

ISSN 2663-3051  
(ISSN 0555-2656 до 2019 р.)

# БІОНІКА ІНТЕЛЕКТУ

ІНФОРМАЦІЯ, МОВА, ІНТЕЛЕКТ

---

---

№ 1 (100)

2024

---

---

НАУКОВО-ТЕХНІЧНИЙ ЖУРНАЛ

Заснований у жовтні 1967 р.

Засновник та видавець  
*Харківський національний університет радіоелектроніки*

Періодичність видання – 2 рази на рік



Науково-технічний журнал  
«БІОНІКА ІНТЕЛЕКТУ»

ISSN 2663-3051

Заснований Харківським національним університетом  
радіоелектроніки у 1967 році

Реферування та індексування:

Google Scholar



INDEX  COPERNICUS  
I N T E R N A T I O N A L



---

Журнал включено до списку наукових спеціалізованих видань України  
з технічних та фізико-математичних наук  
згідно з наказом Міністерства освіти і науки України № 820 від 11.07.2016  
(внесено зміни згідно з наказом МОНУ № 920 від 26.06.2024)

С. М. Неронов<sup>1</sup>, Г. А. Плехова<sup>2</sup>, М. В. Костікова<sup>3</sup><sup>1</sup>ХНАДУ, м. Харків, Україна, sernikner@gmail.com, ORCID iD: 0000-0003-2381-1271<sup>2</sup>ХНАДУ, м. Харків, Україна, plehovaanna1@gmail.com, ORCID iD: 0000-0002-6912-6520<sup>3</sup>ХНАДУ, м. Харків, Україна, kmv\_topaz@ukr.net, ORCID iD: 0000-0001-5197-7389

## УДОСКОНАЛЕННЯ АРХІТЕКТУРИ NEURONET АВТОТРАНСФЕРУ

Для створення єдиного інформаційного простору транспортних послуг без зайвих капітальних витрат на створення спеціальної ІТ-інфраструктури, що є основою підвищення конкурентної спроможності транспортних та дорожніх організацій надані практичні рекомендації з удосконалення архітектури Neuronet автотрансферу. АРХІТЕКТУРА, ВІРТУАЛІЗАЦІЯ, КЛАСТЕР, СЕРВЕР, САЙТ

S. M. Neronov, G. A. Pliekhova, M. V. Kostikova. **Improvement of Neuronet autotransfer architecture.** In order to create a unified information space for transport services without unnecessary capital costs for the creation of a special IT infrastructure, which is the basis of increasing the competitiveness of transport and road organizations, practical recommendations are provided for improving the Neuronet architecture of auto transfer.

ARCHITECTURE, VIRTUALIZATION, CLUSTER, SERVER, SITE

### Вступ

Рекомендації з удосконалення архітектури Neuronet автотрансферу базуються на твердженні про те, що програмна платформа Neuronet автотрансферу є гетерогенною розподіленою системою. Вертикальне масштабування такої системи можна здійснити завдяки збільшенню серверної потужності базової системи. Робочу версію транспортного порталу (з урахуванням попереднього тестування елементів програмних, технічних та апаратних рішень взаємодії бортового інформаційно-комунікаційного комплексу з транспортним порталом та результатів обчислювального експерименту з визначення інформаційних потоків та обсягу комп'ютерних ресурсів, необхідних для моніторингу транспортних ситуацій) не зважаючи на існування більш сучасних зразків було розгорнуто на сервері на базі двох процесорів Intel Xeon 3.00 GHz, 2 Гбайт оперативної пам'яті, двох жорстких дисків за технологією SCSI 7 Gb у апаратному RAID-масиві [1, 2], але це не вплинуло на швидкодію системи.

### Виклад основного матеріалу

Завдяки експериментальним дослідженням при навантаженні серверу запитами користувачів визначено, що основним недоліком обраної архітектури був надто малий об'єм оперативної пам'яті, що є критичним у разі серверної віртуалізації. Тому для здійснення масштабування серверних ресурсів було придбано та введено до експериментальної дії сервер самостійної зборки української компанії (рис. 1) на базі процесору Quad-Core 3 ГГц INTEL Xeon QC E3-1240V2, материнської плати INTEL S1200BTLR, 4-ох модулів пам'яті Kingston 8 Gb із загальним об'ємом 32 Гбайт DDR3, 3-ох накопичувачів HDD 1 Тбайт Seagate ES ST1000N 3, оптичного пристрою DVD-RW ASUS DRW-24B5ST та корпусу серверного 2U CSV UNI (400 Вт).



Рис. 1. Сервер R-Line на базі INTEL Xeon E3

Така архітектура є масштабованою «вертикально», тобто дозволяє додавати кількість модулів пам'яті, накопичувачі. Однак цей тип масштабування звичайно є обмеженим за «горизонтальне», що передбачає додавання однотипних чи гетерогенних вузлів до ферми або кластеру транспортного порталу.

Поруч із цим масштабування на рівні застосування технологій віртуалізації дозволяє виконувати фактичне прозоре збільшення ресурсів як за моделлю «вертикального» масштабування, так й «горизонтального». Для виконання завдань платформи віртуалізації транспортного порталу обрано систему віртуалізації Proxmox VE, яка використовує тільки x86\_64 архітектуру. Proxmox VE базується на дистрибутиві Debian Linux та є Web-інтерфейсом до застосування технологій віртуалізації типу KVM та LXC.

Для використання KVM-віртуалізації центральний процесор повинен підтримувати апаратну віртуалізацію (Intel VT або AMD-V). Система KVM (Kernel-based Virtual Machine) дозволяє запускати у середовищі віртуальних машин фактично будь-яку операційну систему, наприклад, Linux, Windows чи FreeBSD. Для LXC (Linux Containers) апаратна віртуалізація не

потрібна, оскільки контейнери Linux функціонують як запуск одного ядра операційної системи Linux для кожного контейнеру. Контейнери є механізмом захисту Web-додатків та створюють оточення навколо кожного додатку. Це призводить до покращення надійності системи та дозволяє проектувати за модульним принципом вміст віртуальної машини.

Установка Proxmox VE є дуже простою та швидкою. Упевнившись, що перший сервер працює, слід виконати установку додаткового сервера або вузла, процес установки ідентичний процесу установки «першого», тільки вказуємо відповідне ім'я сервера та його IP-адресу. У самому кластері можна заходити на кожен вузол та отримати ідентичний Web-інтерфейс керування системою. Таким чином можна створити на базі системи віртуалізації кластер високої доступності. Позначаються аббревіатурою HA (англ. High Availability – висока доступність). Створюються для забезпечення високої доступності сервісу, що надається кластером. Надмірна кількість вузлів, що входять в кластер, гарантує надання сервісу у разі відмови одного або декількох серверів. Типове число вузлів – два, це мінімальна кількість, що приводить до підвищення доступності. У разі кластеру високої доступності слід застосовувати декілька фізичних

систем з Proxmox VE (за даними: <http://habrahabr.ru/post/122338>).

Принцип дії кластеру розподілу навантаження будується на розподілі запитів через один або кілька вхідних вузлів, які перенаправляють їх на обробку в інші, обчислювальні вузли. Початкова мета такого кластера – продуктивність, однак, у них часто використовуються також і методи, що підвищують надійність. Подібні конструкції називаються серверними фермами. Для такого типу кластеру можна визначити ряд віртуальних машин на одному фізичному сервері з Proxmox VE, однак, у такому разі слід пам'ятати про загрозу відмови системи в цілому.

Комбінація кластерів на базі віртуальних машин дозволить значно підвищити надійність роботи транспортного порталу, а також його продуктивність та готовність до навантажень. На рис. 2 наведено зовнішній вигляд панелі управління Proxmox VE. Відповідна система дозволяє уявити всі інформаційні ресурси та віртуальні мережі у якості певного віртуалізованого центру обробки даних. Перевагою застосування кластеру на базі Proxmox VE є можливість швидкої міграції віртуальних машин у різні обчислювальні середовища та розташування їх накопичувачів у певних мережевих сховищах даних.

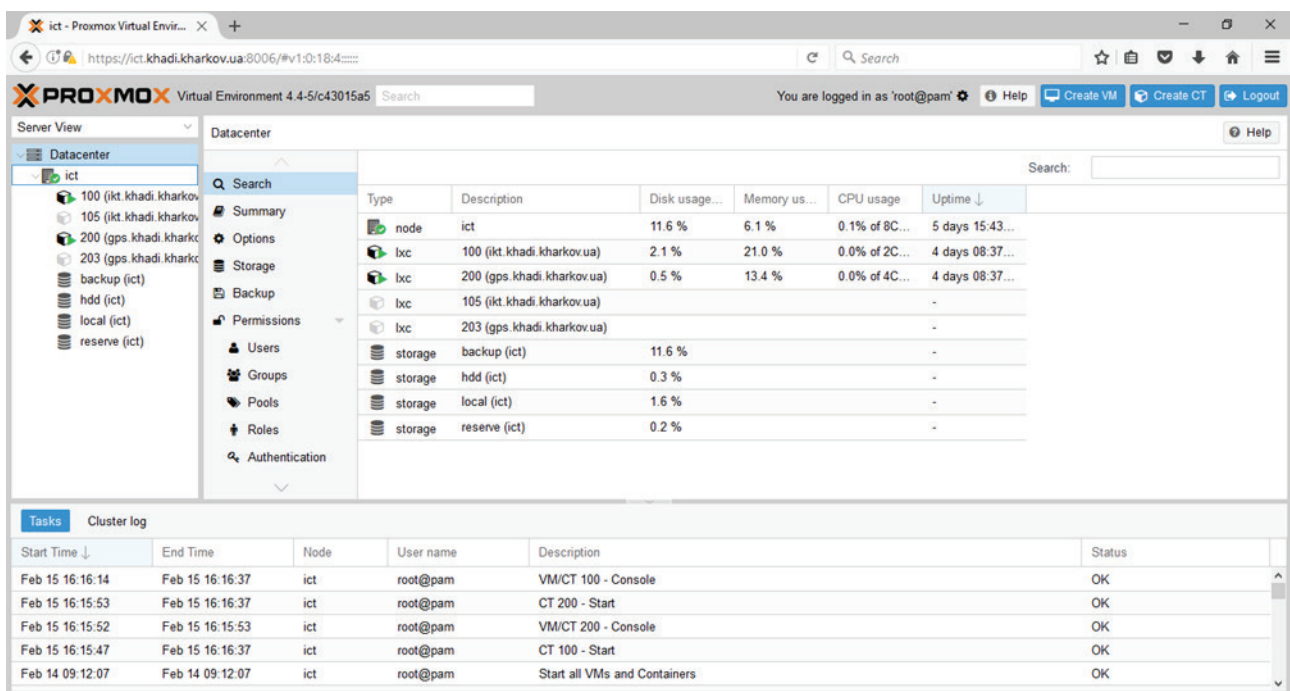


Рис. 2. Середовище адміністрування Proxmox VE

Слід розділяти поняття: масштабованість – здатність своєчасно реагувати на безперервне зростання навантаження і непередбачені напливи користувачів; доступність – надання доступу до додатку навіть у разі надзвичайних обставин; продуктивність – навіть найменша затримка в завантаженні сторінки може залишити негативне враження у

користувача (за матеріалами: <http://www.insight-it.ru/masshtabiruemost/masshtabiruemye-veb-arkhitektury/>).

