ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СУСПІЛЬСТВО

INFORMATION TECHNOLOGY AND SOCIETY

МАУП

ISSN 2786-5460 (Print)
ISSN 2786-5479 (Online)

МІЖРЕГІОНАЛЬНА АКАДЕМІЯ УПРАВЛІННЯ ПЕРСОНАЛОМ

INTERREGIONAL ACADEMY OF PERSONNEL MANAGEMENT

НАУКОВІ ПРАЦІ МІЖРЕГІОНАЛЬНОЇ АКАДЕМІЇ УПРАВЛІННЯ ПЕРСОНАЛОМ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СУСПІЛЬСТВО

SCIENTIFIC WORKS OF INTERREGIONAL ACADEMY OF PERSONNEL MANAGEMENT

INFORMATION TECHNOLOGY AND SOCIETY

Випуск 3 (9), 2023
Issue 3 (9), 2023

Видавничий дім
«Гельветика»
2023

Рекомендовано до друку Вченою радою
Міжрегіональної Академії управління персоналом
(протокол № 9 від 18 жовтня 2023 року)

Інформаційні технології та суспільство / [головний редактор О. Попов]. – Київ : Міжрегіональна Академія управління персоналом, 2023. – Випуск 3 (9). – 92 с.

Журнал «Інформаційні технології та суспільство» є науковим рецензованим виданням, в якому здійснюється публікація матеріалів науковців різних рівнів у вигляді наукових статей з метою їх поширення як серед вітчизняних дослідників, так і за кордоном.

Редакційна колегія не обов'язково поділяє позицію, висловлену авторами у статтях, та не несе відповідальності за достовірність наведених даних і посилань.

Свідоцтво про державну реєстрацію друкованого засобу масової інформації
«Інформаційні технології та суспільство» Серія КВ № 24815-14755Р від 27.04.2021 р.

Відповідно до Наказу МОН України № 1290 від 30 листопада 2021 року (додаток 3) журнал включено до Переліку наукових фахових видань України (категорія Б) зі спеціальностей 121 – Інженерія програмного забезпечення, 122 – Комп'ютерні науки, 123 – Комп'ютерна інженерія, 124 – Системний аналіз, 125 – Кібербезпека, 126 – Інформаційні системи та технології.

Усі електронні версії статей журналу оприлюднюються на офіційній сторінці видання
http://journals.maup.com.ua/index.php/it

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

Recommended for publication
by Interregional Academy of Personnel Management
(Minutes No. 9 dated 18 October 2023)

Information Technology and Society / [chief editor Oleksandr Popov]. – Kyiv : Interregional Academy of Personnel Management, 2023. – Issue 3 (9). – 92 p.

Journal «Information Technology and Society» is a peer-reviewed scientific edition, which publishes materials of scientists of various levels in the form of scientific articles for the purpose of their dissemination both among domestic researchers and abroad.

Editorial board do not necessarily reflect the position expressed by the authors of articles, and are not responsible for the accuracy of the data and references.

Print media registration certifi cate «Information Technology and Society»
series KV No. 24815-14755P dated 27.04.2021

According to the Decree of MES No. 1290 (Annex 3) dated November 30, 2021, the journal was included in the List of scientific professional publications of Ukraine (category B) in specialties 121 – Software engineering, 122 – Computer sciences, 123 – Computer engineering, 124 – Systems analysis, 125 – Cybersecurity, 126 – Information systems and technologies.

All electronic versions of articles in the collection are available on the official website edition
http://journals.maup.com.ua/index.php/it

The articles were checked for plagiarism using the software
StrikePlagiarism.com developed by the Polish company Plagiat.pl.

ЗМІСТ

CONTENTS

УДК 621.396:004.89
DOI https://doi.org/10.32689/maup.it.2023.3.1

Роберт БЄЛЯКОВ
кандидат технічних наук, доцент, докторант науково-організаційного відділу Військового інституту телекомунікацій та інформатизації імені Героїв Крут, вул. Князів Острозьких 45/1, Київ, Україна, індекс 01011 (france417@gmail.com)

ORCID: 0000-0001-9882-3088

Олексій ФЕСЕНКО
Викладач кафедри технічного та метрологічного забезпечення факультету інформаційних технологій Військового інституту телекомунікацій та інформатизації імені Героїв Крут, вул. Князів Острозьких 45/1, Київ, Україна, індекс 01011 (rocimean123@gmail.com)

ORCID: 0000-0002-2114-5327

Robert BIELIAKOV
Candidate of Technical Sciences, Associate Professor, Doctoral Student Of The Scientific And Organizational Department of the Military Institute of Telecommunications and Information technologies named after Heroes of Kruty, Knyaziv Ostrozkih str. 45/1, Kyiv, Ukraine, postal code 01011 (france417@gmail.com)

Oleksii FESENKO
Lecturer at the Department of Technical and Metrological Support of the Faculty of Information Technologies of the Military Institute of Telecommunications and Information technologies named after Heroes of Kruty, Knyaziv Ostrozkih str. 45/1, Kyiv, Ukraine, postal code 01011 (rocimean123@gmail.com)

Бібліографічний опис статті: Бєляков, Р., Фесенко, О. (2023). Модель інтелектуального управління ресурсами наземної комунікаційної мережі класу MANET. *Інформаційні технології та суспільство*, 3, 6–14. DOI: https://doi.org/10.32689/maup.it.2023.3.1

Bibliographic description of the article: Bieliakov, R., Fesenko, O. (2023). Model intelektualnoho upravlinnia resursami nazemnoi komunikatsiinoi merezhi klasu MANET [A model of intelligent resource management of class MANET terrestrial communication network]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 6–14. DOI: https://doi.org/10.32689/maup.it.2023.3.1

МОДЕЛЬ ІНТЕЛЕКТУАЛЬНОГО УПРАВЛІННЯ РЕСУРСАМИ НАЗЕМНОЇ КОМУНІКАЦІЙНОЇ МЕРЕЖІ КЛАСУ MANET

Анотація. *Стаття присвячена розробці моделі інтелектуального управління ресурсами наземної комунікаційної мережі класу MANET. Управління мережами MANET є складним завданням через їхню динамічну природу, високу мобільність вузлів, обмежені ресурси: енергію батареї, технічні характеристики комунікаційних засобів, протоколи різних рівнів моделі OSI, та потребу реалізації функцій управління на вузловому і мережевому рівні в умовах відсутності централізованого контролю. Наукова новизна розробленої математичної моделі полягає у впровадженні алгоритмів машинного навчання з підкріпленням для управління процесом формування управляючих рішень на мережевому рівні, та сукупності нейромереж для забезпечення вимог щодо якості інформаційного обміну на вузловому рівні (реалізація користувальницьких цілей). Процес навчання нейромереж включає використання нової моделі мобільності, що враховує фізичні параметри вузлів мережі, а в сукупності із метриками радіозв'язності та метриками маршрутизації відповідно до вибраного протоколу забезпечується адаптація до змін у мережі в реальному часі. Крім того, використання роботизованих платформ – мобільних базових станцій може підвищити гнучкість та адаптивність мережі.*

Використання дворівневої інтелектуальної системи управління дозволяє розділити процес на два етапи. На першому – пропонується застосовувати ейромережі для оптимізації окремих цільових функцій на вузловому рівні. На другому - використання алгоритму Q-навчання з підкріпленням для адаптації до змін у середовищі з використанням прогнозу винагороди за максимальний виграш у реалізації цільових функцій управління, за рахунок представлення процесу управління ресурсом як Марківський процес.

Використання такого підходу може забезпечити ефективне управління мережею, адаптуючись до змін у середовищі та враховуючи різні цільові функції.

Ключові слова: *комунікаційна мережа, MANET, система інтелектуального управління, нейромережі, машинне навчання із підкріпленням, Марківський процес, прогнозування.*

A MODEL OF INTELLIGENT RESOURCE MANAGEMENT OF CLASS MANET TERRESTRIAL COMMUNICATION NETWORK

Abstract. *The article is devoted to the development of a model of intelligent resource management of the ground MANET class communication network. Management of MANET networks is a difficult task due to their dynamic nature, high mobility of nodes, limited resources: battery power, technical characteristics of communication devices, protocols of different levels of the OSI model, and the need to implement management functions at the node and network level in the absence of centralized control. The scientific novelty of the developed mathematical model consists in the introduction of machine learning algorithms with reinforcement for managing the process of forming control decisions at the network level, and a set of neural networks to ensure the requirements for the quality of information exchange at the node level (implementation of user goals). The learning process of neural networks includes the use of a new mobility model that takes into account the physical parameters of network nodes, and in combination with radio connectivity metrics and routing metrics according to the selected protocol, adaptation to changes in the network is ensured in real time. In addition, the use of robotic platforms - mobile base stations can increase the flexibility and adaptability of the network.*

The use of a two-level intelligent control system allows dividing the process into two stages. At the first stage, it is proposed to use Eurogrids to optimize individual target functions at the nodal level. On the second - the use of the Q-learning algorithm with reinforcement for adaptation to changes in the environment using the forecast of the reward for the maximum gain in the implementation of the target management functions, due to the representation of the resource management process as a Markov process.

Using this approach can provide effective network management by adapting to changes in the environment and considering different objective functions.

Key words: *communication network, MANET, intelligent control system, neural networks, machine learning with reinforcement, Markov process, prediction.*

Вступ. Наземна комунікаційна мережа MANET (Mobile Ad-Hoc Networks) – мобільна радіомережа (МР) самоорганізованої архітектури, що передбачає відсутність фіксованої мережевої інфраструктури (базових станцій) та централізованого управління. В контексті сучасних безпроводових комунікацій побудова таких мереж спеціального призначення (IoT – Internet of Things, рятувальники, військові та ін.), окрім очевидних переваг гнучкої і швидко налаштовуваної мережі, пов'язана із вирішенням ряду проблемних наукових та прикладних питань.

Так, управління мережами MANET є складним завданням через їхню динамічну природу, високу мобільність вузлів, обмежені ресурси (наприклад, енергію батареї, технічні характеристики комунікаційних засобів, протоколи різних рівнів моделі OSI, тощо) та відсутність централізованого контролю. Це вимагає розробки моделей, методів, методик, алгоритмів для підвищення продуктивності мобільних комунікаційних мереж спеціального призначення.

Для моделювання процесу функціонування мереж такого класу необхідно [1]: по-перше правильно вибрати модель мобільності серед існуючих або розробити нову що з максимальною реалістичністю описує процес переміщення мобільних вузлів, і відповідатиме умовам функціонування; по-друге визначити сукупність вузлових, канальних, мережевих ресурсів цієї мережі, визначити протоколи маршрутизації задані технічно для вузлових елементів мережі. Проте, результат аналізу публікацій в даній предметній області показав, що на сьогоднішній день залишається актуальним завдання розробки нового або удосконалення існуючого науково-методичного апарату для інтелектуалізації процесу управління мережами класу MANET, для скорочення обсягів службової інформації, зменшення часу розгортання мереж, забезпечення процесів оперативного управління мережею в реальному часі тощо.

Тому, **метою роботи** є розробка математичної моделі інтелектуального управління наземною комунікаційною мережею спеціального призначення для забезпечення інформаційного обміну заданої якості з урахуванням наявних ресурсів (вузлових, канальних, мережевих).

Наукова новизна. Вперше побудовано наближену математичну модель інтелектуального управління наземною комунікаційною мережею спеціального призначення, описано особливості реалізації функцій управління (класу переміщення і комунікаційних) роботизованих мобільних базових станцій.

Аналіз останніх досліджень і публікацій. Моделювання Ad-Hoc мереж є предметом досліджень низки науковців, зокрема автори наукових праць [1-11]. В статті [10] обговорюється ієрархічний принцип управління, в якому головний вузол визначає параметри для підлеглих вузлів. Це може допомогти зменшити обсяг службової інформації, але в динамічних середовищах цей принцип може не забезпечувати оптимальну продуктивність мережі. Разом з тим, автори не зосереджуються на описі способів досягнення енергоефективності або акцентуванні динаміці топології, хоча вони згадують енергоефективність як один з можливих параметрів для оптимізації, вони не обговорюють специфічні стратегії або методи інтелектуального управління. В статті [8] показано підхід планування маршруту в телекомунікаційній мережі на основі нейронної мережі Хопфілда. Автори описують енергетичну функцію, яка використовується для визначення стану вузлів мережі, визначають рекомендації щодо вибору

початкових ваг нейронів та швидкості навчання. Переваги цього підходу полягають у здатності нейронних мереж до навчання та адаптації, що може бути корисним для планування маршрутів в складних або динамічних мережах, однак, до недоліків можна віднести вимоги до обчислювальних ресурсів [2], дослідження ефективності цього підходу в реальних системах динамічної топології. Крім того, нейронна мережа Хопфілда використовується як централізована система управління, приймає вхідні дані від кожного вузла і видає команди для керування, що може вплинути на часові показники функціонування в реальному часі. Разом з тим, було проаналізовано ряд наукових публікацій, присвячених інтелектуалізації систем управління мережами класу MANET шляхом застосування алгоритмів машинного навчання [6-9, 10]. Так, техніки машинного навчання використовуються для управління параметрами вузлів, в тому числі і для маршрутизації даних. Встановлено, що на ряду із перевагами відносно традиційних протоколів маршрутизації, машинне навчання вимагає великих обсягів вихідних даних, що може викликати суттєві ресурсні витрати.

Таким чином, оптимізація процесів інформаційного обміну в мобільних комунікаційних мережах шляхом удосконалення інтелектуальних систем управління є перспективним напрямком наукових досліджень.

Основна частина.

Вихідні умови функціонування мережі. Розглядається наземна комунікаційна мобільна радіомережа класу MANET розмірністю до сотні мобільних вузлів, зони переміщення яких попередньо задані відповідно до визначених рангів [1]. Мобільні вузли поділяються на дві основні підгрупи: підгрупа мобільних користувачів (переносні радіозасоби, возимі радіозасоби) та підгрупа мобільних базових станцій (наземна комунікаційна роботизована платформа). Наземна комунікаційна мережа не має прив'язки до будь-якої телекомунікаційної інфраструктури загального користування. Кожен мобільний комунікаційний вузол (МКВ) оснащений системою позиціонування, комунікаційним обладнанням, антенними пристроями, системою живлення та відповідною системою управління. Кожен мобільний комунікаційний вузол може бути представлений як об'єкт із системою управління представленою сукупністю агентів навчання із специфічними для кожної підгрупи (МК, МБС) цільовими функціями відповідно до двох основних класи задач: перший – комунікаційні $U_{\text{ком}}$; другий – переміщенням вузлів МБС

$$U(t) = \left\{U_{\text{ком}}(t), U_{\text{пер}}(t)\right\}. \tag{1}$$

Управління комунікаційною складовою НМ в цілому включає наступні основні функції

$$U_{\text{ком}}(t) = \{U^{\text{топ}}(t), U^{мар}(t), U^{ен}(t), U^{нав}(t), U^{бо}(t), U^{QoS}(t)\},$$

де $U^{\text{топ}}$ *– управління топологією* – в залежності від цільової функції управління здійснюється побудова топології мережі для отримання покриття і для отримання зв'язності;

$U^{мар}$ *– управління маршрутизацією* – здійснюється побудова та підтримка маршрутів передачі інформації заданої якості при виконані вимог до їх функціонування (мінімізації службового трафіку, зменшення витрат енергії батарей, вибір протоколу маршрутизації тощо).

$U^{ен}$ *– управління енергоспоживанням* – метою управління є мінімізація витрат енергії вузлами мережі (максимізація "часу життя" мережі – часу роботи мережі до моменту відмови заданої кількості вузла через нульову ємність їх батарей, та/або прогноз "часу життя" КА з метою ротації за заздалегідь встановленим алгоритмом);

$U^{нав}$ *– управління навантаженням* – метою управління є зменшення або перерозподіл навантаження між окремими вузлами НПМ з урахуванням ресурсних обмежень;

$U^{бо}$ *– управління безпекою обслуговування при передачі інформації* – управління набором параметрів (методи шифрування, криптозахисту та інші);

$U^{рад}$ *– управління радіоресурсом* – управління вибором частот, сигнально-кодовою конструкцією, методами модуляції тощо;

U^{QoS} *– управління якістю обслуговування при передачі інформації* – управління набором параметрів (пропускна спроможність, затримка доставки пакетів, джиттер, BER – (Bit Error Rate) або PL (Packet Loss), для певного типу трафіку (мова, дані, відео).

В загальному вигляді задачі управління переміщенням МБС

$$U_{\text{пер}}(t) = \{U^{\text{МБС}}(t)\},$$

передбачає наступні функції

$$U_{\text{пер}}(t) = \{U^{\text{вп}}(t), U^{пок}(t), U^{рух}(t)\},$$

де $U^{\text{вп}}$ –визначення положення МБС на місцевості для досягнення певних цільових функцій управління;

$U^{\text{пок}}$ – визначення зони покриття мобільних користувачів (встановлення потужності передачі, управління діаграмою направленості антени, режимами роботи та ін.);

$U^{руx}$ – управління рухом (траєкторією та параметрами переміщення МБС).

Кількість і конкретні задачі оперативного управління визначаються характеристиками і умовами функціонування мережі, а також прийнятими технологічними рішеннями на етапі її проектування [1-3, 5].

Управління переміщенням мобільних користувачів не здійснюється, так як вони змінюють місце розташування разом із носіями відповідно до цільових завдань згідно штатно-посадових функціональних обов'язків. У процесі функціонування кожен комунікаційний вузол здійснює збір, обробку та формування управляючих впливів за рахунок інтелектуальної системи управління.

Мобільна базова станція за допомогою інтелектуальної системи управління формує оптимальні керуючі впливи із керування траєкторією переміщення з урахуванням перешкод, особливостей рельєфу та реалізує комунікаційні цільові функції відповідно до вимог інформаційного обміну мобільних користувачів (рис. 2).

Обмеження та вимоги:

– траєкторія переміщення мобільних користувачів формується в вигляді кортежу координат точок, що описується функцією переміщення вузлів певного рангу у визначеній зоні переміщення з урахуванням рельєфу місцевості та фізичних перешкод відповідно до моделі мобільності [1];

– інформація про параметри стану вузлів (координати розміщення, рівень енергії батарей, об'єм переданих пакетів, час затримки повідомлень) збирається при розгортанні мережі ТА, надалі інформація про стан вузлів оновлюється під час кожного сеансу інформаційного обміну;

– МБС та мобільні користувачі мають радіозасоби з однаковим МАС-протоколом, який дозволяє змінювати швидкість передачі даних в залежності від співвідношення сигнал/шум та регулювати потужність передачі (витрати енергії на передачу);

– обсяги оперативної пам'яті мобільних комунікаційних вузлів (МК та МБС), достатні для збору, обробки і формування управляючих рішень в режимі реального часу;

– значення енергії батареї МК та МБС достатнє на час моделювання (процеси дозаряджання засобів МК, ротації МБС не розглядались).

МК та МБС обладнані антенно-фідерними пристроями із круговою діаграмою направленості.

Таким чином, траєкторія переміщення МБС визначається за рахунок прогнозування координат її розміщення на основі зібраних даних переміщення мобільних користувачів при виконанні (реалізації) цільових функцій (ЦФ) управління відповідно до заданих вимог:

– мінімізація часу затримки передачі пакетів відповідно до вимог QoS передачі заданого типу трафіку

$$T_з^* \to \min T_з \left(T_з^* \le T_з^{QoS} \right); \tag{2}$$

– мінімізація джиттера передачі пакетів

$$J^* \to \min J \left(J^* \le J^{QoS} \right); \tag{3}$$

максимум пропускної здатності (каналу, маршруту, мережі)

$$S^* \to \max S \left(S^* \ge S^{QoS} \right); \tag{4}$$

– мінімізація коефіцієнту втрати переданих пакетів Packet Loss

$$PL^* \to \min PL \left(PL^* \le PL^{QoS} \right) = \begin{cases} \min BER_{ij}, \\ \max SINR_{ij}, \quad i, j = \overline{1, N_{МКВ}} \end{cases}; \tag{5}$$

– максимізація часу функціонування (вузла) мережі

$$T_{func} \to \max, \tag{6}$$

за рахунок оптимізації енергоефективності мобільних користувачів та мобільних базових станцій.

При обмеженнях Ω на:

– тип роботизованої платформи МБС; максимальну швидкість, характеристики прохідності (коефіцієнт маневреності), час автономної роботи – *v*=[*v*min, *v*max]; *Kman*; *tfunc*МБС ≤ *tfunc*МБСmax;

– кількість МБС – $N_i^{МБС} \le N_i^{МБСmax}$;

– початкову енергію батарей мобільних користувачів $ei \leq e$МКmax і МБС eМБС $\leq e$МБСmax;

– дальність радіозв'язності МК-МБС – di-МБС $\leq d$max;

– радіус площі зони радіопокриття МБС – Rmin $\leq R \leq R$max;

Необхідно: описати процес формування управляючих рішень для реалізації двох класів задач управління НКМ (задачі інформаційного обміну та переміщення).

Методологія дослідження. Було зроблено припущення що інтелектуальне управління цільовими функціями мобільних комунікаційних вузлів дозволить покращити продуктивність мережі, а визначена завчасно модель мобільності, дозволить прискорити процес розгортання і скоротити час навчання СУ мобільних вузлів.

В загальному випадку, управління мобільною комунікаційною мережею можна поділити на етапи: етап планування, розгортання та етап оперативного управління.

Етап планування включає, в основному, ряд організаційно-технічних заходів із визначення складу сил та засобів, вихідних даних умов функціонування тощо.

Етап розгортання в класичному розумінні включає збір інформації про середовище функціонування, наявні вузлові (комунікаційні) ресурси, аналіз цілей управління, розрахунок мережевих ресурсів для прогнозування можливості забезпечення задач із інформаційного обміну із заданим рівнем якості QoS, та обґрунтування організаційно-технічних рішень щодо корегування складу основних засобів для виконання завдань, попереднього налаштування і підготовки мережі тощо.

Етап оперативного управління включає процеси із підтримання реалізації вузлових та мережевих цільових функцій.

В [2] наведено класифікацію цілей управління мобільними радіомережами класу MANET, що поділяються на дві основні групи: перша – мережеві, друга – користувальницькі (рис. 1).

Рис. 1. Класифікація цілей управління МР

У [10] визначено, що критерії ЦФ можуть мати різну фізичну природу і частина з них потребує максимізації (пропускна спроможність та час функціонування МР), а частина – мінімізації (потужність передачі та час затримки передачі пакетів). Тобто, в один і той же момент часу ЦФ різних мобільних вузлів можуть суперечити одна-одній, що призводить до появи внутрішньосистемних конфліктів двох видів:

– *внутрішньорівневих* – виникають через несумісність користувальницьких цілей окремих вузлів у рамках заданої структури МР і обмежень;

– *міжрівневих* (між вузлом-координатором та підлеглими вузлами) – пов'язані з тим, що системна ЦФ, визначена вузлом-координатором, не збігається з ЦФ підлеглих вузлів.

Такі конфлікти викликають потребу постійної ситуаційної обізнаності про фактичний стан мобільних комунікаційних вузлів, що в умовах динамічного середовища, часто стохастичної природи, призводить до стрімкого збільшення службового трафіку мережі. Наприклад, у [10] задача зводилась до розробки методу координації цільових функцій, що фактично являє собою комбінацією ітеративних (на вузловому рівні) та безітеративних (на мережевому рівні) методів оптимізації.

В загальному випадку, таке рішення дійсно призводить до оптимізації використання ресурсів обох рівнів, але разом з тим має суттєвий недолік – обмеження на реалізацію вузлових цільових функцій вектором управляючих рішень вищестоящого вузлового елемента. Інакше кажучи, для деякої багаторівневої системи управління скоординоване управляюче рішення визначається як перетин множини дій та станів в заданих умовах (при дії обмежень на розмірність мережі, фізичні параметри вузлів та середовища тощо).

Концептуально наявність цього недоліку вказує на необхідність врахування етапу управління мережею, тобто «ітеративну частину» методу координації цільових функцій необхідно реалізувати на етапі проектування (розгортання) мережі.

Таким чином, пропонується модель системи управління наземної комунікаційної мережі вузловими і мережевими ресурсами побудувати за загальною структурою, зображеною на рисунку 2.

В процесі ітеративного обміну інформацією формується база знань шляхом збору статистичних даних, що відображають множину станів мобільних користувачів, набутих в результаті реакції на дії (множину відповідних керуючих дій для забезпечення користувальницьких цілей управління – рисунок 1) в середовищі функціонування. В свою чергу це середовище буде визначатися моделлю мобільності [1], технічними характеристиками мобільних користувачів, моделями шумових характеристик та завад різного роду, навантаження, методами маршрутизації тощо.

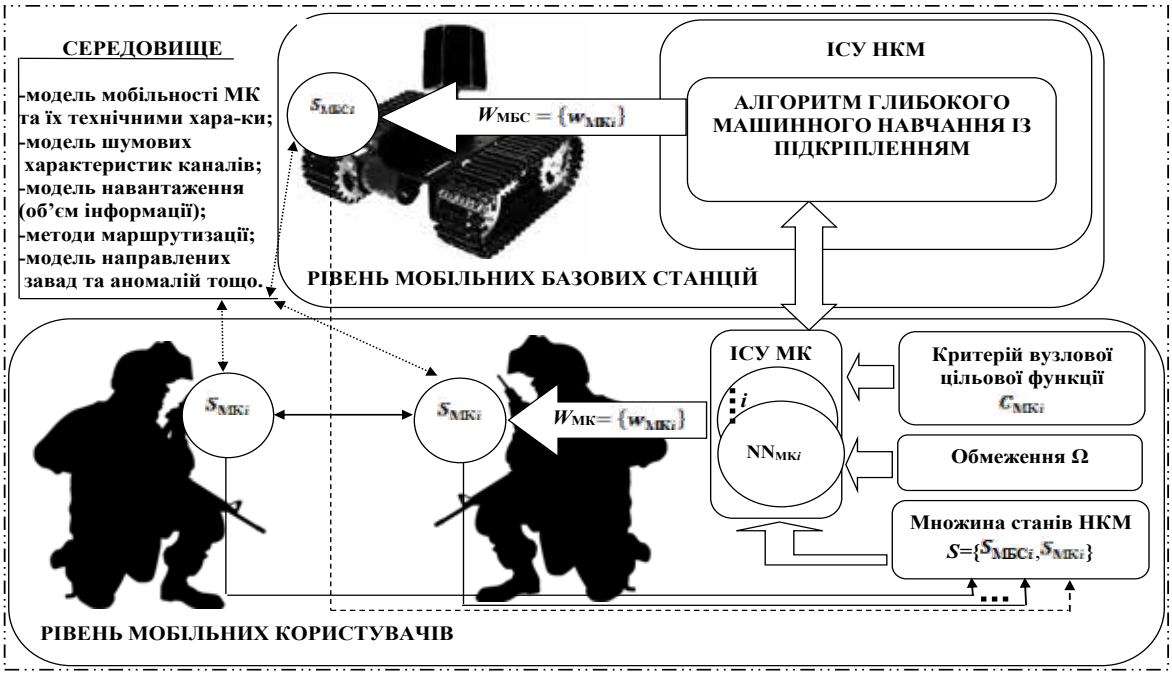


Рис. 2. Загальна структура моделі системи управління наземної комунікаційної мережі вузловими і мережевими ресурсами

Тобто, інтелектуальна система управління (ІСУ) мобільних користувачів, умовно кажучи, системно об'єднує підсистему поповнення бази знань про процеси реалізації користувальницьких цільових функцій, що є основою для вибору оптимального управляючого впливу $W(U_{t+1})$ шляхом прогнозування стану об'єкту управління (МК) із підсистемою реалізації рішень на вузловому або мережевому рівні. Кожна із нейронних мереж *NNi*МК ІСУ МК реалізує виконання однієї або декількох взаємозалежних користувальницьких цільових функцій (таблиця 1), крім того, до таблиці винесено основні напрямки (способи) реалізації математичних співвідношень, вимоги QoS для передачі відповідно до типу трафіку.

Для вирішення поставленої задачі пропонується представити процес формування управляючих рішень для реалізації цільових функцій як Марківський процес (Markov decision process – MDP).

MDP визначається чотирма складовими: множиною станів *S*, множиною дій *A*, функцією переходу *P*, та функцією винагороди *R*:

1. *Множина станів S.* Кожен стан *Si* ∈ *S* може бути вектором, що складається з координат МБС та МК, пропускної здатності наземних комунікаційних вузлів, залишкової енергії кожного вузла в мережі, та інших метрик відповідно до протоколу маршрутизації.

2. *Множина дій A.* Фактично множиною дій є аргументи цільових функцій управління, тобто $U \in A$. І може включати можливі напрямки переміщення МБС, з метою забезпечення радіопокриття, пропускної здатності маршрутів, часу затримки доставки пакетів тощо.

3. *Функція переходу P.* Функція переходу P(*s*'|*s*, *U*) визначає ймовірність переходу в стан *s*' після виконання дії *U* в стані *s*.

4. *Функція винагороди R.* Функція винагороди *R(s, U, s')* визначає винагороду, отриману після виконання дії *U* в стані *s* і переходу в стан *s*'. В контексті управління мережею це збільшення загальної пропускної здатності мережі або зменшення загальної затримки маршрутів, збільшення часу «життя» вузлів тощо.

Таблиця 1

Перелік математичних моделей (функцій) для реалізації ЦФ

Критерій вузлової (МК) ЦФ ($C_{МКi}$)	**Математичні моделі (функції та рівняння)**	**Вимоги QoS передачі заданого типу трафіку**		
		Відео передача	**Голосова передача (VoIP)**	**Передача даних**
Мінімізація часу затримки передачі пакетів (1)	Рівняння затримки передачі Рівняння затримки обробки Рівняння затримки маршрутизації Загальне рівняння затримки	до 150 ms	до 150 ms	може варіюватися: для більшості даних до 250 ms
Мінімізація джиттера передачі пакетів (2)	Рівняння джиттера Загальне рівняння затримки із урахуванням джиттера	до 30 ms	до 30 ms	для більшості даних до 50 ms
Максимум пропускної здатності (каналу, маршруту, мережі) (3)	Рівняння пропускної здатності (формула Шеннона) Модель каналу зв'язку Модель шуму Модель направлених завад	Висока: від 500 Kbps до 5 Mbps в залежності від якості (роздільна здатність, частота кадрів та ін.)	Нижча: приблизно 64 Kbps до 128 Kbps	Залежить від типу даних: може бути від низької до високої
Мінімізація коефіцієнту втрати переданих пакетів Packet Loss (4)	Рівняння втрати пакетів через помилки передачі Рівняння втрати пакетів через перевантаження Рівняння втрати пакетів через маршрутизацію Загальна втрата пакетів	до 1%	до 0.5%	Залежить від типу даних: для більшості даних до 1%
Максимізація часу функціонування вузла (мережі) (5)	Модель енергоспоживання комунікаційного вузла (мережі)	-	-	-

Метою агенту є знаходження політики π, яка максимізує суму винагород, отриманих протягом часу навчання. Політика π визначає ймовірність вибору кожної дії *U* в кожному стані *s*. В контексті глибокого навчання з підкріпленням, політика π часто апроксимується нейронною мережею, параметри якої оновлюються під час процесу навчання.

Математично, метою агента є максимізація очікуваної суми винагород:

$$Q^{t+1} \leftarrow Q^t\left(s_t, a_t\right) + \alpha \cdot \left(r_t + \gamma \cdot \max Q^{t+1}\left(s_{t+1}, a_t\right) - Q^t\left(s_t, a_t\right)\right) \quad (7)$$

де γ є фактором дисконтування, що визначає, наскільки агент «цінує» майбутні винагороди порівняно з нинішніми.

Алгоритми глибокого навчання з підкріпленням, такі як Deep Q-Network (DQN), використовують політики на основі взаємодії агента з середовищем для оновлення параметрів (рис. 3).

Так, в контексті Q-навчання (7) агенту, з метою реалізації цільових функцій (1) для виконання вимог із якості інформаційного обміну (2-6), маємо:

$$Q\left(s_t, U\right) = R(s_t, U) + \gamma \cdot \sum_{S_{t+1}} P(s_{t+1} \mid s_t, U) \max_{U_{t+1}} Q\left(s_{t+1}, U_{t+1}\right), \quad (8)$$

де $R(s_t,U)$ – це винагорода, отримана після виконання дії U в стані s_t, $P(s_{t+1}|s_t,U)$ – це імовірність переходу в стан s_{t+1} після виконання дії U в стані s_t, $\max_{U_{t+1}} Q(s_{t+1},U_{t+1})$ – це максимальна Q-цінність для наступного стану s_{t+1}, що відображає найкращу дію U_{t+1}, яку можна виконати в цьому стані.

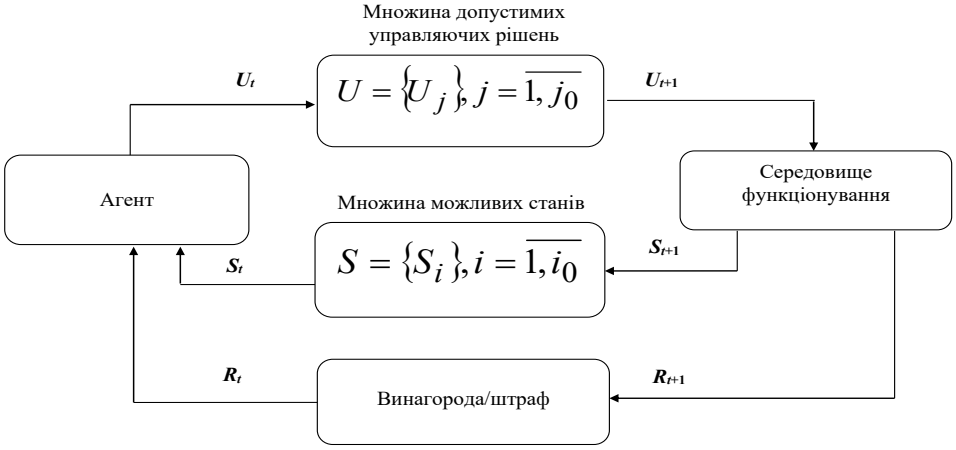


Рис. 3. Загальний процес алгоритму Q-навчання з підкріпленням

Висновки. Використання дворівневої інтелектуальної системи управління дозволяє розділити процес на два етапи: планування та виконання, що може допомогти в управлінні складністю задачі.

На першому рівні, ви використовуєте нейромережі для оптимізації окремих цільових функцій. Це дозволить використовувати нейромережі для моделювання складних функцій і знаходження оптимальних рішень на вузловому рівні.

На другому рівні, використання алгоритму Q-навчання з підкріпленням для адаптації до змін у середовищі в реальному часі дозволяє використовувати інформацію, отриману на першому рівні, для прийняття оптимальних рішень в онлайн режимі.

Використання такого підходу може забезпечити ефективне управління мережею, адаптуючись до змін у середовищі та враховуючи різні цільові функції. Однак, варто зазначити, що реалізація такої системи може бути досить складною, і може накладати часові обмеження на налаштування та тестування різних компонентів системи, щоб досягти оптимальних результатів.

Тому, напрямком подальших досліджень є імітаційне моделювання запропонованої моделі інтелектуального управління ресурсами наземної комунікаційної мережі класу MANET для дослідження її адекватності, крім того необхідно дослідити показники ефективності системи інтелектуального управління у порівнянні із відомими рішеннями.

Список використаних джерел:

1. Бєляков Р. О., Фесенко О. Д. Модель мобільності наземної комунікаційної мережі спеціального призначення. *Computer-Integrated Technologies: Education, Science, Production*. 2023. № 51. С. 130–138. DOI: 10.36910/6775-2524-0560-2023-51-17.

2. Романюк В. А. Концепція ієрархічної побудови інтелектуальних систем управління тактичними радіомережами класу MANET. *НВЧ-техніка та телекомунікаційні технології* : Збірник тез доповідей Міжнар. кримська конф. КриМіКо. Севастополь, 2012. С. 265.

3. Романюк В. А. Цільові функції оперативного управління тактичними радіомережами. *Збірник наукових праць ВІТІ НТУУ «КПІ»*. 2012. № 1. С. 109–117.

4. Романюк В. А. Архітектура системи оперативного управління тактичними радіомережами. *Збірник наукових праць ВІТІ НТУУ «КПІ»*. 2009. № 3. С. 70–76.

5. Deep Reinforcement Learning Aided Packet-Routing for Aeronautical Ad-Hoc Networks Formed by Passenger Planes / D. Liu et al. *IEEE Transactions on Vehicular Technology.* 2021. Vol. 70, no. 5. P. 5166–5171. DOI: 10.1109/tvt.2021.3074015.

6. Enhancing Vehicular Ad Hoc Networks' Dynamic Behavior by Integrating Game Theory and Machine Learning Techniques for Reliable and Stable Routing / N. Phull et al. *Security and Communication Networks*. 2022. Vol. 2022. P. 1–11. DOI: 10.1155/2022/4108231.

7. Implementation of Mobility Management Methods for MANET / J. Hosek et al. *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems*. 2012. Vol. 1, no. 2-3. DOI: 10.11601/ijates.v1i2-3.39.

8. Model of data flow control subsystem of the MANET class mobile radio network control system / Yu. Kramska et al. *Scientific Journal of TNTU*. 2022. Vol. 107. No 3. P. 51–59.

9. Romaniuk V. A., Bieliakov R. O. Objective control functions of FANET communication nodes of land-air network. *Computer-Integrated Technologies: Education, Science, Production.* 2023. No. 50. P. 125–130. DOI: 10.36910/6775-2524-0560-2023-50-19.

10. The Hierarchical Model Of Interaction Between Intelligent Agents In The Manet Control Systems / O. Y. Sova et al. *Information and Telecommunication Sciences*. 2015. No. 1. P. 21–28. DOI: 10.20535/2411-2976.12015.21-28.

11. Жебка В.В. Дослідження методів машинного навчання та їх застосування для прогнозування відтоку користувачів телекомунікаційних послуг. *Зв'язок.* 2020. № 4. DOI: 10.31673/2412-9070.2020.042231.

References:

1. Bieliakov, R., & Fesenko, O. (2023). Model mobilnosti nazemnoi komunikatsiinoi merezhi spetsialnoho pryznachennia [Mobility model of a special purpose terrestrial communication network]. *Computer-Integrated Technologies: Education, Science, Production*, 51, 130–138.

2. Romaniuk, V. A. (2012). Kontseptsyia yerarkhycheskoho postroenyia yntellektualnikh system upravlenyia taktycheskymy radyosetiamy klassa MANET [The concept of hierarchical construction of intelligent management systems for tactical radio networks of the MANET class]. Proceedings from KryMyKo: *Mezhdunarodna Krymska konferentsiia «NVCh-tekhnyka ta telekommunykatsiyni tekhnolohii» – International Crimean conference «Microwave technology and telecommunication technologies»*. (p. 265). Sevastopol: KryMyKo [in Ukrainian].

3. Romaniuk, V. A. (2012). Tsilovi funktsii operatyvnoho upravlinnia taktychnymy radiomerezhamy [Target functions of operational management of tactical radio networks]. *Zbirnyk naukovykh prats VITI NTUU «KPI» – Scientific works collection of MITIT NTUU "KPI"*, 1, 109–117. [in Ukrainian].

4. Romaniuk, V. A. (2009). Arkhitektura systemy operatyvnoho upravlinnia taktychnymy radiomerezhamy [Architecture of the operational management system of tactical radio networks]. *Zbirnyk naukovykh prats VITI NTUU «KPI» – Scientific works collection of MITIT NTUU "KPI"*, 3, 70–76. [in Ukrainian].

5. Liu, D., Cui, J., Zhang, J., Yang, C. Y., & Hanzo, L. (2021). Deep Reinforcement Learning Aided Packet-Routing for Aeronautical Ad-Hoc Networks Formed by Passenger Planes. *IEEE Transactions on Vehicular Technology*, *70*(5), 5166–5171. https://doi.org/10.1109/tvt.2021.3074015.

6. Phull, N., Singh, P., Shabaz, M., & Sammy, F. (2022). Enhancing Vehicular Ad Hoc Networks' Dynamic Behavior by Integrating Game Theory and Machine Learning Techniques for Reliable and Stable Routing. *Security and Communication Networks*, *2022*, 1–11. https://doi.org/10.1155/2022/4108231.

7. Hosek, J., Vajsar, P., Bartl, M., & Molnar, K. (2012). Implementation of Mobility Management Methods for MANET. *International Journal of Advances in Telecommunications, Electrotechnics, Signals and Systems*, *1*(2-3). https://doi.org/10.11601/ijates.v1i2-3.39.

8. Kramska, Y., Salnyk, S., Vasylenko, S., & Mavrina, O. (2022). Model of data flow control subsystem of the manet class mobile radio network control system. *Visnyk Ternopilskoho natsionalnoho tekhnichnoho universytetu – Scientific journal of the Ternopil national technical university*, *107*(3), 51–59. https://doi.org/10.33108/visnyk_tntu2022.03.051.

9. Romaniuk, V. A., & Bieliakov, R. O. (2023). Objective control functions of FANET communication nodes of land-air network. *Computer-Integrated Technologies: Education, Science, Production*, (50), 125–130. https://doi.org/10.36910/6775-2524-0560-2023-50-19.

10. Sova, O. Y., Romanyuk, V. A., Minochkin, D. A., & Polshchykov, K. O. (2015b). The Hierarchical Model Of Interaction Between Intelligent Agents In The Manet Control Systems. *Information and Telecommunication Sciences*, (1), 21–28. https://doi.org/10.20535/2411-2976.12015.21-28.

11. Zhebka, V. V. (2020). Doslidzhennia metodiv mashynnoho navchannia ta yikh zastosuvannia dlia prohnozuvannia vidtoku korystuvachiv telekomunikatsiinykh posluh [Research of machine learning methods and their application for forecasting use outflow by telecommunications services]. *Zviazok – Connectivity*, 146(4). https://doi.org/10.31673/2412-9070.2020.042231. [in Ukrainian].

УДК 004.056.3
DOI https://doi.org/10.32689/maup.it.2023.3.2

Микола ВАСИЛЕНКО
доктор фізико-математичних наук, доктор юридичних наук, професор, професор кафедри кібербезпеки, Національний університет «Одеська юридична академія», вул. Рішельєвська, 28, Одеса, Україна, індекс 65011 (vasylenko.it@journals.maup.kiev.ua)

ORCID: 0000-0002-8555-5712

Валерія СЛАТВІНСЬКА
викладач кафедри кібербезпеки, Національний університет «Одеська юридична академія», вул. Рішельєвська, 28, Одеса, Україна, індекс 65011, Україна, індекс 65000 (slatvinskaya_valeriya@ukr.net)

ORCID: 0000-0002-6082-981X

Світлана СИСОЄНКО
кандидат технічних наук, доцент, доцент кафедри інформаційної безпеки та комп'ютерної інженерії, Черкаський державний технологічний університет, бульв. Шевченка, 460, Черкаси, Україна, індекс 18006 (s.sysoienko@chdtu.edu.ua)

ORCID: 0000-0002-0009-337X

Nikolai VASILENKO
Doctor of Physical and Mathematical Sciences, Doctor of Law, Professor, Professor at the Department of cybersecurity, National University "Odessa Law Academy", 28 Richelevskaya str., Odessa, Ukraine, postal code 65011 (vasylenko.it@journals.maup.kiev.ua)

Valeriia SLATVINSKA
Assistant Professor at the Department of Cybersecurity, National University "Odessa Law Academy", 28 Richelevskaya str., Odessa, Ukraine, postal code 65011 (slatvinskaya_valeriya@ukr.net)

Svitlana SYSOIENKO
Candidate of Technical Sciences, Associate Professor, Associate Professor at the Department of Computer Information Systems and Technologies, Cherkasy State Technological University Shevchenko blvd., 460, Cherkasy, Ukraine, postal code 18006, (s.sysoienko@chdtu.edu.ua)

Бібліографічний опис статті: Василенко, М., Слатвінська, В., Сисоєнко, С. (2023). Особливості створення сценарію для оболонки bash при створенні резервних копій. *Інформаційні технології та суспільство*, 3, 15–22. DOI: https://doi.org/10.32689/maup.it.2023.3.2

Bibliographic description of the article: Vasilenko M., Slatvinska V. & Sysoienko S. (2023). Osoblyvosti stvorennia stsenariiu dlia obolonky bash pry stvorenni rezervnykh kopii [Features of creating a script for the bash shell when creating backups]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 15–22. DOI: https://doi.org/10.32689/maup.it.2023.3.2

ОСОБЛИВОСТІ СТВОРЕННЯ СЦЕНАРІЮ ДЛЯ ОБОЛОНКИ BASH ПРИ СТВОРЕННІ РЕЗЕРВНИХ КОПІЙ

Анотація. *У даній статті пропонується сценарій на мові bash для створення резервних копій файлів і каталогів. Взаємодія користувача з операційною системою здійснюється через оболонку. В Linux існує ряд різних оболонок, найпопулярнішої з яких є саме bash. Створення резервної копії даних надає можливість виконати відновлення інформації при втраті оригіналу, з якого було створено резервну копію. При цьому під втратою треба розуміти настання події, що призвела до зміни даних, після чого вони втратили цінність або були видалені з носія. Сценарій має наступні функції: Сценарій приймає параметри командного рядка, що є іменами каталогів, які треба додати до резервної копії. Командний рядок може стати ідеальним інструментом для забезпечення кібербезпеки. Неймовірна гнучкість і абсолютна доступність перетворюють стандартний інтерфейс командного рядка на фундаментальне рішення, яке надає швидко створювати та моделювати складні функції за допомогою лише одного рядка конвеєрних команд. Якщо параметри не задано, створюється резервна копія домашнього каталогу. Перевіряється наявність каталогу "archives" у домашньому каталозі користувача. Якщо його немає, він створюється. Отримується поточна дата і час для використання у назві архіву. Якщо параметри командного рядка задано, додавання вказаних каталогів до резервної копії. Перевіряється наявність каталогу перед додаванням. Навички та уміння ефективно працювати з командним рядком – найважливіша навичка для фахівців з безпеки та адміністрування. Створюється резервна копія кожного каталогу, включаючи його підкаталоги, за допомогою команди tar. Виводиться повідомлення про створену резервну копію разом зі шляхом, за яким вона зберігається. Стаття також надає інструкції щодо перевірки*

працездатності скрипта та налаштування періодичного запуску за допомогою cron і ефективності використання командного рядка для поліпшення наявного функціоналу.

