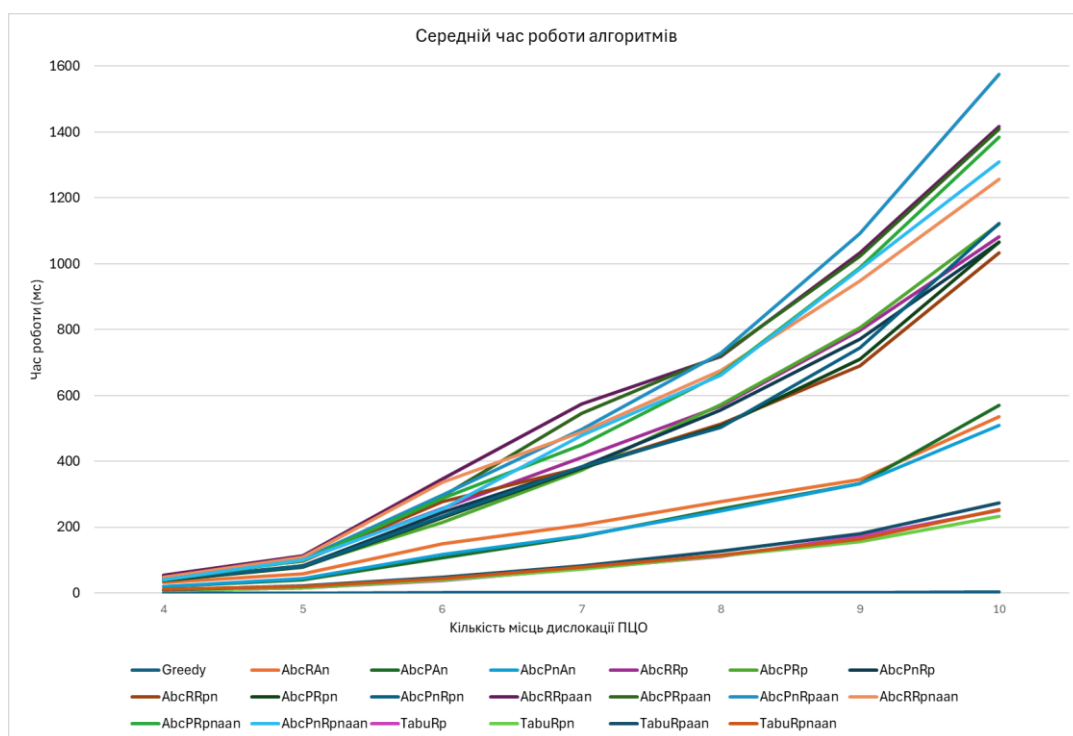


Рис. 2. Середній час роботи алгоритмів

Графік середнього часу роботи алгоритмів, зображений на рисунку 2, показує, що збільшення кількості місць дислокації ПЦО веде до стрімкого, але без суттєвих коливань, зростання часу необхідного бджолиним та алгоритмам табу-пошуку для знаходження маршрутів польоту. Алгоритми табу-пошуку показали себе як найшвидші, Tabu-RPAAAN виявився у майже 5 разів швидше за ABC-PN-RPNAAN, отже поставлену в даному дослідженні гіпотезу було доведено.

Можна зазначити, що використання алгоритму ABC-PN-RPNAAN призводить до найкращий розв'язків у порівнянні з розробленими. Tabu-RPAAAN є значно швидшим за попередній, але відрізняється трохи гіршим середнім відхиленням значення ЦФ від ЦФ жадібного.

Висновки. Розроблено та експериментально досліджено ефективність алгоритмів штучної бджолиної колонії та табу-пошуку. Встановлено, що алгоритм ABC-PN-RPNAAN гарантує найкращий розв'язок серед запропонованих, а Tabu-RPAAAN виявився значно, майже в 5 разів, швидшою альтернативою попередньому, проте відрізняється трохи гіршою, в середньому менш ніж на 4%, якістю побудови маршрутів.

Напрямами покращення якості роботи алгоритмів визначено наступні способи побудови маршрутів польоту: реверсивний (від кінцевого підмаршруту), випадковий (кожну ітерацію підмаршрут обирається випадково), у напрямку зустрічі (почергово з початку та кінця), повний (одразу весь маршрут польоту). Для побудови початкових допустимих розв'язків, а також локальної оптимізації, можна задіяти численну кількість існуючих евристичних правил або розробити їх самостійно.

Наукова новизна отриманих результатів полягає у вдосконаленні методів розв'язання задач складання маршрутів польоту для БПЛА з ПЦО. Практичне значення одержаних результатів зводиться до можливого використання розроблених алгоритмів у маршрутизації БПЛА з ПЦО в умовах реального світу.

Список бібліографічного опису

1. Zear A., Ranga V. (2019). Path planning of unmanned aerial vehicles: Current State and future challenges. *First International Conference on Sustainable Technologies for Computational Intelligence*, 409–419. DOI: 10.1007/978-981-15-0029-9_32.
2. Aggarwal S., Kumar N. (2020). Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges. *Computer Communications*, 149, 270–299. DOI: 10.1016/j.comcom.2019.10.014.
3. Hooshyar M., Huang, Y.-M. (2023). Meta-heuristic algorithms in UAV path planning optimization: A systematic review (2018–2022). *Drones*, 7 (12), 687. DOI: 10.3390/drones7120687.
4. Cakir M. (2015). 2D path planning of UAVs with genetic algorithm in a constrained environment. 2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO). DOI: 10.1109/icmsao.2015.7152235.
5. Roberge V., Tarbouchi M., Labonte G. (2013). Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning. *IEEE Transactions on Industrial Informatics*, 9 (1), 132–141. DOI: 10.1109/tii.2012.2198665.
6. Liu J. (2022). An improved genetic algorithm for Rapid UAV Path Planning. *Journal of Physics: Conference Series*, 2216 (1), 012035. DOI: 10.1088/1742-6596/2216/1/012035.
7. Shen L., Wang Y., Liu K., Yang Z., Shi X., Yang X., Jing K. (2020). Synergistic path planning of multi-UAVs for air pollution detection of ships in ports. *Transportation Research Part E: Logistics and Transportation Review*, 144, 102128. DOI: 10.1016/j.tre.2020.102128.
8. Asma A., Sadok B. (2017). Dynamic distributed PSO joints elites in multiple robot path planning systems: Theoretical and practical review of new ideas. *Procedia Computer Science*, 112, 1082–1091. DOI: 10.1016/j.procs.2017.08.128.
9. Li B., Qi X., Yu B., Liu L. (2020). Trajectory planning for UAV based on improved ACO algorithm. *IEEE Access*, 8, 2995–3006. DOI: 10.1109/access.2019.2962340.
10. Perez-Carabaza S., Besada-Portas E., Lopez-Orozco J. A., de la Cruz J. M. (2018). Ant colony optimization for multi-UAV minimum time search in uncertain domains. *Applied Soft Computing*, 62, 789–806. DOI: 10.1016/j.asoc.2017.09.009.
11. Yu S., Li Y., Ma C. (2021). Research on UAV trajectory planning based on artificial bee colony algorithm. 2021 9th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC). DOI: 10.1109/icwoc52624.2021.9530209.
12. Tian G., Zhang L., Bai X., Wang B. (2018). Real-time dynamic track planning of multi-UAV formation based on improved artificial bee colony algorithm. 2018 37th Chinese Control Conference (CCC). DOI: 10.23919/chicc.2018.8482622.

References

1. Zear A., Ranga V. (2019). Path planning of unmanned aerial vehicles: Current State and future challenges. *First International Conference on Sustainable Technologies for Computational Intelligence*, 409–419. DOI: 10.1007/978-981-15-0029-9_32.

2. Aggarwal S., Kumar N. (2020). Path planning techniques for unmanned aerial vehicles: A review, solutions, and challenges. *Computer Communications*, 149, 270–299. DOI: 10.1016/j.comcom.2019.10.014.
3. Hooshyar M., Huang, Y.-M. (2023). Meta-heuristic algorithms in UAV path planning optimization: A systematic review (2018–2022). *Drones*, 7 (12), 687. DOI: 10.3390/drones7120687.
4. Cakir M. (2015). 2D path planning of UAVs with genetic algorithm in a constrained environment. *2015 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO)*. DOI: 10.1109/icmsao.2015.7152235.
5. Roberge V., Tarbouchi M., Labonte G. (2013). Comparison of parallel genetic algorithm and particle swarm optimization for real-time UAV path planning. *IEEE Transactions on Industrial Informatics*, 9 (1), 132–141. DOI: 10.1109/tii.2012.2198665.
6. Liu J. (2022). An improved genetic algorithm for Rapid UAV Path Planning. *Journal of Physics: Conference Series*, 2216 (1), 012035. DOI: 10.1088/1742-6596/2216/1/012035.
7. Shen L., Wang Y., Liu K., Yang Z., Shi X., Yang X., Jing K. (2020). Synergistic path planning of multi-UAVs for air pollution detection of ships in ports. *Transportation Research Part E: Logistics and Transportation Review*, 144, 102128. DOI: 10.1016/j.tre.2020.102128.
8. Asma A., Sadok B. (2017). Dynamic distributed PSO joints elites in multiple robot path planning systems: Theoretical and practical review of new ideas. *Procedia Computer Science*, 112, 1082–1091. DOI: 10.1016/j.procs.2017.08.128.
9. Li B., Qi X., Yu B., Liu L. (2020). Trajectory planning for UAV based on improved ACO algorithm. *IEEE Access*, 8, 2995–3006. DOI: 10.1109/access.2019.2962340.
10. Perez-Carabaza S., Besada-Portas E., Lopez-Orozco J. A., de la Cruz J. M. (2018). Ant colony optimization for multi-UAV minimum time search in uncertain domains. *Applied Soft Computing*, 62, 789–806. DOI: 10.1016/j.asoc.2017.09.009.
11. Yu S., Li Y., Ma C. (2021). Research on UAV trajectory planning based on artificial bee colony algorithm. *2021 9th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC)*. DOI: 10.1109/icwoc52624.2021.9530209.
12. Tian G., Zhang L., Bai X., Wang B. (2018). Real-time dynamic track planning of multi-UAV formation based on improved artificial bee colony algorithm. *2018 37th Chinese Control Conference (CCC)*. DOI: 10.23919/chicc.2018.8482622.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-32>

УДК 004.75

Якобчук Богдан Анатолійович, здобувач вищої освіти
кафедри комп'ютерної інженерії та кібербезпеки

<https://orcid.org/0009-0002-6979-725X>

Христинець Наталія Анатоліївна, к.т.н., доцент

<https://orcid.org/0000-0002-4836-7632>

Луцький національний технічний університет, м. Луцьк, Україна

СТВОРЕННЯ ПРОСТОГО ГОЛОСОВОГО ПОМІЧНИКА У ДОДАТКУ ДЛЯ ПЛАТФОРМИ ANDROID

Якобчук Б.А., Христинець Н.А. Створення простого голосового помічника у додатку для платформи **Android**. У статті представлено вичерпний посібник зі створення простого голосового помічника для Android-додатків. Голосовий помічник, названий «ExpertAvon», призначений для надання користувачам функціональностей взаємодії з додатком, у який інтегрований голосовий помічник або з іншими додатками чи компонентами смартфона. Процес розробки передбачає інтеграцію технології розпізнавання мовлення та подальшу реакцію на отримані голосові команди. Стаття містить покрокову інструкцію зі створення ExpertAvon, включаючи налаштування інструментів, бібліотек, реалізацію перетворення аудіо в текст та подальшу взаємодію з компонентами додатку.

Ключові слова: голосовий помічник, Android-додаток, розпізнавання мовлення, ExpertAvon, інтеграція голосових команд, взаємодія з додатками, перетворення аудіо в текст, розробка мобільних додатків.

Yakobchuk B., Khrystynets N. **Creating a simple voice assistant in an application for the Android platform.** The article presents a comprehensive guide to creating a simple voice assistant for Android applications. The voice assistant, named «ExpertAvon», is designed to provide users with the functionality of interacting with the application into which the voice assistant is integrated or with other applications or components of the smartphone. The development process involves the integration of speech recognition technology and the subsequent response to received voice commands. The article contains step-by-step instructions for creating ExpertAvon, including setting up tools, libraries, implementing audio-to-text conversion, and further interaction with application components.

Keywords: voice assistant, Android application, speech recognition, ExpertAvon, voice command integration, interaction with applications, audio to text conversion, mobile application development.

Постановка наукової проблеми. Проблемою дослідження є вирішення ряду ключових питань у розробці простого голосового помічника в середовищі Android Studio засобами мови програмування Java. Ці дослідження спрямовані на створення процесу розпізнавання голосових команд, через перетворення аудіо в текст.

Проблему створення оптимального функціоналу перетворення аудіо даних в текстові дані, було вирішено через впровадження у додаток хмарного сервісу розпізнавання мовлення Google Speech Recognition API.

Завдання дослідження полягає у вирішенні технічних та концептуальних аспектів створення голосових помічників, а також ознайомлення із процесом реалізації функції розпізнавання мовлення, шляхом перетворення аудіо в текст.

Загалом, розробка голосового помічника призначена для опрацювання певних завдань і предстала, як зручний спосіб взаємодії з компонентами окремого додатку чи самого пристрою Android без допомоги рук.

Аналіз досліджень. Було оглянуто статті, навчальні матеріали, онлайн-форуми, інтернет-джерела [1-4], відео матеріали [5-6] та з'ясовано, що усі вище розглянуті попередні дослідження акцентують увагу на ключових функціональностях голосових помічників для Android, таких як розпізнавання мовлення, взаємодія з іншими додатками та компонентами смартфона, а також реалізація мовлення в текст.

Проаналізована оцінка ефективності реалізації функцій у різних середовищах та виявлено, що інтеграція голосового помічника з існуючими додатками та сервісами є ключовою для покращення користувацького досвіду та функціональності. При використанні даного підходу проектування була вирішена проблема складності взаємодії користувача з додатком через голосові команди, оскільки основні механіки впроваджені з урахуванням простоти використання. Також, результати дослідження допомогли у виборі середовища та мови програмування, для розробки простого голосового помічника.

У дослідженні [7] розглянуто можливості перетворення аудіо інформації в текст через хмарний сервіс Google Speech Recognition API, що дозволяє на базі цього перетворення створити простий голосовий помічник. Також, ця публікація підкреслює гнучкість та адаптивність голосових помічників до різних потреб користувачів.

Таким чином, проаналізовані джерела підтверджують доцільність використання сучасних технологій розпізнавання мовлення та обробки природної мови для створення простого голосового помічника у додатку для платформи Android. Важливими аспектами при розробці такого помічника є інтеграція з іншими сервісами, налаштування користувацьких команд та забезпечення високої точності розпізнавання мовлення.

Мета роботи полягає у наданні покрокового керівництва зі створення базового голосового помічника для Android-додатків, зосередившись на розробці простого голосового помічника під назвою «ExpertAvon».

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження.

Розробка голосового помічника для платформи Android, зокрема «ExpertAvon», включає кілька ключових етапів, кожен з яких вимагає ретельного підходу та використання сучасних технологій. Процес розробки охоплює інтеграцію голосового розпізнавання, перетворення аудіо в текст і реалізацію відповідних реакцій додатка на отримані голосові команди. Алгоритм роботи голосового помічника можна переглянути на рисунку 1.

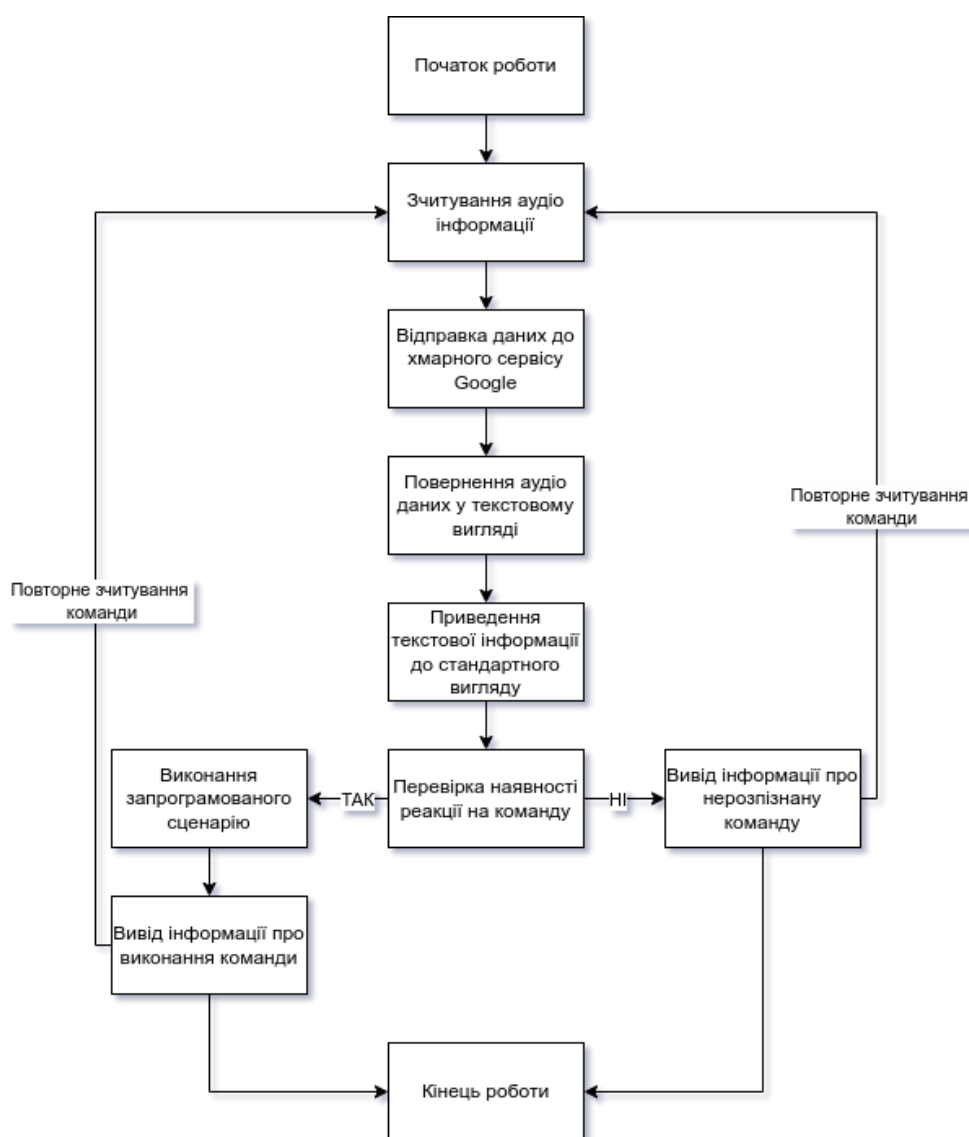


Рис. 1. Алгоритм роботи голосового помічника

На першому етапі відбувається налаштування середовища розробки. Використовується Android Studio та мова програмування Java, що забезпечує високу сумісність з платформою Android. В незалежності від того, чи інтегрується помічник у готовий додаток, чи реалізовується як окрема програма, алгоритм розробки буде однаковим. Для реалізації функціоналу розпізнавання мовлення інтегрується хмарний сервіс Google Speech Recognition API, який надає можливості точного перетворення аудіо в текст.

Наступним кроком є додавання бібліотек, необхідних для роботи з Google Speech Recognition API. Однією з таких бібліотек є бібліотека «speech.RecognizerIntent». Також, було додано декілька інших базових бібліотек. Використання цих бібліотек дозволяє додатку отримувати доступ до мікрофона пристрою, записувати аудіо та надсилати його до сервісу для розпізнавання. Важливим аспектом є налаштування дозволів та обробка можливих виключень, що може виникнути під час роботи з аудіо.

Далі йде етап розпізнавання мовлення та обробки тексту. Код який відповідає за таку обробку представлено на лістингу 1. Аудіо дані, отримані від користувача, передаються на сервер Google для обробки. Після розпізнавання, текстові дані повертаються додатку, де здійснюється їх подальша обробка. Основним завданням на цьому етапі є забезпечення високої точності розпізнавання та мінімізації помилок. Для цього можуть використовуватися різні алгоритми попередньої обробки аудіо, такі як фільтрація шуму та нормалізація гучності, однак для простого голосового помічника вони не потрібні.

Лістинг 1 - Код для налаштування зчитування голосових команд

```
213 private void startVoiceRecognitionActivity() {
214     Intent intent = new Intent(RecognizerIntent.ACTION_RECOGNIZE_SPEECH);
215     intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE, value: "uk-UA");
216     intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE_MODEL, RecognizerIntent.LANGUAGE_MODEL_FREE_FORM);
217     intent.putExtra(RecognizerIntent.EXTRA_PROMPT, (CharSequence) null);
218     startActivityForResult(intent, VOICE_RECOGNITION_REQUEST_CODE);
219 }
```

Кінець лістингу 1

На завершальному етапі реалізується взаємодія з іншими компонентами додатка або смартфона. Отриманий текст використовується для виконання відповідних команд, наприклад, запуск інших додатків, виконання пошукових запитів або управління функціями смартфона. Цей етап вимагає ретельного проектування та тестування, щоб забезпечити коректну роботу всіх функцій. Приклад коду, який відповідає за реакції на голосові команди можна переглянути на лістингу 2. Цей приклад було взято з додатку «Emergency App», у який інтегровано голосовий помічник «ExpertAvon».

Лістинг 2 - Приклад коду, який відповідає за реакцію на команду

```
221 @Override
222 protected void onActivityResult(int requestCode, int resultCode, Intent data) {
223     super.onActivityResult(requestCode, resultCode, data);
224
225     if (requestCode == VOICE_RECOGNITION_REQUEST_CODE && resultCode == RESULT_OK) {
226         ArrayList<String> matches = data.getStringArrayListExtra(RecognizerIntent.EXTRA_RESULTS);
227
228         if (matches.contains("включи телевізор")) {
229             TVImageOn.setVisibility(View.VISIBLE);
230             Toast.makeText(context: this, text: "Телевізор увімкнено", Toast.LENGTH_SHORT).show();
231             mediaPlayerNewsTV.setLooping(true);
232             mediaPlayerNewsTV.start();
233         } else {
234             Toast.makeText(context: this, text: "Невідома команда", Toast.LENGTH_SHORT).show();
235         }
236     }
237 }
238 }
```

Кінець лістингу 2

Розглянемо аспекти створення ефективного голосового помічника. Графічний інтерфейс додатку має бути інтуїтивно зрозумілим та легким у використанні. Визначення основних характеристик і функцій голосового помічника є важливим етапом перед початком розробки.

Голосовий помічник «ExpertAvon» включає в себе дві кнопки: кнопка початку зчитування команд та кнопка інформаційної довідки, яка виводить діалогове вікно з переліком доступних голосових команд. Кнопка початку роботи голосового помічника має код показаний на лістингу 3. Цей код запускає активність голосового помічника при натисненні

Лістинг 3 - Код для кнопки запуску голосового помічника

```
132 microphoneButton.setOnClickListener(new View.OnClickListener() {  
133     @Override  
134     public void onClick(View v) { startVoiceRecognitionActivity(); }  
137 });
```

Кінець лістингу 3

Особливості платформи Android дозволяють реалізувати високоякісну інтеграцію голосових команд з різними додатками та сервісами. Важливими характеристиками такого підходу є можливість адаптації до різних потреб користувачів, що забезпечує більш широкий спектр застосування голосового помічника. Інтеграція голосового помічника з іншими сервісами дозволяє значно покращити користувацький досвід та функціональність додатку.

Таким чином, розробка простого голосового помічника для платформи Android потребує системного підходу та використання сучасних технологій розпізнавання мовлення. Важливими аспектами є інтеграція з хмарними сервісами, налаштування користувацьких команд та забезпечення високої точності розпізнавання мовлення. Результати дослідження підтверджують ефективність запропонованих підходів та демонструють можливості створення зручного та функціонального голосового помічника.

Висновки.

У статті було розглянуто процес розробки простого голосового помічника для Android-додатків, названого «ExpertAvon». Основними етапами розробки стали налаштування середовища розробки, інтеграція технології розпізнавання мовлення, обробка аудіо даних та реалізація відповідних реакцій на голосові команди. Використання Google Speech Recognition API дозволило забезпечити високу точність перетворення аудіо в текст.

Дослідження підтвердило доцільність застосування сучасних технологій розпізнавання мовлення для створення простих голосових помічників. Інтеграція голосового помічника з іншими додатками та сервісами значно покращує користувацький досвід і функціональність додатків. Основними результатами роботи є створення покрокового керівництва з розробки голосового помічника, а також забезпечення його ефективної взаємодії з компонентами додатка та смартфона.

Таким чином, розробка простого голосового помічника для платформи Android, зокрема «ExpertAvon», демонструє можливості створення зручних та функціональних інструментів для взаємодії з мобільними додатками. Результати дослідження можуть бути корисними для розробників, що прагнуть інтегрувати голосові команди у свої Android-додатки.

Список бібліографічного опису

1. Голосові помічники: що це і навіщо вони HR [Ok Google, Alexa, Siri, Cortana]. HURMA. URL: <https://hurma.work/blog/voice-assistants-shho-cze-i-navishho-voni-hr-ok-google-alexa-siri-cortana/> (дата звернення: 16.05.2024).
2. Добірка найзручніших голосових помічників – bit.ua. bit.ua Медіа про життя і технології в ньому. URL: <https://bit.ua/2021/07/dlya-sproshhennya-roboty-dobirka-najzruchnishyh-golosovyh-pomichnykiv/> (дата звернення: 16.05.2024).
3. Building a voice assistant for an android java or kotlin app. Alan docs. Alan AI. URL: <https://alan.app/docs/tutorials/android/integrating-java-kotlin/> (дата звернення: 16.05.2024).
4. How to create virtual assistant apps like siri and google assistant. Cleveroad Inc. - Web and App development company. URL: <https://www.cleveroad.com/blog/how-to-create-virtual-assistant-apps-like-siri-and-google-assistant/> (дата звернення: 17.05.2024).
5. GoXR3Plus Studio. Java speech recognition tutorial (part 1), 2016. YouTube. URL: <https://www.youtube.com/watch?v=R8vsXKFTee0> (дата звернення: 17.05.2024).

6. Infybuzz. Create your own voice assistant in java. Just like google assistant, siri OR amazon alexa, 2021. YouTube. URL: <https://www.youtube.com/watch?v=6Ip42tbq4DM> (дата звернення: 17.05.2024).
7. How to convert speech to text in android? - geeksforgeeks. GeeksforGeeks. URL: <https://www.geeksforgeeks.org/how-to-convert-speech-to-text-in-android/> (дата звернення: 17.05.2024).

References

1. Voice assistants: what they are and why HR needs them [Ok Google, Alexa, Siri, Cortana]. HURMA. URL: <https://hurma.work/blog/voice-assistants-shho-cze-i-navishho-voni-hr-ok-google-alexa-siri-cortana/> (accessed May 16, 2024).
2. A selection of the most convenient voice assistants - bit.ua. bit.ua Media about life and technology in it. URL: <https://bit.ua/2021/07/dlya-sproshhennya-roboty-dobirka-najzruchnishyh-golosovyh-pomichnykiv/> (accessed on May 16, 2024).
3. Building a voice assistant for an android java or kotlin app. Alan docs. Alan AI. URL: <https://alan.app/docs/tutorials/android/integrating-java-kotlin/> (accessed May 16, 2024).
4. How to create virtual assistant apps like siri and google assistant. Cleveroad Inc. is a Web and App development company. URL: <https://www.cleveroad.com/blog/how-to-create-virtual-assistant-apps-like-siri-and-google-assistant/> (accessed on May 17, 2024).
5. GoXR3Plus Studio. Java speech recognition tutorial (part 1), 2016. YouTube. URL: <https://www.youtube.com/watch?v=R8vsXKFTee0> (accessed May 17, 2024).
6. Infybuzz. Create your own voice assistant in java. Just like google assistant, siri OR amazon alexa, 2021. YouTube. URL: <https://www.youtube.com/watch?v=6Ip42tbq4DM> (accessed 17.05.2024).
7. How to convert speech to text in android? - geeksforgeeks. GeeksforGeeks. URL: <https://www.geeksforgeeks.org/how-to-convert-speech-to-text-in-android/> (accessed May 17, 2024).

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-33>

УДК 621.35

Баховський Петро Федорович, к.т.н., доцент

<https://orcid.org/0000-0001-7923-3293>

Євсюк Микола Миколайович, к.т.н., доцент

<https://orcid.org/0000-0002-3768-8959>

Луцький національний технічний університет, м Луцьк, Україна

АНАЛІЗ ЕНЕРГЕТИЧНОГО БАЛАНСУ РАДІОЛІНІЙ ПІД ЧАС ОБМІНУ ДАНИМИ В СИСТЕМАХ SAE / EPS

Баховський П.Ф., Євсюк М.М. Аналіз енергетичного балансу радіоліній під час обміну даними в системах SAE / EPS. В роботі проведено аналіз енергетичного балансу висхідної та низхідної радіоліній у системах мобільного зв'язку під час обміну даними, розраховано межу розміру зони обслуговування для фіксованої швидкості передачі даних. Розрахована межа може відповідати або не відповідати критеріям межі стільника, яка формується таким чином, щоб забезпечити енергетичний баланс на визначеній зоні функціонування систем мобільного зв'язку з якісними параметрами системи SAE/EPS. Мета аналізу низхідної радіолінії полягає у визначенні допустимих втрат під час розповсюдженні радіохвиль із забезпеченням необхідної швидкості передачі даних і підтримці зони обслуговування, обмеженої енергетичним балансом висхідної радіолінії. Результати аналізу енергетичного балансу системи SAE/EPS можна використовувати для випадку коли проєктування виконується з метою забезпечення заданих розміру зони обслуговування й швидкості передачі даних та для планування зон обслуговування групи додаткових каналів радіозв'язку.

Ключові слова: радіолінія, радіозв'язок, енергетичний баланс, радіоканал, зона обслуговування, сектор.

Bakhovskiy P., Yevsyuk M. Analysis of the energy balance of radio lines during data exchange in SAE / EPS systems.

The paper analyzes the energy balance of uplink and downlink radio lines in mobile communication systems during data exchange, calculates the size limit of the service area for a fixed data transfer rate. The calculated limit may or may not meet the criteria of the cell limit, which is formed in such a way as to ensure the energy balance in the defined area of operation of mobile communication systems with quality parameters of the SAE/EPS system. The purpose of the analysis of the downlink radio is to determine the permissible losses during the propagation of radio waves while ensuring the necessary data transfer rate and maintaining the service area limited by the energy balance of the uplink radio link. The results of the analysis of the energy balance of the SAE/EPS system can be used for the case when the design is carried out in order to ensure the specified size of the service area and data transfer rate and for planning the service areas of a group of additional radio communication channels.

Keywords: radio line, radio communication, energy balance, radio channel, service area, sector..

Постановка наукової проблеми. Розрахунок енергетичного балансу двох радіоліній систем мобільного зв'язку дає можливість визначити розмір зони обслуговування такими системами. Вказана межа має відповідати фізичним параметрам стільника, яка задається таким чином щоб підтримати необхідну швидкість передачі по периметру стільника з вищими швидкостями передачі, але лише в межах її внутрішньої області. Якщо швидкість передачі даних задана, то розмір зони обслуговування під час обміну даними задається з умовою, що всі мобільні абоненти, які знаходяться в її межах, працюють з цією швидкістю передачі із використанням додаткового каналу [1]. Опрацювання наукових джерел [1, 2, 4] показало, що методика аналізу таких радіоліній під час прийому/передачі інформації може застосовуватись для режиму обміну даними за визначених умов:

1. Додаткові канали, які підтримуються радіоінтерфейсом, будуть практично безперервно та постійно зайняті, так як вони розподілені між користувачами.
2. Інформаційна швидкість буде більшою за ту, яка відповідає швидкості передачі даних для межі стільника.
3. При передачі матимуть місце додаткові витрати енергії електромагнітних хвиль в біологічній структурі користувача (близько 2 дБ).

Користувач терміналу доступу з антеною, що знаходиться за межами будівельного об'єкту або транспортного засобу, може вносити малі втрати або їх зовсім не вносити, тому в подальшому аналізі при обміні інформаційними даними ці втрати будуть прирівнюватись до нульових значень. Необхідне співвідношення енергії корисного сигналу до сумарної потужності теплових шумів приймача та радіозавад, призначене для забезпечення заданої якості та надійності роботи радіоліній стосовно додатку обміну даними менше, ніж при передачі розмови. Це пояснюється тим, що додаток обміну даними фактично не працює в режимі реального часу, тому кадри, прийняті з перешкодами, будуть змушені повертатись назад з вимогою забезпечення безпомилкового прийому/передачі даних. Мета формування системи проєкту SAE/EPS полягає в забезпеченні саме якісного обслуговування з передачі даних на високій швидкості.

Аналіз досліджень. Проведені результати моделювання в [2] показали, що неочікуване збільшення допустимої невизначеності передачі символу в кадрі не викликає істотного зменшення пропускної здатності під час обміну даними за протоколом TSP/IP. Для радіочастотного проєктування необхідно враховувати аналіз енергетичного балансу радіоліній, що базується на швидкості передачі додатковим каналом. В такому випадку фізична межа стільника визначатиметься межею зони обслуговування для передачі даних. Якщо використовується аналіз енергетичного балансу радіолінії в режимі передачі мови, то сервіс обміну даними з високою швидкістю буде доступний з відповідною якістю забезпечення зв'язку, а реалізоване зростання об'єму передачі пакетних даних (або імовірність досягнення вищої швидкості передачі даних) зменшиться, коли мобільний абонент знаходитиметься близько до граничної межі стільників. Максимальні допустимі втрати під час розповсюдження радіохвиль у режимі обміну пакетними даними можуть бути збільшені за використання користувачами терміналів доступу антен з вищим коефіцієнтом підсилення і більшою потужністю передавача. Необхідно зауважити, що інтерференційний запас є стабільним (становить 5,5 дБ не зважаючи на те, що число додаткових каналів, доступних для кожної фіксованої швидкості передачі даних, зменшується так само, як і при збільшенні швидкості передачі). Зменшення числа додаткових каналів може збільшити навантаження під час формування стабільності її функціонування.