Розробка фактично кожного сайту (порталу) спрямована на функціонування із максимальною стабільністю, тобто мати доступність для абсолютної всіх потенційних відвідувачів у кожен момент часу, однак завжди трапляються непередбачені ситуації,

які можуть стати причиною тимчасової недоступності. Для мінімізації потенційного збитку доступності додатка необхідно уникати наявності компонентів в системі, потенційний збій у яких привів би до недоступності, зменшення функціональності або втрати даних (сайту в цілому або частини порталу). Таким чином, кожен сервер або будь-який інший компонент системи повинен мати хоча одного дублера (не важливо в якому режимі вони працюватимуть: паралельно або один «страхує» інший, перебуваючи при цьому в пасивному режимі), а дані повинні бути реплікованими як мінімум у двох примірниках (причому бажано не на рівні RAID, а на різних фізичних машинах). Зберігання декількох резервних копій даних десь окремо від основної системи (наприклад, на спеціальних сервісах або на окремому кластері) також допоможе уникнути багатьох проблем. Не варто забувати і про фінансову сторону питання: «страхування» на випадок збоїв вимагає додаткових істотних вкладень в устаткування, які має сенс намагатися мінімізувати.

Масштабованість прийнято розділяти на два напрямки: вертикальна масштабованість – збільшення продуктивності кожного компонента системи з метою підвищення загальної продуктивності та горизонтальна масштабованість – розбиття системи на більш дрібні структурні компоненти та рознесення їх по окремих фізичних машинах (або їх груп) та/чи збільшення кількості серверів, які паралельно виконують одну й ту ж функцію.

При розробці стратегії росту системи слід шукати компроміс між ціною, часом розробки, підсумкової продуктивності, стабільністю та багатьма іншими критеріями. З фінансової точки зору вертикальна масштабованість є далеко не самим привабливим рішенням, оскільки ціни на сервера з великою кількістю процесорів завжди ростуть практично експоненціально щодо кількості процесорів. Саме по цьому найбільш цікавий є горизонтальний підхід. Але вертикальна масштабованість має право на існування, особливо в ситуаціях, коли основну роль відіграє час і швидкість вирішення завдання, а не фінансове питання. Це продиктовано тим, що купити потужний сервер істотно швидше, ніж практично заново розробляти додатки, адаптуючи їх до роботи на великій кількості паралельно працюючих серверів.

Архітектуру транспортного порталу обрано таким чином, що кожен компонент системи є окремим та незалежним від інших. Тому актуальною є задача рівномірного розподілення запитів між доступними серверами додатків. Запуск типового сайту починається з дуже простої архітектури – один Web-сервер (звичайно в його ролі виступає Apache), який виконує всю роботу по обслуговуванню HTTP-запитів, що надходять від відвідувачів. Він віддає клієнтам так звану «статичку», тобто файли, що лежать на диску сервера і

не потребують обробки: картинки (gif, jpg, png), листи стилів (css), клієнтські скрипти (javascript). Той же сервер відповідає на запити, що вимагають обчислень – це формування html-сторінок, хоча іноді динамічно створюються і зображення та інші документи. Найчастіше відповіді на такі запити формуються скриптами, написаними на java, php або іншими мовами (за матеріалами: <http://habrahabr.ru/post/15362/>).

Недоліком простої схеми роботи сайту в тому, що різні за характером запити (віддача файлів з диска і обчислювальна робота скриптів) обробляються одним й тим же Web-сервером. Обчислювальні запити потребують збереження в пам'яті сервера багато інформації (інтерпретатор скриптової мови, самі скрипти, дані, з якими вони працюють) та можуть займати багато обчислювальних ресурсів. Видача статички, навпаки, не вимагає багато ресурсів процесора, але може займати тривалий час, якщо у клієнта низька швидкість зв'язку. Внутрішній устрій сервера Apache припускає, що кожне з'єднання обробляється окремим процесом. Це зручно для роботи скриптів, однак неоптимально для обробки простих запитів. Відповідно «важкі» процеси Apache багато часу проводять в очікуванні (спочатку при отриманні запиту, потім при формуванні відповіді), не оптимально займаючи пам'ять сервера.

Вирішення цієї проблеми – розподіл роботи по обробці запитів між двома різними програмами – тобто поділ на frontend (сторона користувача ресурсу або зовнішнє представлення) та backend (бізнес-логіка роботи додатка). Легкий frontend-сервер виконує завдання по віддачі статички, а решта запитів перенаправляє на backend (застосовує режим проксі), де виконується формування сторінок. Очікування повільних клієнтів також бере на себе frontend, і якщо він використовує мультиплексування (коли один процес обслуговує кількох клієнтів, наприклад, nginx або NArгоху), то очікування практично не впливає на навантаження обладнання серверу (рис. 3).

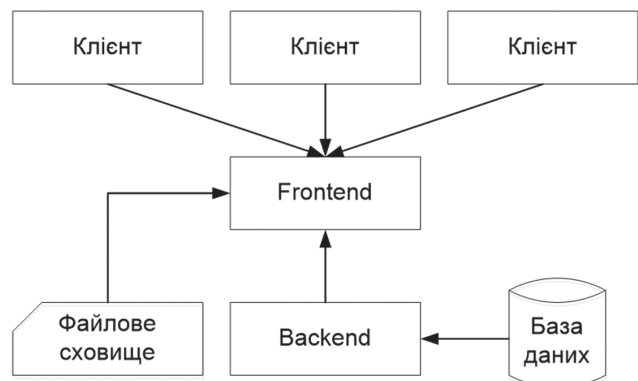


Рис. 3. Типова архітектура Web-ресурсу

Слід зазначити, що сучасний Web-ресурс фактично завжди застосовує базу даних, в якій зазвичай зберігаються основні дані системи, наприклад, MySQL

та PostgreSQL. Часто окремо можна виділити окреме сховище бінарних файлів, де містяться картинки (наприклад, ілюстрації до статей сайту та фотографії) або інші файли.

У більшості випадків на початку життєвого циклу сайту всі компоненти його архітектури розташовуються на одному сервері. Якщо він перестає справлятися з навантаженням, то є просте рішення – винести найбільш легко відокремлювані частини на інший сервер.

Найбільш продуктивним рішенням є застосування статичних сайтів (тільки frontend, що автоматично генерується одноразово на стороні серверу чи засобів розробника), однак, хоча вони є найбільш швидкісними, для рішення порталу такий підхід не задовольнить. У разі рішення масштабу Web-порталу звичайно будуть застосовані стандартні технології рівня frontend-backend із можливістю кешування статички.

Типовою ситуацією для зростаючого сайту є стан, коли база даних вже винесена на окрему машину і виконано поділ на frontend та backend, однак навантаження продовжує збільшуватися і backend не встигає обробляти запити. Це означає, що необхідним є розподіл обчислень на кілька серверів. Зробити це відносно просто – достатньо ввести до експлуатації другий сервер (певну віртуальну машину) і розгорнути на ньому програми та скрипти, що необхідні для роботи backend. Після цього треба зробити так, щоб запити користувачів розподілялися (балансувалися) між наявними серверами.

Крім цього мають місце менш поширені варіанти, засновані на DNS, тобто в процесі визначення клієнтом IP-адреси сервера з необхідним йому Інтернет-ресурсом адрес для нього формується з урахуванням навантаження на доступні сервери, а також з певних географічних міркувань (рис. 4).

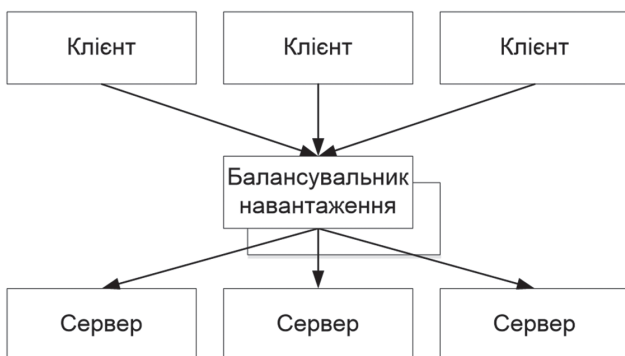


Рис. 4. Типова архітектура балансування навантаження Web-ресурсу

Мережеве устаткування, що дозволяє розподіляти навантаження між декількома серверами, зазвичай коштує досить значні суми, але серед інших варіантів саме цей підхід пропонує найвищу продуктивність та стабільність (в основному завдяки якості, також таке обладнання іноді поставляється парами, що

працюють за принципом HeartBeat або визначення працездатності вузлів). Іншим варіантом є застосування програмного забезпечення, зазвичай це просто HTTP-сервери, що перенаправляють запити іншим Web-серверам на інших вузлах замість відправки безпосередньо на обробку інтерпретатору мови програмування. Всі запити проходять через балансувальник, який визначає кому з серверів віддати запит на обробку. При отриманні запиту від клієнта балансувальнику потрібно визначити якому з Web-серверів переслати запит. Алгоритм прийняття рішення називається методом або стратегією балансування (за матеріалами: <http://www.phphighload.com/2012/08/blog-post.html>; <http://habrahabr.ru/post/147390/>). Найбільш поширені стратегії:

- Round Robin де з доступних серверів будуються черги і балансувальник вибирає перший в черзі. Після виконання запиту сервер переміщається в кінець черги;
- менша кількість з'єднань. Балансувальник веде облік кількості незакритих з'єднань та обирає той сервер, у якого кількість таких з'єднань менша;
- використання «ваги» серверів. Кожному серверу в залежності від потужності присвоюється вага, яка використовується для ранжирування.

Стратегія балансування, що не включає перевірку стану серверів або хоча б працездатності, не придатна для використання, так як не гарантує обробку запиту. Тому, балансувальнику слід перевіряти стан сервера, його завантаженість і вибирати найбільш здібний.

При балансуванні часто виникає проблема зберігання сесій, адже сесія доступна тільки на тому сервері, який створив її. Це слід враховувати в алгоритмі перенаправлення запиту або додаток повинен зберігати сесії на окремому сервері або в базі даних.

Для реалізації програмного балансувальника є багато рішень, наприклад: nginx – вільний Web-сервер і проксі-сервер, який має версії для сімейства Unix-подібних операційних систем (FreeBSD, Linux та ін.) й Microsoft Windows. Основні функції – це HTTP-сервер для обслуговування статичних запитів, акселерування проксіювання з підтримкою кешування, акселерована підтримка FastCGI та memcached серверів, має простий розподіл навантаження і відмовостійкість, модульність. Apache HTTP-сервер – відкритий веб-сервер для UNIX-подібних, Microsoft Windows та інших операційних систем. На сьогодні є найуживанішим Web-сервером мережі Інтернет.