Ключові слова: *скрипт, копіювання, командний рядок, резервна копія.*

FEATURES OF CREATING A SCRIPT FOR THE BASH SHELL WHEN CREATING BACKUPS

Abstract. *This article provides a bash script for creating backup copies of files and directories. The user interacts with the operating system through a shell. In Linux, there are a number of different shells, the most popular of which is bash itself. Creating a backup copy of data provides an opportunity to restore information in case of loss of the original from which the backup copy was created. At the same time, loss should be understood as the occurrence of an event that led to a change in data, after which they lost their value or were deleted from the medium. The script has the following functions: the script accepts command line parameters, which are the names of the directories to be added to the backup. The command line can be an ideal tool for ensuring cybersecurity. Unbelievable flexibility and absolute accessibility transform the standard command-line interface into a fundamental solution that allows you to quickly create and model complex functions using only one line of pipelined commands. If no parameters are specified, a backup copy of the home directory is created. If no parameters are specified, a backup copy of the home directory is created. If it does not exist, it is created. The current date and time are obtained to be used in the archive name. If command line parameters are specified, the specified directories are added to the backup. The existence of the directory is checked before adding it. The ability to work effectively with the command line is an essential skill for security and administration professionals. Each directory, including its subdirectories, is backed up using the tar command. A message about the created backup copy is displayed along with the path in which it is saved. The article also provides instructions on how to test the functionality of the script and how to set up a periodic run using cron and how to use the command line effectively to improve existing functionality.*

Key words: *script, copy, command line, backup copy.*

Актуальність проблеми. Розвиток інформатизації у суспільстві останніх років висунув на одне з перших місць проблему захисту величезної кількості інформації, що формується, оброблюється і передається в комп'ютерних системах, а також створення копії даних з носіїв за допомогою резервного копіювання або (бекап) та призначений для відновлення цих даних у разі їх пошкодження або видалення щодо різних дестабілізуючих факторів.

Основні причини створення резервних копій включають:

Захист від втрати даних: Резервні копії дозволяють відновити важливі файли та дані у випадку їх втрати, помилкового видалення або пошкодження. Це особливо важливо для бізнесу та користувачів, які зберігають цінні дані.

Відновлення системи: Резервні копії допомагають відновити систему у разі непередбачених ситуацій, таких як аварії жорсткого диска, атаки зловмисників або вірусів.

Міграція та переміщення даних: Резервні копії дозволяють легко переміщувати дані між різними середовищами або системами. Вони також корисні для міграції на нову апаратну або програмну інфраструктуру.

Відновлення попередніх версій: Резервні копії дозволяють повернутися до попередніх версій файлів або каталогів, які були змінені або видалені. Це може бути корисним у випадку помилкових змін або потреби відновити стару версію.

Забезпечення безпеки даних: Резервні копії є важливим елементом стратегії захисту даних. Вони дозволяють уникнути втрати важливих даних через випадкові або навмисні дії.

Одним із способів полегшити цю проблему з технічної точки зору є корисним інструментом вивчення мови bash. Без цього системні адміністратори та фанати терміналів опинилися б на терміналі годинами. Замість цього ми можемо писати сценарії bash для автоматизації Linux. Навчившись використовувати можливості Bash, ми можемо писати всі ці складні операції та швидко запускати їх за допомогою сценарію.

Оболонка чудова, досліджуючи її, можемо запропонувати шляхи автоматизації та покращення внутрішньої роботи системи Linux.

Аналіз останніх досліджень і публікацій. Для стабільної роботи компаній які працюють з інформацією є процес резервного копіювання даних. Не залежно від того, чи це банк, чи сайт по продажам господарських товарів, у разі несистемного збою та пошкодження основних серверів або баз даних –в найкоротший термін, системна помилка повинна бути виправлена а резервні дані відновлені. *Враховуючи активний розвиток шкідливого ПЗ, що випереджає розвиток антивірусів, найбільш раціонально будувати IT-безпеку навколо системи резервного збереження інформації. Замість того, щоб фокусуватися виключно на запобіганні атакам та боротьбі з вірусами, набагато простіше, дешевше і легше підняти систему з резервних копій* [1]. Авторами роботи [2] спроектовано інтелектуальну систему дедублікації та розподілу даних у хмарних сховищах, яка при доробці систем перевір-

ки цілісності резервних копій та оптимізації дискового простору після видалення старих резервних копій може використовуватись на практиці. У сфері IT оболонка – це програмне забезпечення, яке з'єднує користувача з комп'ютером. Оболонка дозволяє користувачеві дізнатися про систему або файли в ній або керувати системою. Оболонка зазвичай є частиною операційної системи. Існує два типи оболонок.

Командно-орієнтована оболонка, наприклад, OS X термінал.

Оболонка з графічним інтерфейсом користувача, наприклад шукач для OS X або всім знайома Windows.

Термінал є командно-орієнтованою оболонкою, тому що тут ми вводимо команди безпосередньо, а не натискаємо кнопки миші або вводимо інформацію в поля. Коли ми знаходимося на терміналі, у нас працює процес Bash, який надає нам оболонку Bash. Якщо ми починаємо виконувати сценарій, він фактично не виконується в цьому процесі, а замість цього запускає новий процес для виконання всередині. Командно-орієнтовані оболонки також мають безліч різновидів – оболонка Bourne, Bash, оболонка Z тощо. Як стандарт на комп'ютерах Mac, термінал відкриває оболонку Bash. Bash є дуже потужним, оскільки він може спростити певні операції, які важко ефективно виконати за допомогою графічного інтерфейсу користувача. Користуючись підтримкою численних прихильників програмного забезпечення з відкритим кодом, Linux набула великої популярності, а разом з тим – і значного технологічного удосконалення [3]. В Linux існує ряд різних оболонок, найпопулярнішої з яких є bash. Оболонка bash стала ідеальним інструментом для операцій із забезпечення безпеки, оскільки вона має міжплатформні методи і сценарії. Поширеність bash також дає певну перевагу фахівцям, які тестують на стійкість до атак і вторгнень, оскільки в багатьох випадках їм не потрібно встановлювати в системі додаткову інфраструктуру підтримки або інтерпретатор [4 – 5].

У [6-7] автори діляться деякими корисними інструментами та ресурсами, які можуть допомогти у вивченні командного інтерпретатора. Написання сценаріїв для оболонки – це потужна та універсальна навичка, яка може допомогти автоматизувати завдання, маніпулювати даними та налаштовувати інтерфейс командного рядка (CLI). Незалежно від того, чи використовується Bash, Zsh або інша оболонка.

Робота [8] охоплює програмування в командному рядку з акцентом на Linux в командному рядку Bash; що забезпечує достовірну, реальну актуальність, а також надає гнучкі інструменти для негайного початку роботи та вирішення різноманітних реальних завдань. Робота [9] присвячена багатьом поширеним програмам, що використовуються у командному рядку, а також більш складним темам. З огляду на коло досліджень, знаня загального синтаксису bash та вміня писати на ньому – критичний скіл для розробника. Дані дослідження обов'язково станутьв нагоді як початківцям, так і досвідченим розробникам, які могли забути деякі особливості цієї надпотужної командної оболонки.

Метою статті є створення сценарію для оболонки Bash, який автоматизує процес створення резервних копій файлів і каталогів. Основна мета такого сценарію полягає в забезпеченні збереження важливих даних шляхом регулярного створення резервних копій.

Виклад основного матеріалу. Напишемо сценарій для оболонки bash для створення резервних копій згідно таких вимог:

Сценарій приймає довільну кількість параметрів командного рядка. Якщо жодного з параметрів не задано, повинна створюватись резервна копія (tar.gz або tar.bz2) домашнього каталогу користувача. Якщо параметри задані, вони розглядаються як імена каталогів, які треба додати до резервної копії (при цьому домашній каталог у цілому за умовчанням не додається).

Якщо імена каталогів задані параметрами командного рядка, слід додати перевірку їх наявності. Сценарій повинен повідомляти про відсутність певних каталогів, але продовжувати роботу з архівації інших каталогів із списку.

Архів повинен створюватись у підкаталозі archives домашнього каталогу. Сценарій повинен перевірити наявність цього каталогу, і якщо його немає, створити його. Ім’я створюваного архіву повинно містити поточну дату у форматі YYYYMMDDHHMM.

Сценарій повинен коректно відпрацьовувати помилки, такі як некоректні імена каталогів і помилки доступу (зокрема, відсутність права доступу до певних файлів). При цьому сценарій повинен видавати діагностику помилок [3, с. 50]. Перша частина сценарію– це шебанг (#!). Ця послідовність дозволяє скрипту повідомляти інтерпретатору, що він повинен використовувати для розуміння коду. Далі напишемо коментар. Це дозволить кожному, хто використовує сценарій, зрозуміти, для чого призначений код. Коментарі можна додати до скрипту, поставивши символ # [10].

Результати досліджень. Для вирішення завдання напишемо скрипт на мові bash:

```
#!/bin/bash

# Перевірка наявності каталогу "archives" у домашньому каталозі
archive_dir="$HOME/archives"
if [[ ! -d $archive_dir ]]; then
  mkdir $archive_dir
fi

# Отримання поточної дати та часу
current_date=$(date +"%Y%m%d%H%M")

# Якщо параметри командного рядка не задано, створення резервної копії домашнього каталогу
if [[ $# -eq 0 ]]; then
  backup_filename="$archive_dir/home_backup_$current_date.tar.gz"
  tar -zcf $backup_filename -C $HOME .
  echo "Створено резервну копію домашнього каталогу: $backup_filename"
  exit 0
fi

# Якщо параметри командного рядка задано, додавання вказаних каталогів до резервної копії
for dir in "$@"; do
  # Перевірка наявності каталогу
  if [[ ! -d $dir ]]; then
    echo "Помилка: Каталог '$dir' не існує. Пропускаю..."
    continue
  fi

  # Отримання імені каталогу
  dir_name=$(basename "$dir")

  # Створення резервної копії каталогу
  backup_filename="$archive_dir/$dir_name"_"$current_date.tar.gz"
  tar -zcf $backup_filename -C $dir .
  echo "Створено резервну копію каталогу '$dir': $backup_filename"
done
```

Цей скрипт отримує параметри командного рядка, які є іменами каталогів, які потрібно додати до резервної копії. Якщо параметри не задано, створюється резервна копія домашнього каталогу користувача. Сценарій перевіряє наявність каталогів, повідомляє про відсутність деяких каталогів, але продовжує створювати резервну копію інших каталогів зі списку. Резервні копії зберігаються у каталозі "archives" у домашньому каталозі користувача з іменем, що містить поточну дату та час.

Для налаштування періодичного запуску скрипта за допомогою cron, ви можете відкрити cron-редактор командою **crontab -e** і додати наступний рядок:

0 0 * * * /шлях/до/скрипта.sh

Цей рядок вказує, що скрипт має бути запущений щодня о 00:00. Ви можете змінити ці значення відповідно до своїх потреб. Замість **/шлях/до/скрипта.sh** вкажіть повний шлях до файлу скрипта. Після збереження змін у cron-редакторі, cron почне виконувати скрипт зазначеним розкладом.

Щоб перевірити, чи працює скрипт резервного копіювання, було виконано наступні кроки:

1. Збережіть скрипт у файл з розширенням**.sh** (наприклад, **backup_script.sh**) і надайте йому права на виконання, використовуючи команду **chmod +x backup_script.sh**.

2. Запустіть скрипт, вказавши параметри командного рядка (якщо необхідно) або без них. Наприклад:

./backup_script.sh # Без параметрів – створення резервної копії домашнього каталогу;

./backup_script.sh dir1 dir2 # 3 параметрами – додавання вказаних каталогів до резервної копії.

На рисунку 1 наведено фрагмент виконання цих кроків в Git Bash на Windows і в результаті маємо наступне:

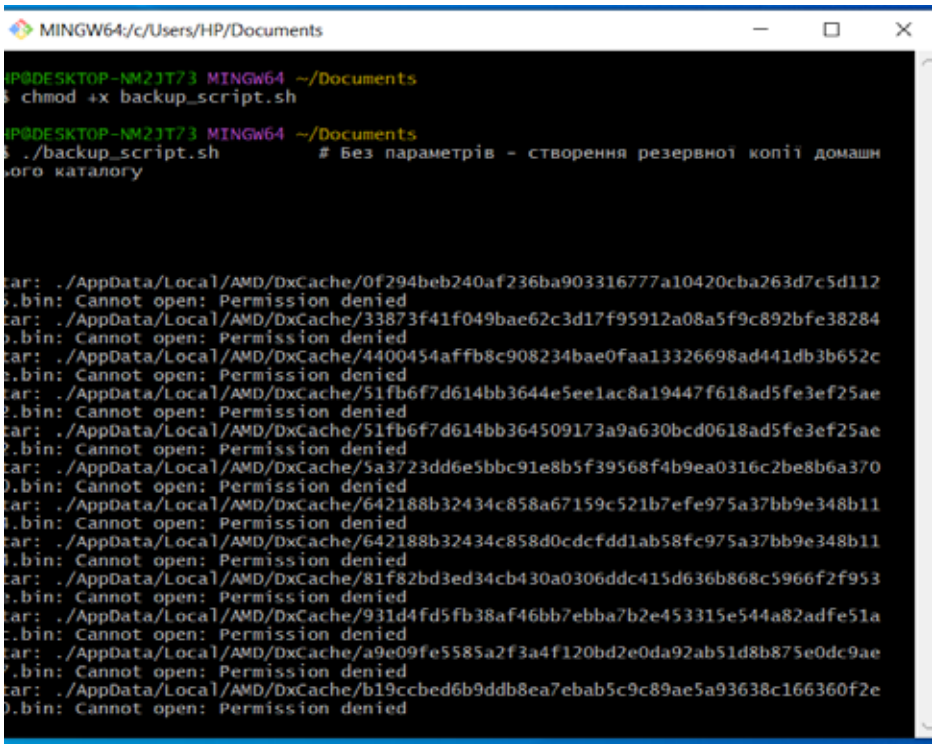


Рис. 1. Фрагмент виконання скрипту резервного копіювання

Після запуску скрипту з'явилося повідомлення про створення резервної копії та шлях, за яким зберігається архів (рисунок 2).

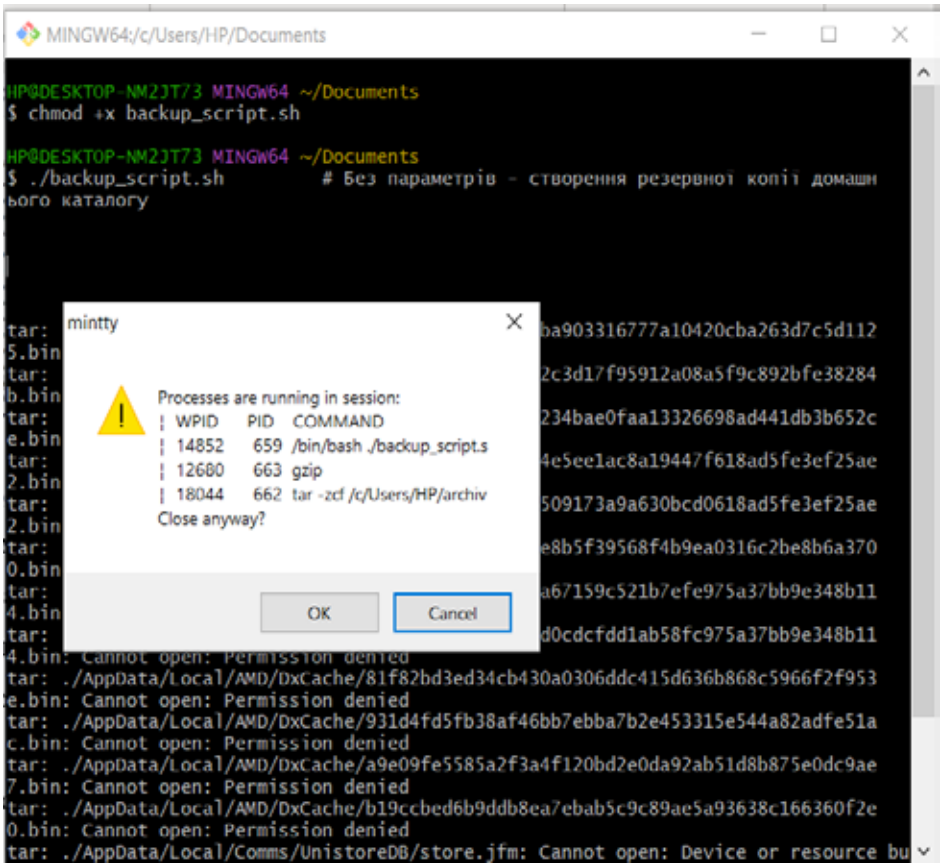


Рис. 2. Фрагмент створення та зберігання резервної копії

Потім було перевірено каталог "archives" у власному домашньому каталозі. Там з'явились нові архіви, які створив скрипт.

Архів містить необхідні файли та каталоги (рисунок 3), і їх можна відновити при необхідності.

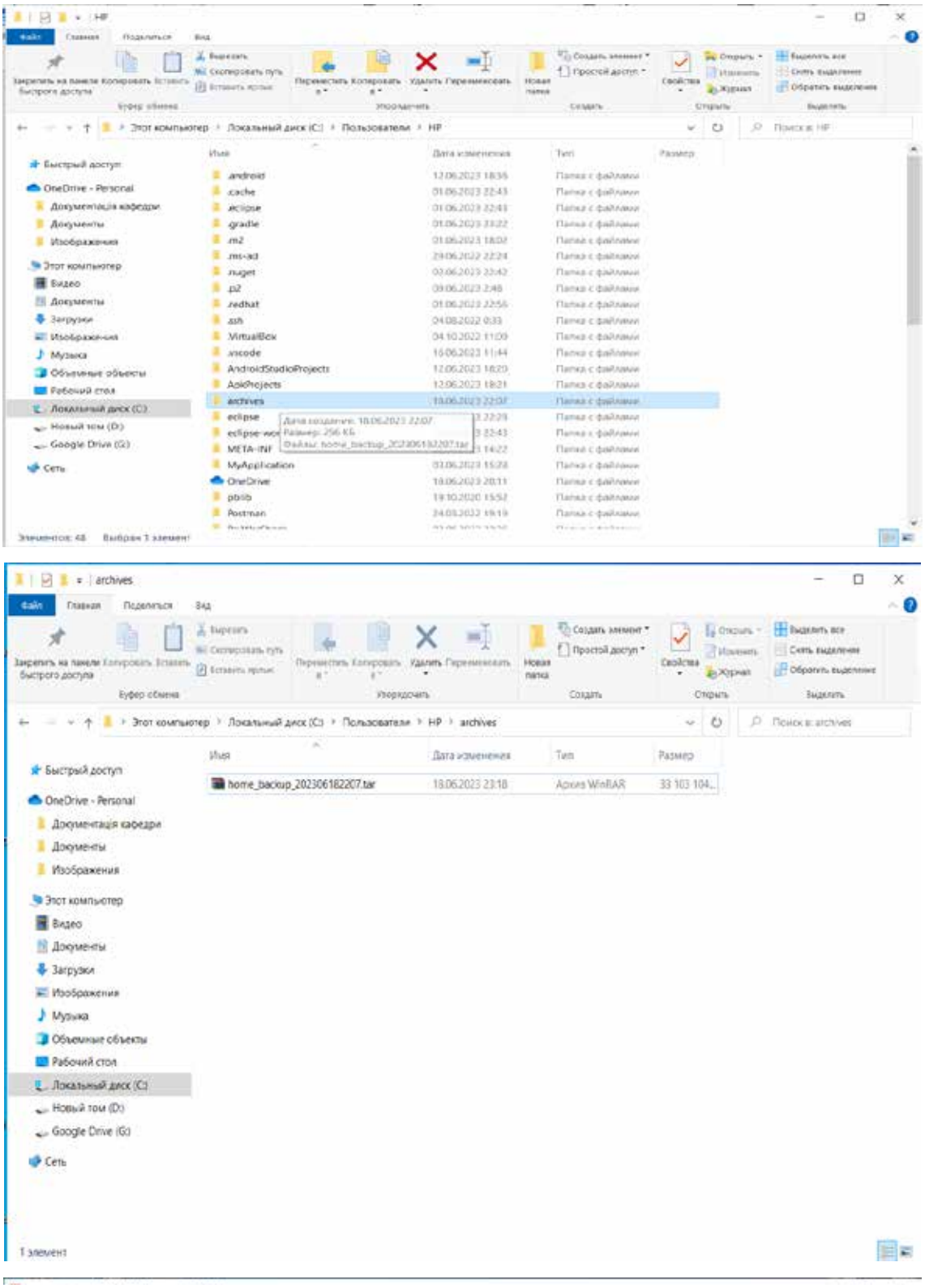


Рис. 3. Архів файлів та каталогів створених скриптом

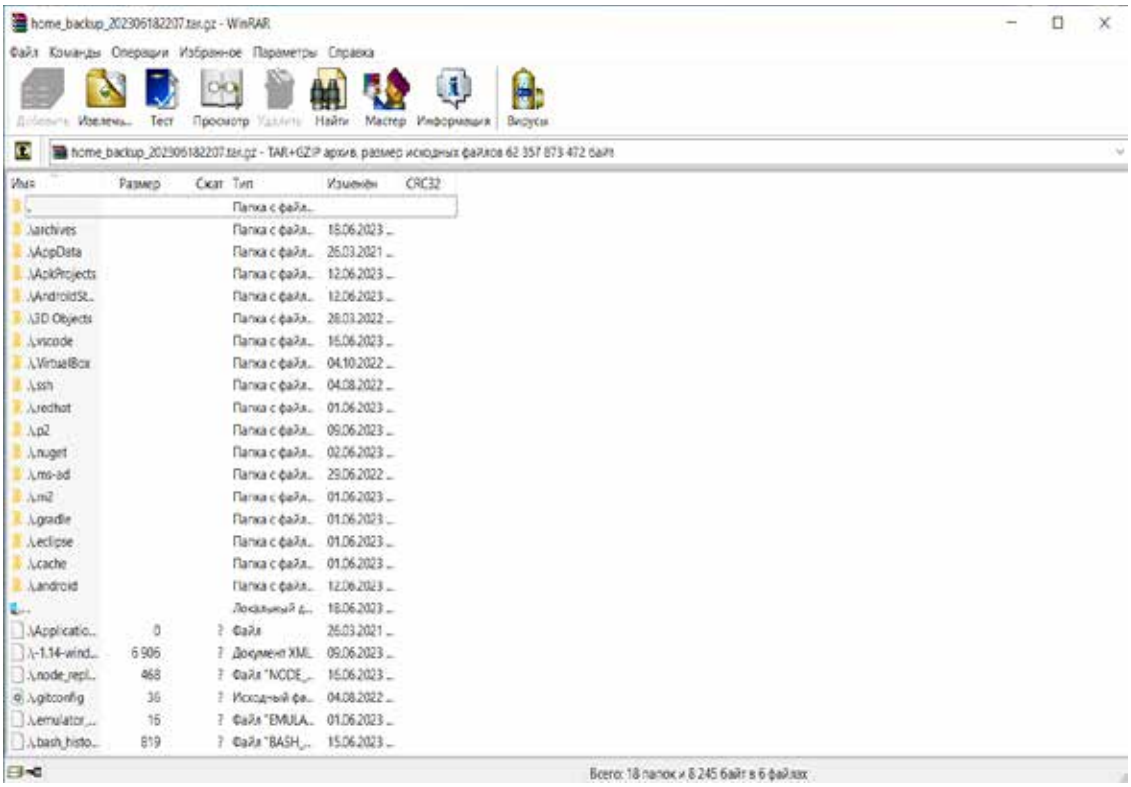


Рис. 3 (закінчення)

Обговорення результатів. Отже, сценарій для оболонки Bash для створення резервних копій має на меті автоматизувати процес створення резервних копій файлів та каталогів. Основна мета такого сценарію – забезпечити збереження важливих даних шляхом регулярного створення резервних копій.

Основні причини використання такого сценарію:

1. Захист від втрати даних: Резервні копії дозволяють відновити важливі файли та дані у випадку їх втрати, помилкового видалення або пошкодження. Це особливо важливо для бізнесу та користувачів, які зберігають цінні дані.

2. Відновлення системи: Резервні копії допомагають відновити систему у разі непередбачених ситуацій, таких як аварії жорсткого диска, атаки зловмисників або вірусів.

3. Міграція та переміщення даних: Резервні копії дозволяють легко переміщувати дані між різними середовищами або системами. Вони також корисні для міграції на нову апаратну або програмну інфраструктуру.

4. Відновлення попередніх версій: Резервні копії дозволяють повернутися до попередніх версій файлів або каталогів, які були змінені або видалені. Це може бути корисним у випадку помилкових змін або потреби відновити стару версію.

5. Забезпечення безпеки даних: Резервні копії є важливим елементом стратегії захисту даних. Вони дозволяють уникнути втрати важливих даних через випадкові або навмисні дії.

Висновки та перспективи подальших досліджень. Практична значущість отриманих результатів полягає у створенні сценарію для оболонки Bash, що допомагає автоматизувати процес створення резервних копій, забезпечуючи надійний та ефективний спосіб збереження даних. Командний рядок і пов'язані з ним можливості та інструменти створення сценаріїв є безцінним ресурсом для фахівця з кібербезпеки. Це можна порівняти з мультиінструментом із нескінченнозмінюваною конфігурацією. Поєднуючи продуману послідовність команд у конвеєр, можна створити однорядковий сценарій, що виконує надзвичайно складні функції. Для збільшення функціональності можна створювати багаторядкові сценарії.

Хоча у bash є свої переваги, такі як простий виклик інших програм або з'єднання послідовностей інших програм. У нього також є свої недоліки: bash не може виконувати операції над числами з плаваючою крапкою і не підтримує (навіть частково) складні структури даних. Але на сьогоднішній день величезна кількість програмістів працюють над постійним удосконаленням Linux: пишуть різні програми, що працюють в ній, розробляють різновиди і нові версії цієї ОС. Мета цієї статті полягає в тому,

щоб розглянути основні особливості мови BASH в ракурсі її використання для написання сценаріїв в операційній системі Linux. Linux на сьогоднішній день є єдиною альтернативою операційній системі Windows від розробників Microsoft. Подальші дослідженнях плануються в напрямку написання сценаріїв для оболонки, яка може допомогти автоматизувати завдання, маніпулювати даними та налаштовувати інтерфейс командного рядка (CLI) при використанні інших оболонок.

Список використаних джерел:

1. Види резервного копіювання: повний, інкрементальний та диференціальний бекап. URL: https://www.sim-networks.com/ukr/blog/backup-full-increment-differential

2. Русин Б. П., Погрелюк Л. В., Висоцька В. А., Осипов М. М., Варецький Я. Ю., і Капшій О. В. Архітектура системи дедублікації та розподілу даних у хмарних сховищах під час резервного копіювання. Інформаційні технології та комп'ютерна інженерія. 2019. Т. 45, Вип. 2. С. 40–63.

3. Операційні системи: Методичні вказівки до комп'ютерного практикуму: навч. посіб. для студ. спец. 113 «Прикладна математика», 125 «Кібербезпека» / КПІ ім. Ігоря Сікорського ; уклад.: М. В. Грайворонський, В. В. Демчинський. Електронні текстові дані (1 файл: 1,44 Мбайт). Київ : КПІ ім. Ігоря Сікорського, 2021. 74.

4. Paul Troncone, Carl Albing. Cybersecurity Ops with bash: Attack, Defend, and Analyze from the Command Line 1st Edition. *O'Reilly Media 1st edition* (April 2, 2019). 504 p. ISBN 978-1492041313.

5. Jason Cannon. Shell Scripting: How to Automate Command Line Tasks Using Bash Scripting and Shell Programming. *CreateSpace Independent Publishing Platform* (September 14, 2015). 99 p. ISBN 151738043X.

6. What are some useful tools or resources for learning and improving your shell scripting skills? URL: https://www.linkedin.com/advice/3/what-some-useful-tools-resources-learning-improving-your-shell

7. What is a Bash Script? URL: https://ryanstutorials.net/

8. Steve Parker. Shell Scripting: Expert Recipes for Linux, Bash, and more 1st Edition. *Wrox; 1st edition* (August 30, 2011), 608 p. ISBN-101118024486.

9. Shotts, W. E., Book, A. L. (2009). The Linux command line. Lulu. Com, 2009. 555 p.

10. Гайд для початківців: як писати Shell скрипти URL: https://blog.iteducenter.ua/guides/shell-scripting-for-beginners/

References:

1. Vydy rezervnoho kopiiuvannia: povnyi, inkrementalnyi ta dyferentsialnyi bekap. [Types of backups: full, incremental, and differential backups]. Retrievet from: https://www.sim-networks.com/ukr/blog/backup-full-increment-differential [in Ukrainian].

2. Rusyn, B.P., Pohreliuk, L.V., Vysotska, V.A., Osypov, M.M., Varetskyi, Ya. Yu. & Kapshii, O. V. (2019). Arkhitektura systemy dedublikatsii ta rozpodilu danykh u khmarnykh skhovyshchakh pid chas rezervnoho kopiiuvannia, [Architecture of the system for deduplication and distribution of data in cloud storage during backup]. *Informatsiini tekhnolohii ta kompiuterna inzheneriia – Information technology and society*, 45 (2), 40–63. [in Ukrainian].

3. M. V. Hraivoronskyi, V. V. Demchynskyi. (2021). Operatsiini systemy: Metodychni vkazivky do kompiuternoho praktykumu, [Operating systems: Methodological instructions for a computer workshop]. navch. posib. dlia stud. spets. 113 «Prykladna matematyka», 125 «Kiberbezpeka» / KPI im. Ihoria Sikorskoho ; uklad.: – Elektronni tekstovi dani (1 fail: 1,44 Mbait). – Kyiv : KPI im. Ihoria Sikorskoho, 74. [in Ukrainian].

4. Paul Troncone, Carl Albing. Cybersecurity Ops with bash: Attack, Defend, and Analyze from the Command Line 1st Edition. *O'Reilly Media 1st edition*, 504 p. ISBN 978-1492041313.

5. Jason Cannon. Shell Scripting: How to Automate Command Line Tasks Using Bash Scripting and Shell Programming. *CreateSpace Independent Publishing Platform*, 99 p. ISBN 151738043X.

6. What are some useful tools or resources for learning and improving your shell scripting skills? Retrievet from: https://www.linkedin.com/advice/3/what-some-useful-tools-resources-learning-improving-your-shell.

7. What is a Bash Script? – Retrievet from: https://ryanstutorials.net/.

8. Steve Parker. Shell Scripting: Expert Recipes for Linux, Bash, and more 1st Edition. *Wrox; 1st edition*, 608 p. ISBN-101118024486.

9. Shotts, W. E., & org Book, A. L. (2009). The Linux command line. Lulu. Com, 555 p.

10. Haid dlia pochatkivtsiv: yak pysaty Shell skrypty. Retrievet from: https://blog.iteducenter.ua/guides/shell-scripting-for beginners/ [in Ukrainian].

УДК 007.3+331.108.45+331.445+658.5.011
DOI https://doi.org/10.32689/maup.it.2023.3.3

Станіслав ГОРБАЧЕНКО
доктор економічних наук, професор, завідувач кафедри кібербезпеки Національного університету "Одеська юридична академія", вул. Рішельєвська, 28, Одеса, Україна, індекс 65045 (stasgorbachenko@gmail.com)
ORCID: 0000-0001-8442-9581

Віктор БОЙКО
кандидат технічних наук, доцент кафедри кібербезпеки Національного університет "Одеська юридична академія", вул. Рішельєвська, 28, Одеса, Україна, індекс 65045 (boyko-work@ukr.net)
ORCID: 0000-0001-5929-657X

Stanislav HORBACHENKO
Doctor of Economic Science, Professor, Head of the Department of Cybersecurity at National university "Odesa Law Academy", 28 Rishelievskaya, Odesa, Ukraine, postal code 65045 (stasgorbachenko@gmail.com)

Victor BOYKO
Candidate of Technical Sciences, Senior Lecturer at the Department of Cybersecurity at National university "Odesa Law Academy", 28 Rishelievskaya, Odesa, Ukraine, postal code 65045 (boyko-work@ukr.net)

Бібліографічний опис статті: Горбаченко, С., Бойко, О. (2023). Тестування на проникнення як ефективний інструмент менеджменту кібербезпеки. *Інформаційні технології та суспільство*, 3, 23–29. DOI: https://doi.org/10.32689/maup.it.2023.3.3

Bibliographic description of the article: Horbachenko, S., Boyko, O. (2023). Testuvannya na proniknennya yak efectivniy instrument manajmentu kiberbezpeki [Penetration testing as an effective tool for cybersecurity management]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 23–29. DOI: https://doi.org/10.32689/maup.it.2023.3.3

ТЕСТУВАННЯ НА ПРОНИКНЕННЯ ЯК ЕФЕКТИВНИЙ ІНСТРУМЕНТ МЕНЕДЖМЕНТУ КІБЕРБЕЗПЕКИ

Анотація. *В статті розглянуто теоретичні аспекти розбудови та функціонування системи управління, спрямованої на захист комп'ютерних систем, мереж, даних, та інформації підприємства чи організації від кіберзагроз. В кінцевому рахунку вказана система має трансформуватися у менеджмент кібербезпеки. Останню категорію визначено як процес планування, розробки, впровадження та керування заходами, які спрямовані на захист комп'ютерних систем, мереж і даних від наявних та потенційних кіберзагроз. Виявлено, що основні функції класичного менеджменту, а саме, планування, організування, мотивування та контроль, відповідають і завданням менеджменту кібербезпеки.*

Серед конкретних інструментів які має використовувати менеджмент кібербезпеки виокремлено моніторинг подій, моніторинг мережі, інтрузійне виявлення, системи виявлення аномалій, етичний хакінг тощо.

Зроблено висновок, що поряд з безпосередньо технічними питаннями, істотну частину проблем сучасного кіберзахисту підприємств та організацій спричиняє «людський фактор». Відтак однією з нагальних проблем у цій галузі, поряд із загальним підвищенням технічної грамотності та підготовленості персоналу, є підтримання високого рівня корпоративної пильності та алертності.

З огляду на це, як один з ефективних інструментів підвищення рівня пильності та алертності персоналу, в статті пропонується використовувати процедуру тестування на проникнення у формі Cyber Red Team, при якій тестування виконується зовнішньою командою. Доведено, що тестування на проникнення сприяє розумінню з боки персоналу факту, що вторгнення може здійснюватися різними шляхами, а кіберризики є об'єктивною реальністю. В результаті запропоновано практичні рекомендації щодо проведення пентестів, які мають забезпечити збільшення ефективності та результативності менеджменту кібербезпеки на рівні окремих підприємств та організацій.

Ключові слова: *кіберзахист, менеджмент кібербезпеки, контроль, алертність, тестування на проникнення.*

PENETRATION TESTING AS AN EFFECTIVE TOOL FOR CYBERSECURITY MANAGEMENT

Abstract. *In the modern information environment, special attention is paid to cyber security issues. The growing threat of cyber attacks and unauthorized access to computer systems, networks, data and information of enterprises and organizations presents them with the task of effective cyber security management. Accordingly, the article examines the theoretical aspects of the development and functioning of the management system aimed at protecting against cyber threats.*

One of the key concepts is the transformation of this system into cyber security management. It has been studied that the classical functions of management – planning, organization, motivation and control – are properly correlated with the tasks of cyber security. This allows for an integrated approach to cyber security management, providing effective protection against cyber threats.

Paradoxically, the cornerstone of modern cyber defense is the "human factor". It was found that, in addition to technical aspects, an important role in ensuring cyber security is played by the preparedness and awareness of personnel. Therefore, increasing technical literacy and corporate vigilance become extremely important tasks.

In this context, the article suggests the use of a penetration testing procedure known as the Cyber Red Team. According to this concept, an expert team conducts testing using external methods, carrying out attacks on the system, which allows identifying weak points and gaps in cyber defense.

Finally, the article provides practical recommendations for conducting pentests aimed at improving the effectiveness of cyber security management at the level of individual enterprises and organizations. These measures will contribute to improving the level of vigilance and alertness of personnel, which is an important step in ensuring a high degree of cyber security.

In conclusion, the article highlights the current aspects of the development of the cyber security management system, focusing on the importance of the "human factor" and proposing innovative approaches to increasing the level of cyber security at enterprises and organizations.

Key words: *cybersecurity, cybersecurity management, control, alertness, penetration testing.*

Постановка проблеми у загальному вигляді. Безпека, в різних її проявах, має для сучасного менеджменту велике значення, оскільки дозволяє зменшити ризики втрати активів, порушення операцій, втрати довіри клієнтів і репутації, а також може допомогти підприємству чи організації бути більш стійкими та успішними в непередбачуваних умовах. Серед складових безпеки чільне місце займає інформаційна безпека, тобто захист інформації, даних, конфіденційної інформації та інтелектуальної власності від несанкціонованого доступу, втрати або викрадення.

У свою чергу розвиток інформаційних технологій призвів до повноцінного відокремлення від інформаційної безпеки кола питань пов'язаних із захистом комп'ютерних систем, мереж, даних, програмного забезпечення та інформації від кіберзагроз, кібератак, зловмисних дій та несанкціонованого доступу, тобто, напряму кібербезпеки. На макрорівні кібербезпеку розглядають як найважливішу складову національної та економічної безпеки, адже від початку повномасштабної війни кіберзагрозам підлягають, в першу чергу, об'єкти критичної (промислової та міської) інфраструктури. Атаки таких об'єктів можуть завдавати значних збитків, до того ж часто їхнє технічне та організаційне оснащення або відстає, або знаходиться в процесі оновлення та перебудови.

В процесі побудови системи кіберзахисту суттєву роль завжди відіграє «людський фактор», оскільки будь-який недбалий і недосвідчений співробітник (або, ще гірше – управлінець) здатний звести нанівець заходи технічного характеру. З огляду на це серед інструментів менеджменту кібербезпеки, спрямованих на підготовку співробітників (та й всієї організації в цілому) до можливих вторгнень чи не найважливіше місце займає тестування на проникнення.

Аналіз останніх досліджень і публікацій. Навіть в умовах загального розуміння в науковому середовищі важливості досліджень в сфері забезпечення кіберзахисту, на практиці основна увага зосереджується на, так званій, SHN-тріаді («software-hardware-network»). В цьому сенсі можна, зокрема, виділити наукові праці Дж. Боема, П. Меррата, Л. Шентона, А. Самнера, Х. Алдавуда, Дж. Скіннера, В. Бенсона, Д. Боссарта, С. Ванга та інших. Деякі із зазначених авторів доречно вказували на зростання кількості випадків використання інструментів соціальної інженерії для здійснення вторгнень і важливість відповідної роботи з персоналом. Однак саме управлінському аспекту формування ефективної системи кіберзахисту все ще приділяється недостатньо уваги.

Формулювання мети статті. Мета даної статті полягає в аналізі та обгрунтуванні ефективності тестування на проникнення як ключового інструменту в області менеджменту кібербезпеки. Досліджено основні аспекти та переваги застосування тестування на проникнення, включаючи його вплив на виявлення потенційних вразливостей, забезпечення недоступності для зловмисників, та зниження ризиків кібератак. Також, стаття має на меті надати читачам практичні рекомендації щодо впровадження та оптимального використання тестування на проникнення для підвищення рівня кібербезпеки в організаціях.

Виклад основного матеріалу. Вагомість проблематики кібербезпеки підтверджується тим, що протягом 2022 року на Україну було здійснено понад 7 тис. кібератак що у 2,8 разів більше, ніж у 2021 році. Серед опрацьованих протягом 2022 року фахівцями CERT-UA 2194 кіберінцидентів 120 стосувалися фінансового сектору, 156 – комерційних організацій та 92 – сектору телекомунікацій і розробки програмного забезпечення. Крім того за результатами 2022 р. Україна займає 2 місце серед найбільш атакованих країн світу після США [1].

В огляду на наявність вищевказаних проблем на державному рівні основним завданням кібербезпеки виступає захист життєво важливих інтересів людини і громадянина, суспільства та держави під час використання кіберпростору, за якої забезпечуються сталий розвиток інформаційного суспільства та цифрового комунікативного середовища, своєчасне виявлення, запобігання і нейтралізація реальних і потенційних загроз національній безпеці України у кіберпросторі [2].

В той самий час для окремих підприємств та організацій кібербезпека – це сукупність заходів, процедур, стратегій та політик, спрямованих на захист від ризиків, загроз і небезпек, що спричиняються кіберсередовищем і можуть впливати на операції, активи, працівників, імідж, фінансовий стан, репутацію тощо.

Забезпечення ефективного кіберзахисту потребує не тільки сучасних технологічних рішень, а й відповідної управлінської складової. Адже основним джерелом загроз все частіше виступає «людський чинник». Як управлінська категорія менеджмент кібербезпеки – це процес планування, розробки, впровадження та керування заходами, які спрямовані на захист комп'ютерних систем, мереж і даних від наявних та потенційних кіберзагроз.

Робота з персоналом та зменшення впливу «людського фактору» має здійснюватися у відповідності з основними функціями менеджменту, які транслюються й на менеджмент кібербезпеки. Такими функціями є планування, організування, мотивування та контроль.

Планування, як функція менеджменту – це процес розробки системи заходів, спрямованих на досягнення певних цілей. Основною одиницею планування в менеджменті кібербезпеки найчастіше виступають задача (з такими характеристиками як: тип, пріоритет, компоненти, зміст, вартість, обмеження тощо) та подія (яка характеризується ймовірністю та можливими втратами).

Організаційна функція менеджменту кібербезпеки спрямована на забезпечення впорядкування процесу управління в цілому. З одного боку, менеджменту кібербезпеки не притаманні складні організаційні структури. Навпаки, в ньому превалює проєктний підхід, зокрема, в процесі управління розробкою та супроводу програмних продуктів. З іншого боку проєкти кібербезпеки є комплексними і відрізняються такими характеристиками як складність, масштабність і різноманітність [3, с. 107].

Мотивація передбачає управлінську діяльність, спрямовану на спонукання до вибору співробітниками того або іншого типу поведінки в залежності від сили впливу стимулів, мотивів і очікуваних результатів. Вона дозволяє розкрити потенційні можливості персоналу, в тому числі й креативні здібності, а також збільшити продуктивність праці. Це обумовлюється тим, що роботодавець усвідомлює свою залежність від фахових працівників, він готовий мотивувати і берегти команду, а тому враховує матеріальні, соціальні і кар'єрні запити [4]. При вирішенні питань кібербезпеки використовують такі інструменти мотивації як розуміння персоналом загальної мети захисту інформації, а також можливих втрат, у тому числі фінансових та репутаційних, створення відповідної корпоративної культури, забезпечення індивідуального підходу до окремих працівників.