Мета роботи. Аналітичне опрацювання енергетичного балансу висхідної та низхідної радіоліній в системі SAE/EPS (System Architecture Evolution/Evolved Packet System - еволюція системної архітектури/розвинена пакетна система) під час обміну даними для визначення можливостей розрахунку розміру зони обслуговування з фіксованою швидкістю передачі даних, яку необхідно забезпечити на межі зони стільника передавача/передавача базової станції.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Занижені вимоги щодо відношення енергії корисного сигналу до сумарної потужності теплових шумів приймача та радіозавад, необхідні для забезпечення заданої якості та надійності роботи радіолінії під час обміну даними, можуть бути використанні для збільшення максимальних допустимих втрат під час розповсюдження радіохвиль, а отже і збільшення розміру зони обслуговування. Відомо, що за мінімальної швидкості передачі даних допустимі втрати під час розповсюдження радіохвиль для двох заданих таких режимів роботи системи проєкту SAE/EPS дещо відрізняються, і зі збільшенням швидкості обміну даними ця відмінність зростатиме і досягає 6 дБ. Таким чином, максимальний розмір зони обслуговування для сервісу високошвидкісного обміну даними буде менший в порівнянні з аналогічним параметром для передачі розмови.

Опускаючи громіздкі об'єми обчислення, зупинимося на результатах проведеного аналізу енергетичного балансу. Як приклад використана висхідна радіолінія системи проєкту SAE/EPS з швидкостями передачі пакетних даних від 19,2 кБіт/с до 153,6 кБіт/с, які суттєво відрізняються від одержаних результатів, що для швидкості передачі 9,6 кБіт/с, визначені як довідкові дані. Таким чином, зовнішній фізичний периметр стільника може бути визначений з використанням балансу радіолінії для швидкості передачі даних 9,6 кБіт/с. В межах такого периметру можна розрахувати додаткові обмеження, що накладаються на допустимі втрати енергії електромагнітних хвиль у децибелах з більшою швидкістю обміну даними [4].

Результати аналізу енергетичного балансу системи SAE/EPS можна використовувати для випадку коли проєктування виконується з метою забезпечення заданих розміру зони обслуговування й швидкості передачі даних та для планування зон обслуговування для групи додаткових каналів [3].

Зменшення максимальних допустимих втрат під час розповсюдження радіохвиль із збільшенням швидкості передачі, призведе до зменшення розміру зони обслуговування, що наочно підтверджує ефект «дихання» стільника, характерний для систем зв'язку з кодовим розподіленням каналів [1]. Таке співвідношення енергії якісного сигналу до сумарної потужності теплових шумів приймача та радіозавад, необхідне для забезпечення заданої якості та надійності роботи радіолінії. Отриманий при цьому вигравш (близько 3 дБ) пояснюється застосуванням турбокодів з можливістю повторної передачі/повернення інформаційних даних.

За результатами матеріалів [2, 5] встановлено, що аналіз балансу радіолінії найдоцільніше проводити методами числового моделювання, проте обчислюване навантаження, що виникає під час таких розрахунків, досить велике, що і приводить до необхідності пошуку простіших засобів [5].

Енергетичний баланс низхідної і висхідної радіолінії з використанням методу симетричного аналізу передбачає, що всі мобільні абоненти розміщені симетрично на межі стільника. Мета аналізу низхідної радіолінії полягає у визначенні допустимих втрат під час розповсюдженні радіохвиль із забезпеченням необхідної швидкості передачі даних і підтримці зони обслуговування, яка обмежена енергетичним балансом висхідної радіолінії [3].

Для аналізу вибрана швидкість передачі даних в додатковому каналі, яка підтримується на межі стільника за умови [2]:

- діюче обладнання визначається єдиною швидкістю передачі даних, наприклад 76,8 кБіт/с. Ця швидкість передачі даних відповідає фізичному краю по периметру стільника та таким чином визначає розмір зони обслуговування для заданої швидкості.

- концентричне розташування визначене двома швидкостями передачі даних, наприклад 76,8 кБіт/с і 9,6 кБіт/с. Нижча швидкість передачі даних визначає зовнішню межу стільника для цієї швидкості (рис. 1). Вища швидкість визначає внутрішню межу зони обслуговування, в границях якої може підтримуватися вища швидкість передачі даних (рис. 1).

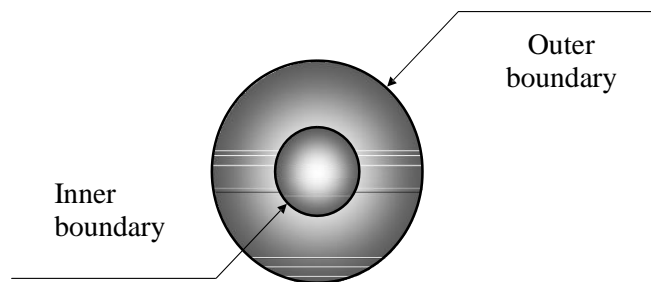


Рис. 1 - Концентричне розташування терміналу мобільного абонента в межах стільника (внутрішні та зовнішні межі визначені двома швидкостями передачі даних): Inner boundary – внутрішня межа, наприклад 76,8 кБіт/с; Outer boundary – зовнішня межа, наприклад 9,6 кБіт/с .

В ході аналізу розрахунковим методом необхідно визначити, чи є достатньою потужність передавача базової станції для досягнення необхідного відношення енергії інформаційного біта до сумарної потужності теплового шуму та інтерференції в приймачі мобільного абонента з урахуванням повільних завмирань електромагнітних хвиль під час їх розповсюдження. Швидкості передачі даних у всіх радіолініях під час обміну даними не є ідентичними. В аналізі повинні розглядатися мобільні абоненти, що використовують обидва канали – основний низькошвидкісний і додатковий високошвидкісний. Програмне забезпечення автоматичного перемикавання секторів доступне лише для основного каналу і не використовується в додаткових каналах низхідної радіолінії.

У системах мобільних телекомунікацій SAE/EPS високошвидкісні додаткові канали динамічно призначаються, коли здійснюється передача пакетних даних, і відключаються, коли така передача закінчується. Для управління рівнем системної інтерференції одночасно активується обмежене число додаткових каналів. Ці високошвидкісні канали можуть розглядатися як сервери в моделі Ерланга [3].

Пакетні дані, які підлягають передачі, або передаються відразу, або зберігаються в очікуванні передачі. У висхідній радіолінії зберігання даних забезпечується в мобільному пристрої обміну даними. У низхідній радіолінії зберігання даних забезпечується в буфері стільника. Ця ситуація відповідає черзі стосовно моделі Ерланга, [3] коли отриманий запит на обслуговування чекає доступу до серверів. В обох випадках може бути збережено велику кількість повідомлень, які надійшли, тоді довжина черги, як і визначає модель Ерланга буде наближатися до безкінечності. Співвідношення між завантаженням, кількістю серверів і середнім часом очікування в черзі також визначаються цією моделлю трафіку. Як тільки перераховані значення визначені, може бути розрахована повна пропускна здатність із використанням наступного виразу:

$$K = \sum_{N=0}^{N-1} n\mu r(n) + N\mu r$$
, де K – повна пропускна здатність; $r(n)$ – час затримки залежно від навантаження [5].

Трафік обміну даними в мережах проекту SAE/EPS передбачається планувати по наведеному в [5] алгоритму:

1. Визначення числа серверів і швидкості обміну даними, які забезпечують радіоінтерфейс у межах сектору.
2. Завдання середнього необхідного часу очікування в черзі на обслуговування.
3. Розрахунок пропускної можливості, яку може забезпечити сектор.
4. Оцінка пропускної можливості, яка може створюватись групою абонентів шляхом визначення загального навантаження.
5. Порівняння пропускної можливості для групи абонентів з ймовірною пропускною можливістю з метою визначення кількості числа секторів.

Наприклад, якщо один сектор може забезпечити підтримку пропускної здатності 100 кБіт/с для заданого обмеженого часу очікування, то необхідно 10 секторів для реалізації загальної пропускної здатності в 1000 кБіт/с. В дійсності ж для виконання ряду інших вимог, наприклад таких, як забезпечення необхідної зони обслуговування, може знадобитися ще більше число секторів, а відповідно необхідно буде збільшити кількість каналів для прийому/передачі інформації.

Далі, за числом секторів можна оцінити залежність інтерференційного співвідношення від розташування мобільних абонентів (МА) в межах стільника. При цьому допускається, що значення і статистика зміни інтерференційного відношення ідентичні як для основних, так і для додаткових каналів. Це визначено симетричним розміщенням всіх МА на межі стільника під час передачі даних і є або точним або консервативним залежно від використовуваної конфігурації низхідної радіолінії. Під час передачі МА, що знаходяться у внутрішній межі стільника, програмне забезпечення серверу терміналу здійснює автоматичне перемикавання секторів; відношення потужності корисного сигналу, що приймається від головного обслуговуючого сектора, до потужності інтерференційних сигналів інших секторів максимізується і фактичне інтерференційне співвідношення погіршується (для розглянутого випадку з (-4 дБ) під час передачі розмови до (+2 дБ) під час обміну даними з використанням додаткового каналу). Наведений приклад досить консервативний, оскільки в ньому використовується вмонтоване розташування, при якому симетричне позиціонування МА на межі стільника забезпечує умови для прийому досить потужних інтерференційних радіосигналів від сусідніх стільників за невеликої потужності сигналу від головного обслуговуючого серверу.

Допустимі втрати під час розповсюдження радіохвиль у низхідній радіолінії складають 128 дБ, це дещо менше, ніж у висхідній радіолінії (не менше 144,4 дБ), але при цьому забезпечується підтримка передачі даних у низхідній радіолінії в основних та додатковому каналах (наприклад зі швидкостями 9,6 і 153,6 кБіт/с відповідно).

Отже, радіус зони обслуговування з передачі даних для заданого трафіку визначається енергетичним потенціалом низхідної радіолінії, що пояснюється діленням потужності передавача базової станції між мобільними абонентами.

Висновки. Аналізуючи вище наведене визначимо, що в даних умовах на межі зони обслуговування може підтримуватися одночасна робота одинадцяти основних каналів у режимі передачі даних і одого додаткового каналу, який забезпечує передачу максимального об'єму трафіку (коефіцієнт використання каналу - одиниця). При цьому 21,8% потужності передавача базової станції витрачається на передачу сигналів у службових каналах, а 78,2% - припадають на основні та додаткові канали трафіку (для розглянутого прикладу повна пропускна здатність низхідної радіолінії складає 259,2 кБіт/с).

Список бібліографічного опису

1. Баховський П.Ф., Євсюк М.М. Кількісний опис вирашу за рахунок м'якого автоматичного перемикавання секторів у висхідній радіолінії систем мобільного зв'язку // Радіотехнічні поля, сигнали, апарати та системи. Міжнародна науково-технічна конференція НТУУ КПІ. Київ, 2015. С. 151-153.
2. Баховський П.Ф., Євсюк М.М. Окремі аспекти розрахунку енергетичного балансу радіоліній проекту SAE/EPS в мережах мобільного зв'язку. // Наукові нотатки. Міжвузівський збірник. - Випуск 37, м. Луцьк, 2012. С.14-22.
3. Баховський П.Ф., Євсюк М.М. Застосування моделі Ерланга до аналізу обміну даними у системах мобільного зв'язку // Радіотехнічні поля, сигнали, апарати та системи. Міжнародна науково-технічна конференція НТУУ КПІ. Київ, 2014, С.59-61с.
4. Баховський П.Ф., Євсюк М.М., Селепина Й.Р. Деякі особливості розрахунку характеристик радіоліній радіоінтерфейсу SAE/EPS в мобільних телекомунікаційних мережах. // 69-та науково-технічна конференція професорсько-викладацького складу, науковців, аспірантів та студентів. Одеська національна академія зв'язку ім. О.С.Попова. Одеса, 2014, 3-5 грудня. С. 9-12.

5. Баховський П.Ф., Євсюк М.М. Окремі аспекти розвитку мобільних телекомунікацій // Перспективні технології та прилади. Збірник наукових праць. Луцьк: Луцький НТУ, 2018. №13. С. 25-33.

References

1. Bakhovskyi P.F., Yevsyuk M.M. Quantitative description of the gain due to soft automatic switching of sectors in the uplink of mobile communication systems // Radio engineering fields, signals, devices and systems. International scientific and technical conference of NTUU KPI. - Kyiv, 2015. P. 151-153.
2. Bakhovskyi P.F., Yevsyuk M.M. Some aspects of calculating the energy balance of radio lines of the SAE/EPS project in mobile communication networks. // Scientific notes. Interuniversity collection. - Issue 37, Lutsk, 2012. P. 14-22.
3. Bakhovskyi P.F., Yevsyuk M.M. The application of the Erlang model to the analysis of data exchange in mobile communication systems // Radio engineering fields, signals, devices and systems. International scientific and technical conference of NTUU KPI. - Kyiv, 2014, P. 59-61.
4. Bakhovskyi P.F., Yevsyuk M.M., Selepina Y.R. Some features of calculating the characteristics of SAE/EPS radio interface radio lines in mobile telecommunication networks. // 69th scientific and technical conference of professors and teaching staff, scientists, graduate students and students. Odesa National Academy of Communication named after O.S.Popova. - Odesa, 2014, December 3-5. P. 9-12.
5. Bakhovskyi P.F., Yevsyuk M.M. Some aspects of the development of mobile telecommunications // Promising technologies and devices. Collection of scientific works. – Lutsk: Lutsk National Technical University, 2018. No. 13. P. 25-33.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-34>

УДК 621.391

Васильківський Микола Володимирович, к.т.н., доцент

<https://orcid.org/0000-0002-6586-2563>

Коломієць Альона Анатоліївна, д.пед.н., доцент

<https://orcid.org/0000-0003-2120-7644>

Грабчак Назарій Віталійович, аспірант

Грицаюк Дмитро Юрійович, магістр

Костянін В'ячеслав Юрійович, бакалавр

Вінницький національний технічний університет, м. Вінниця, Україна

ПРОГРАМНІ ЗАСОБИ PYTHON МОДЕЛЮВАННЯ ТЕЛЕКОМУНІКАЦІЙНИХ ПРИБОРІВ

Васильківський М.В., Коломієць А.А., Грабчак Н.В., Грицаюк Д.Ю., Костянін В.Ю. Програмні засоби Python моделювання телекомунікаційних пристроїв. Здійснено дослідження комп'ютерної моделі телекомунікаційного пристрою модулятора програмно керованої мережі сформувано передумови для розробки та випробування програмного забезпечення для такого пристрою з метою оптимізації його функціональності, продуктивності та надійності. Разом з тим, визначено необхідний інструментарій для розробки ефективного та гнучкого пристрою модуляції сигналів у програмно керованих мережах зв'язку. Основною метою проекту було створення ефективною комп'ютерної моделі модулятора, що здатна адаптуватися до зміни умов роботи мережі та забезпечувати високу якість передачі даних. Результати тестування показали, що модель демонструє високу продуктивність, стабільність у роботі та адаптивність до різних сценаріїв використання. В роботі передбачено використати модель, яка не тільки ефективно працює на поточному етапі, але й може бути легко адаптована під майбутні потреби та технології. Аналіз сумісності та інтеграційного тестування підтверджує, що модель має високу ступінь масштабованості та може бути інтегрована з різноманітними мережевими ком-понентами. Забезпечення стійкості та надійності моделі було критично важливим для гарантування її ефективного роботи в реальних умовах. Результати тестування на стійкість демонструють, що модель здатна працювати стабільно протягом три-вального часу, без втрати продуктивності або виникнення помилок. Виявлено декілька областей для оптимізації, але загалом модель відповідає вимогам надійності. Підсумковою метою роботи була розробка моделі, яка легко інтегрується з існуючими телекомунікаційними системами та програмно керованими мережами. Результати показали, що модель сумісна з різними операційними системами та мережевими обладнанням, що спрощує процес її впровадження та використання в реальних проектах. Комп'ютерна модель демонструє високу ефективність, адаптивність, надійність та легкість інтеграції, відповідаючи по-требам сучасних телекомунікаційних систем. Проте, виявлені потенційні напрямки для подальшого удосконалення вказують на необхідність продовження роботи над дослідженням, зокрема, оптимізація продуктивності при високих навантаженнях та розширення функціональних можливостей моделі.

Ключові слова: модулятор телекомунікаційних сигналів, об'єктно-орієнтована мова програмування, програмно-керована мережа, програмування телекомунікаційних систем.

Vasylykivskiy M., Kolomiets A., Hrabchak N., Hrytsaiuk D., Kostianin V. Python software tools for modeling telecommunication devices. The study of the computer model of a telecommunications device modulator of a software-controlled network formed the prerequisites for the development and testing of software for such a device in order to optimize its functionality, performance, and reliability. At the same time, the necessary tools for the development of an efficient and flexible signal modulation device in software-controlled communication networks were identified. The main goal of the project was to create an effective computer model of the modulator that can adapt to changes in network conditions and ensure high quality data transmission. The test results showed that the model demonstrates high performance, stability in operation, and adaptability to various usage scenarios. The project sought to develop a model that not only works effectively at the current stage, but can also be easily adapted to future needs and technologies. Compatibility analysis and integration testing confirms that the model has a high degree of scalability and can be integrated with a variety of network components. Ensuring the resilience and reliability of the model was critical to ensure its effective operation in real-world conditions. The results of the robustness testing demonstrate that the model is able to operate stably for a long time, without loss of performance or errors. Several areas for optimization were identified, but overall the model meets the reliability requirements. The final goal of the project was to develop a model that can be easily integrated with existing telecommunication systems and software-defined networks. The results showed that the model is compatible with different operating systems and network equipment, which simplifies the process of its implementation and use in real projects. The computer model demonstrates high efficiency, adaptability, reliability, and ease of integration, meeting the needs of modern telecommunication systems. Nevertheless, the identified potential areas for further improvement indicate the need to continue working on the project, in particular, optimizing performance under high loads and expanding the model's functionality.

Keywords: telecommunication signal modulator, object-oriented programming language, software-controlled network, programming of telecommunication systems.

Постановка наукової проблеми. Створення ефективною та гнучкою системи модуляції відкриває шлях до оптимізації роботи телекомунікаційних мереж та забезпечення їхньої ефективності. Аналіз існуючих рішень та робота з теоретичними основами модуляції сигналів допомагають зрозуміти поточний стан справ у галузі та визначити області для подальших

досліджень та покращень. Архітектура та програмна реалізація розробленої моделі на мові Python відкривають широкі можливості для налаштування та оптимізації телекомунікаційних систем, що є ключовим аспектом в сучасному інформаційному суспільстві [1-2].

Демонстрація потенціалу програмної моделі в реальних умовах відображає практичну значущість дослідження і підтверджує можливості подальшого розвитку цієї галузі. Цей проект може відігравати важливу роль у покращенні телекомунікаційних технологій і сприяти розвитку зв'язку та інформаційних технологій взагалі. Розробка програмного забезпечення на мові Python для моделювання роботи модуля-тора в програмно керованих мережах - це важливий крок для покращення та оптимізації телекомунікаційних систем. Такий підхід відкриває широкі можливості для налаштування та експериментів, що дозволяє збільшити ефективність та якість зв'язку.

Демонстрація потенціалу програмної моделі в реальних умовах є ключовим кроком, оскільки вона підтверджує практичну придатність та ефективність розробленого рішення. Це також формує передумови для подальшого розвитку та застосування цієї технології в різних областях телекомунікацій.

Метою роботи є: дослідження особливостей реалізації та тестування програмного забезпечення, що використовується для моделювання різних аспектів роботи телекомунікаційних пристроїв із врахуванням принципів функціонування телекомунікаційних систем та практичних навичок розробки складних програмних систем.

Аналіз досліджень. Розвиток телекомунікаційних технологій, зокрема програмно керованих мереж (SDN), є важливим фактором у сучасному цифровому світі. SDN дозволяє централізовано управляти мережевими ресурсами, що робить їх більш гнучкими та масштабованими, що відповідає потребам сучасних користувачів, які вимагають швидкого та надійного з'єднання. Проте, ефективне використання SDN також потребує вдосконалення інших елементів мережі, таких як модулятори сигналів. Модуляція сигналів в телекомунікаційних системах є критично важливим етапом, оскільки вона визначає швидкість та ефективність передачі даних. Таким чином, для досягнення повного потенціалу SDN необхідно вдосконалювати й інші компоненти мережі, у тому числі модулятори сигналів, щоб забезпечити оптимальну пропускну здатність, мінімальну затримку та надійність передачі даних [3].

Сучасні телекомунікаційні системи потребують високоефективних та адаптивних модуляторів, які здатні швидко адаптуватися до змінних умов передачі даних. Динамічні зміни в мережі можуть включати зміни в каналних умовах, швидкості передачі даних, потоку користувачів тощо. Оптимальні модулятори сигналів мають бути здатні враховувати ці зміни та адаптуватися до них для забезпечення ефективної передачі даних. Важливо також враховувати, що статичні модулятори можуть бути неефективними в умовах динамічних мереж, оскільки вони можуть не враховувати зміни умов передачі даних. Це може призводити до зниження ефективності передачі даних та збільшення витрат на обслуговування мережі [4].

Розробка комп'ютерної моделі модулятора для програмно керованих мереж, яка б задовольняла сучасні потреби телекомунікаційних систем у плані адаптивності, ефективності та масштабованості, є дуже важливим завданням. Така модель може стати ключовим інструментом для оптимізації процесу модуляції сигналів у динамічно змінюваних умовах мережі. Під час розробки цієї моделі можна врахувати різноманітні параметри та умови, які можуть впливати на процес модуляції, такі як швидкість передачі даних, каналні умови, потік користувачів тощо. Це дозволить побудувати модель, яка буде адаптивною до змін у мережі та забезпечить ефективну передачу даних.

Використання програмного управління в такій моделі відкриває можливості для автоматизації та оптимізації процесу модуляції в реальному часі. Це дозволить забезпечити швидку реакцію на зміни в мережі та максимально використовувати її ресурси. Крім того, така модель може служити інструментом для тестування та оцінки різних алгоритмів та стратегій модуляції, що допоможе вибрати найбільш ефективні рішення для конкретних умов мережі. Отже, розробка комп'ютерної моделі модулятора для програмно керованих мереж відповідає сучасним вимогам телекомунікаційного сектору та може значно покращити ефективність та надійність мережевих інфраструктур.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Вивчення основних принципів модуляції сигналів є ключовим для розуміння технології передачі даних у телекомунікаційних системах. Модуляція полягає в зміні одного або декількох параметрів

носія сигналу (наприклад, частоти, амплітуди або фази) для кодування інформації. Це дозволяє передавати дані через канали з обмеженими характеристиками та максимально використовувати доступну пропускну здатність. Особливості програмно керованих мереж (SDN) полягають у тому, що управління мережею розділене від самої мережі, і програмне забезпечення може централізовано керувати ресурсами мережі через відкритий інтерфейс. Це забезпечує більшу гнучкість та швидкість управління мережею, а також дозволяє ефективніше використовувати її ресурси.

Аналіз існуючих рішень включає вивчення різних підходів до модуляції сигналів та їх вплив на продуктивність та ефективність мережі. Недоліки можуть включати обмежену адаптивність до змінних умов мережі, низьку швидкість реакції на зміни, складність управління та інші проблеми, які обмежують ефективність мережі. Цей аналіз допоможе зрозуміти потреби та вимоги до розробки моделі модулятора для програмно керованих мереж, а також визначити напрямки подальшої роботи над її удосконаленням.

При визначенні вимог до комп'ютерної моделі модулятора для програмно керованих мереж (SDN), ключовим є забезпечення адаптивності, ефективності та сумісності із принципами SDN. Враховуючи ці вимоги, комп'ютерна модель модулятора буде здатна ефективно функціонувати в програмно керованих мережах і відповідати вимогам сучасних телекомунікаційних систем.

При проектуванні архітектури моделі модулятора для програмно керованих мереж необхідно ретельно розглянути вибір алгоритмів модуляції, методів програмного управління та інтерфейсів взаємодії між компонентами. Проектування архітектури повинно забезпечити не лише правильність роботи окремих компонентів, але і їх взаємодію для досягнення загальної мети проекту. Детальне проектування архітектури забезпечить ефективність та надійність моделі модулятора в програмно керованих мережах.

Після проектування архітектури наступним кроком є реалізація програмного коду для комп'ютерної моделі модулятора на мові програмування Python. Під час програмної реалізації необхідно дотримуватися визначених вимог та архітектурного плану. Реалізація програмного коду на Python забезпечить функціональність моделі модулятора та її взаємодію з програмно керованими мережами. Після розробки програмного коду моделі модулятора на мові Python необхідно провести тестування для перевірки її роботи у різних сценаріях використання. Після завершення тестування моделі та аналізу результатів, ви зможете впевнено використовувати її у програмно керованих мережах, забезпечуючи високу адаптивність, ефективність та стабільність передачі даних.

Після проведення тестування моделі модулятора і аналізу результатів можна зробити оцінку її ефективності та потенціалу для впровадження в реальні телекомунікаційні системи. На основі задіяних критеріїв і отриманих висновків можна прийняти рішення щодо подальшого використання та розвитку розробленої моделі модулятора для телекомунікаційних систем.

Підготовка технічної документації є важливою частиною розробки будь-якого програмного продукту, включаючи модель модулятора для телекомунікаційних систем. Документація має бути структурованою та легко зрозумілою для будь-якого користувача чи розробника, який має намір використовувати чи розширювати модель. Це допоможе забезпечити ефективне використання моделі та сприятиме подальшому розвитку проекту. Так, виконання цих завдань сприятиме досягненню мети проекту і реалізації його ключових цілей. Розробка ефективної та адаптивної комп'ютерної моделі модулятора відповідно до вимог сучасних програмно керованих телекомунікаційних мереж є ключовою для оптимізації роботи телекомунікаційних систем та забезпечення їхньої ефективності.

Завдання, описані вище, охоплюють всі аспекти розробки моделі, починаючи від теоретичного аналізу і визначення вимог, через проектування архітектури та програмну реалізацію, до тестування, оцінки результатів і підготовки документації. Кожен етап є важливим для забезпечення якості та ефективності розробленої моделі. Розробка такої моделі дозволить не лише вдосконалити роботу конкретних телекомунікаційних систем, а й сприятиме загальному розвитку сфери телекомунікацій та впровадженню передових технологій у цей сектор. Такий підхід відповідає сучасним тенденціям у розвитку інформаційних технологій і забезпечує основу для майбутнього розвитку програмно керованих телекомунікаційних мереж.

Розвиток систем автоматизованого програмування інфокомунікаційних мереж. Системи автоматизованого програмування інфокомунікаційних мереж є ключовими компонентами для управління та оптимізації сучасних телекомунікаційних інфраструктур. Основною метою цих систем є автоматизація процесів конфігурації, моніторингу, управління та відлагодження інфраструктури мережі для забезпечення надійної та ефективної роботи. Основні функції систем

автоматизованого програму-вання інфокомунікаційних мереж включають [6]:

- Автоматизоване конфігурування мережі;
- Моніторинг та діагностика мережі;
- Автоматизоване управління пропускнуою здатністю;
- Резервне копіювання та відновлення конфігурації;
- Інтеграція з принципами SDN;
- Автоматизоване виявлення та вирішення проблем.

Ці системи є важливим інструментом для операторів телекомунікаційних мереж, оскільки вони до-звояють підтримувати високу продуктивність та надійність мережі при мінімальному втручанні опера-тора.

Отже, програмно керовані мережі (SDN) відкривають нові можливості для мережевого управління та реалізації технологій, проте вони також вносять свої виклики і недоліки.

Централізоване управління SDN може створювати нові точки вразливості, оскільки контролер стає однією з ключових мішеней для атак. Додатково, перехід до SDN може потребувати перегляду і перепроєктування заходів безпеки. Оскільки контролер стає центральним елементом управління мережею, він має бути вкрай надійним. Відмова контролера може призвести до серйозних перебоїв у роботі мережі.

Перехід до SDN вимагає від адміністраторів мережі вивчення нових концепцій, протоколів та ін-струментів управління. Це може бути складно для тих, хто вже працює з традиційними мережами. Пере-хід до SDN може потребувати оновлення або заміни існуючого обладнання та програмного забезпечен-ня, що може бути дорогим та складним процесом [7].

Незважаючи на ці виклики, SDN продовжують розвиватися та стають все більш популярними за-вдяки своїм перевагам у гнучкості, масштабованості та ефективності управління мережею. З плином ча-су, багато з цих проблем можуть бути вирішені завдяки розвитку технологій та вдосконаленню практич-ного досвіду у роботі з SDN [8].

Дослідження технологій побудови модуляторів сучасних телекомунікаційних систем. Модулятори є важливою складовою телекомунікаційних систем, оскільки вони відповідають за перетворення цифрової інформації на радіосигнали, які можуть бути передані через канали зв'язку. Вони забезпечують ефективне використання радіочастотного спектра та забезпечують надійну передачу даних.

Сучасні технології модуляції пропонують широкий спектр методів, включаючи амплітудну моду-ляцію (AM), частотну модуляцію (FM), фазову модуляцію (PM), квадратурну амплітудну модуляцію (QAM) та інші. Вибір конкретного методу модуляції залежить від вимог до пропускнуої здатності каналу зв'язку, якості зв'язку, вартості обладнання та інших факторів. Крім того, сучасні модулятори також включають різноманітні технології для оптимізації передачі даних, такі як цифрове перетворення сиг-налу (DAC), алгоритми корекції помилок, а також методи управління потоком даних. Ці технології сприяють підвищенню ефективності та надійності передачі даних у телекомунікаційних системах. Використання передових алгоритмів кодування та модуляції дійсно покращує стійкість сигналу до перешкод та знижує ймовірність помилок при передачі. Це особливо важливо в умовах шумного середовища або на довгих дистанціях передачі [1, 4].

Сучасні методи модуляції, які надають високу ефективність та гнучкість, можуть мати свої склад-нощі при реалізації в практиці. Складність реалізації може виникати через потребу у високоякісному апаратному та програмному забезпеченні, що може підвищувати вартість та складність телекомуніка-ційних систем. Наприклад, для підтримки високошвидкісної передачі даних за допомогою складних ме-тодів модуляції можуть знадобитися спеціалізовані обчислювальні пристрої та швидкість обробки даних.

Потреба в високій потужності обчислень також може бути проблемою, оскільки алгоритми адаптивної модуляції та кодування можуть вимагати значних обчислювальних ресурсів для аналізу умов каналу та оптимізації параметрів передачі в реальному часі. Деякі методи модуляції також можуть бути чутливими до параметрів каналу, таких як затухання та флуктуації сигналу. Це може призвести до вимог до високоякісної інфраструктури для досягнення оптимальної продуктивності, що може збільшувати вартість та складність реалізації системи.

Особливості розроблення комп'ютерної моделі. При розробці комп'ютерної моделі модулятора для програмно керованої мережі дуже важливо врахувати ряд ключових вимог, які забезпечать її ефективність, адаптивність та інтегрованість в сучасні телекомунікаційні системи.

Модель повинна мати здатність динамічно адаптуватися до змінних умов мережі та

параметрів пе-редачі даних, таких як швидкість передачі, рівень шуму та стан каналу зв'язку. Модель повинна бути оптимізована для ефективного використання обчислювальних та мережевих ресурсів. Це включає уникнення зайвих операцій та оптимізацію алгоритмів обробки даних. Модель повинна бути сумісною з принципами програмно керованих мереж, забезпечуючи можливість централізованого управління та взаємодії з іншими компонентами мережі. Модель повинна бути відповідно захищеною від помилок та перешкод, що можуть виникнути в мережі. Це може включати в себе використання методів помилкового виявлення та корекції, а також механізмів автоматичного відновлення після збоїв. Модель повинна мати зрозумілий та легкий у використанні інтерфейс, що дозволить легко інтегрувати її в існуючі телекомунікаційні системи та програми управління мережею. Забезпечення виконання цих вимог допоможе створити реалістичну та практичну модель, яка буде відповідати потребам користувачів та вимогам сучасних телекомунікаційних систем [2, 4].

Архітектура комп'ютерної моделі модулятора для програмно керованої мережі (SDN) відображає сучасні потреби телекомунікаційних систем у високій адаптивності, ефективності та інтегрованості. Вказана комп'ютерна модель може бути розроблена з урахуванням конкретних потреб і характеристик телекомунікаційної системи, забезпечуючи гнучкість та широкі можливості для оптимізації роботи системи передачі даних.

Вибір версії Python є важливим аспектом при розробці програмного забезпечення. Python 3 має ба-гато переваг і його використання рекомендується для нових проєктів, які спрямовані на досягнення високої продуктивності, ефективності та безпеки.

Бібліотеки NumPy і SciPy є незамінними для обробки та аналізу числових даних, а також для вико-нання наукових обчислень. NumPy: надає потужні структури даних, такі як масиви (ndarray), які дозво-ляють ефективно працювати з великими обсягами даних; містить широкий набір функцій для математичних операцій, векторизації та роботи з масивами; є основою для багатьох інших бібліотек та інструментів для наукових обчислень у Python. SciPy: містить розширений набір алгоритмів для наукових обчислень, таких як оптимізація, інтеграція, інтерполяція, оптимізація функцій, розв'язання диференціальних рівнянь та багато іншого; є високорівневим інтерфейсом до операцій NumPy та інших бібліотек, що полегшує роботу з числовими даними та виконання складних наукових обчислень. Обидві ці бібліотеки часто використовуються для моделювання модуляційних процесів, оскільки вони надають інструменти для обробки, аналізу та оптимізації числових даних, які часто виникають у сфері сигнальної обробки та телекомунікацій. Вони можуть бути використані для виконання різних завдань, від обробки сигналів до моделювання та аналізу модуляційних схем.

Matplotlib є потужною бібліотекою для візуалізації даних у Python. Вона надає широкий спектр можливостей для створення різноманітних типів графіків, включаючи лінійні графіки, стовпчаті діаграми, кругові діаграми, гістограми, контурні графіки, графіки розсіювання та багато інших. Matplotlib дозволяє налаштовувати практично всі аспекти графіків, включаючи колір, стиль ліній, розмір шрифту, маркери точок та багато іншого. Вона інтегрується з іншими бібліотеками Python, такими як NumPy, що дозволяє легко візуалізувати дані, оброблені в NumPy масивах. Разом з тим, можна використовувати Matplotlib для створення візуальних зображень, які допоможуть у зрозумінні результатів моделювання, аналізу даних та представленні висновків з ефективності модуляції та діагностики можливих проблем у моделі телекомунікаційної системи [5].