Більш простим рішенням є застосування проксі-серверу, наприклад, HAProxy. Для його конфігурування потрібно вказати IP-адреси веб-серверів. Налаштування балансування зводиться до запису в конфігурації проксі-сервера `/etc/haproxy/haproxy.cfg` – «mode http» режиму роботи проксі, який стосується обробки пакетів, «balance leastconn» – алгоритм

вибору сервера (перенаправляємо запит серверу з найменшою кількістю з'єднань). Параметр «cookie serv insert» застосовується для забезпечення роботи сесій. Таким чином клієнт завжди буде потрапляти на один сервер і мати доступ до даних сесії. Параметр «option httpclose» автоматично закривати з'єднання після завершення передачі даних. Запускається проксі сервер командою: `sudo haproxy-f/etc/haproxy/haproxy.cfg`. Конфігурація HAProxy звичайно сильно залежить від завдання, наприклад, для роботи із статичним контентом або роботи сайту без сесій немає потреби прив'язувати конкретний сервер до клієнта.

Тому не випадково все більшою популярністю користуються Web-сервери, наприклад, nginx, які можуть працювати автономно або в зв'язці з Apache в якості кешу для статичних запитів. Однак можна розглянути дещо інший підхід, пов'язаний з використанням HTTP-прискорювача Varnish, який дозволяє помітно розвантажити Web-сервери і зменшити загальний час відгуку (згідно електронному ресурсу: <http://webperformance.ru/2011/06/03/varnish-speed-up/>).

Varnish є безкоштовним рішенням для кешування як статичного, так і динамічного контенту. Працює він як балансувальник до будь-якого Web-сервера або сервера додатків і повністю орієнтований на високу продуктивність, багатопоточність та максимально ефективно використання можливостей операційних систем сімейства Linux.

Альтернативою до визначених підходів виступають так звані Content Delivery Network (CDN) – зовнішні сервіси, що забезпечують доступність контенту користувачам. Перевага очевидна – немає необхідності організовувати власну інфраструктуру для вирішення цього завдання, але з'являється інша додаткова стаття витрат.

Таким чином, для рішення завдань балансування навантаження на транспортний портал оптимальним вибором є система Varnish, що забезпечує як балансування, так й кеш даних. Поставивши між користувачем і Web-сервером прозорий проксі-сервер, можна надавати користувачу дані з кеша проксі (який може бути як в оперативній пам'яті, так і дисковим), не доводячи запити навіть до HTTP-серверів. У більшості випадків цей підхід актуальний тільки для статичного контенту, в основному різних форм медіа-даних: зображень, відео і тому подібного. Це дозволяє Web-серверам зосередитися тільки на роботі з самими сторінками.

У разі здійснення заходів оптимізації додаток все одно може не впоратись з навантаженням. У такому випадку рішенням проблеми, очевидно, може послужити рознесення його по декільком вузлах, з метою збільшення загальної продуктивності додатка за рахунок збільшення доступних ресурсів.

Більшість Web-додатків априорі є розподіленими, оскільки в їх архітектурі можна виділити мінімум три

шари: Web-сервер, бізнес-логіка (додаток), дані (база даних, статика). Кожен із цих шарів можна масштабувати. Тому, якщо у системі додаток і база даних розміщені на одному хості – першим кроком, безсумнівно, має стати рознесення їх по різних хостам (за матеріалами: <http://habrahabr.ru/post/113992/>, <http://www.phphighload.com/2012/10/mysql-scaling-strategies.html>).

Починаючи масштабування системи, варто визначити, який із шарів є «вузьким місцем» – тобто працює повільніше за решту системи. Для цього можна скористатися простими утилітами типу `top` (`htop`) для оцінки завантаження процесора/пам'яті та `df`, `io-stat` – для оцінки продуктивності дискової підсистеми серверу. Однак, бажано виділити окремий хост, з емуляцією промислового навантаження (наприклад, за допомогою Apache JMeter), на якому можна буде профілювати роботу додатка. Для виявлення вузьких запитів до бази даних можна скористатися утилітами на основі логів з журналу сервера, що знаходиться у промислової експлуатації.

Більшість залежить від архітектури Web-додатку, але найбільш вірогідними кандидатами на «вузьке місце» в загальному випадку є бази даних та код. Якщо додаток працює з великим об'ємом даних користувача, то «вузьким місцем», відповідно, швидше за все буде зберігання статистики.

Часто вузьким місцем в сучасних додатках є бази даних. Проблем з базами даних діляться, як правило, на два класи: продуктивність і необхідність зберігання великої кількості даних. Знизити навантаження на базу даних можна за умови поширення її на декілька вузлів. При цьому гостро постає проблема синхронізації між вузлами.

## Висновки

Результат полягає в усуненні протиріч наявності загальних вартісних обмежень та потрібних комп'ютерних ресурсів раціональної організації клієнт-серверної технології перевізного процесу. Своєрідною вільною новою нішею відповідних розробок є синергетичний, інформаційний розвиток ринку транспортних послуг.

## Список літератури:

- [1] Alekseyev O. Development of automotive computer systems based on the virtualization of transportation processes management / O. Alekseyev, V. Alekseyev, D. Klets, V. Khabarov et al. // Eastern-European Journal of Enterprise Technologies. – 2017. – Vol. 6, No 3 (90). – P. 14 – 25. – DOI: 10.15587/1729-4061.2017.116351.
- [2] Алексієв О. П. Інформаційний розвиток порталу віртуального управління процесами транспортного обслуговування / О. П. Алексієв, В. О. Алексієв, // Інформаційні технології: проблеми та перспективи: монографія [Текст]. – Х.: Вид-во: Рожко С. Г., 2017. – Розд. 2. – С. 32 – 47. URI: <http://www.repository.hneu.edu.ua/jspui/handle/123456789/16051>.

*Надійшла до редколегії 15.01.2024*



В. В. Литвин<sup>1</sup>, І. Р. Пелешчак<sup>2</sup>, В. А. Висоцька<sup>3</sup>, В. М. Мотика<sup>4</sup>, Р. М. Пелешчак<sup>5</sup>

<sup>1</sup>НУ «Львівська політехніка», м. Львів, Україна, vasyi.v.lytvyn@lpnu.ua,  
ORCID iD: 0000-0002-9676-0180

<sup>2</sup>НУ «Львівська політехніка», м. Львів, Україна, ivan.r.peleshchak@lpnu.ua,  
ORCID iD: 0000-0002-7481-8628

<sup>3</sup>НУ «Львівська політехніка», м. Львів, Україна, victoria.a.vysotska@lpnu.ua,  
ORCID iD: 0000-0001-6417-3689

<sup>4</sup>НУ «Львівська політехніка», м. Львів, Україна, volodymyr.motyka.mnsam.2023@lpnu.ua,  
ORCID iD: 0009-0009-8086-7619

<sup>5</sup>НУ «Львівська політехніка», м. Львів, Україна, roman.m.peleshchak@lpnu.ua,  
ORCID iD: 0000-0002-0536-3252

## ПАСИВНИЙ ПІДХІД ДО ВИЯВЛЕННЯ ТА КЛАСИФІКАЦІЇ МІН ІЗ ВИКОРИСТАННЯМ МАГНІТНИХ ПОЛІВ І ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ

В наш час для виявлення мін широко використовуються активні детектори, що забезпечують виявлення мін з високою точністю. Однак їхній принцип роботи може бути небезпечним, оскільки він має потенціал спричинити детонацію прихованих мін. Це дослідження представляє інноваційний метод виявлення та класифікації мін, виготовлених з різних матеріалів (металевих, напівметалевих, пластикових), за допомогою згорткової нейронної мережі. Мережа аналізує аномалії, отримані з унікально розробленої матриці ознак, досягаючи значної точності у класифікації. Процес класифікації базується на вхідному векторі ознак, який включає: показники напруги від магнітометричного датчика FLC-100, який виявляє магнітні аномалії біля мін із чутливістю в діапазоні від  $10^{-10}$  до  $10^{-4}$  Тесла; дані з шести типів ґрунту з урахуванням їх вологості; висоту розташування датчика над міною. Оскільки міни, виготовлені з різних матеріалів та мають унікальні магнітні властивості, використання даних про магнітні аномалії дозволяє ефективно класифікувати міни. Точність класифікації оцінювалась за допомогою різних показників, таких як F1, Precision, Accuracy, FPR, FNR, Recall та TNR.

МАГНІТНІ АНОМАЛІЇ, КЛАСИФІКАЦІЯ МІН, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ (CNN), СЕНСОР FLC-100, БЕЗПЕЧНЕ РОЗМІНУВАННЯ.

V.V. Lytvyn, I.R. Peleshchak, V.A. Vysotska, V.M. Motyka, R.M. Peleshchak. *A Passive Approach to Mine Detection and Classification Utilizing Magnetic Fields and Convolutional Neural Networks*. Active detectors are widely used today for mine detection, providing high accuracy. However, their operating principle can pose a danger, as it has the potential to trigger the detonation of hidden mines. This study introduces an innovative method for detecting and classifying mines composed of different materials (metallic, semi-metallic, plastic) through the use of a convolutional neural network. The network analyzes anomalies derived from a uniquely designed feature matrix, achieving notable accuracy in classification. The classification process is based on a feature input vector that includes: voltage readings from the FLC-100 magnetic field sensor, which detects magnetic field anomalies near mines with sensitivities ranging from  $10^{-10}$  to  $10^{-4}$  Tesla; data from six soil types, considering their moisture content; and the sensor's height above the mine. Since mines made from different materials exhibit distinct magnetic properties, utilizing magnetic field anomaly data allows for effective material-based mine classification. The classification accuracy was evaluated using various metrics, such as F1, Precision, Accuracy, FPR, FNR, Recall, and TNR.

MAGNETIC ANOMALIES, MINE CLASSIFICATION, CNN, FLC-100 SENSOR, SAFE DEMINING.

### Вступ

Виявлення наземних мін залишається критичною та глобальною проблемою, яка загрожує мільйонам життів. У 2016 році щодня в середньому 23 особи гинули або зазнавали поранень через міни. Наразі близько 61 країни та регіону продовжують стикатися з постійною загрозою від цих прихованих вибухових пристроїв, що становить постійну небезпеку для життя людей [1]. Виявлення та знешкодження мін є важливими кроками для забезпечення безпеки в постконфліктних зонах.

Традиційні методи виявлення мін більше не забезпечують необхідного рівня ефективності та надійності. Активні детектори, що генерують електричні сигнали для виявлення мін, можуть випадково спровокувати спрацьовування вибухового механізму, що призводить до небезпечних наслідків. З цієї причини пасивні детектори, засновані на вимірюванні магнітних аномалій у місці розташування міни, стають більш безпечним варіантом. Сучасні автоматизовані методи, такі як нейронні мережі, виявляються ефективними у вирішенні цього завдання.



Важливість використання нейронних мереж у поєднанні з даними про магнітні аномалії пояснюється такими факторами:

*Безпека військових і цивільних:* Виявлення мін є ключовим для порятунку життів як під час бойових дій, так і при розмінуванні територій після завершення конфліктів.