І, нарешті, функція контролю передбачає управлінську діяльність, спрямовану на виявлення, виправлення та попередження відхилень досягнутих результатів від намічених параметрів та цілей. Для менеджменту кібербезпеки контроль передбачає постійний моніторинг об'єкту та процесів з метою перевірки відповідності поточного стану об'єкта його прогнозованому стану, згідно з вимогами клієнтів, технологічними можливостями та законодавчими обмеженнями. Щодо конкретних інструментів які використовує менеджмент кібербезпеки в процесі контролю слід, зокрема, відзначити моніторинг подій, моніторинг мережі, інтрузійне виявлення, системи виявлення аномалій, етичний хакінг тощо.

На мікрорівні суб'єктом менеджменту кібербезпеки виступає будь-яка організація або підприємство, незалежно від розміру або сфери діяльності. Головною ознакою є те, що вони використовують комп'ютерні системи, мережі та залежать від цифрових технологій та мережі Інтернет, а також потребують захисту інформації, активів та інфраструктури від кіберзагроз і кібератак відразу з декількох вимірів.

Перший вимір – це категоріальний розподіл заходів щодо кіберзахисту. Адже ефективний кіберзахист будується як комплексне поєднання кількох підсистем з різними «зонами відповідальності». Зазвичай виділяють три основні зони відповідальності: фізичну, зовнішній периметр, внутрішній простір.

Системи, задіяні у фізичній зоні, визначають що робити, щоб не допустити противника фізичної присутності противника чи зловмисника на території організації або його доступу до матеріальних ресурсів та носіїв організації. Це може включати як очевидні (фізична охорона периметра, системи сигналізації та відеоспостереження), так і неочевидні аспекти, наприклад, охорону та захист комунікацій між філіями організації, питання використання співробітниками фізичних носіїв інформації.

Сюди має входити захист від вардрайвінгу (wardriving) та інших варіантів злому бездротових мереж. З одного боку, їх можна було б віднести до інформаційних атак, на зовнішній периметр, які відбуваються без матеріального підключення до мережі. Однак, вони не можуть бути зроблені без підключення до точок бездротового доступу, яке в більшості випадків обмежено фізичною відстанню, а отже, вимагає присутності на території або в околиці периметра організації.

Кошти та системи захисту зони відповідальності зовнішнього інформаційного периметра націлені на запобігання проникненню зловмисника через зовнішній інформаційний периметр організації. Найбільш поширеним, але не єдиним засобом захисту такого класу зазвичай є система фаєрволів (firewall) або брандмауерів (brandmauer), яка дозволяє фільтрувати інформаційний трафік, що проходить через задані заздалегідь точки периметра.

Захист внутрішнього простору спрямовано зменшення втрат, або повне їх виняток у припущенні, що фізичний чи зовнішній інформаційний периметр був компрометований і зловмисник виявився "всередині" периметра. До засобів захисту насамперед відносять політику поділу прав та ролей, засоби резервного копіювання даних тощо.

Перераховані вище зони відповідальності з одного боку дозволяють розподілити ресурси та засоби при організації кіберзахисту, з іншого слід розуміти, що як захист, так і атака може бути комплексною – наприклад, у класичному випадку хробака Stuxnet були задіяні способи атаки одночасно через фізичну (заражений флешдрайв) і внутрішню (прохід крізь захисні системи Windows) зони безпеки [5].

Комплексний підхід можна розглядати з погляду іншої категорії системи. Насамперед – поділ заходів на технічні та організаційні. Відповідно, категоріями аналізованого захисту може з одного боку виступати інформаційно-комунікаційна система як апаратно-програмний комплекс, а з іншого – користувачі та оператори цих систем.

У внутрішній зоні захисту користувачі можуть ігнорувати рекомендації щодо резервного копіювання (особливо, якщо не впроваджено автоматизовану систему бекапів), обмінюватися паролями «бо так зручніше», залишати собі шпаргалки зі складними паролями поруч із робочим місцем. Відповідно до досліджень витоків паролів – досі досить поширена хрестоматійна помилка, коли як пароль використовується слово «пароль», «qwerty» та ще кілька занадто простих та розповсюджених варіантів паролів [6].

У зовнішньому периметрі захисту частою помилкою є неправильна конфігурація фаєрволла, або повна його відсутність – оскільки оператору не вистачає кваліфікації для його налаштування та тестування. Оборотною стороною непрофесіоналізму можуть бути занадто закриті політики ланцюжків (chains), що ускладнюють функціонування організації. Також часто зустрічається стереотип мислення, що передбачає, що «air gap» – повна відсутність провідного підключення до зовнішніх мереж – забезпечує надійний захист від зовнішнього вторгнення. Як показує приклад описаного вище хробака Stuxnet, що цілком успішно подолав захист такого типу, це досить небезпечна помилка.

Помилки у фізичному захисті на перший погляд цілком очевидні – наприклад відсутність контролю доступу (вахти) на вході, недбале ставлення до замків і ключів, проте, коли йдеться про контроль інформаційних ресурсів, можливі помилки іншого роду, оскільки фізичні системи вторгнення іноді дозволяють «вторгатися не вторгаючись» – наприклад, шляхом злому бездротового роутера, доступ до якого є за межами фізичних кордонів, контрольованих організацією.

При цьому слід розуміти, що механічне ускладнення та посилення заходів безпеки, наприклад, шляхом введення додаткових заходів захисту у вигляді ускладнення паролів та завдання двофакторної автентифікації найчастіше не призводить до поліпшення ситуації. Наприклад, за оцінками [7] на 2018 рік близько 16 мільйонів пар паролів (включаючи 30% змінених після витоку паролів) можуть бути зламані лише за 10 спроб. Дослідження 2022 [8] виявило використання в паролі особистої інформації (56%), повторне використання пароля (69%), використання поширених шаблонів (81,3%). Ці цифри можуть відрізнятися від дослідження до дослідження, однак загальний зміст той самий – низька поінформованість і недостатня технічна кваліфікація користувачів тягне за собою появу вразливостей навіть у найдосконаліших системах [9].

Ступінь і надійність заходів безпеки (наприклад, складність використовуваних паролів) часто входить у конфлікт із зручністю використання системи (необхідністю частої авторизації), тому таке механічне посилення заходів безпеки зазвичай працює лише до певної (найчастіше – недостатньої) межі, за якою користувачі починають обходити систему, спрощуючи собі життя. Наприклад, є приклади підприємств та організацій, в яких система поділу прав і ролей була нівельована тим, що використовувався один обліковий запис користувача і один пароль на всю організацію – таким чином зловмисник отримав доступ до одного облікового запису, отримував доступ відразу до всієї системи.

Таким чином, для управління людським фактором, як складовою системи кібернетичного захисту організації, потрібен особливий підхід і впровадження правил і політик роботи організації тут є лише одним із аспектів проблеми.

Однією з очевидних і нагальних проблем у цій галузі, поряд із загальним підвищенням технічної грамотності та підготовленості персоналу, є підтримка високого рівня пильності (vigilance) та алерт-

ності (alertness) організації. У тому числі – культивування серйозного ставлення до заходів безпеки та уваги до різноманітних інцидентів, які на перший погляд можуть бути незначними (наприклад, вхід до системи під логіном співробітника, який зараз перебуває у відпустці).

Складність підтримки такого рівня пов'язана по-перше з тим, що з одного боку такі якості (режим роботи мозку) можуть вимагати витрат з погляду ресурсів [10] і можуть вступати в конфлікт із прямими обов'язками співробітника з виконання основних завдань у створенні, і навіть із загальними особливостями сучасного життя. Як приклад – у роботах [11],[12] розглянуто дію депривації сну (sleep deprivation) на загальну алертність працівників організації.

Другим джерелом складності підтримки пильності та алертності є загальна властивість адаптивності психіки людини [13], [14]. Завдяки цій властивості будь-який стимул так чи інакше «втрачає новизну», що знижує рівень алертності та уваги при його пред'явленні. Наслідком вказаної ситуації є неможливість постійно підтримувати пильність на максимально високому рівні.

Також слід брати до уваги, що обидва ефекти (кількість доступного ресурсу та загальний рівень «новизни» стимулів) досить складно оцінюються, особливо в умовах повсякденного функціонування підприємства чи організації.

В цьому сенсі ефективним управлінським інструментом перевірки рівня кіберзахисту є тестування на проникнення або pentesting. Воно може проводитись різними засобами та різними командами і відігравати велику роль не тільки в суто технічному аспекті, як розглянуто у багатьох наукових працях [15], [16], але й як інструмент менеджменту кібербезпеки.

Періодичне проведення тестування на проникнення може виконувати у тому числі функцію ефективної підтримки комплексної готовності персоналу. При цьому слід виділити кілька різних форм вказаного тестування – ручний, автоматизований, виконуваний власним персоналом, виконуваний фахівцями, залученими із боку і таке інше.

Функціонал тестування на проникнення корисно розглядати, як частину спільної задачі, спрямованої на підвищення організаційно-соціального аспекту кіберзахисту організації: підтримання високого рівня пильності та алертності персоналу організації, підвищення технічної грамотності персоналу, впровадження безпечних практик тощо; увагу в першу чергу на соціальну складову (фішинг, соціальна інженерія, доксинг, OSINT), які набувають все більшого значення: велика кількість інформації, яку можна отримати в сучасних мережах, технології підробки та імітації особистості (зокрема – deepfake), які можна реалізувати в атаці тощо.

У практиці тестування на проникнення склалося своє найменування команд – за аналогією з військовими навчаннями: «своя команда» – сині (CBT – Cyber Blue Team), «чужі» – червоні (CRT – Cyber Red Team). При цьому для CRT виділяють основні напрямки вектора атаки, що частково збігаються і перетинаються з описаними вище: фізична атака – на фізичні ресурси та матеріальні об'єкти обмеження – проникнення в будівлю, приміщення і таке інше; вектор атак «ззовні», спрямований на впровадження в інформаційні системи організації зовні; внутрішні атаки припускають, що атака «ззовні» пройшла успішно і доступ всередину периметра вже отримано; гібридна атака, що поєднує всі перелічені види атак у різних пропорціях.

Імітація атаки реальних зловмисників крім інших своїх переваг дозволяє тримати персонал у формі – розуміння факту, що вторгнення може йти різними шляхами дозволяє робити його значущим психологічно, що у свою чергу позначається на пильності та алертності персоналу. Узагальнена загроза кібератаки – це «звичний ворог», який не викликає необхідної віддачі та сприймається як незначний. При цьому інформація про те, що відбувається реальне тестування з реальними атаками, буде свіжою і сприймається психологічно по-іншому, що також безпосередньо впливає на персонал організації.

Як інструмент менеджменту кібербезпеки CRT pentesting характеризується наступними перевагами:

– використання CRT дозволяє виявляти та ідентифікувати слабкі місця та вразливості організації на всіх рівнях – у тому числі організаційно-соціальному;

– дозволяє проводити незалежне тестування;

– дозволяє виявляти потенційні ризики, які в іншому випадку залишилися б поза увагою (навіть zero day, якщо CRT є достатньо кваліфікованою) і дає можливість скасувати або змінити потенційно небезпечні рішення (наприклад, винесення чутливих даних у потенційно небезпечний хмарний сервіс);

– тренує команду «від захисту» CBT в умовах, максимально наближених до реальних, дозволяє організовувати та підтримувати творчу та здорову конкурентну атмосферу;

– дозволяє уникнути рутинних атак і застою у плануванні заходів кіберзахисту, оскільки CRT зазвичай не обмежені (або слабко обмежені) у виборі заходів, сценаріїв та інструментів;

– впливає на пильність і алертність всього персоналу організації, адже наслідки та висновки з тестування мають не тільки впливати на суто технічні рішення, а й змінювати організаційні підходи менеджменту кібербезпеки і формувати відповідну корпоративну культуру.

Однак, за всієї привабливості «повноконтактного пентесту», слід обмежувати та планувати заздалегідь діяльність CRT – пентестери схильні захоплюватися і можуть навмисно чи ненавмисно завдати організації шкоди. Повинні вживатися заходи безпеки для того, щоб позитивний результат тестування на проникнення не виявився «надто хорошим» і не призвів до виникнення проблем (наприклад, блокування робочого процесу, або витоку конфіденційної інформації).

Планування процесів тестування на проникнення має виконуватися з максимальним залученням всіх стейкхолдерів, а результати проведення пентесту, хоча б частково, мають ставати предметом обговорень та розбору не лише CBT, а й усього колективу за участю вищого керівництва.

Паралельно із цим слід правильно підбирати частоту та серйозність проведення тестувань. Занадто рідкісне проведення пентестів не дозволяє виконати достатню кількість перевірок, хоча все одно краще ніж повна відсутність тестів на проникнення. З іншого боку, занадто часте проведення пентесту пов'язане із втратами продуктивності та простоями основного робочого процесу. Тобто, воно призводить до перевитрати ресурсів, а також до звикання та перетворення пентесту на «знайому загрозу», що знищує всі організаційно-соціальні переваги від проведення таких процедур.

Так само й питання «звикання та адаптації» робить бажаною ротацію CRT, адже коли персонал CRT вичерпає свої можливості він має передати максимум досвіду та порад CBT. Цінною в цьому сенсі є практика purple teaming – поєднання фахівців із CRT/CBT, які працюють разом. Фахівці CBT, що беруть участь в атаці зможуть краще вчитися працювати на захист. Аналогічно для CRT розуміння процесів захисту та логіки дій CBT дозволяє більш креативно підходити до процесів тестування на проникнення. І саме правильно побудована взаємодія та обмін досвідом, знаннями та інформацією взаємно збагачуватиме обидві команди.

Висновки. Більшу частину проблем кіберзахисту будь-якого підприємства чи організації ще й досі спричиняє «людський фактор», оскільки технічна неграмотність і недбалість персоналу може звести нанівець ефект від будь-яких технічних засобів захисту. Як один з ефективних інструментів підвищення рівня пильності та алертності персоналу, мотивуючого фактору для впровадження ефективних політик та практик безпеки пропонується використовувати тестування на проникнення, яке дотепер розглядається швидше як чисто технічний інструмент контролю кібербезпеки.

Найбільш ефективною процедурою тестування на проникнення є CRT, або purple teaming, що проводиться відповідно до запропонованих вище рекомендацій. Використання означеного інструменту дозволяє зменшити кіберризики в основі яких знаходиться «людська складова» та, одночасно, збільшити загальну ефективність всієї системи менеджменту кібербезпеки.

Список використаних джерел:

1. У 2022 році кількість кібератак на Україну зросла майже втричі. 2023. URL: https://forbes.ua/news/v-2022-rotsi-kilkist-kiberatak-na-ukrainu-zrosla-mayzhe-vtrichi-90-khakerskikh-grup-z-rf-kontrolyuyut-siloviki-04052023-13454

2. Про основні засади забезпечення кібербезпеки України : закон України від 05.10.2017 р. № 2163-VIII / Верховна Рада України. 2017. URL: http://zakon3.rada.gov.ua/laws/show/2163-19

3. Сметанюк О.А., Бондарчук А. В. Особливості системи управління проєктами в it-компаніях. *Агросвіт*. 2020. № 10. С. 105–111.

4. Орлова О.М. Особливості управління персоналом в IT-сфері. URL: http://www.visnyk-econom.uzhnu.uz.ua/archive/11_2017ua/28.pdf

5. Barzashka I. Are Cyber-Weapons Effective? Assessing Stuxnet's Impact on the Iranian Enrichment Programme. *The RUSI Journal. Taylor & Francis*. 2013. Vol. 158, № 2. P. 48–56.

6. Jaeger D., et al. Analysis of Publicly Leaked Credentials and the Long Story of Password. 2016. P. 1–19.

7. Wang C., et al. The Next Domino to Fall */ Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy. CODASPY18 The 8th ACM Conference on Data; Application Security; Privacy March 19 – 21.* 2018. Tempe, AZ, USA. P. 196–203.

8. Tanni T., et al. Is My Password Strong Enough? : A Study on User Perception in The Developing World. EAI Endorsed Transactions on Creative Technologies. *European Alliance for Innovation n.o*. 2022. Vol. 9, № 30. P. 1–12.

9. Hitchcock K. Linux System Administration for the 2020s : The Modern Sysadmin Leaving Behind the Culture of Build and Maintain. Apress, 2022. P. 328.

10. Aston-Jones G. Brain structures and receptors involved in alertness. *Sleep Medicine. Elsevier BV*. 2005. Vol. 6. P. 3–7.

11. Caldwell J.A., Caldwell J.L., Schmidt R.M. Alertness management strategies for operational contexts URL: https://pubmed.ncbi.nlm.nih.gov/18359253/

12. Niu S.F., Chung M.H., Chen C.H., Hegney D., OBrien A., Chou K.R. The Effect of Shift Rotation on Employee Cortisol Profile, Sleep Quality, Fatigue, and Attention Level. *Journal of Nursing Research. Ovid Technologies (Wolters Kluwer Health)*. 2011. Vol. 19. № 1. P. 68–81.

13. Oken B., et al. Vigilance state fluctuations and performance using braincomputer interface for communication. *Brain-Computer Interfaces. Informa UK Limited*. 2018. Vol. 5, № 4. P. 146–156.

14. Langner R., Eickhoff S. B. Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. *Psychological Bulletin. – American Psychological Association (APA).* 2013. Vol. 139, № 4. P. 870–900.

15. Zakaria M. N., et al. Review of Standardization for Penetration Testing Reports and Documents. *2019 6th International Conference on Research and Innovation in Information Systems (ICRIIS).* 2019. P. 1–5.

16. Shebli H.M.Z.A., Beheshti B.D. A study on penetration testing process and tools. 2018 IEEE Long Island Systems, Applications and Technology Conference (LISAT). 2018. P. 1–7.

References:

1. U 2022 rotsi kilkist kiberatak na Ukraini zrosla mayje vtichi. [In 2022, the number of cyberattacks on Ukraine increased almost threefold]. Retrieved from https://forbes.ua/news/v-2022-rotsi-kilkist-kiberatak-na-ukrainu-zrosla-mayzhe-vtrichi-90-khakerskikh-grup-z-rf-kontrolyuyut-siloviki-04052023-13454 [in Ukrainian].

2. Pro osnovni zasadi zabezpechennya kiberbezpeki Ukraini : zakon Ukraini vid 05.06.2017 r. № 2163-VIII / Verhovna Rada Ukraini. [On the main principles of ensuring cyber security of Ukraine: Law of Ukraine dated 05.10.2017 No. 2163-VIII / Verkhovna Rada of Ukraine]. 2017. Retrieved from http://zakon3.rada.gov.ua/laws/show/2163-19 [in Ukrainian].

3. Smetaniuk, O.A., Bondarchuk, A.V. (2020). *Osoblivosti sistemi upravlinnya proyektami v it-companiyah. Agrosvit. [Peculiarities of the project management system in IT companies. Agroworld].* № 10. P. 105–111. Ukraine. [in Ukrainian].

4. Orlova, O.M. Osoblivosti upravlinnya personalom v IT-sferi. [Peculiarities of personnel management in the IT sphere]. Ukraine. Retrieved from http://www.visnyk-econom.uzhnu.uz.ua/archive/11_2017ua/28.pdf [in Ukrainian].

5. Barzashka, I. (2013). Are Cyber-Weapons Effective? Assessing Stuxnet's Impact on the Iranian Enrichment Programme. USA: The RUSI Journal. Taylor & Francis. Vol. 158, № 2. P. 48–56.

6. Jaeger, D., Pelchen, C., Graupner, H., Cheng, F., Meinel, C. (2016). Analysis of Publicly Leaked Credentials and the Long Story of Password (Re-)use. USA. P. 1–19.

7. Wang, C., Jan, S.TK., Hu, H., Bossart, D., Wang, G. (2018). The Next Domino to Fall / *Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy. CODASPY18 The 8th ACM Conference on Data; Application Security; Privacy March 19 – 21*. USA: Tempe, AZ. P. 196–203.

8. Tanni, T., Taharat, T., Parvez, M., Rumee, S., Zaber, M. (2022). Is My Password Strong Enough?: A Study on User Perception in The Developing World. USA: EAI Endorsed Transactions on Creative Technologies. – European Alliance for Innovation n.o. Vol. 9, № 30. P. 1–12.

9. Hitchcock, K. (2022). Linux System Administration for the 2020s: The Modern Sysadmin Leaving Behind the Culture of Build and Maintain. Apress. USA. P. 328.

10. Aston-Jones, G. (2005). Brain structures and receptors involved in alertness. USA: Sleep Medicine. Elsevier BV. Vol. 6. P. S3–S7.

11. Caldwell, J.A., Caldwell, J.L., Schmidt, R.M. Alertness management strategies for operational contexts. USA. Retrieved from https://pubmed.ncbi.nlm.nih.gov/18359253/.

12. Niu, S.F., Chung, M.H., Chen, C.H., Hegney, D., OBrien, A., Chou, K.R. (2011). The Effect of Shift Rotation on Employee Cortisol Profile, Sleep Quality, Fatigue, and Attention Level. USA: *Journal of Nursing Research. Ovid Technologies (Wolters Kluwer Health)*. Vol. 19. № 1. P. 68–81.

13. Oken, B., Memmott, T., Eddy, B., Wiedrick, J., Fried-Oken. M. (2018). Vigilance state fluctuations and performance using braincomputer interface for communication. United Kingdom: Brain-Computer Interfaces. Informa UK Limited. Vol. 5 № 4. P. 146–156.

14. Langner, R., Eickhoff, S. B. (2013). Sustaining attention to simple tasks: A meta-analytic review of the neural mechanisms of vigilant attention. USA: *Psychological Bulletin. American Psychological Association (APA).* Vol. 139, no. 4. P. 870–900.

15. Zakaria, M.N., Phin, P.A., Mohmad, N., Ismail, S.A., Kama, M.N., Yusop, O.A. (2019). Review of Standardization for Penetration Testing Reports and Documents. *USA: 6th International Conference on Research and Innovation in Information Systems (ICRIIS)*. P. 1–5.

16. Shebli, H.M.Z.A., Beheshti, B.D. (2018). A study on penetration testing process and tools. USA: 2018 IEEE Long Island Systems, Applications and Technology Conference (LISAT). P. 1–7.

УДК 004.021
DOI https://doi.org/10.32689/maup.it.2023.3.4

Роман ЗОЛОТУХА
аспірант зі спеціальності 122 – Комп'ютерні науки, факультету інформаційних технологій, Національного Університету Біоресурсів та Природокористування України, вул. Героїв Оборони, 16А, Київ, Україна, індекс 03041 (remko740@gmail.com)

ORCID: 0000-0003-3099-722X

Олена ГЛАЗУНОВА
доктор педагогічних наук, професор, декан факультету інформаційних технологій, Національного Університету Біоресурсів та Природокористування України, вул. Героїв Оборони, 16А, Київ, Україна, індекс 03041 (o-glazunova@nubip.edu.ua)

ORCID: 0000-0002-0136-4936

Roman ZOLOTUKHA
Postgraduate Student of 122 specialty – Computer sciences, Faculty of Information Technology, National University of Life and Environmental Sciences of Ukraine, 16A Heroiv Oborony str., Kyiv, Ukraine, postal code 03041 (remko740@gmail.com)

Olena GLAZUNOVA
Doctor of Pedagogical Sciences, Professor, Dean of the Faculty of Information Technologies, National University of Life and Environmental Sciences of Ukraine, 16A Heroiv Oborony Str., Kyiv, Ukraine, postal code 03041 (o-glazunova@nubip. edu.ua)

Бібліографічний опис статті: Золотуха, Р., Глазунова, О. (2023). Алгоритм розпізнавання тексту з PDF-резюме для автоматизації підбору кандидатів в IT проекти. *Інформаційні технології та суспільство*, 3, 30–38. DOI: https://doi.org/10.32689/maup.it.2023.3.4

Bibliographic description of the article: Zolotukha, R., Glazunova O. (2023). Alhorytm rozpiznavannia tekstu z PDF-reziume dlia avtomatyzatsii pidboru kandydativ v IT proekty [Text recognition algorithm from PDF resumes to automate the selection of candidates for IT projects]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 30–38. DOI: https://doi.org/10.32689/maup.it.2023.3.4

АЛГОРИТМ РОЗПІЗНАВАННЯ ТЕКСТУ З PDF-РЕЗЮМЕ ДЛЯ АВТОМАТИЗАЦІЇ ПІДБОРУ КАНДИДАТІВ В IT ПРОЕКТИ

Анотація. *У теперішній час, розвиток технологій та зростання індустрії інформаційних технологій супроводжується безпрецедентним попитом на висококваліфікованих інженерів та спеціалістів у галузі IT. У даній статті автори розглядають проблему автоматизації підбору кандидатів в IT проекти та пропонуємо алгоритм розпізнавання тексту з PDF-резюме, який, за допомогою мови Python, значно спрощує та прискорює процес відбору кандидатів. Алгоритм використовує сучасні інструменти обробки природної мови (Natural Language Processing, NLP) та бібліотеки для роботи з PDF-файлами з метою виділення ключової інформації з резюме кандидатів. Він розпізнає важливі дані, такі як освіта, навички, контактна інформація та інше, і структурує їх в легкозрозумілий формат. Результати нашого дослідження вказують на ефективність запропонованого алгоритму та його здатність до швидкого та точного аналізу великої кількості резюме. Це відкриває широкі можливості для впровадження автоматизованих систем підбору кандидатів в IT галузі, що збільшує продуктивність та сприяє ефективному використанню ресурсів у сфері HR. У статті ми також обговорюємо потенціал розвитку даного алгоритму, включаючи можливість розширення підтримуваних мов, вдосконалення процесу розпізнавання навичок та врахування специфіки окремих IT-галузей. Ми підкреслюємо важливість інтеграції машинного навчання для поліпшення точності та аналізу патернів у резюме кандидатів. Автори статті ставлять перед собою завдання вдосконалення та спрощення процесу підбору кандидатів в IT проекти, що допоможе підприємствам більш ефективно використовувати свій інтелектуальний потенціал та забезпечити наявність висококваліфікованих спеціалістів у сфері інформаційних технологій.*

Ключові слова: *розпізнавання тексту, PDF-резюме, автоматизація підбору кандидатів, IT, NLP, HR технології.*

TEXT RECOGNITION ALGORITHM FROM PDF RESUMES TO AUTOMATE THE SELECTION OF CANDIDATES FOR IT PROJECTS

Abstract. *Nowadays, the development of technology and the growth of the IT industry are accompanied by an unprecedented demand for highly qualified engineers and IT specialists. In this article, the authors consider the problem of automating the selection of candidates for IT projects and propose a text recognition algorithm from PDF resumes that, using the Python*

language, greatly simplifies and speeds up the candidate selection process. The algorithm uses modern Natural Language Processing (NLP) tools and libraries to work with PDF files to extract key information from candidates' resumes. It recognizes important data, such as education, skills, contact information, etc., and structures it into an easily understandable format. The results of our study indicate the effectiveness of the proposed algorithm and its ability to quickly and accurately analyze a large number of resumes. This opens up wide opportunities for the implementation of automated candidate selection systems in the IT industry, which increases productivity and promotes the efficient use of HR resources. In the article, we also discuss the potential for the development of this algorithm, including the possibility of expanding the supported languages, improving the skill recognition process, and taking into account the specifics of individual IT industries. We emphasize the importance of integrating machine learning to improve the accuracy and analysis of patterns in candidates' resumes. The authors of the article aim to improve and simplify the process of recruiting candidates for IT projects, which will help enterprises to use their intellectual potential more efficiently and ensure the availability of highly qualified IT specialists.

Key words: *text recognition, PDF resume, automation of candidate selection, IT, NLP, HR technologies.*

Актуальність проблеми. У сучасному світі інформаційних технологій індустрія надзвичайно динамічна та швидкозмінна. З кожним роком з'являються нові технології, інструменти та професійні навички, які стають необхідними для успішної кар'єри в цьому секторі. Ця стійка динаміка вимагає від підприємств залучення висококваліфікованих IT-фахівців, що відповідають сучасним вимогам ринку.

В умовах повномасштабного вторгнення рф український IT ринок показав неабияку стійкість, проте і він суттєво втратив, особливо в кількості активних вакансій. Ринок кандидата змінився на ринок мопаній, а з цією зміною і суттєво збільшилось навантаження на HR департаменти. Згідно статистики інтернет порталу DOU середня кількість вакансій в порівнянні з 2022 роком впала в більше ніж 2 рази, а кількість відгуків на вакансію зросла в 5 разів. [5]

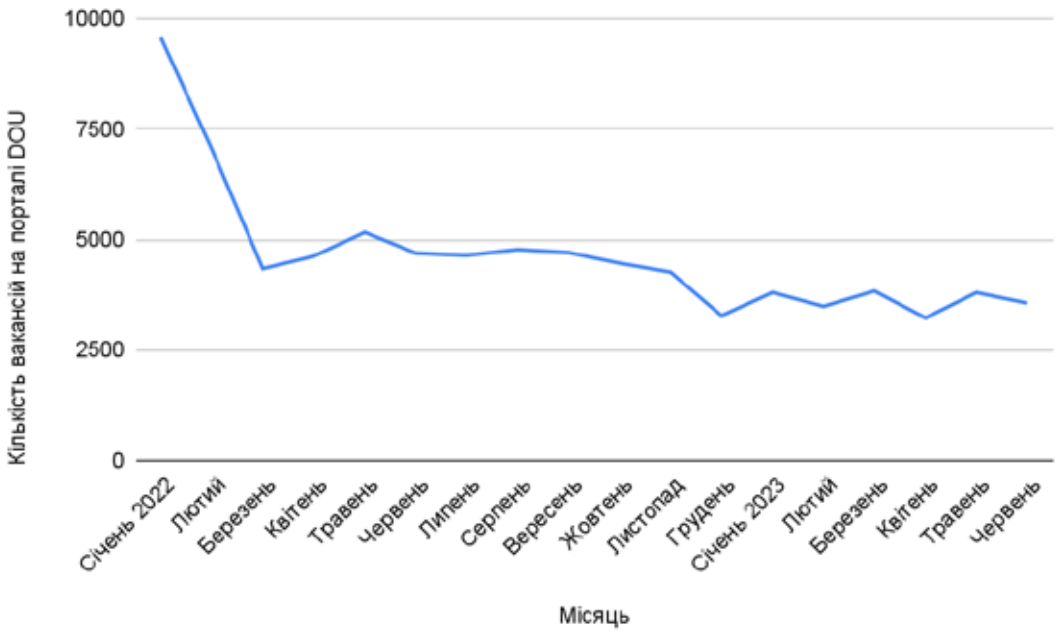


Рис. 1. Кількість IT вакансій на порталі DOU

В умовах постійного зростання кількості заявок на робочі позиції в галузі IT, завдання відбору кращих кандидатів стає вкрай складним і вимагає нових підходів. Інтелектуальні системи, здатні аналізувати та оцінювати професійні навички та досвід кандидатів, стають ключовими інструментами для оптимізації цього процесу.

У даній науковій статті ми досліджуємо проблему автоматизації підбору кандидатів в IT проекти та пропонуємо алгоритм розпізнавання тексту з PDF-резюме. Цей алгоритм дозволяє ефективно виділяти ключову інформацію з резюме кандидатів, таку як освіта, навички та контактна інформація, та автоматично структурувати її для подальшого аналізу або зберігання в профільних HR базах.

Наша робота спрямована на вирішення важливої проблеми в сфері HR та рекрутингу в IT галузі. Ми розглянемо деталі алгоритму, його архітектуру, методи та результати використання. Крім того, ми проаналізуємо можливості подальшого розвитку та вдосконалення цього підходу для оптимізації відбору кандидатів у сучасному IT середовищі.

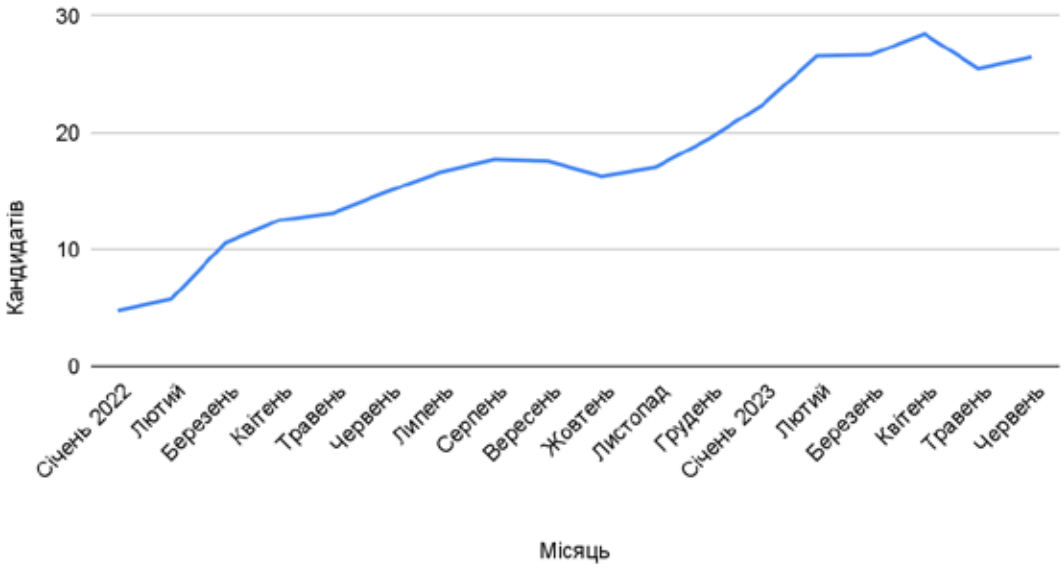


Рис. 2. Середня кількість надісланих резюме кандидатів

Аналіз останніх досліджень та публікацій. Проблемі формування команд для реалізації проектів, а також управлінню персоналом присвячено багато вітчизняних та іноземних праць. Зокрема у роботі [1] розглянуто існуючі в науковій літературі методи розробки стратегій управління персоналом, а також запропоновано методичні положення розробки загальної стратегії управління персоналом. У статті [2] авторами досліджено та запропоновано прототип математичної моделі управління кадрами з урахуванням специфіки діяльності ІТ-компаній з цільовою функцією моделі спрямованою на оптимізацію часу, який витрачають HR-менеджери на роботу з підбору кадрів в ІТ-команди.

Дослідженню питанню автоматизації роботи з текстовими файлами присвячують свої роботи все більше науковців. У статті [3] визначає ефективність використання інструментів обробки природної мови та систем штучного інтелекту для аналізу тексту в інтернеті. Автор розглядає, які інструменти можуть бути використані для виявлення ключових слів у тексті оголошень, а також як ці інструменти можуть бути поєднані зі спеціалізованими моделями машинного навчання для виявлення шахрайської та зловмисної інформації. Стаття [4] присвячена визначенню області проведення досліджень, пов'язаних з підключенням зацікавлених осіб до процесу розробки програмного забезпечення через оцінювання сприйманої якості сервіс-орієнтованих систем в контексті їхнього використання, коли початкова специфікація системи задана природною мовою. Пропонується використання технології аналізу природної мови для отримання інформації про область застосування з цієї специфікації у форматі спеціальних моделей перепроектування, які є сумісними з основними модулями імітаційного рішення.

Також питанню автоматизації процесу підбору присвячена велика кількість наукових статей, зокрема у статті [6] розглянута можливість використання підходу з застосуванням скінченного автомата Мура для створення 3D резюме у процедурах підбору, сегментації та навчання кадрів. У даній роботі 3D резюме розглядається як документ, який представляє професійні досягнення особи у трьох вимірах. Використання автомата пропонується для автоматизації перебору питань та завдань інтерактивного резюме в залежності від попередніх відповідей респондента.

Метою статті є розробка та представлення алгоритму розпізнавання тексту з PDF-резюме для автоматизації процесу підбору кандидатів в ІТ проекти.

Виклад основного матеріалу.

Для розробки алгоритму нами була розроблена модель потоків даних процесу подачі резюме кандидатом. Кандидат подає заявку на вакансію, прикріплюючи резюме у форматі PDF. Резюме обробляється системою і виділяє ключову інформацію про кандидата: контактну інформацію, місце навчання та навички. Ця інформація структурується і вноситься в базу даних HR, де HR-спеціаліст може бачити потрібних кандидатів, використовуючи задані фільтри. Варто зазначити, що запропонований процес руху даних від кандидата до HR значно оптимізує часові витрати порівняно зі стандартним підбором кандидатів. Він виключає механічне вивчення HR-фахівцем резюме всіх кандидатів, які подали заявки на вакансію, а також зменшує кількість контактів, які не відповідають вакансії. У цій статті ми зосередимося на розробці алгоритму, який зчитуватиме необхідну інформацію з PDF-резюме і конвертуватиме її в JSON-файли.

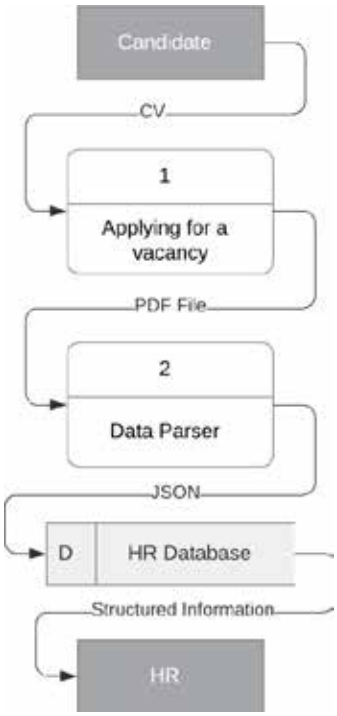
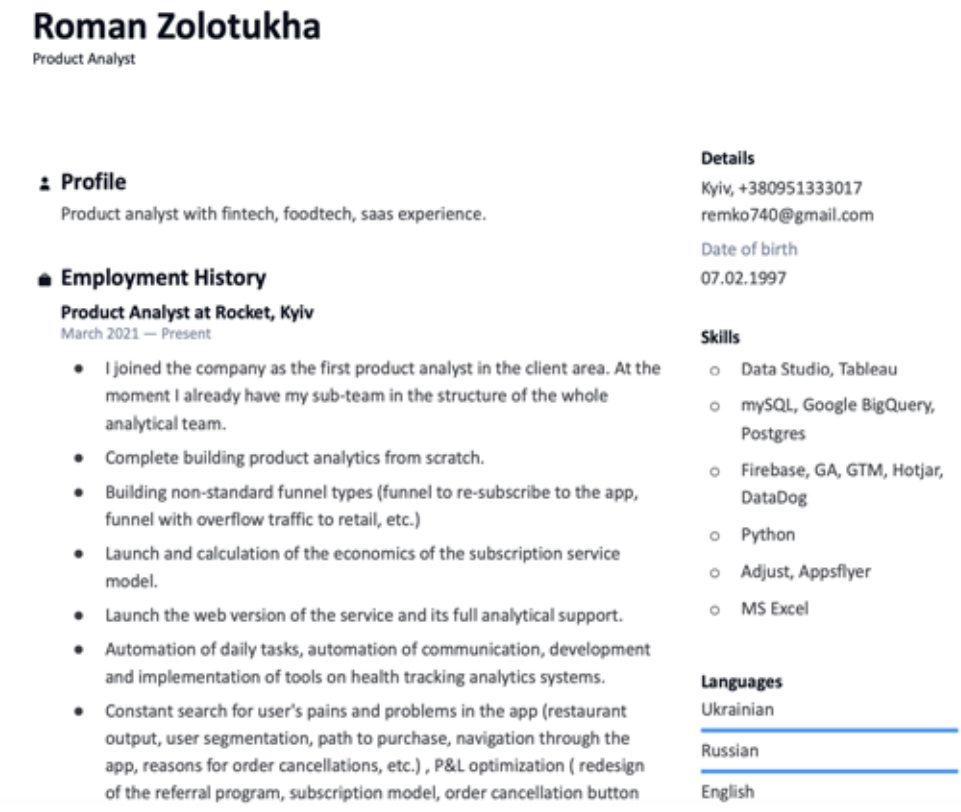


Рис. 3. Схема руху даних від кандидата до HR

Щоб виділити ключові розділи в резюме, ми згенерували 30 резюме, доступних на сайтах-агрегаторах резюме. Базове резюме містить контактну інформацію про кандидата, інформацію про його попередній досвід роботи, місце навчання, перелік технологій, якими володіє кандидат, та рівень володіння іноземними мовами. [7] Базовий вигляд резюме показано на рисунку 4.



Roman Zolotukha

Product Analyst

Profile

Product analyst with fintech, foodtech, saas experience.

Employment History

Product Analyst at Rocket, Kyiv

March 2021 — Present

- I joined the company as the first product analyst in the client area. At the moment I already have my sub-team in the structure of the whole analytical team.
- Complete building product analytics from scratch.
- Building non-standard funnel types (funnel to re-subscribe to the app, funnel with overflow traffic to retail, etc.)
- Launch and calculation of the economics of the subscription service model.
- Launch the web version of the service and its full analytical support.
- Automation of daily tasks, automation of communication, development and implementation of tools on health tracking analytics systems.
- Constant search for user's pains and problems in the app (restaurant output, user segmentation, path to purchase, navigation through the app, reasons for order cancellations, etc.) , P&L optimization (redesign of the referral program, subscription model, order cancellation button

Details

Kyiv, +380951333017

remko740@gmail.com

Date of birth

07.02.1997

Skills

- Data Studio, Tableau
- mySQL, Google BigQuery, Postgres
- Firebase, GA, GTM, Hotjar, DataDog
- Python
- Adjust, Appsflyer
- MS Excel

Languages

Ukrainian

Russian

English

Рис. 4. Базовий приклад вигляду резюме в експерименті

Для розробки такого алгоритму необхідне розуміння наборів сучасних IT навичок, які користуються високим попитом і на які часто звертають увагу особи, що шукають роботу в IT-індустрії [8; 9; 10].

Наша методологія передбачала ретельне вивчення цих наукових ресурсів, в ході якого ми визначили та каталогізували ключові слова та фрази, що вказують на технічні навички, мови програмування, інструменти, фреймворки та інші релевантні IT-навички. Такий системний підхід дозволив нам скласти вичерпний перелік потенційних навичок, що підлягають дослідженню. Проте варто зауважити, що список характеристик не є сталим явищем та змінюється в залежності від динаміки розвитку IT індустрії та появи там нових професій та сфер діяльності. На 2023 рік нами був підготовлений список характеристик навичок кандидата який включає в себе 38590 скілів, які може вказати кандидат у своєму CV.

Для досягнення мети дослідження і розробки алгоритму інструментом реалізації задуму була обрана мова програмування високорівнева інтерпретована мова програмування Python. Ця мова програмування використовується в різних галузях, включаючи веб-розробку, наукові дослідження, штучний інтелект, обробку природної мови, аналіз даних та багато інших. Вона є однією з найпопулярніших мов програмування у світі завдяки своїм перевагам та широкому спектру застосувань. Ми обрали її для розробки алгоритму на основі ряду ключових переваг, які роблять Python відмінним інструментом для вирішення задачі автоматизації підбору кандидатів в IT галузі:

– Популярність та спільнота розробників: Python є однією з найбільш популярних мов програмування у світі, що призводить до наявності великої та активної спільноти розробників. Це означає доступ до безлічі ресурсів, бібліотек та форумів для підтримки та вирішення проблем.

– Простота: Python відомий своєю простотою та читабельністю коду. Це дозволяє розробникам легко створювати та розуміти складні алгоритми, що є важливим для розробки алгоритму розпізнавання тексту.

– Багатофункціональність: Python має багату екосистему бібліотек та модулів для роботи з текстом, обробки PDF-файлів, обчислювального аналізу, та багато іншого. Ця багатофункціональність дозволяє ефективно реалізувати всі етапи алгоритму.

– Портативність: Python підтримується на багатьох платформах, що дозволяє легко переносити алгоритм на різні операційні системи та середовища.

– Велика кількість розширень: Python має багато сторонніх розширень та бібліотек, які спрощують роботу з текстом та PDF-файлами. Наприклад, бібліотеки pdfplumber та spacy дозволяють легко витягти текст та виконати обробку природної мови.

– Активний розвиток: Python постійно розвивається, і нові версії мови регулярно надають нові функції та покращення.

Завдяки цим перевагам мови програмування Python вона відмінно підходить для розробки алгоритму розпізнавання тексту з PDF-резюме для автоматизації підбору кандидатів в IT проекти.