Jupyter Notebook є потужним інструментом для розробки та демонстрації комп'ютерних моделей у Python, особливо у випадках, коли потрібно поєднати код, текстові пояснення та візуалізації в одному документі.

Бібліотеки PyQt та Tkinter є популярними інструментами для розробки графічного інтерфейсу користувача (GUI) у Python. Обидві бібліотеки мають свої переваги та недоліки, і вибір між ними залежить від потреб, рівня зручності використання та функціональних вимог проєкту. Якщо потрібен потужний та багатофункціональний інтерфейс, PyQt може бути кращим вибором. Однак, якщо потрібно щось просте та швидкозрозуміле, Tkinter може бути оптимальним варіантом.

Git є важливим інструментом для розробника програмного забезпечення. Дозволяючи відстежувати зміни, зберігати різні версії програмного коду, співпрацювати з іншими членами команди та впроваджувати нові функції, Git робить процес розробки більш організованим та ефективним. Крім того, Git забезпечує можливість відновлення коду в разі виникнення проблем та гарантує безпеку телекомунікаційного проєкту.

Планування та реалізація комп'ютерної моделі модулятора для програмно керованої мережі

вимагає систематичного та уважного підходу. Python є потужним інструментом для реалізації таких моделей, оскільки він має велику кількість бібліотек для наукових обчислень, а також простий та зрозумілий синтаксис.

Створення комп'ютерної моделі модулятора для програмно керованої мережі вимагає ретельного планування та реалізації програмного коду. Використання Python дозволяє ефективно впоратися з цим завданням завдяки його високому рівню абстракції та великій кількості доступних наукових та інженерних бібліотек. Нижче наведено огляд ключових компонентів програмного коду для реалізації такої моделі.

Першим кроком у реалізації моделі є ініціалізація необхідних бібліотек та встановлення параметрів модуляції. Це включає імпорт бібліотек NumPy для числових обчислень та Matplotlib для візуалізації, а також визначення основних параметрів, таких як частота несучої хвилі, швидкість передачі даних, та тип модуляції.

```
import numpy as np
import matplotlib.pyplot as plt
# Параметри модуляції
carrier_freq = 1e3 # Частота несучої хвилі
data_rate = 1e2 # Швидкість передачі даних
modulation_type = 'QAM' # Тип модуляції
```

Далі необхідно згенерувати інформаційний сигнал, який буде модульовано, та несучу хвилю. Інформаційний сигнал може бути представлений у вигляді випадкового двійкового потоку, тоді як несуча хвиля генерується як синусоїда з заданою частотою.

```
# Генерація інформаційного сигналу
data_length = 100 # Кількість бітів даних
data = np.random.randint(0, 2, data_length)
# Генерація несучої хвилі
t = np.linspace(0, data_length / data_rate, num=int(data_length * 100))
carrier_wave = np.sin(2 * np.pi * carrier_freq * t)
```

На наступному етапі реалізується власне процес модуляції, в залежності від вибраного типу модуляції. Наприклад, для QAM модуляції сигнал модулюється шляхом зміни амплітуди та фази несучої хвилі.

```
def qam_modulate(data, carrier_wave, data_rate):
    # Прикладна реалізація QAM модуляції
    # З цього місця може бути розширена для конкретної логіки QAM
    modulated_signal = carrier_wave * data # Упрощена демонстрація
    return modulated_signal
modulated_signal = qam_modulate(data, carrier_wave, data_rate)
```

Візуалізація інформаційного сигналу, несучої хвилі та модульованого сигналу допомагає в аналізі процесу модуляції та діагностиці потенційних проблем.

```
plt.figure(figsize=(15, 5))
plt.subplot(1, 3, 1)
plt.plot(t[:1000], data[:10], label="Data Signal")
plt.title("Data Signal")
plt.subplot(1, 3, 2)
plt.plot(t, carrier_wave, label="Carrier Wave")
plt.title("Carrier Wave")
plt.subplot(1, 3, 3)
plt.plot(t, modulated_signal, label="Modulated Signal")
plt.title("Modulated Signal")
plt.show()
```

Цей базовий приклад демонструє основні етапи розробки програмного коду для моделі

модулятора телекомунікаційної системи.

Методологія тестування функціональних вузлів телекомунікаційних систем та аналіз результатів

Для забезпечення надійності та ефективності комп'ютерної моделі модулятора для програмно-керованої мережі, ретельне тестування є критично важливим. Методологія тестування повинна охоплювати різні аспекти моделі, від індивідуальних компонентів до загальної функціональності та продуктивності. Нижче описані ключові етапи методології тестування [6].

Юніт-тестування полягає в перевірці окремих компонентів моделі, таких як алгоритми модуляції, генерації сигналів та демодуляції, на предмет коректності їх роботи. Цей етап включає написання тестових сценаріїв, які викликають різні функції моделі з контрольованими вхідними даними та перевіряють, чи відповідають вихідні дані очікуваним результатам.

Після того, як окремі компоненти були протестовані, наступним кроком є інтеграційне тестування, де перевіряється взаємодія між різними частинами системи. Це дозволяє виявити проблеми у взаємодії компонентів, такі як неправильна передача даних або конфлікти в ресурсах.

Тестування продуктивності оцінює, наскільки ефективно модель виконує свої задачі під навантаженням. Це включає вимірювання часу обробки сигналів, споживання пам'яті та процесорного часу при різних рівнях навантаження. Цей тип тестування допомагає виявити потенційні "вузькі місця" в продуктивності та визначити оптимальні конфігурації для роботи моделі.

Оскільки модель модулятора може використовуватися в різних мережевих середовищах, важливо забезпечити її сумісність з різними платформами та протоколами. Тестування сумісності включає перевірку роботи моделі в різних операційних системах, з різними версіями Python та іншим мережевим обладнанням.

Тестування на стійкість перевіряє здатність моделі продовжувати ефективно функціонувати протягом тривалого часу під стандартним або підвищеним навантаженням. Це допомагає виявити потенційні проблеми з пам'яттю, витоками ресурсів або збоями, що можуть виникати при тривалій роботі.

Після застосування зазначеної методології тестування до комп'ютерної моделі модулятора для програмно-керованої мережі, отримано наступні результати, що відображають ефективність, надійність та готовність моделі до використання.

Юніт-тестування показало, що всі індивідуальні компоненти моделі, включаючи алгоритми модуляції, генерації сигналів та демодуляції, працюють коректно відповідно до специфікацій. Всі тестові сценарії, що охоплюють різноманітні вхідні дані та умови, були успішно пройдені без виявлених помилок або непередбачених поведінок.

Інтеграційне тестування виявило високу сумісність та ефективну взаємодію між різними компонентами системи. Дані успішно передаються між модулями без втрат або спотворень, а змінені параметри модуляції коректно впливають на роботу моделі. Тестування також підтвердило, що система адаптується до змін умов мережі, забезпечуючи оптимальну продуктивність [7].

Під час тестування продуктивності модель продемонструвала здатність ефективно обробляти великі обсяги даних з мінімальними затримками. Модель показала стабільну роботу при різних рівнях навантаження, хоча і було виявлено незначне зниження продуктивності при екстремально високому обсязі даних, що вказує на потенційні напрямки для оптимізації.

Тестування сумісності підтвердило, що модель коректно функціонує в різних операційних системах та середовищах. Взаємодія з програмно керованими мережевими компонентами відбувається без помилок, демонструючи готовність моделі до інтеграції у складні телекомунікаційні системи.

Довготривале тестування на стійкість показало, що модель здатна безперервно працювати протягом тривалого часу без втрати продуктивності або стабільності. Не було виявлено витоків пам'яті або інших проблем, що можуть призвести до збоїв у системі.

Загалом, результати тестування свідчать про високу готовність комп'ютерної моделі модулятора до застосування в реальних умовах. Проте, ідентифіковані області для подальшої оптимізації та поліпшення продуктивності при високому навантаженні вказують на необхідність продовження роботи над проектом.

Аналіз отриманих даних після тестування комп'ютерної моделі модулятора. Отримані результати тестування комп'ютерної моделі модулятора надають важливу інформацію для оцінки її ефективності, надійності та готовності до інтеграції у програмно керовану мережу. Аналіз цих даних

дозволяє визначити сильні сторони моделі, виявити потенційні проблеми та напрямки для подальшого удосконалення.

Результати юніт-тестування та інтеграційного тестування підтвердили високу ефективність окремих компонентів моделі та їх здатність коректно взаємодіяти один з одним. Це свідчить про успішну реалізацію основних алгоритмів модуляції та обробки сигналів. Також, адаптивні можливості моделі, які були перевірені під час інтеграційного тестування, демонструють її здатність оптимально функціонувати в динамічних умовах мережі.

Незважаючи на загалом позитивні результати, тестування продуктивності виявило зниження ефективності при екстремально високих рівнях навантаження. Це вказує на потенційні "вузькі місця" в архітектурі моделі, які можуть потребувати оптимізації для підтримки більш високих обсягів даних.

Результати тестування на стійкість демонструють, що модель здатна працювати стабільно та безперервно протягом тривалого часу, що є критично важливим для її застосування в реальних телекомунікаційних системах. Відсутність витоків пам'яті та інших критичних проблем підтверджує високу якість розробки та ефективність використання ресурсів [8].

Тестування сумісності підтвердило, що модель може бути успішно інтегрована у різноманітні про-грамно-керовані мережеві середовища, вказуючи на її універсальність та гнучкість. Це розширює потенційні сфери застосування моделі та забезпечує її сумісність з існуючими та майбутніми технологіями в телекомунікаціях.

По завершенню розробки та тестування комп'ютерної моделі модулятора для програмно-керованої мережі, важливо оцінити, наскільки успішно були досягнуті поставлені цілі проекту. Цей аналіз дозволяє зрозуміти ефективність виконаної роботи та визначити напрямки для подальшого розвитку.

Після успішного завершення основних етапів розробки та тестування комп'ютерної моделі модулятора для програмно керованої мережі, визначення перспектив подальшого розвитку проекту є важливим для забезпечення його тривалої актуальності та вдосконалення. Один з основних напрямків подальшого розвитку полягає у розширенні функціональних можливостей моделі. Це може включати інтеграцію додаткових алгоритмів модуляції, які можуть підтримувати новітні стандарти та вимоги до телекомунікаційних систем. Також важливим є розробка механізмів для більш глибокої адаптації моделі під змінні умови мережі. Незважаючи на вже досягнуту високу продуктивність, завжди існує потенціал для її покращення. Оптимізація коду, використання більш ефективних алгоритмів обчислень та паралелізація процесів можуть сприяти збільшенню швидкості обробки даних та зменшенню затримок.

Для зручності використання моделі кінцевими користувачами, важливо зосередитися на розробці та покращенні інтерфейсу користувача. Це може включати створення графічного інтерфейсу, який дозволяє легко налаштовувати параметри модуляції, візуалізувати процеси модуляції та аналізувати результати в реальному часі. Для підвищення універсальності та практичної цінності моделі, важливо працювати над її інтеграцією з іншими телекомунікаційними та інформаційними системами. Це може включати розробку API для легкої інтеграції з програмно керованими мережами, системами управління даними та іншими інструментами аналізу.

У сучасному цифровому світі питання безпеки даних є критично важливим. Розробка додаткових механізмів шифрування та захисту даних, які передаються за допомогою модулятора, є важливим напрямком для подальших досліджень і вдосконалення моделі. Вказані напрямки подальшого розвитку не лише сприятимуть покращенню існуючої моделі модулятора, але й забезпечать її адаптацію до майбутніх технологічних викликів та потреб користувачів, відкриваючи нові можливості для її застосування в телекомунікаціях.

Висновки та перспективи подальшого дослідження. Розвиток високоефективних та адаптивних модуляторів є важливим напрямком для покращення телекомунікаційних систем і забезпечення їхньої ефективності та надійності в умовах сучасного цифрового світу.

Розробка ефективного та адаптивного модулятора, який би був інтегрований у програмно-керовані мережі, дійсно може вирішити багато викликів, з якими стикаються сучасні телекомунікаційні системи. Підтримка високої адаптивності та ефективності у передачі даних дозволить оптимізувати використання мережевих ресурсів, зменшувати затримки у передачі даних та підвищувати загальну продуктивність телекомунікаційних систем. Це особливо важливо в умовах зростаючих потреб користувачів та швидко-змінюваних вимог цифрового світу. Розв'язання

цієї проблеми може відкрити шлях до створення мереж нового покоління, які будуть готові ефективно адаптуватися до змінних потреб користувачів та динамічно мінятися, щоб забезпечити найвищу якість обслуговування. Це є ключовим аспектом розвитку телеко-мунікаційного сектору в майбутньому.

Список бібліографічного опису

1. Васильківський, М. В. Програмні технології в інфокомунікаційних системах. Навчальний посібник для студентів спеціальності 172 «Телекомунікації та радіотехніка» : електронний навчальний посібник комбінованого (локального та мережного) використання [Електронний ресурс] / Васильківський М. В., Бортник Г. Г., Кичак В. М. – Вінниця : ВНТУ, 2023. – 141 с.
2. Основи програмування (Python, Java) : лабораторний практикум / Смотр О., Придатко О., Малець І. – Львів : ЛДУ БЖД, 2019. – 134 с.
3. Програмування мовою Python / О.М. Васильєв. — Тернопіль: Навчальна книга – Богдан, 2019. — 504 с.; іл.
4. Васильківський, М., Коломієць, А., & Грабчак, Н. (2022). Дослідження функціональних параметрів інфокомунікаційних мереж 6G. Вісник Хмельницького національного університету, (6), 46–52. <https://www.doi.org/10.31891/2307-5732-2022-315-6-46-52>
5. Шарадкін Д.М., Субач І.Ю., Микитюк А.В. Інструментальні засоби Python для моделювання та системного аналізу часових рядів при вирішенні задач кіберзахисту інформаційно-комунікаційних систем: навч. пос. / Шарадкін Д.М., Субач І.Ю., Микитюк А.В.; ІСЗЗІ КПІ ім. Ігоря Сікорського. – Київ : КПІ ім. Ігоря Сікорського, 2023.- 139 с.
6. Березовський В. Є. Чисельні методи з прикладами реалізації мовою Python / В. Є. Березовський, Л. Є.Ковальов, М. О.Медведева : навчальний посібник. Умань : ВПЦ «Візаві», 2023. 88 с.
7. Васильківський, М., Прикмета, А., Олійник, А., & Нікітович, Д. (2023). Оптимізація інтелектуальних телекомунікаційних мереж. Вісник Хмельницького національного університету, Технічні науки. – 2023. – № 1. (317). – С. 33–41. doi: 10.31891/2307-5732-2023-317-1-33-41
8. М. Васильківський, О. Городецька, Б. Климчук, і В. Говорун, «Стратегії технологічного розвитку апаратного забезпечення інфокомунікаційних радіомереж», ІТКІ, вип. 56, вип. 1, с. 83–91, Бер 2023.

References

1. Vasykivskiy, M. V. Prohramni tekhnolohii v infokomunikatsiinykh systemakh. Navchalnyi posibnyk dlia studentiv spetsialnosti 172 «Telekomunikatsii ta radiotekhnika» : elektronnyi navchalnyi posibnyk kombinovanoho (lokalnoho ta merezhnoho) vykorystannia [Elektronnyi resurs] / Vasykivskiy M. V., Bortnyk G. G., Kychak V. M. – Vinnytsia : VNTU, 2023. – 141 s. [in Ukrainian].
2. Osnovy prohramuvannia (Python, Java) : laboratornyi praktykum / Smotr O., Prydatko O., Malets I. – Lviv : LDU BZhD, 2019. – 134 s. [in Ukrainian].
3. Prohramuvannia movoiu Python / O.M. Vasyliiev. — Ternopil: Navchalna knyha – Bohdan, 2019. — 504 s.; il. [in Ukrainian].
4. Vasykivskiy, M., Kolomiets, A., & Hrabchak, N. (2022). Doslidzhennia funktsionalnykh parametriv infokomunikatsiinykh merezh 6G. Visnyk Khmelnytskoho natsionalnoho universytetu, (6), 46–52. <https://www.doi.org/10.31891/2307-5732-2022-315-6-46-52> [in Ukrainian].
5. Sharadkin D.M., Subach I.Iu., Mykytiuk A.V. Instrumentalni zasoby Python dlia modeliuvannia ta systemnoho analizu chasovykh riadiv pry vyrishenni zadach kiberzakhystu informatsiino-komunikatsiinykh system: navch. pos. / Sharadkin D.M., Subach I.Iu., Mykytiuk A.V.; ISZZI KPI im. Ihoria Sikorskoho. – Kyiv : KPI im. Ihoria Sikorskoho, 2023.- 139 s. [in Ukrainian].
6. Berezovskyi V. Ye. Chyselni metody z prykladamy realizatsii movoiu Python / V. Ye. Berezovskyi, L. Ye.Kovalov, M. O.Medvedieva : navchalnyi posibnyk. Uman : VPTs «Vizavi», 2023. 88 s. [in Ukrainian].
7. Vasykivskiy, M., Prykmeta, A., Oliinyk, A., & Nikitovych, D. (2023). Optyimizatsiia intelektualnykh telekomunikatsiinykh merezh. Visnyk Khmelnytskoho natsionalnoho universytetu, Tekhnichni nauky. – 2023. – № 1. (317). – S. 33–41. doi: 10.31891/2307-5732-2023-317-1-33-41 [in Ukrainian].
8. M. Vasykivskiy, O. Horodetska, B. Klymchuk, i V. Hovorun, «Stratehii tekhnolohichnoho rozvytku aparatnoho zabezpechennia infokomunikatsiinykh radiomerezh», ITKI, vyp. 56, vyp. 1, s. 83–91, Ber 2023. [in Ukrainian].

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-35>

УДК 519.876.5; 621.31.33

Лишук Віктор Васильович, к.т.н., доцент

<https://orcid.org/0000-0003-4049-8467>

Мороз Сергій Анатолійович, к.т.н., доцент

<https://orcid.org/0000-0003-4677-5170>

Заблоцький Валентин Юрійович, к.т.н., доцент

<https://orcid.org/0000-0002-2921-0031>

Євсюк Микола Миколайович, к.т.н., доцент

<https://orcid.org/0000-0002-2421-1844>

Гриценюк Віталій Віталійович, студент

Луцький національний технічний університет, м. Луцьк, Україна

АЛГОРИТМИ КЕРУВАННЯ АВТОНОМНИМИ ІНВЕРТОРАМИ НАПРУГИ ПЕРЕТВОРЮВАЧІВ ЧАСТОТИ

Лишук В.В., Мороз С.А., Заблоцький В.Ю., Євсюк М.М., Гриценюк В.В. Алгоритми керування автономними інверторами напруги перетворювачів частоти. У статті запропоновано розробку алгоритмів керування автономними інверторами напруги перетворювачів частоти. Ключовими елементами трифазного інвертора є IGBT-транзистори. В порівнянні з іншими електронними ключами вони мають більшу швидкодію та можливість виготовлення модуля з шести транзисторів та діодів. Запропоновані алгоритми перетворювача дадуть змогу розуміти фізичні процеси в перетворювачах, правильно проектувати, діагностувати та експлуатувати такі пристрої. Перетворювач частоти виготовляється на основі новітніх інтегрованих модулів потужності IPM (Intelligent Power Module), що містять 6 IGBT-транзисторів та їх кола керування з системою захисту від коротких замикань і температурним захистом. Керування модуляцією вихідної напруги реалізується програмно за допомогою мікроконтролерів. У ПЧ застосована модифікована модуляція, так званої „орієнтації вектора напруги". Це сприяє повному використуванню можливостей ПЧ. Пристрій може працювати в режимі лінійної або квадратичної характеристики U/f . Ці характеристики дають змогу змінювати швидкість обертання асинхронних двигунів в широкі межі, що важливо для приводу різноманітних механізмів і в кінцевому варіанті покращить економічні та енергетичні показники електроприводу. В даний час основними приладами силової електроніки в області струмів комутованих до 50 А є біполярні транзистори з ізольованим затвором (IGBT).

Ключові слова: перетворювач частоти, IGBT-транзистор, інвертор, електронний ключ.

Lyshuk V., Moroz S., Zablotzkyi V., Yevsiuk M., Hryceniuk V. Algorithms for controlling autonomous voltage inverters of frequency converters. The article proposes the development of control algorithms for autonomous voltage inverters. The key elements of the frequency converter inverter are IGBT transistors. Compared to other electronic keys, they have a higher speed and the possibility of manufacturing a module of six transistors and diodes. The proposed converter algorithms will make it possible to understand physical processes in converters, to correctly design, diagnose and operate such devices. The frequency converter is made on the basis of the latest integrated power modules IPM (Intelligent Power Module), which contain 6 IGBT transistors and their control circuits with a system of protection against short circuits and thermal protection. Output voltage modulation control is implemented programmatically using microcontrollers. The inverter uses a modified modulation, the so-called "voltage vector orientation". This contributes to the full use of the inverter's capabilities. The device can operate in the mode of linear or quadratic U/f characteristics. These characteristics make it possible to vary the rotation speed of asynchronous motors within wide limits, which is important for drive various mechanisms and will ultimately improve the economic and energy performance of the electric drive. Currently, the main power electronics devices in the field of switching currents up to 50 A are insulated gate bipolar transistors (IGBT).

Keywords: mathematical model, frequency converter, IGBT-transistor, inverter, electronic key.

Постановка наукової проблеми. Регулювання частоти напруги електричної мережі є сучасним напрямом в області силової електроніки та електроприводу. Перетворювачі частоти (ПЧ) є інтелектуальними модулями, що автоматизують технологічні процеси на виробництві. Це дає змогу суттєво скоротити витрати на енергоспоживання, оскільки графіки електричних навантажень різко змінні. Наприклад, теплоелектроцентралі, що слугують опаленням для населення, використовують такі перетворювачі і дають змогу регулювати подачу гарячої води в різну пору доби за рахунок зміни швидкості обертання приводних асинхронних двигунів-насосів. Вони можуть під'єднуватись як до однофазної так і трифазної мережі. ПЧ є складним електронним пристроєм, в якому мережна напруга з частотою 50 Гц перетворюється в змінну напругу з регульованими амплітудою і частотою [4].

Електроніка ПЧ живиться стабілізованими напругами від блоку живлення, підключеного в ланку постійного струму, яка зберігає працездатність при фазній напрузі мережі в межах від 190 до 250 В. Так як електронна частина ПЧ живиться постійним струмом, то тут забезпечується стійка робота системи при коливаннях і короткочасному зникненні напруги мережі. Кола

керування інвертора гальванічно розв'язані від кіл процесора і випростувача. Повна гальванічна роз'язка процесора, як від кіл інвертора, так і від вхідної частини, забезпечує велику завадостійкість мікропроцесора.

ПЧ може керуватися напругою величиною 2-10 В або струмом величиною 4-20 мА. В режимі роботи з «плаваючим нулем» зниження рівня вхідного сигналу до величин нижче 2 В або 4 мА викликає припинення роботи ПЧ. Система забезпечена розширеною системою діагностики, налаштувань і блокувань, що захищають ПЧ і навантаження від пошкоджень.

Перетворювачі з регульованою вихідною частотою застосовуються в електроприводах змінного струму, а перетворювачі зі стабілізованою вихідною частотою – у технологічних установках, джерелах живлення, включаючи і джерела безперебійного живлення [2].

До перетворювачів частоти, призначених для частотного керування електроприводами, висуваються такі основні вимоги:

- незалежне регулювання величини та частоти вихідної напруги;
- можливість двостороннього обміну енергією між навантаженням та мережею живлення;
- стійкість перебігу динамічних режимів частотного керування електроприводу;
- швидкодіючий захист та ефективна діагностика;
- високі коефіцієнт корисної дії та симетрія фазних напруг і струмів;
- відсутність постійних складових та субгармонік у перетворених напругах та струмах;
- мінімальні спотворення напруги мережі живлення, низький рівень радіоперешкод і шуму;
- підтримка з необхідною точністю частоти, напруги (струму) в усталених режимах роботи електроприводу.

За принципом роботи перетворювачі частоти поділяються на перетворювачі частоти з ланкою постійного струму та перетворювачі частоти без ланки постійного струму або перетворювачі частоти з безпосереднім зв'язком кіл навантаження та мережі живлення.

ПЧ із ланкою постійного струму мають у своєму складі випрямляч та інвертор. Випрямляч перетворює електричну енергію змінного струму на електричну енергію постійного струму, а інвертор перетворює електричну енергію постійного струму на електричну енергію змінного струму. Очевидною перевагою ПЧ зі ланкою постійного струму є незалежність частоти вихідної напруги на виході ПЧ від частоти мережі живлення [3]. На рисунку 1 показано структурну схему перетворювача частоти.

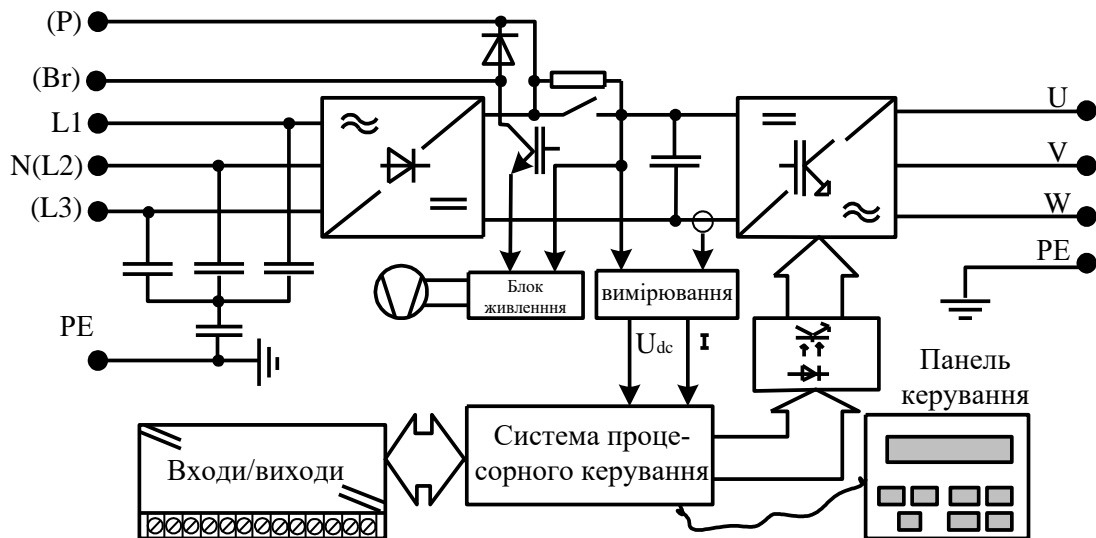


Рис. 1 – ПЧ з керованим випрямлячем та інвертором напруги

На індикаторі панелі керування відображаються наступні аварії:

- коротке замикання або струм, перевищуючий допустимий, на виході ПЧ;
- напруга в колі постійного струму, перевищує допустиму;
- знижена напруга в колі постійного струму;
- температура радіатора вище гранично допустимої;

- термічний захист двигуна;
- відсутність живлячої фази.

Регулювання частоти вихідної напруги ПЧ здійснюється шляхом зміни частоти перемикання напівпровідникових ключів (транзисторів або керованих тиристорів) автономного інвертора. Величина вихідної напруги у цих схемах може регулюватися шляхом регулювання величини вихідної напруги керованого випрямляча (амплітудний спосіб).

Аналіз досліджень. На сьогодні перспективним є застосування широтно-імпульсного методу регулювання вихідної напруги інвертора, що реалізується шляхом застосування відповідного алгоритму керування вентилями автономного інвертора. Розрахункові співвідношення встановлюють зв'язок між напругою кола постійного струму напруги інвертора і напругою змінного струму на виході інвертора [3]. Діюче значення лінійної вихідної напруги ІН із синусоїдною ШІМ визначається з рівняння (1).

$$U_{\text{нр.л}} = \frac{\sqrt{3}}{2\sqrt{2}} \mu U_{\text{п}}, \quad (1)$$

де $U_{\text{п}}$ – середнє значення напруги на вході інвертора;
 μ – коефіцієнт модуляції.

Для електроприводів змінного струму, у яких випадки рекуперації електричної енергії в мережу живлення досить рідкісні, можливе застосування схеми ПЧ, наведеної на рисунку 2. У цій схемі енергія навантаження, що рекуперується в коло постійного струму, розсіюється на баластному опорі R_b при ввімкненні транзистора VT. Транзистор VT, який називають чопером, вмикається сигналом системи керування мікропроцесора в тому випадку, коли напруга на конденсаторі фільтра підвищується вище встановленої межі. Підвищення напруги на конденсаторі відбувається при переході асинхронного двигуна в генераторний режим [2].

При живленні інвертора напруги від джерела постійної напруги з односторонньою провідністю (від випрямляча) виникає необхідність встановлення на вході інвертора компенсуючого конденсатора, який повинен приймати енергію в моменти часу, коли струм спрямований від інвертора до джерела живлення. Ємність компенсуючого конденсатора може бути знайдена за формулою (2).

$$C_0 = \frac{\int_t^{t+\Delta t} i_d dt}{\Delta U_c}, \quad (2)$$

де Δt – інтервал часу, протягом якого струм i_d кола постійного струму напрямлений від інвертора до джерела;

ΔU_c – допустима перенапруга на конденсаторі.

На рисунку 2 показано ПЧ з рекуперацією електричної енергії.

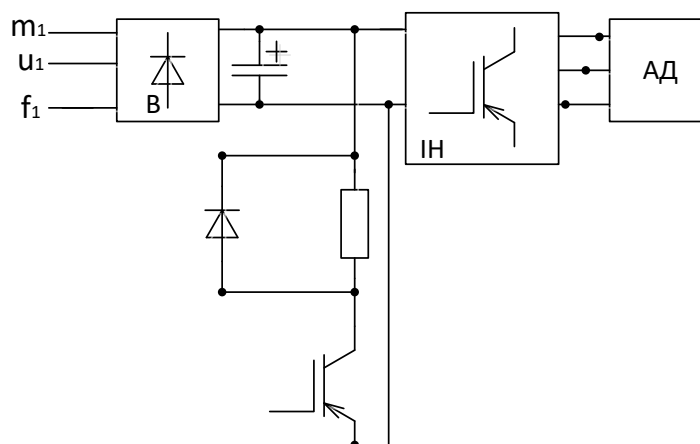


Рис. 2 – Перетворювач частоти з рекуперацією електричної енергії в коло чопера

Вирішуючи рівняння (2), отримаємо формулу для розрахунку величини ємності компенсуючого конденсатора:

$$C_0 = \frac{\sqrt{3}}{2} \frac{\mu I_{нзм}}{f_{нес} \Delta U_c} \sin \frac{\varphi_{нр(1)} - \pi / 6}{2}, \quad (3)$$

де μ – коефіцієнт модуляції ($0 < \mu < 1$);

$I_{нзм}$ – амплітудне значення струму навантаження;

$f_{нес}$ – несуча частота ШІМ;

$\varphi_{нр(1)}$ – фазовий кут між першими гармоніками напруги та струму.

Вираз (3) показує, що ємність компенсуючого конденсатора не залежить від вихідної частоти. Ця обставина дозволяє використовувати інвертори з ШІМ для роботи на низьких частотах. Слід відмітити, що ємність компенсуючого конденсатора обернено пропорційна несучій частоті. Завдяки тому, що несуча частота досить висока, то ємність компенсуючого конденсатора в інверторах напруги з ШІМ завжди менша, ніж у інверторів без ШІМ.

Виклад основного матеріалу й обґрунтування отриманих результатів дослідження. Регулювання вхідної напруги може здійснюватися зміною моменту закривання того ключа, який (які) в даний момент часу відкритий (чи закриті). Структура силової частини інвертора при алгоритмі переключення ключів не змінюється [1].

Розглянемо порядок розробки алгоритмів керування ключами інвертора для трифазної мостової схеми інвертора, при з'єднанні навантаження зіркою, що є характерним при під'єднанні асинхронного двигуна в трифазну мережу (рис.3).

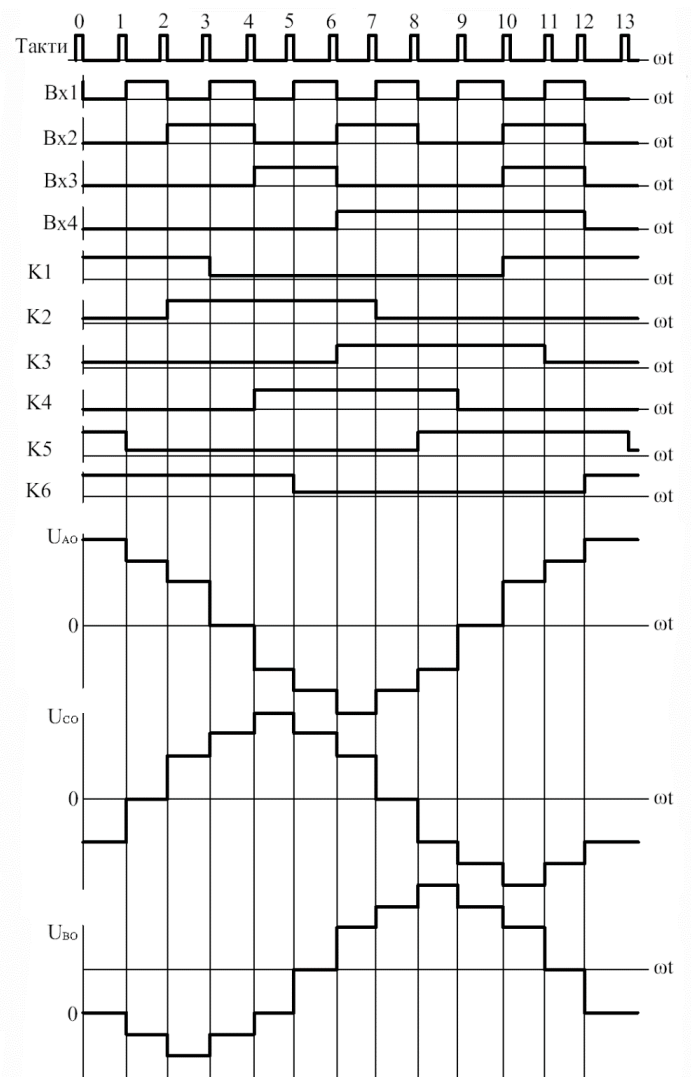


Рис. 3 – Алгоритм переключення ключів у трифазному мостовому інверторі при симетричному навантаженні, з'єднаному зіркою

Регулювання змінної напруги U_{\sim} та f_{\sim} на виході інвертора може здійснюватись або регулюванням вхідної напруги (вхідної напруги випростувача), або регулюванням, як було сказано вище, часу ввімкнення t_i кожного ключа інвертора, при цьому $U_{\sim} = U_d \cdot \gamma$ на кожному інтервалі регулювання.