*Швидкість і точність:* Датчики магнітного поля сприяють швидкому виявленню закопаних мін, що дозволяє оперативно реагувати та розпізнавати міни. Нейронні мережі підвищують ефективність цього процесу, точно класифікуючи сигнали і фільтруючи шум або нерелевантні об'єкти.

*Зниження ризиків:* Автоматизовані системи на основі нейронних мереж зменшують ймовірність людських помилок та суттєво підвищують точність виявлення і класифікації мін.

У цій роботі представлено метод виявлення та класифікації мін із різних матеріалів (металевих, напівметалевих, пластикових) з використанням згорткової нейронної мережі. Підхід базується на даних від магнітного сенсора, який здатен виявляти магнітні аномалії з чутливістю в діапазоні від  $10^{-10}$  до  $10^{-4}$  Тесла, при цьому враховується тип ґрунту та висота сенсора над закопаним об'єктом. Ці дані перетворюються в спеціалізовану матрицю ознак, після чого згорткова нейронна мережа аналізує дані з цієї графічної матриці.

## 1. Огляд та аналіз літератури

До класичних засобів виявлення та розпізнавання вибухонебезпечних пристроїв можна віднести такі групи: апаратура з проникаючим випромінюванням, детектори парів та вибухових речовин (газоаналізатори, засоби хімічного аналізу), магнітомеханічні засоби, засоби пошукової техніки [2]. Апаратура з проникаючим випромінюванням може використовувати рентгенівське, іонізуюче випромінювання. У цьому напрямку розробляються ядерно-фізичні методи, основою яких є джерела нейтронів та виявлення гамма-випромінювання після взаємодії нейтрона з об'єктом пошуку [2]. Методи на основі хімічного аналізу дедалі частіше інтегруються в робототехнічні технології [3]. Окрім цього, пошукові пристрої, як-от металодетектори, активно вдосконалюються, і в сучасних розробках акцент робиться на створенні автоматизованих систем на їхній основі. Ці інновації підвищують безпеку та покращують точність виявлення, інтегруючи металодетектори у роботизовані платформи [4].

Детектори мін можна поділити на активні та пасивні. Активні детектори випромінюють сигнал у напрямку цілі та аналізують відбитий сигнал для виявлення мін. Незважаючи на високу точність цих пристроїв, вони мають ризик випадкового

спрацьовування вибухового механізму міни. Пасивні детектори, навпаки, засновані на виявленні порушень у магнітному полі, викликаних присутністю мін. Хоча вони менш точні, але загалом вважаються безпечнішими [5, 6]. Існує декілька подібних досліджень, які включають використання алгоритмів машинного навчання для класифікації виявлених мін за допомогою аномалії магнітного поля. Наприклад, у одному з досліджень використовували автоенкодер зі спеціальним налаштуванням, досягнувши точності класифікації 93% і показника AUC 98% для виявлення прихованих мін [7]. В іншому дослідженні [8] застосовувався алгоритм k-NN з даними від мережі з 32 магнітних датчиків, і було досягнуто точності 91,66% у виявленні мін.

Інноваційність нашого дослідження полягає в розробці згорткової нейронної мережі, здатної забезпечити високу точність класифікації мін, виготовлених із різних матеріалів. Процес класифікації базується на вхідному векторі ознак, який включає: показники напруги від магнітометричного датчика FLC-100; інформацію про шість різних типів ґрунтів із різним рівнем вологості; та висоту сенсора над міною.

## 2. Матеріали і методи

У цьому дослідженні розглядаються архітектури нейронних мереж, що включають компоненти, такі як повнозв'язані шари, згорткові шари, механізми пулінгу та різні функції активації.

Повнозв'язаний шар у нейронній мережі з'єднує кожен нейрон з усіма нейронами попереднього шару. Це забезпечує те, що кожен вихід з попереднього шару пов'язаний з кожним нейроном у поточному шарі.

Згортковий шар є основним компонентом згорткових нейронних мереж (CNN). Він застосовує згорткові операції до вхідного зображення за допомогою набору фільтрів (ядер), щоб виділити візуальні ознаки, такі як краї, кути, форми та візерунки [9–11]. Використовуючи малі квадратні ядра, ці фільтри дозволяють виявити різні характеристики у вхідних даних.

Згортка дозволяє мережі захоплювати ознаки вхідного зображення, зберігаючи просторові зв'язки між пікселями, що дає можливість мережі навчатися ознакам зображення за допомогою невеликих фрагментів вхідних даних. У згортковому шарі нові ознаки виявляються шляхом виконання матричного множення між ядром і вхідним зображенням. Розмір ядра може варіюватися, найчастіше використовуються  $2 \times 2$ ,  $3 \times 3$  або  $5 \times 5$ , і вони представлені у вигляді матриць відповідного розміру.

Застосоване ядро витягує різні ознаки з вхідних даних, і вибір та застосування цих ядер є критично важливим. Великий розмір ядра може призвести

до втрати певних характеристик у вхідних даних. Коефіцієнти ядра представляють ваги в згортковій нейронній мережі. Під час кожної ітерації коефіцієнти ядра коригуються в процесі навчання. У міру того, як значення навчаються, коефіцієнти оновлюються, що призводить до створення нових ядер.

Для кожного пікселя у вхідних даних ядро ковзає по всіх пікселях, виконуючи множення між коефіцієнтом вхідних даних і коефіцієнтом ядра. Ця операція проводиться незалежно для кожного каналу вхідних даних. Результати з усіх каналів підсумовуються, і кінцеве значення присвоюється відповідній точці у вихідних даних.

Згортка – це математичний процес, який включає вхідні дані, такі як матриця зображення, і ядро. Операцію згортки ілюструє рис. 1. Розмір виходу відображає загальну кількість різних ядер, які можуть бути отримані з зображення.

Розміри вхідного зображення:  $h \times w \times d$ .

Розміри ядра:  $f_h \times f_w \times d$ .

Розміри вихідної матриці:  $(h - f_h + 1) \times (w - f_w + 1) \times 1$ .

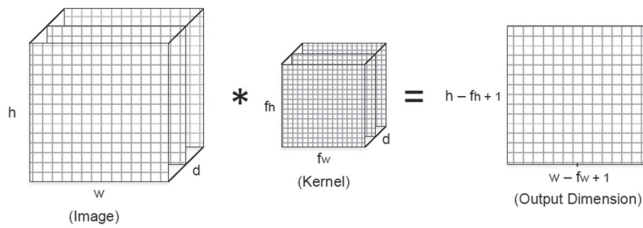


Рис. 1. Процес операції згортки

Шар пулінгу в глибоких нейронних мережах зменшує розмірність вхідних даних, зменшуючи кількість параметрів та усуваючи незначні деталі з зображення або вектору ознак. Він діє як фільтр, який обчислює статистичні властивості (такі як максимальні або середні значення) для кожного пікселя, фрагмента зображення або групи ознак у векторі.

Під час навчання моделі шар пулінгу зменшує кількість параметрів і допомагає зменшити ризик перенавчання. Шар пулінгу можна додати після кожного згорткового шару для зменшення розміру вхідних даних і відкидання менш важливих деталей.

Шар пулінгу зменшує розмір зображення, зберігаючи важливі ознаки. Пулінг зменшує висоту та ширину вхідних даних, що прискорює процес навчання нейронної мережі, хоча це може призвести до деякої втрати ознак. Однією з обмежень пулінгу є те, що він не вводить навчальні параметри, оскільки розмір ядра та крок є заздалегідь визначеними та фіксованими.

Одним з найбільш поширених методів пулінгу є макс-пулінг. У цій техніці обирається вікно розміром  $n \times n$ , де  $n$  менше розміру зображення, і вибирається найбільше значення в цьому вікні. Потім вікно переміщується по зображенню з певною довжиною кроку,

і ця операція повторюється. Рис. 2 ілюструє операцію макс-пулінгу.

Отже, шари пулінгу відіграють вирішальну роль у зменшенні розміру даних, запобіганні перенавчанню та забезпеченні швидших обчислень, зберігаючи найважливішу інформацію.

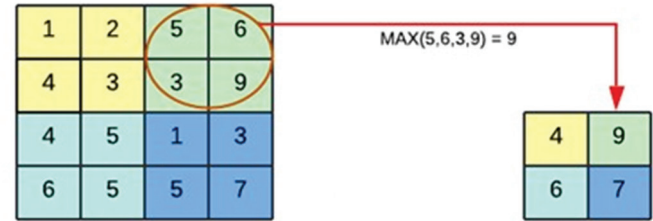


Рис. 2. Приклад максимального пулінгу з ядром  $2 \times 2$

Метриками для моделей машинного навчання у цьому дослідженні були обрані:

F1 Score: метрика є гармонійним середнім значенням Accuracy та Recall. Вона фіксує компроміс між цими двома аспектами, пропонуючи всебічну оцінку ефективності моделі.

Precision: метрика оцінює відсоток правильних позитивних прогнозів, зроблених моделлю. Вища точність свідчить про меншу кількість хибнопозитивних прогнозів.

Accuracy (точність): метрика оцінює частку всіх прогнозів моделі, які є правильними. Вона відображає загальну ефективність моделі, незалежно від класу, який прогнозується.

FPR: метрика вимірює частку негативних прикладів (класу, який не прогнозується), які були неправильно класифіковані моделлю як позитивні.

FNR: метрика оцінює відсоток фактичних позитивних випадків, які модель помилково ідентифікувала як негативні.

Recall: метрика оцінює відсоток фактичних позитивних випадків, які модель успішно розпізнала.

TNR: метрика оцінює відсоток негативних випадків, які модель правильно класифікувала. Вона показує, наскільки добре модель може розрізнити негативні класи.

### 3. Набір даних та передобробка даних

Це дослідження зосереджується на класифікації мін на основі характеристик аномалій магнітного поля, використовуючи дані з [12], параметри яких детально описані в Таблиці 1, Таблиці 2 та Таблиці 3. Крім того, було проведено кореляційний аналіз між аномаліями магнітного поля, складом ґрунту та відстанню сенсора до міни. Також були виявлені типові патерни магнітних аномалій для кожного типу міни. Для виявлення магнітних аномалій поблизу закопаних мін використовувався ферозондовий сенсор FLC100 [13, 14].