Для реалізації даного алгоритму ми використали універсальні можливості мови програмування Python. Надійна екосистема бібліотек Python полегшила наші зусилля в розборі та вилученні релевантної інформації зі складної структури резюме кандидатів. Щоб розібратися в тонкощах PDF-документів, ми використовували бібліотеку "pdfplumber", яка дозволила нам точно витягувати текстовий контент. Ця бібліотека надає можливість переглядати макет PDF-резюме та виокремлювати необхідні текстові сегменти для подальшого аналізу. Для обробки природної мови та розпізнавання текстових шаблонів ми використовували бібліотеку "spacy". Цей потужний інструмент НЛП дозволив токенізувати, тегувати та аналізувати текст, що дало нам змогу виявити шаблони та сутності, важливі для виокремлення навичок та освіти. Модуль "Matcher" у складі "spacy" допоміг виявити конкретні лінгвістичні патерни, впорядкувавши наш процес визначення ключової інформації. Щоб полегшити організацію та зберігання наших результатів, ми використали модуль "csv" для створення та управління структурованими наборами даних. Бібліотека "pandas" запропонувала нам ефективний засіб для маніпулювання та аналізу цих наборів даних, що дозволило нам отримати уявлення та тенденції з отриманої інформації. Як невід'ємну частину нашої реалізації ми використали можливості вбудованої бібліотеки Python під назвою "json". Ця бібліотека надала нам необхідні інструменти для легкого перетворення складних структур даних у формат JSON (JavaScript Object Notation), що дозволяє безперешкодно серіалізувати та зберігати дані. [11; 12; 13]

Запропонований нами процес перетворення PDF документу, яка містить інформацію про кандидата в JSON файл готовий до внесення в базу даних можна візуалізувати наступним чином:

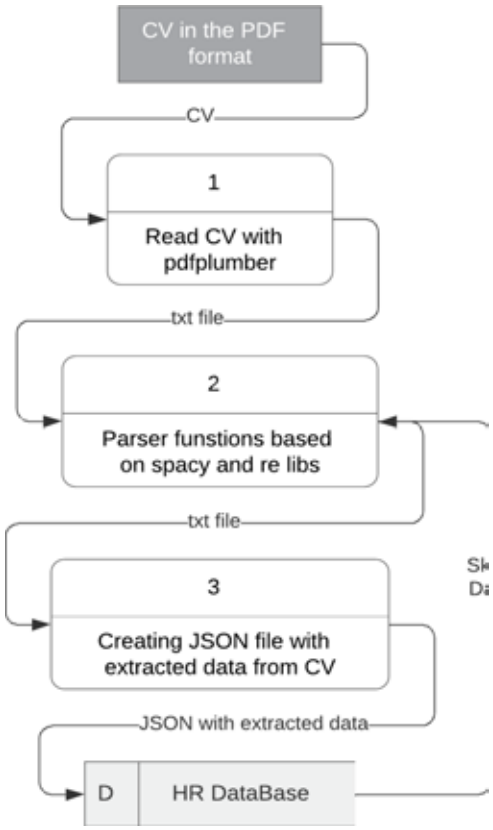


Рис. 5. Покрокова схема алгоритму

```
if __name__ == '__main__':
    resume_text = extracted_text

    name = extract_name(resume_text)
    if name:
        print("Name:", name)
    else:
        print("Name not found")

    contact_number = extract_contact_number_from_resume(resume_text)
    if contact_number:
        print("Contact Number:", contact_number)
    else:
        print("Contact Number not found")

    email = extract_email_from_resume(resume_text)
    if email:
        print("Email:", email)
    else:
        print("Email not found")

    skills_list = csv_data_list
    extracted_skills = extract_skills_from_resume(resume_text, skills_list)
    if extracted_skills:
        print("Skills:", extracted_skills)
    else:
        print("No skills found")

    extracted_education = extract_education_from_resume(resume_text)
    if extracted_education:
        print("Education:", extracted_education)
    else:
        print("No education information found")
```

```
Name: Roman Zolotukha
Contact Number: 380951333017
Email: remko740@gmail.com
Skills: ['computer science', 'mysql', 'health', 'p', 'api', 'mailchimp', 'economics', 'gmail', 'automation', 'pytho
n', 'subscribe', 'analytics', 'google analytics', 'firebase', '2014', 'communication', 'improvement', '.com', 'engl
ish', 'travel', 'british', 'data collection', 'web', 'digital', 'navigation', 'com', 'facebook', 'russian', 'analyt
```

Рис. 6. Частина коду реалізованого алгоритму в середовищі Jupyter Notebook

Для даного дослідження ми використали середовище розробки Jupyter Notebook. У нашому дослідженні ми використали середовище розробки Jupyter Notebook з кількох ключових причин:

– Інтерактивний аналіз даних: Jupyter Notebook дозволяє нам виконувати код блоками і бачити результати негайно після виконання. Це надає можливість проводити ітеративний аналіз даних та швидко реагувати на зміни та експерименти.

– Об'єднання коду та тексту: У Jupyter ми можемо поєднувати код з пояснювальним текстом, що робить наші дослідження більш зрозумілими та легкими для сприйняття.

– Візуалізація результатів: Ми можемо легко вставляти графіку та візуалізації результатів в наші документи, що робить наші дослідження більш інформативними.

Цей JSON-файл є наочним представленням результатів нашого дослідження і полегшує організоване зберігання та обмін відповідною інформацією. JSON-файл має кілька окремих полів, які відображають ключові аспекти профілю кандидата:

– Ім'я: Це поле містить повне ім'я кандидата, що дозволяє його чітко ідентифікувати та посилатися на нього.

– Контактний номер: Поле контактного номера містить наданий кандидатом номер телефону, що забезпечує безперебійний зв'язок та комунікацію.

– Електронна пошта: У полі електронної пошти вказується адреса електронної пошти кандидата, яка слугує основним засобом для листування.

– Навички: У цьому полі ретельно реєструється цілий ряд навичок. Ці навички охоплюють різноманітні технічні навички, мови програмування, інструменти, фреймворки та інші компетенції, пов'язані з IT, якими володіють кандидати.

– Освіта: Сфера освіти містить інформацію про академічні досягнення кандидата.

Результатом роботи нашого алгоритму є ретельно структурований JSON-файл, зображений на рисунку нижче:


```
{
    "Name": "Roman Zolotukha",
    "Contact Number": "380951333017",
    "Email": "remko740@gmail.com",
    "Skills": [
        "computer science",
        "mysql",
        "mailchimp",
        "economics",
        "gmail",
        "automation",
        "python",
        "analytics",
        "google analytics",
        "firebase",
        "data collection",
        "web",
        "digital",
        "facebook",
        "analytical",
        "excel",
        "data studio",
        "dashboards",
        "sql",
        "analyzing data",
        "tableau",
        "mysql",
        "search",
        "models",
        "data",
        "analysis",
        "computer science",
        "ms excel",
        "data collection",
        "output",
        "calculations",
        "hypothesis",
        "analytical support",
        "excel",
        "analytics",
        "api",
        "facebook api",
        "sql",
        "tableau",
        "python",
    ],
    "Education": [
        "PhD in Computer Science"
    ]
}
```


Рис. 7. Приклад JSON-файлу з інформацією з PDF CV

Висновки. Результатом нашого дослідження є розроблений алгоритм на мові програмування Python, який перетворює дані, витягнуті з PDF-резюме кандидатів, в агреговану, готову до використання інформацію. Цей трансформаційний процес усуває необхідність додаткових ручних зусиль, спрощуючи інтеграцію даних про кандидатів у бази даних HR. Використовуючи передові технології, такі як обробка природної мови та аналіз даних, наш алгоритм досягає автоматизації агрегації даних.

Цей алгоритм має величезний потенціал для HR-відділів, оскільки він не тільки прискорює процес рекрутингу, але й зменшує ручну працю, вишукування необхідної інформації серед великого тексту у резюме кандидата. Здатність алгоритму самостійно витягувати, класифікувати та структурувати інформацію про кандидатів дає змогу HR-фахівцям швидко оцінювати придатність кандидатів. Це означає оптимізацію використання ресурсів та більш ефективне прийняття рішень.

Подальшим розвитком дослідження цього алгоритму є розробка інтерфейсу для HR спеціалістів, а також створення та наповнення бази кандидатів та подальшої імплементації цього алгоритму в робоче середовище та застосування на практиці.

Список використаних джерел:

1. Криворучко О. М., Водолажська Т. О. Розробка стратегій управління персоналом методом концептуального абстрагування. *Економіка транспортного комплексу*. 2016. № 28. С. 69–83.

2. Івченко І.Ю., Лінгур Л.М., Філатова Т.В. Моделювання управління кадрами на ІТ-ринку праці. *Вісник Харківського національного університету імені В. Н. Каразіна серія «Економічна»*. 2021. № 101. С. 101-112. URL: https://doi.org/10.26565/2311-2379-2021-101-10

3. Круціцький В. Я., Сугоняк І. І. Оцінка ефективності використання інструментів NLP та систем AI для аналізу рекламних оголошень у системах обміну інтернет-рекламою. *Технічна інженерія*. 2023. № 1(91). С. 161–165. URL: https://doi.org/10.26642/ten-2023-1(91)-161-165

4. Шевцов В. А., Баженов Н. А. Використання НЛП для визначення обсягу оцінки стейкхолдерами імітованих якостей послуг. *Штучний інтелект*. 2010. № 3. С. 161–169.

5. Огляд ІТ-ринку праці, червень 2023. *DOU.ua*. URL: https://dou.ua/lenta/articles/it-job-market-june-2023/

6. Ющенко К.С. Підхід до автоматизації процесу підбору кадрів за допомогою 3D резюме. *Математичні машини і системи*. 2022. № 2. С. 29–39.

7. Eggert M. Perfect CV (Perfect). Random House Books, 2007. 192 p.

8. Martin R. Clean Code: A Handbook of Agile Software. Craftsmanship. – Pearson, 2008. 497 p.

9. Gamma E. Design patterns: Elements of reusable object oriented software. Reading, USA : Addison-Wesley, 1995. 395 p.

10. Peter A. Machine Learning: The Art and Science of Algorithms that Make Sense of Data. Cambridge University Press. 2012. 416 p.

11. McKinney W. Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Jupyter. O'Reilly Media, 2022. 550 p.

12. Muellera A Introduction to Machine Learning with Python: A Guide for Data Scientists. O'Reilly Media, 2016.

13. Yu D., Deng L. Automatic Speech Recognition: A Deep Learning Approach. Springer-Verlag Longon, 2015.

References:

1. Krivoruchko, O. M., & Vodolazhska, T. O. (2016). Rozrobka stratehii upravlinnia personalom metodom kontseptualnoho abstrahuvannia [Development of personnel management strategies by the method of conceptual abstraction]. *Ekonomika transportnogo kompleksa – Economy of the transport complex,* 28, 69–83. [in Ukrainian].

2. Ivchenko, I.Y., Lingur, L.M., Filatova T.V. (2021). Modeliuvannia upravlinnia kadramy na IT-rynku pratsi [Modeling of personnel management in the IT labor market]. *Visnyk Kharkivskoho natsionalnoho universytetu imeni V. N. Karazina seriia «Ekonomichna» – Bulletin of V. N. Karazin Kharkiv National University, Economic Series*, 101, 101–112. Retrieved from https://doi.org/10.26565/2311-2379-2021-101-10 [in Ukrainian].

3. Krutsitsky, V. Y., & Sugonyak, I. I. (2023). Otsinka efektyvnosti vykorystannia instrumentiv NLP ta system AI dlia analizu reklamnykh oholoshen u systemakh obminu internet-reklamoiu [Evaluation of the effectiveness of using NLP tools and AI systems for analyzing advertisements in online advertising exchange systems]. *Tekhnichna inzheneriia – Technical Engineering*, 1(91), 161-165. Retrieved from https://doi.org/10.26642/ten-2023-1(91)-161-165 [in Ukrainian].

4. Shevtsov, V. A., & Bazhenov, N. A. (2010). Vykorystannia NLP dlia vyznachennia obsiahu otsinky steikkholderamy imitovanykh yakostei posluh [Using NLP to determine the extent of stakeholder evaluation of simulated service qualities]. *Shtuchnyi intelekt – Artificial intelligence*, 3, 161-169. [in Ukrainian].

5. Ohliad IT-rynku pratsi, cherven 2023 [IT labor market, June 2023]. Retrieved from https://dou.ua/lenta/articles/it-job-market-june-2023/ [in Ukrainian].

6. Yushchenko, K. S. (2022). Pidkhid do avtomatyzatsii protsesu pidboru kadriv za dopomohoiu 3D reziume [An approach to automating the recruitment process using 3D resumes]. *Matematychni mashyny i systemy – Mathematical machines and systems*, 2, 29-39.

7. Eggert, M. (2007). Perfect CV (Perfect). Random House Books. 192 p.

8. Martin, R. C. (2008). Clean Code: A Handbook of Agile Software Craftsmanship. Prentice Hall. 497 p.

9. Gamma, E. (1995). Design patterns: Elements of reusable object orientated software. Addison-Wesley. 395 p.
10. Peter, A. (2021). Machine Learning: The Art and Science of Algorithms that Make Sense of Data. 416 p.
11. McKinney, W. (2022). Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. 3rd edition. O'Reilly Media. 550 p.
12. Müller, A. C., Guido, S. (2016). Introduction to Machine Learning with Python: A Guide for Data Scientists. 1st edition. O'Reilly Media.
13. Yu, D., Deng, L. (2015). Automatic Speech Recognition: A Deep Learning Approach. Springer-Verlag London. DOI: 10.1007/978-1- 4471-5779-3.

УДК 004.89
DOI https://doi.org/10.32689/maup.it.2023.3.5

Едуард КІНШАКОВ
аспірант кафедри інформаційних технологій, Сумський державний університет, вул. Миколи Сумцова, 2, Суми, Україна, індекс 40007 (edikkinshakov@gmail.com)
ORCID: 0000-0001-7116-7244

Юлія ПАРФЕНЕНКО
кандидат технічних наук, доцент, доцент кафедри інформаційних технологій, Сумський державний університет вул. Миколи Сумцова, 2, Суми, Україна, індекс 40007 (yuliya_p@cs.sumdu.edu.ua)
ORCID: 0000-0003-4377-5132

Eduard KINSHAKOV
Postgraduate Student at Information Technology Department, Sumy State University, 2 Mykoly Sumtsova str., Sumy, Ukraine, postal code 40007 (edikkinshakov @gmail.com)

Yuliia PARFENENKO
Candidate of Technical Sciences, Associate Professor, Senior Lecturer at Information Technology Department, Sumy State University, 2 Mykoly Sumtsova str., Sumy, Ukraine, postal code 40007 (yuliya_p@cs.sumdu.edu.ua)

Бібліографічний опис статті: Кіншаков, Е., Парфененко, Ю. (2023). Застосування алгоритмів сегментації для пошуку контурів захворювання на ділянках шкіри. *Інформаційні технології та суспільство*, 3, 39–46. DOI: https://doi.org/10.32689/maup.it.2023.3.5

Bibliographic description of the article: Kinshakov, E., Parfenenko, Yu. (2023). Zastosuvannia alhorytmiv sehmentatsii dlia poshuku konturiv zakhvoriuvannia na diliankakh shkiry [Application of Segmentation Algorithms for Searching Contours of Disease on Skin Areas]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 39–46. DOI: https://doi.org/10.32689/maup.it.2023.3.5

ЗАСТОСУВАННЯ АЛГОРИТМІВ СЕГМЕНТАЦІЇ ДЛЯ ПОШУКУ КОНТУРІВ ЗАХВОРЮВАННЯ НА ДІЛЯНКАХ ШКІРИ

Анотація. *Стаття присвячена дослідженню застосування сегментації, котра допоможе виявити та виділити локалізацію захворювання на ділянці шкіри. Об'єктом дослідження є підбір оптимального алгоритму сегментації зображення з чітким відокремленням ділянки та контурів хвороби незалежно від її форми. Актуальність дослідження обумовлена тим, що сучасні методи сегментації та локалізації захворювань широко використовуються для покращення точності та чіткості навчання нейронної мережі. Алгоритми дозволяють виявити та зафіксувати саме ту ділянку шкіри, яка потрібна для подачі до нейронної мережі.*

Мета роботи – розробити алгоритм сегментації та пошуку контурів, який зможе виявити та виділити локальну частину хвороби на зображенні шкіри, наданому користувачем. Алгоритм повинен бути точним і ефективним, незалежно від зовнішніх факторів зображення.

У роботі продемонстровано застосування методів сегментації зображень, таких як сегментація за пороговим значенням, алгоритм морфологічної обробки та алгоритм watershed. Для експериментів було використано зображення атипової родимки з набору даних DermNet. Сегментація зображення виконувалась за допомогою бібліотеки Skimage, яка також включає в себе алгоритми пошуку контурів.

За результатами поставлених експериментів, де всі алгоритми отримували одне і теж саме зображення, чіткість виявлення хвороби було продемонстровано за допомогою сегментації за методом watershed, котрий на відміну від інших зміг визначити локалізацію хвороби, чітко відокремити від загального зображення та помітно використати затухання, яке не перешкоджає подальшій колаборації з алгоритмом пошуку контурів.

У результаті дослідження було встановлено, що даний метод є придатним для вирішення задач сегментації та обробки зображень в дерматології. Це пов'язано з тим, що він ефективно виділяє ділянки шкіри, уражені хворобою, і не вступає в конфлікт з алгоритмом локалізації контурів на базі бібліотеки Skimage при стандартних параметрах.

Подальша робота полягає у подачі до загорткових нейронних мереж у тому вигляді, в котрому будуть оброблені зображення, дослідження поведінки алгоритмів мережі при різних зображень та виявлення точності при різних обробках.

Ключові слова: *сегментація, локалізація, watershed, дані, обробка, зображення, морфологічна обробка, порогове значення.*

APPLICATION OF SEGMENTATION ALGORITHMS FOR FINDING DISEASE CONTOURS ON SKIN AREAS

Abstract. *This paper investigates the application of segmentation to identify and highlight the location of a disease on a skin area. The object of the study is to select the optimal image segmentation algorithm with clear separation of the disease area and contours, regardless of its shape. The relevance of the study is due to the fact that modern methods of segmentation and localization of diseases are widely used to improve the accuracy and clarity of training a neural network. The algorithms allow to identify and fix the exact area of the skin that is needed to be fed to the neural network.*

The purpose of the work is to develop an algorithm for segmentation and contour finding that can identify and highlight a local part of a disease on a skin image provided by the user. The algorithm should be accurate and efficient, regardless of the image's external factors.

The paper demonstrates the application of image segmentation methods, such as threshold segmentation, morphological processing algorithm, and watershed algorithm. For experiments, an image of atypical nevus from the DermNet dataset was used. Image segmentation was performed using the Skimage library, which also includes contour finding algorithms.

Based on the results of the set experiments, where all algorithms received the same image, the clarity of disease detection was demonstrated using watershed segmentation. Unlike others, it was able to determine the location of the disease, clearly separate it from the overall image, and significantly use attenuation, which does not harm further collaboration with the contour finding algorithm.

The study found that this method is suitable for solving segmentation and image processing problems in dermatology. This is due to the fact that it effectively highlights areas of the skin affected by the disease and does not conflict with the Skimage library-based contour localization algorithm at standard parameters.

Key words: *segmentation, localization, watershed, data, processing, image, morphological processing, threshold.*

Постановка проблеми. У сучасну епоху широке поширення хвороб шкіри суттєво впливає на фізичний стан та майбутнє здоров'я людей. Багато населених пунктів України мають обмеження у доступі до невідкладної медичної допомоги та амбулаторій, де доступна лише базова медична допомога, і вони не здатні приймати спеціалістів, зокрема, дерматологів і фахівців з інших вузьких галузей. Це ставить питання про важливість впровадження сучасних технологій дистанційної діагностики хвороб шкіри, що стає надзвичайно актуальним у наш часі [1].

У той же час кількість осіб, що користуються смартфонами та мають доступ до Інтернету, стабільно росте. Світ спостерігає зростаючий інтерес до телемедицини, яка надає можливість здійснювати дистанційні консультації з лікарем, надавати необхідну медичну допомогу в невідкладних ситуаціях та планувати подальше лікування. Україна розвиває телемедицину, проте вона є відносно витратною, що обумовлено, зокрема, відсутністю відповідного інформаційного забезпечення для прийняття рішень [2; 3].

Це вказує на необхідність створення інформаційної системи, яка допоможе людям реагувати на різні симптоми шкірних захворювань і визначати ступінь їх серйозності, запобігаючи погіршенню стану пацієнтів.

Одним із ключових завдань досліджень є розробка інформаційно-інтелектуальної системи, котра буде розпізнавати та класифікувати хворобу за зображеннями ділянок шкіри не найвищої якості. Задача даної роботи полягає в тому, щоб провести попередні експерименти над зображеннями, котрі будуть подаватися до нейронної мережі, з метою підвищення достовірності розпізнавання шкірних захворювань. А саме поставити експерименти над алгоритмами сегментації та пошуку контурів, котрі будуть виявляти та локалізувати хворобу. Таким чином якість точності нейронної мережі перейде на більш високий рівень.

З огляду на широке використання методів сегментації у сфері медицини, необхідно провести наукові дослідження для визначення ефективності таких методів.

Аналіз досліджень і публікацій. Сегментації зображень з проявами хвороб шкіри приділено багато уваги авторів. Більшість авторів використовують сегментацію не як метод для подальшої локалізації контурів хвороби, а як визначення покриття хвороби. В роботі [4] автор присвячує увагу методам сегментації котрий побудований на основі кластеризації кольорів.

Це не поганий метод, щоб викреслити із дослідження максимально попередню обробку зображень. Але вважається, що кластеризація може мати велику похибку у випадку великого радіусу хвороби. Дане дослідження [5] представляє алгоритм для вирішення проблеми глобальної оптимізації сегментації раку шкіри меланоми. Алгоритм заснований на згладжуванні допоміжної функції, яка будується за допомогою відомого локального мінімізатора та згладжується за допомогою кривих Безьє.

Результати показують, що запропонований алгоритм демонструє високу точність, чутливість і специфічність порівняно з іншими методами. Основною метою дослідження [6] є пошук оптимального методу сегментації зображень уражень шкіри. У цьому дослідницькому документі розроблено модель, яка охоплює такі градації, першою є попередня обробка для зменшення небажаних частин, таких як волосся, освітлення чи багато іншого, за допомогою покращеної техніки з використанням граничних

і морфологічних операцій для досягнення вищої точності інтелектуальної системи. Але, виходячи з результатів, автори помітили, що запропонована модель дає середнє значення точності величезної кількості зображень раку. У цих роботах автори [7; 8] представили метод виявлення раку шкіри на основі методу **watershed** з контрольованим маркером.

Запропонована техніка була застосована та випробувана на кількох зображеннях різних типів раку шкіри, які були зібрані з Інтернету, а також із набору даних Kaggle. Щоб оцінити цінність досягнутих результатів, автори використали кілька показників оцінки, як чутливість, специфічність, а також подібність Жаккара, які показали хороші та задовільні результати.

Виклад основного матеріалу. Дані для дослідження було взято з всесвітнього дерматологічного ресурсу DermNet. Надана ресурсом база зображень має 22 класи хвороби. Всі зображення мають формат jpg. Але для зрозумілої демонстрації результатів було взято атиповий невус. На даному зображені можна чітко зрозуміти та побачити ефективність сегментації та потім, як вона допоможе алгоритму пошуку контурів знайти потрібну ділянку.

В попередніх публікаціях вже було продемонстровано метод кластеризації для обробки зображень, головна задача котрого повинна була підняти якість прогнозування нейронної мережі. Покращення були зафіксовані, але не значні. Тому було прийняте рішення розширити методи сегментації та на їх фоні зробити пошук контурів. Дана задача дозволить виявити необхідну нам ділянку шкіри на котрій є локалізація хвороби. Це допоможе відсіяти непотрібні шуми та зменшити розмірність, таким чином якість нейронної мережі збільшиться.

У даній роботі використовуються три алгоритми сегментації, одним з них є метод граничного значення або, як ще його називають методом Оцу (Otsu) [9]. Використання головної ідеї методу Оцу полягає в тому, щоб обрати поріг сегментації так, щоб дисперсія (внутрішнього розсіювання) в класі об'єктів була мінімальною, а розсіювання між класами (між об'єктами і фоном) була максимальною.

Це робиться шляхом розгладжування всіх можливих граничних значень і обчислень внутрішнього і міжкласового розсіяння для кожного порогу.

Міжкласове розсіювання – це міра розсіювання всередині кожного класу (об'єктів та фону). Вона вказує, наскільки значення пікселів в межах кожного класу розподілені відносно середнього значення цього класу. Функція для внутрішнього розсіяння (intra-class variance) обчислюється наступним чином [10] :

$$\sigma_{\omega}^2(t) = \omega_0(t) * \sigma_0^2(t) + \omega_1(t) * \sigma_1^2(t) \quad (1)$$

де:

σ_{ω}^2 – внутрішнє розсіяння для порогу t;

$\omega_0(t)$ – вага першого класу (фону) для порогу t;

σ_0^2 – дисперсія першого класу (фону) для порогу t;

$\omega_1(t)$ – вага другого класу (об'єктів) для порогу t.

$\sigma_1^2(t)$ – дисперсія другого класу для порогу t.

Результати методу загальної граничної сегментації зображено на (рис. 1), де можемо побачити область родимки, але можна спостерігати нестандартне розсіювання.

Рис. 1. Сегментація за допомогою граничного значення

Далі переходимо до наступного методу морфологічної обробки. Морфологічна обробка зображення включає в себе групу методів обробки зображень, які базуються на математичних операціях, що використовують структурні елементи (ядро) для зміни форми та структури об'єктів на зображенні [11; 12]. Дані операції дозволяють виокремлювати особливості об'єктів, виділяти їх контури, видаляти шум, об'єднувати та ділити об'єкти, покращувати якість.

Перелік операцій допоможе вирішити проблему з шумами, оскільки попередній метод може виявити та прибрати шуми та розсіювання, котрі можуть впливати на якість сегментації.

Математично морфологічні операції використовують операцію морфологічного відображення (morphological transformation) над зображенням A з використанням структурного елементу B [13]. Операції можна виразити наступним чином:

$$(A \oplus B)(x, y) = \bigcup^*_{(i,j)\in B} A(x+i, y+j) \quad (2)$$

де:

A – це початкове зображення, над яким виконується операція розширення;

B – структурний елемент (іноді називається ядром), який визначає форму і розмір розширення;

(i,j) – координати пікселів в структурному елементі B.

Отже, результати методу морфологічної обробки зображено на (рис. 2), де чітко можемо побачити область родимки, підкреслені контури, а також деяке подавлення шумів.

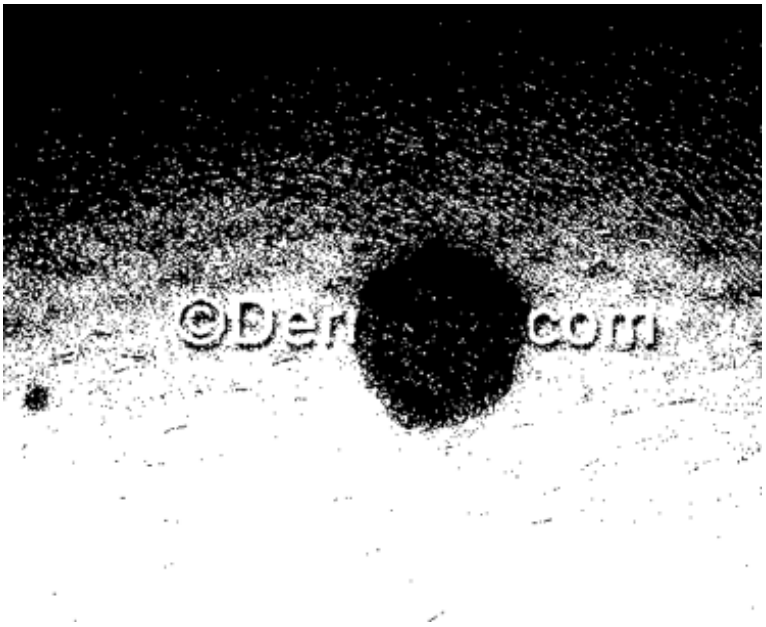


Рис. 2. Сегментація за допомогою морфологічної обробки

Найрезультативніший алгоритм Watershed зміг реалізувати поставлену задачу, завдяки котрому було досягнуто результату пошуку контурів. Будь-яке відтінкове зображення може бути розглянуто як топографічна поверхня, де високий інтенсивний світловий сигнал вказує на вершини, тоді як низький інтенсивний світловий сигнал вказує на інші частини зображення. Спочатку заповнюється кожна окрема частина (локальні мінімуми) різнокольоровими мітками [14].

Поступово піднімаючи рівень мітки, залежно від навколишніх градієнтів, мітки з різних частин, очевидно, різнокольорова, починає об'єднуватися. Щоб цього уникнути, створюються бар'єри на місцях об'єднання міток. Потім продовжується робота із наповненням та створення бар'єрів до тих пір, поки всі вершини опиняться під міткою. Потім бар'єри, які були створені, надають результат сегментації. Алгоритм watershed важко виразити однією математичною формулою, оскільки він базується на обчисленнях на основі географічних аналогій та дискретних зображень. Проте можна продемонструвати загальну ідею та кроки алгоритму [15] :

$$I = \sqrt{(I_x)^2 + (I_y)^2} \quad (3)$$

де:

I_x – похідна зображення за напрямом x;

I – зображення;

I_y – похідна зображення за напрямом y.

Результати методу (watershed) зображено на рис. 3, де чітко можемо побачити область родимки, підкреслені контури більш чітко виділяються, ніж при використанні попереднього методу.

Рис. 3. Сегментація за допомогою watershed алгоритму

Після того, як було поставлено експерименти на зображенні за допомогою сегментації, наступним кроком є побудова алгоритму, котрий буде включати в себе один із методів сегментації для пошуку контурів захворювання. Алгоритм пошуку контурів сприймає всі зображення в чорно-білому вигляді, тобто бінарному, а повертає вже у кольорі з локалізацією. Для бінарного зображення $|(x,y)|$, де $|(x,y)|=0$ або $|(x,y)|$ (де 0 представляє чорний колір, а 1 – білий колір), і порогового значення T, процес пошуку контурів здійснюється наступним чином. Першочергово ініціюється порожній список контурів C [16]. Для кожного пікселя $|(x,y)|$ відбуваються наступні операції:

1. Формування нового контуру, який представляє границю об'єкта.
2. Створюється порожній список точок для поточного контуру, який позначається як P.
3. Будується функція обходу для пошуку з'єднаних білих пікселів, починаючи з пікселя (x, y).
4. Здійснюється обхід сусідніх пікселів у пошуку з'єднаних компонентів, який продовжується, поки не буде оброблено кожен з'єднаний білий піксель об'єкта.
5. Після завершення формування контуру список точок P додається до списку контурів C.

Цей процес пошуку контурів дозволяє виділити границі об'єктів на бінарному зображенні та представити їх у вигляді наборів точок (координат x, y), які формують контури [17]. Результати пошуку контурів після сегментації методом watershed зображено на (рис. 4). Результати алгоритмів граничного значення та морфологічної обробки показали візуально однаковий результат.

Результати пошуку контурів зображення за цим методом показано на рис. 5-6.

Висновки. У результаті проведеного дослідження встановлено, що кожний метод сегментації має свої переваги та недоліки. Це залежить від поставленої задачі та самого зображення. У певних випадках морфологічна обробка та метод граничного значення можуть бути ефективними, але на окремих ділянках сегментації. Як можна побачити з результатів застосування цих методів, шуми так і не були подавлені, більше того можна побачити затухання на окремих ділянках зображення. Найкращий результат продемонстрував метод watershed, котрий зміг чітко сегментувати зображення, знайти градієнти і в деяких випадках подавити шуми. Контур родимки після методу watershed було виділено чітко. Наступним етапом дослідження буде розробка інформаційної системи діагностування шкірних

захворювань з використанням нейронних мереж. На вхід нейронної мережі буде здійснюватися подача зображення після обробки методом watershed та знаходження контурів. Наступними діями дослідження буде перевірка ефективності нейромережевої моделі діагностування шкірних захворювань на оброблених зображеннях.

Рис. 4. Результат пошуку контурів після сегментації методом watershed

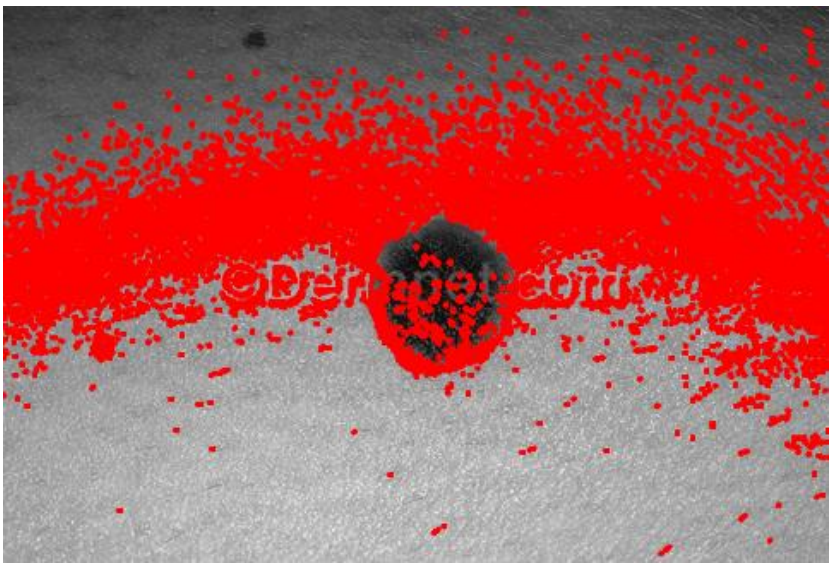


Рис. 5. Результат пошуку контурів після морфологічної обробки та граничного значення

Список використаних джерел:

1. Haleem A., Javaid M., Singh R. P., and. Suman R. «Telemedicine for healthcare: Capabilities, features, barriers, and applications». *Sensors International*. Vol. 2. 2021. DOI: 10.1016/j.sintl.2021.100117.

2. Zhang C., «Smartphones and telemedicine for older people in China: Opportunities and challenges». *Digital Health*. Vol. 8. 2022. DOI: 10.1177/20552076221133695.

3. Allaert F.A., Legrand L., Abdoul Carime N., and Quantin C. «Will applications on smartphones allow a generalization of telemedicine?». *BMC Medical Informatics and Decision Making*, Vol. 20. № 1. 2020. DOI: 10.1186/s12911-020-1036-0.

4. Joseph S. and Olugbara O. O., «Preprocessing Effects on Performance of Skin Lesion Saliency Segmentation». *Diagnostics*, Vol. 12, № 2. Feb. 2022, DOI: 10.3390/diagnostics12020344.

5. Masoud Abdulhamid I.A., Sahiner A., and Rahebi J. «New Auxiliary Function with Properties in Nonsmooth Global Optimization for Melanoma Skin Cancer Segmentation». *Biomed Res Int*. Vol. 2020. DOI: 10.1155/2020/534592

6. Garg S. and Jindal B. «Skin Lesion Segmentation in Dermoscopy Imagery». *International Arab Journal of Information Technology*. Vol. 19. № 1. P. 29–37. Jan. 2022. DOI: 10.34028/iajit/19/1/4.

7. Wang H. *et al.* «Watershed segmentation of dermoscopy images using a watershed technique». *Skin Research and Technology*. 2010. Vol. 16. № 3. P. 378–384. DOI: 10.1111/j.1600-0846.2010.00445.x.

8. Moussaoui H., N. El Akkad, and Benslimane M. «A hybrid skin lesions segmentation approach based on image processing methods». *Statistics, Optimization and Information Computing*. Vol. 11. № 1. P. 95–105. DOI: 10.19139/soic-2310-5070-1549.

9. Shahabi F., Poorahangaryan F., Edalatpanah S. A., and Beheshti H. «A Multilevel Image Thresholding Approach Based on Crow Search Algorithm and Otsu Method». *Int J Comput Intell Appl*. 2020. Vol. 19. № 2. DOI: 10.1142/S1469026820500157.

10. Pitoy P. A. and Suputra I. P. G. H. «Dermoscopy Image Segmentation in Melanoma Skin Cancer using Otsu Thresholding Method». *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana)*. Vol. 9. № 3. P. 397. DOI: 10.24843/jlk.2021.v09.i03.p11.

11. Lumini A., L. Nanni A., Codogno A., and Berno F. «Learning morphological operators for skin detection». *Journal of Artificial Intelligence and Systems*. 2019. Vol. 1. № 1. DOI: 10.33969/ais.2019.11004.

12. Rew J., Kim H., and Hwang E. «Hybrid segmentation scheme for skin features extraction using dermoscopy images». *Computers, Materials and Continua*. 2021. Vol. 69. № 1. DOI: 10.32604/cmc.2021.017892.

13. Prabha Devi D. and Iniya Shree S. «Recognition and investigation of skin cancer using morphological operations». *International Journal of Recent Technology and Engineering*. 2019. Vol. 7. № 4.

14. Wang H. *et al.* «Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images». *Computerized Medical Imaging and Graphics*. 2011. Vol. 35. № 2. DOI: 10.1016/j.compmedimag.2010.09.006.

15. Das A. and Ghoshal D. «Human Skin Region Segmentation Based on Chrominance Component Using Modified Watershed Algorithm». *Procedia Computer Science*. 2016. DOI: 10.1016/j.procs.2016.06.072.

16. Shalu and Kamboj A. «A Color-Based Approach for Melanoma Skin Cancer Detection». *ICSCCC 2018 – 1st International Conference on Secure Cyber Computing and Communications*. 2018. DOI: 10.1109/ICSCCC.2018.8703309.

17. Ashour A. S., Nagieb R. M., El-Khobby H. A., Abd Elnaby M. M., and Dey N. «Genetic algorithm-based initial contour optimization for skin lesion border detection» *Multimed Tools Appl*. Vol. 80. № 2. 2021. DOI: 10.1007/s11042-020-09792-8.

References:

1. Haleem, A., Javaid, M., Singh, R. P., & Suman, R. (2021). Telemedicine for healthcare: Capabilities, features, barriers, and applications. *Sensors International*, *Vol. 2.* Retrieved from https://doi.org/ 10.1016/j.sintl.2021.100117

2. Zhang, C. (2022). Smartphones and telemedicine for older people in China: Opportunities and challenges. *Digital Health, Vol. 8.* Retrieved from https://doi.org/10.1177/20552076221133695

3. Allaert, F. A., Legrand, L., Abdoul Carime, N., & Quantin, C. (2020). Will applications on smartphones allow a generalization of telemedicine? *BMC Medical Informatics and Decision Making*, *Vol. 20(1).* Retrieved from https://doi.org/10.1186/s12911-020-1036-0

4. Joseph, S., & Olugbara, O. O. (2022). Preprocessing Effects on Performance of Skin Lesion Saliency Segmentation. *Diagnostics*, *12(2).* Retrieved from https://doi.org/10.3390/diagnostics12020344

5. Masoud Abdulhamid, I. A., Sahiner, A., & Rahebi, J. (2020). New Auxiliary Function with Properties in Nonsmooth Global Optimization for Melanoma Skin Cancer Segmentation. *BioMed Research International*. Retrieved from https://doi.org/10.1155/2020/5345923

6. Garg, S., & Jindal, B. (2022). Skin Lesion Segmentation in Dermoscopy Imagery. *International Arab Journal of Information Technology, 19(1), 29–37.* Retrieved from https://doi.org/10.34028/iajit/19/1/4

7. Wang, H., Moss, R. H., Chen, X., Stanley, R. J., Stoecker, W. V., Celebi, M. E., Malters, J. M., Grichnik, J. M., Marghoob, A. A., Rabinovitz, H. S., Menzies, S. W., & Szalapski, T. M. (2011). Modified watershed technique and post-processing for segmentation of skin lesions in dermoscopy images. *Computerized Medical Imaging and Graphics, 35(2).* Retrieved from https://doi.org/10.1016/j.compmedimag.2010.09.006

8. Moussaoui, H., El Akkad, N., & Benslimane, M. (2023). A hybrid skin lesions segmentation approach based on image processing methods. *Statistics, Optimization and Information Computing, 11(1), 95–105*. Retrieved from https://doi.org/10.19139/soic-2310-5070-1549

9. Shahabi, F., Poorahangaryan, F., Edalatpanah, S. A., & Beheshti, H. (2020). A Multilevel Image Thresholding Approach Based on Crow Search Algorithm and Otsu Method. *International Journal of Computational Intelligence and Applications, 19(2).* Retrieved from https://doi.org/10.1142/S1469026820 500157

10. Pitoy, P. A., & Suputra, I. P. G. H. (2021). Dermoscopy Image Segmentation in Melanoma Skin Cancer using Otsu Thresholding Method. *JELIKU (Jurnal Elektronik Ilmu Komputer Udayana), 9(3), 397.* Retrieved from https://doi.org/10.24843/jlk.2021.v09.i03.p11

11. Lumini, A., Nanni, L., Codogno, A., & Berno, F. (2019). Learning morphological operators for skin detection. *Journal of Artificial Intelligence and Systems, 1(1).* Retrieved from https://doi.org/10.33969/ais.2019.11004

12. Rew, J., Kim, H., & Hwang, E. (2021). Hybrid segmentation scheme for skin features extraction using dermoscopy images. *Computers, Materials and Continua*, 69(1). Retrieved from https://doi.org/10.32604/cmc.2021.017892

13. Prabha Devi, D., & Iniya Shree, S. (2019). Recognition and investigation of skin cancer using morphological operations. *International Journal of Recent Technology and Engineering, 7(4).*

14. Wang, H., Moss, R. H., Chen, X., Stanley, R. J., Stoecker, W. V., Celebi, M. E., Malters, J. M., Grichnik, J. M., Marghoob, A. A., Rabinovitz, H. S., Menzies, S. W., & Szalapski, T. M. (2011). Modified watershed technique and post-processing for

segmentation of skin lesions in dermoscopy images. *Computerized Medical Imaging and Graphics, 35(2).* Retrieved from https://doi.org/10.1016/j.compmedimag.2010.09.006

15. Das, A., & Ghoshal, D. (2016). Human Skin Region Segmentation Based on Chrominance Component Using Modified Watershed Algorithm. *Procedia Computer Science, 89.* Retrieved from https://doi.org/10.1016/j.procs. 2016.06.072

16. Shalu, & Kamboj, A. (2018). A Color-Based Approach for Melanoma Skin Cancer Detection. *ICSCCC 2018 – 1st International Conference on Secure Cyber Computing and Communications*. Retrieved from https://doi.org/10.1109/ICSCCC.2018.8703309

17. Ashour, A. S., Nagieb, R. M., El-Khobby, H. A., Abd Elnaby, M. M., & Dey, N. (2021). Genetic algorithm-based initial contour optimization for skin lesion border detection. *Multimedia Tools and Applications*, *80(2)*. https://doi.org/10.1007/s11042-020-09792-8

УДК 004.896
DOI https://doi.org/10.32689/maup.it.2023.3.6

Павло КОЗОЛУП
аспірант кафедри комп'ютерних наук, Сумський державний університет, Суми, Україна, індекс (pavlo.kozolup@student.sumdu.edu.ua)

ORCID: 0009-0000-1303-3424

Володимир ЛЮБЧАК
кандидат фізико-математичних наук, доцент, завідувач кафедри кібербезпеки, Сумський державний університет, Суми, Україна, індекс (v.liubchak@dcs.sumdu.edu.ua)

ORCID: 0000-0002-7335-6716

Pavlo KOZOLUP
Postgraduate Student at Department of Computer Science, Sumy State University, Sumy, Ukraine, postal code (pavlo.kozolup@student.sumdu.edu.ua)

Volodymyr LIUBCHAK
PhD in Physics and Mathematics, Associate Professor, Head of the Cybersecurity Department, Sumy State University, Sumy, Ukraine, postal code (v.liubchak@dcs.sumdu.edu.ua)

Бібліографічний опис статті: Козолуп П., Любчак В. (2023). Огляд методів та інструментів для розробки інформаційного сервісу обліку особистих активів. *Інформаційні технології та суспільство*, 3, 47–53. DOI: https://doi.org/10.32689/maup.it.2023.3.6

Bibliographic description of the article: Kozolup P., Liubchak V. (2023). Ohliad metodiv ta instrumentiv dlia rozrobky informatsiinoho servisu obliku osobystykh aktyviv [Review of Methods and Tools for Developing an Information Service for Personal Asset Management]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 47–53. DOI: https://doi.org/10.32689/maup.it.2023.3.6

ОГЛЯД МЕТОДІВ ТА ІНСТРУМЕНТІВ ДЛЯ РОЗРОБКИ ІНФОРМАЦІЙНОГО СЕРВІСУ ОБЛІКУ ОСОБИСТИХ АКТИВІВ

Анотація. *Розробка функціональної моделі інформаційного сервісу та створення додатку-помічника, призначеного для задоволення персональних потреб користувачів у керуванні особистими запасами, є науково актуальним завданням, що відображає сучасні тенденції в області інформаційних технологій та управління. Ця проблематика допускає проведення досліджень, спрямованих на вдосконалення методів розробки алгоритмів управління запасами, прогнозування потреб та вибору оптимальних стратегій для керування особистими запасами користувачів.*

У цьому контексті особлива увага приділяється аналізу та розробці різноманітних алгоритмів управління запасами для особистого використання. Управління особистими запасами є важливою складовою для забезпечення ефективності володіння та використання особистих ресурсів, включаючи товари і ресурси в побуті. В даному контексті досліджуються техніки, що базуються на математичних моделях, статистичних методах та оптимізаційних алгоритмах з метою підвищення раціонального використання та оптимізації особистих запасів.