В роботі як базові використовуються інвертори, в яких формується як однопульсна, двопульсна так і п'ятипульсна за половину періоду вхідної напруги форма кривої фазних напруг U_A, U_B, U_C .

Алгоритм перемикання ключів VT1-VT6 відповідає закону зміни керуючої напруги, що подається на ці ключі (напрузі база-емітер ключів VT1-VT6) на рисунку 3 полярністю. Період (частота) напруги інвертора змінюється часом вмикання-вимикання керуючих імпульсів на відповідні ключі.

Алгоритм керування автономними інверторами напруги реалізується шляхом програмування мікропроцесорної системи керування, що забезпечує також подачу імпульсів керування відповідної напруги, потужності імпульсів на ключі (транзисторі інвертора).

Керування транзисторами ІН здійснюється мікроконтролером. На мікроконтролер через АЦП надходять два сигнали керування (рис.4):

- керування величиною вихідної напруги інвертора (резистор R32);
- керування частотою вихідної напруги інвертора (резистор R31).

Крім імпульсів керування транзисторів інвертора мікроконтролер виробляє імпульси керування транзистора VT37. Завдання транзистора VT37 полягає у підключенні гальмівного резистора R33 паралельно конденсатору фільтра С31 для запобігання перенапруги на конденсаторі С31 при генераторному режимі роботи двигуна [5].

На рисунку 3 показано схему підключення датчика частоти (R31) в схемі керування інвертора напруги.

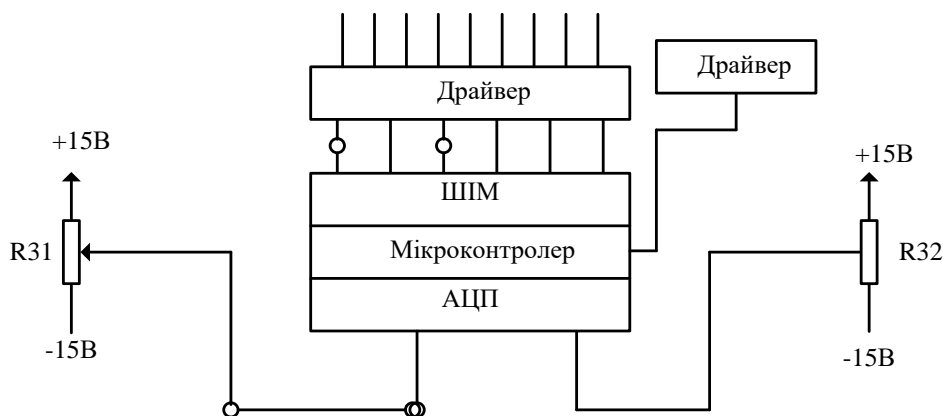


Рис. 4 – Схема підключення датчика частоти (R31) в схемі керування інвертора напруги

Схема підключення драйвера до силових транзисторів одного плеча інвертора напруги наведена на рисунку 5.

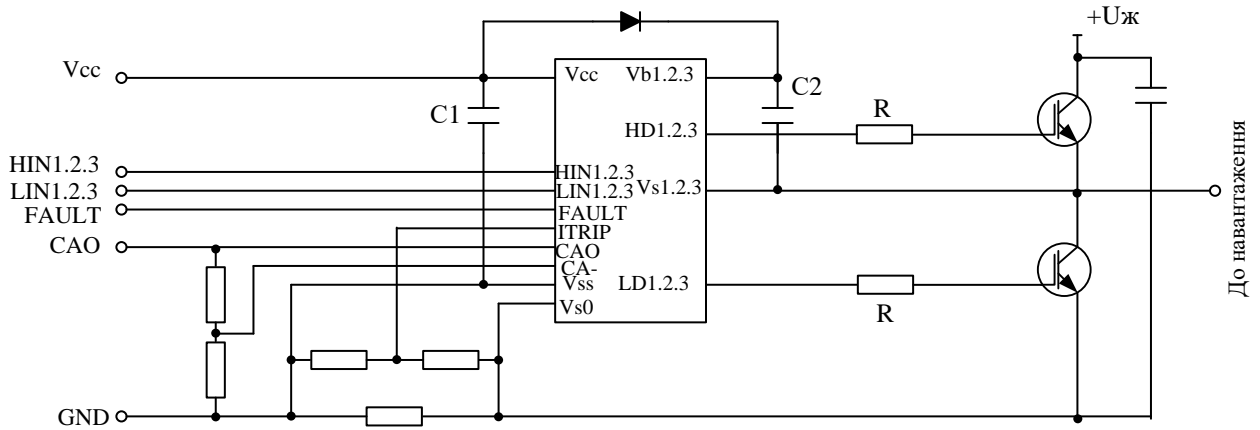


Рис. 5 – Схема підключення драйвера до силових транзисторів одного плеча інвертора

Транзисторний ключ (ТК) на базі IGBT-технології керує процесами перетворення електричної енергії. Специфіка протікання цих електричних процесів вимагає детальнішого розгляду принципів роботи ТК та її елементної бази задля забезпечення надійності ПЧ [4].

Розглянемо класичну схему одного плеча перетворювача. Зауважимо, що ПЧ містить три такі плеча. На рисунку 5 представлена схема такого плеча та показані електромагнітні процеси, що протікають у ньому при включенні та виключенні транзистора. Класична теорія динамічних процесів виділяє чотири комутаційних інтервали при роботі плеча на активно-індуктивне навантаження, два при включенні транзистора і два при вимкненні. При включенні транзистора виділяються етап відновлення діода у фазі високої зворотної провідності (t_1 на рис. 6) та етап встановлення стаціонарного стану силового високовольтного транзистора (t_2 на рис. 6).

В області комутованих струмів до 50 А основними приладами силової електроніки є силові модулі на базі біполярних транзисторів та силові модулі на базі IGBT технології.

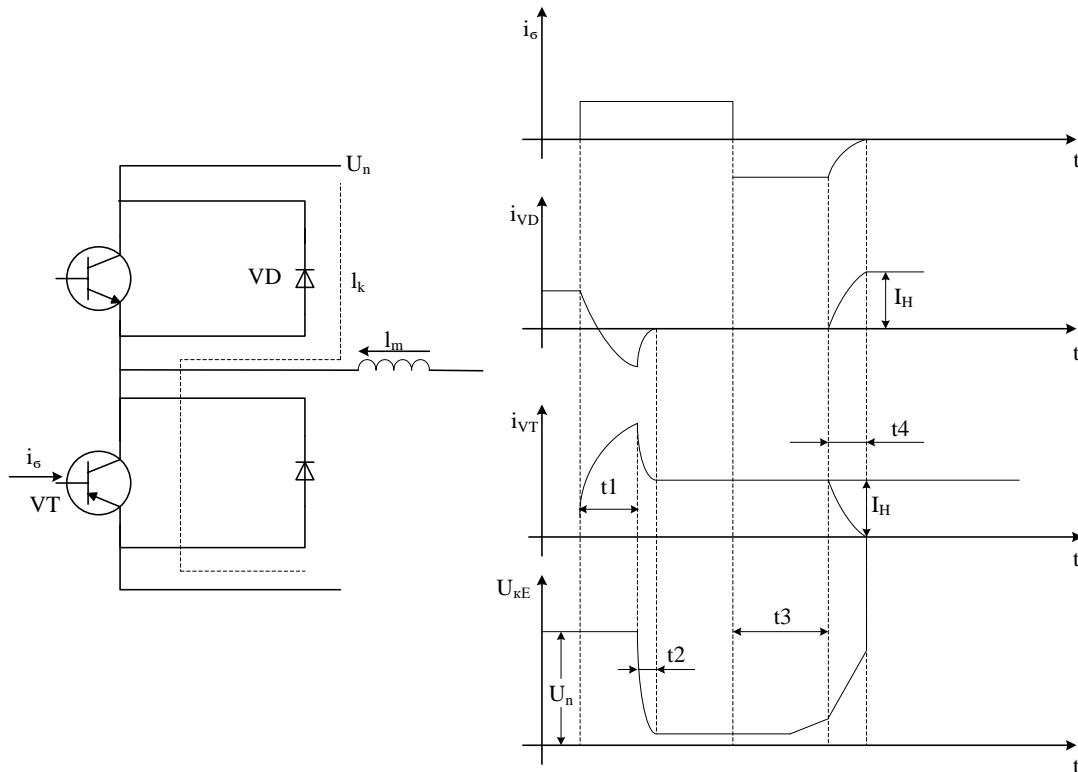


Рис. 6 – Динамічні процеси при перемиканні силового транзисторного ключа

На першому із зазначених етапів кола «транзистор-діод» протікає значний струм, який може перевищити номінальний у кілька разів. При цьому напруга на транзисторі залишається рівною напрузі живлення. Цей етап є найнебезпечнішим для транзистора. На другому етапі струм зменшується до номінального при одночасному зменшенні напруги на транзисторі. При вимкненні транзистора виділяються етап розсмоктування неосновних носіїв заряду в колекторі силового високовольтного транзистора (t_4 на рис. 6) та етап спаду струму колектора силового транзистора та включення діода (t_4 на рис. 6). На всіх зазначених інтервалах комутації у транзисторі та діоді виділяється значна потужність [1].

Висновки та перспективи подальшого дослідження. Перетворювачі частоти, побудовані на трифазних інверторах, є одними з перспективних напівпровідникових перетворювачів енергії з безпосереднім зв'язком джерела напруги і навантаження, дають змогу реалізовувати робочі режими електро-приводу так і режими рекуперативного гальмування. Їх переваги полягають в тому, що вони реалізують двосторонній обмін між джерелом і навантаженням, можливість формування струму синусоїдної форми, мають високі ККД та електромагнітну сумісність. Використання повністю керованих транзисторних ключів з високими динамічними характеристиками та двосторонньою симетричною провідністю дає змогу забезпечити досить широкий діапазон зміни частоти напруги мережі.

Список бібліографічного опису

1. Жуйков В.Я. Енергетичні процеси в електричних колах з ключовими елементами / В.Я. Жуйков, С.П. Денисюк. – К.: Текст, 2010. – 264 с.
2. Кириленко О.В. Системи силової електроніки та методи їх аналізу / О.В. Кириленко, В.Я. Жуйков, С.П. Денисюк, О.Б. Рибіна. – К.: Текст, 2006. – 488 с.
3. Лишук В. В., Мороз С. А., Погинець А. Я., Кмитко Н. О., Барташук Р. І. Математична модель інвертора перетворювача частоти. Науковий журнал Комп'ютерно-інтегровані технології: освіта, наука, виробництво. Луцьк, 2023. Випуск № 52. с.140-146.
4. Махно О.О. Автономні перетворювачі. Навч. посібник / О.О. Махно, В.В. Семенов, О.В. Будьонний, Н.А. Омелчук. – Запоріжжя: ЗДІА, 2009. – 126 с.
5. Шавьолкін О. О. Енергетична електроніка: навч. посібник / О.О. Шавьолкін. – К.: КНУТД, 2017. – 396 с.

Referenses

1. Zhujkov V.Ya. Energetichni procesi v elektrichnih kolah z klyuchovimi elementami / V.Ya. Zhujkov, S.P. Denisyuk. – K.: Tekst, 2010. – 264 s.
2. Kirilenko O.V. Sistemi silovoyi elektroniki ta metodi yih analizu / O.V. Kirilenko, V.Ya. Zhujkov, S.P. Denisyuk, O.B. Ribina. – K.: Tekst, 2006. – 488 s.
3. Lyshuk V. V., Moroz S. A., Pohynets A. Ya., Kmytko N. O., Bartashchuk R. I. Matematychna model invertora peretvoriuvacha chastoty. Naukovyi zhurnal Kompiuterno-intehrovani tekhnolohii: osvita, nauka, vyrobnytstvo. Lutsk, 2023. Vypusk № 52. s.140-146.
3. Mahno O.O. Avtonomni peretvoryuvachi. Navch. posibnik / O.O. Mahno, V.V. Semenov, O.V. Budonnij, N.A. Omelchuk. – Zapo-rizhzhya: ZDIA, 2009. – 126 s.
4. Shavolkin O. O. Energetichna elektronika: navch. posibnik / O.O. Shavolkin. – K.: KNU TD, 2017. – 396 s.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-36>

УДК 004.77

Мадінов Микола Леонідович, аспірант

<https://orcid.org/0009-0005-5910-8774>

Державний університет інформаційно-комунікаційних технологій, м. Київ, Україна

ОПТИЧНІ ЛІНІЇ ЗВ'ЯЗКУ

Мадінов М.Л. Оптичні лінії зв'язку. Передача сигналів паралельними шляхами є ключовим аспектом сучасних ліній зв'язку, особливо тих, що включають використання оптичних волокон. Просторовий паралелізм, фундаментальна концепція в цій області, відіграє вирішальну роль у підвищенні пропускної здатності та ефективності волоконно-оптичних мереж зв'язку. Завдяки одночасному передаванню сигналів через кілька просторових каналів просторовий паралелізм значно підвищує швидкість передачі та пропускну здатність даних. Цей підхід дозволяє одночасно використовувати різні шляхи в межах оптичного волокна, ефективно збільшуючи доступну пропускну здатність і підвищуючи загальну продуктивність мережі. Завдяки ретельному аналізу стає очевидним, що фактор паралельності, який представляє кількість просторових каналів, створених за допомогою мультиплексування, є ключовим визначальним фактором для оптимізації спектральної ефективності та сукупної швидкості передачі даних. Наголошується, що роль просторового мультиплексування полягає в тому, щоб ефективно та якомога рівномірніше поєднувати світло від пучка одномодових волокон на ядра багатоядерного волокна або моди багатоходових або когерентних ядерних волокон, і навпаки. У випадку мультиплексування для одномодового багатоядерного або когерентного волокна механізм з'єднання є простим і відображає масив одномодових вхідних променів на масив одномодових вихідних променів. Загалом, робота містить уявлення про роль просторового паралелізму в сучасних волоконно-оптичних системах зв'язку. Вона досліджує, як передача сигналів через паралельні шляхи підвищує швидкість передачі та пропускну здатність даних. У статті кількісно оцінено внесок просторового мультиплексування у швидкість передачі даних та спектральну ефективність. Вона підкреслює важливість методів просторового мультиплексування, таких як поляризація, спектральне, модальне та базове мультиплексування для оптимізації продуктивності мережі.

Ключові слова: просторовий паралелізм, волоконно-оптичний зв'язок, мультиплексування, передача даних, сміність мережі, ефективність зв'язку.

Madinov M. Optical communication lines. The transmission of signals through parallel paths is a key aspect of modern communication lines, especially those involving the use of optical fibers. Spatial parallelism, a fundamental concept in this field, plays a crucial role in increasing the capacity and efficiency of fiber optic communication networks. By simultaneously transmitting signals through multiple spatial channels, spatial parallelism significantly enhances transmission speed and data throughput. This approach allows for the simultaneous utilization of different paths within the optical fiber, effectively increasing the available bandwidth and improving the overall network performance. Through careful analysis, it becomes evident that the parallel factor, representing the number of spatial channels created through multiplexing, is a key determinant for optimizing spectral efficiency and aggregate data transmission speed. It is emphasized that the role of spatial multiplexing is to efficiently and as uniformly as possible couple light from a bundle of single-mode fibers onto the cores of a multi-core fiber or the modes of a few-mode, multi-mode, or coherent-core fiber, and vice versa. In the case of multiplexing for single-mode multi-core or coherent fiber, the connection mechanism is straightforward, reflecting an array of single-mode input beams onto an array of single-mode output beams. The work provides insights into the role of spatial parallelism in modern optical fiber communication systems. It examines how transmitting signals through parallel paths enhances transmission speed and data throughput. The article quantitatively evaluates the contribution of spatial multiplexing to data transmission speed and spectral efficiency. It highlights the importance of spatial multiplexing techniques such as polarization, spectral, modal, and core multiplexing in optimizing network performance. Overall, the abstract emphasizes the significance of spatial parallelism in improving the efficiency and capacity of optical fiber communication networks.

Key words: spatial parallelism, fiber optic communication, multiplexing, data transmission, network capacity, communication efficiency.

Вступ та постановка проблеми. У загальному розумінні цього терміну, оптичні лінії зв'язку являють собою системи, які використовують світло, що зазвичай передається через оптичні волокна, забезпечуючи передачу інформації на великі відстані з високою пропускну здатністю та низькими втратами сигналу.

Оптичні лінії зв'язку широко використовуються в різних сферах застосування, яке наприклад реалізується у вигляді хребта Інтернету, який являє собою основні маршрути передачі даних між великими стратегічно пов'язаними комп'ютерними мережами та основними маршрутизаторами Інтернету. Дана концепція також реалізується завдяки прокладанню підводних волоконно-оптичних кабелів. У сфері телекомунікацій телефонні компанії покладаються на волоконно-оптичні лінії зв'язку між центральними офісами, надаючи високошвидкісні послуги передачі голосу та даних споживачам і підприємствам. Так само в кабельному телебаченні оптичні волокна передають телевізійні сигнали високої чіткості в громадське житло. Великі центри обробки даних використовують оптичні лінії зв'язку для внутрішньої та зовнішньої передачі даних, забезпечуючи швидке та ефективно з'єднання між серверами та системами зберігання. У

мегаполісах міста оптичні лінії зв'язку використовуються для з'єднання різних локальних мереж, забезпечуючи високошвидкісний Інтернет і послуги зв'язку в межах столичного регіону. У сфері медичної візуалізації та діагностики оптичні волокна є невід'ємною частиною медичного обладнання, що дозволяє виконувати візуалізацію та діагностичні процедури з високою роздільною здатністю, прикладом якого є магнітно-резонансна томографія. Наведені приклади підкреслюють універсальність і важливість оптичних ліній зв'язку для підвищення можливостей підключення та передачі даних у різних секторах.

З одного боку, завдяки меншому загасанню та зменшенню перешкод оптичне волокно пропонує значні переваги перед мідним дротом для застосування на великих відстанях із високою пропускнуою здатністю. З іншого боку, розвиток інфраструктури в містах є відносно складним і трудомістким, а волоконно-оптичні системи є складнішими та дорожчими для встановлення та експлуатації. Через ці проблеми волоконно-оптичні системи зв'язку в основному впроваджуються для додатків на великі відстані, завдяки чому використовується їхня повна пропускна здатність, що виправдовує вищі витрати. Крім того, під час передачі сигналів у волоконно-оптичних мережах може виникнути загасання багатопроменевого поширення через вплив загасання сигналу та відстані передачі, більш того, електромагнітні хвилі сигналів легко піддаються впливу електричних і магнітних полів навколишнього середовища, що призводить до великої неоднозначності у вихідних сигналах зв'язку. Тому існує потреба в оптимізації конструкції передачі сигналів в волоконно-оптичних мережах зв'язку.

Таким чином, існує метод паралельної передачі сигналу і, зокрема, просторовий паралелізм, який полягає у передачі сигналів шляхом надсилання кількох сигналів одночасно через різні канали або частоти, що, у свою чергу може вирішити проблему швидкості передачі даних у оптично-волоконній лінії зв'язку.

Аналіз останніх досліджень і публікацій. У науково-дослідницькому просторі сьогодення з'являються роботи, присвячені винаходу та аналізу методології по розробці методів формування високої напруги шляхом каскадування підсилювачів та розширення робочої напруги окремих каскадів.

Дослідження роботи [1] було зосереджене на розробці методу точної передачі сигналів релейного захисту у волоконно-оптичних мережах зв'язку на основі змінного у часі придушення багатопроменевого завмирання та адаптивного формування променя. Система аналізує джерела сигналів перешкод бездротового сигналу болю на великій відстані, запроваджує технології захисту від перешкод, такі як двовимірні спільна обробка (Space Time Adaptive Processing – STAP), забезпечує алгоритми захисту від перешкод і пов'язаний аналіз посилення, а також проводить моделювання посилення обробки сигналу за допомогою мови програмування MATLAB. На підставі аналізу всебічних результатів моделювання при заданій довжині символу пропускна здатність сигналу збільшується, а коефіцієнт підсилення обробки нескінченно наближається до заданого теоретичного граничного значення, а не зростає нелінійно. Причина полягає в тому, що на канал впливає шум, а значення оцінки каналу та сполучене множення сигналу створюють квадратичний член шуму. У цей момент оцінене значення каналу когерентної області зменшується під впливом шуму, а втрата відношення сигнал/шум, спричинена квадратичним членом шуму, зменшується, тому посилення обробки збільшується. У процесі нескінченного збільшення смуги пропускання сигналу відношення потужності вхідного сигналу до шуму приймача прагне зменшитися до нескінченного значення, обмеженого розміром когерентної області. Значення оцінки каналу зростає під впливом шуму, а квадратичний член шуму є основним фактором, що впливає на потужність вихідного шуму. За умови більшої довжини символу у порівнянні з когерентним часом, діє наступне: чим менший максимальний доплерівський зсув частоти і чим більша область когерентного виявлення, тим більше коефіцієнт збільшення обробки даних.

Крім того, варто зазначити праці наступних науковців: Агте О., Образін Дж., Ана П. [2], Горан Н., Ходжич М., Олія А., Муйчич А. [3], Льв З., Шен Л., Лі Ю., Чжан М., Тіан Р., Руан Л., Чжан Ю., Чжао Дж. [4], Фіхтнер А., Богріс А., Боуден Д., Лентас До., Меліс Н., Нікас Т., Сімос До., Сімос І., Смолінський До. [5], Бортник Р., Кичак Ст., Василевський О., Васильківський М., Слободяник О., Романюк Р., Смолярж А., Єрлієва Б. [6], Деревяшкін Ст., Віркунін А., Горлов Н., Малашенко А., Сидоренко А. [7], Кононов С., Кононова О., Самотній І. [8], Ву Р., Ян Ф., Сунь Ю., Ченг Н., Ван Дж., Вей Ф., Гуй Ю., Хайвень К. [9], Нагшваріан Джахромі М., Кумар С., Дін М. Дж., Іва Т., Кімура К., Йосіда М., Хіроока Т., Наказова М. [10], Педяш Ст [11], Малюк І., Куліш Н. [12], Утацу Х., Оцука

Х. [13], Мехта Дж., Моралес Би., Росмейсл Дж., Дебінські Дж., Райландер К. [14], Като Т., Беппу С., Сома Д., Муранака Х., Окада С., Іріє Х., Вакаяма Ю., Ёсікане Н., Цурітані Т., Танака Ю., Хошида Т. [15] та інших.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з методологією по розробці поліпшених методів передачі даних з використанням волоконно-оптичних ліній зв'язку все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Постановка завдання. Метою роботи є розробка поліпшених методів передачі даних з використанням волоконно-оптичних ліній зв'язку.

Викладення основного матеріалу дослідження. При волоконно-оптичній передачі дані передаються за допомогою керованого середовища. Кероване середовище передачі використовує кабельну систему, яка спрямовує сигнали даних по певному шляху; такі носії також можна назвати зв'язаними носіями. Інформаційно-комунікаційний транспорт такого типу постійно нарощує швидкість і рік від року збільшує свою пропускну здатність. Пропускна спроможність оптичних систем зв'язку зі спектральним поділом може досягти 32-400 тбіт/с до 2037 року, а при використанні просторового мультиплексування навіть 5-100 Пбіт/с. Це показує, що паралельна передача сигналів у фізичному просторі може значно збільшити швидкість передачі у 150-250 разів.

Модифікації спектральної ефективності передачі даних по оптичному волокну ґрунтуються на загальному підході до визначення ефективності W технічної системи, яка оцінюється за допомогою співставлення величини досягнутого результату G виконаної операції із відносною кількістю використаних ресурсів H :

$$W = G/H \quad (1)$$

У ролі властивостей у даному випадку розглядаються пропускна здатність оптичного волокна R та агрегована спектральна ефективність передачі даних γ_{agr} :

$$G = \{G_i | i = \overline{1,2}, G_1 = R, G_2 = \gamma_{agr}\} \quad (2)$$

$$H = \{R_i | i = \overline{1,2}, R_1 = V, R_2 = B\} \quad (3)$$

де V – бітрейт а B – символна швидкість у бодах.

Після цього шукається результат проведеної операції з визначеним спектральним ресурсом S :

$$H = \{H_i | i = \overline{1,2}, H_1 = S, H_2 = A\} \quad (4)$$

$$S = \{S_i | i = \overline{1,2}, S_1 = F, S_2 = \Lambda\} \quad (5)$$

де F являє собою ширину діапазону частот, який використовується з метою передачі сигналів; Λ – ширина діапазону довжини хвиль.

Таким чином, просторовий ресурс залежить від того, якої площі є поперечний розріз оптичного волокна (optical fibre – OF) у зоні оптичної оболонки з діаметром D_{OF} :

$$A = \pi D_{OF}^2 / 4 \quad (6)$$

Наступним кроком є визначення паралельної передачі сигналів, яка відбувається шляхом надсилання кількох сигналів одночасно через різні канали або частоти. Одночасна передача кількох потоків даних збільшує загальну пропускну здатність мережі, дозволяючи передавати більше даних за той самий проміжок часу. Таким чином, просторове мультиплексування являє собою особливу форму паралельної передачі сигналу, при якій кілька оптичних каналів використовуються одночасно в одному оптоволоконному кабелі або пучку оптоволокон. Паралельна передача сигналів через різні канали або частоти збільшує загальну пропускну здатність мережі та дозволяє передавати більше даних за той самий проміжок часу. Ця паралельна передача сигналу реалізується шляхом застосування методів мультиплексування, таких як поляризаційне, спектрального, модального та ядерного мультиплексування.

Роль просторового мультиплексування полягає в тому, щоб ефективно і якомога більш унітарно поєднувати світло від пучка одномодового волокна на ядра мультіядерного волокна або моди маломодового, мультимодового або когерентного волокон, і навпаки. У випадку мультиплексування для одномодового мультіядерного або когерентного волокна механізм з'єднання є простим і відображає масив одномодових вхідних променів на масив одномодових вихідних променів.

У випадку маломодових волокон, мультиплексування поєднує просторово розділені одномодові вхідні промені на моди маломодових волокон. Цей зв'язок не обов'язково має відбуватися з будь-якою конкретною модовою основою, доки достатньою мірою зберігається ортогональність просторових каналів, тобто до тих пір, поки мультиплексування реалізує унітарне перетворення між його вхідними фундаментальними модами одномодового волокна і його вихідними керованими модами маломодового волокна. Кожен коефіцієнт мультиплексування K_i впливає на паралельність передачі інформації в оптичних мережах. До таких коефіцієнтів мультиплексування належать:

- $K_1 = p$ відноситься до кількості поляризацій, що відбулися;
- $K_2 = m$ означає кількість спектральних каналів у волоконно-оптичній системі передачі інформації зі спектральним розділенням;
- $K_3 = \mu$ дорівнює числу просторових мод, які передаються;
- $K_4 = \psi$ відповідає числу ядер у оптичному волокні.

Враховуючи бітрейт у спектральному каналі перед поляризаційним мультиплексуванням V_{in} , результуюча швидкість V_{OF} у оптично-волоконному кабелі зазвичай становить:

$$V_{OF} = V_{in} \times \prod_{i=1}^4 K_i | K_i = p, m, \mu, \psi \quad (7)$$

Надалі вводиться термін коефіцієнт паралелізму, який представляє збільшення бітної швидкості передачі даних. Відповідний параметр P представляє добуток використаних коефіцієнтів мультиплексування K_i :

$$P \times \prod_{i=1}^q K_i | q \leq q_{max} = 4 \quad (8)$$

У концептуальній схемі мультиплексування, де дані об'єднуються та передаються на різних етапах застосовуються чотири види паралелізму: поляризаційний, спектральний, модальний і просторовий. Сигнал спектрального каналу до поляризаційного мультиплексування з бітрейтом V_{in} реалізується за допомогою квадратурної амплітудної модуляції (Quadrature Amplitude Modulation – QAM), яка є методом модуляції, при якому і амплітуда, і фаза несучого сигналу модулюються для передачі кількох бітів на символ. Вона поєднує амплітудну модуляцію з фазовою модуляцією в одному складному сигналі. Змінюючи амплітуду та фазу, сигнали QAM можуть передавати кілька бітів в одному символі, що призводить до ефективного використання частотного спектру. Після першого етапу мультиплексування утворюються звичайні для високошвидкісних систем передачі даних формати модуляції з двома поляризаціями.

Наступні етапи мультиплексування утворюють просторовий паралелізм передачі даних в оптичному волокні за рахунок застосування просторового мультиплексування (SDM – space division multiplexing).

Коефіцієнти просторового паралелізму – модовий (mod), ядерний (cor) та комплексний (comp) відповідно:

$$P_M = \mu; P_{cor} = \psi; P_{spat} = P_{mod} \times P_{cor} = \mu \times \psi = M \times \quad (9)$$

де M відповідає кількості просторових каналів.

Наступним чином вводиться інтегральний коефіцієнт паралелізму, який визначається так:

$$P_{int} = p \times m \times \mu \times \psi \quad (10)$$

Приведені величини коефіцієнтів паралелізму слід розглядати в якості максимальних показників швидкості передачі даних задля оцінки пропускної здатності оптичного волокна.

Таким чином просторовий паралелізм впливає на величини бітрейту V та агрегованої спектральної ефективності γ_{agr} . Агрегована спектральна ефективність збільшується коефіцієнтом просторового паралелізму, а спектральна ефективність визначається відношенням каналної швидкості V_c до ширини каналного інтервалу F_{ci} :

$$\gamma_{agr} = \gamma \times P_{spat} \quad (11)$$

$$\gamma = V_c / F_{ci} \quad (12)$$

По такому ж принципу спектрально-хвильова γ_λ ефективність визначається відношенням каналної швидкості V_c до хвильового каналного інтервалу Λ_{ci} , а агрегована спектрально-хвильова ефективність $\gamma_{\lambda agr}$ збільшується коефіцієнтом просторового паралелізму P_{spat} .

Таким чином агрегована швидкість передачі даних по оптичному волокну визначається за допомогою формули:

$$V_{OF} = \begin{cases} F_{ci} \times \gamma \times m \times \mu \times \psi \\ \Lambda_{ci} \times \gamma_\lambda \times m \times \mu \times \psi \end{cases} \quad (13)$$

При застосуванні умови, що канална швидкість дорівнює:

$$V_c \times m = F_{ci} \times \gamma \times m = \Lambda_{ci} \times \gamma_\lambda \times m \quad (12)$$

діє наступне:

$$V_{OF} = V_c \times m \times P_{spat} \quad (12)$$

При гіпотетично наданих даних, що канална швидкість $V_c = 100$ Гбіт/с, ширина каналного інтервалу $F_{ci} = 50$ Гц, хвильовий каналний інтервал $\Lambda_{ci} = 0,4$ нм, $m, \mu, \psi = 2$ діють наступні обчислення:

$$\gamma = \frac{V_c}{F_{ci}} = \frac{100}{50} = 2 \text{ біт/Гц} \quad (13)$$

$$\gamma_{agr} = \gamma \times P_{spat} = 2 \times 4 = 8 \text{ біт/Гц} \quad (14)$$

$$V_{OF} = F_{ci} \times \gamma_{agr} \times m \times \mu \times \psi = 50 \times 8 \times 2 \times 2 \times 2 = 3200 \text{ Гбіт/с} \quad (15)$$

Отже, за допомогою просторового мультиплексування сукупна швидкість передачі даних через оптичне волокно стає 3200 Гбіт/с, що значно перевищує швидкість одноканальної передачі 100 Гбіт/с.

Висновки. У даному дослідженні був запропонований метод паралельної передачі даних, який ґрунтується на методі мультиплексуванні із використанням просторового паралелізму. Загалом, модернізація традиційної одномодової оптоволоконної системи за допомогою підсилювачів з мультиплексуванням просторового розподілу може дати значні переваги, якщо багатоядерні підсилювачі досягнуть підвищення ефективності

Розглядається кількісна оцінка внеску просторового мультиплексування у швидкість передачі даних і спектральну ефективність. У роботі підкреслюється концепція просторового мультиплексування як форми паралельної передачі сигналу, при якій кілька оптичних каналів одночасно використовуються в одному оптичному волокну або пучку волокон. Загалом, було продемонстровано, як методи просторового мультиплексування сприяють ефективному використанню ресурсів оптичного волокна, що призводить до значного покращення швидкості передачі та спектральної ефективності, прокладаючи таким чином шлях для розвитку високошвидкісних мереж передачі даних.