Таблиця 1

Параметри даних				
Параметри				
	Вхідні дані			Вихідні дані
	Вольтаж (V)	Висота (H)	Тип ґрунту (S)	Тип міни (M)
Означення	Вихідна напруга сенсора FLC	Висота датчика над землею	Шість різних типів ґрунтів	П'ять різних класів мін
Граничні значення/ Класи	[0 V, 10.6 V]	[0 cm, 20 cm]	Посушливий і піщаний	Null
			Посушливий і багnistий	Протитанкова
			Посушливий і крейдяний	Протипіхотна
			Мокрий і піщаний	Протипіхотна мінована пастка
			Вологий і гумусний	M14 Протипіхотна
			Вологий і крейдяний	

Таблиця 2

Зв'язок між аномаліями магнітного поля для мін і типом ґрунту

Тип ґрунту	Null, V	Протитанкова, V	Протипіхотна, V	Протипіхотна мінована пастка, V	M14 Протипіхотна, V
Посушливий і піщаний	3.570	10.410	3.840	5.600	3.210
Посушливий і багnistий	3.510	7.510	3.930	5.600	4.430
Посушливий і крейдяний	3.730	10.410	6.900	2.416	5.000
Мокрий і піщаний	3.790	10.410	6.230	4.500	5.240
Вологий і гумусний	3.360	10.410	5.060	2.780	4.210
Вологий і крейдяний	3.620	10.410	5.970	4.410	4.560

Таблиця 3

Вплив відстані від сенсора до землі на аномалії магнітного поля біля мін

Висота (cm)	Null, V	Протитанкова, V	Протипіхотна, V	Протипіхотна мінована пастка, V	M14 Протипіхотна, V
0.00	3.7	10.5	4.2	6.0	6.3
1.83	3.5	10.5	4.1	5.6	4.8
3.65	3.5	10.5	3.9	5.1	3.7
5.46	2.9	10.5	4.0	4.5	3.6
7.28	3.0	9.6	3.7	4.4	3.2
9.10	2.8	8.4	3.5	4.3	3.0
10.92	3.0	7.1	3.5	4.3	2.9
12.74	2.7	6.5	3.5	4.1	2.8
14.56	2.6	6.3	3.6	4.0	2.7
16.37	2.7	4.9	3.7	3.3	2.6
18.19	2.7	4.7	3.3	3.3	2.6
20.00	2.5	4.6	3.3	3.2	2.6

Набір даних містить інформацію про п'ять типів мін у шести різних типах ґрунту, з висотою сенсора над землею, що варіюється від 0 см до 20 см. Наступним кроком є створення матриці 101x6, яка відображає дані з навчального набору (рис. 3).

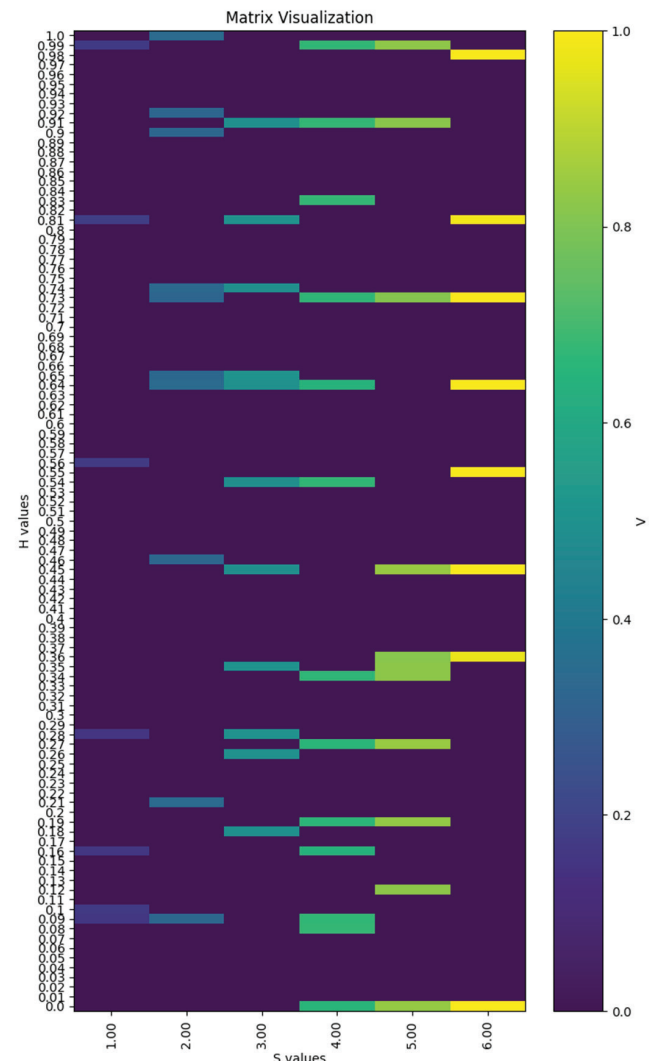


Рис. 3. Матриця візуалізації даних

Процес створення матриці починається з визначення осі X, використовуючи унікальні значення типів ґрунту (S), які варіюються від 1 до 6. Ось Y встановлюється за значеннями висоти (H), які коливаються від 0,00 до 20,00, але були нормалізовані до діапазону від 0,00 до 1,00 з кроком 0,01.

Ініціалізується порожня матриця, заповнена нулями, де кількість рядків відповідає значенням осі Y, а кількість стовпців відповідає значенням осі X.

Функція проходить через кожен рядок DataFrame, визначаючи відповідний стовпець на основі типу ґрунту (S) та правильного рядка на основі висоти (H). Для кожного запису значення напруги (V) розміщується в клітинці матриці у вказаному рядку та стовпці.

Якщо кілька записів мають однакові значення S та H, останнє значення напруги переписує попереднє в матриці.

Це призводить до створення структурованої матриці, яка представляє значення напруги для різних

типів ґрунту та висот. Нарешті, завершена матриця повертається для подальшої обробки. Цей систематичний підхід організовує дані таким чином, щоб підкреслити взаємозв'язки між напругою, типом ґрунту та висотою.

#### 4. Модель CNN та результати

Для порівняння були використані дві структури згорткових нейронних мереж. Для створення та оцінки ефективності моделі були використані Python 3.9 та бібліотека Keras 2.10.

##### Structure #1

1@101x6 - 16@3x3/2x2 - 32@3x3/2x2 - F128 - F64 - F32 - S5

Вхідний шар, який приймає матрицю розміром 101x6, позначається як 1@101x6. Повністю з'єднаний шар з 128 нейронами позначається як F128, тоді як S5 вказує на шар з активаційною функцією Softmax. Згортковий шар з 16 фільтрами, ядром згортки 3x3 та ядром пулінгу 2x2 представлений як 16@3x3/2x2.

Таблиця 4

Результати навчання мережі

№	Структура	Критерій	Оптимізатор	Параметри оптимізатора	F1	Accuracy	Precision	Recall	FPR	FNR	TNR
1	1@101x6 - 16@3x3/2x2 - 32@3x3/2x2 - F128 - F64 - F32 - F5	Cross Entropy Loss	Adam	lr = 0.001	0.9215	0.9230	0.9230	0.9230	0.0443	0.0770	0.9557
2	1@101x6 - 16@3x3/2x2 - 32@3x3/2x2 - F128 - F64 - F32 - S5	Cross Entropy Loss	Adam	lr = 0.001	0.933	0.9347	0.9345	0.9347	0.0845	0.0973	0.9155
3	1@101x6 - 8@3x3/2x2 - 16@3x3/2x2 - F128 - F64 - F32 - S5	Cross Entropy Loss	Adam	lr = 0.001	0.9012	0.9006	0.9043	0.9006	0.1308	0.0628	0.8692
4	1@101x6 - 32@3x3/2x2 - 64@3x3/2x2 - F128 - F64 - F32 - S5	Cross Entropy Loss	Adam	lr = 0.001	0.9074	0.9078	0.9179	0.9078	0.2268	0.0959	0.7732
5	1@101x6 - 16@3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - S5	Cross Entropy Loss	Adam	lr = 0.001	0.9347	0.9363	0.9376	0.9363	0.0813	0.1015	0.9187
6	@101x6 - 16@4x4/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - S5	Cross Entropy Loss	Adam	lr = 0.0001	0.9037	0.9091	0.9277	0.9091	0.1917	0.0595	0.8083
7	1@101x6 - 16@3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - S5	NLLoss	Adam	lr = 0.001	0.929	0.9307	0.9362	0.9307	0.1532	0.0304	0.8468
8	1@101x6 - 16@3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	Adam	lr = 0.001	0.9375	0.9387	0.9385	0.9387	0.0978	0.082	0.9022
9	1@101x6 - 16@3x3/2x2 - 24@T3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	Adam	lr = 0.001	0.9271	0.9291	0.9307	0.9291	0.0463	0.1545	0.9537

№	Структура	Критерій	Оптимізатор	Параметри оптимізатора	F1	Accuracy	Precision	Recall	FPR	FNR	TNR
10	1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	Adam	lr = 0.001	0.939	0.9408	0.9419	0.9408	0.081	0.1063	0.919
11	1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	AdamW	lr = 0.001	0.9312	0.9326	0.9313	0.9326	0.0825	0.0932	0.9175
12	1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	NAdam	lr = 0.001	0.9251	0.9304	0.9395	0.9304	0.0634	0.2752	0.9366
13	1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	RAdam	lr = 0.001	0.9361	0.9382	0.9375	0.9382	0.0683	0.1315	0.9317
14	1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5	NLLoss	RAdam	lr = 0.002	0.9121	0.9134	0.9157	0.9134	0.1887	0.0441	0.8113

*Structure #2*

1@101x6 - 16@T3x3/2x2 - 24@3x3/2x2 - F128 - F64 - F32 - LS5

У 16@T3x3/2x2 літера 'T' вказує на те, що матриці транспонуються перед подачею в згортковий шар. LS5 відноситься до повністю з'єднаного шару з активувальною функцією LogSoftMax.

Результати навчання для кожної мережі, разом із відповідними метриками продуктивності, представлені в Таблиці IV. Комплексний аналіз тестових результатів підкреслює складні взаємовідносини між архітектурою мережі, критеріями втрат, оптимізаторами та їх параметрами, а також те, як ці фактори разом впливають на метрики продуктивності. Проектування архітектури є вирішальним для здатності моделі навчатися та ефективно узагальнювати інформацію з даних. Наприклад, моделі з глибшими згортковими шарами, такі як Тест #2 та Тест #5, які включають кілька згорткових та пулінгових шарів, а також повністю з'єднані шари, послідовно досягають вищих показників F1 та точності. Це свідчить про те, що більш складна архітектура з достатньою кількістю шарів може захоплювати складніші шаблони в даних.

Вибір критерію втрат також суттєво впливає на продуктивність моделі. Моделі, які використовують втрати крос-ентропії, зазвичай перевершують ті, що використовують NLLoss, з точки зору F1 та точності. Це може бути пов'язано з тим, що крос-ентропійні втрати забезпечують більш інформативні градієнти, що веде до більш ефективного процесу навчання.

Однак деякі конфігурації з NLLoss, такі як Тест #8 та Тест #10, все ще демонструють конкурентоспроможні результати, вказуючи на те, що за відповідних структурних налаштувань та параметрів оптимізатора NLLoss може бути ефективно використано.