Крім того, наукова спільнота активно вивчає методи прогнозування особистих потреб і ресурсів. Це важливий компонент управління особистими запасами, оскільки точне прогнозування допомагає уникнути надмірного або недостатнього запасу товарів для особистого використання. Тут використовуються різноманітні аналітичні методи, включаючи методи машинного навчання та статистичні моделі.

Окремою частиною наукового дослідження є вибір оптимальних стратегій управління особистими запасами. Це охоплює визначення оптимальних підходів та методів для керування запасами на особистому рівні, враховуючи індивідуальні особливості користувачів.

Невід'ємною частиною досліджень є вибір інструментів і середовища для розробки додатку-помічника, спрямованого на підтримку користувачів у керуванні особистими запасами. Це включає в себе оцінку якості, масштабованості та зручності інструментів з метою створення зручного та ефективного інформаційного сервісу для користувачів.

У висновку, цей науковий підхід ставить перед собою завдання створення інноваційного додатку-помічника, спрямованого на задоволення потреб користувачів у керуванні особистими запасами. Результати дослідження та розробки такого додатку можуть значно полегшити процеси управління особистими запасами та підвищити ефективність управління особистими ресурсами.

Ключові слова: *інформаційний сервіс; додаток-помічник; прогнозування потреб; управління запасами; системи прийняття рішень; функціональна модель сервісу.*

REVIEW OF METHODS AND TOOLS FOR DEVELOPING AN INFORMATION SERVICE FOR PERSONAL ASSET MANAGEMENT

Abstract. *The development of a functional model of an information service and the creation of an assistant application designed to meet the personal needs of users in managing their personal stocks is a scientifically relevant task that reflects modern trends in the field of information technology and management. This issue allows for research aimed at improving methods of developing inventory management algorithms, forecasting needs, and selecting optimal strategies for managing users' personal stocks.*

In this context, special attention is paid to the analysis and development of various inventory management algorithms for personal use. Managing personal stocks is an important component to ensure the efficiency of owning and utilizing personal resources, including goods and household resources. In this context, techniques based on mathematical models, statistical methods, and optimization algorithms are investigated to increase the rational use and optimization of personal stocks.

Additionally, the scientific community actively studies methods for forecasting personal needs and resources. This is an important component of managing personal stocks, as accurate forecasting helps avoid excessive or insufficient stocks of goods for personal use. Various analytical methods, including time series analysis, machine learning methods, and statistical models, are used in this regard.

A separate part of the research focuses on selecting optimal strategies for managing personal stocks. This involves determining optimal approaches and methods for inventory management at the personal level, taking into account individual user characteristics.

An integral part of the research is the selection of tools and environments for developing an assistant application aimed at supporting users in managing their personal stocks. This includes evaluating the quality, scalability, and convenience of tools to create a convenient and efficient information service for users.

In conclusion, this scientific approach aims to create an innovative assistant application aimed at satisfying the needs of users in managing personal stocks. The results of research and development of such an application can significantly simplify the processes of managing personal stocks and enhance the efficiency of managing personal resources.

Key words: *information service; assistant application; demand forecasting; inventory management; decision support systems; functional service model.*

Актуальність теми дослідження. Інформаційні системи, які забезпечують автоматизований облік матеріальних, фінансових та інших активів, є широко поширеними та традиційними у великих підприємствах. Ці системи призначені для обслуговування процесів закупівель, складського господарства, обліку нерухомості, фінансових операцій та багатьох інших. До таких систем можна віднести, наприклад, системи MRP (Material Requirements Planning) та ERP (Enterprise Resource Planning) [1].

Аналогічні задачі та потреби є характерними не лише для великих підприємств, а й для малих приватних підприємств, а також у повсякденній діяльності особистості. Існує нагальна потреба в додатках, які б допомагали користувачам з обліком персональних товарів. З поняттям "особисті товари/активи" зазвичай пов'язують товари, які призначені для особистого використання. Це можуть бути продукти харчування, одяг, косметика, електроніка, книги, нематеріальні (інформаційні) продукти та інші товари, що задовольняють особисті потреби людини. В нагоді був би сервіс-помічник для спрощення закупівлі побутових товарів. Такий помічник (програма) повинен відслідковувати регулярність закупівель, яка існує в нашому повсякденному житті.

Деякі з аналогічних сервісів вже використовуються. Наприклад, додаток Fishka від ОККО дозволяє користувачам відстежувати інформацію про покупки пального, включаючи дату, обсяг та інші деталі. Цей додаток спрямований на допомогу бізнесу в продажу супутніх товарів, крім пального, і на стимулювання користувачів до покупок. Крім того, досить поширені додатки, такі як банківський додаток UKRSIB ONLINE [2], допомагають користувачам відстежувати свої фінанси.

З урахуванням сучасних потреб та існуючих пропозицій ми плануємо проектування та створення персонального додатку-помічника з метою спрощення процесу пошуку, покупки повсякденних товарів та зменшення необхідності користувача слідкувати за цим.

Постановка проблеми. Завдання полягає у розробці алгоритмів для автоматизації регулярних процесів обліку, пошуку та закупівлі товарів, а також їх програмної реалізації. Планується, що сервіс буде включати інструмент для нагадування та повідомлення користувачеві про необхідність купівлі певних товарів у відповідний час. Для досягнення цілей сервісу необхідно реалізувати наступні функції:

1) Прогнозування потреб: Необхідно прогнозувати час та обсяг замовлення відповідних товарів. Метод та алгоритм вирішення цієї задачі повинні бути простими та ефективними, не перевантажувати систему. Розрахунок потреби відбуватиметься щоденно після актуалізації залишків.

2) Оптимальний вибір: Сервіс повинен застосовувати відповідні методи оптимізації для формування оптимального плану закупівель. Вибір може базуватися на найнижчих цінах, вищій якості товарів або на попередніх покупках користувача. Користувач має мати можливість налаштувати свої вподобання в додатку.

3) Управління запасами: Користувач задає мінімальну та максимальну кількість товарів, яка йому необхідна. Сервіс аналізує щоденні потреби та пропонує оптимальну кількість товарів для закупівлі. Для розрахунку оптимальної кількості використовуються динамічні дані про залишки та замовлення на протязі кількох циклів закупівлі та використання.

4) Прийняття рішень: Сервіс приймає рішення про створення нового замовлення на основі інформації про залишки товару, період використання та побажання користувача. Наприклад, користувач може відмовитись від замовлення певного товару.

Для вирішення цих завдань використовуються методи дослідження, такі як спостереження, аналіз даних та моделювання. Планується отримати відгуки користувачів та зібрати статистичні дані, які допоможуть розуміти їхні потреби. Моделювання допоможе створити математичну модель функціонування сервісу з використанням методів підтримки прийняття рішень. Методи підтримки прийняття рішень дозволяють приймати обґрунтовані рішення в умовах невизначеності, складності та багатокритеріальності. Для моделювання функціонування сервісу та застосування методів підтримки прийняття рішень можуть бути використані такі підходи:

1) Аналітична ієрархічна процедура (Analytic Hierarchy Process, AHP): Використовується для розрахунку вагових коефіцієнтів критеріїв та альтернатив для прийняття обґрунтованих рішень на основі їх значущості та взаємного порівняння [3].

2) Аналітична мережа процесів (Analytic Network Process, ANP): Розширена версія AHP, яка враховує зв'язки та взаємозалежності між критеріями та альтернативами [4].

3) Мультикритеріальний аналіз (Multi-Criteria Decision Analysis, MCDA): Включає набір методів для порівняння альтернатив на основі різних критеріїв та їх вагомості [5].

З огляду на існуючи потреби та наявні пропозиції, пропонується проєктування та створення персонального додатку – помічника. Такий додаток має на меті спростити процес пошуку та купівлі повсякдених товарів. Також буде використаний інструмент для нагадування, нотифікації користувачу. Програма надішле повідомлення з інформацією: коли, та в якому об’ємі необхідно зробити покупку товару. Помічник має нагадувати користувачу про потрібні товари, які раніше купував користувач, у відповідний час. Додаток повинен бути зі спрощеним функціоналом, та мінімальним набором елементів. Це надасть можливість користуватись додатком більшій аудиторії. Додаток має бути «не нав’язливим» в плані нагадувань. Додаток має містити комфортну графічну та текстову складову. Схематично, модель процесу обробки даних можно зобразити таким чином:

Рис. 1. Схематична модель процесу обробки даних

Виділення недосліджених частин загальної проблеми. На сьогоднішній день немає подібних програм, які би володіли схожим функціоналом та спрямованістю досліджень. В цілому, функціональність, яка б допомагала користувачам у повсякденному житті, є досить обмеженою. В основному, існують програми, що охоплюють такі напрямки, як спорт, банкінг, розваги та замовлення послуг. Проте, саме створення додатку-помічника збільшує кількість програм, спрямованих на спрощення та покращення повсякденного життя звичайних людей. Розробка такого додатку має великий потенціал для інновацій. Наприклад, можливість автоматичного замовлення товарів, що базується на власних уподобаннях та попередніх покупках користувача, може значно спростити процес шопінгу та зекономити його час. Додаток також може забезпечити рекомендації щодо оптимальних продуктів або послуг на основі аналізу великої кількості даних.

Метою роботи є аналіз матеріалів, які будуть використовуватися при розробці додатку-помічника, а також аналіз доступних та якісних інструментів для створення алгоритмів автоматичного замовлення товарів. Описаний алгоритм передбачає використання в додатку-помічнику, який збиратиме дані для перерахунку та удосконалення розрахунків потреб користувача. Для поліпшення функціональності будуть використані принципи теорії прийняття рішень.

Виклад основного матеріалу

1. Інформаційний огляд результатів досліджень та розробок

Стаття має оглядовий характер, тому надалі ми розглядаємо наукові публікації та інструменти, які корисні та необхідні для вирішення завдань роботи.

Питання управління матеріальними запасами та моделі оптимального визначення обсягу активів для зберігання та поповнення розглядалися в роботах [6], [7].

Багато досліджень в сфері інвентаризації та обліку зводяться до теми, як покращити та спростити облік матеріальних цінностей для бізнесу: обробку та аналіз інформації для великих об'ємів закупок, як покращити облік закуплених товарів та їх перепродаж.

Деякі з цих методів та алгоритмів можуть бути використані в нашій розробці. Як приклад, дослідження Development of Decision Support System for Ordering Goods using Fuzzy Tsukamoto [8] розглядає проблему оптимальної кількості різноманітних товарів для підприємства. В статті розглядається проблема оптимізації закупок за допомогою алгоритмів та програмного забезпечення.

Темою цього дослідження є розробка системи підтримки прийняття рішень для замовлення товарів за допомогою нечіткої логіки Тсукамото. Це дозволяє моделювати нечіткість та неоднозначність прийнятих рішень. Застосування нечіткої логіки дозволяє системі враховувати різноманітність факторів прийняття рішень та забезпечувати ефективне замовлення товарів з урахуванням різних критеріїв. Для розробки системи були використані методи аналізу та моделювання процесів прийняття рішень, а також програмування та тестування системи. Система підтримки прийняття рішень для замовлення товарів за допомогою нечіткої логіки може допомогти компаніям в ефективному управлінні запасами та замовленнями. Для розробки та створення нашого сервісного додатку, система прийняття рішень має стати невід'ємною її частиною, і розглянутий у роботі [8] алгоритм може бути використаний та дозволить поліпшити функціональність нашої розробки.

Наукова стаття «Online Procurement and Inventory Technology Based on Cloud Computing System» [9] належить до галузі інформаційних технологій та управління запасами і присвячена розробці технології онлайн закупівель та управління запасами на основі системи хмарних обчислень. Розроблена технологія включає в себе модулі для замовлення товарів, управління запасами та контролю якості продуктів. Були використані методи програмування та тестування хмарних систем, а також методи управління запасами та процесами закупівлі. Отже, загальна мета статті полягає в розробці технології онлайн закупівель та управління запасами на основі системи хмарних обчислень, що може допомогти компаніям зберігати та управляти запасами, забезпечуючи при цьому ефективне використання ресурсів та зменшення витрат на управління запасами. Ця стаття була обрана як гарний приклад вирішення проблеми навантаження на кінцевого споживача, можливості прискорити роботу додатку від завантаження сторінок, до обробки даних. Але ресурси на розробку додатку та вартість розробки зростуть, тому планується впровадити на другій стадії розробки для її удосконалення.

Наукова стаття «Demand forecasting tool for inventory control smart systems» [10] пропонує методологію для прогнозування попиту на товари в інвентарних системах. Автори використали техніку глибинного навчання (Deep Learning) на основі нейронних мереж для прогнозування попиту на різні товари в магазинах. Дослідження показало, що запропонований метод дозволяє з точністю прогнозувати попит на товари з різних категорій та з урахуванням впливу різних факторів, таких як сезонність, рекламні акції, зміни у споживацькому попиті і т.д. Крім того, використання такої методології дозволяє покращити планування запасів товарів, що може позитивно вплинути на оптимізацію логістики та зменшення витрат на зберігання товарів. Дані зі статті показують використання глибинного навчання для прогнозування попиту та є ефективним методом для управління запасами товарів в інвентарних системах, і може призвести до значного покращення в ефективності цих систем. Але такі технології та ціна їх використання завеликі для нашого проєкту додатку- помічника.

В праці «Теоретичні та практичні аспекти розробки додатку помічника для смартфонів» [11] розглядається процес розробки та створення алгоритмів що мають удосконалити процеси виконання завдань та скоротити час на вирішення цих завдань. Стаття більш націлена на подолання проблем виконання завдань в сфері аграрного бізнесу, але має деякі цікаві ідеї для нашої роботи. Як приклад, це голосовий помічник, що допомагає створювати задачі для працівників. Така ідея, може спростити процесс створення замовлень за допомогою голосових повідомлень в додатку, але сервіси такого типу зазвичай платні та дорогі. Саме за огляд та використання голосового помічника була обранна ця стаття до розгляду. Як опція, цей функціонал можна передбачити в платній версії додатку помічника. Також, таку функцію можна використати для створення графіків на замовлення користувача не через натискання кнопки, а за голосовою командою. Функціонал для створення графіків використання, витрат залишків та іншого будуть присутні в додатку.

В статті «Дискретно-контекстуальна модель управління збутом у реальному маштабі часу» [12] досліджується стратегія динамічного управління цінами при реалізації продукції на фіксованому інтервалі часу. Враховуються два фактори попиту: поток потенційних споживачів за законом Пуассона та індивідуальні споживачі, ймовірність купівлі яких залежить від ціни. Споживачі мають обмежену потребу в одиниці товару і незалежні оцінки його споживчої вартості. Оптимальна цінова стратегія формалізується як задача оптимального управління, що вирішується за допомогою динамічного програмування. Результати дають систему диференційних рівнянь типу Ріккаті, що визначають оптимальну цінову політику залежно від часу і рівня не реалізованої продукції.

Додаток помічник не потребує розраховувати ціни, але ми можемо використовувати модель розрахунку. Наприклад, місцем для динамічного розрахунку має бути отримання інформації з різних сервісів про ціну одного товару, та надавати пріоритет дешевшим, або акційним товарам, при створення замовлення. Такий підхід підвищить конкурентність продавців товарів, та зменшить витрати користувача додатку.

Також оглядаючи статтю «Проект інформаційної системи для збору, обробки та аналізу метаданих» [13] було знайдено декілька цікавих і корисних речей. Метою статті є розробка проекту інформаційної системи для збору, обробки та аналізу метеоданих великого обсягу. Реалізація поставленої мети передбачає вирішення низки завдань:

1) визначення та характеристика специфіки побудови метеорологічних прогнозів та обробки статистичних даних;

2) розробка концепції роботи системи та її архітектури;

3) формалізація та проведення тестовою обробки даних для виявлення сильних сторін у створеній інформаційній системі.

Стаття була обрана завдяки корисній опції для обробки та зберігання даних. У статті розглядається можливість використання низки алгоритмів машинного навчання, зокрема моделей штучних нейронних мереж різних архітектурних типів, для автоматизації процесу вирішення завдання регресії та при роботі з часовими рядами даних. Дана концепція є перспективною, але складною та витратною для імплементації на даний час і буде використана в майбутньому для удосконалення нашої розробки.

2. Огляд інструментальних засобів розробки

Наразі для програмування та реалізації додатків існують багато різноманітних мов та середовищ програмування. Для розробки великих за масштабом, кількістю користувачів, кількістю клієнтів, додатків часто використовують декілька мов програмування. Взагалі, великий та середній бізнес купує різноманітні підсистеми, додатки, API тощо, та покладає технічну частину та підтримку готового продукта на виконавця. Саме виконавець, погодивши це з замовником вирішує, яку мову програмування використовувати.

Розглянемо декілька можливих мов програмування які більш підходять для розробки нашого сервісу. Для нашої розробки не потрібне щось принципово нове. Наразі маємо дві мови які найбільш підходять для цього. Це PHP та JAVA. За матеріалами «Top 4 Software Development Methodologies» маємо такі порівняльний аналіз мов [14]:

PHP – це мова програмування, завданням якої є HTML-сторінок на стороні веб сервера. Є однією з найпоширеніших та популярних мов програмування. До переваг якої відносять швидкість завантаження сторінок, різноманітні можливості до підключення к базам даних, багато довідкових матеріалів. Як інструмент розробки може бути використаний для створення помічника.

JAVA – мова програмування, яка підтримує принципи ООП, та має історію розвитку понад 25 років. Може бути використана як для розробки простих та великих проектів. Також добре підходить як мова для розробки помічника. Розглянемо переваги та недоліки цих мов для розробки додатків та програм.

1) Вважається що Java є більш швидкою ніж PHP.

2) Java більш пристосована для використання багатьма користувачами ніж PHP.

3) Java має безліч фреймворків, які є безкоштовними та мають дуже широкий функціонал. Тому вибір пав саме на Java як основної мови програмування.

Як допоміжним звеном для розробки додатка помічника мною було обрано Spring Framework [15]. Це потужний інструмент розробника на Java. Він допомагає, спрощує реалізацію доступу до баз даних за допомогою JDBC або ORM. Також надає можливість управління транзакціями та координує операції між Java objects.

Spring Security надає зручний та потужний інструмент для роботи з безпекою (авторизацією та аутентифікацією користувачів). Механізм має безліч різноманітних налаштувань, що допомагає бути гнучким в реалізації безпекової частини додатку.

Spring MVC [15] це веб-фреймворк, структура для створення слабо пов'язаних веб-додатків, що розділяє основні аспекти їхньої розробки: об'єкти, бізнес-логіку та зовнішній вигляд програми. Він дозволяє створювати веб-сайти або RESTful сервіси (наприклад JSON/XML) та добре інтегрується в екосистему Spring.

Висновки. Розглянуто підходи до вирішення проблеми ефективного управління особистими активами, наприклад запасами товарів для окремого користувача. Запланований сервіс допоможе користувачу приймати обґрунтовані рішення щодо вибору об'єму товарів для закупівлі, дозволить їм скоротити час на вирішення цих завдань та покращити процес управління запасами. На основі проведених досліджень можемо зробити висновок, що використання відповідних алгоритмів та методів дозволить отримувати оптимальні результати з точки зору вибору об'єму товарів для закупівлі.

В статті проведено інформаційний огляд та розглянуто різні аспекти інвентаризації та контролю запасами, прогнозування попиту для автоматизації процесів управління запасами. Зокрема, було проведено аналіз існуючих систем управління та прогнозування, визначено їхні переваги та недоліки. Визначені корисні для застосування методи, такі як методи статистичного аналізу, нейронні мережі та алгоритми. На основі проведеного аналізу, буде розроблено систему управління запасами, яка дозволяє здійснювати ефективний контроль за запасами. Буде використано методи прогнозування, зокрема, алгоритм Тсукамото на основі нечіткої логіки, який дозволяє отримати більш точні результати прогнозування. Розробка інструментів прогнозування є важливим кроком у покращенні ефективності управління запасами. Використання хмарних технологій також може забезпечити швидкий та надійний доступ до даних про запаси.

Отже, з огляду на проведені дослідження, можна зробити висновок, що розробка систем закупівель є важливим завданням для бізнесу та може бути використана як зразок при розробці обліку та закупівлі товарів для окремих споживачів. Використання інноваційних технологій, таких як хмарні обчислення та методи прогнозування попиту, також може допомогти підвищити ефективність цих систем та знизити витрати на управління запасами.

Розглядаються розширення додатку та покращення функціоналу. Планується додати голосове управління, залучити більш якісні та потужні механізми. Як приклад, це може бути chatGPT API. Також є можливість проаналізувати дані та використовувати їх для розробки більш якісних алгоритмів, які зможуть враховувати індивідуальні потреби користувача. Таким чином звільнити час для більш нагальних проблем користувача.

Список використаних джерел:

1. Кавецький В. В., Ратушняк О. Г. Сучасні системи управління плануванням та організацією виробництва. Ефективна економіка. 2021. DOI: 10.32702/2307-2105-2021.12.94

2. ПАТ "УКРСИББАНК". UKRSIB ONLINE. URL: https:/ /online.ukrsibbank.com/ibank/

3. Saaty T. L., Vargas L. G. Prediction, projection and forecasting: applications of the analytic hierarchy process in economics, finance, politics, games, and sports. Springer Science & Business Media. 1991. URL: https://link.springer.com/book/10.1007%2F978-94-010-9961-1

4. Al-Sharafi A.A., Ahmad N. Analytic network process (ANP): A systematic literature review. *Decision Science Letters.* 2019. 8(1). P. 89-102. URL: https://doi.org/10.5267/j.dsl.2018.12.003

5. Figueira J., Greco S., Ehrgott M. Multiple Criteria Decision Analysis: State of the Art Surveys (2nd ed.). Springer. 2005. ISBN 978-0-387-24122-3.

6. Круш П.В., Орлюк Ю. В. Теоретичні основи управління матеріальними запасами підприємств. *Економічний вісник Національного технічного університету України «Київський політехнічний інститут».* 2017. DOI: 10.20535/2307-5651. 14.2017.108775

7. Стецьків Т. Оптимізаційні моделі організації управління товарними запасами торговельних мереж. Управління соціально-економічним розвитком в умовах глобалізації. ІФННІМ ТНЕУ. 2015. URL: http://dspace.wunu.edu.ua/handle/316497/5824

8. Mousavi S. M., Mortazavi S. M., Behdadfar M. *Development of Decision Support System for Ordering Goods using Fuzzy Tsukamoto.* In 5th International Conference on Industrial Engineering and Applications (ICIEA). 2018. P. 136-141. DOI: 10.1109/ICIEA.2018.8397745

9. Syahputra M.F.D., Suryono A. *Online Procurement and Inventory Technology Based on Cloud Computing System.* In International Conference on Computer Science, Information Technology and Electrical Engineering. (ICOMITEE). 2020. P. 308-313. DOI: 10.1109/ICOMITEE50132.2020.9328331

10. Ahmed A., Ali A., Ahsan M. R. Demand forecasting tool for inventory control smart systems. In International Conference on Electrical, Communication, and Computer Engineering (ICECCE). 2020. P. 1-5. DOI: 10.1109/ICECCE50272.2020.9323892

11. Корнієнко Ю., Франциан В. Теоретичні та практичні аспекти розробки додатку помічника для смартфонів. Автоматизація технологічних та бізнес-процесів. 2021. 13(2). С. 28-36. URL: https://doi.org/10.15673/atbp.v13i2.2054

12. Мельников О. Дискретно-контекстуальна модель управління збутом у реальному масштабі часу. Національний технічний університет «Харківський політехнічний інститут», Україна. 2022. DOI: https://doi.org/10.20998/2079-0023.2022.02.10

13. Рудниченко М., Гришин С., Шибаєв Д., Петров І., Носов М. Проект інформаційної системи для збору, обробки та аналізу метаданих. Інформаційні технології та суспільство. 2021. 2(2). С. 34-41. DOI: https://doi.org/10.32689/maup.it.2021.2.4.

14. Synopsys. Top 4 Software Development Methodologies. URL: https://www.synopsys.com/blogs/software-security/top-4-software-development-methodologies/.

15. Spring Framework. (n.d.). Guides. URL: https://spring.io/guides

References:

1. Kavetsky, V.V., & Ratushnyak, O. G. (2021). Modern production planning and organization management systems. Retrieved from DOI: 10.32702/2307-2105-2021.12.94

2. JSC "UKRSIBBANK". (n.d.). UKRSIB ONLINE. Retrieved from https://online.ukrsibbank.com/ibank/

3. Saaty, T.L., & Vargas, L. G. (1991). Prediction, projection and forecasting: applications of the analytic hierarchy process in economics, finance, politics, games, and sports. Springer Science & Business Media. Retrieved from https://link.springer.com/book/10.1007%2F978-94-010-9961-1

4. Al-Sharafi, A.A., & Ahmad, N. (2019). Analytic network process (ANP): A systematic literature review. *Decision Science Letters,* 8(1), 89-102. Retrieved from https://doi.org/10.5267/j.dsl.2018.12.003

5. Figueira, J., Greco, S., & Ehrgott, M. (2005). Multiple Criteria Decision Analysis: State of the Art Surveys (2nd ed.). Springer. ISBN 978-0-387-24122-3.

6. Krush, P.V., & Orlyuk, Yu.V. (2017). Theoretical foundations of inventory management at enterprises. *Economic Bulletin of the National Technical University of Ukraine «Kyiv Polytechnic Institute».* Retrieved from DOI: 10.20535/2307-5651.14.2017.108775

7. Stetskiv, T. (2015). Optimization models of inventory management organization in trade networks. Management of socio-economic development in the conditions of globalization, IFNIM TNEU. Retrieved from http://dspace.wunu.edu.ua/handle/316497/5824

8. Mousavi, S. M., Mortazavi, S. M., & Behdadfar, M. (2018). Development of Decision Support System for Ordering Goods using Fuzzy Tsukamoto. *In 5th International Conference on Industrial Engineering and Applications (ICIEA), 2018* (pp. 136-141). Retrieved from DOI: 10.1109/ICIEA. 2018.8397745

9. Syahputra, M.F.D., & Suryono, A. (2020). Online Procurement and Inventory Technology Based on Cloud Computing System. *In International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), 2020* (pp. 308-313). Retrieved from DOI: 10.1109/ICOMITEE50132.2020.9328331

10. Ahmed, A., Ali, A., & Ahsan, M. R. (2020). Demand forecasting tool for inventory control smart systems. *In International Conference on Electrical, Communication, and Computer Engineering (ICECCE), 2020,* (pp. 1-5). Retrieved from DOI: 10.1109/ICECCE50272.2020.9323892

11. Kornienko, Y., & Frantsyan, V. (2021). Theoretical and practical aspects of developing an assistant application for smartphones. *Automation of Technological and Business Processes,* 13(2), 28-36. Retrieved from https://doi.org/10.15673/atbp.v13i2.2054

12. Melnykov, O. (2022). Discrete-contextual model of real-time sales management. National Technical University «Kharkiv Polytechnic Institute», Ukraine. Retrieved from DOI: https://doi.org/10.20998/2079-0023.2022.02.10

13. Rudnychenko, M., Hryshyn, S., Shybayev, D., Petrov, I., & Nosov, M. (2021). Design of an information system for metadata collection, processing, and analysis*. Information Technologies and Society,* 2(2), 34-41. Retrieved from DOI: https://doi.org/10.32689/maup.it.2021.2.4

14. Synopsys. (2021). Top 4 Software Development Methodologies Retrieved from https://www.synopsys.com/blogs/software-security/top-4-software-development-methodologies/

15. Spring Framework. (n.d.). Guides. Retrieved from https://spring.io/guides

UDC 004.75
DOI https://doi.org/10.32689/maup.it.2023.3.7

Lyubov KRESTYANPOL
Candidate of Technical Sciences, Senior Lecturer at Department of Applied Linguistics, Lesya Ukrainka Volyn National University, 13,Voli, Lutsk, Ukraine, postal code 43025 (krestyanpol.luybov@vnu.edu.ua)
ORCID: 0000-0003-3617-7900

Stanislav NOVACHEVSKYI
Postgraduate Student at the Department of Computer science and cyber security, Lesya Ukrainka Volyn National University 13,Voli, Lutsk, 43025, Ukraine, postal code 43025 (stan.novachevskyi@gmail.com)
ORCID:

Mykola HRANOVSKYI
Dataloop, Sapir Str. 2, Herzliya, Israel (nickolaxranovsky@gmail.com)
ORCID:

Любов КРЕСТЬЯНПОЛЬ
кандидат технічних наук, доцент кафедри прикладної лінгвістики, Волинський національний університет імені Лесі Українки, пр.т. Волі 13, Луцьк, Україна, індекс 43025 (Krestyanpol.Luybov@vnu.edu.ua)

Станіслав НОВАЧЕВСЬКИЙ
аспірант кафедри компютерних наук та кібербезпеки, Волинський національний університет імені Лесі Українки пр.т. Волі 13, Луцьк 43025, Україна, індекс 43025 (stan.novachevskyi@gmail.com)

Микола ХРАНОВСЬКИЙ
інженер технічної ідтримки, Dataloop, Sapir Str. 2, Herzliya, Ізраїль (nickolaxranovsky@gmail.com)

Bibliographic description of the article: Kresyanpol, L., Novachevskyi, S., Hranovskyi, M. (2023). Zastosuvannia blokchein – tekhnolohii dlia zberezhennia ta poshyrennia kulturnoi spadshchyny cherez NFT [Application of blockchain technologies for preservation and dissemination of cultural heritage through NFT]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 54–62. DOI: https://doi.org/10.32689/maup.it.2023.3.7

Бібліографічний опис статті: Крестьянполь, Л., Новачевський, С., Храновський, М. (2023). Застосування блокчейн – технологій для збереження та поширення культурної спадщини через NFT. *Інформаційні технології та суспільство,* 3, 54–62. DOI: https://doi.org/10.32689/maup.it.2023.3.7

APPLICATION OF BLOCKCHAIN TECHNOLOGIES FOR PRESERVATION AND DISSEMINATION OF CULTURAL HERITAGE THROUGH NFT

Abstract. *Non-fungible tokens (NFT) represent one of the most important technologies in the space of Web3. Thanks to NFTs, digital or physical assets can be tokenized to represent their ownership through the use of smart contracts and blockchains. The first generation of this technology, called NFT 1.0, considers static tokens described by a set of metadata that cannot be changed after token creation. The static nature prevents their wide spread as they do not support any meaningful user interaction. For this reason, its evolution, called NFT 2.0, has been proposed to make tokens interactive and dynamic and enhance user experience, opening the possibility to use NFTs in more ways and scenarios. Cultural heritage assets are in danger of extinction or damage due to lack of publicity and financial problems. Technological advances can play a role in their preservation and promotion.*

The preservation and dissemination of cultural heritage is a serious challenge, especially in the age of digital technologies. Non-fungible tokens (NFTs) offer a promising solution for managing and authenticating digital assets. Using blockchain technology, NFTs can help preserve and promote the visibility of these important cultural artifacts.

This research aims to create a blockchain-based cultural heritage protection system. Cultural assets are transformed into unique digital items using NFT technology. An autonomous working system is provided by smart contracts. Supporters of the project make donations to receive their share of the rights to protect and maintain cultural heritage sites.

The use of NFT in this field raises important issues of preservation, digitization, and promotion of Ukrainian culture. The purpose of the article is to analyze the blockchain technology used for NFT, developing a unique collection of digital assets and publication it on the marketplace.

Key words: *Non-fungible token, blockchain technology, cultural heritage of Ukraine, digital assets.*

ЗАСТОСУВАННЯ БЛОКЧЕЙН – ТЕХНОЛОГІЙ ДЛЯ ЗБЕРЕЖЕННЯ ТА ПОШИРЕННЯ КУЛЬТУРНОЇ СПАДЩИНИ ЧЕРЕЗ NFT

Анотація. *Незамінні токени (NFT) представляють одну з найважливіших технологій у просторі Web3. Завдяки NFT цифрові або фізичні активи можна токенізувати, щоб представляти їх власність за допомогою смарт-контрактів і блокчейнів. Перше покоління цієї технології під назвою NFT 1.0 розглядає статичні токени, описані набором метаданих, які неможливо змінити після створення токена. Статичний характер перешкоджає їх широкому розповсюдженню, оскільки вони не підтримують жодної значущої взаємодії з користувачем. З цієї причини була запропонована його еволюція, яка називається NFT 2.0, щоб зробити токени інтерактивними та динамічними та покращити взаємодію з користувачем, відкриваючи можливість використовувати NFT у більшій кількості способів і сценаріїв. Об'єкти культурної спадщини знаходяться під загрозою зникнення або пошкодження через відсутність публічності та фінансові проблеми. Технологічний прогрес може зіграти певну роль у їх збереженні та просуванні.*

Збереження та поширення культурної спадщини є серйозним викликом, особливо в епоху цифрових технологій. Невзаємозамінні токени (NFT) пропонують багатообіцяюче рішення для управління та автентифікації цифрових активів. Використовуючи технологію блокчейн, NFT можуть допомогти зберегти та сприяти видимості цих важливих культурних артефактів. Це дослідження спрямоване на створення системи захисту культурної спадщини на основі блокчейну. Культурні активи перетворюються на унікальні цифрові елементи за допомогою технології NFT. Смарт-контрактами забезпечена автономна робоча система. Прихильники проекту роблять пожертви, щоб отримати свою частку прав на захист і підтримку обєктів культурної спадщини.

Використання NFT у цій сфері ставить важливі питання збереження, оцифрування та популяризації української культури. Метою статті є аналіз технології блокчейн, яка використовується для NFT, розробка унікальної колекції цифрових активів та розміщення її на маркетплейсі.

Ключові слова: *невзаємозамінний токен, блокчейн технолгія, культурна спадщина України, цифрові активи.*

Introduction.

The use of non-fungible tokens (NFTs) to preserve and disseminate Ukrainian cultural heritage is an innovative and promising approach that can improve the visibility and accessibility of these important cultural artifacts, as well as raise funds for the restoration and rehabilitation of cultural sites damaged during the Russian-Ukrainian war. Non-fungible tokens (NFTs) have emerged as a way to collect digital data as well as a means of investment. Despite being popularized only recently, NFT markets have witnessed several high-profile asset sales and a huge increase in trading volumes over the past year [1]. The study of the NFT phenomenon, the study of key concepts and principles related to the metaverse, blockchain, decentralized programs and user interaction, and the analysis of NFT trading platforms are described in [2]. An interesting approach to analyzing the market dynamics and security issues of the multibillion-dollar NFT ecosystem is described in [1]. The author examines the NFT ecosystem and identifies three main participants: markets, external organizations, and users.

Literature review.

Non-funny token (NFT) is a data unit stored in the blockchain. From the technical point of view, we can find that art is facing unprecedented virtual and realistic problems [3, 4]. Through the discussion of digital identity, fair trade and gameplay, this paper regards NFT art as a Metaverse pointing to the future, and discusses the generation and evolution of NFT as a digital mechanism from the creation and acceptance of NFT art.

Public attention towards NFTs has exploded in 2021, when their market has experienced record sales, but little is known about the overall structure and evolution of its market. In in the article [5] made analyse data concerning 6.1 million trades of 4.7 million NFTs between June 23, 2017 and April 27, 2021, obtained primarily from Ethereum and WAX blockchains.

Interesting approach in investigate the predictability of NFT sales using simple machine learning algorithms and find that sale history and, visual features are good predictors for price described in works 5-6.

Undoubtedly, the digitization of cultural heritage is the next step in the digitalization of people, society, and the state, and the use of blockchain technologies in this area will significantly accelerate this process. In In paper 7-9 provided an overview of this technology in the sale of digital art and how it influences both the process itself and what are the challenges and risks of using it today.

Based on the value co-creation theory, in paper [10] studies the motivation of adopting gamification in art NFT, analyzes its connection with user engagement, loyalty and self-brand connection, and provides a theoretical basis and future development direction for the development of gamification in art NFT.

The study [11] describes the construction of a cryptographic method of interaction Museum Art Exchange Protocol (MAXP) for museum digital collections based on blockchain technology. Using this method, a digital collection exchange system on Ethereum is created to implement an online exchange of digital collections between two museums. Blockchain technology also has a significant impact on the transportation of cultural heritage, offering dynamic real-time remote control for the transportation of cultural relics [12, 13, 14], as well as providing regulation of the exchange of transportation of physical collections. Traditionally, there are three ways to use blockchain technology to record copyrights. The first is to record copyrights for off-grid design schemes [15], using a decentralized data management structure to protect user privacy. The second is to control the copyright

of user data protocol records under a master-slave paradigm [16] to realize digital copyright management. The third is a copyright-aware encryption algorithm [17], which uses digital watermarking technology to enhance the reliability of the encryption algorithm, creating a digital copyright blockchain management scheme. Although a number of digital copyright schemes for museums have been proposed based on blockchain technology, the regulatory mechanism for sharing digital collections is not ideal. Three typical methods are relatively mature in blockchain regulatory technology research: a blockchain transaction tracking mechanism [16], a blockchain address collection mechanism [18], and a blockchain certificate management mechanism [19, 20].

Problem statement. Preserving and disseminating Ukrainian cultural heritage is a serious challenge, especially in the digital age. Non-fungible tokens (NFTs) offer a promising solution for managing and authenticating digital assets. Using blockchain technology, NFTs can help preserve and promote the visibility of these important cultural artifacts. The use of NFTs in this area raises important issues of preservation, digitization, and promotion of Ukrainian culture.

The purpose of the study. The purpose of the article is to analyze the blockchain technology used for NFT. Development of NFT tokens for the preservation and promotion of the cultural heritage of Ukraine.

Anatomy of the NFT.

Non-fungible tokens (NFTs) are digital representations of data that contain digital authenticity and ownership and support the creation of intangible assets. NFTs are recorded on a blockchain, but not all NFTs are the same in terms of how they are stored. In this article, we will take a deeper look at the technical side of NFTs and how they are stored. Generally, content such as files and metadata can be stored on or off the blockchain. While it is a common understanding that blockchain technology allows NFTs to be immutable and permanent, this is not entirely true for all NFTs. This is where the difference between on-chain and off-chain NFTs becomes important. Before we begin, there are a few important concepts we should take into account.

A server is a computer or computer program that manages access to a centralized resource or service on a network. NFTs are sometimes stored on servers. Nowadays, it is more common to place them on IPFS or Arweave.

Hosting is the storage of a website on a server or other computer as a service for accessing it via the Internet.

Metadata is a set of data that describes and provides information about other data. Metadata helps servers find, process, and host data more efficiently. In the case of NFTs, metadata provides more information about the characteristics of the digital asset, such as the token name, the token description, and any special properties the creator wants to add (e.g., size, color, etc.).

A hash is the result of applying an algorithmic function to information to transform it into a random string of numbers and letters. This acts as a digital fingerprint of that information. Hash is commonly used to encrypt data in an efficient and verifiable manner.

Smart Contracts – A smart contract is a contract in which the terms of the agreement between the parties are directly written into code. The code and the terms of the contract stored there exist on a distributed, decentralized blockchain network. The lines of code ensure execution, and transactions are traceable, transparent, and irreversible. Without the need for a central authority, legal system, or external enforcement mechanism, smart contracts allow for trusted transactions and agreements. NFTs are minted and exchanged using smart contracts.

On-Chain NFT. NFTs exist entirely on the blockchain and have all the metadata and smart contracts stored on the chain. This means that the data of these NFTs is stored on the blockchain.

On-chain NFTs include two important data components:

– metadata, which is the basic information of a particular NFT – the name of the token, special properties of the NFT, where the digital copy is stored, etc. Since this metadata is integrated with networked NFTs, the information itself also lives on the blockchain.

– smart contracts that can generate digital assets on the network.

Since the metadata, smart contract, and NFT are stored on the blockchain, we eliminate the need to rely on external systems or a third party. Therefore, if the blockchain is up and running, the NFT collection will always be available. However, storing entire image files requires significantly more computer memory. It is therefore extremely expensive and therefore not yet widespread.

Off-Chain NFT. While on-chain NFTs exist entirely on the blockchain, off-chain NFTs do not. They exist in two parts – the smart contract and the metadata for the actual illustration. Most often, the smart contract of an NFT that is deployed on the blockchain does not store the digital work/image of the physical work or any metadata on the blockchain, but only stores a pointer to a storage location outside the blockchain, hence these NFTs are called Off-Chain. The pointer is basically a link that leads to a file containing the relevant metadata of the NFT. This metadata file contains, among other information, a link to an image file of the represented digital work or an image of the represented physical asset. The smart contract exists on the blockchain containing a set of rules that assist the transaction and serve as a digital description of the content. The smart contract also contains a link that points to a server that stores digital artwork, which may be stored offline rather than on

the blockchain. While such off-chain NFTs are currently the most common way to link NFTs to the assets they represent (e.g., an image file), this approach is technically quite risky. There is no guarantee that the file will not be replaced or overwritten by a file with the same name in the future. There is also the potential risk that the link leading to the work or even the metadata file will be tampered with, as the server hosting the image or metadata may no longer be operational. The record on the blockchain will still exist and the NFT holder will still own the NFT as it consists of the tokenID and the alphanumeric address of the smart contract. However, the NFT holder will only have the token ID, but will not have an image of the physical asset or metadata. In this case, the NFT will no longer fulfill its purpose and will not be able to associate the NFT with the represented asset.

There is indeed a risk of communication failure. This can happen if a web server company can no longer operate its servers due to bankruptcy or other reasons. Even today, there are links to NFTs that lead nowhere. To avoid the risks associated with this type of technical setup of an NFT, there are alternatives to linking an image to an NFT. For example, the NFT can point to an IPFS (Interplanetary File System) location. IPFS is a peer-to-peer network. This means that uploaded files are stored and shared on multiple computers, allowing the network to operate in a decentralized and secure manner. Each file uploaded to IPFS is assigned a unique URL, a "fingerprint" of the file. This means that pointing an NFT to a specific IPFS URL ensures that the asset behind the NFT cannot be modified (as the modified asset will be assigned a different IPFS URL). When using IPFS, any user can access a file and compare the hash value of the file with the hash value of the NFT. Therefore, anyone can verify that the NFT refers to the original file. Another solution is to use Arweave. At its core, Arweave is a decentralized storage network (DSN) that connects people who have extra available disk space with those who need more storage. It is designed to provide scalable, cost-effective, and persistent data storage, and it is built on a blockchain-like data structure called blockweave.

There are several other options for offline storage, such as Google Drive, AWS, iCloud, or centralized storage on a hardware server. However, there are some problems with this approach. First, we rely on the work of third parties, just like any centralized server or storage. Secondly, there is a security risk as these servers can be hacked and their information can be tampered with. Nevertheless, IPFS provides a more secure method of data storage because it is a P2P, distributed and decentralized storage network with no central point of failure. In the event that a storage location fails, it will be backed up by a peer. But if the NFT holder decides to delete the NFT file from off-chain storage, it will break the connection between the file and the blockchain.

Research results.

Using non-fungible tokens (NFTs) to preserve and disseminate Ukrainian cultural heritage is an innovative and promising approach that can improve the visibility and accessibility of these important cultural artifacts, preserve their digital copy, as the real object may be destroyed during the fighting in the Russian-Ukrainian war. Also, the digitization of cultural monuments and their conversion into non-fungible tokens (NFTs) allows for the creation of a certain ecosystem for the management and authentication of digital assets that can be used to restore and rebuild damaged and destroyed cultural monuments. The study by Dipanjan Das [1] considers the NFT sector as a certain ecosystem with markets, external organizations, and users. Users. NFTs are often used to sell digital collectibles and works of art. Users in the NFT ecosystem can fall into the following categories: content creator, seller, and buyer (Fig.1).