Список бібліографічного опису

1. Wu J., Jin C., Wang Z. Linear Anti-interference Algorithm for Digital Signal Transmission in Fiber Optic Communication Networks based on Link Analysis. *Scalable Computing: Practice and Experience*. 2024. №25. P. 920-927. DOI:10.12694/scpe.v25i2.2607.
2. Atte O., Obidinnu J., Ana P. Fiber Optic Transmission: Architectures, Technologies and Innovations, Applications, Types, Testing and Troubleshooting. 2021. №8. P. 88-93.
3. Goran N., Hodzic M., Maslo A., Mujcic A. Modeling and Simulation of Fiber Optic Transmission Links. *Science, Engineering and Technology*. 2022. №2. P. 16-23. DOI:10.54327/set2022/v2.i1.22.
4. Lv Z., Sheng L., Li Y., Zhang M., Tian G., Ruan L., Zhang Y., Zhao J. An analog fiber optic link developed for electrical pulse signal transmission. *Review of Scientific Instruments*. 2023. №94. DOI:10.1063/5.0132828.
5. Fichtner A., Bogris A., Bowden D., Lentas K., Melis N., Nikas T., Simos C., Simos I., Smolinski K. Sensitivity kernels for transmission fiber optics. *Geophysical Journal International*. 2022. №231. DOI:10.1093/gji/ggac238.
6. Bortnyk G., Kychak V., Vasilevskyi O., Vasylykivskyi M., Slobodyanik A., Romaniuk R., Smolarz A., Yeraliyeva B. Digital restoration of signals in fiber optic transmission systems. *Proceedings of SPIE - The International Society for Optical Engineering*. (2022).
7. Derevyashkin V., Virkunin A., Gorlov N., Malashenko A., Sidorenko A. Marine-based Fiber-optic Systems for Data Collection and Transmission. *The Herald of the Siberian State University of Telecommunications and Information Science*. 2023. №18. P. 48-58. DOI:10.55648/1998-6920-2024-18-1-48-58.
8. Kononov V., Kononova O., Odinkii I. Designing The Principal Transmission Scheme Fiber Optic Device Fiber Optical Gyroscope. 2019. №6. P.28-32. DOI:10.26906/SUNZ.2019.6.028.
9. Wu R., Yang F., Sun Y., Cheng N., Wang J., Wei F., Gui Y., Haiwen C. Absolute phase marking technology and fiber-optic remote coherent phase transmission. *Optics Express*. 2021. №29. DOI:10.1364/OE.419695.
10. Naghshvarian Jahromi M., Kumar S., Deen M.J., Iwaya T., Kimura K., Yoshida M., Hirooka T., Nakazawa M. Software-Defined Fiber Optic Communications for Ultrahigh-Speed Optical Pulse Transmission Systems. *IEEE Journal of Selected Topics in Quantum Electronics*. 2022. №28. P. 1-11. DOI:10.1109/JSTQE.2022.3190885.
11. Pedyash V. Mathematical Modeling of Fiber-Optic Transmission System With Intensity Modulation. *Herald of Khmelnytskyi National University. Technical sciences*. 2023. №317. P. 162-166. DOI:10.31891/2307-5732-2023-317-1-162-166.
12. Malysh I., Kulish N. Fiber Optic Information Transmission Systems. *The National Transport University Bulletin*. 2021. №1. P. 203-207. DOI:10.33744/2308-6645-2021-1-48-203-207.
13. Utatsu H., Otsuka H. Performance Analysis of Fiber-Optic Relaying With Simultaneous Transmission and Reception on the Same Carrier Frequency. *IEICE Transactions on Communications*. 2019. DOI:10.1587/transcom.2018EBP3298.
14. Mehta J., Morales B., Rossmeisl Jr J., Debinski W., Rylander C. Solid Fiber Inside of Capillary and Modified Fusion-Spliced Fiber Optic Microneedle Devices for Improved Light Transmission Efficiency. 2022. №16. DOI:10.1115/1.4055607].
15. Kato T., Beppu S., Soma D., Muranaka H., Okada S., Irie H., Wakayama Y., Yoshikane N., Tsuritani T., Tanaka Y., Hoshida T. U-band WDM Transmission over 90-km Deployed Fiber-optic Cable Leveraged by S+C+L-band WDM Channels. 2024. DOI:10.1364/OFC.2024.W4D.3.
16. Klaus W., Winzer P., Nakajima K. The Role of Parallelism in the Evolution of Optical Fiber Communication Systems. *Proceedings of the IEEE*. 2022. №110. P. 1619-1654. DOI:10.1109/JPROC.2022.3207920.

References

17. Wu J., Jin C., Wang Z. Linear Anti-interference Algorithm for Digital Signal Transmission in Fiber Optic Communication Networks based on Link Analysis. *Scalable Computing: Practice and Experience*. 2024. №25. P. 920-927. DOI:10.12694/scpe.v25i2.2607.
18. Atte O., Obidinnu J., Ana P. Fiber Optic Transmission: Architectures, Technologies and Innovations, Applications, Types, Testing and Troubleshooting. 2021. №8. P. 88-93.
19. Goran N., Hodzic M., Maslo A., Mujcic A. Modeling and Simulation of Fiber Optic Transmission Links. *Science, Engineering and Technology*. 2022. №2. P. 16-23. DOI:10.54327/set2022/v2.i1.22.
20. Lv Z., Sheng L., Li Y., Zhang M., Tian G., Ruan L., Zhang Y., Zhao J. An analog fiber optic link developed for electrical pulse signal transmission. *Review of Scientific Instruments*. 2023. №94. DOI:10.1063/5.0132828.
21. Fichtner A., Bogris A., Bowden D., Lentas K., Melis N., Nikas T., Simos C., Simos I., Smolinski K. Sensitivity kernels for transmission fiber optics. *Geophysical Journal International*. 2022. №231. DOI:10.1093/gji/ggac238.
22. Bortnyk G., Kychak V., Vasilevskyi O., Vasylykivskyi M., Slobodyanik A., Romaniuk R., Smolarz A., Yeraliyeva B. Digital restoration of signals in fiber optic transmission systems. *Proceedings of SPIE - The International Society for Optical Engineering*. (2022).
23. Derevyashkin V., Virkunin A., Gorlov N., Malashenko A., Sidorenko A. Marine-based Fiber-optic Systems for Data Collection and Transmission. *The Herald of the Siberian State University of Telecommunications and Information Science*. 2023. №18. P. 48-58. DOI:10.55648/1998-6920-2024-18-1-48-58.
24. Kononov V., Kononova O., Odinkii I. Designing The Principal Transmission Scheme Fiber Optic Device Fiber Optical Gyroscope. 2019. №6. P.28-32. DOI:10.26906/SUNZ.2019.6.028.
25. Wu R., Yang F., Sun Y., Cheng N., Wang J., Wei F., Gui Y., Haiwen C. Absolute phase marking technology and fiber-optic remote coherent phase transmission. *Optics Express*. 2021. №29. DOI:10.1364/OE.419695.

26. NaghshvarianJahromi M., Kumar S., Deen M.J., Iwaya T., Kimura K, Yoshida M., Hirooka T., Nakazawa M. Software-Defined Fiber Optic Communications for Ultrahigh-Speed Optical Pulse Transmission Systems. *IEEE Journal of Selected Topics in Quantum Electronics*. 2022. №28. P. 1-11. DOI:10.1109/JSTQE.2022.3190885.
27. Pedyash V. Mathematical MOdeling of Fiber-Optic Transmission System With Intensity Modulation. *Herald of Khmelnytskyi National University. Technical sciences*. 2023. №317. P. 162-166. DOI:10.31891/2307-5732-2023-317-1-162-166.
28. Malysh I., Kulish N. Fiber Optic Information Transmission Systems. *The National Transport University Bulletin*. 2021. №1. P. 203-207. DOI:10.33744/2308-6645-2021-1-48-203-207.
29. Utatsu H., Otsuka H. Performance Analysis of Fiber-Optic Relaying With Simultaneous Transmission and Reception on the Same Carrier Frequency. *IEICE Transactions on Communications*. 2019. DOI:10.1587/transcom.2018EBP3298.
30. Mehta J., Morales B., Rossmesl Jr J., Debinski W., Rylander C. Solid Fiber Inside of Capillary and Modified Fusion-Spliced Fiber Optic Microneedle Devices for Improved Light Transmission Efficiency. 2022. №16. DOI:10.1115/1.4055607].
31. Kato T., Beppu S., Soma D., Muranaka H., Okada S., Irie H., Wakayama Y., Yoshikane N., Tsuritani T., Tanaka Y., Hoshida T. U-band WDM Transmission over 90-km Deployed Fiber-optic Cable Leveraged by S+C+L-band WDM Channels. 2024. DOI:10.1364/OFC.2024.W4D.3.
32. Klaus W., Winzer P., Nakajima K. The Role of Parallelism in the Evolution of Optical Fiber Communication Systems. *Proceedings of the IEEE*. 2022. №110. P. 1619-1654. DOI:10.1109/JPROC.2022.3207920.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-37>

УДК 004.4'2

Дейнеко Лев Миколайович, магістр

<https://orcid.org/0009-0000-7203-1883>

Кунанець Наталія Едуардівна, д.н.соц.комун., професор

<https://orcid.org/0000-0003-3007-2462>

Національний університет «Львівська політехніка», м. Львів, Україна

РОЛЬ UML-ДІАГРАМ У ПЛАНУВАННІ ПРОЄКТУ ІНФОРМАЦІЙНОЇ СИСТЕМИ НА ПРИКЛАДІ СИСТЕМИ ПІШИХ ТУРИСТИЧНИХ МАРШРУТІВ

Дейнеко Л.М., Кунанець Н.Е. Роль UML-діаграм у плануванні проєкту інформаційної системи на прикладі системи піших туристичних маршрутів. Будь-який процес вимагає попереднього аналізу та планування. Сучасні інформаційні системи стають шоразу більш складними, а, відповідно, етап планування потребує додаткового часу та уваги. Одним з інструментів, що активно використовується фахівцями з розроблення IT-продукції, є UML - мова моделювання, що дозволяє описати та відобразити логіку системи. У цій статті розглянуто використання UML-діаграм на етапі планування та вплив даного інструменту на загальний результат розробленого рішення. Дане дослідження базується на прикладі системи піших туристичних маршрутів та розробці окремої реальної системи. Обґрунтовано важливість використання такого підходу в інженерії вимог, що дозволить прослідкувати залежність між наявністю моделей та якістю системи. У даній роботі проаналізовано не тільки переваги моделювання, а й виклики та складнощі, що можуть виникати під час аналізу нефункціональних вимог до системи. Це дослідження дозволило підсумувати роль UML-діаграм та окреслити основоположні принципи, що значною мірою покращать досвід використання такого підходу на різних етапах життєвого циклу інформаційних систем. Дана робота закладає підвалини для подальших досліджень та дозволить масштабувати вже здобуті результати для виявлення додаткових особливостей і деталей застосування моделювання. Варто підкреслити, що такі дослідження можуть стосуватися різних областей та аспектів UML, наприклад, у якості інструменту для пошуку недоліків чи покращення комунікації в команді спеціалістів загалом. Отримані результати можна застосувати у сфері тестування, бізнес-аналізу та під час розробки архітектури програмних рішень.

Ключові слова: UML; моделювання; діаграма; інформаційна система; дефект; планування.

Deineko L., Kunanets N. The role of UML diagrams in planning an information system project on the example of a hiking tourist routes system. Any process requires preliminary analysis and planning. Modern information systems are getting more and more complex, and, accordingly, the planning stage requires additional time and attention. One of the tools actively used by specialists in the development of IT products is UML – a modeling language that allows you to describe and display the logic of the system. This article outlines the use of UML diagrams at the planning stage and its impact on the overall result of the developed solution. This study is based on the example of a hiking tourist routes system and the development of a separate real solution. The importance of using such an approach in requirements engineering is substantiated, which will enable tracking the dependence between models' availability and the quality of the system. This paper analyzes not only the advantages of modeling but also the challenges and difficulties that may arise during the analysis of non-functional requirements for the system. This research made it possible to summarize the role of UML diagrams and outline the basic principles that will significantly improve the experience of using this approach at various stages of the information systems life cycle. This work lays the foundation for further research and will enable scaling the already obtained results to reveal additional features and details of the modeling application. It is worth emphasizing that such studies can pertain to different areas and aspects of UML, such as a tool for identifying the flaws or improving communication in a team of engineers in general. The obtained results can be applied in the field of testing, business analysis, and during the development of the architecture of software solutions.

Keywords: UML; modeling; diagram; information system; defect; planning.

Постановка проблеми. Інформаційні системи (ІС) стали невід'ємною складовою сучасного життя, проникаючи у різні сфери діяльності. Вони значно полегшують та видозмінюють велику кількість процесів – від електронної комерції до управління великими корпораціями, від медичних засобів та устаткування до автомобільних технологій [6]. Зростання складності ІС та потреби у якісному управлінні проєктами приводять до пошуку ефективних інструментів планування, аналізу та розробки ІС.

Серед таких інструментів можна виокремити «уніфіковану мову моделювання» (UML), що зарекомендувала себе як один з найефективніших та найпоширеніших способів моделювання різного роду систем. Даний засіб не просто став стандартом у світі розробки програмного забезпечення, а також є невід'ємною частиною інженерії вимог, їх аналізу та проєктування ІС [2].

Існує чимало розбіжностей щодо практичного застосування цієї мови та відмінностей, що накладаються іншими мовами графічного моделювання. Для спрощення ситуації можна виокремити 3 найпопулярніші сценарії застосування UML: підготовка ескізу, робота над проєктом та, власне, програмування. У першому випадку UML використовується для загальної візуалізації між деякими

компонентами системи. У другому – під час планової або зворотної розробки, тобто підготовка UML-діаграм перед початком написання коду, або ж навпаки – побудова діаграми на основі вже існуючого коду (системи) для кращого його розуміння [2].

Для більш ефективної відповіді на поточні виклики, ця робота базуватиметься на дослідженні та аналізі ролі UML-діаграм під час життєвого циклу програмного забезпечення на прикладі проєкту інформаційної системи піших туристичних маршрутів. Огляд практичного застосування, оцінка позитивних аспектів та виявлення можливих ризиків допоможуть відповісти на різноманітні питання, наприклад: «Як зрозуміти необхідний обсяг робіт та як вони вплинуть на імплементацію?», або «Які ризики можуть виникнути під час роботи над конкретною частиною застосунку?».

Аналіз останніх досліджень та публікацій. Моделювання за допомогою UML не є новий підходом у сфері розробки програмного забезпечення. Існує чимало досліджень, що детальніше розглядають застосування UML, особливості такого процесу та його переваги.

У своїй роботі С. Денисенко [8] вивчає проблеми, що виникають під час формалізації бізнес-процесу. Дослідження спирається на техніки IDEF та UML, де UML здобуває перевагу на етапах опрацювання складної та деталізованої інформації. Український професор та доктор технічних наук Є. Буров [7] запропонував інтелектуальну систему проєктування та керування бізнес-процесами. Це дослідження охоплює різні способи моделювання включно з UML та аналізує процес адаптації систем задля зменшення витрат на супровід. У науковій публікації В. Яковина та Ю. Парфенюк [9] проаналізували UML як засіб прогнозування надійності програмного забезпечення, а також розробили підхід для визначення надійності на етапі описання архітектури. Запропонований метод дозволяє оцінити ймовірність сценаріїв використання програмної системи.

На основі проведеного аналізу можна зробити висновок, що сучасні дослідження підкреслюють різні аспекти моделювання та його застосування у багатьох сферах. Однак, досі є велика потреба у ґрунтовних дослідженнях, які б прослідкували роль моделювання за допомогою UML на етапі планування проєкту та порівняли отримані показники з кінцевими результатами впровадження.

Обґрунтування важливості UML-діаграм у плануванні проєктів. «Планування» є першим етапом розробки у розрізі життєвого циклу продукту. Відповідно, планування визначає якість виконаного продукту і є фундаментальним етапом у розробленні ПЗ. Ефективно виконане планування дозволяє виявити й усунути можливі ризики та проблеми на ранніх етапах розроблення, сприяючи успішному завершенню проєкту в обумовлений термін та бюджет. Недоліки у плануванні можуть призвести до затримок, перевищення бюджету (через залучення додаткових ресурсів) та низької якості продукту.

Таке твердження можна аргументувати дослідженням у роботі Хатчінсона [3], який стверджував, що перші 20% часу на проєкті присвячуються розробці моделей. У такому випадку обсяги моделювання можна відобразити наступним способом (Рис. 1):

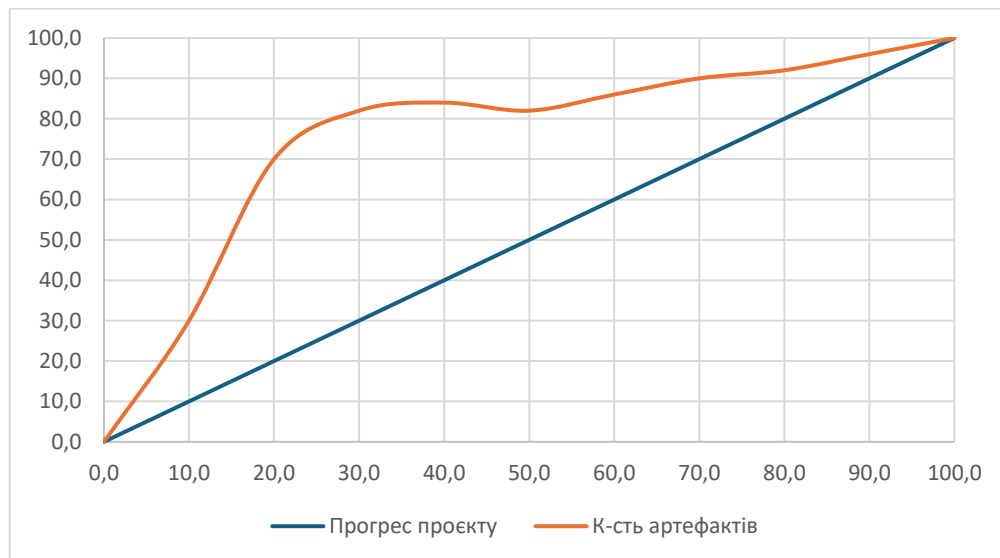


Рис. 1. Обсяги артефактів під час імплементації системи

Граничні значення прогресу та кількості артефактів прийняті за 100%. Оскільки переважна більшість моделювання відбувається на ранніх етапах, то в подальшому змін дуже мало, що свідчить про те, що впродовж цього періоду робота полягає переважно у внесенні правок до існуючих діаграм. Отже, графік можна інтерпретувати наступним чином: зазвичай, проект розпочинається з визначення загального бачення продукту, і тільки згодом додаються деталі та незначні зміни.

Для кращого розуміння такого підходу необхідно проаналізувати фактичний стан інформаційної системи, що розроблялася з червня 2023 року до квітня 2024 (Рис. 2). У даній ситуації не було надано належної уваги використанню UML-моделювання, відповідно, ризик недоліків значно більший і це підтверджує діаграма на Рисунку 2. Незважаючи на спроби команди виправити недоліки, загальна кількість неточностей в системі невпинно зростає і обсяг технічного боргу збільшується. Проаналізувавши обидві діаграми (Рис.1 та Рис. 2), цілком доцільним буде припущення, що брак належного моделювання на етапі планування погіршує загальну якість продукту, що, у свою чергу, впливає на подальшу розробку та супутні витрати. Отже, нижча якість розроблених моделей також негативно впливає на якість коду.

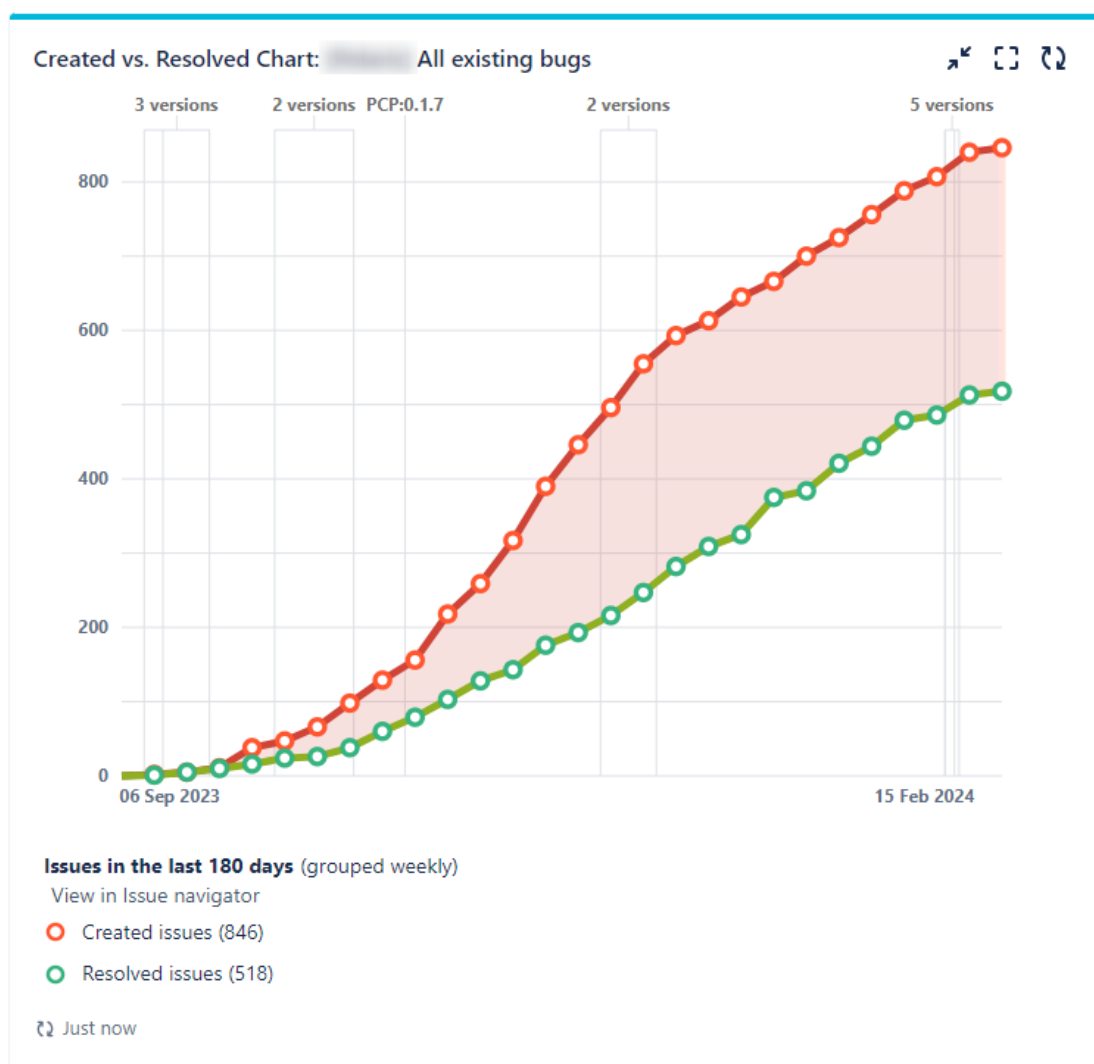


Рис. 2. Співвідношення загальної кількості дефектів та вирішених недоліків

Враховуючи вищесказане, розглянемо приклад діаграми використання інформаційної системи піших туристичних маршрутів (Рис. 3).

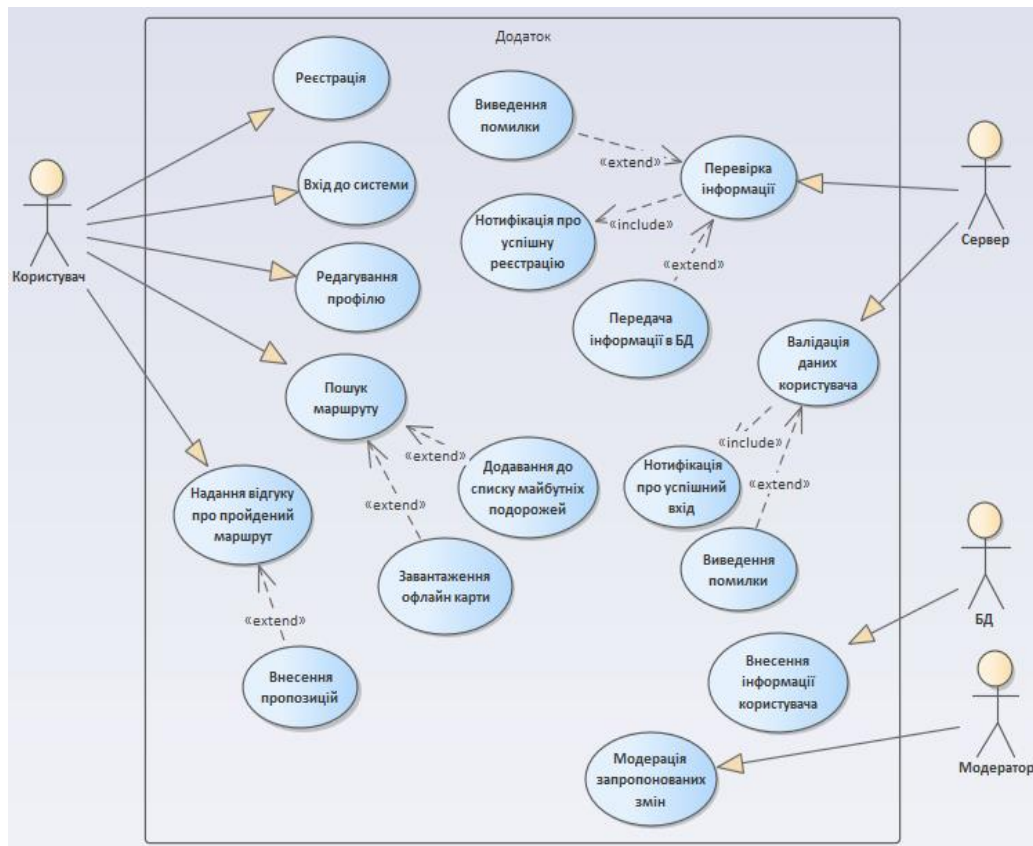


Рис. 3. Діаграма використання інформаційної системи піших туристичних маршрутів

Діаграма використання ІС піших туристичних маршрутів є одним з ключових інструментів для подальшого розроблення: вона ідентифікує основних акторів, відображає їх головні взаємодії через варіанти використання (use cases). Тобто, якщо було отримано даний артефакт на початку проєкту – це дозволить використати діаграму для ідентифікації функціональних вимог, визначити основні сценарії, що стане базою для подальшого аналізу, планування розробки та потенційного зменшення неточностей.

Наступні кроки повинні враховувати дуже важливий аспект, визначення на якому етапі розроблених моделей буде достатньо. Це вимагатиме певного компромісу щодо зусиль, які замовник готовий вкласти у високоякісні і деталізовані артефакти та додаткового ресурсу, що необхідний для виправлення неточностей у моделях. Вартість змін продемонстровано на графіку (Рис. 4). Відтак, можна зробити висновок, що економічно вигідно виправляти будь-які недоліки якомога швидше. Типовими видами недоліків у моделях UML є невідповідності, неповнота та навіть дублювання [4]. Якщо враховувати, що однією з головних цілей UML-діаграм є донесення спільного та однозначного контексту до усіх членів команди – це доволі серйозний показник того, що однозначність та завершеність моделей підлягає постійному контролю.

Для дотримання практики збалансування розподілу зусиль між якістю моделювання та зниженням ризиків необхідно впроваджувати належний рівень тестування, що сприятиме правильній витраті ресурсів. Наприклад, критичні і складніші компоненти потребуватимуть детальнішого тестування, а простіші – переважно поверхневого. Це вимагає правильної класифікації та розрізнення критичних елементів системи, що більше належить до площини тестування і належної комунікації між командою та бізнесом, де часто постає питання правильного виваження знайдених недоліків. Тобто, моделювання напряму впливатиме на хід роботи усього проєкту, а правильний розподіл зусиль допоможе визначити належний рівень кількості артефактів та їх якість.

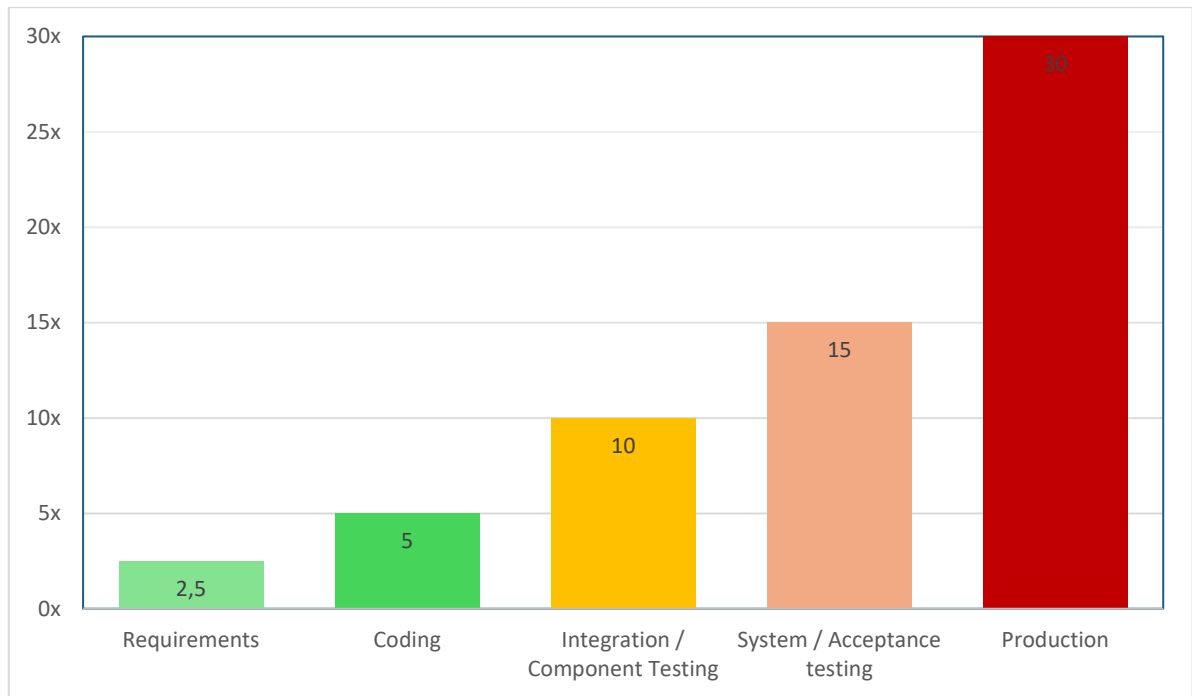


Рис. 4. Ріст вартості змін під час життєвого циклу розробки ПЗ

У книзі С. Макконелла наведено порівняльну характеристику різних засобів, що сприяють виявленню дефектів [5]. З точки зору тестування ПЗ, моделювання можна трактувати як один з можливих та ефективних способів пошуку неточностей під час реалізації життєвого циклу системи (Рис. 5).

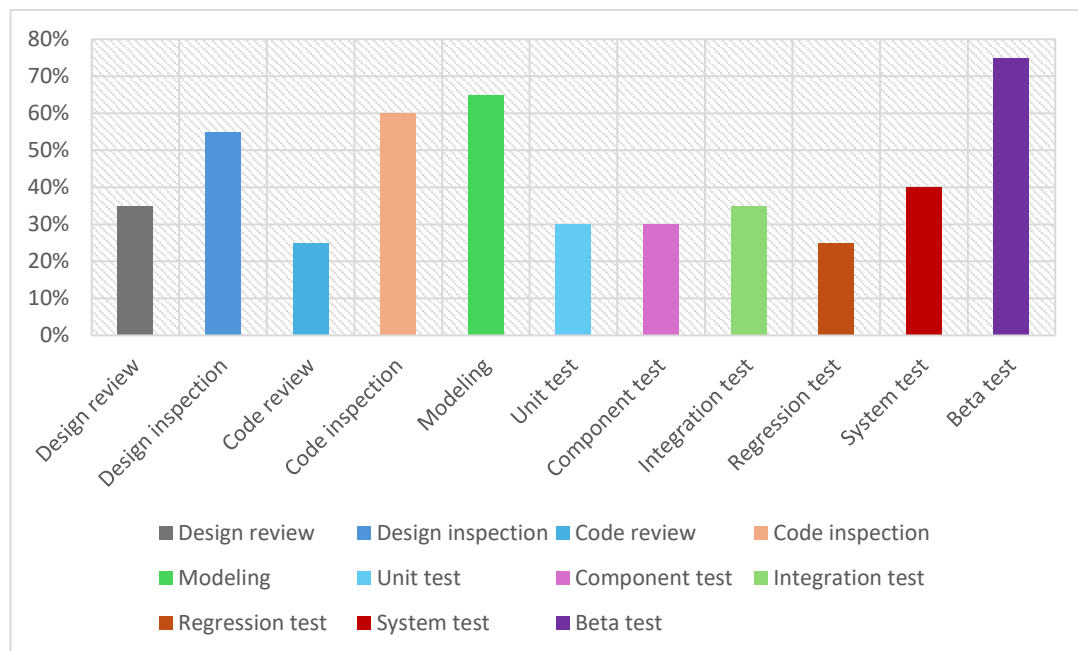


Рис. 5. Ефективність методів пошуку дефектів

Оскільки моделі розробляються переважно перед початком імплементації, їх можна використати для ґрунтовного оцінювання дизайну системи, а не чекати завершення етапу розробки, коли дефекти почнуть проявлятися під час модульного та інтеграційного тестування.

Окрім переваг моделювання, доцільно також висвітлити і деякі його недоліки, що можуть по-різному впливати на процес впровадження та на продукт загалом. Повертаючись до раніше зазначеної діаграми використання інформаційної системи піших туристичних маршрутів (Рис. 3)

можна припустити, що така модель не враховує мотивацію та досвід потенційних кінцевих користувачів. Тобто, розроблений продукт може цілком і повністю відповідати усім вимогам, але не гарантовано, що цей продукт задовольнить кінцевого користувача з погляду валідації. Також варто розглянути діаграму активності даної ІС (Рис. 6). У цьому випадку додаток розроблявся з урахуванням наступних можливостей:

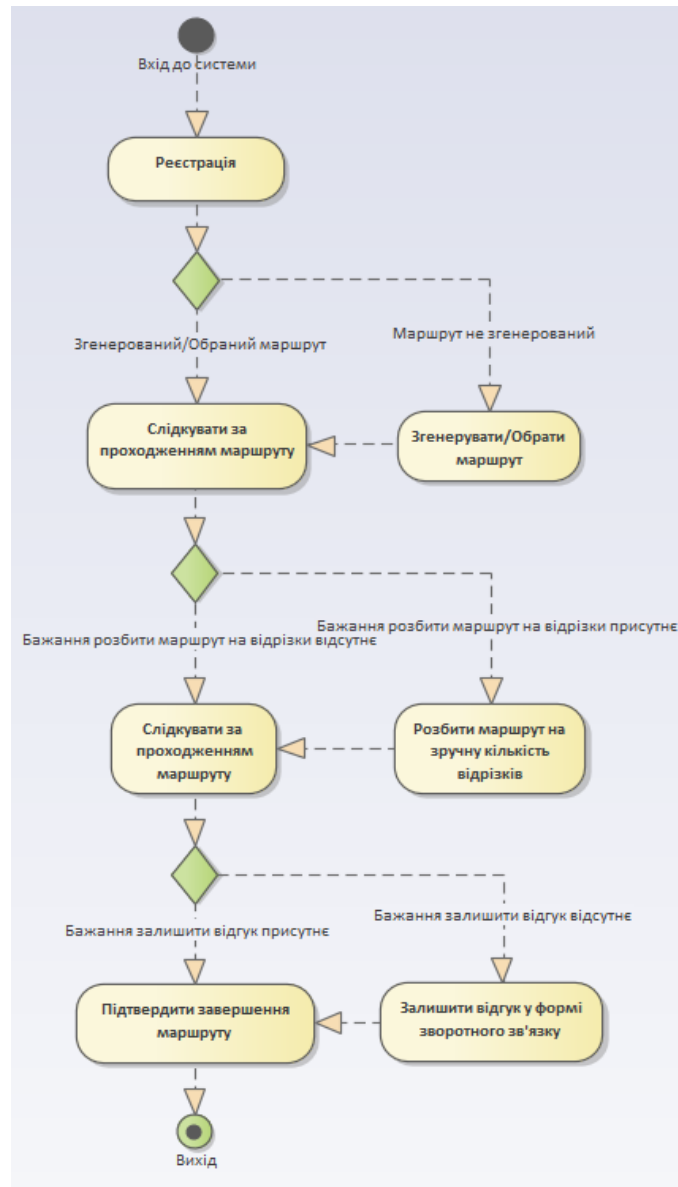


Рис. 6. Діаграма активності системи

Користувач:

- заходить в додаток та реєструється/вводить дані наявного профілю;
- генерує необхідний маршрут;
- у разі потреби розділяє маршрут на декілька етапів/відрізків;
- при бажанні залишає відгук стосовно згенерованого/пройденого маршруту.