Вибір оптимізатора та його параметри є критично важливими для тонкого налаштування процесу навчання. Adam з швидкістю навчання 0.001 виявляється найбільш ефективним у кількох тестах, постійно забезпечуючи високі значення точності та F1. Наприклад, Тест #2 та Тест #5, які використовують Adam з цією швидкістю навчання, демонструють хороші результати. Навпаки, альтернативні оптимізатори, такі як NAdam та RAdam, показують змішані результати; NAdam у Тесті #12 досягає високого TNR, але низького FPR, в той час як RAdam у Тесті #13 ефективно балансує точність і Recall. Крім того, нижча швидкість навчання, як видно в Тесті #6 з Adam, встановленим на 0.0001, призводить до більш стабільного навчання, але з трохи нижчою продуктивністю, що підкреслює компроміс між швидкістю навчання та продуктивністю моделі.

Аналіз метрик продуктивності показує, що висока точність і Recall корелюють з високими показниками F1, що є показником збалансованої та ефективної класифікації. Тести #2, #5 та #8 виділяються з F1, що перевищують 0.93, демонструючи їх надійність у правильному визначенні як позитивних, так і негативних класів. Частота хибнопозитивних прогнозів (FPR) і частота хибнонегативних прогнозів (FNR)

надають додаткове розуміння специфічних помилок, які допускають моделі. Моделі з нижчими FPR і FNR, такі як Тест #10 та #14, демонструють кращий баланс між чутливістю та специфічністю, що робить їх більш надійними для практичного застосування.

В цілому, запропонований метод, що використовує згорткову нейронну мережу, може ефективно класифікувати міни з точністю 94.08% та F1 93.75%. Однак може знадобитися додаткове дослідження та налаштування параметрів, щоб адаптувати цей метод для інших наборів даних.

### Висновки

Запропонований метод класифікації мін, оснований на даних магнітних аномалій, зібраних сенсором FLC-100, демонструє високу ефективність і безпеку в порівнянні з активними детекторами, які можуть становити ризик детонації. Модель згорткової нейронної мережі (CNN) дозволяє точно класифікувати міни, побудовані з різних матеріалів, таких як метал, напівметал і пластик.

Дослідження показало, що використання магнітного поля як основи для аналізу аномалій дозволяє ефективно розпізнавати міни, враховуючи тип ґрунту та висоту сенсора. Комплексний аналіз результатів продемонстрував, що найкраща продуктивність досягається за допомогою складних архітектур мереж і оптимізаторів, таких як Adam, з швидкістю навчання 0.001.

Найвища точність (94.08%) та F1 (93.75%) підкреслюють надійність моделі, що робить її обіцяючою для реальних застосувань. Однак для забезпечення високої продуктивності на інших наборах даних може знадобитися додаткове налаштування моделі, враховуючи специфічні умови навколишнього середовища та параметри ґрунту.

Запропонований метод не тільки забезпечує високу точність класифікації, але й підвищує безпеку при розмінуванні, оскільки використовує пасивний підхід виявлення, усуваючи ризик випадкової детонації.

### Подяка

Дана стаття підготована завдяки грантової підтримки Національного Фонду Досліджень України, реєстраційний номер проєкту 273/0024 від 1/08/2024 (2023.04/0024) «Методи та засоби активного та пасивного розпізнавання мін на основі глибоких нейронних мереж» за конкурсом «Наука для зміцнення обороноздатності України».

### Список літератури:

[1] International Campaign to Ban Landmines, Landmine Monitor 2023 (ICBL-CMC: January 2023). [Online]. Available: [www.the-monitor.org](http://www.the-monitor.org).

[2] Ishchenko, A.V.; Kobets, M.V. "Means and Methods of Detecting Explosive Substances and Devices in the Fight against Terrorism: Training and Practice," NAVSU Publishing House: Kyiv, Ukraine, 2005; p. 144.

[3] Krausa, M.; Massong, H.; Rabenecker, P.; Ziegler, H. "Chemical methods for the detection of mines and explosives. In Detection of Explosives and Landmines," Eds.; NATO Science Series; Springer: Dordrecht, The Netherlands, 2002; Volume 66. [https://doi.org/10.1007/978-94-010-0397-1\\_1](https://doi.org/10.1007/978-94-010-0397-1_1).

[4] Masunaga, S.; Nonami, K. "Controlled metal detector mounted on mine detection robot," Int. J. Adv. Robot. Syst. 2007, 4, pp. 237–245. <https://doi.org/10.5772/5692>

[5] Yilmaz, C.; Kahraman, H.T.; Söyler, S. "Passive Mine Detection and Classification Method Based on Hybrid Model," IEEE Access 2018, 6, 47870–47888. <https://doi.org/10.1109/ACCESS.2018.2866538>.

[6] Peleshchak R.M., Lytvyn V.V., Nazarkevych M.A., Peleshchak I.R., Nazarkevych H.Y. "Influence of the Symmetry Neural Network Morphology on the Mine Detection Metric," Symmetry. 2024; 16(4):485. <https://doi.org/10.3390/sym16040485>

[7] Bestagini, P., Lombardi, F., Lualdi, M., Picetti, F., & Tubaro, S. "Landmine Detection Using Autoencoders on Multi-polarization GPR Volumetric Data," 2018. [Online]. Available: <https://arxiv.org/pdf/1810.01316.pdf>

[8] S. Gürkan, M. Karapinar and S. Doğan, "Classification of explosives materials detected by magnetic anomaly method," 2017 4th International Conference on Electrical and Electronic Engineering (ICEEE), Ankara, Turkey, 2017, pp. 347–350, doi: 10.1109/ICEEE2.2017.7935848.

[9] Xiaohui Wei, Changbao Zhou, Hengshan Yue, Joey Tianyi Zhou. "TC-SEPM: Characterizing soft error resilience of CNNs on Tensor Cores from program and microarchitecture perspectives," Journal of Systems Architecture, Volume 145, 2023, 103024, ISSN 1383-7621. <https://doi.org/10.1016/j.sysarc.2023.103024>.

[10] Martin Tetard, Veronica Carlsson, Mathias Meunier, Taniel Danelian. "Merging databases for CNN image recognition, increasing bias or improving results?" Marine Micropaleontology, Volume 185, 2023, 102296, ISSN 0377-8398. <https://doi.org/10.1016/j.marmicro.2023.102296>.

[11] Huanhua Liu, Wei Wang, Hanyu Liu, Shuheng Yi, Yonghao Yu, Xunwen Yao. "A Degradation Type Adaptive and Deep CNN-Based Image Classification Model for Degraded Images," CMES - Computer Modeling in Engineering and Sciences, Volume 138, Issue 1, 2023, Pages 459–472, ISSN 1526-1492. <https://doi.org/10.32604/cmescs.2023.029084>.

[12] Dataset Land Mines - UCI Machine Learning Repository, 2022. [Online]. Available: <https://archive.ics.uci.edu/dataset/763/land+mines-1>

[13] C. Yilmaz, Y. Sönmez, H. T. Kahraman, S. Soyler, and U. Güvenç, "Developing of decision support system for land mine classification by metaheuristic classifier," in Proc. Int. Symp. Innov. Intell. Syst. Appl. (INISTA), Aug. 2016, pp. 1–5.

[14] A. Barnawi, K. Kumar, N. Kumar, B. Alzahrani, and A. Almansour "A Deep Learning Approach for Landmines Detection Based on Airborne Magnetometry Imaging and Edge Computing," Comput. Model. Eng. Sci., vol. 139, no. 2, pp. 2117–2137. 2024. <https://doi.org/10.32604/cmescs.2023.044184>



Б. С. Карпішен

ХНАДУ, м. Харків, Україна, karpishen.bogdan@gmail.com,  
ORCID iD: 0009-0001-1790-9048

## ПОКРАЩЕННЯ СПОЖИВАННЯ ЕНЕРГІЇ ЗА ДОПОМОГОЮ ADAS В АВТОНОМНОМУ ВОДІННІ КОМУНІКАЦІЙНОЇ СИСТЕМИ

У статті проведений аналіз тенденцій розвитку технологій енергозбереження та рекуперації на електричному автомобільному транспорті. Розглянуто та визначено перспективні напрямки розвитку технологій енергозбереження зокрема з використанням систем допомоги водієві. Збільшення наповненості автомобілів електронними системами допомоги водієві збільшує загальні витрати енергії. Наведені рекомендації з вибору оптимальних технологій та методів енергозбереження для автотранспорту. Для деяких методів економія спожитої енергії є доповненням до основного напрямку роботи. Поєднання інтелектуальних транспортних засобів та відповідних засобів організації дорожнього руху може сприяти подальшій реалізації транспортних переваг інтелектуальних транспортних засобів.

ЕКОНОМІЯ ЕНЕРГОСПОЖИВАННЯ, ЕЛЕКТРОТРАНСПОРТ, АВТОНОМНІ АВТОМОБІЛІ, ADAS, ЕКО-ВОДІННЯ, РЕКУПЕРАЦІЯ

**B. S. Karpishen. Improving energy consumption with ADAS in autonomous driving communication system.** The article analyses the trends in the development of energy saving and recovery technologies in electric motor vehicles. Promising directions for the development of energy saving technologies, in particular with the use of driver assistance systems, are considered and identified. Increasing the occupancy of cars with electronic driver assistance systems increases the overall energy consumption. Recommendations for choosing the optimal technologies and methods of energy saving for motor vehicles are given. For some methods, saving energy is a complement to the main area of work. The combination of intelligent vehicles and appropriate traffic management tools can help to further realise the transport benefits of intelligent vehicles.

ENERGY SAVING, ELECTRIC TRANSPORT, AUTONOMOUS VEHICLES, ADAS, ECO-DRIVING, RECOVERY

### Вступ

Перехід від палива з нафтопродуктів, до інших видів палива зумовлений постійно зростаючою кількістю транспортних засобів і сучасною потребою в більш ефективних і екологічно чистих видах палива. Альтернативні види палива, наприклад водень, більше підходять до теперішніх та застарілих видів авто, а заміна ДВЗ електродвигуном більше притаманна новому автопрому та майбутнім ринковим надходженням.

Електрифікація стала одним з основних напрямків розвитку автомобільної промисловості і, як очікується, стане провідним видом енергії для транспортних засобів у майбутньому [1].

Економічні та екологічні міркування спонукають автомобільну промисловість переходити до використання електромобілів. Однак поява та розповсюдження електромобілів невизначені. Розвиток і розповсюдження електромобілів мають певні перешкоди. Серед усіх цих бар'єрів велику увагу привернули дальність і ефективність цих транспортних засобів.

Також це зумовлено підвищенням рівня автоматизації автомобіля, автономні вантажівки і електротаксі, це все несе за собою нові виклики. Підвищення автоматизації веде до збільшення потреби в енергоресурсі на одиницю транспорту [2].

### 1. Аналіз публікацій

Дослідженням пов'язаним з розробкою та експлуатацією присвячено багато робіт в зарубіжних та вітчизняних виданнях. Це стосується як вантажних і легкових автомобілів так і особливо електротранспорту за принципом доставки останньої милі.

Питання споживання енергії можна вирішити або шляхом підвищення ефективності електродвигунів, за рахунок підвищення ККД акумуляторів шляхом термоконтролю, контролю процедури зарядки, роблячи їх ефективнішими за допомогою електричних компонентів і напівпровідників з кращою потужністю [3]. Інший підхід до підвищення ефективності електромобілів полягає в оптимізації споживання енергії за допомогою стратегій управління енергією під час різних циклів руху.