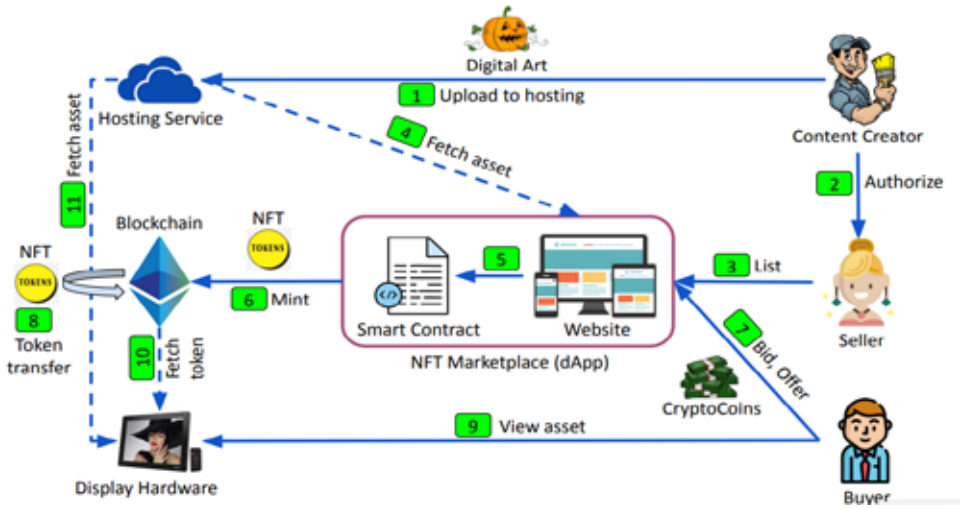


Fig. 1. Ecosystem of NFT functioning

Source: [1]

Content creators, in turn, create and upload artwork to hosting platforms where it is published and tokenized. In this ecosystem, hosting platforms play the role of external organizations that act as intermediaries between sellers and buyers. Hosting platforms are a combination of a web interface and smart contracts concluded between users.

Considering NFT as one of the tools for preserving and popularizing Ukrainian cultural heritage, we have developed a collection of NFT art aimed at preserving cultural and architectural monuments that were damaged or destroyed during the Russian-Ukrainian war. Similar projects have already been implemented and have shown their effectiveness. For example, the National Art Museum of Ukraine (NAMU), the oldest museum in Kyiv, created the NAMU NFT collections to support two goals: to raise funds for the NAMU and other museums in Ukraine that were affected by the war, and to promote Ukrainian culture around the world (Fig. 2). This project involved extensive consultation with Ukrainian cultural communities and included clear provisions to protect confidentiality and cultural sensitivity.

Fig. 2. Cossack Mamai as NFT at Patron-of-art NFT marketplace

Source: https://collect.patron-of-art.com/release/2

The collection we have developed includes 5 architectural monuments that have been damaged and need to be restored. In particular, the Holy Dormition Sviatohirsk Lavra, the first mention of which dates back to the 15th century as a monastery complex that began to form in these places during the period of settlement of the "Wild Field" by the Cossacks and the creation of the first watch fortifications (Fig. 2).

Fig. 3. NFT Sviatohirsk Lavra

Source: compilete by authors

The next project was the Donetsk Drama Theater in Mariupol (Fig. 4).

The center of cultural life and the only professional theater in Mariupol, which since March 16, 2022 has been a symbol of the humanitarian catastrophe caused by Russia. The theater became a mass grave for hundreds of people who hid within its walls from Russian shelling. The building of the Mariupol Drama Theater was built in 1956-1960 in the style of monumental classicism.

Fig. 4. NFT Donetsk Drama Theater

Source: compilete by authors

Fig. 5. NFT Regional Youth Library in Chernihiv

Source: compilete by authors

Regional youth library in Chernihiv. The neo-Gothic building was built in the late nineteenth century. It housed the premises of an orphanage and the Vasyl Tarnovsky Museum of Antiquities, which kept items from the Cossack era and Taras Shevchenko's personal belongings. Since 1978, the library has been housed here.

Fig. 6. NFT Local History Museum in Okhtyrka Sourse: compilete by authors

The building of the local history museum in Okhtyrka was built in 1861 and is a cultural landmark of local significance.

The manor was built in 1762 in the neoclassical style with Baroque elements.

Digital art in combination with blockchain technologies forms a specific cultural practice, depends on the development of technology, and responds to the Internet not only as a medium, but also as a technologically dependent, socio-cultural, political, and economic ecosystem that broadcasts cultural codes [11]. The developed collection History_Memory_and_Identity aims to preserve cultural and architectural monuments destroyed

during the war in the digital space. The proceeds from the sale of tokens are planned to be used for social projects to restore Ukraine's cultural heritage.

Fig. 7. NFT The estate of L.E. Koenig in Trostianets

Source: compilete by authors

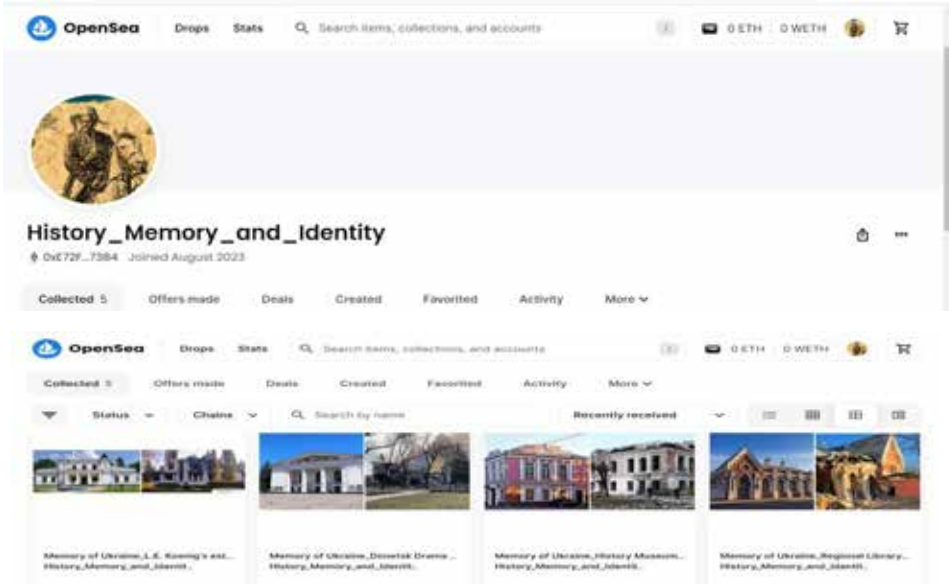


Fig. 8. The NFT collection is available on the OPENSEA marketplace

Source: compilete by authors

Conclusion.

Cultural heritage sites, especially monuments of architecture and urban planning, are of great importance and value as authentic documents of history. They are a source of information on history, cultural and socio-economic development, philosophy of perception of the world; they also convey the level of artistic and scientific thought of different generations. That is why the preservation of architectural and urban heritage sites is a very relevant issue for today, given the realities of today.

The introduction of digital blockchain technology is a promising area for the development of creative collaborations. The resource allows not only to preserve the cultural heritage object in a digitized form for future generations, but also opens up new opportunities, serves as a model for restoration work, helps to track the process of their loss, etc. The digital system is best suited for obtaining and registering rights to a digitized unique immovable object, its creation in the form of NFT tokens, and blockchain technology allows us to develop promising strategies for further interactive use of cultural heritage monuments. The available experience allows us to recognize the creation of NFT tokens based on a cultural heritage object as an act of popularization of Ukrainian castles and fortresses far beyond its borders, expansion of target audiences and involvement of those segments of the population that have never been interested in the topic of Ukraine's real estate heritage in general.

However, the problem raised requires further in-depth theoretical and practical research, in particular: the creation of a platform for digitized domestic immovable cultural heritage, a regulatory framework for the regulation of digital technologies and the crypto market.

Bibliography:

1. Dipanjan Das, Priyanka Bose, Nicola Ruaro, Christopher Kruegel, and Giovanni Vigna. Understanding Security Issues in the NFT Ecosystem. *In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22).* Association for Computing Machinery, New York, NY, USA. (2022). 667–681. URL: https://doi.org/10.1145/3548606.3559342

2. Deshpande I., Sangitrao R., Panchal L. Study of NFT Marketplace in the Metaverse. *Data Management, Analytics and Innovation. ICDMAI 2023. Lecture Notes in Networks and Systems*, vol 662. Springer, Singapore. 2023. URL: https://doi.org/10.1007/978-981-99-1414-2_59

3. Wang J. (2023). Research on NFT Art Creation and Acceptance from Metaverse Perspective. *Highlights in Art and Design*, 2(3), 98–100. 2023. URL: URL: https://doi.org/10.54097/hiaad.v2i3.8093

4. Jones N. How scientists are embracing NFTs. *Nature*. 2021. URL: https://doi.org/10.1038/d41586-021-01642-3

5. Nadini M., Alessandretti L., DiGiacinto Fl., Martino M., Aiello L.M., Baronchelli A. Mapping the NFT revolution: market trends, trade networks, and visual features. 2021. URL: https://osf.io/wsnzr/?view_only=319a53cf1bf542bbbe538aba37916537

6. Chen Z., Guo Y., Wang Z. The Future Trends of NFT: Evidence from Art and Brand Industries. *BCP Business & Management*, 28, 58–67. 2022. URL: https://doi.org/10.54691/bcpbm.v28i.2216

7. Mitu FG., Bota M. Exploratory Research on Using NFT for Selling Digital Art. *Remodelling Businesses for Sustainable Development. ICMTBHT 2022. Springer Proceedings in Business and Economics. Springer*, *Cham.* 2022. URL: https://doi.org/10.1007/978-3-031-19656-0_4

8. Nadini M., Alessandretti L., Giacinto F.D. Mapping the NFT revolution: market trends, trade networks and visual features. *Papers*. 2021.

9. Jinkook A.H.N. The meaning and limitation of nft in the art regime: crypto art in the dialectical movement of digital art. *J. Korean Modern Contemp. Art History*. 2022. 43. P. 395–429.

10. Wang J., Jiang H., Sun J., Deng L. Motivation to Adopt Gamification in NFT. *Design, User Experience, and Usability. HCII 2023. Lecture Notes in Computer Science*, vol 14030. Springer, Cham. 2023. URL: https://doi.org/10.1007/978-3-031-35699-5_44

11. Zhao L, Zhang J, Jing H, Wu J and Huang Y. A Blockchain-Based cryptographic interaction method of digital museum collections. *Journal of Cultural Heritage*. 2022. P. 69-82. Online publication date: 1-Jan-2023. URL: https://linkinghub.elsevier.com/retrieve/pii/S1296207422001923 DOI: 10.1016/j.culher.2022.11.001. 59.

12. Zhang J., Guo M., Li B., Lu R. A transport monitoring system for cultural relics protection based on blockchain and internet of things. *Journal of Cultural Heritage*. 2021. 50. P. 106-114.

13. Volynets V. NFT as a Symbiosis of Modern Digital Art and Blockchain Technology. *Bulletin of the National Academy of Culture and Arts Managers: of science magazine*. 2023. No. 1. P. 42–47. [in Ukrainian]

14. Guljajeva V., Sola M. C. NFT Shop and Making Sense of the NFT Art Market. Is NFT a blessing or a curse to digital art? *Academy of Managerial Staff of Culture and Arts Herald*: Science journal. 2023. 1. 42–47. URL: https://doi.org/10.31235/osf.io/xm3jbal

15. Vishwa A., Hussain F. K. A blockchain based approach for multimedia privacy protection and provenance. In *2018 IEEE symposium series on computational intelligence (SSCI)*. 2018. 1941-1945.

16. Ma Z., Huang W., Bi W., Gao H., Wang Z. A master-slave blockchain paradigm and application in digital rights management. *China Communications*. 2018. 15(8). 174-188.

17. Zhaofeng M., Weihu, H., Hongmin G. A new blockchain-based trusted DRM scheme for built-in content protection. *EURASIP Journal on Image and Video Processing*. 2018. (1). 1-12.

18. Zhu P., Hu J., Li X., Zhu Q. Using blockchain technology to enhance the traceability of original achievements. *IEEE Transactions on Engineering Management*. 2021.

19. Zheng B., Zhu L., Shen M., Du X., Guizani M. Identifying the vulnerabilities of bitcoin anonymous mechanism based on address clustering. *Science China Information Sciences*. 2020. 63. 1-15.

20. Dumpeti N. K., Kavuri R. WITHDRAWN: A framework to manage smart educational certificates and thwart forgery on a permissioned blockchain. 2021.

References:

1. Dipanjan Das, Priyanka Bose, Nicola Ruaro, Christopher Kruegel, and Giovanni Vigna. (2022). Understanding Security Issues in the NFT Ecosystem. *In Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security (CCS '22).* Association for Computing Machinery, New York, NY, USA, 667–681. Retrieved from https://doi.org/10.1145/3548606.3559342

2. Deshpande, I., Sangitrao, R., Panchal, L. (2023). Study of NFT Marketplace in the Metaverse. *Data Management, Analytics and Innovation. ICDMAI 2023. Lecture Notes in Networks and Systems*, vol 662. Springer, Singapore. Retrieved from https://doi.org/10.1007/978-981-99-1414-2_59

3. Wang, J. (2023). Research on NFT Art Creation and Acceptance from Metaverse Perspective. *Highlights in Art and Design*, 2(3), 98–100. Retrieved from https://doi.org/10.54097/hiaad.v2i3.8093

4. Jones, N. (2021). How scientists are embracing NFTs. *Nature*. Retrieved from https://doi.org/10.1038/d41586-021-01642-3

5. Nadini, M., Alessandretti, L., DiGiacinto, Fl., Martino, M., Aiello, L. M.& Baronchelli, A. (2021). Mapping the NFT revolution: market trends, trade networks, and visual features. Retrieved from https://osf.io/wsnzr/?view_only=319a53cf1bf542bbbe538aba37916537

6. Chen, Z., Guo, Y., & Wang, Z. (2022). The Future Trends of NFT: Evidence from Art and Brand Industries. *BCP Business & Management*, 28, 58–67. Retrieved from https://doi.org/10.54691/bcpbm.v28i.2216

7. Mitu, FG. & Bota, M. (2022). Exploratory Research on Using NFT for Selling Digital Art. *Remodelling Businesses for Sustainable Development. ICMTBHT 2022. Springer Proceedings in Business and Economics. Springer*, *Cham.* Retrieved from https://doi.org/10.1007/978-3-031-19656-0_4

8. Nadini, M., Alessandretti, L. & Giacinto, F.D. (2021). Mapping the NFT revolution: market trends, trade networks and visual features. *Papers*.

9. Jinkook, A.H.N. (2022). The meaning and limitation of nft in the art regime: crypto art in the dialectical movement of digital art. *J. Korean Modern Contemp. Art History*, 43, 395–429.

10. Wang, J., Jiang, H., Sun, J. &Deng, L. (2023). Motivation to Adopt Gamification in NFT. *Design, User Experience, and Usability. HCII 2023. Lecture Notes in Computer Science*, vol 14030. Springer, Cham. Retrieved from https://doi.org/10.1007/978-3-031-35699-5_44

11. Zhao, L., Zhang, J.,Jing, H., Wu Jand Huang, Y. (2022). A Blockchain-Based cryptographic interaction method of digital museum collections.*Journal of Cultural Heritage*. Retrieved from https://linkinghub.elsevier.com/retrieve/pii/S129620742200 1923

12. Zhang, J., Guo, M., Li, B., & Lu, R. (2021). A transport monitoring system for cultural relics protection based on blockchain and internet of things. *Journal of Cultural Heritage*, *50*, 106-114.

13. Volynets V. (2023). NFT as a Symbiosis of Modern Digital Art and Blockchain Technology. *Bulletin of the National Academy of Culture and Arts Managers: of science magazine*. No. 1. P. 42–47. (in Ukrainian)

14. Guljajeva, V., & Sola, M. C. (2023). NFT Shop and Making Sense of the NFT Art Market. Is NFT a blessing or a curse to digital art? *Academy of Managerial Staff of Culture and Arts Herald*: Science journal, 1, 42–47. Retrieved from https://doi.org/10.31235/osf.io/xm3jbal

15. Vishwa, A., & Hussain, F. K. (2018). A blockchain based approach for multimedia privacy protection and provenance. In *2018 IEEE symposium series on computational intelligence (SSCI)*, 1941-1945.

16. Ma, Z., Huang, W., Bi, W., Gao, H., & Wang, Z. (2018). A master-slave blockchain paradigm and application in digital rights management. *China Communications*, *15*(8), 174-188.

17. Zhaofeng, M., Weihua, H., & Hongmin, G. (2018). A new blockchain-based trusted DRM scheme for built-in content protection. *EURASIP Journal on Image and Video Processing*, (1), 1-12.

18. Zhu, P., Hu, J., Li, X., & Zhu, Q. (2021). Using blockchain technology to enhance the traceability of original achievements. *IEEE Transactions on Engineering Management*.

19. Zheng, B., Zhu, L., Shen, M., Du, X., & Guizani, M. (2020). Identifying the vulnerabilities of bitcoin anonymous mechanism based on address clustering. *Science China Information Sciences*, *63*, 1-15.

20. Dumpeti, N. K., & Kavuri, R. (2021). WITHDRAWN: A framework to manage smart educational certificates and thwart forgery on a permissioned blockchain. 2021.

UDC 004.65
DOI https://doi.org/10.32689/maup.it.2023.3.8

Valerii NIKITIN
Postgraduate Student at the Department of Information Systems and Technologies, Igor Sikorsky Kyiv Polytechnic Institute (19valeranikitin96@gmail.com)

ORCID: 0000-0002-4509-1204

Evgen KRYLOV
Candidate of Technical Sciences, Senior Lecturer at the Department of Information Systems and Technologies, Igor Sikorsky Kyiv Polytechnic Institute (ekrylov1964@ gmail.com)

ORCID: 0000-0003-4313-938X

Валерій НІКІТІН
аспірант кафедри інформаційних систем та технологій, КПІ ім. Ігоря Сікорського (19valeranikitin96@gmail.com)

Євген КРИЛОВ
кандидат технічних наук, доцент кафедри інформаційних систем та технологій, КПІ ім. Ігоря Сікорського (ekrylov1964@gmail.com)

Bibliographic description of the article: Nikitin, V., Krylov, E. (2023). Mekhanizm Active Anti-Entropy na osnovi spektralnoho filtru Bluma ta PH-2 alhorytmu kheshuvannia dlia uzghodzhennia replik u nereliatsiinykh rozpodilenykh dokumentooriientovanykh bazakh danykh [Active Anti-Entropy mechanism based on Spectral Bloom Filter and PH-2 hash algorithm for reconcilation of replicas of NoSQL distributed document oriented databases]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 63–67. DOI: https://doi.org/10.32689/maup.it.2023.3.8

Бібліографічний опис статті: Нікітін В., Крилов Є. Механізм Active Anti-Entropy на основі спектрального фільтру Блума та PH-2 алгоритму хешування для узгодження реплік у нереляційних розподілених документоорієнтованих базах даних. *Інформаційні технології та суспільство.* 3, 63–67. DOI: https://doi.org/10.32689/maup.it.2023.3.8

ACTIVE ANTI-ENTROPY MECHANISM BASED ON SPECTRAL BLOOM FILTER AND PH-2 HASH ALGORITHM FOR RECONCILATION OF REPLICAS OF NOSQL DISTRIBUTED DOCUMENT ORIENTED DATABASES

Abstract. *Information systems are used in many areas of human activity, which are not limited to one country or continent. This may require horizontal scaling for the system to function properly. Ignoring this can affect performance and availability, which in turn can lead to a loss of reputation and users.*

Horizontal scaling increases the number of database replicas, which creates the need for data reconciliation, since writing operations to different nodes increases entropy. There are various technologies aimed at reducing it, including Active Anti-Entropy. Its essence is to detect inconsistencies and start the reconciliation process between replicas. It is actively used in a database such as Riak and uses the Merkle Tree data structure, which is based on the use of hashing algorithms. The speed of inconsistency identification depends on the chosen hashing algorithms and the number of documents in the collection. An increase in the number of documents or even their size can worsen the even distribution and lead to an increase in the number of collisions. The occurrence of collisions increases the time period of data inconsistency, because the system cannot detect the inconsistency in time.

In addition to the collisions that can occur, you need to consider the delay due to data transfer over the network when nodes interact, and remember that such verification is not a one-time operation, but requires constant computation on replicas and sending for verification. Minimizing the time of these operations will speed up the data reconciliation process.

Critically important data must be reconciled with minimal delay, as an untimely or incorrectly made decision can lead to material or even human losses. To prevent this, there must be a solution that will minimize the delay of matching such data.

Key words: *NoSQL, distributed system, Active Anti-Entropy, Spectral Bloom Filter, consistency, PH2 hash algorithm.*

МЕХАНІЗМ ACTIVE ANTI-ENTROPY НА ОСНОВІ СПЕКТРАЛЬНОГО ФІЛЬТРУ БЛУМА ТА PH-2 АЛГОРИТМУ ХЕШУВАННЯ ДЛЯ УЗГОДЖЕННЯ РЕПЛІК У НЕРЕЛЯЦІЙНИХ РОЗПОДІЛЕНИХ ДОКУМЕНТО-ОРІЄНТОВАНИХ БАЗАХ ДАНИХ

Анотація. *Інформаційні системи використовуються у багатьох сферах діяльності людини, які не обмежуються однією країною або континентом. Це може призводити до необхідності горизонтального масштабування, щоб система могла нормально функціонувати. Ігнорування цього може впливати на швидкодію та доступність, що у свою чергу призведе до втрати репутації та користувачів.*

При горизонтальному масштабуванні збільшується кількість реплік бази даних, що створює необхідність в узгодженні даних, оскільки операції запису до різних вузлів збільшує ентропію. Є різні технології, які направлені на її зменшення, серед яких Active Anti-Entropy. Суть її полягає у тому, щоб виявити неконсистентність та розпочати процес узгодження між репліками. Вона активно використовується у такій базі даних, як Riak та використовує структуру даних Merkle Tree, яка базується на використанні алгоритмів хешування. Швидкість ідентифікування неузгодженості залежить від обраних алгоритмів хешування та кількості документів в колекції. Збільшення кількості документів або навіть їх розмір може погіршувати рівномірний розподіл та призводити до збільшення кількості колізій. Виникнення колізій збільшує проміжок часу неузгодженості даних, оскільки система не може вчасно виявити неконсистентність.

Окрім колізій, які можуть виникати, потрібно враховувати затримку через передачу даних мережею при взаємодії вузлів та пам'ятати, що така перевірка не є поодинокою операцією, а вимагає постійного обчислення на репліках та відправки для перевірки. Мінімізація часу виконання цих операцій дозволить пришвидшити процес узгодження даних.

Критично важливі дані повинні бути узгоджені з мінімальною затримкою, оскільки невчасно або неправильно прийняте рішення може призвести до матеріальних, або навіть людських втрат. Для запобігання цьому, повинно існувати рішення, яке дозволить мінімізувати затримку узгодження таких даних.

Ключові слова: *нереляційна база даних, розподілена система, активна антієнтропія, спектральний фільтр Блума, консистентність, алгоритм хешування PH2.*

Problem statement. Horizontal scaling of information systems increases speed and availability due to increased computing power, but in turn, creates additional tasks. One of these tasks is the reconciliation of data on different replicas of a distributed database [1].

Since digitalization permeates almost all spheres of human activity, there may be completely different cases that emphasize the need to pay attention to it. It can be the coordinates of objects in space, and various financial transactions that can be carried out in different point of the Earth. Given the development of mobile technologies, which cause an increase in bandwidth and radius of coverage, the opportunities for automating processes that could not be automated before are increasing. In addition, such a direction as the Internet of Things is actively developing, which requires a large number of sensors, which in turn increase the amount of information in cyber-physical systems. It should be noted that technologies are actively being implemented in the goverment sector of countries and information about citizens should be consistent for all public services from different sources. In addition, the consistency of certain data on citizens can be useful for partner countries to simplify the work of border services, perhaps even to find the necessary qualified engineers or scientists [2].

Latest research and publications analysis. Distributed databases can use different techniques to maintain consistency. They can represent not only additional mechanisms, but be provided at the level of the architecture of the database itself and its CRUD operations [3].

One method may be to centralize write operations, which guarantees synchronous writing to replicas. Data is read from replicas, which significantly reduces the load from the central node. In addition, the presence of replicas improves the availability of the database, because if one node goes down, it is possible to continue work using copies.

Ведення версій записів також є методом вирішення конфліктів, яке використовує вектор часу. Вектор часу – це послідовність пар, яка описує порядок поновлення цього запису. Перевагами вектору часу є відсутність єдиної точки відмови системи, тому що при використанні тимчасових міток в записах необхідно виконувати точну синхронізацію часу з одним еталоном. Недоліками вектору часу є відсутність можливості автоматично вирішувати конфлікти, а також збільшення довжини вектору часу при багаторазовому оновленні запису. Однак в NoSQL існують механізми урізання вектору часу. Наприклад, в системі Riak можна задавати частоту обрізання вектору на рівні сегмента, а також максимальний розмір (довжину) вектора часу [4].

Active Anti-Entropy mechanism consists in detecting and correcting inconsistency. The active antientropy process involves periodically comparing and synchronizing data between nodes to detect any differences. This ensures that copies of data on different nodes of a distributed system always remain consistent with each other [5].

A Merkle tree is a data structure used in distributed systems to effectively check the integrity of data in the Active Anti-Entropy mechanism. It is created by recursively hashing pairs of data, creating a tree-like structure where each leaf represents a particular piece of data and each non-leaf represents the hash of its children. The highest level of the tree, known as the root node, contains a single hash value, often called the root Merkle hash [6].

Merkle trees are often used to ensure data integrity in blockchain technologies. By comparing root Merkle hashes between different nodes, it is possible to effectively determine whether the data between these nodes is consistent or whether there are differences.

Bloom filter is a probabilistic data structure that is used to efficiently determine whether a certain element belongs to a set of data. Use this filter to quickly respond to queries about the presence or absence of certain data in a set without having to store all of that data separately [7].

One of the varieties of this data structure is the Spectral Bloom filter, which is a vector of counters. The value of the counter increases when it is accessed accordingly. The peculiarity is that it becomes possible to add and remove elements that have been added to the filter, unlike the classic Bloom filter [8].

Aim of the research. The main goal of the research is to present the proposed Active Anti-Entropy mechanism for matching critical data in distributed NoSQL document-oriented databases, as existing methods are vulnerable due to insufficient collision resistance in the context of critical data.

Active Anti-Entropy mechanism based on Spectral Bloom Filter and PH-2 algorithm. The Active Anti-Entropy mechanism is a process that occurs in the background of a running distributed database. The essence of the mechanism is to search for entropy and perform the process of matching data between different replicas.

The diagram of the Active Anti-Entropy mechanism using the Spectral Bloom filter and the PH-2 algorithm is shown in Fig. 1.

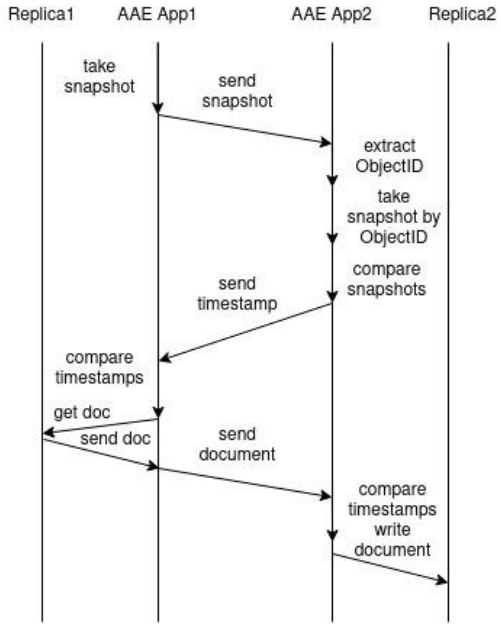


Fig. 1. Scheme of operation of the proposed Active Anti-Entropy mechanism

The essence of the mechanism is that there is a constant exchange of snapshots between the nodes, which are compared with each other. Consider the mechanism on the example of a distributed NoSQL document-oriented database, which consists two replicas.

The first replica takes a snapshot for a particular document and sends it to another replica. When the second replica received this snapshot, it calculates the snapshot of the same document by the document ID and compares it with what it received. If the snapshots are the same, the second replica simply ignores the received snapshot. In this case, this node can send a certain message to the 1st node that the data is agreed and no additional actions are required. If the snapshots are different, the second node sends the document identifier and the timestamp of its last modification. This is necessary so that the first node can determine which replica stores the most recent data. If, after comparing the timestamps, it turns out that the first replica contains a newer document, then the document is taken from it and sent to the second node. The second node receives the document, compares the timestamps, and writes it to the database if the time of the last modification of the local document is older than the time of the received document.

If the first replica contains an outdated document, it simply ignores that the timestamps are different because the second node performs the same actions in parallel, providing simultaneous monitoring.

Since the exchange of messages occurs continuously, the following requirements for messages arise:

– the snapshot must be collision-resistant;

– the snapshot calculation speed should be maximum;

– the size of the snapshot should be minimal to reduce the time required for transmission between nodes over the computer network;

Collision resistance is a critically important indicator, as it depends on how quickly the mechanism can detect inconsistencies and start the data reconciliation process.

Alternatively, cryptographic hashing algorithms could be applied, but these have more security-related properties. For consistency purposes, this is irrelevant, but can negatively affect the needs of fast computation and size. In other words, they do not meet the second and third requirements. In this case, non-cryptographic hashing algorithms can be used, but the issue of collision resistance remains open for them.

Another possible option is to use a spectral Bloom filter. In its classic form, it is a vector of counters that is formed from input data. The counter is an unsigned integer. Each input block of data is hashed, a digest is obtained and the position in the vector is calculated using it. When addressing a certain element of this vector, the value of the counter is increased by one.

It would be advisable to use it if you change the algorithm of forming this filter and avoid usage of hash functions. Also, it would be advisable to use such an algorithm, which would target data that in most cases is stored in databases to increase collision resistance.

In addition to the snapshot, the message may include the operation ID, document ID, PH2 hash, and timestamp. The structures of possible messages are shown in figures 2 and 3.

39 bytes			
1 byte	24 bytes	8 bytes	6 bytes
OpCode	ObjectID	Spectral Bloom Filter	PH2-48 hash

Fig. 2. Structure of message to check documents

52 bytes		
1 byte	24 bytes	27 bytes
OpCode	ObjectID	UTC Timestamp

Fig. 3. Structure of message to compare timestamps

The operation identifier is 1 byte in size and can take three values:

– if a message is sent to terminate the process, the field is set to 0;

– if a message is sent with a snapshot, the field takes the value 1;

– if a message is sent with a timestamp, the field takes the value 2.

These identifiers allow you to identify the type of message and process it in the appropriate way.

In the figures, there is an ObjectID that identifies a particular document in the database. Its size and format depends on the database for which mechanism is applied.

Figure 3 has a UTC Timestamp, which represents the time stamp when the last changes were made to the document. It is needed to determine the document that is newer on replicas. Its size and format can also be represented differently and depends on the precision with which time is described. For example, you can use a timestamp with seconds only, or you can also include microseconds or nanoseconds. Since not all databases have a dedicated API to retrieve this value, the mechanism must store it separately, or developers must add it to those documents that comply.

There is also a field for PH2-48 hash value, which is required to avoid collision situations. This algorithm is sensitive to changes in the size of the input data, which makes it useful in cases where the filters are the same when formed from different data arrays, but they are differ in size [9].

It should be noted that there is also a process termination operation when there is no need for negotiation, but the message in this case will consist only of the operation code, which is equivalent in size to 1 byte.

The mechanism uses UDP and TCP transport layer protocols for communication, which reduces the processing time of messages on the sender and receiver side. The use of such protocols as HTTP, HTTPs is impractical, as it creates an unnecessary load when encapsulating and decapsulating packets [10].

The UDP protocol is used for all operations. The TCP protocol ensures reliable transmission of documents when inconsistencies are detected. For the mechanism to work, you need to listen two ports at the same time to ensure operation state.

Discussion of the results and conclusions. Thus, the proposed method of the Active Anti-Entropy mechanism can solve the problem of consistency of critical data in distributed NoSQL document-oriented databases. The next step is the implementation of this mechanism for the existing NoSQL document-oriented database with the search for optimal means that will allow to achieve maximum collision resistance, message calculation speed and minimize message size for less delay due to transmission by computer networks. The implementation of such a subsystem will allow conducting experimental studies and identifying the strengths and weaknesses of the method itself.

Bibliography:

1. Changlin H. Survey on NoSQL Database Technology. *Journal of Applied Science and Engineering Innovation.* 2015. 2, 50-54. URL: http://www.jasei.pub/PDF/2-2/2-50-54.pdf

2. Muniswamaiah M., Agerwala T., C. Tappert C. Performance of databases in IoT applications. *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom).* 2020. (190-192). New York, NY, USA : IEEE. URL: https://doi.org/10.1109/CSCloud-EdgeCom49738.2020.00041

3. K. Aguilera M., B. Terry D. The Many Faces of Consistency. *IEEE Database Engineering Bulletin.* 2016. 3-13. URL: http://sites.computer. org/debull/A16mar/p3.pdf

4. Belous R., Krylov E. TIME OPTIMIZATION OF PROCESS OF DATA CONSISTENCY IN NOSQL. *Herald of the Khmelnytskyi National University. Series: "Technical Sciences".* 2023. 3, 37-42. URL: http://journals. khnu.km.ua/vestnik/wp-content/uploads/2023/07/vknu-ts-2023-n3321-37-42.pdf

5. Nikitin V., Krylov E. A collision-resistant hashing algorithm for maintaining consistency in distributed NoSQL databases. *Adaptive Systems of Automatic Control Interdepartamental scientific and technical collection.* 2022. 2, 45-57. URL: https://doi.org/10.20535/1560-8956.41.2022.271338

6. Tarkoma S., Rothenberg C., Lagerspetz E. Theory and Practice of Bloom Filters for Distributed Systems. *IEEE Communications Surveys & Tutorials.* 2011. 14, 131-155. URL: https://doi.org/10.1109/SURV.2011.031611.00024

7. Cohen S., Matias Y. Spectral Bloom Filters. *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data.* 2003. 1-12. URL: http://dx.doi.org/10.1145/872757.872787

8. Nikitin V., Krylov E. Comparison of hashing methods for supporting of consistency in distributed databases. *Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection.* 2022. 1, 48-53. URL: http://asac.kpi.ua/article/view/261646/258069

9. Al-Dhief F., Sabri N., Latiff N., Obaid O. Performance comparison between TCP and udp protocols in different simulation scenarios. *International Journal of Engineering & Technology.* 2018. 7, 172-176. URL: https://doi.org/10.14419/ijet.v7i4.36.23739

References:

1. Changlin, H. (2015). Survey on NoSQL Database Technology. *Journal of Applied Science and Engineering Innovation,* 2, 50-54. Retrieved from http://www.jasei.pub/PDF/2-2/2-50-54.pdf

2. Muniswamaiah, M., Agerwala, T., & C. Tappert, C. (2020). Performance of databases in IoT applications. *2020 7th IEEE International Conference on Cyber Security and Cloud Computing (CSCloud)/2020 6th IEEE International Conference on Edge Computing and Scalable Cloud (EdgeCom),* (190-192). New York, NY, USA : IEEE. Retrieved from https://doi.org/10.1109/CSCloud-EdgeCom49738.2020.00041

3. K. Aguilera, M., & B. Terry, D. (2016). The Many Faces of Consistency. *IEEE Database Engineering Bulletin,* 3-13. Retrieved from http://sites.computer.org/debull/A16mar/p3.pdf

4. Belous, R., & Krylov, E. (2023). TIME OPTIMIZATION OF PROCESS OF DATA CONSISTENCY IN NOSQL. *Herald of the Khmelnytskyi National University. Series: "Technical Sciences",* 3, 37-42. Retrieved from http://journals.khnu.km.ua/vestnik/wp-content/uploads/2023/07/vknu-ts-2023-n3321-37-42.pdf

5. Nikitin, V., & Krylov, E. (2022). A collision-resistant hashing algorithm for maintaining consistency in distributed NoSQL databases. *Adaptive Systems of Automatic Control Interdepartamental scientific and technical collection,* 2, 45-57. Retrieved from https://doi.org/10.20535/1560-8956.41.2022.271338

6. Tarkoma, S., Rothenberg, C., & Lagerspetz, E. (2011). Theory and Practice of Bloom Filters for Distributed Systems. *IEEE Communications Surveys & Tutorials,* 14, 131-155. Retrieved from https://doi.org/10.1109/SURV.2011.031611.00024

7. Cohen, S., & Matias, Y. (2003). Spectral Bloom Filters. *Proceedings of the 2003 ACM SIGMOD International Conference on Management of Data,* 1-12. Retrieved from http://dx.doi.org/10.1145/872757.872787

8. Nikitin, V., & Krylov, E. (2022). Comparison of hashing methods for supporting of consistency in distributed databases. *Adaptive Systems of Automatic Control Interdepartmental scientific and technical collection,* 1, 48-53. Retrieved from http://asac.kpi.ua/article/view/261646/258069

9. Al-Dhief, F., Sabri, N., Latiff, N., & Obaid, O. (2018). Performance comparison between TCP and udp protocols in different simulation scenarios. *International Journal of Engineering & Technology,* 7, 172-176. Retrieved from https://doi.org/10.14419/ijet.v7i4.36.23739

УДК 004.9
DOI https://doi.org/10.32689/maup.it.2023.3.9

Євген РЕХЛЕЦЬКИЙ
кандидат економічних наук, доцент кафедри комп'ютерних технологій, Львівський торговельно-економічний університет, вул. Туган-Барановського, 10, Львів, Україна, індекс 79005 (rea@lute.lviv.ua)
ORCID: 0000-0002-8879-4161

Василь ПЛЕША
ст. викладач кафедри комп'ютерних технологій, Львівський торговельно-економічний університет, вул. Туган-Барановського, 10, Львів, Україна, індекс 79005 (plesha_v@i.ua)
ORCID: 0000-0001-5321-9602

Леся ХМІЛЯРЧУК
ст. викладач кафедри комп'ютерних технологій, Львівський торговельно-економічний університет, вул. Туган-Барановського, 10, Львів, Україна, індекс 79005 (_lkh@ukr.net)
ORCID: 0000-0002-1753-6472

Марія КРУТЯК
асистент кафедри комп'ютерних технологій, Львівський торговельно-економічний університет, вул. Туган-Барановського, 10, Львів, Україна, індекс 79005 (kmb55@lute.lviv.ua)
ORCID: 0009-0000-8544-8733

Evhen REKHLETSKYY
Candidate of Economic Sciences, Senior Lecturer at the Department of Computer Technologies, Lviv University of Trade and Economics, Faculty of Economics, 10 Tugan-Baranovskoho Street, Lviv, Ukraine, postal code 79005 (rea@lute.lviv.ua)

Vasyl PLESHA
Assistant professor of the Department of Computer Technologies, Lviv university of trade and economics, St. 10 Tugan-Baranovskoho Street, Lviv, Ukraine, postal code 79005 (plesha_v@i.ua)

Lesia KHMILIARCHUK
Assistant professor of the Department of Computer Technologies, Lviv university of trade and economics, St. 10 Tugan-Baranovskoho Street, Lviv, Ukraine, postal code 79005 (_lkh@ukr.net)

Mariia KRUTIAK
Assistant of the Department of Computer Technologies, Lviv University of Trade and Economics, St. 10 Tugan-Baranovskoho Street, Lviv, Ukraine, postal code 79005, (kmb55@lute.lviv.ua)

Бібліографічний опис статті: Рехлецький, Є., Плеша, В., Хмілярчук, Л., Крутяк, М. (2023). Вплив інформаційних технологій на еволюцію системи освіти. *Інформаційні технології та суспільство*, 3, 68–71. DOI: https://doi.org/10.32689/maup.it.2023.3.9

Bibliographic description of the article: Rekhletskyi, E., Plesha, V., Khmylarchuk, L., Krutyak, M. (2023). Vplyv informatsiinykh tekhnolohii na evoliutsiiu systemy osvity [The influence of information technologies on the evolution of the education system]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 68–71. DOI: https://doi.org/10.32689/maup.it.2023.3.9

ВПЛИВ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ НА ЕВОЛЮЦІЮ СИСТЕМИ ОСВІТИ

Анотація. *Освіта має найдавнішу історію серед всіх складових науки, на протязі всієї історії людства серед всіх верст населення освіта мала особливий статус і до неї був особливий підхід. Якщо еволюцію освіти поділити умовно на 7 етапів розтягнуті на 6 тисяч років, то останні 3 охоплюватимуть трохи більше 100 років, відтак можна зробити висновок, що еволюція освіти значно прискорилася. Значною мірою це пов'язано зі стрімким розвитком науки та технологій, але якщо висока динаміка в науки це більше позитивний процесс, то висока динаміка у освіти це нестабільність і невпевненість, тобто криза. Для освіти зараз багато викликів. Серед них глобальні виклики – такі, як зміна світового порядку, і такі, які не змінилися на протязі історії людства – соціальна нерівність і доступність. Деякі виклики пов'язані з передовим технологіями, а саме з високою динамікою їх розвитку – технології, що запозичуються до освітнього процесу морально застарівають з тею ж швидкістю, що й самі технології, тобто занадто*

швидко. За деякими тенденціями можна лише спостерігати, а деякі вже бажано використовувати. Одну з таких проблем, на стику інформаційних технологій та освіти і розглядає дана робота.

***Метою роботи** є привернути увагу до можливостей розповсюджених технологій, які можна і потрібно залучати вже сьогодні.*

*Галузь інформаційних технологій з надзвичайно високою динамікою генерує нові технологічні підходи, тому шукати рішень для проблем галузі освіти у новітніх технологіях логічно і **актуально.***

***Висновки.** Деякі проблеми сучасної освіти можна вирішити використанням розповсюджених технологій, і вартує додати до їх використання більше зусиль.*

***Ключові слова:** освіта, інформаційні технології.*

THE INFLUENCE OF INFORMATION TECHNOLOGIES ON THE EVOLUTION OF THE EDUCATION SYSTEM

***Abstract.** Education has the oldest history among all components of science, throughout the entire history of mankind, among all segments of the population, education had a special status and there was a special approach to it. If the evolution of education is conventionally divided into 7 stages spanning 6,000 years, the last 3 will cover a little more than 100 years, so it can be concluded that the evolution of education has accelerated significantly. To a large extent, this is related to the rapid development of science and technology, but if high dynamics in science is more of a positive process, then high dynamics in education is instability and uncertainty, that is, a crisis. There are many challenges for education today. Among them are global challenges – such as a change in the world order, and those that have not changed throughout the history of mankind – social inequality and affordability. Some challenges are associated with advanced technologies, namely with the high dynamics of their development – technologies borrowed from the educational process become morally obsolete at the same speed as the technologies themselves, that is, too quickly. Some trends can only be observed, and some are desirable to use. This work examines one of these problems at the junction of information technologies and education.*

The purpose of the work is to draw attention to the possibilities of widespread technologies, which can and should be used today.

The field of information technology with extremely high dynamics generates new technological approaches, so it is logical and relevant to look for solutions to the problems of the field of education in the latest technologies.

***Conclusions**. Some problems of modern education can be solved by the use of widespread technologies, and it is worth adding more efforts to their use.*

***Key words:** education, information technologies, software licenses.*

Наука, за визначенням, це сфера діяльності людини, спрямована на отримання (вироблення і систематизацію) нових знань про навколишній світ.

Освіта, за визначенням, це процес набуття та вдосконалення знань, навичок, цінностей, переконань і звичок. Навіть за визначенням зрозуміло, що освіта – це частина світу науки, яка, до того ж, має найдавнішу історію.

Цей процес передбачає, що є вчитель та учень. Один передає інформацію, другий сприймає. Крім отримання інформації, треба її вірно застосувати у контексті розуміння дійсності, тобто знати що робити і як робити. У філософському сенсі можна припустити, що вчитель передає суб'єктивну картину свого бачення світу. Відтак особистість вчителя дуже важлива. Принаймні так було раніше.

Еволюціяо світи була динамічним процесом, на який впливали різноманітні історичні, культурні, технологічні та соціальні фактори. Античний і класичний періоди – у стародавніх цивілізаціях, таких як Єгипет, Месопотамія, Греція та Рим, освіта була прерогативою еліти, і було зосереджено на філософії, риториці, математиці та фізичному вихованні.

Середньовічний період – освіта зосереджувалася навколо релігійних установ, таких як монастирі та кафедральні школи, було суто релігійною, з наголосом на богослов'ї та навичках латинської мови.

Відродження і Просвітництво – ця епоха (Відродження) викликала відновлення інтересу до класичного навчання, що призвело до заснування гуманітарних шкіл, які зосереджувалися на предметах поза теологією.

Промислова революція та масова освіта – потребувала більш освіченої робочої силі, що призвело до прийняття законів про обов’язкову, масову освіту в багатьох країнах.

Прогресивний освітній рух, що виник наприкінці 19-го та на початку 20-го століть наголошував на практичному і індивідуальному навчанні та практичних навичках.

Освіта після Другої світової війни, тобто від середині 20-го століття – освітня експансія була спрямована на розширення доступу до освіти для всіх соціально-економічних груп.

Наприкінці 20-го століття і надалі прогрес призвів до інтеграції технологій в освіту, що змінило спосіб навчання учнів і навчання вчителів – онлайн-навчання, цифрові ресурси та дистанційна освіта стали важливими компонентами сучасних освітніх систем.