Система:

- опрацьовує дані зареєстрованого/наявного користувача;
- відображає згенерований маршрут/можливі альтернативні шляхи;
- відображає стан проходження обраного маршруту;
- приймає відгук користувача через форму зворотного зв'язку.

У певній мірі це дає загальне розуміння обсягу робіт та можливих способів досягнення результату, але на даному етапі розробки немає впевненості, що це остаточна версія, котра не потребуватиме подальших покращень опісля досліджень із користувачами.

До цього списку також входить і зручність використання. Не існує систематизованого та узагальненого підходу щодо покриття нефункціональних вимог у контексті діаграм використання. Діаграми класів переважно використовуються для відображення системи в статистиці. Однак, часом необхідно проаналізувати і динамічні аспекти роботи системи, що, вочевидь, вимагатиме додаткових зусиль. Діаграми послідовності відображають переважно один загальний сценарій, відповідно, у випадках зі складною логікою системи необхідно збирати різні діаграми даного типу разом, щоб відобразити різні можливі варіанти поведінки. Це означає, що інтеграція сценаріїв є однією з критичних проблем у використанні моделювання цього типу. Діаграми послідовності потребують об'єктів у вигляді вхідних даних та станів об'єктів – як вихідних, а це означає, що діаграма класів впливатиме на процес створення такої моделі.

На основі проведеного дослідження можна виокремити декілька важливих принципів у моделюванні. По-перше, «вибір моделі має фундаментальний вплив на формування вирішення задачі» [1]. Тобто, вкрай важливо правильно обрати модель – це підкреслить проблемні місця та рішення, що можна не помітити за інших умов. Ретельне та повторюване тестування моделей дозволить досягнути значно вищого рівня впевненості у тому, що змодельоване рішення поводитиметься максимально близько до очікуваного результату. У такому випадку більшість нестандартних та неочевидних випадків стане помітною для всіх членів команди, що дозволить провести додаткові обговорення та запланувати необхідні дії. Слід відзначити, що спеціальність колег також впливатиме на вихідний результат обраної моделі у розробці ПЗ, адже кожен член команди матиме власне бачення, що базуватиметься на персональному досвіді та кваліфікації. Проте, за умов правильно організованої співпраці це буде радше перевагою, ніж недоліком.

По-друге, «будь-яка діаграма може мати різні рівні трактування» [1]. У будь-якому випадку найкращим вибором буде той, що дозволить застосувати необхідний ступінь деталізації, оскільки аналітики, кінцеві користувачі чи розробники цікавитимуться певними аспектами інформаційної системи у різний час.

По-третє, «найкращим варіантом є той, що найбільш близький до реальності» [1]. Якщо модель тільки частково передає справжній стан речей – користі з неї дуже мало. Найкраще обирати моделі, що максимально передають реальний хід подій, а також шукати конкретні чинники, що негативно впливають на це. Спрощення реальності є хорошим фактором, але тільки за умови, що такі спрощення не приховують жодних важливих деталей. Нездатність подолати такий аспект призводить до того, що з часом фактична реалізація системи все більше відрізнятиметься від першопочаткового задуму.

По-четверте, «жодна діаграма не може вичерпно покрити систему. Краще для будь-якої нетривіальної системи розробляти невеликий набір майже незалежних моделей» [1]. В практичному застосуванні ідея полягає у тому, щоб наявні діаграми залишалися взаємопов'язаними, але водночас дозволяли достатньо детально вивчити конкретну сферу. Певною мірою це альтернативне рішення діаграм декомпозиції, що розбивають складні компоненти на менші та зрозуміліші.

Висновки та перспективи подальшого дослідження. Проведений аналіз засвідчує та обґрунтовує важливість використання UML-діаграм як одного з інструментів моделювання та документування, що посідає чільне місце у роботі над проектами та розробкою рішень. Додаткові приклади базуються на інформаційній системі піших туристичних маршрутів. Результати дослідження окреслюють значення та роль UML в інженерії вимог на різних етапах життєвого циклу. Детальне вивчення результатів впровадження системи без активного використання UML-моделювання засвідчує значну кількість недоліків, що негативно впливатиме на супутні витрати.

Окреслено принципи моделювання, що слугують коротким і водночас зрозумілим посібником з успішного застосування моделювання.

Ця робота дає змогу покращити вже набуті знання та закладає підвалини для нового аналізу сфери моделювання у контексті роботи над інформаційними системами.

Напрямок подальших досліджень є масштабування вже здобутих результатів, розширення вибірки проектів та інформаційних систем, що допоможе виявити додаткові особливості використання моделювання під час розроблення програмних рішень.

Список бібліографічного опису

1. Booch, G., Rumbaugh, J.E., & Jacobson, I. (2005). *The Unified Modeling Language User Guide*. Addison-Wesley.
2. Fowler, M. (2004). *UML distilled: A brief guide to the standard object modeling language*. Addison-Wesley.
3. Hutchinson, J., Whittle, J., Rouncefield, M., & Kristoffersen, S. (2011). Empirical assessment of MDE in industry. In *Proceedings of the 33rd International Conference on Software Engineering* (pp. 471-480). New York, NY; ACM.
4. Lange, C., Chaudron, M.R.V., Muskens, J., Somers, L.J., & Dortmans, H.M. (2003). An empirical investigation in quantifying inconsistency and incompleteness of UML designs. In *Proceedings of the Workshop on Consistency Problems in UML-based Software Development* (pp. 26-34). San Francisco, CA; Blekinge Institute of Technology.
5. McConnell, S. (2004). *Code Complete: A practical handbook of software construction* (2nd ed.). Microsoft Press.
6. Rainer, R. K., & Prince, B. (2019). *Introduction to information systems: Supporting and Transforming Business* (8th ed.). John Wiley & Sons, Inc.
7. Буров, С. (2008). Система моделювання інтелектуальної мережі бізнес-процесів. *Вісник Національного університету «Львівська політехніка»*, 610, 34 - 39.
8. Денисенко, С. (2010). Проблемні моменти моделювання бізнес-процесів з метою проведення подальшого реінжинірингу на базі стандартів IDEF та UML-моделювання. *Вісник Національного університету «Львівська політехніка»* 691, 41 - 46.
9. Яковина, В., Парфенюк, Ю. (2013). Використання засобів UML для прогнозування надійності програмного забезпечення на етапі його проектування. *Вісник Національного університету «Львівська політехніка»* 773, 151 - 156.

References

1. Booch, G., Rumbaugh, J.E., & Jacobson, I. (2005). *The Unified Modeling Language User Guide*. Addison-Wesley.
2. Fowler, M. (2004). *UML distilled: A brief guide to the standard object modeling language*. Addison-Wesley.
3. Hutchinson, J., Whittle, J., Rouncefield, M., & Kristoffersen, S. (2011). Empirical assessment of MDE in industry. In *Proceedings of the 33rd International Conference on Software Engineering* (pp. 471-480). New York, NY; ACM.
4. Lange, C., Chaudron, M.R.V., Muskens, J., Somers, L.J., & Dortmans, H.M. (2003). An empirical investigation in quantifying inconsistency and incompleteness of UML designs. In *Proceedings of the Workshop on Consistency Problems in UML-based Software Development* (pp. 26-34). San Francisco, CA; Blekinge Institute of Technology.
5. McConnell, S. (2004). *Code Complete: A practical handbook of software construction* (2nd ed.). Microsoft Press.
6. Rainer, R. K., & Prince, B. (2019). *Introduction to information systems: Supporting and Transforming Business* (8th ed.). John Wiley & Sons, Inc.
7. Burov, Ye. (2008). Systema modeliuвання intelektualnoi merezhi biznes-protsesiv. *Visnyk Natsionalnoho universytetu «Lvivska politekhnik»*, 610, 34 - 39.
8. Denysenko, S. (2010). Problemni momenty modeliuвання biznes-protsesiv z metoiu provedennia podalshoho reinzhynirynhu na bazi standartiv IDEF ta UML-modeliuвання. *Visnyk Natsionalnoho universytetu «Lvivska politekhnik»* 691, 41 - 46.
9. Yakovyna, V., Parfeniuk, Yu. (2013). Vykorystannia zasobiv UML dlia prohnouzuvannia nadiinosti prohramnoho zabezpechennia na etapi yoho proektuvannia. *Visnyk Natsionalnoho universytetu «Lvivska politekhnik»* 773, 151 - 156.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-38>

УДК 004.77

Садовий Ян Станіславович, здобувач ступеня доктора філософії

<https://orcid.org/0000-0003-0387-6772>

Державний університет «Житомирська політехніка», м. Житомир, Україна

ПОРІВНЯЛЬНИЙ АНАЛІЗ РОЗПОДІЛЕНИХ СУФІКСНИХ ДЕРЕВ І ТРАДИЦІЙНИХ МЕТОДІВ УПРАВЛІННЯ ДАНИМИ

Садовий Я.С. Порівняльний аналіз розподілених суфіксних дерев і традиційних методів управління даними. Дана стаття містить огляд різних методів реалізації суфіксних дерев, порівнюючи їх із традиційними методами подання даних, такими як префіксні дерева та перетворення Берроуза-Вілера (BWT). Переглянуті методи включають наївний підхід, алгоритм Укконена та алгоритм розділення та запису лише зверху вниз (Partition and Write Only Top Down – PWOTD). Серед розглянутих методів алгоритм PWOTD виявляється найефективнішим для реалізації суфіксних дерев. Використовуючи розбиття суфіксів, цей підхід значно зменшує вимоги до основної пам'яті для побудови дерева суфіксів, дозволяючи створювати незалежні піддерева повністю в основній пам'яті. Незважаючи на збільшення обчислювальних витрат, пов'язаних із розділенням, алгоритм PWOTD пропонує суттєве скорочення вимог до простору для суфіксів і тимчасових масивів, а також масиву дерева. Порівняльний аналіз показує, що суфіксні дерева перевершують інші методи завдяки ефективному зберіганню та представленню всіх суфіксів певного рядка в компактному та доступному форматі. Крім того, суфіксні дерева підтримують динамічні оновлення з логарифмічною часовою складністю, полегшуючи додавання або видалення рядків з вихідного набору даних без необхідності повної реконструкції, на відміну від префіксних дерев. У підсумку, ця стаття підкреслює ефективність алгоритму PWOTD у реалізації суфіксальних дерев і підкреслює переваги суфіксальних дерев над традиційними методами представлення даних. Універсальність, ефективність і динамічні можливості дерев суфіксів роблять їх незамінними для широкого спектру завдань обробки рядків у різних областях, включаючи біоінформатику, індексування тексту та обробку природної мови.

Ключові слова: суфіксні дерева, алгоритм PWOTD, репрезентація даних, обробка рядків, перетворення, обчислювальна ефективність.

Sadovyi I. The Comparative Analysis Of Distributed Suffix Trees And Traditional Data Management Methods.

This article provides a comprehensive review of various methods for implementing suffix trees, comparing them with traditional data representation techniques such as prefix trees and the Burrows-Wheeler transformation (BWT). The reviewed methods include the naive approach, the Ukkonen algorithm, and the Partition and Write Only Top Down (PWOTD) algorithm. Among the methods discussed, the PWOTD algorithm emerges as the most effective for implementing suffix trees. By employing suffix splitting, this approach significantly reduces the main memory requirements for constructing a suffix tree, enabling the creation of independent subtrees entirely in main memory. Despite the increased computational costs associated with partitioning, the PWOTD algorithm offers substantial reductions in space requirements for suffixes and temporary arrays, as well as the tree array. Comparative analysis reveals that suffix trees outperform other methods by efficiently storing and presenting all suffixes of a given string in a compact and accessible format. Moreover, suffix trees support dynamic updates with logarithmic time complexity, facilitating the addition or removal of rows from the original dataset without necessitating complete reconstruction, unlike prefix trees. In summary, this article highlights the effectiveness of the PWOTD algorithm in implementing suffix trees and underscores the advantages of suffix trees over traditional data representation methods. The versatility, efficiency, and dynamic capabilities of suffix trees make them indispensable for a wide range of string processing tasks across various domains, including bioinformatics, text indexing, and natural language processing.

Key words: suffix trees, PWOTD algorithm, data representation, string processing, transformation, computational efficiency.

Вступ та постановка проблеми. Суфіксні дерева являють собою стиснуті таблиці, які зберігають всі суфікси заданого текстового рядка. Ця структура даних інтенсивно використовується для зіставлення шаблонів на рядках і деревах із широким діапазоном застосувань, таких як молекулярна біологія, обробка даних, редагування тексту, переписування термінів, проектування інтерпретатора, пошук інформації, абстрактні типи тощо. Структура даних суфіксного дерева має центральне значення в обробці рядків. Воно забезпечує компактне представлення всіх суфіксів рядка та дає можливість ефективного вирішення проблеми найдовшого загального підрядка, яка полягає у визначенні найдовшого рядка, який є безперервною послідовністю символів, присутніх у двох або більше вхідних рядках. Зокрема, для двох рядків мета полягає в тому, щоб знайти підрядок максимальної довжини, який з'являється в обох рядках. На відміну від підпоследовностей, підрядки повинні займати послідовні позиції в оригінальних рядках. Проблема є важливою для різних застосувань, включаючи біоінформатику для аналізу послідовності ДНК, обробку тексту для виявлення плагіату та стиснення даних для виявлення повторюваних шаблонів. Крім того суфіксні дерева здатні вирішувати проблеми точної відповідності рядків.

З іншого боку, суфіксні дерева являють собою ніщо інше, як метод представлення рядків. Великою проблемою цього типу представлення є те, що він вимагає багато місця для зберігання.

Загалом, суфіксні дерева вимагають у декілька разів більше пам'яті, ніж якби рядки зберігалися без використання подібного типу структурування. Інша проблема полягає в тому, що конструювання суфіксного дерева вимагає великих часових ресурсів. Іншими словами будувати таке дерево ще не означає наявності необхідних даних. На фоні цього, існує необхідність порівняння методів представлення даних у вигляді суфіксних дерев з одного боку та традиційних методів представлення даних, які включають: префіксні дерева та перетворення Барроуза-Вілера, з іншого, що і є метою поточного дослідження.

Аналіз останніх досліджень і публікацій. У науково-дослідницькому просторі сьогодення з'являються роботи, присвячені винаходу та аналізу методології по розробці поліпшених методів управління даними.

У роботі [1] досліджувалися алгоритми побудови суфіксних дерев у широкому спектрі джерел даних і розмірів. Було показано, що на сучасних процесорах кеш-ефективний алгоритм зі складністю $O(n^2)$ перевершує такі популярні алгоритми, як Укконена та Мак-Крайта, навіть для побудови в пам'яті. Для більших наборів даних вимога до дискового вводу-виводу швидко стає слабким місцем у продуктивності кожного алгоритму. Щоб вирішити цю проблему, було описано два підходи. Спочатку було представлено стратегію керування буфером для алгоритму $O(n^2)$. Отриманий новий алгоритм, який має назву «Top Down Disk-based» (TDD), масштабується до розмірів, набагато більших, ніж описано раніше в літературі. Цей підхід значно перевершує найкращі відомі методи конструювання на основі дисків. По-друге, було представлено новий дисковий алгоритм побудови суфіксного дерева, який базується на парадигмі сортування-злиття.

Робота [2] була зосереджена на побудові суфіксного масиву для набору рядків. Було представлено `gsufsort`, що являє собою програмне забезпечення з відкритим вихідним кодом для побудови суфіксних масивів і відповідних структур індексування даних для колекції рядків із N символів за $O(N)$ час. Інструмент написаний мовою ANSI/C і базується на алгоритмі `gSACA-K`, найшвидшому алгоритмі для створення суфіксних масивів для колекцій рядків. Інструмент підтримує великі файли `fasta`, `fastq` і текстові файли з кількома рядками як вхідні дані. Експерименти показали хорошу продуктивність на різних типах рядків.

Крім того, варто зазначити праці наступних науковців: Фарруджіа Андреа, Гегі Тревіс, Наварро Гонсало, Пуглісі Сімон, Сірен Джуні [3], Фунакосі Міцуру, Мієно Такуя, Накашима Юто, Іненага Сюнсукі, Баннаї Хідео, Такеда Масаюкі [4], Касерес Мануель, Еварлена Цюса [5], Кішор Чандра, Джаганатан Субаш Чандра Бос, Азат М., Решмі Т., Марина Нінослав [6], Білазугі Джамаль, Косолобов Дмитро, Пуглісі Сімон, Раман Раджив [7], Буше Христина, Цвачо Ондржей, Гегі Тревіс, Голуб Ян, Манзіні Джованні, Наварро Гонсало, Россі Массіміліано [8], Барські Марина, Стег Ульріке, Томо Алекс [9], Укконен Е [10], Аяд Лоррейн, Лукідес Григоріос, Пісіс Солон, Вербек Хільде [11], Аджеро Дональд, Тімоті Белл, Мукерджи Амар [12], Рабеа Зеаніб, Ель-Метваллі Сара, Ельмугі Самір, Закарія Магді [13], Ву І, Нонг Ге, Чан Вей, Хань Лін [14], Флік Патрік, Алуру Срінівас [15] та інших.

Проте, беручи до уваги вище зазначену наукову документацію, питання, пов'язане з методологією по розробці поліпшених методів управління даними все ще залишається недостатньо дослідженим та потребує подальшого опрацювання.

Постановка завдання. Метою роботи є проведення порівняльного аналізу розподілених суфіксних дерев і традиційних методів управління даними.

Викладення основного матеріалу дослідження. Побудова суфіксного дерева може бути досягнута за допомогою різних алгоритмів. Ці алгоритми відрізняються за своїми підходами та ефективністю, задовольняючи різні обчислювальні потреби. Наївний підхід, хоч і є простим методом, є неефективним для великих рядків. Більш складні методи включають алгоритм Мак-Крейта, який поступово будує дерево, і алгоритм Укконена, відомий своєю складністю в лінійному часі.

Наївна побудова суфіксного дерева, яка зображена схематично на рисунку 1, передбачає ітеративне вставлення кожного суфікса рядка в дерево по одному. В якості прикладу береться рядок `хавхас$`. Початок являє собою побудову дерева суфіксів для всього рядка, оскільки суфікс `хавхас$` починається з позиції рядка 1. Таким чином рядок проходиться символ за символом.

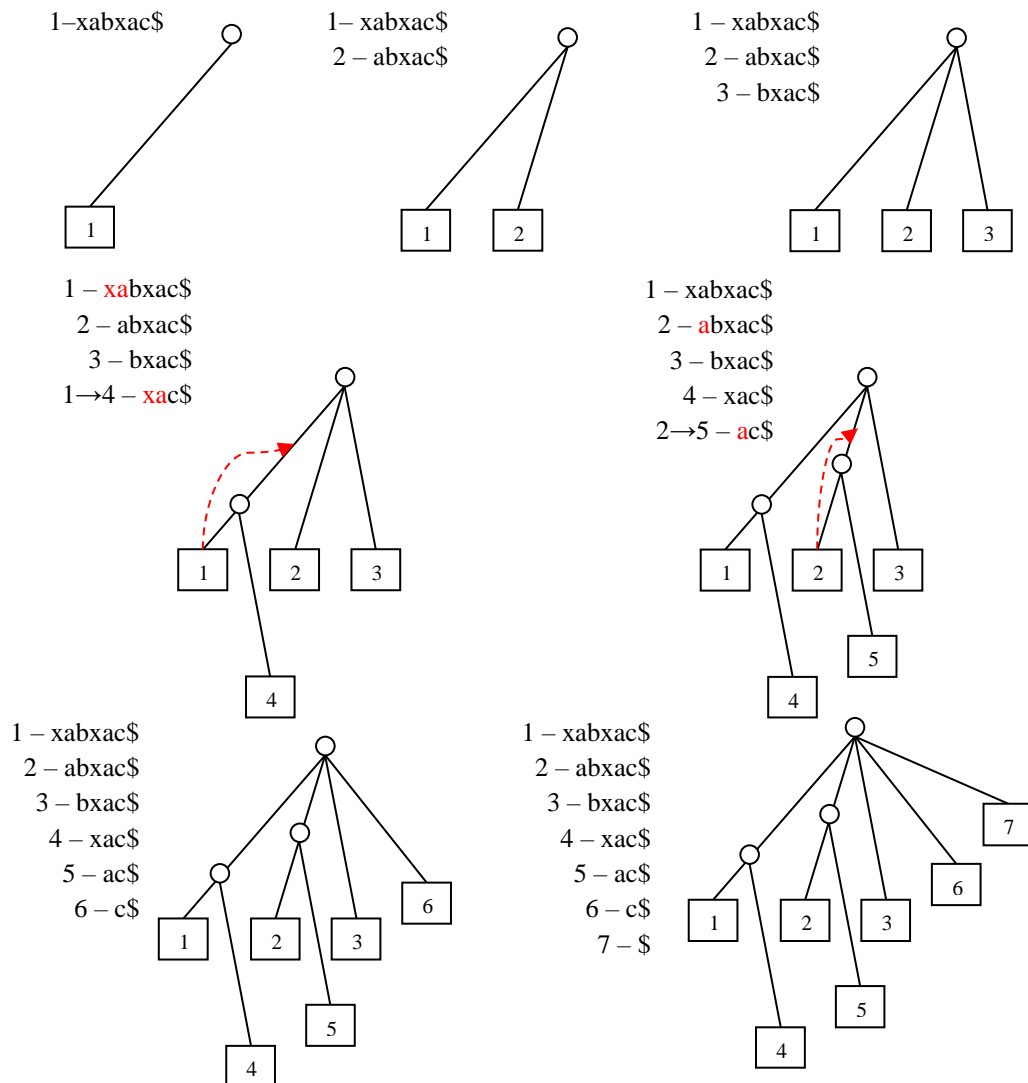


Рис.1. Схема побудови системи суфіксних дерев за допомогою наївного методу

Однак, якщо виникає суфікс, початкові символи якого вже з'являються в дереві суфіксів, існуюче ребро розбивається на два ребра (слова) саме в тому місці, де закінчується рівність між уже існуючим суфіксом і суфіксом, який потрібно вставити. Головною перевагою використання даного методу є його простота, яка полягає в ітеративному вставленні кожного суфіксу рядка. Однак цей метод має значні недоліки, які обмежують його практичність для більш масштабних застосувань, як наприклад складання коротких послідовностей ДНК у повний геном. Найбільш помітним недоліком є його неефективність, наївний підхід має квадратичну часову складність, тобто час побудови різко збільшується з довжиною рядка. Ця неефективність виникає через те, що кожна вставка суфікса часто вимагає переходу від кореня, що призводить до повторних і надлишкових операцій.

Наступним розглядається алгоритм Укконена, який полягає у поступовому будівництві дерева, додаючи по одному символу до вхідного рядка під час створення дерева на льоту. Це інкрементний алгоритм, який починає з дерева для пустого слова і поступово додає нові символи до слова, одночасно оновлюючи дерево суфіксів. Кожен символ додається за амортизований постійний час. З цієї причини даний алгоритм будує суфікс не дерево для слова σ за час $O(|\sigma|)$. Припускається, що ребра до нащадків одного вузла можуть бути проіндексовані їхніми початковими символами при умові, що алфавіт є фіксованим та малим. Коли слово σ розширюється до $s\sigma$, дерево змінюється наступним чином:

– всі існуючі вузли дерева (включаючи приховані) відповідають підсловом σ . Це також підслова $s\sigma$, тому вони також зустрічаються як вузли нового дерева;

– якщо β було розгалуженим підсловом, воно залишається розгалуженим – отже, внутрішні вузли залишаються такими;

– кожен новий суфікс β отримується шляхом розширення вихідного суфікса β .

Якщо β був некладеним, тобто листом, β також буде некладеним. Мітки їхніх країв мають бути подовжені на a . Щоб зробити це ефективним, вводяться відкриті ребра, мітки яких індексують σ від заданої позиції до кінця. Якщо β був вкладеним суфіксом (тобто внутрішнім або прихованим вузлом), тоді або β присутній у σ , і β є вкладеним суфіксом нового слова, і дерево не потребує змін; або β не зустрічається в σ . Потім необхідно створити новий лист з відкритим краєм i , можливо, новим внутрішнім вузлом, під яким буде з'єднаний новий лист.

Для того щоб визначити чи є суфікс вкладеним, чи ні, припускається, що якщо α є вкладеним суфіксом слова σ , а β є суфіксом слова α , то β також є вкладеним суфіксом слова σ . Слово σ містить як αx , так і αy для деяких різних символів x і y . Оскільки α закінчується на β , де y у σ повинні бути і βx , і βy , тому β є вкладеним. З цієї причини достатньо зберегти найдовший вкладений суфікс слова σ . Він називається активним суфіксом і позначається $\alpha(\sigma)$. Кожен суфікс $\beta \subseteq \sigma$ є вкладеним коли $|\beta| \leq |\alpha(\sigma)|$. Активний суфікс розмежує некладені та вкладені суфікси.

Потім, необхідно підтримувати умову $\alpha = \alpha(\sigma)$ і, додаючи новий символ a , перевірити, чи αa залишається вкладеним. Якщо так, нічого не зміниться. Якщо цього не відбувається, додається новий аркуш i , можливо, також новий внутрішній вузол; потім видаляється перший символ a і продовжується перевірка. Додавання символу до σ викликає амортизовану постійну кількість модифікацій дерева. Щоб виконувати кожну модифікацію в (амортизованому) постійному часі, нам потрібне відповідне представлення слова α , яке підтримує ефективні додавання, вирізання першого символу та перевірки існування відповідного вузла в дереві.

Наступний етап полягає у введенні поняття референтної пари. Референтна пара для слова $\alpha \subseteq \sigma$ являє собою пару (π, τ) , де π є вузлом дерева, τ – довільним словом, а $\pi\tau = \alpha$. Крім того, $\tau \subseteq \sigma$, тому τ може бути представлено парою індексів у σ . Референтна пара є канонічною, якщо немає ребра з вузла π з міткою, яка є префіксом τ . Кожне слово $\alpha \subseteq \sigma$ має одну канонічну референтну пару, що представляє його. Заднє ребро $back(\pi)$ веде від внутрішнього вузла π до внутрішнього вузла, який представляє $\pi[1 :]$.

Активний суфікс α представляється референтною парою (π, τ) . При додаванні a до τ , отримується референтна пара для αa , але вона не обов'язково є канонічною. Перед додаванням необхідно перевірити, чи присутній α в дереві. Якщо π не є коренем дерева, π замінюється на $back(\pi)$ і зберігається τ . В іншому випадку π є порожнім рядком, внаслідок чого перший символ видаляється із τ (збільшується відповідний початковий індекс). Наступним кроком проводиться канонізація, оскільки попередні дві операції можуть створити неканонічну пару. Отже, перевіряється, чи є пара канонічною: для $\tau \neq \varepsilon$ перевіряється, чи є ребро з π , індексоване значенням τ , достатньо коротким, щоб бути префіксом τ .

Таким чином остаточний вид алгоритму формулюється так. Є вхідна послідовність $\alpha = \alpha(\sigma)$, представлена канонічною референтною парою (π, τ) , суфіксним деревом T для σ , задніми ребрами $back$ і новим символом a . Перевіряється, чи є αa в дереві, і за потреби створюється у ньому. Якщо $\tau = \varepsilon$: ($\alpha = \pi$ – внутрішній вузол) Якщо є ребро з вузла π , мітка якого починається з a , то наявне αa . Якщо такого ребра немає, його немає, створюється нове відкрите ребро, що йде від π до нового листа. У протилежному випадку $\tau \neq \varepsilon$: (α – прихований вузол) Далі шукається ребро з π , мітка якого починається з τ . Якщо після τ у мітці стоїть a , то присутнє αa . В іншому випадку його немає, тому ребро ділиться, та створюється новий вузол так, що ребро від π до нового вузла позначається τ . Новий вузол матиме відкритий край для нового дочірнього вузла. Якщо αa не було присутнім, видаляється перший символ a і операцій починається знову з перевірки наявності αa в дереві. При присутності αa , оновлюється референтна пара, щоб представляти αa та перераховуються задні ребра. Наприкінці відомо, що $\alpha = \alpha(\sigma a)$ представляється як канонічна референтна пара (π, τ) , суфіксне дерево T наявне для σa та задніх ребер $back$.

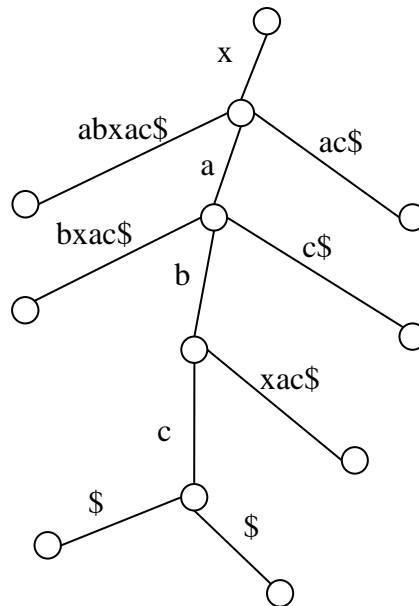


Рис.2. Схематичне зображення побудованого суфіксного дерева завдяки алгоритму Укконена

Узагальнюючи, необхідно відмітити, що однією з переваг алгоритму Укконена є його лінійна часова складність, $O(n)$, що робить його надзвичайно ефективним для обробки довгих рядків. Ця ефективність досягається завдяки розумному використанню суфіксальних посилань і активних точок, що дозволяє алгоритму уникати надлишкової роботи та гарантувати швидку обробку кожної фази. Крім того, алгоритм Укконена поступово створює дерево суфіксів, дозволяючи йому обробляти потокові дані та оновлювати дерево, коли нові символи додаються до рядка. Ця поетапна конструкція особливо корисна в програмах, які вимагають оновлення в реальному часі, таких як текстові редактори та системи пошуку в реальному часі. Крім того, здатність алгоритму обробляти великі алфавіти без значного зниження продуктивності ще більше підвищує його універсальність.

Незважаючи на ці сильні сторони, алгоритм Укконена також має деякі недоліки, які полягають у складності його реалізації та потреба у великій кількості пам'яті, особливо для дуже великих рядків, через необхідність зберігати численні суфіксні посилання та вузли дерева.

Останнім у даній статті розглядатиметься алгоритм Partition and Write Only Top Down (PWOTD). Суфікси вхідного рядка розбиваються на розділи $|X|^{prefixlen}$, де $|A|$ являє собою розмір рядка в алфавіті, а $prefixlen$ – це глибина розділення. Етап розділення виконується наступним чином. Вхідний рядок сканується зліва направо. У кожній позиції індексу i наступні символи $prefixlen$ використовуються для визначення одного з розділень $|X|^{prefixlen}$. Потім цей індекс i записується в буфер обчисленого розділу. Наприкінці сканування кожен розділ міститиме покажчики суфіксів для суфіксів, які мають однаковий префікс розміру $prefixlen$. Необхідно зауважити, що кількість розділів ($|X|^{prefixlen}$) є набагато меншою за довжину рядка. Розділення рядка XABXAC\$ використовуючи $prefixlen$ зі значенням 1 створило б чотири розділи суфіксів, по одному для кожного символу в алфавіті, оскільки ігнорується остаточний розділ, що складається лише з символу закінчення рядка \$. Суфіксне розділення для символу X буде {0,3}, що представляє суфікси {XABXAC\$, XAC\$}. Розділення суфікса для символу A буде {1,4}, що представлятиме суфікси {ABXAC\$, AC\$}.

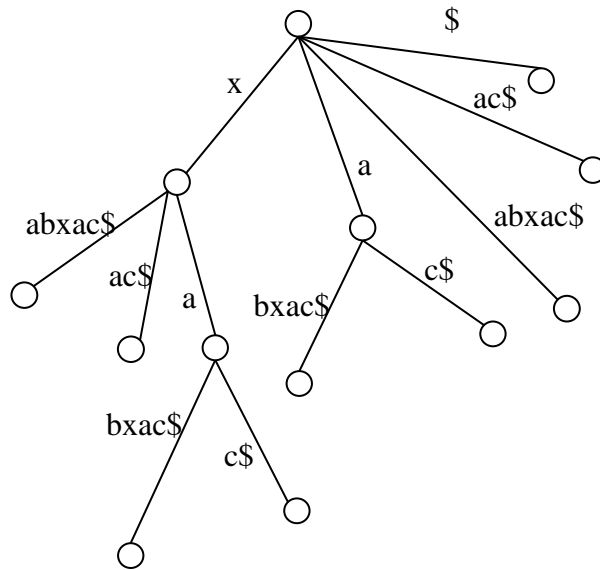


Рис. 3. Схематична побудова суфіксного дерева за допомогою алгоритму PWOTD

У підсумку, Алгоритм PWOTD пропонує ефективний підхід до побудови дерев суфіксів для великих рядків шляхом поділу вхідного рядка та роботи з підмножинами суфіксів. Це розділення зменшує вимоги до основної пам'яті для побудови дерева суфіксів, дозволяючи будувати незалежні піддерева повністю в основній пам'яті. З іншої сторони, однак розділення зменшує вимоги до простору для суфіксів і тимчасових масивів, а також масиву дерева, це відбувається за рахунок збільшення обчислювальних витрат на етапі розділення.