У міру того, як електромобілі набувають популярності, виробники все частіше використовують функції Advanced Driver Assistance Systems (ADAS) для вирішення проблем, унікальних для електричної мобільності [4].

Автономність транспортного засобу досягається завдяки використанню функцій ADAS, які вимагають безперервних вимірювань і обчислень у транспортному засобі. Функції ADAS включають адаптивний круїз-контроль, класифікацію об'єктів, виявлення перешкод, картографування, планування шляху та одночасну локалізацію та картографування

(SLAM). Ці програми зазвичай покладаються на камери, LiDAR і радарні датчики, які також генерують великі обсяги необроблених даних для розрахунку і аналізу обчислювальними блоками транспортного засобу. Також вони можуть використовувати зв'язок між іншим транспортним засобом (V2V) і зв'язок між інфраструктурою (V2I). Автомобіль оснащений цими потужними технологіями, може мати доступ до великої кількості даних і може взаємодіяти з іншими автомобілями.

Дослідження [5] показує необхідність складних розрахункових схем для точного контролю та забезпечення ефективної роботи нових надійних та економічно ефективних напівпровідникових технологій. ADAS системи мають потенціал для значного підвищення енергоефективності та можуть зайняти цю нішу.

Тим часом збільшення наповненості автомобілів електронними системами допомоги водієві збільшує загальні витрати енергії особливо зі збільшенням кількості і комплексності самої системи за кількістю технологій і загальною обчислювальною потужністю. Звісно зі збільшенням рівня автоматизації потреба в енергії також зростає [6].

Дослідження [7] показують, що бортове споживання енергії перевищує 1000 Вт, а загальне споживання енергії одним умовним автоматизованим транспортним засобом, що об'єднує всі три категорії, може становити близько 2500 Вт-год на 100 км водіння. Високе бортове споживання енергії зумовлене використанням алгоритмів і пристроїв обробки, які потребують інтенсивних обчислень, таких як графічні процесори, необхідні для перцептивних і візуальних програм.

## 2. Мета та постановка задачі

Метою роботи є аналіз та обґрунтування сучасного електротранспорту та висвітлення різних методів економії енергії, їх переваг та недоліків та оцінки потенційного впливу на розвиток і впровадження електромобілів.

Для досягнення поставленої мети необхідно:

- проаналізувати перспективні підходи до енергозбереження в з урахуванням використання сучасних технологій;
- визначити переваги і недоліки кожного з підходів;
- зробити висновки, що до доцільності використання та майбутні перспективи кожного методу.

## 3. Класифікація підходів енергозбереження

Надалі пропонуються підходи до енергозбереження які базуються на автономному транспорті (АТ) та деякі підходи до їх вирішення.

- Дорожня карта і оптимізація маршруту. Інформація про прогнозування доріг може допомогти

АТ оптимізувати споживання енергії на основі інформації про дорожній рух, кривизни дороги та висоти за допомогою 3D-карт доріг. Дані передаються через зв'язок V2I. Їх можна детально проаналізувати, щоб зменшити споживання енергії. В мінімальному випадку можна використовувати заздалегідь завантажену карту навігатора.

– Часовий трафік і гармонізація дорожнього руху. Кожен транспортний засіб у парку або в трафіку може оптимізувати свій цикл водіння, щоб зменшити кількість операцій гальмування та прискорення. Існує кілька підходів для планування транспортних засобів у русі або перед перехрестями. Перший заснований на зміні швидкості та прискорення, але він може бути досягнутий лише шляхом передбачення без співпраці. Другий метод заснований на оптимізації споживання енергії під час попутного руху, для якого підходить транспортний засіб, оскільки виїзд автомобіля на іншу смугу вплине на рух попереду. В цьому можуть допомогти зв'язок V2I та радари. Ідеальний сценарій має на увазі кооперативний метод керування [8, 9].

– Слідування за автомобілем. Існує велика кількість рішень круїз-контролю з функцією слідування за автомобілем попереду. Тому в цьому випадку потрібно приділити увагу кооперативному адаптивному круїз-контролю. Системи мають потенціал для збільшення пропускної здатності руху, дозволяючи зменшити прохід між транспортними засобами та узгодити їх швидкість. Хоча це більш важливо для вантажівок, оскільки ці транспортні засоби страждають від високого коефіцієнта лобового опору через їхні габарити. Однак ця технологія добре підходить і для громадського транспорту великих міст та шосе.

– «Ідеальне» перехрестя. Завдяки попередженню кожен транспортний засіб отримує часовий сигнал з перехрестя і може зменшити витрату енергії. Цього можна досягти за рахунок меншої кількості непотрібних прискорень і гальмувань. У режимі співпраці система керування перехрестям також може зробити набагато більше для керування попутним транспортом, зменшивши час очікування. У найкращому випадку, при наявності 100% АТ на перехресті, червоне світло може бути пропущене за допомогою зв'язку V2I і V2V.

– Формування корисних навичок. Впровадження автономного транспорту дозволить простіше впровадити і навчати водіїв корисним звичкам економічного водіння. Еко-водіння один з таких принципів, що дозволить автономному транспорту отримуючи інформацію про навколишню обстановку оптимізувати дорожній рух. Якщо транспортний засіб, долучений до мережі, може передбачати швидкість і наміри інших транспортних засобів через зв'язок V2V, він зможе робити це більш ефективно. Отримавши інформацію



про час руху, швидкість автомобіля та пройдено відстань, можна оптимізувати споживання енергії.

– Внутрішня комплектація. Вибір апаратної частини комплектації є важливою складовою подальшого використання енергії транспортом. Технології рекуперації енергії допоможуть отримувати енергію від самого автомобіля наприклад під час спуску, кочення, гальмування і т. д. Правильне налаштування роботи приладів та оптимізація їх роботи допоможе використувати їх максимально ефективно та економічно.

#### 4. Споживання енергії

Першим кроком у визначенні споживання енергії транспортним засобом є визначення рівняння, що лежить в основі. Основне енергетичне рівняння для автомобіля  $E_t = E_c$ , де  $E_t$  – тягова енергія, що генерується мотором або двигуном і передається на колеса, а  $E_c$  – енергія, споживана автомобілем.

Загальне споживання енергії визначається як:

$$E_c = E_{\text{пов}} + E_{\text{коч}} + E_{\text{пр}} + E_{\text{в}}, \quad (1)$$

де  $E_{\text{пов}}$  – аеродинамічна енергія,  $E_{\text{коч}}$  – енергія, що витрачається на подолання сили опору коченню;  $E_{\text{пр}}$  – енергія, що витрачається на прискорення транспортного засобу;  $E_{\text{в}}$  – енергія, що витрачається на зміну висоти транспортного засобу, а саме, на підйом транспортного засобу [8].

В автомобілях з електродвигуном ця формула виглядає трохи інакше через те що через інерційний ефект частина енергії може бути збережена в умовах рекуперативного гальмування

$$E_c = E_{\text{пов}} + E_{\text{коч}} + E_{\text{пр}} + E_{\text{в}} - E_{\text{рг}}, \quad (2)$$

де  $E_{\text{рг}}$  – відновлена енергія за допомогою рекуперативного гальмування. Кількість відновленої енергії може бути ключовою для вирішення та мінімізації споживання енергії транспортним засобом.

Наприклад для зменшення аеродинамічних втрат середня швидкість та її відхилення є єдиними параметрами, які можуть мінімізувати цей енергетичний елемент за допомогою динаміки автомобіля. Варто зазначити, що мінімізація середньої швидкості призведе до збільшення часу в дорозі та зменшення транспортного потоку, що в більшості випадків небажано, але зменшення відхилення швидкості є більш доцільним.

#### 5. Дорожня карта і оптимізація маршруту

Планування маршруту руху було і досі залишається ефективним способом економії палива та енергії. Важливим етапом планування маршруту є так звані доставки останньої милі, бо вони в логістиці вважаються найдорожчими. У логістиці ці питання вирішуються за допомогою перегляду варіантів доставки та оцінкою компромісів, пов'язаних із конкретним маршрутом [10, 11].

Іншим варіантом економії живлення може слугувати система контролю відстеження траєкторії руху, для зменшення відхилення поперечного та поздовжнього положення між фактичною траєкторією руху та еталонною траєкторією. Якщо відхилення занадто велике, це призведе до більших витрат електроенергії силової батареї та скорочення запасу ходу.

Побудована [12] модель показує деяке зменшення енергоспоживання батареї та підвищення точності відстеження траєкторії. Хоча різниця зекономленої енергії доволі мала, перевагами цього методу є те, що він підходить до автомобілів різного рівня автоматизації. Це також не є його основною задачею, а лише доповненням до контролю положення.

#### 6. Часовий трафік і гармонізація дорожнього руху

Протягом останніх кількох років розглядалися різні аспекти технології формування автоколони. Транспортні засоби у взводі можуть отримати вигоду від підвищення ефективності за рахунок зменшення сили лобового опору, методів оптимізації споживання палива на основі даних з V2V і V2I з'єднань, а також від співпраці при формуванні парку транспортних засобів.

Деякі роботи, такі як [13], показують підвищення паливної ефективності при моделюванні 100%-го ефекту колони до 10 машин з використанням модельного прогнозованого управління (МПУ), що зображено на рис. 1. Повідомляється про збільшення ефективності електродвигунів до 50% у деяких їздових циклах.

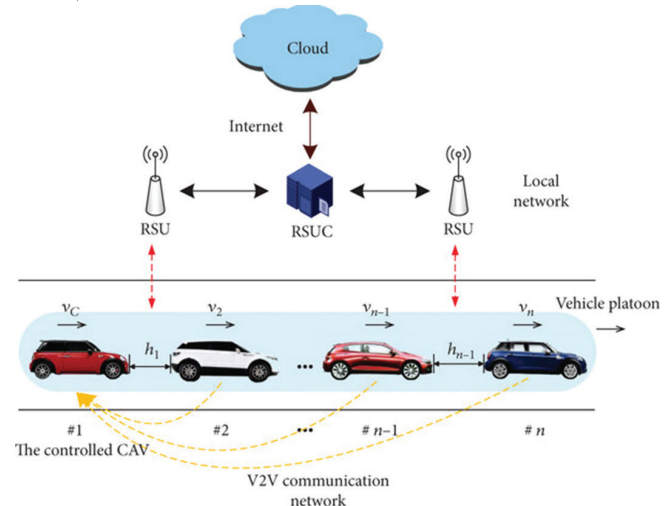


Рис. 1. Автоколонна авто об'єднаних в одну мережу з використанням V2V зв'язку

Гармонізація руху зменшить випадки гальмування і призведе до більш плавного циклу руху з дуже невеликою різницею в швидкості прискорення та меншого відхилення від середньої швидкості.

Оскільки транспортні засоби, що їдуть попереду, можуть отримувати інформацію про умови дорожнього руху через V2V, транспортні засоби, що їдуть

позаду, можуть адаптуватися до дорожніх умов. Зменшений інтервал між автомобілями в колоні призведе до меншого аеродинамічного опору.