Стрімкий розвиток штучного інтелекту дозволяє припускати, що тандем “вчитель – учень” вже може не бути настільки обов’язковим атрибутом освіти вже у досяжному майбутньому. Розвиток технологій вніс свою частку у системну кризу освіти, і він же і підштовхує на певні шляхи розв’язку, цієї

кризи. Окрім технологічного шляху корекції кризи, технологічний прогресс (та процеси глобалізації) підштовхнув і адміністративну ділянку освітньої системи; освітні системи почали фокусувати зусилля на таких навичках (умовно назвемо ці фокусовані напрями навичками 21-го століття), як критичне мислення, креативність, співпраця та спілкування; цей період освіти заснований на "компетенціях" і персоналізованому навчанні, що набули популярності під гаслом "пристосування освіти до індивідуальних потреб студента".

Еволюція освітніх систем є безперервним процесом під впливом культурних змін, технологічного прогресу, освітніх досліджень і суспільних потреб. Хочемо звернути увагу (і таким цином підвести риску), що у історичній ретроспективі сучасний стан освіти належить до відносно маленького часового відрізку (менше двухсот років), через це можна зробити висновок про надзвичайно високу динаміку цього процесу. Звідси постійна нестабільність освітньої системи в цілому. До тепер різні країни та регіони переживають власні унікальні траєкторії, що призвело до появи різноманітних освітніх моделей і практик у всьому світі.

Сучасні освітні системи стикаються з різними проблемами, які можуть вплинути на ефективність викладання та навчання. Те що більшість цих проблем носять поширений та глобальний характер свідчить про наявну системну кризу в освіті. Коротко про поширені (глобальні) проблеми:

1. Відсутність індивідуалізації: наявним системам освіти важко пристосуватися до різноманітних стилів і темпів навчання окремих учнів.

2. Інтеграція технологій, а саме їх несистемне впровадження: хоча технології можуть покращити навчальний досвід, їх інтеграція в навчальний процес може бути нерівномірною, а отже не системною.

3. Відповідність реальним навичкам – одне з критичних зауважень до сучасних систем освіти не відповідність отримуваних освітніх практичних навичок та знань викликам реального світу, наприклад фінансова грамотність, або навички вирішення проблем у спілкуванні та співпраці.

4. Нерівність і доступність: зберігаються розбіжності в освітніх ресурсах, фінансуванні та доступі до якісної освіти – наслідок соціальної та економічної нерівності, оскільки студенти з недостатньо успішних сімей можуть не отримати тих самих можливостей, що й їхні більш привілейовані однолітки.

5. Тиск і психічне здоров'я: Академічний тиск, конкуренція та високі очікування.

6. Прогалина в інноваціях: сектор освіти може повільно впроваджувати інноваційні методи та підходи до навчання, великою мірою через догматичність системи в цілому.

7. Глобалізація та культурне розмаїття: у світі, що дедалі більше глобалізується, зростає потреба у супроводі культурного розмаїття та забезпеченні освіти, яка є актуальною та включає різні точки зору та досвідів.

Крім того, окремо і коротко окреслимо деякі глобальні проблеми, що у національній системі мають (на нашу думку) вагому частку:

1. Стандартизоване тестування та надмірне значення оцінок: багато освітніх систем зловживають стандартизованим тестуванням та оцінками, як мірилом успішності учнів.

2. Закостенілість навчальної системи: фіксовані та негнучкі навчальні програми можуть (як правило) не встигати за швидкими темпами суспільних змін і прогресу в різних сферах.

3. Проблема низької конкурентності національних ВУЗів, кажучи простіше, їх забагато. За інформацією Держстату, станом на 2018-2019 навчальні роки в Україні налічувалося 282 університети, академії та інститути, де навчалися 1,3 млн студентів. Порівнюючи з кількістю ВУЗів у розвинутих країнах Європи, в Україні (на душу населення, якщо рахувати 40 мільйонів) вузів більше в сім разів, ніж в Великій Британії і Німеччині, в 5 разів більше ніж у Франції.

Навіть, якщо не докладати жодних зусиль до подолання кризи у системі освіти, система, яка має таку довгу історію еволюціює на наступний рівень, або це призведе до певних революційних змін. Всі ми спостерігаємо історичні зміни, які можна назвати перебудовою світового порядку, і це не зможе не вплинути на подальші зміні і у освіті. Спрогнозувати їх важка задача, але можна підготуватися до чих змін, або допомагати позитивним тенденціям.

Деякі тенденції можна оговорювати як такі, що фактично оформилися. Наприклад, явним є надання уподобань молодшого покоління до смартфонів, та застосування їх більшою мірою у виконанні задач, які ще 5 років перед тим були прерогативою персональних комп'ютерів. Так само можна констатувати їх більш "прохолодне" відношення до телебачення, та й взагалі до формату великих екранів. Оскільки глобальний ринок продажу смартфонів за 2022 рік показав приріст у 12%, а продаж телевізорів за останні 4 роки впали на 4%, можна опосередковано зробити висновок, що це прояви глобальних тенденцій.

Ще варто згадати AI (Artificial intelligence) – штучний інтелект, який прямо зараз вривається у багато різних аспектів людського життя, і вже виявив вплив на освітній процес. Хоча і рано ще робити прогно-

зи про вплив саме штучного інтелекту на освіту, але вже можна побачити перспективи застосування цього явища, і саме у освіті він може позитивно себе проявити.

Зробимо висновки. Навіть не змінюючи надмірне застосування тестових систем та оцінок, можна перенести акцент з тестування на персональних комп'ютерах до тестувань на смартфонах. В цьому є явні плюси, оскільки смартфон є готовим засобом ідентифікації користувача, що є важливим, при проведенні тестів та екзаменів. Крім того, використання смартфону сьогодні вже є еквівалентно участі у мінімум одній соціальній мережі. Тобто вже є готові і звичні для молоді програмні рішення для підвищення інклюзивності освітнього процесу.

Список використаних джерел:

1. Alenezi M. Deep Dive into Digital Transformation in Higher Education Institutions. Education Sciences. 2021. 11(12). 770. URL: https://doi.org/10.3390/educsci11120770

2. Burbules N.C., Fan G., Repp P. Five trends of education and technology in a sustainable future. Geogr. Sustain. 2020. 1. 93–97. URL: https://www.sciencedirect.com/science/article/pii/S2666683920300213?via%3Dihub

3. Grimus M. Emerging technologies: Impacting learning, pedagogy and curriculum development. In *Emerging Technologies and Pedagogies in the Curriculum*; Springer: Berlin/Heidelberg, Germany. 2020. P. 127–151. URL: https://www.researchgate.net/publication/338349901_Emerging_Technologies_Impacting_Learning_Pedagogy_and_Curriculum_Development

4. Haleem A., Javaid M., Qadri M. A., Suman R. Understanding the role of digital technologies in education: A review. Sustainable Operations and Computers. 2022. 3. P. 275-285. URL: https://doi.org/ 10.1016/j.susoc.2022.05.004

5. Heine S., Krepf M., Kunig J. Digital resources as an aspect of teacher professional digital competence: One term, different definitions – a systematic review. Education and Information Technologies. 9. 2022. URL: https://doi.org/10.1007/s10639-022-11321-z

6. Hollister B., Nair P., Hill-Lindsay S., Chukoskie L. Engagement in Online Learning: Student Attitudes and Behavior During COVID-19. Frontiers in Education. 2022. URL: https://doi.org/10.3389/feduc.2022.851019

7. Li Y., Kim M., Palkar J. Using emerging technologies to promote creativity in education: A systematic review. Int. J. Educ. Res. Open 2022, 3, 100177. URL: https://www.sciencedirect.com/science/article/pii/S266637402200053X

8. Vagg T., Balta J. Y., Bolger A., Lone M. Multimedia in education: What do the students think? Health Professions Education. 2020. 6. P. 325-333. URL: https://doi.org/10.1016/j.hpe.2020.04.011

9. Zhang K., Aslan A.B. AI technologies for education: Recent research & future directions. *Comput. Educ. Artif. Intell.* 2021. 2. URL: https://www.sciencedirect.com/science/article/pii/S2666920X21000199

10. Zhanga K., Aslan A. AI technologies for education: Recent research & future directions. Computers and Education: Artificial Intelligence. 2021. 2. 100025. URL: https://doi.org/10.1016/j.caeai.2021.100025

References:

1. Alenezi, M. (2021). Deep Dive into Digital Transformation in Higher Education Institutions. Education Sciences, 11(12), 770. Retrieved from https://doi.org/10.3390/educsci11120770

2. Burbules, N.C.; Fan, G.; Repp, P. Five trends of education and technology in a sustainable future. Geogr. Sustain. 2020, 1, 93–97. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666683920300213?via%3Dihub

3. Grimus, M. Emerging technologies: Impacting learning, pedagogy and curriculum development. In *Emerging Technologies and Pedagogies in the Curriculum*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 127–151. Retrieved from https://www.researchgate.net/publication/338349901_Emerging_Technologies_Impacting_Learning_Pedagogy_and_Curriculum_Development

4. Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. Sustainable Operations and Computers, 3, 275-285. Retrieved from https://doi.org/10.1016/j.susoc. 2022.05.004

5. Heine, S., Krepf, M., & Kunig, J. (2022). Digital resources as an aspect of teacher professional digital competence: One term, different definitions – a systematic review. Education and Information Technologies, 9. Retrieved from https://doi.org/10.1007/s10639-022-11321-z

6. Hollister, B., Nair, P., Hill-Lindsay, S., & Chukoskie, L. (2022). Engagement in Online Learning: Student Attitudes and Behavior During COVID-19. Frontiers in Education. Retrieved from https://doi.org/10.3389/feduc.2022.851019

7. Li, Y.; Kim, M.; Palkar, J. Using emerging technologies to promote creativity in education: A systematic review. Int. J. Educ. Res. Open 2022, 3, 100177. Retrieved from https://www.sciencedirect.com/science/article/pii/S266637402200053X

8. Vagg, T., Balta, J. Y., Bolger, A., & Lone, M. (2020). Multimedia in education: What do the students think? Health Professions Education, 6, 325-333. Retrieved from https://doi.org/10.1016/j.hpe.2020.04.011

9. Zhang, K.; Aslan, A.B. AI technologies for education: Recent research & future directions. *Comput. Educ. Artif. Intell.* 2021, *2*, 100025. Retrieved from https://www.sciencedirect.com/science/article/pii/S2666920X21000199

10. Zhanga, K., & Aslan, A. (2021). AI technologies for education: Recent research & future directions. Computers and Education: Artificial Intelligence, 2, 100025. Retrieved from https://doi.org/10.1016/j.caeai.2021.100025

УДК 004.912
DOI https://doi.org/10.32689/maup.it.2023.3.10

Яків ЮСИН
доктор філософії з інженерії програмного забезпечення, асистент кафедри програмного забезпечення комп'ютерних систем, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Берестейський проспект, 37, Київ, Україна, індекс 03056 (yusyn@pzks.fpm.kpi.ua)
ORCID: 0000-0001-6971-3808

Yakiv YUSYN
PhD in Software Engineering, Assistant of the Computer Systems Software Department, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Prospect Beresteiskyi, 37, Kyiv, Ukraine, postal code 03056 (yusyn@pzks.fpm.kpi.ua)

Бібліографічний опис статті: Юсин, Я. (2023). Автоматизація процесу аналізу та перевірки студентських кваліфікаційних робіт. *Інформаційні технології та суспільство*, 3, 72–79. DOI: https://doi.org/10.32689/maup.it.2023.3.10

Bibliographic description of the article: Yusyn, Y. (2023). Avtomatyzatsiia protsesu analizu ta perevirky studentskykh kvalifikatsiinykh robit [Automation of the process of analysis and checking of student qualification papers]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3, 72–79. DOI: https://doi.org/10.32689/maup.it.2023.3.10

АВТОМАТИЗАЦІЯ ПРОЦЕСУ АНАЛІЗУ ТА ПЕРЕВІРКИ СТУДЕНТСЬКИХ КВАЛІФІКАЦІЙНИХ РОБІТ

Анотація. Проблематика. *Поширеною формою атестації здобувачів вищої освіти є виконання кваліфікаційної роботи, невід'ємною складовою якої є її текстова частина. На сьогодні процес перевірки текстової частини є практично не автоматизованим і ґрунтується на зусиллях працівників випускних кафедр.* ***Мета дослідження.*** *Зменшити показник середньої кількості помилок у кваліфікаційних роботах за рахунок можливості виконання перевірки студентами самостійно, під час написання робіт, шляхом автоматизації процесу перевірки.* ***Методика реалізації.*** *Помилки, що можуть допускатися студентами при написанні кваліфікаційних робіт, розподілено на три групи: помилки технічного оформлення, помилки логічного оформлення, граматичні помилки. Для кожної групи розглянуто існуючі інструменти, які можуть запобігати та виявляти помилки цієї групи, і визначено ті класи помилок, на запобіганні та виявленні яких інформаційні системи автоматичної перевірки мають спеціалізуватися. В якості прикладу, за допомогою хмарних технологій, розроблено таку інформаційну систему, що реалізовує 9 правил виявлення помилок.* ***Результати дослідження.*** *Оцінювання ефективності автоматизації процесу перевірки кваліфікаційних робіт виконано на прикладі розробленої інформаційної системи за допомогою аналізу шістдесят чотирьох випускних робіт двох минулих років, що виконувались студентами двох різних структурних підрозділів. Для переважної більшості реалізованих правил (сім із дев'яти) отримано ненульову кількість істинних спрацювань. При цьому, кількість істинних спрацювань менше кількості хибних спрацювань лише для одного правила серед цих семи.* ***Висновки.*** *Експериментальна перевірка розробленої інформаційної системи доказала ефективність (за критерієм зменшення показника середньої кількості помилок) автоматизації процесу аналізу кваліфікаційних робіт, адже всі виявлені помилки могли би бути виявленими ще під час написання робіт студентами самостійно.*

Ключові слова: *кваліфікаційна робота; аналіз кваліфікаційних робіт; перевірка граматики; перевірка оформлення; нормоконтроль; хмарні технології; безсерверна архітектура.*

AUTOMATION OF THE PROCESS OF ANALYSIS AND CHECKING OF STUDENT QUALIFICATION PAPERS

Abstract. Background. *A common form of attestation for students of higher education is the performance of a qualification work, an integral component of which is its textual part. Today, the process of checking the text part is practically not automated and is based on the efforts of employees of graduate departments.* ***Objective.*** *To reduce the average number of errors in qualifying papers due to the possibility of self-checking by students during the writing of papers, by automating the checking process.* ***Methods.*** *Errors that can be made by students when writing qualification papers are divided into three groups: errors of technical design, errors of logical design, and grammatical errors. For each group, the existing tools that can prevent and detect errors of this group are considered, and those classes of errors are defined, on which prevention and detection of automatic verification information systems should be specialized. As an example, with the help of cloud technologies, such an information system was developed that implements 9 error detection rules.* ***Results.*** *The evaluation of the effectiveness of the automation of the process of checking qualification papers was carried out on the example of the developed information system through the analysis of sixty-four qualification papers of the past two years, which were performed by students of two different structural units. For the vast majority of implemented rules (seven out of nine), a non-zero number of true activations was obtained. At the same time, the number of true activations is less than the number of false activations for only one rule among these seven.* ***Conclusions.*** *The experimental verification of the developed information system proved the effectiveness (according to the criterion of reducing the average number of errors)*

of the automation of the process of analysis of qualifying papers because all the detected errors could have been detected even during the writing of the papers by the students themselves.

Key words: *qualification paper; analysis of qualification papers; spell checking; design checking; normative control; cloud technologies; serverless architecture.*

Вступ.

Однією із форм атестації здобувачів вищої освіти, що використовується в Україні, є виконання кваліфікаційної роботи. «Кваліфікаційна робота – це засіб діагностики ступеня сформованості компетентностей щодо вирішення типових завдань діяльності згідно з вимогами стандартів вищої освіти. Видами кваліфікаційних робіт є: дипломний проєкт, дипломна робота, магістерська дисертація» [1]. Виконання кваліфікаційної роботи передбачає «… систематизацію, закріплення, розширення теоретичних і практичних знань зі спеціальності та застосування їх при вирішенні конкретних наукових, технічних, економічних виробничих та інших завдань» [2]. Вибір форми атестації здобувачів вищої освіти у вигляді написання кваліфікаційної роботи визначається освітньою програмою, на основі вимог стандартів вищої освіти.

Обов'язковою складовою виконання кваліфікаційної роботи є написання та оформлення текстової частини – пояснювальної записки та іншої супровідної документації. На сьогодні, якість текстової частини (відповідність оформлення і змісту встановленим вимогам) забезпечується працею самого студента, контролем його керівника та нормоконтролюванням. Єдиним процесом забезпечення якості документації кваліфікаційної роботи, який наразі широко автоматизований, є процес перевірки дотримання принципів академічної доброчесності за допомогою спеціалізованих інформаційних систем [3; 4]. Всі інші процеси виконуються вручну, тому потребують багато людських зусиль та часу, є не захищеними від помилок. Також це призводить до великої затримки надання зворотного зв'язку студенту: проміжок часу між написанням текстової частини та виявленням помилок у ній в найкращому випадку складає дні, а в найгіршому випадку може сягати місяців (якщо певні помилки будуть знайдені лише під час завершального нормоконтролювання).

Таким чином, розроблення спеціалізованих інформаційних систем для автоматизації процесу перевірки кваліфікаційних робіт є актуальною задачею, вирішення якої підвищить якість робіт, а також зменшить навантаження на науково-педагогічних працівників підрозділів, які є керівниками та/або виконують нормоконтролювання.

Постановка задачі.

Метою роботи є удосконалення процесу перевірки студентських кваліфікаційних робіт шляхом автоматизації цього процесу, що зменшить показник середньої кількості помилок у роботах за рахунок можливості виконання перевірки студентами самостійно, під час написання робіт, та за рахунок зменшення людського фактору.

Теоретичні аспекти аналізу кваліфікаційних робіт.

Всі помилки, які можуть бути наявними у кваліфікаційних роботах, можливо розділити на три групи: помилки технічного оформлення, помилки логічного оформлення та граматичні помилки. Розглянемо кожну із груп помилок та програмні засоби, які можуть застосовуватися для запобігання та/або виявлення таких помилок.

До помилок технічного оформлення можливо віднести всі помилки, які пов'язані з порушенням технічних вимог до документації кваліфікаційної роботи – розмір та поля сторінок, відступи, шрифти, міжрядкові інтервали і так далі. Дані вимоги задаються галузевими стандартами, розробленими на їх основі положеннями окремих вищих навчальних закладів або їх підрозділів, тощо. Для запобігання таких помилок можуть застосовуватися шаблони оформлення кваліфікаційних робіт (наприклад, Word або TeX), які будуть містити правильно задані стилі всіх структурних елементів роботи, та які будуть надаватися студентам. Проте, як показує аналіз, практика підготовки таких шаблонів не є поширеною; також, надавання шаблонів не гарантує те, що студент буде його правильно використовувати чи використовувати взагалі – тому це потребує додаткової ручної перевірки.

До помилок логічного оформлення можливо віднести всі помилки, які пов'язані з порушенням встановлених вимог до оформлення документації кваліфікаційної роботи, але які не пов'язані з технічними вимогами. Наприклад, такими вимогами може бути наявність певних структурних елементів документації, їх іменування або присвоєння шифрів, наявність посилань у тексті роботи на всі джерела із списку використаних джерел, тощо. Існуючі на сьогодні програмні засоби здатні запобігати або виявляти лише незначну частку від можливих помилок цієї групи. Серед них можливо виділити наступне:

- Системи перевірки на плагіат можуть перевіряти правильність цитування джерел, наявність посилань на список використаних джерел, але це не є їх основною функцією та перевірка за допомогою таких систем виконується в кінці написання роботи.
- Деякі текстові процесори надають можливість керування джерелами, що включає у себе автоматичне створення та форматування бібліографії, посилань, підсвітку невикористаних джерел; також для цього існує спеціалізоване програмне забезпечення.
- Також текстові процесори можуть надавати інші функціональні можливості, що можуть запобігти виникненню помилок даної групи, такі як автоматичне створення змісту, підписів рисунків (з їх нумерацією), тощо.

Проте, на практиці забезпечити використання вище описаних функціональних можливостей текстових процесорів (або спеціалізованого програмного забезпечення) може бути неможливим з різних причин: відсутність підтримки у всіх текстових процесорах, несумісність наявного у студента апаратного чи програмного забезпечення з текстовим процесором, тощо.

До граматичних помилок можливо віднести всі помилки, які пов'язані з порушенням правил правопису (орфографії та пунктуації), стилістичні помилки. На сьогодні більшість текстових процесорів надають вбудовані можливості перевірки правопису, у тому числі українською мовою, а для систем комп'ютерної типографії (таких як TeX) існують відповідні розширення. Перевірка правопису у текстових процесорах ввімкнута за замовчуванням, тому вона здатна запобігти більшості помилок даної групи (хоча вона також може мати певні недоліки – наприклад, не відповідати останній прийнятій редакції правопису).

На основі даного огляду можемо зробити висновок, що у інформаційній системі аналізу кваліфікаційних робіт доцільно зосередитись на підтримці виявлення помилок перших двох груп. Також доцільною є підтримка виявлення помилок третьої групи, але не у повному обсязі: варто зосередитись на специфічних для предметної галузі правилах та правилах, які не підтримуються текстовими процесорами (наприклад, відповідність останній редакції правопису).

За наявності інформаційної системи аналізу та перевірки кваліфікаційних робіт, робочий процес дипломного керівника із студентом може будуватися наступним чином:

1. Студент завершує написання документації кваліфікаційної роботи або її частини (наприклад, певний розділ).
2. Студент самостійно перевіряє свій.docx файл з текстом за допомогою інформаційної системи аналізу.
3. Студент виправляє знайдені помилки.
4. Студент надсилає файл керівнику на перевірку.
5. Керівник встановлює факт перевірки студентом свого тексту.
6. Керівник перевіряє документацію фокусуючись на тих аспектах роботи, що автоматично не перевіряються.

Перед проходженням нормоконтролю, після завершення кваліфікаційної роботи, може відбуватися фінальна перевірка.

Огляд розробленої інформаційної системи аналізу та перевірки.

В якості прикладу інформаційної системи аналізу та перевірки кваліфікаційних робіт розглянемо систему, розроблену для робіт з ІТ спеціальностей.

Архітектура.

Розроблена інформаційна система реалізована у вигляді класичного вебзастосунку. Даний вибір обґрунтовано тим, що вебзастосунок взагалі не потребує окремого встановлення студентами (потрібен лише сучасний веббраузер), тому помилки сумісності з наявним у студента апаратним чи програмним забезпеченням виникнути не можуть. Крім того, у випадку вебзастосунку значно спрощується процес оновлення, оскільки є лише одне місце розгортання (сервер).

Інформаційна система розроблена з використанням хмарних технологій та безсерверної архітектури (так звані безсерверні обчислення). В контексті предметної галузі аналізу документації кваліфікаційних робіт, безсерверна архітектура має як технічні (спрощення апаратної інфраструктури та операцій з нею, спрощення розгортання), так і економічні переваги (оскільки сплачуються лише де-факто спожиті ресурси під час виконання аналізу) [5; 6].

Основна логіка аналізу реалізована на платформі.NET мовою програмування C#. Графічний інтерфейс користувача реалізовано за допомогою HTML/CSS, з інтерактивною логікою реалізованою на мові програмування JavaScript. Схематичне зображення розробленої системи, розбитої на компоненти, наведено на рис. 1.

Головною одиницею розбиття логіки аналізу та перевірки на компоненти є сервіс. В контексті розробленого програмного забезпечення, сервіс – це компонент, який реалізовує одне або декілька правил

аналізу, об'єднаних одним предметом, що аналізується. Наприклад, це може бути сервіс аналізу посилань, сервіс аналізу пунктуації, тощо. Сервіси між собою не взаємодіють, тому для спрощення повторного використання загального коду між сервісами, такий код винесено в окремі компоненти.

Кожному сервісу аналізу відповідає одна безсерверна функція, яка слугує публічно доступним інтерфейсом до нього. Кожна безсерверна функція запускається за допомогою HTTP-запиту, в тілі якого передається.docx файл для аналізу.

Рис. 1. Схематичне зображення компонентів розробленої системи

Враховуючи вищеописані особливості серверної частини, клієнт (графічний інтерфейс) не містить обчислювально складної логіки: лише викликає кожну безсерверну функцію (надсилаючи обраний користувачем.docx файл) та відображає отримані результати.

Серверна частина розробленої інформаційної системи розгорнута за допомогою хмарного провайдеру Azure [7], а клієнтську частину розгорнуто на сервісі розміщення статичних сторінок GitHub Pages [8]. Обидві частини розгортаються автоматично при внесені змін до головної гілки репозиторію коду, де ведеться розроблення інформаційної системи [9].

Реалізовані правила.

На момент написання даної роботи, розроблена інформаційна система аналізу та перевірки кваліфікаційних робіт з ІТ спеціальностей реалізовує дев'ять правил, згрупованих по семи сервісам.

Сервіс аналізу посилань. Сервіс аналізу посилань реалізовує одне правило REF01, яке перевіряє наявність посилань на використані джерела у тексті розділу чи кваліфікаційної роботи загалом. Правило підтримує посилання за стандартами ДСТУ [10] та APA [11], виконуючи їх пошук за допомогою регулярних виразів.

Сервіс аналізу розмітки документу. Даний сервіс реалізовує два правила LAY01 та LAY02, які перевіряють базові налаштування розмітки документів – встановлені розмір (відповідність формату А4) та поля сторінки (відповідність встановленим вимогам).

Сервіс аналізу правопису 2019-го року. Даний сервіс реалізовує правило ORT01, яке перевіряє відповідність тексту кваліфікаційної роботи змінам, що були внесені з новим українським правописом 2019-го року [12]. Доцільність реалізації даного правила зумовлена тим, що з новим правописом були внесені зміни у написання слів «проєкт» та префіксу «веб-», які дуже часто зустрічаються у документації кваліфікаційних робіт з ІТ спеціальностей. Пошук застарілого написання слів виконується за допомогою списку регулярних виразів.

Сервіс аналізу використання кальки англомовних слів. Даний сервіс реалізовує правило ORT02, яке перевіряє відсутність жаргонізмів у тексті кваліфікаційної роботи. Такі жаргонізми – утворені за допомогою кальки англомовних слів, наприклад, «деплой» – поширені в усній мові в індустрії ІТ, проте для письмової документації їх використання є неприпустимим. Пошук жаргонізмів реалізовано за допомогою списку регулярних виразів, що доповнюється при зустрічі нових форм на практиці.

Сервіс аналізу використання слів у правильному контексті. Даний сервіс реалізовує правило ORT03. Дане правило є подібним до правила ORT02, але тут перевіряються певні україномовні слова, які часто вживаються студентами у неправильному значені або контексті (наприклад, слово «додаток» у значенні «застосунок»). На даний час, це правило реалізовано тільки за допомогою регулярних виразів.

Сервіс аналізу змішування символів різних алфавітів. Даний сервіс реалізовує правило ORT04, яке перевіряє відсутність змішування символів різних алфавітів у одному слові (наприклад, викори-

стання «а» латиницею замість кирилиці). Таке змішування може виникати з різних причин (це необов'язково є спробою обманути системи перевірки плагіату), тому є доцільним включення такої перевірки до розробленої інформаційної системи, для попередньої перевірки. Дане правило реалізовано за допомогою регулярних виразів.

Сервіс аналізу пунктуації. Даний сервіс реалізовує два правила PNC01 та PNC02, які перевіряють використану у тексті кваліфікаційної роботи пунктуацію, а саме: коректне виділення розділових знаків пробілами; використання лапок-ялинок, які є найбільш поширеними в українській мові.

Експеримент.

Опис експерименту.

Для експериментальної перевірки ефективності розробленої інформаційної системи вирішено виконати перевірку робіт минулих років. Якщо така перевірка здатна виявити достатню кількість помилок у вже перевірених людьми роботах, то, відповідно, розроблене програмне забезпечення здатне виявити всі ці помилки в автоматичному режимі під час написання робіт, таким чином запобігши їх появі.

Недоліком такого експерименту є те, що розподіл виявлених помилок між правилами буде відрізнятися від того, який буде отриманий з використанням розробленої інформаційної системи під час написання кваліфікаційних робіт. Це пов'язано з тим, що певні типи помилок людині виявити легше, ніж інші; відповідно, у вже прийнятих роботах такі типи помилок будуть зустрічатися рідше. Проте, даний недолік не має великого значення, адже ціллю експерименту є принципова перевірка ефективності розробленого програмного забезпечення, тобто здатність знаходити помилки в цілому.

Для проведення експерименту використано актуальну на той момент версію розробленого програмного забезпечення 1.0.0-beta.2.2.1.

Експериментальні дані та їх підготовка.

В якості вхідних даних для проведення експерименту використано 64 бакалаврських дипломних проєктів, що виконувалися у 2021 та 2022 роках у двох різних структурних підрозділах.

Оскільки бакалаврські дипломні проєкти було представлено у форматі.pdf, аналіз якого не підтримується розробленою інформаційною системою (адже його метою є аналіз вихідних файлів під час написання кваліфікаційної роботи, а формат.pdf є кінцевим форматом для подання результатів), всі файли конвертовано у формат.docx за допомогою стороннього застосунку. Також з отриманих файлів. docx було видалено елементи, які не підлягають аналізу реалізованими правилами – титульні аркуші, списки використаних джерел, додатки, тощо.

Результати.

Отримана статистика по кількості спрацювань кожного із підтримуваних розробленою інформаційною системою правил наведено на рис. 2.

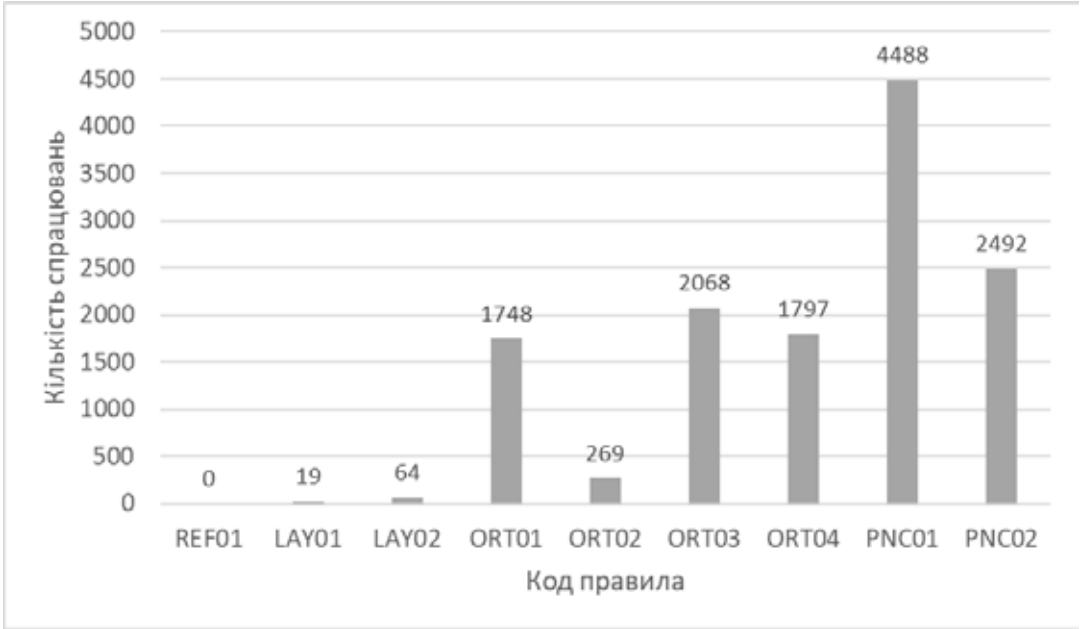


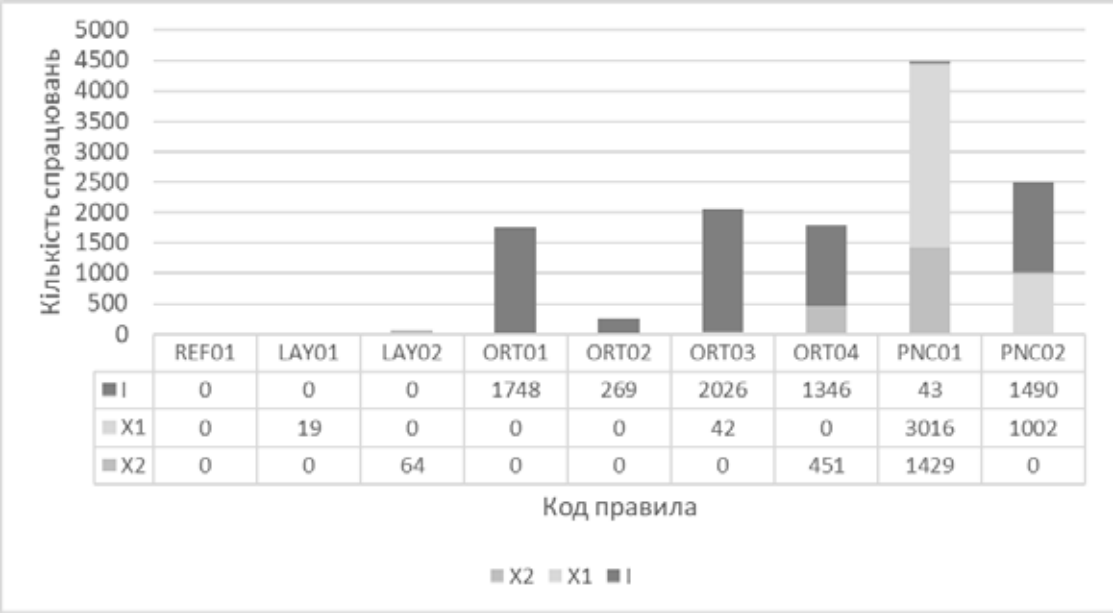
Рис. 2. Результати аналізу

Далі кожне спрацювання правил було проаналізовано і розподілено до однієї з трьох наступних категорій:

- істинне спрацювання (I);
- хибне спрацювання, викликане недоліками розробленої програмної системи (X1);
- хибне спрацювання, викликане недоліками вхідних даних (X2).

Наявність третьої категорії викликана тим, що дані для проведення експерименту конвертовано з кінцевих файлів формату.pdf, а не використовувалися вихідні файли формату.docx. Процес конвертації не завжди може правильно відтворити оригінальну структуру документу, тому за умови використання вихідних файлів формату.docx, цих спрацювань не було би отримано взагалі (детальний опис таких спрацювань наведено далі).

Отримані результати класифікації всіх спрацювань за описаними категоріями наведено на рис. 3.



	REF01	LAY01	LAY02	ORT01	ORT02	ORT03	ORT04	PNC01	PNC02
I	0	0	0	1748	269	2026	1346	43	1490
X1	0	19	0	0	0	42	0	3016	1002
X2	0	0	64	0	0	0	451	1429	0

Рис. 3. Розподіл отриманих спрацювань за трьома категоріями

Обговорення результатів.

Розглянемо отримані результати для кожного із реалізованих правил.

Відсутність посилань (REF01). Кількість отриманих спрацювань даного правила дорівнює нулю. Нульова кількість спрацювань даного правила на вже прийнятих роботах є очікуваною, адже відсутність у тексті посилань на використані джерела легко перевіряється людиною, і така робота не буде допущеною до захисту.

Розмір сторінки (LAY01). Перевірка робіт на відповідність розміру сторінки формату А4 спрацювала 19 разів. Після аналізу дані спрацювання віднесено до категорії «X1», тобто викликаних недоліками розробленої інформаційної системи. Це пов'язано з особливостями внутрішнього подання.docx файлу, у якому розміри сторінки зберігаються у двадцятих частинах пункту [13].

Оскільки одна двадцята частина пункту приблизно дорівнює 0.0017 сантиметра (1/1440 дюйму), розроблене програмне забезпечення повинно перевіряти відповідність розміру сторінки з досить великим допустимим відхиленням (у двадцятих частинах пункту) від значень, що відповідають розмірам формату А4.

Поля сторінки (LAY02). Дане правило спрацювало для кожної роботи із набору експериментальних даних. Всі спрацювання віднесено до хибних спрацювань, викликаних недоліками вхідних даних (категорія «X2»), оскільки, як показав аналіз, використане для конвертації з.pdf у.docx програмне забезпечення неправильно розпізнало поля документів.

Правопис 2019-го року (ORT01), використання кальки англомовних слів (ORT02). Всі отримані спрацювання двох правил є істинними.

Використання слів у неправильному контексті (ORT03). Більшість отриманих спрацювань даного правила є істинними. Невелика частина (близько 2%) спрацювань віднесено до категорії «X1», тобто викликаних недоліками розробленого програмного забезпечення.

Для виправлення зазначених хибних спрацювань, необхідно удосконалити використовуваний регулярний вираз для слова «додаток», відсікаючи посилання на додатки до роботи, що пронумеровані літерами.

Змішування символів різних алфавітів (ORT04). Близько 33% спрацювань даного правила викликано недоліками програмного забезпечення, використаного для конвертації файлів.pdf у.docx, які призвели до склеювання декількох слів у одне слово.

Виділення пунктуації (PNC01). Спрацювання, віднесені до категорії «X2», викликані вище описаним недоліком конвертації файлів у формат.docx, який призводить до склеювання слів. За їх виключенням, переважна більшість спрацювань (98%) віднесено до категорії «X1», у зв'язку з хибним спрацюванням правила на шифри робіт, що наводяться на кожній сторінці.

Використані лапки (PNC02). Близько 40% спрацювань даного правила віднесено до категорії «X1». Ці хибні спрацювання викликані лістингами коду, які автори робіт наводили у тексті – оскільки мови програмування підтримують лише один вид лапок для рядкових літералів, це призводить до хибних спрацювань. Виправити це можливо за допомогою додаткового автоматичного попереднього оброблення тексту роботи, що буде знаходити та видаляти лістинги коду.

Висновки.

Проведений аналіз поточного стану автоматизації процесу перевірки кваліфікаційних робіт показав доцільність розроблення спеціалізованих інформаційних систем для автоматизації цього процесу. У дальній перспективі такі інформаційні системи, у тому числі, можуть повністю замінити процес ручного нормоконтролювання для кваліфікаційних робіт.

В даному дослідженні розглянуто класифікацію помилок, що можуть допускатися студентами при написанні кваліфікаційних робіт, та можливість їх запобігання та виявлення існуючими інструментами. На даній основі сформульовано вимоги до інформаційних систем автоматичної перевірки кваліфікаційних робіт, а саме визначено ті класи помилок, на запобіганні та виявленні яких такі ІС мають спеціалізуватися.

В якості прикладу розроблено інформаційну систему аналізу кваліфікаційних робіт з ІТ спеціальностей, що реалізовує 9 правил виявлення помилок. Інформаційна система побудована у вигляді класичного вебзастосунку, з використанням хмарних безсерверних технологій. Такий вибір обґрунтований технічними (спрощення розгортання, відсутність необхідності в налаштуванні та моніторингу серверу), організаційними (студентам не потрібно встановлювати програмне забезпечення для аналізу, немає необхідності у вирішенні проблем сумісності) та економічними (більшість часу ІС не використовується, а безсерверна технологія дає можливість сплачувати лише за фактично спожиті ресурси) перевагами такої архітектури.

Проведена експериментальна перевірка ефективності розробленої інформаційної системи аналізу кваліфікаційних робіт полягала у виконанні аналізу кваліфікаційних робіт минулих років. Перевірка шістдесят чотирьох робіт, виконаних студентами протягом двох років у двох різних структурних підрозділах, виявила достатню кількість істинних спрацювань більшості реалізованих правил. Це означає, що використання спеціалізованої інформаційної системи ефективно виявило і запобігло би цим помилкам ще під час написання кваліфікаційних робіт.

Розроблена інформаційна система аналізу кваліфікаційних робіт має значний потенціал для подальшого розвитку, наприклад:

- зменшення кількості хибних спрацювань існуючих правил, наприклад, за рахунок розроблення механізму автоматичного виключення програмного коду з аналізу;
- розроблення нових правил для виявлення та запобігання помилок оформлення, граматичних помилок, у тому числі, зі застосуванням методів оброблення природної мови;
- впровадження нових типів аналізу – наприклад, виявлення текстів, написаних за допомогою засобів штучного інтелекту;
- розширення супровідної до аналізу функціональності – наприклад, автоматичне виправлення деяких помилок.

Список використаних джерел:

1. Положення про екзаменаційну комісію та атестацію здобувачів вищої освіти в КПІ ім. Ігоря Сікорського. *КПІ ім. Ігоря Сікорського*. Київ, 2020.

2. Положення про організацію освітнього процесу у Київському національному університеті імені Тараса Шевченка (друга редакція). *КНУ ім. Тараса Шевченка*. Київ, 2022.

3. Положення про систему виявлення та запобігання академічному плагіату у Київському національному університеті імені Тараса Шевченка. *КНУ ім. Тараса Шевченка*. Київ, 2020.

4. Положення про систему запобігання академічному плагіату в КПІ ім. Ігоря Сікорського. *КПІ ім. Ігоря Сікорського*. Київ, 2020.
5. Roberts M. Serverless Architectures // martinFowler.com. 2018. URL: https://martinfowler.com/articles/serverless.html.
6. Wen J. et al. Rise of the Planet of Serverless Computing: A Systematic Review // *ACM Transactions on Software Engineering and Methodology*. 2023. 5 (32). P. 1-61.
7. Azure Functions – Serverless Apps and Computing. *Microsoft Azure*. 2016. URL: https://azure.microsoft.com/en-us/services/functions/.
8. GitHub Pages. *GitHub*. 2009. URL: https://pages.github.com/.
9. GitHub. Features. GitHub Actions. *GitHub*. 2018. URL: https://github.com/features/actions.
10. ДСТУ 8302:2015. Інформація та документація. Бібліографічне посилання. *ДСТУ*. 2015.
11. Publication Manual of the American Psychological Association, Seventh Edition. *American Psychological Association*. 2020.
12. Український правопис. *Українська національна комісія з питань правопису*. 2019.
13. ISO/IEC 29500-1:2016. Information technology – Document description and processing languages – Office Open XML File Formats – Part 1: Fundamentals and Markup Language Reference. *ISO*. 2016.

References:

1. Igor Sikorsky Kyiv Polytechnic Institute. (2020). Polozhennia pro ekzamenatsiinu komisiiu ta atestatsiiu zdobuvachiv vyshchoi osvity v KPI im. Ihoria Sikorskoho [Regulations on the examination board and attestation of applicants for higher education at Igor Sikorsky Kyiv Polytechnic Institute]. Kyiv [in Ukrainian].
2. Taras Shevchenko Kyiv National University. (2022). Polozhennia pro orhanizatsiiu osvitnoho protsesu u Kyivskomu natsionalnomu universyteti imeni Tarasa Shevchenka (druha redaktsiia) [Regulations on the organization of the educational process at Taras Shevchenko Kyiv National University (second edition)]. Kyiv [in Ukrainian].
3. Taras Shevchenko Kyiv National University. (2020). Polozhennia pro systemu vyiavlennia ta zapobihannia akademichnomu plahiatu u Kyivskomu natsionalnomu universyteti imeni Tarasa Shevchenka [Regulations on the system of detection and prevention of academic plagiarism at Taras Shevchenko Kyiv National University]. Kyiv [in Ukrainian].
4. Igor Sikorsky Kyiv Polytechnic Institute. (2020). Polozhennia pro systemu zapobihannia akademichnomu plahiatu v KPI im. Ihoria Sikorskoho [Regulations on the system of prevention of academic plagiarism at Igor Sikorsky Kyiv Polytechnic Institute]. Kyiv [in Ukrainian].
5. Roberts, M. (2018). Serverless Architectures. martinFowler.com: https://martinfowler.com/articles/serverless.html.
6. Wen, J., Chen, Z., Jin, X., & Liu, X. (2023). Rise of the Planet of Serverless Computing: A Systematic Review. ACM Transactions on Software Engineering and Methodology, 32(5), pp. 1-61. https://doi.org/10.1145/3579643.
7. Microsoft. (2016). Azure Functions – Serverless Apps and Computing. Microsoft Azure: https://azure.microsoft.com/en-us/services/functions/.
8. GitHub. (2009). GitHub Pages. GitHub: https://pages.github.com/.
9. GitHub. (2018). Features. GitHub Actions. GitHub: https://github.com/features/actions.
10. DSTU. (2015). DSTU 8302:2015. Informatsiia ta dokumentatsiia. Bibliohrafichne posylannia [Information and documentation. Bibliographic reference] [in Ukrainian].
11. American Psychological Association. (2020). *Publication Manual of the American Psychological Association, Seventh Edition.* American Psychological Association.
12. Ukrainian National Commission on Orthography. (2019). Ukrainskyi pravopys [Ukrainian orthography]. [in Ukrainian].
13. ISO. (2016). ISO/IEC 29500-1:2016. Information technology – Document description and processing languages – Office Open XML File Formats – Part 1: Fundamentals and Markup Language Reference.