Одним з основних традиційних методів представлення даних є префіксні дерева, які являють собою спеціальні дерева пошуку для зберігання кількох рядків символів одночасно. Дані типи дерев включають тип стиснення даних, оскільки загальні префікси рядків символів зберігаються лише один раз. Вони будуються з використанням набору довільних рядків символів. Кожне вихідне ребро вузла в префіксному дереві анується одним символом, так що шлях, починаючи від кореня до листа в дереві, представляє один із рядків символів, з яких будується дерево. Порівняно з ними, суфіксні дерева дозволяють виконувати запити швидше та використовують менше пам'яті, особливо для текстів із великою кількістю суфіксів.

Далі, перетворення Барроуза-Вілера представляються як оборотне перетворення вхідного рядка s з додаванням до нього унікального символ кінця файлу. Потім формується концептуальна матриця, що містить усі циклічні зсуви s . Далі сортується рядки цієї матриці в лексикографічному порядку справа наліво і встановлюється значення $bw(s)$ як перший стовпець відсортованої матриці з кінцем – символ файлу видаляється. Необхідно зауважити, що цей процес еквівалентний сортуванню s з використанням, як ключа сортування для кожного символу, його контексту, тобто набору символів, що йому передують. Результатом перетворення Барроуза-Вілера є рядок $bw(s)$ та індекс I у відсортованій матриці рядка, що починається з символу кінця файлу. Приклад даного метода зображений у таблиці 1. Уваги потрубує той факт що, на відміну від суфіксних дерев, найбільшим недоліком алгоритмів, заснованих на перетвореннях Барроуза-Вілера, є те, що вони не працюють в режимі on-line, тобто вони повинні обробити велику частину вхідних даних, перш ніж можна буде створити один вихідний біт.

Тим не менш, основна перевага перетворень Барроуза-Вілера порівняно з суфіксними деревами полягає в його простоті реалізації. Зазвичай BWT потребує менше пам'яті та обчислювальних ресурсів порівняно з суфіксними деревами, що робить його більш придатним для програм з обмеженою пам'яттю чи обчислювальною потужністю.

Таблиця 1 – Приклад перетворення Берроуза-Вілера. $bw(\text{mississippi}) = \text{mssiprissii}$. Матриця справа отримана шляхом сортування рядків у лексикографічному порядку справа наліво.

mississippi•		m ississippi•
ississippi•m		s sissippi•mi
ssissippi•mi		• mississippi
sissippi•mis		s sippi•missi
issippi•miss		s sippi•missi
ssippi•missi	→	i ssissippi•m
sippi•missis		i ssissippi•m
ippi•mississ		i ssissippi•m
ppi•mississi		i ssissippi•m
pi•mississip		s ippi•missis
i•mississipp		i sippi•miss
•mississippi		i ppi•mississ

Крім того, BWT є особливо ефективним для завдань стиснення, що особливо помітно для повторюваних або структурованих даних, де перетворена послідовність має тенденцію мати довгі цикли повторюваних символів, що дозволяє використовувати ефективні алгоритми стиснення, такі як Move-to-Front (MTF) і Run-Length Encoding (RLE).

Висновки. У даній роботі був проведений огляд основних методів реалізації суфіксних дерев, які включали наївний метод, алгоритм Укконена та алгоритм Partition and Write Only Top Down (PWOTD) та їх порівняння з урахуванням традиційних методів репрезентації даних, таких як префіксні дерева та перетворення Берроуза-Вілера.

При порівнянні різних методів реалізації суфіксних дерев, найбільш ефективним виявився метод на основі алгоритму PWOTD, оскільки він пропонує поділ вхідних рядка та роботу з підмножинами суфіксів. Розділення суфіксів зменшує вимоги до основної пам'яті для побудови дерева суфіксів, дозволяючи будувати незалежні піддерева повністю в основній пам'яті. Тим не менш, дане розділення зменшує вимоги до простору для суфіксів і тимчасових масивів, а також масиву дерева за рахунок збільшення обчислювальних витрат.

При порівнянні суфіксних дерев до інших методів, виявляється що вони, на відміну від інших методів здатні ефективно зберігати та представляти всі суфікси заданого рядка в компактній та легкодоступній формі. Крім того, суфіксні дерева можна модифікувати в логарифмічній часовій складності, дозволяючи ефективно додавати або видаляти рядки з вихідного набору даних без необхідності повної реконструкції, на відміну від префіксних дерев.

Список бібліографічного опису

1. Hankins R., Tata S., Patel J., Tian Y. Practical methods for constructing suffix trees. *The VLDB Journal*. 2005. №14. DOI:10.1007/s00778-005-0154-8.
2. Louza F., Telles G., Gog S., Prezza N., Rosone G. gsufsort: constructing suffix arrays, LCP arrays and BWTs for string collections. *Algorithms for molecular biology : AMB*. 2020. №15. №18. DOI:10.1186/s13015-020-00177-y.
3. Farruggia A., Gagie T., Navarro G., Puglisi S., Sirén J. Relative Suffix Trees. *The Computer Journal*. 2018. №61. P.773-788. DOI:10.1093/comjnl/bxx108.
4. Funakoshi M., Mieno T., Nakashima Y., Inenaga S., Bannai H., Takeda M. Computing palindromes and suffix trees of dynamic trees. 2024. DOI:10.21203/rs.3.rs-4167645/v1.
5. Caceres M., Navarro G. Faster Repetition-Aware Compressed Suffix Trees based on Block Trees. *Information and Computation*. 2021. №285. DOI:10.1016/j.ic.2021.104749.
6. Cirkoska E., Kishore C., Jaganathan S., Azath M., Reshmi T., Marina N. A Study on Suffix Trees and Their Applications in Genome Sequences Using MUMmer. 2021. DOI:10.1007/978-981-15-9647-6_2.
7. Belazzougui D., Kosolobov D., Puglisi S., Raman R. Weighted Ancestors in Suffix Trees Revisited. 2021.
8. Boucher C., Cvacho O., Gagie T., Holub J., Manzini G., Navarro G., Rossi M. PFP Compressed Suffix Trees. 2021. DOI:10.1137/1.9781611976472.5.
9. Barsky M., Stege U., Thomo A. Suffix trees for inputs larger than main memory. *Inf. Syst.* 2011. №36. P.644-654. DOI:10.1016/j.is.2010.11.001.
10. Ukkonen E. On-line construction of suffix trees. *Algorithmica*. 1995. №14(3) P.249–260.
11. Ayad L., Loukides G., Pissis S., Verbeek H. Sparse Suffix and LCP Array: Simple, Direct, Small, and Fast. 2024. DOI:10.1007/978-3-031-55598-5_11.
12. Adjeroh D., Bell T., Mukherjee A. Suffix trees and suffix arrays. 2008. DOI:10.1007/978-0-387-78909-5_4.
13. Rabea Z., El-Metwally S., Elmougy S., Zakaria M. A fast algorithm for constructing suffix arrays for DNA alphabets.

Journal of King Saud University - Computer and Information Sciences. 2022. №34. DOI:10.1016/j.jksuci.2022.04.015.
 14. Wu Y., Nong G., Chan W., Han L. Checking Big Suffix and LCP Arrays by Probabilistic Methods. *IEEE Transactions on Computers*. 2017. P. 1. DOI:10.1109/TC.2017.2702642.
 15. Flick P., Aluru S. Distributed enhanced suffix arrays: efficient algorithms for construction and querying. *SC '19: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2019. P.1-17. DOI:10.1145/3295500.3356211.

References:

1. Hankins R., Tata S., Patel J., Tian Y. Practical methods for constructing suffix trees. *The VLDB Journal*. 2005. №14. DOI:10.1007/s00778-005-0154-8.
2. Louza F., Telles G., Gog S., Prezza N., Rosone G. gsufsort: constructing suffix arrays, LCP arrays and BWTs for string collections. *Algorithms for molecular biology : AMB*. 2020. №15. №18. DOI:10.1186/s13015-020-00177-y.
3. Farruggia A., Gagie T., Navarro G., Puglisi S., Sirén J. Relative Suffix Trees. *The Computer Journal*. 2018. №61. P.773-788. DOI:10.1093/comjnl/bxx108.
4. Funakoshi M., Mieno T., Nakashima Y., Inenaga S., Bannai H., Takeda M. Computing palindromes and suffix trees of dynamic trees. 2024. DOI:10.21203/rs.3.rs-4167645/v1.
5. Caceres M., Navarro G. Faster Repetition-Aware Compressed Suffix Trees based on Block Trees. *Information and Computation*. 2021. №285. DOI:10.1016/j.ic.2021.104749.
6. Cirkoska E., Kishore C., Jaganathan S., Azath M., Reshmi T., Marina N. A Study on Suffix Trees and Their Applications in Genome Sequences Using MUMmer. 2021. DOI:10.1007/978-981-15-9647-6_2.
7. Belazzougui D., Kosolobov D., Puglisi S., Raman R. Weighted Ancestors in Suffix Trees Revisited. 2021.
8. Boucher C., Cvacho O., Gagie T., Holub J., Manzini G., Navarro G., Rossi M. PFP Compressed Suffix Trees. 2021. DOI:10.1137/1.9781611976472.5.
9. Barsky M., Stege U., Thomo A. Suffix trees for inputs larger than main memory. *Inf. Syst.* 2011. №36. P.644-654. DOI:10.1016/j.is.2010.11.001.
10. Ukkonen E. On-line construction of suffix trees. *Algorithmica*. 1995. №14(3) P.249–260.
11. Ayad L., Loukides G., Pissis S., Verbeek H. Sparse Suffix and LCP Array: Simple, Direct, Small, and Fast. 2024. DOI:10.1007/978-3-031-55598-5_11.
12. Adjero D., Bell T., Mukherjee A. Suffix trees and suffix arrays. 2008. DOI:10.1007/978-0-387-78909-5_4.
13. Rabea Z., El-Metwally S., Elmougy S., Zakaria M. A fast algorithm for constructing suffix arrays for DNA alphabets. *Journal of King Saud University - Computer and Information Sciences*. 2022. №34. DOI:10.1016/j.jksuci.2022.04.015.
14. Wu Y., Nong G., Chan W., Han L. Checking Big Suffix and LCP Arrays by Probabilistic Methods. *IEEE Transactions on Computers*. 2017. P. 1. DOI:10.1109/TC.2017.2702642.
15. Flick P., Aluru S. Distributed enhanced suffix arrays: efficient algorithms for construction and querying. *SC '19: Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. 2019. P.1-17. DOI:10.1145/3295500.3356211.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-40>

UDC 519.8:004.42

Dymova Hanna, Candidate of Technical Sciences, Phd., Associate Professor

<https://orcid.org/0000-0002-5294-1756>

Oksana Larchenko, Candidate of Agricultural Sciences, Associate Professor

<https://orcid.org/0000-0001-7857-0802>

Kherson State Agrarian and Economic University, Kherson, Ukraine

USING THE MAX FLOW PROBLEM FOR BUSINESS PROCESSES

Dymova H., Larchenko O. Using the Max Flow Problem for Business Processes. The article is devoted to the use of optimization methods for managing business processes. Problems of synthesizing the structure and choosing parameters of systems, as well as problems of optimal control of them, are most often solved by reducing them to problems of finding a number of parameters so that the extremum of the quality function is achieved under a set of restrictions. A separate group of problems related to the choice of topology can be solved using the apparatus of graph theory. Problems of choosing channel capacity can be reduced to classical optimization problems, such as linear programming, transport problem, nonlinear programming, etc. The minimum flow problem is a special case of the transport problem, which is related to linear programming problems.

The paper analyzes the use of the maximum flow problem in economic systems, namely for planning business processes. The use of tasks to optimize business processes in various industries, such as logistics, production and finance, and personnel management, is described. A review of the developments and research of the maximum flow problem is carried out, a number of algorithms for solving this problem and their main points are given. The problems, the mathematical formulation of the problem, the algorithm for placing marks and the calculation of arc flows for each iteration are described in detail. An example of the flagging algorithm shows how maximum flow models can be used to analyze and optimize production processes. Maximum flow problems can identify bottlenecks in a production chain, optimize resource utilization, and plan production. Using the maximum flow problem when planning business processes will improve efficiency and increase production, help reduce costs and improve profitability, and ensure that products are in stock to meet demand.

Keywords: business processes, maximum flow, Ford-Fulkerson algorithm, source, stock, augmentative flow path.

Димова Г.О., Ларченко О.В. Використання задачі про максимальний потік для бізнес-процесів. Стаття присвячена використанню методів оптимізації для управління бізнес-процесами. Задачі синтезу структури та вибору параметрів систем, а також задачі оптимального управління ними найчастіше розв'язують, зводячи до задач знаходження низки параметрів так, щоб досягався екстремум функції якості при наборі обмежень. Окрему групу задач, які зв'язані з вибором топології, можна розв'язати із застосуванням апарату теорії графів. Задачі вибору пропускної здатності каналів можуть бути зведені до класичних задач оптимізації, таких як лінійного програмування, транспортної задачі, нелінійного програмування тощо. Задача про мінімальний потік являється окремим випадком транспортної задачі, що відноситься до задач лінійного програмування.

В роботі проаналізовано використання задачі про максимальний потік в економічних системах, а саме для планування бізнес-процесів. Описано використання задачі для оптимізації бізнес-процесів в різних галузях, таких як логістика, виробнича та фінансова сфери, управління персоналом. Зроблено огляд розробок та досліджень задачі про максимальний потік, наведена низка алгоритмів розв'язання цієї задачі та основні їх моменти. Детально наведено опис проблеми, математична постановка задачі, алгоритм розставлення поміток та розрахунок дугових потоків для кожної ітерації. Приклад роботи алгоритму розставлення поміток демонструє, як моделі максимального потоку можуть бути використані для аналізу та оптимізації виробничих процесів. Задачі про максимальний потік можуть визначити вузькі місця у виробничому ланцюгу, оптимізувати використання ресурсів та планувати виробництво. Використання задачі про максимальний потік при плануванні бізнес-процесів дозволить покращити ефективність та збільшення виробництва, допоможе зменшити витрати та покращити прибутковість, забезпечити продукцією на складі для задоволення попиту.

Ключові слова: бізнес-процеси, максимальний потік, алгоритм Форда-Фалкерсона, джерело, стік, аугментальний шлях потоку.

Formulation of the problem. Business process management is aimed at providing quality service to consumers (clients) [1]. For companies with a high degree of business diversification and a variety of partnerships, business process optimization provides solutions to the following tasks:

- determining the optimal sequence of functions performed, which leads to a reduction in the cycle time for the production and sale of goods and services, customer service, which results in an increase in capital turnover and an increase in all economic indicators of the company;
- optimization of the use of resources in various business processes, as a result of which production and distribution costs are minimized and an optimal combination of various types of activities is ensured;
- building adaptive business processes aimed at quickly adapting to changes in the needs of end consumers of products, production technologies, the behavior of competitors in the market and, consequently, improving the quality of customer service in a dynamic external environment;

– determination of rational schemes for interaction with partners and clients, and as a result, profit growth, optimization of financial flows.

The maximum flow problem is a special case of the transport problem and can be used to optimize business processes in various industries [2]. In logistics, maximum flow can be used to plan optimal delivery routes for goods or allocate resources between warehouses. This will help reduce delivery times and optimize the use of vehicles. In manufacturing, the maximum flow problem is used to optimize the flow of materials through production lines, thereby reducing congestion, increasing productivity, and reducing production costs. In the financial sector, the maximum flow problem helps banks efficiently distribute payment flows and reduce transaction processing time, that is, it is used to optimize the processing of financial transactions. In human resource management, a task can be used to optimize workflow flow and distribute tasks among work groups, allowing for more efficient use of work time and reducing task completion time. Consequently, the maximum flow problem is one of several deterministic flow models that are used to formulate and solve important economic problems.

Research analysis. The maximum flow problem has a rich history of research [2, 3]. The first research in this area was carried out by Lester Ford and Delbert Fulkerson (1955), who developed the Ford-Fulkerson algorithm, which is one of the most famous algorithms for solving the maximum flow problem. Later in 1972, Edmonds and Karp developed the Edmonds-Karp algorithm, which is more efficient than the Ford-Fulkerson algorithm for some types of graphs. There has also been research and development to solve this problem using different methods and algorithms. In 1988, one of the most effective algorithms for solving the maximum flow problem appeared - the push-relabel algorithm, developed by Alexander Rao and Andrew Goldberg. In 1995, Ulrich Feige and Hans-Wolfram Geertz proposed the potential method, which has become a powerful tool for analyzing and solving maximum flow problems. Kevin Busch developed a capacity-constrained flow algorithm in 2000 that can be used to solve maximum flow problems with additional constraints.

The study of the maximum flow problem is still ongoing. New algorithms and heuristics are constantly being developed, and existing methods are being improved. Let us list other researchers who have made significant contributions to this area - S.T. McCarthy, R.T. Rao, D.B. Kuhn, M.V. Frederiksen and S. Ramamurthy. The study of the maximum flow problem has had a significant impact on many fields, including computer science, operations research, mathematics, and engineering.

Presentation of the main material and justification of the obtained results. In the context of business process planning, consider a manufacturing enterprise that produces products at several stages: procurement of raw materials, production, assembly, packaging and shipment. Each stage has its own throughput - the maximum amount of products that can be processed in a certain period of time. The company's goal is to maximize the volume of products that can be shipped to customers.

There are many algorithms for solving the maximum flow problem, such as [1-4]:

- linear programming method: the problem is considered as a linear programming problem, where the variables are the flows along the edges, and the constraints are the conservation of the flow and the capacity limitation;
- Ford-Fulkerson algorithm: here you need to find any increasing path, increase the flow along all its edges by the minimum of their remaining capacities and repeat as long as there is an increasing path;
- Edmonds-Karp algorithm: the Ford-Fulkerson algorithm is executed, each time choosing the shortest increasing path (found by breadth-first search);
- Dinitz algorithm: is an improvement of the Edmonds-Karp algorithm. At each iteration, using breadth-first search, the distances from the source to all vertices in the residual network are determined. A graph is constructed containing only those edges of the residual network at which this distance increases by 1. Next, all dead-end vertices with edges incident to them are excluded from the graph until all vertices become non-dead-end; here a dead-end vertex is a vertex that does not enter or exit from any edge except the source and sink. The shortest increasing path is found on the resulting graph (this will be any path from s to t). Next, the edge with the minimum capacity is excluded from the residual network, dead-end vertices are again excluded, and so on while increasing paths still exist;
- preflow promotion algorithm: the algorithm operates on the preflow instead of a stream. The difference is that for any vertex, except for the source and sink, the sum of the flows entering it must be strictly zero (flow conservation condition), and for a preflow it must be integral. This amount is called excess flow into a vertex, and a vertex with positive excess flow is called overflowing. In

addition, for each vertex the algorithm stores an additional characteristic, height, which is an integral integer. The algorithm uses two operations: advancement and lifting. Promotion is possible when the edge belongs to the residual network, when it leads from a higher vertex to a lower one and the original vertex is full. Ascent is possible when the rising vertex is overcrowded, but none of the vertices to which the edges of the residual network lead from it are lower than it;

- “lift to top” algorithm: this is a variant of the previous algorithm, which in a special way determines the order of the promotion and lifting operations;
- binary blocking thread algorithm.

In this case, we will use the Ford-Fulkerson algorithm.

Let $G = (N, A)$ – be a directed network with one source $s \in N$ and one sink $t \in N$, and let the arcs $(i, j) \in A$ have limited capacity. The maximum flow problem is to find flows along arcs belonging to the set A , such that the resulting flow flowing from source s to sink t is maximum. It is assumed that an unlimited flow can enter the source and that for each intermediate node in the network the flow conservation condition is satisfied. This problem is nontrivial when the capacity U_{ij} of each arc represents a finite upper bound on the flow f_{ij} along that arc.

The maximum flow problem can be formulated as a linear programming problem, so the usual simplex method can be used to solve it. Let's consider another, more effective procedure for finding a solution to this problem. The algorithm starts with some feasible solution. The flagging procedure developed by Ford and Fulkerson [2, 5] is then performed to determine another allowable flow of larger magnitude. In this algorithm, nodes are considered as intermediate points of flow transmission, and arcs are considered as distribution channels. To formally describe the algorithm, it is necessary to introduce two basic concepts – markings and augmenting flow paths.

The node label is used to indicate both the flow value and the source of the flow causing a change in the current flow value along the arc connecting this source to the node in question [6]. If q_i units of flow are sent from node i to node j and cause an increase in flow along this arc, then we will say that node j is marked from node i with the symbol $+q_j$. In this case, node j is assigned the mark $[+q_j, i]$. Similarly, if sending q_j flow units causes flow to decrease along the arc, then node j is marked from node i with the symbol $-q_j$. In this case, node j is assigned the mark $[-q_j, i]$.

The current flow from node i to node j increases when q_j units of additional flow are sent to node j along an oriented arc (i, j) in the direction coinciding with its orientation. In this case, the arc (i, j) is called a direct flow [6]. A corresponding example is shown in Fig. 1.

The current flow from node i to node j is reduced when q_j flow units are sent to node j along an oriented arc (j, i) in the direction coinciding with its orientation. In this case, arc (j, i) is called reverse flow [6]. A corresponding example is shown in Fig. 2.

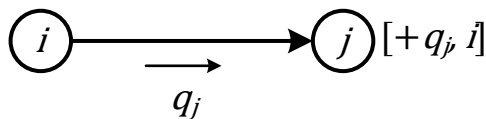


Figure 1 – Direct flow

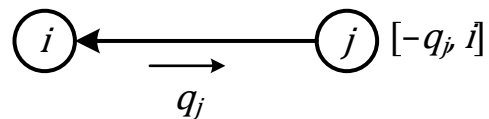


Figure 2 – Reverse flow

If node j is marked from node i and arc (i, j) is straight, then the flow along this arc increases and the value corresponding to the remaining unused capacity of the arc must be adjusted as necessary. This value is usually called the residual arc capacity. If a node is marked and a forward branch is used, then it can only have positive “residual capacity”. In addition, node j can be labeled from node i only after the node has been assigned a mark.

An augmental flow path from s to t is defined as a connected sequence of forward and backward arcs along which several units of flow can be sent from s to t . The flow along each forward arc increases without exceeding its capacity, and the flow along each reverse arc decreases, while remaining negative. The augmented flow path is used to select a method of changing the flow in which the flow at node t increases and at the same time the flow conservation condition is not violated for each internal node of the network.

The maximum flow problem is often encountered in practice, and the number of nodes and arcs in the network often reaches several thousand. Therefore, to solve such problems it is necessary to use an efficient calculation procedure. Due to the simplicity of the formulation of the maximum flow problem, an effective recurrent algorithm for finding the optimal solution (maximum flow) was developed using the labeling procedure. Let us present an algorithm for placing marks for the maximum flow problem.

Let (i, j) be an oriented arc leading from node i to node j . Let us note in what cases the flow along a given arc can be increased [2, 6]. As noted above, the flow can be increased by q_j units if the arc (i, j) is a forward flow and node j is labeled $[+q_j, i]$. It remains to be seen when this occurs. Let us assume that the arc (i, j) is already assigned a flow $f_{ij} \geq 0$ ($f_{ij} \leq U_{ij}$). Obviously, the value q_j cannot exceed the residual capacity $U_{ij} - f_{ij}$. This is not enough to mark node j , since from node i it is not always possible to obtain $U_{ij} - f_{ij}$ flow units. Note that you can send as many flow units to node j as are added to node i , that is, at most q_i . Therefore, the flow along a straight arc (i, j) can be increased by the amount q_j , where $q_j = \min[q_i, U_{ij} - f_{ij}]$.

In the same way, you can mark node j if the arc (j, i) is a reverse flow. Here the following question arises: is it possible to reduce the flow along the arc (j, i) ? Obviously, this is only possible if $f_{ji} > 0$. This flow can be reduced at most by the number of flow units that can be taken from node i , that is, by the value q_i . Consequently, the flow along the reverse arc (j, i) can be reduced by the amount q_j , where $q_j = \min[q_i, f_{ji}]$.

The marking algorithm works as follows. First, the source is assigned the mark $[\infty, -]$, indicating that a flow of infinitely large magnitude can flow from this node. Next, we look for the augmented flow path from the source to the sink, passing through the marked nodes. All nodes other than the source are initially unlabeled. Trying to reach the sink, we walk along forward and backward arcs and sequentially mark the nodes belonging to them [7]. The following two cases are possible:

1. The stock t is labeled $[+q_t, k]$. In this case, the augmented flow path is found and the flow along each arc of this path can be increased or decreased by the amount q_t . After changing the arc flows, the current marks are erased and the entire procedure described above is performed again.

2. Stock t cannot be labeled. This means that the augment flow path cannot be found. Consequently, the constructed arc flows form an optimal solution (maximum flow).

To illustrate the operation of the labeling algorithm for the network shown in Table 1, the maximum flow that can flow from node s to node t will be found. Each arc (i, j) is assigned a mark $[f_{ij}, U_{ij}]$, where f_{ij} is the current value of the arc flow, and U_{ij} is the arc capacity. Initially, the magnitudes of all arc fluxes are assumed to be zero. When performing each iteration, you need to mark the stock t . This problem is solved in 6 iterations, the results of each of which are shown in Table 1.

Table 1 – Operation of the marking algorithm

Iteration number	Steps	Description of the procedure	Graph
1	1	Assign to node s the mark $[\infty, -]$	
	2	Assign to node 2 the mark $[+3, s]$	
	3	Assign to node t the mark $[+2, 2]$	
	4	Changing arc flows: $f_{s2} = 2, f_{2t} = 2$	
2	5	Assign to node s the mark $[\infty, -]$	
	6	Assign to node 1 the mark $[+2, s]$	
	7	Assign to node 2 the mark $[+2, 1]$	

	8	Assign to node 3 the mark $[+1, 2]$	
	9	Assign to node t the mark $[+1, 3]$	
	10	Changing arc flows: $f_{s1} = 1, f_{12} = 1, f_{23} = 1, f_{3t} = 1$	
3	11	Assign to node s the mark $[\infty, -]$	
	12	Assign to node 1 the mark $[+1, s]$	
	13	Assign to node t the mark $[+1, 1]$	
	14	Changing arc flows: $f_{s1} = 2, f_{1t} = 1$	
4	15	Assign to node s the mark $[\infty, -]$	
	16	Assign to node 2 the mark $[+1, s]$	
	17	Assign to node 1 the mark $[-1, 2]$	
	18	Assign to node t the mark $[+1, 1]$	
	19	Changing arc flows: $f_{s2} = 2, f_{12} = 0, f_{1t} = 2$	
Iteration number	Steps	Description of the procedure	Graph
5	20	Assign to node s the mark $[\infty, -]$	
	21	Assign to node 3 the mark $[+1, s]$	
	22	Assign to node t the mark $[+1, 3]$	
	23	Changing arc flows: $f_{s3} = 1, f_{3t} = 2$	
6	24	Assign to node s the mark $[\infty, -]$	
	25	None of the nodes can be marked, so the maximum flow is 6	

Conclusions and prospects for further research. The Ford-Fulkerson algorithm is one of the most famous algorithms for solving maximum flow problems. The paper presents a general description of the problem, a mathematical formulation and a description of the algorithm for solving the problem. Using maximum flow models can help businesses improve their efficiency and profitability, calculate flow patterns to reduce costs and provide better customer service.

References

1. Saaty T.L. & Kearns K.P. (1985) Analytical Planning. The Organization of Systems. Pergamon Press. Oxford, New York, Toronto, Sydney, Frankfurt.
2. Phillips D.T. & Garcia-Diaz A. (1981) Fundamentals of Network Analysis. Prentice-Hall, Inc., Englewood Cliffs, N.J.
3. Anderson J.A. (2001) Discrete Mathematics with Combinatorics. Prentice Hall, Upper Saddle River, New Jersey.
4. Murtagh B.A. (1981) Advanced Linear Programming: Computation and Practice. McGraw-Hill International Book Company, New York, London, Madrid.

5. Dymova H. (2024) Development of a Software Application Algorithm for Solving Computer Network Optimization Problems. *Débats scientifiques et orientations prospectives du développement scientifique: c avec des matériaux de la VI conférence scientifique et pratique internationale*, Paris, 1er Mars 2024. Paris-Vinnytsia: La Fedeltà & UKRLOGOS Group LLC. DOI: <https://doi.org/10.36074/logos-01.03.2024.051>.
6. Dymova H. & Larchenko O. (2020) Rozrobka komp'yuternoyi prohramy rozv'yazannya zadach merezhevoyi optymizatsiyi [Development of a Computer Program for Solving Network Optimization Problems]. *Computer-Integrated Technologies: Education, Science, Production*. Lutsk. No. 41. DOI: <https://doi.org/10.36910/6775-2524-0560-2020-41-23>. [in Ukrainian].
7. Dymova H. (2023) Application of Characterization Analysis Methods to Investigation of Logical Networks Structures. *Theoretical and Empirical Scientific Research: Concept and Trends with Proceedings of the V International Scientific and Practical Conference*. Oxford, United Kingdom: European Scientific Platform. DOI: <https://doi.org/10.36074/logos-23.06.2023.34>.

Список бібліографічного опису

1. Saaty T.L. & Kearns K.P. (1985) *Analytical Planning. The Organization of Systems*. Pergamon Press. Oxford, New York, Toronto, Sydney, Frankfurt.
2. Phillips D.T., Garcia-Diaz A. (1981) *Fundamentals of Network Analysis*. Prentice-Hall, Inc., Englewood Cliffs, N.J.
3. Anderson J.A. (2001) *Discrete Mathematics with Combinatorics*. Prentice Hall, Upper Saddle River, New Jersey.
4. Murtagh B.A. (1981) *Advanced Linear Programming: Computation and Practice*. McGraw-Hill International Book Company, New York, London, Madrid.
5. Думова Н. (2024) Development of a Software Application Algorithm for Solving Computer Network Optimization Problems. *Débats scientifiques et orientations prospectives du développement scientifique: c avec des matériaux de la VI conférence scientifique et pratique internationale*, Paris, 1er Mars 2024. Paris-Vinnytsia: La Fedeltà & UKRLOGOS Group LLC. DOI: <https://doi.org/10.36074/logos-01.03.2024.051>.
6. Думова Г., Ларченко О. (2020) Розробка комп'ютерної програми розв'язання задач мережевої оптимізації. *Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво"*, (41). DOI: <https://doi.org/10.36910/6775-2524-0560-2020-41-23>.
7. Думова Н. (2023) Application of Characterization Analysis Methods to Investigation of Logical Networks Structures. *Theoretical and Empirical Scientific Research: Concept and Trends with Proceedings of the V International Scientific and Practical Conference*. Oxford, United Kingdom: European Scientific Platform. DOI: <https://doi.org/10.36074/logos-23.06.2023.34>.

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-55-41>

UDC 004.056:004.7:004.9

Larchenko Oksana, Candidate of Agricultural Sciences, Associate Professor,

ORCID: <https://orcid.org/0000-0001-7857-0802>

Kherson State Agrarian and Economic University, Kherson, Ukraine

THEORETICAL ASPECTS OF DATA CONFIDENTIALITY IN INFOCOMMUNICATION NETWORKS

Larchenko O. Theoretical Aspects of Data Confidentiality in Infocommunication Networks. Data confidentiality in information and communication networks is a hot topic in the modern digital world. With the development of infocommunications and global information infrastructure, the volumes of transmitted and processed information have acquired an unpredictable scale. In this context, ensuring data privacy becomes a vital concern.

A growing number of cyber threats, including hacking, phishing, data theft and privacy breaches, pose risks to the security and privacy of information transmitted over networks. Known cases of leakage of confidential data, such as personal information, commercial and state secrets, and medical records, demonstrate the need for effective means and methods of protection.

In addition, public awareness regarding data privacy and confidentiality is growing. Users are becoming more aware and aware of the risks associated with storing and transmitting their personal data online. This is driving demands for organizations and companies to ensure strong privacy protections for their customers.

In recent years, there has been a significant increase in the amount of digital data stored and transmitted through information and communication networks, thanks to new technologies such as the Internet of Things, cloud computing, social networks, e-commerce and other digital technologies that have become an integral part of our daily lives. This creates a need to reliably protect these large volumes of data from unauthorized access and malicious activity.

The number of cyber threats is constantly growing, and attackers are constantly developing new methods of attack and cybercrime, which jeopardizes data confidentiality. From hacking attacks, phishing, malware, to identity theft and business disclosure, digital threats are becoming more complex and dangerous. Ensuring data confidentiality is becoming a critical need to avoid financial losses, privacy breaches and other negative consequences.

Keywords: cyber threat, infocommunications, infocommunication networks, confidential information, insider threats.

Ларченко О. Теоретичні аспекти конфіденційності даних в інфокомунікаційних мережах. Конфіденційність даних в інфокомунікаційних мережах є надзвичайно актуальною темою в сучасному цифровому світі. З розвитком інфокомунікацій та глобальної інформаційної інфраструктури, обсяги інформації, що передається та обробляється, набули непередбачуваного масштабу. У цьому контексті забезпечення конфіденційності даних стає життєво важливим завданням.

Зростаюча кількість кіберзагроз, включаючи хакерські атаки, фішинг, крадіжки даних та порушення приватності, створює небезпеку для безпеки та конфіденційності інформації, що передається через мережі. Відомі випадки витоку конфіденційних даних, такі як особисті інформації, комерційні та державні секрети, медичні записи, демонструють потребу в ефективних засобах та методах захисту.

Окрім того, суспільне усвідомлення, щодо приватності та конфіденційності даних росте. Користувачі стають більш обізнаними та свідомими щодо ризиків, пов'язаних зі збереженням та передачею їхніх особистих даних в мережі. Це стимулює вимоги до організацій і компаній забезпечувати надійний захист конфіденційності даних своїх клієнтів.

В останні роки спостерігається значне збільшення кількості цифрових даних, які зберігаються та передаються через інфокомунікаційні мережі, завдяки таким новим технологіям як Інтернет речей, хмарні обчислення, соціальні мережі, електронна комерція та інші цифрові технології, що стали невід'ємною частиною нашого повсякденного життя. Це створює потребу в надійному захисті цих великих обсягів даних від несанкціонованого доступу та зловмисних дій.

Кількість кіберзагроз постійно зростає, зловмисники постійно розвивають нові методи атак та кіберзлочинності, що ставить під загрозу конфіденційність даних. Від атак хакерів, фішингу, шкідливих програм до крадіжок особистих даних та розкриття комерційної інформації - цифрові загрози стають все більш складними та небезпечними. Забезпечення конфіденційності даних стає критичною необхідністю для уникнення фінансових втрат, порушення приватності та інших негативних наслідків.