УДК 378+519.876.2+519.816
DOI https://doi.org/10.32689/maup.it.2023.3.11

Олексій ЧЕРНЮК
студент магістратури, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Навчально-науковий Інститут прикладного системного аналізу, проспект Берестейський, 37-А, м. Київ, Україна, індекс 03056 (cherniuk.oleksii@lll.kpi.ua)
ORCID: 0009-0008-7341-2934

Oleksii CHERNIUK
Master of Science Student, National Technical University of Ukraine"Igor Sikorsky Kyiv Polytechnic Institute", Educational and Research Institute for Applied System Analysis, 37-A Beresteyskyi Avenue, Kyiv, Ukraine, postal code 03056 (cherniuk.oleksii@lll.kpi.ua)
ORCID: 0009-0008-7341-2934

Бібліографічний опис статті: Чернюк, О. (2023). Моделювання впливу чат-ботів на основі штучного інтелекту на якість вищої освіти методами системного аналізу. *Інформаційні технології та суспільство*. Вип. 3(9), 80–90. DOI: https://doi.org/10.32689/maup.it.2023.3.11

Bibliographic description of the article: Cherniuk, O. (2023). Modeliuvannia vplyvu chat-botiv na osnovi shtuchnoho intelektu na yakist vyshchoi osvity metodamy systemnoho analizu [Modeling by methods of system analysis the impact of chatbots based on artificial intelligence on the quality of higher education]. *Informatsiini tekhnolohii ta suspilstvo – Information technology and society*, 3(9), 80–90. DOI: https://doi.org/10.32689/maup.it.2023.3.11

МОДЕЛЮВАННЯ ВПЛИВУ ЧАТ-БОТІВ НА ОСНОВІ ШТУЧНОГО ІНТЕЛЕКТУ НА ЯКІСТЬ ВИЩОЇ ОСВІТИ МЕТОДАМИ СИСТЕМНОГО АНАЛІЗУ

Дане дослідження дозволяє оцінити вплив генеративних чат-ботів на основі штучного інтелекту на якість вищої освіти, враховуючи як академічно доброчесну, так і академічно недоброчесну навчальні діяльності. Основна мета дослідження полягає в тому, щоб запропонувати ефективні стратегії для покращення якості вищої освіти в умовах стрімкого розвитку штучного інтелекту.

У дослідженні використано двоетапний метод модифікованого морфологічного аналізу та метод когнітивного моделювання. Ці системні підходи забезпечили комплексну основу для аналізу складних взаємодій між студентами, генеративними чат-ботами та освітніми практиками. Використання комбінованого підходу сприяло фундаментальному розумінню проблем, що розглядаються, дозволяючи більш ретельно вивчити потенційні рішення.

Унікальність даного дослідження полягає у вивченні конкретного впливу чат-ботів на якість вищої освіти та впровадженні конкретних стратегій для підвищення рівня освітнього процесу. Цікавою і ключовою особливістю є те, що для обох методів системного аналізу використовувались майже ідентичні базові дані. Незважаючи на використання різних підходів, результати обох методів виявилися дивовижно схожими, що відкриває шлях до комбінування та порівняння різних математичних методів системного аналізу для дослідження будь-яких інших галузей знань. Обидва методи вказали на майже ідентичні стратегії для покращення якості вищої освіти, що підтверджує високу якість розроблених моделей і дає можливість однозначно приймати ефективні рішення науково-педагогічному персоналу вищих навчальних закладів.

У результаті дослідження було побудовано ефективні стратегії для покращення якості вищої освіти шляхом мінімізації негативних наслідків, пов'язаних із зловживанням чат-ботами студентами, і максимізації навчальних, практичних і наукових переваг, які можна отримати від взаємодії учасників навчального процесу зі штучним інтелектом чат-ботів. Результатом моделювання є розробка двох взаємопов'язаних моделей, які виявили чисельну ієрархію ефективності освітніх втручань для покращення якості вищої освіти. Ця ієрархія дій може служити цінним інструментом для педагогів і бути рекомендованою для впровадження в систему вищої освіти.

Ключові слова: *якість вищої освіти, морфологічний аналіз, когнітивна карта, когнітивне моделювання, прийняття рішень, системний аналіз.*

MODELING BY METHODS OF SYSTEM ANALYSIS THE IMPACT OF CHATBOTS BASED ON ARTIFICIAL INTELLIGENCE ON THE QUALITY OF HIGHER EDUCATION

The given research allows assessing the impact of artificial intelligence-based generative chatbots on the quality of higher education, taking into account both academically ethical and academically unethical educational activities. The primary aim of the study is to propose effective strategies for enhancing the quality of higher education in the rapidly developing landscape of artificial intelligence.

The research employs a two-stage method of modified morphological analysis and cognitive modeling. These systematic approaches provide a comprehensive foundation for analyzing complex interactions among students, generative chatbots, and educational practices. The use of a combined approach contributes to a fundamental understanding of the addressed issues, enabling a more thorough exploration of potential solutions.

The uniqueness of this research lies in the examination of the specific influence of chatbots on the quality of higher education and the implementation of specific strategies to elevate the level of the educational process. An intriguing and key feature is that both methods of system analysis utilized nearly identical basic data. Despite employing different approaches, the results of both methods turned out remarkably similar, paving the way for combining and comparing various mathematical methods of system analysis in researching other knowledge domains. Both methods indicated almost identical strategies for improving the quality of higher education, affirming the high quality of the developed models and providing clear guidance for effective decision-making by the scientific and pedagogical staff of higher educational institutions.

As a result of the research, effective strategies were formulated to enhance the quality of higher education by minimizing the negative consequences associated with students' misuse of chatbots and maximizing the educational, practical, and scientific benefits derived from the interaction between participants in the educational process and artificial intelligence chatbots. The modeling outcome involves the development of two interconnected models that reveal a numerical hierarchy of the effectiveness of educational interventions for improving the quality of higher education. This hierarchy of actions can serve as a valuable tool for educators and is recommended for implementation in the higher education system.

Key words: *quality of higher education, morphological analysis, cognitive map, cognitive modeling, decision making, system analysis.*

1. Постановка проблеми та мета дослідження.

Вища освіта є критично важливим компонентом суспільного розвитку, що надає людям знання, навички та кваліфікацію, необхідні для особистого та професійного зростання. Одним із визначних аспектів галузі вищої освіти, яка розвивається, є інтеграція технологій в освітні процеси.

Чат-боти, що базуються на технологіях штучного інтелекту (ШІ) і обробки природної мови (NLP), стали цінними інструментами в різних сферах, включаючи вищу освіту. Чат-боти – це комп'ютерні програми, призначені для імітації людської розмови та надання автоматичних відповідей на запити користувачів. Значний технологічний прорив здійснили генеративні чат-боти на основі штучного інтелекту, такі як Chat GPT (OpenAI), Bard (Google), Bing Chat (Microsoft), Perplexity AI, YouChat, Chatsonic, Aria та багато інших. Хоча основною функцією генеративних чат-ботів є імітація співрозмовника-людини, вони можуть виконувати дуже багато завдань. Наприклад, писати і налагоджувати код програм; відповідати на тестові питання з поясненнями; генерувати бізнес-ідеї; писати вірші, твори, статті, тексти пісень; перекладати, переписувати та резюмувати текст; емулювати систему Linux; моделювати цілі чати; грати в такі ігри, як «хрестики-нулики»; змоделювати банкомат; надавати психологічні консультації; розпізнавати зображення; вирішувати математичні задачі та багато іншого [1-5].

Зв'язок чат-ботів на основі штучного інтелекту з вищою освітою є неминучим, тому існує потреба критично вивчити їхній вплив на якість освіти. Це дослідження прагне надати інформацію та рекомендації для навчальних закладів, які, хочуть цього чи ні, змушені взаємодіяти з новою реальністю, коли студент стає одним цілим з чат-ботом, і стає не зрозуміло, де справжні знання студента, і в чому вони полягають.

2. Виклад основного матеріалу дослідження.

2.1. Шкала вимірювання взаємовпливу альтернатив.

Усі впливи вершин одна на одну відбуваються на інтервалі [-1; +1]. Вагу кожного впливу будемо розуміти так (таб. 1):

Таблиця 1

Міри впливу

Дуже сильно негативний вплив	**[-1; -0.8)**
Сильно негативний вплив	[-0.8; -0.6)
Досить негативний вплив	[-0.6; -0.4)
Помірно негативний вплив	[-0.4; -0.2)
Легко негативний вплив	[-0.2; -0)
Прямого впливу немає	{0}
Легко позитивний вплив	(0; +0.2]
Помірно позитивний вплив	(+0.2; +0.4]
Досить позитивний вплив	(+0.4; +0.6]
Сильно позитивний вплив	(+0.6; +0.8]
Дуже сильно позитивний вплив	(+0.8; +1]

2.2. Модель на основі методу двоетапного модифікованого морфологічного аналізу.

2.2.1. Опис параметрів морфологічних таблиць.

Для дослідження нашої тематики застосуємо метод двоетапного модифікованого морфологічного аналізу (ММА) [6-9]. Експериментальна модель об'єкта буде заснована на морфологічних таблицях з двох характеристичних параметрів, кожен з яких має свої альтернативні сценарії.

Усі параметри морфологічної таблиці є якісними за своєю природою, тобто альтернативи таких параметрів принципово (якісно) відрізняються між собою, і між такими альтернативами неможливо встановити відношення переваги, як для кількісних параметрів.

Усі параметри є релевантними, тобто параметр повинен залежати або впливати на хоча б один інший параметр. У рамках деталізації, яка обрана для задачі, з параметрів, на основі яких визначається вплив на якість вищої освіти, було вилучено усі нерелевантні альтернативи.

У державних та університетських документах, підручниках та інших джерелах [10-12] можна знайти декілька десятків параметрів для оцінки якості вищої освіти, але в контексті нашої задачі було підібрано тринадцять релевантних показників ефективності студентської навчальної діяльності (таб. 2).

Таблиця 2

Морфологічна таблиця першого етапу ММА (морфологічна таблиця сценаріїв)

Вплив чат-ботів на основі штучного інтелекту на якість вищої освіти	
Мета використання чат-ботів студентами	Студентські показники ефективності навчальної діяльності
1.1. Списування і халатність у навчанні	2.1. Оцінки студентів
1.2. Саморозвиток, навчальна комунікація, удосконалення навичок	2.2. Розуміння навчального матеріалу
	2.3. Базові практичні навички, здобуті під час навчання
	2.4. Навички критичного, творчого та незалежного мислення
	2.5. Запам'ятовування інформації
	2.6. Швидкість виконання навчальних завдань
	2.7. Дослідницька майстерність
	2.8. Відвідуваність занять та активність на них
	2.9. Здатність ефективно передавати думки усно і письмово
	2.10. Співпраця та командна робота
	2.11. Адаптивність та стійкість до нових ситуацій та викликів
	2.12. Організація часу (тайм-менеджмент)
	2.13. Технологічна компетентність

На основі джерел про генеративні чат-боти на основі ШІ [3-5] та експертного актуального навчального студентського досвіду було відібрано дев'ять способів контролю якості вищої освіти (таб. 3) в контексті використання студентами чат-ботів. Альтернативи параметрів є взаємовиключними, що було враховано при проведенні процедури експертного оцінювання початкових наближень параметрів при експертному оцінюванні попарними порівняннями. Множина альтернатив є повною.

Таблиця 3

Морфологічна таблиця другого етапу ММА (морфологічна таблиця стратегій)

Способи контролю якості вищої освіти в контексті використання студентами чат-ботів
3.1. Використання чат-ботів в навчальних програмах предметів
3.2. Заборона використання гаджетів під час очних контрольних робіт і суворий контроль
3.3. Усні опитування віч на віч
3.4. Оцінювання за творчими індивідуальними унікальними роботами
3.5. Використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт
3.6. Усунення гуманітарних або тестових завдань в якості способів оцінювання студентів
3.7. Вимога дотримання спеціального стилю оформлення робіт, який чат-боти не зможуть повторити
3.8. Написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя
3.9. Суворі часові обмеження для написання контрольних робіт

2.2.2. Перший етап ММА.

Для початку оцінимо початкові ймовірності альтернатив. Головна мета першого етапу морфологічного аналізу – отримати початкові наближення $p_j^{'(i)}$ для ймовірностей кожної з альтернатив $a_j^{(i)}$ характеристичних параметрів. Для альтернатив $a_j^{(i)}$, $j \in \overline{1,n}$ параметра «Мета використання чат-ботів студентами» експертами надається оцінка $\widetilde{p_j^{(i)}}$. Для альтернативи «1.1» обрали значення 0.8, знаючи що в основному студенти використовують чат-боти для списування і халтури, для «2» відповідно значен-

ня 0.2. Для альтернатив параметра «Студентські показники ефективності навчальної діяльності» використали рівномірний розподіл, тобто однакову ймовірність для кожної його альтернативи. Оскільки ми не можемо апріорно отримати адекватні оцінки ймовірностей, використання експертної процедури для цього не є раціональним через значну невизначеність оцінок і через їх близькість. У такому випадку результат роботи МММА над параметром «Студентські показники ефективності навчальної діяльності» буде базуватись виключно на використанні матриці взаємозв'язків альтернатив параметрів і стане одним із важливих результатів даного дослідження.

Наведемо оцінену уже нормовану морфологічну таблицю (таб. 4).

Таблиця 4

Початкові ймовірності альтернатив

Мета використання чат-ботів студентами		**Студентські показники ефективності навчальної діяльності**	
Номер альтернативи	Ймовірність альтернативи	Номер альтернативи	Ймовірність альтернативи
1.1	0.8	2.1	0.076
1.2	0.2	2.2	0.077
		2.3	0.077
		2.4	0.077
		2.5	0.077
		2.6	0.077
		2.7	0.077
		2.8	0.077
		2.9	0.077
		2.10	0.077
		2.11	0.077
		2.12	0.077
		2.13	0.077

Далі оцінимо матрицю взаємозв’язків параметрів першого етапу.

Для врахування зв’язків між параметрами морфологічної таблиці (МТ) на основі таблиці 1 побудуємо числову матрицю взаємозв'язків альтернатив параметрів (таб. 5).

Таблиця 5

Матриця взаємозв’язків параметрів першого етапу

	1.1	*1.2*
2.1	0.4	0.5
2.2	-0.8	0.3
2.3	-0.5	0.15
2.4	-0.3	0.15
2.5	-0.7	-0.2
2.6	0.8	0.2
2.7	-0.7	-0.2
2.8	-0.4	-0.2
2.9	-0.25	0.05
2.10	-0.35	-0.1
2.11	0.4	0.4
2.12	0.15	0.3
2.13	0.5	0.8

Нарешті, проведемо процедури з розрахунку ймовірностей альтернатив і конфігурацій. Щоб отримати остаточні значення ймовірності, необхідно розв'язати задачу розрахунку ймовірностей альтернатив параметрів. Подібні розрахунки повторюємо для всіх 26 конфігурацій, отримуємо результат першого етапу морфологічного аналізу – оцінки ймовірностей альтернатив з урахуванням зв’язків між ними (таб. 6).

Таблиця 6

Результат першого етапу МММА

Мета використання чат-ботів студентами		**Студентські показники ефективності навчальної діяльності**	
Номер альтернативи	Ймовірність альтернативи	Номер альтернативи	Ймовірність альтернативи
1.1	0.74806734	2.1	0.11668415
1.2	0.25193266	2.2	0.03496632
		2.3	0.05244948
		2.4	0.06576998
		2.5	0.03330126
		2.6	0.13986528
		2.7	0.03330126
		2.8	0.05328201
		2.9	0.06743505
		2.10	0.0582772
		2.11	0.1165544
		2.12	0.09823871
		2.13	0.1298749

2.2.3. Другий етап МММА.

Специфіка другого етапу МММА полягає в тому, що вибір альтернатив параметрів морфологічної таблиці стратегій залежить не від випадкових зовнішніх факторів, а від особи, що приймає рішення. Тому на другому етапі для оцінки альтернатив і конфігурацій використовується величина очікуваної результативності, тобто вірогідності того, що вибір цієї альтернативи або конфігурації призведе до бажаних результатів.

Матриця зв'язків, яка співставляє кожну пару альтернатив першого і другого етапів наведена в таблиці 7.

Таблиця 7

Матриця зв'язків альтернатив першого та другого етапів

	3.1	**3.2**	**3.3**	**3.4**	**3.5**	**3.6**	**3.7**	**3.8**	**3.9**
1.1	0.5	1	1	0.65	0.9	0.6	0.15	1	0.4
1.2	0.5	0	0	0	0	0	0.1	0	0
2.1	0	-0.5	-0.4	0	-0.4	0	-0.1	-0.5	-0.5
2.2	0.3	0.5	0.4	0.6	0.3	-0.1	0	0.5	0.3
2.3	-0.25	0.7	0.3	0.5	0.3	-0.1	0	0.5	0.3
2.4	-0.15	0.2	0.3	0.6	0.1	0.2	0	0.2	-0.45
2.5	-0.35	0.5	0.3	0.15	0.3	0.2	0	0.4	0.2
2.6	0.7	-0.5	0	-0.25	0	-0.2	-0.2	-0.4	0.5
2.7	-0.1	-0.2	0	0.4	0	0.3	0	0	0
2.8	0.2	0.5	0.8	-0.2	0	0	0.1	0	0
2.9	-0.2	0	0.45	0.4	0	-0.25	0	0	0
2.10	0.2	0.3	0.15	0	-0.1	0	0.1	0	-0.4
2.11	0.6	0.35	0.4	0.1	0.1	0	0.1	0.25	0.45
2.12	0.55	0.3	0	0.25	0.3	0	0	0.2	0.7
2.13	0.85	-0.1	0	0.4	0.15	0	0.15	0.05	0

У результаті процедури з розрахунку оцінок альтернатив другого етапу отримуємо таблицю 8.

Таблиця 8

Розраховані на другому етапі MMMA оцінки альтернатив

3.1. Використання чат-ботів в навчальних програмах предметів	**0.13713944**
3.3. Усні опитування віч на віч	0.129650906
3.4. Оцінювання за творчими індивідуальними унікальними роботами	0.116804045
3.2. Заборона використання гаджетів під час очних контрольних робіт і суворий контроль	0.116340931
3.5. Використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт	0.114872605
3.8. Написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя	0.112691262
3.6. Усунення гуманітарних або тестових завдань в якості способів оцінювання студентів	0.096571773
3.9. Суворі часові обмеження для написання контрольних робіт	0.096144714
3.7. Вимога дотримання спеціального стилю оформлення робіт, який чат-боти не зможуть повторити	0.079784324

2.2.4. Результати моделювання MMMA.

У результаті першого етапу MMMA отримали початкові наближення для ймовірностей кожної з альтернатив «1.1-1.2», «2.1-2.13». З отриманих даних можна зробити висновки, що приблизно три використання студентами генеративного чат-бота із чотирьох відбувається для списування і халтури. Було отримано різноманітну ієрархію показників студентської ефективності навчальної діяльності саме в контексті використання чат-ботів студентами (рис. 1). Пам'ятаємо що вибіркове середнє рівномірного розподілу появи будь-якого параметра альтернативи «Студентські показники ефективності навчальної діяльності» дорівнює 0,077. Згідно з таблицею 8 можна зробити висновок, що використання чат-ботів студентами позитивно вплинуло на швидкість виконання навчальних завдань, технологічну компетентність, оцінки студентів, адаптивність та стійкість до нових ситуацій, організацію часу. І негативно – на запам'ятовування інформації, дослідницьку майстерність, розуміння навчального матеріалу, базові практичні навички, відвідуваність занять та активність на них, співпрацю та командну роботу, навички критичного, творчого та незалежного мислення, здатність ефективно передавати думки усно і письмово. Порівнявши кожне числове значення ефективності з початковим вибірковим середнім рівномірного розподілу, можна порахувати приблизно, у скільки разів покращилася або погіршилася ситуація з відповідним показником студентської ефективності.

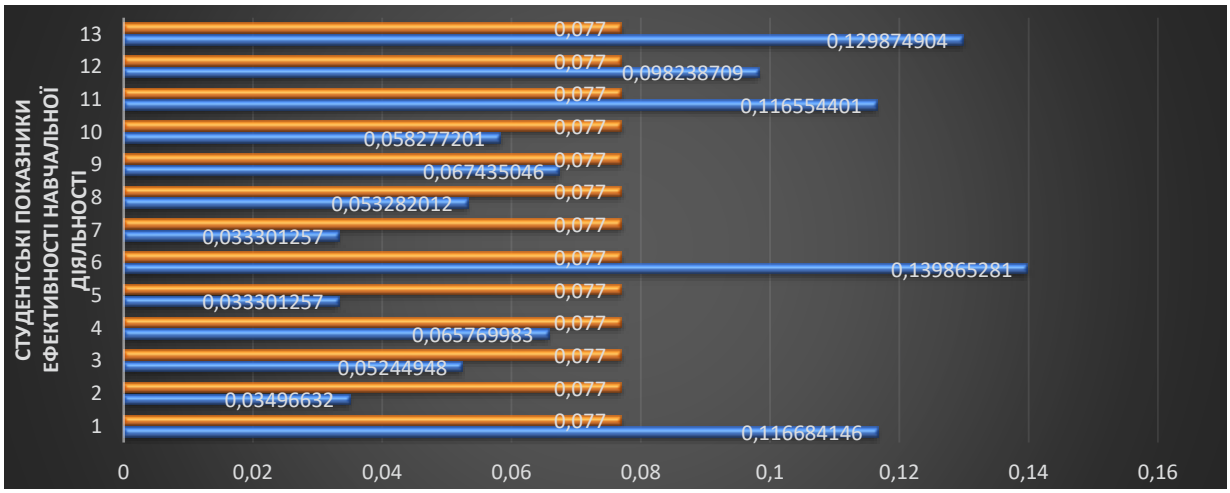


Рис. 1. Порівняння ймовірностей альтернатив за рівнем позитивно-негативного впливу генеративних чат-ботів на студентські показники ефективності навчальної діяльності

У результаті другого етапу MMMA отримали рейтинг способів контролю якості вищої освіти (рис. 2).

Рис. 2. Порівняння ймовірностей результативності альтернатив для способів контролю якості вищої освіти

З діаграми можемо бачити, що майже усі величини очікуваної результативності близькі до вибіркового середнього 0.(11), тобто усі способи мають деякий позитивний вплив на якість вищої освіти в умовах використання чат-ботів студентами. Проте можемо виділити 6 найефективніших способів контролю якості вищої освіти в контексті використання студентами чат-ботів, а саме: використання чат-ботів в навчальних програмах предметів, усні опитування віч на віч, оцінювання за творчими індивідуальними унікальними роботами, заборона використання гаджетів під час очних контрольних робіт і суворий контроль, використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт, написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя. Тобто 4 із 6 найефективніших способів є способами-обмежувачами, які спрямовані на посилений контроль студента. Спосіб використання чат-ботів в навчальних програмах предметів набрав високий рівень ефективності через те, що на відміну від інших способів він має значний вплив на альтернативу «1.2»: студенти активно використовують чат-боти на основі штучного інтелекту для саморозвитку, навчальної комунікації і удосконалення своїх навичок, тобто таке рішення не лише допомагає зменшити шкідливий вплив чат-ботів, а також надає можливість отримувати від них значну користь.

2.3.Когнітивне моделювання [13-16].

Вершини когнітивної карти (КК) повністю співпадають з альтернативами параметрів методу ММА з додаванням однієї цільової вершини «4. Якість вищої освіти», нумерація вершин також співпадає (рис. 3).

Взаємовпливи між вершинами КК засновані на даних з таблиць 5, 8. До матриці суміжності КК (рис. 4) були додані взаємодії параметрів 2.1 – 2.13 з параметром 4.

Усі власні числа отриманої матриці суміжності рівні 0, тому дана КК є імпульсно стійкою і стійкою за значеннями (абсолютно). Над отриманою когнітивною моделлю було проведено експериментальне тестування шляхом направлення на кожну керуючу вершину 3.1 – 3.9 одиничного додатного імпульсу (рис. 5).

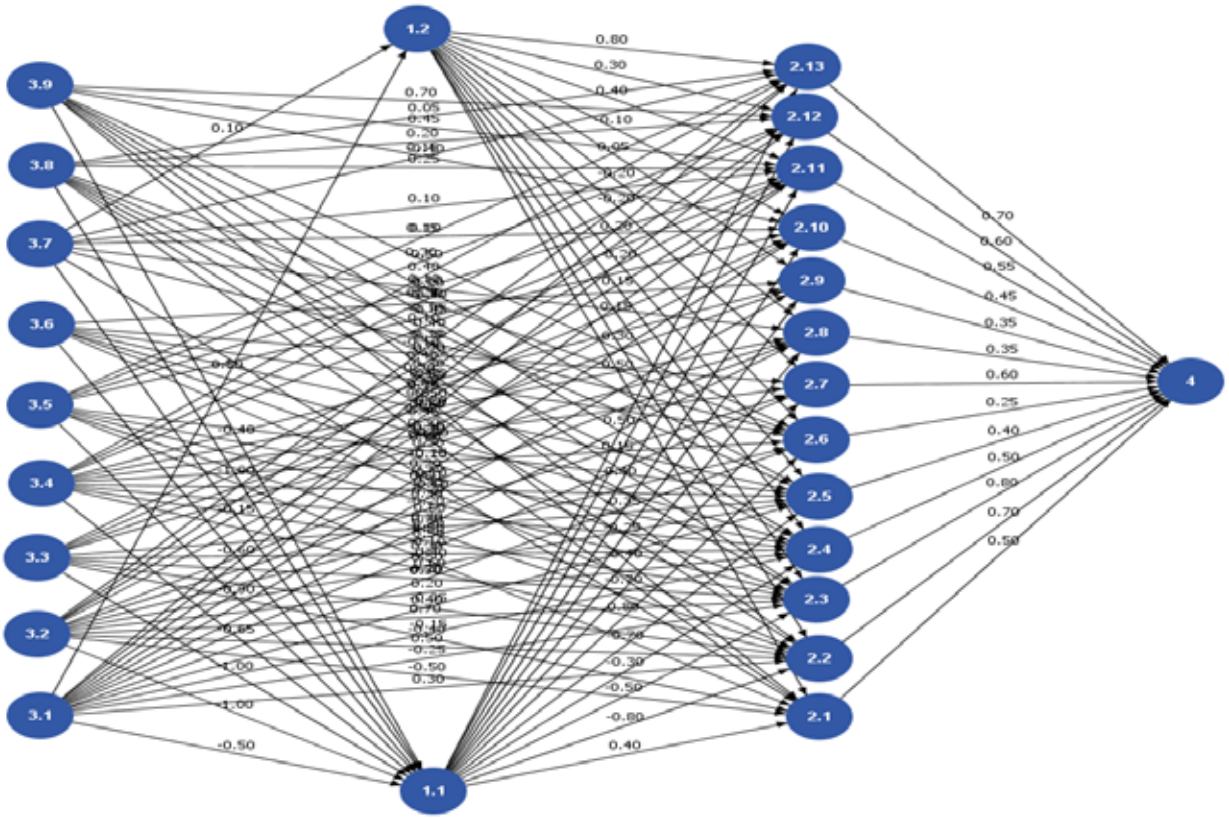


Рис. 3. Когнітивна карта впливу способів контролю якості вищої освіти на студентські показники ефективності навчальної діяльності, які в свою чергу впливають на цільовий параметр «Якість вищої освіти»

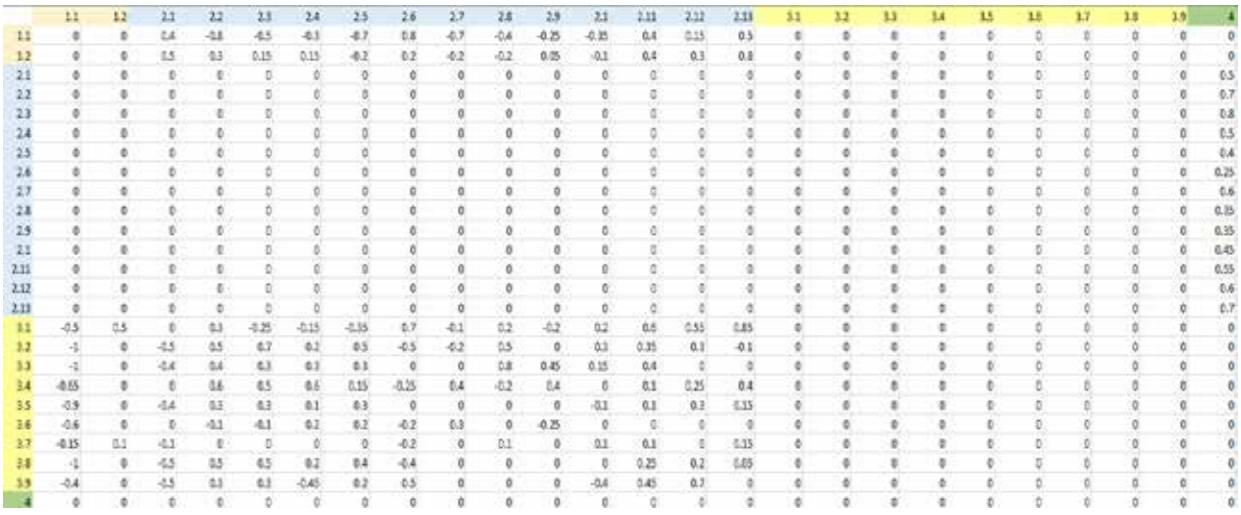


Рис. 4. Матриця суміжності КК

Отже, створена когнітивна модель реагує на зовнішні зміни адекватно, адже при моделюванні вона показала цілком реалістичні закономірності. За результатами моделювання можемо виділити 5 найефективніших способів контролю якості вищої освіти, а саме: використання чат-ботів в навчальних програмах предметів, усні опитування віч-на-віч, заборона використання гаджетів під час очних контрольних робіт і суворий контроль, використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт, написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя.

Рис. 5. Ієрархія впливу способів контролю якості вищої освіти в контексті використання студентами чат-ботів на основі ШІ

3. Результати дослідження.

Результати роботи двох методів для порівняння і наочності були зведені в єдину таблицю (таб. 10), в якій можемо бачити ієрархію способів максимізації позитивного впливу чат-ботів на основі ШІ на якість вищої освіти.

Таблиця 10

Порівняння результатів моделювання методами морфологічного аналізу та когнітивного моделювання

Ієрархія пріоритетності	*Метод двоетапного модифікованого морфологічного аналізу*	*Метод когнітивного моделювання*
1	Використання чат-ботів в навчальних програмах предметів	Використання чат-ботів в навчальних програмах предметів
2	Усні опитування віч на віч	Усні опитування віч на віч
3	Оцінювання за творчими індивідуальними унікальними роботами	Написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя
4	Заборона використання гаджетів під час очних контрольних робіт і суровий контроль	Заборона використання гаджетів під час очних контрольних робіт і суровий контроль
5	Використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт	Використання спеціального контролюючого програмного забезпечення в умовах дистанційного навчання під час контрольних робіт
6	Написання контрольних робіт від руки перед камерою, яка фіксує руки та обличчя	Оцінювання за творчими індивідуальними унікальними роботами
7	Усунення гуманітарних або тестових завдань в якості способів оцінювання студентів	Усунення гуманітарних або тестових завдань в якості способів оцінювання студентів
8	Суворі часові обмеження для написання контрольних робіт	Суворі часові обмеження для написання контрольних робіт
9	Вимога дотримання спеціального стилю оформлення робіт, який чат-боти не зможуть повторити	Вимога дотримання спеціального стилю оформлення робіт, який чат-боти не зможуть повторити

4. Висновки

Отримані результати є досить реалістичними, хоча і засновані на суб'єктивній оцінці. Дане дослідження є досить наочним і актуальним, адже побудоване на оцінці студента, який сам знаходиться у розглянутій ситуації і знає, як це працює зсередини. Була розроблена модель, яка може допомогти проаналізувати способи контролю якості вищої освіти в контексті використання студентами генеративних чат-ботів на основі ШІ.

Новизна роботи полягає в тому, що було досліджено конкретну складову впливу чат-ботів на якість вищої освіти і застосовано конкретні рішення, які дозволять підвищити якість вищої освіти. Особливістю даної роботи є те, що в ній для двох різних методів системного аналізу було використано майже однакові базові дані. Результати виявилися вражаюче подібними, не дивлячись на різні способи їх отримання. Обидва методи показали майже однакові пріоритетні напрями для покращення якості вищої освіти. Такого результату вдалося досягти через експериментальне комбінування властивостей когнітивної і морфологічної моделей. Завдяки комплексному підходу визначення і оцінки параметрів результуючі моделі стали більш структурованими, репрезентативними, повноцінними, логічними, зрозумілими і адекватними.

Результати даного дослідження дозволять викладачам поставити пріоритети в тому, які методи взаємодії зі студентами будуть найбільш ефективними для покращення якості вищої освіти.

Ці моделі можна покращити шляхом уточнення початкових оцінок за допомогою великої кількості компетентних у даній області експертів, або за допомогою використання певних статистичних даних. У методі морфологічного аналізу можна додати нові контекстні параметри. У методі когнітивного моделювання можливо застосувати методи керування імпульсними процесами.

Список використаних джерел:

1. Abdullahi A. 10 Best AI Chatbots 2023. *eWEEK*. 2023. URL: https://www.eweek.com/artificial-intelligence/best-ai-chatbots/#comparison-chart (дата звернення: 10.11.2023).

2. ChatGPT. *Wikipedia, the free encyclopedia.* URL: https://en.wikipedia.org/wiki/ChatGPT (дата звернення: 15.11.2023).

3. Fengchun M., Wayne H. Guidance for generative AI in education and research. 7, place de Fontenoy, 75352 Paris 07 SP, France : *UNESCO*, 2023. 44 c. URL: https://unesdoc.unesco.org/ark:/48223/pf0000386693 (дата звернення: 13.11.2023).

4. Hulick K. How ChatGPT and similar AI will disrupt education. *ScienceNews*. 2023. URL: https://www.sciencenews.org/article/chatgpt-ai-artificial-intelligence-education-cheating-accuracy (дата звернення: 11.10.2023).

5. Kamalov F., Santandreu Calonge D., Gurrib I. New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. *Sustainability*. 2023. Т. 15, № 16. С. 12451. URL: https://doi.org/10.3390/su151612451 (дата звернення: 11.11.2023).

6. Морфологічний аналіз. Теорія, проблеми, застосування : навчальний посібник / Н.Д. Панкратова, І.О. Савченко ; М-во освіти і науки України, НТУУ "КПІ", Ін-т прикладного системного аналізу. Київ : Наукова думка, 2015. – 244 с.

7. Методологічне і математичне забезпечення розв'язання задач передбачення на основі модифікованого методу морфологічного аналізу / І.О. Савченко // Систем. дослідж. та інформ. технології. 2011. № 3. С. 18-28. Бібліогр. : 14 назв. укр. URL: http://dspace.nbuv.gov.ua/handle/123456789/50108 (дата звернення: 1.10.2023).

8. Pankratova, N., Haiko, Hennadii, Savchenko, Illia. Morphological model for underground crossings of water objects. *System research and information technologies*. 2021. 53-67. URL: https://www.researchgate.net/publication/358359378_Morphological_model_for_underground_crossings_of_water_objects (дата звернення: 2.10.2023).

9. Панкратова, Н. Д. Стратегія застосування методу морфологічного аналізу в процесі технологічного передбачення / Н. Д. Панкратова, І. О. Савченко // *Наукові вісті НТУУ «КПІ» : науково-технічний журнал*. 2009. № 2(64). С. 35–44. Бібліогр. : 16 назв.

10. Педагогіка вищої школи [Електронний ресурс] : підручник / В. П. Головенкін ; КПІ ім. Ігоря Сікорського. 2-ге вид., переробл. і доповн. Електронні текстові дані (1 файл: 3,6 Мбайт). Київ : КПІ ім. Ігоря Сікорського, 2019. 290 с.

11. Аннєнкова І. П. Критерії і показники якості освіти у ВНЗ. *Наука і освіта.* 2011. URL: https://scienceandeducation.pdpu.edu.ua/doc/2011/8_2011/1.pdf (дата звернення: 01.11.2023).

12. Гапон Л. О. Показники ефективності освітньої діяльності педагога. *Методист ТКМЦНОІМ Гапон Л. О. Блог учителів української мови і літератури міста Тернополя.* 19 лютого 2020. URL: https://gapon.te.ua/rubryka-metodysta/dorobok-metodysta/metodychni-rekomendatsii/item/1360-mekhanizm-pobudovy-u-zakladi-osvity-vnutrishnoyi-systemy-otsinyuvannya-yakosti-osvity (дата звернення: 21.10.2023).

13. Мілявський, Ю. Л. Ідентифікація та керування складними системами на основі моделей імпульсних процесів когнітивних карт : дис. … д-ра техн. наук. : 01.05.04 Системний аналіз і теорія оптимальних рішень / Мілявський Юрій Леонідович. Київ, 2021. 297 с. URL: https://ela.kpi.ua/handle/123456789/43829 (дата звернення: 5.10.2023).

14. Roberts F. Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems. Englewood Cliffs, Prentice-Hall, 1976. 559 p.

15. Романенко В. Д. Когнітивне моделювання динаміки прийняття рішень для стабілізації нестійких режимів у соціально-навчальному процесі студента / В. Д. Романенко, Ю. Л. Мілявський // *Наукові вісті НТУУ «КПІ» : науко-*

во-технічний журнал. 2016. № 5(109). С. 48–53. Бібліогр.: 10 назв. URL: https://ela.kpi.ua/handle/123456789/18873 (дата звернення: 7.10.2023).

16. В. Б. Мокін, О. В. Бурдейна, К. О. Коваль, А. Р. Ящолт. Метод проектування когнітивної карти для оптимізації профорієнтаційної діяльності ЗВО. *Вісник Вінницького політехнічного інституту*. 2018. № 3. С. 89-99. URL: https://www.researchgate.net/publication/330089909_METOD_PROEKTUVANNA_KOGNITIVNOI_KARTI_DLA_OPTIMIZACII_PROFORIENTACIJNOI_DIALNOSTI_ZVO (дата звернення: 14.10.2023).

References:

1. Abdullahi, A. (2023, September 14). *10 Best AI Chatbots 2023*. eWEEK. Retrieved from https://www.eweek.com/artificial-intelligence/best-ai-chatbots/#comparison-chart.

2. Wikipedia contributors. (2023, December 2). *ChatGPT*. Wikipedia. Retrieved from https://en.wikipedia.org/wiki/ChatGPT.

3. Fengchun M. & Wayne H. (2023). *Guidance for generative AI in education and research*. UNESCO. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000386693.

4. Hulick, K. (2023, May 10). *How ChatGPT and similar AI will disrupt education.* Science News. Retrieved from https://www.sciencenews.org/article/chatgpt-ai-artificial-intelligence-education-cheating-accuracy.

5. Kamalov, F., Calonge, D. S., & Gurrib, I. (2023). *New era of Artificial intelligence in Education: Towards a sustainable Multifaceted Revolution*. Sustainability, *15*(16), 12451. Retrieved from https://doi.org/10.3390/su151612451

6. N.D. Pankratova & I.O. Savchenko. (2015). *Morfolohichnyi analiz. Teoriia, problemy, zastosuvannia: navchalnyi posibnyk [Morphological analysis. Theory, problems, application: study guide]*. Kyiv: Naukova dumka [in Ukrainian].

7. I.O. Savchenko. (2011). Metodolohichne i matematychne zabezpechennia rozv'iazannia zadach peredbachennia na osnovi modyfikovanoho metodu morfolohichnoho analizu [Methodological and mathematical support for solving prediction problems based on a modified method of morphological analysis]. *Systemni doslidzhennia ta informatsiini tekhnolohii – System research and information technologies,* 3, 18-28. Retrieved from http://dspace.nbuv.gov.ua/handle/123456789/50108 [in Ukrainian].

8. Pankratova, N. D., Haiko, H., & Savchenko, I. (2021). Morphological model for underground crossings of water objects. *Systemni doslidzhennia ta informatsiini tekhnolohii,* 4, 53–67. Retrieved from https://doi.org/10.20535/srit.2308-8893.2021.4.04.

9. N.D. Pankratova & I.O. Savchenko. (2009). Stratehiia zastosuvannia metodu morfolohichnoho analizu v protsesi tekhnolohichnoho peredbachennia [The strategy of applying the method of morphological analysis in the process of technological prediction]. *Naukovi visti NTUU «KPI»: naukovo-tekhnichnyi zhurnal – Scientific news of NTUU "KPI": scientific and technical journal.* 2(64). 35-44 [in Ukrainian].

10. V. P. Holovenkin. (2019). *Pedahohika vyshchoi shkoly [Pedagogy of high school]*. Kyiv: KPI im. Ihoria Sikorskoho [in Ukrainian].

11. Annienkova I. P. (2011). Kryterii i pokaznyky yakosti osvity u VNZ [Criteria and indicators of the quality of education in universities]. *Nauka i osvita – Science and education.* Retrieved from https://scienceandeducation.pdpu.edu.ua/doc/2011/8_2011/1.pdf [in Ukrainian].

12. Hapon L. O. (2020, February 19). Pokaznyky efektyvnosti osvitnoi diialnosti pedahoha [Indicators of the effectiveness of the teacher's educational activity]. *gapon.te.ua*. Retrieved from https://gapon.te.ua/rubryka-metodysta/dorobok-metodysta/metodychni-rekomendatsii/item/1360-mekhanizm-pobudovy-u-zakladi-osvity-vnutrishnoyi-systemy-otsinyuvannya-yakosti-osvity [in Ukrainian].

13. Miliavskyi, Yu. L. (2021). Identyfikatsiia ta keruvannia skladnymy systemamy na osnovi modelei impulsnykh protsesiv kohnityvnykh kart [Identification and control of complex systems based on models of impulse processes of cognitive maps]. *Doctor's thesis*. Kyiv: National Technical University of Ukraine"Igor Sikorsky Kyiv Polytechnic Institute". Retrieved from https://ela.kpi.ua/handle/123456789/43829 [in Ukrainian].

14. Nevison, C., & Roberts, F. S. (1977). Discrete Mathematical Models with Applications to Social, Biological, and Environmental Problems. *The American Mathematical Monthly*, 84(10), 834. https://doi.org/10.2307/2322080.

15. V. D. Romanenko & Yu. L. Miliavskyi. (2016). Kohnityvne modeliuvannia dynamiky pryiniattia rishen dlia stabilizatsii nestiikykh rezhymiv u sotsialno-navchalnomu protsesi studenta [Cognitive modeling of decision-making dynamics to stabilize unstable regimes in the student's social-educational process]. *Naukovi visti NTUU «KPI»: naukovo-tekhnichnyi zhurnal – Scientific news of NTUU "KPI": scientific and technical journal.* 5(109). 48-53. Retrieved from https://ela.kpi.ua/handle/123456789/18873 [in Ukrainian].

16. V. B. Mokin, O. V. Burdeina, K. O. Koval & A. R. Yashcholt. (2018). Metod proektuvannia kohnityvnoi karty dlia optymizatsii proforiientatsiinoi diialnosti ZVO [The method of designing a cognitive map to optimize career guidance activities of higher education institutions]. *Visnyk Vinnytskoho politekhnichnoho instytutu – Bulletin of the Vinnytsia Polytechnic Institute.* 3. 89-99. Retrieved from https://www.researchgate.net/publication/330089909_METOD_PROEKTUVANNA_KOGNITIVNOI_KARTI_DLA_OPTIMIZACII_PROFORIENTACIJNOI_DIALNOSTI_ZVO [in Ukrainian].

НОТАТКИ

НАУКОВЕ ВИДАННЯ

ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ТА СУСПІЛЬСТВО

INFORMATION TECHNOLOGY AND SOCIETY

ВИПУСК 3 (9)
ISSUE 3 (9)

2023

Коректура
Ірина Чудеснова

Комп'ютерна верстка
Наталія Кузнєцова

Формат 60x84/8. Гарнітура Cambria.
Папір офсет. Цифровий друк. Ум. друк. арк. 10,70. Замов. № 1123/739. Наклад 300 прим.

Видавництво і друкарня – Видавничий дім «Гельветика»
65101, Україна, м. Одеса, вул. Інглезі, 6/1
Телефон +38 (095) 934 48 28, +38 (097) 723 06 08
E-mail: mailbox@helvetica.ua
Свідоцтво суб'єкта видавничої справи
ДК No 7623 від 22.06.2022 р.