Ключові слова: кіберзагроза, інфокомунікації, інфокомунікаційні мережі, конфіденційна інформація, інсайдерські загрози.

Formulation of the problem. Data privacy protection has become the subject of strict legislation and regulation. Many countries and regions are establishing new regulations, such as the General Data Protection Regulation (GDPR) in the EU, to regulate the collection, storage, processing and transfer of personal data. Companies must meet these requirements because data breaches can have serious consequences, such as fines and loss of reputation.

Infocommunications is a modern concept that combines telecommunications technologies with information and computer systems. This is an inextricable connection between different elements that ensure the transfer and exchange of information. Infocommunications develops through the convergence of various technologies such as signaling, routing, signal conditioning and programming, ensuring optimal information processing.

Infocommunication networks are a complex set of technical means, including elements of infocommunications, structures and routing systems. They are designed to transmit, receive and exchange various types of signals, messages, text, images and sound through various types of communication systems such as radio, wired and optical.

The global information infrastructure includes networks and infocommunication systems that connect communication nodes, computers and electronic devices to ensure the transfer of a variety of information. It creates the basic infrastructure for organizing various infocommunication services and ensures communication between users around the world.

Given the rapid technological development and growing dependence on information systems, data confidentiality in information and communication networks is becoming increasingly important. The need to protect personal information, business data and other sensitive data requires the development of effective security methods and tools on these networks.

In the modern world of information technology and the Internet, the importance of issues related to data confidentiality is growing. Ensuring data confidentiality is becoming a critical issue for ensuring the privacy of users, businesses and governments in the digital environment.

Research analysis. Scientists and specialists in the field of information security and cryptography are actively researching and developing new methods and algorithms to ensure data confidentiality in information and communication networks. Ongoing efforts are aimed at improving encryption, authentication and access control systems to ensure that information is protected from unauthorized access and malicious activity. The purpose of the article is to study data confidentiality in information and communication networks and analyze the means of ensuring it. The object of the study is data confidentiality in information and communication networks. The subject of the study is ways to ensure data confidentiality in information and communication networks. Including methods of encryption, authentication, access control, threat and risk analysis.

Presentation of the main material and justification of the obtained results. Confidential information is a set of data that is in the possession, use or disposal of individual individuals or legal entities and is distributed at their request.

The basic principles of information security are confidentiality, integrity and availability.

1. Confidentiality: This principle deals with the confidentiality of information, which means that only authorized persons have access to confidential data. Information must be protected from unauthorized access or disclosure. This is achieved through encryption, authentication and access control.

2. Integrity: This principle requires maintaining the integrity of information, i.e. protection against unauthorized changes or modifications. Information must be stored in a permanent state and reliably protected from unauthorized interference. To ensure integrity, data integrity monitoring methods, backups, and intrusion detection mechanisms are used.

3. Availability: This principle means that information and resources should be available to authorized users at the right time. Information systems must be resilient to operational failures, technical problems, or malicious attacks to ensure that information is always available. Measures such as backup, replication and recovery systems are used to ensure availability.

According to Article 21 of the Law of Ukraine «On Information», confidential information, together with official and secret information, refers to information with limited access.

Restricted information is a set of data, access to which is limited only to a certain circle of persons, and its disclosure is prohibited by the information manager by law. Restriction of access to information is carried out for the purpose of national security or protection of the legal rights of individuals and legal entities. It is important to note that access to the information itself is restricted, not the document. Thus, if a document contains both open and restricted information, then the open information can be provided as a separate document to the interested party for review. [1].

According to Article 6 of the Law of Ukraine «On Access to Public Information», information with limited access may be as follows:

1. Confidential information: this is data to which access is limited to an individual or legal entity other than those in authority, and can be disseminated in accordance with their wishes and the conditions determined by them.

2. Secret information: this is data to which access is limited and the disclosure of which could cause harm to a person, society or state. A secret is information that contains state, professional, bank secrets, investigative secrets and other secrets provided for by law.

3. Official information: this is data contained in documents of subjects of power and constituting internal official correspondence, memos, recommendations related to the development of the direction of the institution's activities or the implementation of control and supervisory functions by public authorities, the decision-making process, precedes public discussion or adoption decisions. It may also contain data collected in the process of operational-search, counterintelligence activities and in the field of national defense that are not related to state secrets. [1].

Confidential information owned by the state and in the use of state authorities, local governments, enterprises, institutions and organizations of all forms of ownership does not include information about:

1. State of the environment, quality of food products and household items.
2. Accidents, catastrophes, natural hazards and other emergencies that have occurred or may occur and threaten the safety of citizens.
3. The state of health of the population, its standard of living, including food, clothing, housing, medical care and social security, as well as socio-demographic indicators, the state of law and order, education and culture of the population.
4. The state of affairs with the rights and freedoms of man and citizen, as well as facts of their violations.
5. Illegal actions of state authorities, local governments, their officials and officials.

Activities of state and municipal unitary enterprises, business societies, in the authorized capital of which more than 50 percent of the shares (shares) belong to the state or territorial community, as well as business companies, 50 or more percent of the shares (shares) of which belong to the business company, the share of the state or territorial society in which it is 100 percent, information about these enterprises is subject to mandatory disclosure in accordance with the law.

Other information, access to which in accordance with the laws of Ukraine and international treaties, the consent to be bound by which is provided by the Verkhovna Rada of Ukraine, is limited. These are some alternative formulations for defining confidential and restricted information under Ukrainian law.

Confidential information can be divided into information relating to individuals and information relating to legal entities.

The basis for recognizing information as confidential is the desire of an individual or legal entity to consider certain information about himself or another person in his possession as confidential. The terms «personal data» and «identity information» are used interchangeably. Personal information may concern both individuals and legal entities. However, not all information about a person is confidential unless prohibited by law.

Certain categories of information relating to individuals are not considered restricted information, in particular, income statements of individuals and members of their families who are applying for or already occupy an elected position in government or are civil servants, employees of local government bodies of the first or second category.

Information with limited access also does not include income declarations of persons and members of their families who apply for a position or hold an elected position in government bodies or hold the position of a civil servant, employee of a local government body of the first or second category. This information is publicly available and subject to disclosure to ensure transparency and prevent conflicts of interest.

Declarations of income of persons holding important positions in state or local authorities are an important tool for monitoring the distribution and use of public resources. These declarations contain information about income, property, financial obligations and other important information that helps to check the declarant for the absence of conflicts of interest and possible corrupt behavior.

Such disclosure of information promotes transparency in government structures and helps the public monitor possible inconsistencies and abuses in the sphere of public service. It helps to establish trust in authorities and ensures public control over their activities.

Threats to data confidentiality in information and communication networks. Analyzing the potential threats that arise during the processing of personal data, taking into account the wide range of possibilities for gaining access to information, the following classes of threats can be distinguished:

1. Insider or insider threats arising from the capabilities of internal users. These risks are associated with the possibility of theft or alteration of information by persons who have access to information systems through their official positions or work in a company or government agency.
2. External threats that arise when a threat actor gains unauthorized access to the protected object using the capabilities of public networks.

3. Technical threats that arise when using hardware and programs designed to steal electronic information.

All of these types of threats require serious analysis of personnel who have access to sensitive information, as well as a focus on the technical security of data storage and processing facilities using specialized software, information security management systems (DLP), security event and incident management systems (SIEM)) and other means of protecting information resources. [2].

Threats based on different types of equipment used are also taken into account:

1. Data security threats processed by an employee at his automated workstation (AMW), not connected to the Internet.

2. Threats to information processed by an employee at his workplace connected to the Internet.

3. Threats to information arrays processed in local networks not connected to the Internet.

4. Threats to data arrays processed in local networks of enterprises and organizations connected to the Internet.

5. Security threats to data processed in distributed networks of operators, which may or may not be connected to Internet networks.

In the security threat model, there is also a classification associated with different types of technical means used to access protected amounts of information.

This classification includes:

1. Use of malware, viruses, worms and the like that are created for illegal purposes.

2. Loss of data through technical and physical leakage channels.

3. Other special effects.

Based on the types of security and technical vulnerabilities installed in a particular operator, the threat model identifies the following categories:

– Related to system software vulnerabilities.

– Associated with the use of application software deficiencies.

– Related to the possibilities of using hardware bookmarks.

– Related to the use of communication tools and information transfer protocols.

– Associated with the use of technical data transmission channels, such as telephone networks, power supply networks and others.

This classification allows you to identify different types of threats and vulnerabilities associated with the use of technical means and implement appropriate security measures to prevent them. [3].

It follows from practice that most information leaks from protected areas occur through technical data transmission channels. A signal propagating through a physical environment can be acoustic or electromagnetic, and can be intercepted using embedded devices and other methods. These devices can intercept electromagnetic radiation, acoustic and visual information. Protection against this type of interception due to restricted access to the protected object. Unauthorized access to data from employee owners is very weak and can be completely eliminated.

The following methods can be used to harm data:

1. Use of standard software that allows you to access the operating system.

2. Creation of uncontrolled operating conditions that allow the resulting distortions to be used to modify data.

3. Use of malware.

4. Threats related to remote access to the system.

5. Combined threats.

Characteristics of information systems that influence the emergence of new threats and risks include:

1. The volume and content of information contained in the database.

2. System structure and configuration.

3. Connecting the system to public information transmission networks or networks with cross-border communication.

4. Availability and quality of protection systems.

5. Personal data processing mode.

6. Levels of access to data for persons with different powers and tasks.

7. Physical location of technical devices and mode of protection against illegal access.

Among modern cyber threats the following stand out:

1. Social engineering and phishing – the use of manipulative methods to obtain confidential information.
2. Virus software - malicious programs that can damage or hack the system.
3. Using outdated versions of software with known vulnerabilities.
4. Insider threats – threats associated with unauthorized access or use of information by employees of one's own organization.
5. Lack of policies and procedures for handling information resources.

Social engineering. Social engineering is based on exploiting human weaknesses. Through successful psychological manipulation, a sophisticated attacker can gain access to many important aspects of an organization to plan a hack and steal information. These could be physical access control systems, security operating hours, cleaning schedules, location of printers, trash bins, availability of shredders, etc. Preparation for hacking an organization's information system begins with this type of work.

Phishing attacks are a continuation of social engineering and are one of the most common and effective methods of malicious and illegal access to the resources of businesses and organizations of all forms of government. According to various estimates, up to 90% of successful cyber-attacks are related to phishing. This method is quite simple and is based on the use of fragmented electronic sheets that manipulate the victim by running virus programs, for example, under office applications, or by clicking on fragmented messages that lead to forgery of their websites, where you are asked to enter a login and password before shipment or other resources. [4].

Virus software is one of the most important weapons of attackers, against which it is necessary to have an up-to-date version of anti-virus software. There is no universal antivirus software that would provide equally effective protection against all types of viruses. However, practice shows that it is very useful to exchange information between specialists regarding the beginning of attacks and the emergence of new viruses, as well as the use of services for analyzing suspicious files and links to detect worms, Trojans and other malicious software.

Updating your software is key to avoiding cyber threats. *Constantly updating your software* is of great importance, as it will ensure an increased level of security and protection against new vulnerabilities and threats. It is important to consult with staff, make regular updates and make assumptions, and illustrate with specific questions what risks may arise from using an outdated version of the software. [5].

Insider threats are a significant group of threats emanating from one's own employees. This often happens in organizations where there is no high-level control over the provision of access rights and delimitation of access to information resources. The ideal situation would be if there is a systematic assessment of employee loyalty during recruitment, employment and termination. Considering that personnel are always the weak point in the security system, it is important to constantly work on increasing the security consciousness of employees [5].

The lack of policies and procedures for handling information assets creates serious threats. Policies and procedures define the rules for operating an organization's information system, as well as the distribution of roles and responsibilities. The absence of such policies and procedures leads to chaos, possible financial losses due to inefficient use of resources, increased risk of personnel errors, inadvertent disclosure and loss of information. Having the right policies and procedures allows users to make their own decisions and work within clearly defined boundaries. This includes the proper procedure for granting access rights, following basic information security rules, such as limiting access to necessary resources.

With appropriate security policies in place, there are fewer problems with responsibilities, role allocation, resource protection, and there is a clear procedure for reporting cybersecurity incidents and assessing possible business consequences. All this can be predicted ahead of time by conducting a risk assessment procedure and analyzing the consequences of incidents for business [3].

Technical means of ensuring data confidentiality in information communication networks. At the technological level, the information security policy is implemented through the implementation of a complex of modern automated information technologies. [4] Technical means ensure the confidentiality of data in information communication networks by applying various methods of encryption and information protection.

Here are a few such tools:

1. Virtual Private Networks (VPN): VPN provide secure connections and data transfers over unsecured networks by encrypting traffic. They provide data privacy by creating an encrypted tunnel between remote networks or devices.

2. Data Encryption: Encryption uses mathematical algorithms to convert plain text into encrypted text that is unintelligible to unauthorized persons. Various encryption methods are used, such as symmetric encryption (AES, DES) and asymmetric encryption (RSA, ECC).

3. Firewalls: Firewalls are the first level of network protection and control traffic passing through network nodes. They set traffic filtering rules that allow or block certain types of connections and communications based on predefined security rules.

4. Identification and authentication: Technical means, such as access control systems, two-factor authentication (2FA), biometric systems, ensure the identification of users and ensure their authenticity before providing access to confidential data.

5. Integrated Security Systems: These are comprehensive solutions that combine several technical tools, such as intrusion detection systems (IDS), intrusion prevention systems (IPS), antivirus programs, access control systems and others, to provide comprehensive network and data protection.

6. Link Encryption (SSL/TLS): Link encryption is used to protect data transmitted between a client and a server over the Internet. The SSL (Secure Sockets Layer) and TLS (Transport Layer Security) protocols encrypt data and provide server authentication, which avoids data interception and hacking.

7. Other technologies: Other technical means of data privacy protection include access control systems to network resources, file and folder encryption systems, data backup and recovery systems, and others.

These technical means are used to ensure data confidentiality in information communication networks and reduce the risk of unauthorized access to valuable information. A combination of several tools and proper configuration help ensure a high level of data security. [4]

Conclusions and prospects for further research. The article examined the theoretical aspects of data confidentiality in infocommunication networks. The study began with a definition of data privacy, which points to the importance of protecting information in today's digital world. The threats to data confidentiality encountered in information and communication networks are also analyzed. This includes data interception, unauthorized access, password cracking, network protocol vulnerabilities and others.

Understanding these threats allows you to identify the main attack vectors and take effective measures to prevent them.

Technical means to ensure data confidentiality, such as encryption, access control and authentication, are also presented. Virtual private networks (VPN) have been studied as an effective means of ensuring data confidentiality. Various protocols for ensuring data confidentiality and their suitability for use in information communication networks are analyzed. The advantages and disadvantages of existing methods for ensuring data confidentiality and recommendations for their practical application are identified.

Overall, research in the field of data privacy in information communication networks has great potential for improving information security and protecting user privacy. With the advancement of technology and the challenges of the digital age, further research and development in this area is essential.

References

1. Dokument: Zakon Ukrainy «Pro zakhyst personal'nykh danykh» № 2297-VI. URL: <https://www.president.gov.ua/documents/2297vi-11567> [in Ukrainian]
2. Lahun A. E. (2019) Kryptohrafichni systemy ta protokoly: nav. posibnyk L'viv: Vydavnytstvo L'vivs'koyi politekhniki. [in Ukrainian]
3. Hryhorenko O.H., Holub O.S. (2021) Konfidentsiynist' danykh v infokomunikatsiynyykh merezhakh ta zasoby yiyi zabezpechennya. Zbirnyk «Perspektyvy telekomunikatsiy» PT-2023. K.: KPI im. Ihorya Sikors'koho. [in Ukrainian].
4. Konakhovych H.F., Klymchuk V.P., Pauk S.M., Potapov V.H., Chuprin V.M., Horbunov O.O. (2009) Zakhyst informatsiyi v telekomunikatsiynyykh systemakh: Navchal'nyu posibnyk. Kyiv: NAU. [in Ukrainian]
5. Seniv M. M., V.S. Yakovyna. (2019) Bezpeka program ta danykh: navch. posibnyk. L'viv: Vydavnytstvo L'vivs'koyi politekhniki. [in Ukrainian]

Список бібліографічного опису

1. Документ: Закон України «Про захист персональних даних» № 2297-VI. URL: <https://www.president.gov.ua/documents/2297vi-11567>
2. Лагун А. Е. (2019) Криптографічні системи та протоколи: навч. посібник Львів: Видавництво Львівської політехніки.
3. Григоренко О.Г., Голуб О.С. (2021) Конфіденційність даних в інфокомунікаційних мережах і засоби її забезпечення. Збірник «Перспективи телекомунікацій» ПТ-2023. К.: КПІ ім. Ігоря Сікорського.
4. Коначович Г.Ф., Климчук В.П., Паук С.М., Потапов В.Г., Чуприн В.М., Горбунов О.О. (2009) Захист інформації в телекомунікаційних системах: Навчальний посібник. Київ: НАУ.
5. Сенів М. М., В.С. Яковина. (2019) Безпека програм та даних: навч. посібник. Львів: Видавництво Львівської політехніки

CONTENTS

AUTOMATION AND MANAGEMENT	
Aloshyn S.V. Implementation of the educational load calculation in terms of the student count per academic discipline by the SQL language	5
Dubuk V., Kovivchak Ya., Svystun B. Development of an automated system for ordering motorcycle and bicycle parts	15
INFORMATICS AND COMPUTER SCIENCE	
Turchyn O. Implementation of Machine Learning to Increase the Accuracy of Forecasting the Operating Modes of Deep-Sea Pumping Stations	24
Anisimov V., Kunanets N. Transition from Monolithic to Microservice Architecture: Methodology and Implementation Experience	30
Bahniuk N., Khrystynets N., Bykov S. Development of multimedia 3D models for technological process simulation.	42
Vynnyshyn O. A Comprehensive Methodology for Optimizing Vhdl Program Code when Designing a Programmable Integrated Circuit.	48
Volynets L., Harmatiuk N., Derevianko V., Duda O., Kramar T. Skaletskyi P. Development of a museum mobile application concept with elements of augmented reality	55
Horhun R., Saltovskiy B. Interactive Alarm Map of Ukraine.	66
Demchenko A., Fedosova I. Research of information and communication platforms for distance learning	71
Didus A., Tereikovskiy I. Formalization of the keyword recognition process in speech signal.	78
Dmitrenko T. Application Of Uavs Based On The Wavelet Transform Technology In Military Operations: Scouting, Target Location And Navigation	87
Dobryshyn Yu. Formalization of the design process of decision-making support systems for increasing of the level of software protection	93
Zhuravska I., Fisun M. Risks of information leaks from mobile devices.	100
Kolyada K., Rozhko D. Using metaheuristic algorithms to solve the Traveling Salesman Problem.	107
Korovii O., Tereikovskiy I. Conceptual model of the process of determining the emotional tonality of the text	115
Lukashuk M. Standardization Of Web Accessibility Of Information And Communication Technologies As A Way To Attract More Users	124
Maruniak S. Detecting and mitigating security vulnerabilities in dynamic routing protocols: current challenges and solutions.	130
Marchenko O. Method of the MCTS Tree Search Method Parallelization	137
Melnyk V., Bahniuk N., Roiko O., Bortnyk K., Kizym S. Socket performance influence on data processing intensity in a virtual machine cluster with heterogeneous conditions	143
Moroz B., Kruhlyk A. Conceptual model of information flow processing in the system of unmanned aerial vehicles	155
Morokhovych V., Liakh I., Khomyak M., Morokhovych B. Predicting passengers who survived the Titanic disaster using a decision tree.	161
Nyzhnyk O., Yashchynskyy L., Fedosov S. The principle of operation and prospects for the use of holographic displays	167

Pryhoda A. Evaluation Of The Effectiveness Of The Crm System Development And Implementation Project Based On Microservice Architecture.	172
Reshetniak V., Faure E. Eye-tracking as a tool for researching user behavior.	181
Rozlomii I., Symonyuk V., Naumenko S., Mykhailovskyi P. The security model of interconnected computing devices based on a lightweight encryption scheme for IoT	191
Sedinkin O., Derkach M., Skarga-Bandurova I., Matiuk D. Eye tracking system based on machine learning	199
Fil N. Model for Selecting a Temperature and Humidity Sensor for a Module of an Environmental Monitoring System.	206
Khambir V. Automation of mobile application testing processes	213
Khrystynets N., Viktorova Y., Zhukovska S., Radion Yu. OS file manager development.	225
Chybiriak Ya. I., Zakharova A. M., Lavrov E. A., Siryk O. E., Baranova I. V., Tatarchuk R. M. Computer modeling for disaster medicine. An approach to automated analysis of solutions for organizing mobile medical hospitals.	232
Shvalikovskiy D. Two problems of heat equation in two-component systems	243
Shenheliia V., Zhdanova O. Path planning for UAVs with a moving maintenance center.	251
Yakobchuk B., Khrystynets N. Creating a simple voice assistant in an application for the Android platform	260
TELECOMMUNICATIONS AND RADIO ENGINEERING	
Bakhovskyi P., Yevsyuk M. Analysis of the energy balance of radio lines during data exchange in SAE / EPS systems	265
Vasykivskiy M., Kolomiets A., Hrabchak N., Hrytsaiuk D., Kostianin V. Python software tools for modeling telecommunication devices.	270
Lyshuk V., Moroz S., Zablotyskiy V., Yevsiuk M., Hryceniuk V. Algorithms for controlling autonomous voltage inverters of frequency converters	279
Madinov M. Optical communication lines	286
PROJECT MANAGEMENT	
Deineko L., Kunanets N. The role of UML diagrams in planning an information system project on the example of a hiking tourist routes system	293
Sadovyi I. The Comparative Analysis Of Distributed Suffix Trees And Traditional Data Management Methods.	301
Dymova H., Larchenko O. Using the Max Flow Problem for Business Processes	309
Larchenko O. Theoretical Aspects of Data Confidentiality in Infocommunication Networks	315

ВИМОГИ ДО СТРУКТУРИ ТА ОФОРМЛЕННЯ МАТЕРІАЛУ СТАТЕЙ

- **Наукова стаття обов'язково повинна мати наступні необхідні елементи:**
 - 1) **постановка проблеми** у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями;
 - 2) **аналіз останніх досліджень і публікацій**, в яких започатковано розв'язання даної проблеми і на які спирається автор,
 - 3) **виділення невирішених раніше частин загальної проблеми**, котрим присвячується означена стаття;
 - 4) **формулювання мети дослідження** (постановка завдання);
 - 5) **виклад основного матеріалу дослідження** з повним обґрунтуванням отриманих наукових результатів; **висновки** з даного дослідження, у тому числі з науковою новизною;
 - 6) **висновки та перспективи подальших досліджень** у даному напрямку.
- Статтю можна подавати українською або англійською мовами. Вона повинна бути набрана у текстовому редакторі MS WORD. **Нумерацію сторінок** не виконувати. **Обсяг статті** 5 повних сторінок і більше (рекомендовано 5-12 ст.).
- **Параметри сторінки.** Верхнє та нижнє поле –1,5 см, лівє – 2,5 см, правє поле – 2 см. Від краю до верхнього колонтитула – 1,25 см, нижнього – 1,25 см. Дзеркальні поля. Папір розмір А4.
- У верхньому колонтитулі по центру шрифтом Time New Roman розміром 11 пт з одинарним міжрядковим інтервалом текст:

*Науковий журнал "Комп'ютерно-інтегровані технології: освіта, наука, виробництво"
Луцьк, 202*. Випуск № ***
- У нижньому колонтитулі по лівому краю шрифтом Time New Roman розміром 10 пт з одинарним міжрядковим інтервалом прізвища авторів статті: @ Прізвище1 І.П., Прізвище2 І.П...
- **Шапка статті містить наступні рядки** шрифтом Time New Roman розміром 11 пт з одинарним міжрядковим інтервалом та вирівнюються по лівому краю:

DOI: (заповнюється редактором),
УДК: (заповнюється автором самостійно),
Прізвище, ім'я та по батькові 1 автора (напівжирним виділенням), вч.ступінь, вч.звання
ORCID 1 автора (у форматі <http://orcid...>),
Прізвище, ім'я та по батькові 2 автора (жирним накресленням), вч.ступінь, вч.звання
ORCID 2 автора (у форматі <http://orcid...>)
і т.д. (кількість авторів однієї публікації не більше 5)
Назва організації (у форматі- назва повністю, місто, країна)
- **Назва статті** розміщується через один рядок нижче назви організації (усі великі літери, розмір шрифту Time New Roman 11 пт з напівжирним виділенням та вирівнюванням по центру).
- **Анотації** (українською та англійською мовами) повинні містити прізвища та ініціали авторів, назву статті та короткий її зміст і розміщуються через один рядок нижче назви статті та набираються з абзацного відступу 1,25 см шрифтом Time New Roman розміром 9 пт з одинарним міжрядковим інтервалом і вирівнюються по ширині. Кожна анотація обсягом 200 слів та формується одним абзацом. Нижче анотацій обов'язково вказуються **ключові слова** (кількість від 5 до 10). У англійській анотації вказуються прізвища та ініціали імені авторів (без по батькові).
- **Основний текст** статті розміщується через один рядок нижче анотацій, набирається з абзацного відступу 1,25 см шрифтом Time New Roman розміром 11 пт з одинарним міжрядковим інтервалом та вирівнюється по ширині.
- **Формули** набираються у редакторі формул MS WORD (використовувати шрифти: Symbol, Time New Roman Суг; розміри шрифтів: звичайний 12 пт, крупний індекс 7 пт, дрібний індекс 5 пт, крупний символ 18 пт, дрібний символ 12 пт). Формула вирівнюється по центру і не повинна займати більше 5/6 ширини рядка.
- **Ілюстрації**, що присутні у статті, необхідно розташовувати у тексті по центру, вирівнюючи підписи по центру (Рис. 1. Назва). Ілюстрації повинні бути чіткими та контрастними.
- **Таблиці** потрібно розташовувати у тексті по центру, причому їх ширина повинна бути на 1 см менша ширини рядка. Над таблицею ставиться її порядковий номер і назва (Таблиця 1. Назва) та вирівнюється по ширині з абзацного відступу.

- **Посилання** на ту чи іншу роботу повинні позначатися в тексті у квадратних дужках за порядковим номером у списку літератури в кінці статті; посилання на джерела статистичних даних обов'язкові; посилання на публікації дослідників обов'язкові; посилання на власні публікації допускаються тільки у випадку крайньої необхідності.
- **Список бібліографічного опису та References.** У даному виданні не публікуються статті, що ґрунтуються на посиланнях країни-агресора. Список літератури («References») потрібно приводити повністю окремим блоком, повторюючи список літератури, який подається українською мовою, незалежно від того, є в ньому іноземні джерела чи ні. Тобто, після статті подається 2 списки: «Список бібліографічного опису» (звичайний список літератури) і «References» (список для міжнародних БД). Необхідно в опис джерела вносити всіх авторів, не скорочуючи їх до трьох, як це рекомендовано діючими у нас державними стандартами. References – повинен бути укладений англійською мовою або транслітерований. Рекомендуємо скористатись сервісом: <https://www.grafiati.com/uk/>. При оформленні інтернет-джерел задля запобігання довгих нерозривних лінків, користуйтеся сервісом скорочення URL-адрес, наприклад: <https://hyperhost.ua/tools/uk>
- В кінці статті вказується напрямок публікації відповідно до спеціальностей наукового журналу (у журналі виділено такі тематики досліджень: автоматика та управління, інформатика та обчислювальна техніка, телекомунікації та радіотехніка, управління проектами)
- **Стаття обов'язково переслається електронною поштою за адресою: cit@lntu.edu.ua**
- Рукописи, що не відповідають вище вказаним вимогам, не розглядаються і до друку не приймаються.
- **Усі рукописи проходять перевірку на плагіат!**

Довідки з питань публікації та прийому матеріалів у науковий журнал «Комп'ютерно-інтегровані технології: освіта, наука, виробництво» можна отримати у відповідального секретаря – к.т.н. Христинець Наталії Анатоліївни за тел. (0332) 74-61-15.

Адреса: 43018, м. Луцьк, вул. Львівська, 75, ауд. 141

Автор статті отримує електронний примірник збірника.

Вартість однієї сторінки становить 40 грн. для працівників Луцького НТУ, 50 грн – для інших авторів. Окремо кожній статті присвоюється DOI (digital object identifier) – ідентифікатор цифрового об'єкту, що веде за собою додаткову оплату 60 грн за одну статтю.

Кошти можна перерахувати на рахунок: Луцький національний технічний університет <https://lntu.edu.ua/uk> – Онлайн оплата – Інші платні послуги – р/р UA86 820172 0 3132 4 1 002 2 02 017820 – призначення платежу: «За публікацію в журналі КІТ №п від _____ (ПІБ автора)» Без попереднього рецензування та перевірки на плагіат самостійно кошти на рахунок не надсилати.

Зразок оформлення статті на наступній сторінці

DOI:

УДК 004:02

Ковалець Віктор Геннадійович¹, д.т.н., професор

<https://orcid.org/2525-0006-0146-081X>

Золотар Ольга Петрівна², к.т.н., доцент

<https://orcid.org/0005-0005-3007-2462>

Сулій Павло Павлович¹, магістрант

Ліпук Микола Сергійович², студент

¹Національний університет «Львівська політехніка», м. Львів, Україна

²Тернопільський національний технічний університет, м. Тернопіль, Україна

(якщо у статті усі автори з однієї організації, то верхні індекси (цифри) не потрібно, лише вказати під прізвищами одну назву ЗВО, місто та країну)

ПЕРЕХІД ВІД МОНОЛІТНОЇ ДО МІКРОСЕРВІСНОЇ АРХІТЕКТУРИ: АПАРАТНІ МЕТОДИ ВПРОВАДЖЕННЯ

Ковалець В. Г., Золотар О.П., Сулій П.П., Ліпук М.С. Перехід від монолітної до мікросервісної архітектури: апаратні методи впровадження. У статті розглядається проблема переходу від монолітної архітектури до мікросервісної з розподіленою базою... (200 слів)

Ключові слова: монолітна архітектура... (5-10 слів)

Kovalets V., Zolotar O., Suliy P., Lipuk M. Transition from Monolithic to Microservice Architecture: Hardware Implementation Methods. The article addresses the problem of transitioning

Keywords: monolithic architecture

Постановка наукової проблеми. У сучасному світі інформаційних технологій, де системи постійно розвиваються та ускладнюються, питання ефективної архітектури та управління даними стає все більш актуальним.

Таблиця 1. Характеристики мікроконтролерів та апаратних шифраторів

Параметр	Криптографічні бібліотеки	Апаратні шифратори
Продуктивність	Помірна, залежить від потужності мікроконтролера	Висока, завдяки спеціалізованому обладнанню

Крім імпульсів керування транзисторів інвертора мікроконтролер виробляє імпульси керування транзистора VT37. Завдання транзистора VT37 полягає у підключенні гальмівного резистора R33 паралельно конденсатору фільтра C31 (рис. 2).

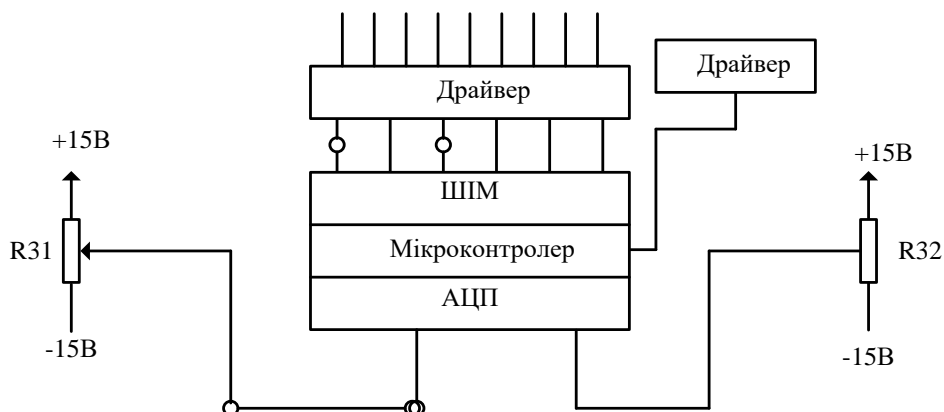


Рис.2. Схема підключення датчика частоти

Висновки та перспективи подальшого дослідження. У даній статті було розглянуто проблему переходу від мікро-архітектури системи команд до мікро-архітектури спрощених вказівок...

Список бібліографічного опису:

1. Долгов О. М. Композиційні матеріали. Дніпро: Дніпр. політехніка, 2022. 67 с. URL: <http://surl.li/swvwq> (дата звернення: 16.03.2024).
2. Марчук В. Алгоритм оцінювання якості 3D моделі для адитивного виробництва. Матеріали та технології в інженерії (МТІ-2023): Зб. наук. доп. міжнар. конф., м. Луцьк, 16–18 трав. 2023 р. Луцьк, 2023. С. 178–180. URL: <http://surl.li/swwge> (дата звернення: 02.09.2024).

References:

1. Dolgov O. M. Composite materials. Dnipro: Dnipro. polytechnic, 2022. 67 p. URL: <http://surl.li/swvwq> (access date: 16.03.2024).
2. Marchuk V. 3D model quality assessment algorithm for additive manufacturing. Materials and technologies in engineering (MTI-2023): Collection. of science add. international conference, Lutsk, May 16–18. 2023. Lutsk, 2023. P. 178–180. URL: <http://surl.li/swwge> (access date: 02.09.2024).

Міністерство освіти і науки України
Луцький національний технічний університет

Колектив авторів

КОМП'ЮТЕРНО-ІНТЕГРОВАНІ ТЕХНОЛОГІЇ: ОСВІТА, НАУКА, ВИРОБНИЦТВО

Науковий журнал

Підп. до друку 20.06.2024. Формат А4. Папір офс.
Гарн. Таймс. Ум. друк. арк. 15.25 Обл. – вид. арк. 15.75
Тираж 20 прим. Зам. № 14/22

Комп'ютерний набір та верстка:

Н.А. Христинець

Зареєстровано Національною радою України з питань телебачення і
радіомовлення як суб'єкт у сфері друкованих медіа
(рішення №40 від 11.01.2024 р., ідентифікатор медіа R30-02456)
43018, м. Луцьк, вул. Львівська, 75