Ці фактори призведуть до зниження споживання енергії та викидів порівняно з транспортними засобами без V2V [14].

Слід зазначити, що зниження споживання енергії для групи з кількох транспортних засобів на великій відстані є більш ефективним, ніж зменшення споживання енергії на короткій відстані лише для обмеженої кількості транспортних засобів.

### 7. Слідування за автомобілем

Адаптивний круїз контроль автомобілів може скоротити середній час у дорозі та зменшити середнє споживання енергії для одного транспортного засобу в більшості сценаріїв. Встановлено [15], що АТ (рівня 2 і вище [16]) можуть скоротити середній час у дорозі та зменшити середнє споживання енергії в більшості сценаріїв у перерахунку на один транспортний засіб. Але що стосується загальних переваг, то спостерігається негативний вплив на споживання енергії через збільшення фактичної пропускної здатності доріг, що приблизно еквівалентно збільшенню інтенсивності руху. Коли потік трафіку перевищує межу насичення, досягаються позитивні економічні вигоди, які стають все більшими в міру розповсюдження на ринок АТ.

Хоча середнє споживання енергії в багатьох сценаріях зменшується, зростання інтенсивності руху збільшує загальне споживання енергії, що нівелює переваги енергозбереження.

Однак тема може бути розширена глобальними і локальними стратегіями регулювання адаптивного круїз-контролю, що теоретично [17] перетворюється на реальність, оскільки електромобілі поступово стають все більш популярними в міру підвищення рівня автоматизації.

Коли рівень автоматизації підвищується до 90%, результати моделювання змішаної моделі кращі, ніж результати моделі без АТ. Таким чином, в умовах одного виїзду, при використанні локальної стратегії прокладання маршруту, модель змішаного транспортного потоку спрямована на поліпшення умов руху на дорозі, а частка автономних транспортних засобів також є умовною. Лише коли частка автономних транспортних засобів перевищує 90%, гібридна модель транспортного потоку може покращити ефективність руху та енергоспоживання краще, ніж традиційні моделі транспортних засобів.

Подальші розробки в безпілотних автомобілях включають використання нейронних мереж на основі навчання з підкріпленням для зменшення коливань трафіку, спричинених водіями, та покращення споживання енергії. Експериментальні результати показують, що модель на основі навчання з підкріпленням значно

зменшує обмеження симуляції та має можливості самонавчання та само коригування. Ефективність подорожей підвищиться шляхом пом'якшення негативних наслідків коливань трафіку та зниження середнього споживання електроенергії [18].

### 8. «Ідеальне» перехрестя

Перехрестя є одним з найскладніших ділянок дороги, доволі часто такі ділянки стають епіцентрами дорожніх небезпек і заторів. Максимально ефективне проходження ділянки дасть найбільший ефект плавності поїздки, а отже і економічний ефект. Успішний прохід перехрестя потребує великої кількості даних, що потребує розрахункової потужності, наприклад загальної мережі керування або локального серверу перехрестя для обробки даних та спілкування з АТ.

Для ефективного подолання перехрестя можна використовувати оптимізацію методом змішаного цілочисельного лінійного програмування (MILP) [19]. За допомогою цього методу було зафіксовано зниження споживання енергії до 20%. За допомогою MILP контролер створює реальне зображення перехрестя на основі з'єднання V2V, швидкості та місцезнаходження кожного транспортного засобу та завчасно надає доступ кожному транспорту, щоб мінімізувати споживання енергії. У [19] бажаний час прибуття транспортних засобів на перехрестя додається до мети оптимізації, щоб уникнути затримок і покращити продуктивність і безпеку перехрестя.

Однак реалізація потребує футуристичного середовища та максимального розповсюдження АТ та підключених користувачів. Що і нівелює хороші результати.

### 9. Формування корисних навичок

Одним із способів покращити економію палива є використання зворотного зв'язку водія для формування ефективних звичок (еко-водіння). Імплементация ADAS в автотранспорт допоможе в просуванні екологічного водіння.

Цей принцип на сьогоднішній день є важливим впровадженням який є основою для збереження енергії. Вищий рівень автоматизації відповідно дає кращий рівень еко-водіння особливо в містах завдяки більш плавному водінню. Також рівень економії електродвигунів зберігається вищим ніж у бензинових при низьких швидкостях. До основних прийомів можна віднести:

- уникання надмірного прискорення та різкого гальмування;
- запобігання холостому ходу;
- дотримання стабільного швидкісного режиму;
- оптимізоване планування маршруту;
- мінімізація коротких зупинок.

Параметр еко-водіння можна розрахувати використовуючи так званий «фактор агресивного водіння» [20].

$$AF = \frac{1}{D} [E_w - E_w^*] = \frac{\int (Av + Bv^2 + Cv^3 + Mav)^+ dt}{\int v dt} - (A + B\bar{v} + C\bar{v}^2), \quad (3)$$

$$E_{wheel} = \int (Av + B\bar{v}^2 + C\bar{v}^3 + Mav)^+ dt, \quad (4)$$

$$E_{wheel}^* = (Av + B\bar{v}^2 + C\bar{v}^3) dt = (A + B\bar{v} + C\bar{v}^2) \int \bar{v} dt. \quad (5)$$

де  $AF$  – різниця в споживанні енергії на колесах між дотриманням профілю руху  $E_{wheel}$  і руху зі сталою середньою швидкістю  $E_{wheel}^*$  нормалізовано за відстанню шляху  $D$ . Зі зменшенням  $AF$  росте екологічність водіння.

$A(N)$ ,  $B(N)$ ,  $C(N)$  – коефіцієнти вибігу, що відповідають коефіцієнтам кочення, обертання та аеродинамічного опору відповідно. Коефіцієнти вибігу це специфічні параметри надані базою даних палива «Агентства з охорони навколишнього середовища» [21]. Надалі  $a$  ( $\text{м/с}^2$ ),  $v$  ( $\text{м/с}$ ) і  $\bar{v}$  ( $\text{м/с}$ ) – миттєве позитивне прискорення, швидкість і середня швидкість у дорозі відповідно.  $M$  ( $\text{кг}$ ) маса транспортного засобу. Для прикладу було створено три варіанти сценаріїв еко-водіння (низький, середній, високий рівні), що порівнювались зі звичайним водінням.

Сценарій з низьким рівнем екологічного водіння зменшує параметр агресивного водіння на 29% (середнє значення) завдяки плавній їзді, але не змінює середню швидкість у дорозі. Середній сценарій зменшує  $AF$  на 38% і збільшує середню швидкість поїздки на 10% порівняно з вихідними профілями швидкості. Сценарій високого рівня зменшує автофокус на 61% і збільшує середню швидкість поїздки на 20% (рис. 2).

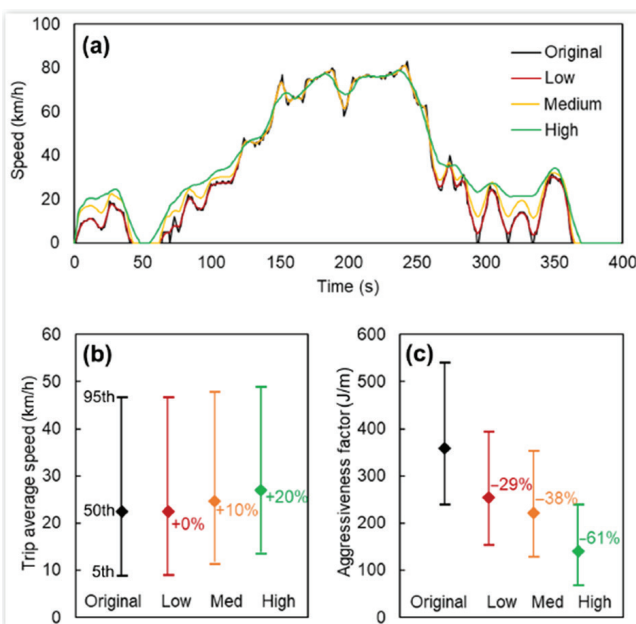


Рис. 2. Данні сценаріїв еко-водіння.

Данні показують середні значення показників  $AF$  для близько 8000 поїздок [22]

Вплив автоматизації на енергоспоживання менший для електромобілів, ніж для бензинових, але різниця в перевагах екологічного водіння відносно невелика. Відомо, що автоматизовані транспортні засоби витрачають до 20% більше енергії на низьких швидкостях порівняно з неавтоматизованими електромобілями. Ця динаміка стає меншою зі збільшенням швидкості.

За деякими розрахунками [20] еко-водіння може зменшити потребу в енергії на 10 – 21% для електромобілів в залежності від рівня «агресивності» водіння.

## 10. Внутрішня комплектація

Традиційний ADAS відстежує та використовує дані, які можуть допомогти визначити ситуації, коли транспортному засобу може знадобитися сповільнити швидкість. Майбутній поворот зазвичай свідчить про зниження швидкості. Значні повороти на дорозі або повороти на  $90^\circ$  гарантують зниження швидкості, що робить функцію повороту необхідною для оптимального прогнозування маршруту.

Гальмівна система відома як джерело розсіювання енергії транспорту. Двигун, який можна застосовувати як електродвигун і генератор, має чотири квадрантні робочі функції. Коли двигун працює в другому або четвертому квадранті (тобто в режимі генератора), інші форми енергії можуть бути перетворені в електричну; наприклад, під час рекуперативного гальмування кінетична енергія перетворюється на електричну енергію, а рекуперация відпрацьованого тепла, що виділяється вихлопним газом, перетворюється на електричну енергію. Використання двигунів у системах підвіски також перетворює механічну та вібраційну енергію в електричну.

Рекуперативна гальмівна система (РГС) може збільшити запас ходу транспортних засобів і зменшити кількість енергії, яка розсіюється в навколишнє середовище у вигляді тепла. РГС може зберігати регеновану енергію в різних формах, наприклад, у вигляді маховика, що обертається, або у вигляді гідравлічної чи пружинної системи. Однак найбільш поширеною і розвинутою формою рекуперативної енергії є система накопичення енергії, яка може легко виробляти електричну енергію для подальшого використання. Суперконденсатори та акумулятори можуть зберігати та видавати електричну енергію з високою ефективністю. Гібридна конструкція, що використовує як батареї, так і супер/ультраконденсатори, є одним з підходів до підвищення ефективності зберігання регенованої енергії [23].

Послідовне гальмування з максимальною рекуперативною енергією: У цьому методі комбінованого гальмування метою є максимальна рекуперация енергії. Тому, коли гальмівне зусилля менше за максимально

можливе, автомобіль використовує гальмівне зусилля РГС на ведених колесах (у більшості випадків передніх). Якщо гальмівного зусилля РГС недостатньо для забезпечення необхідного гальмівного сповільнення, задіюється механічна гальмівна система на задніх колесах, щоб забезпечити додаткове зусилля, необхідне для досягнення бажаного сповільнення.

Ефективність роботи системи значно залежить від типу дорожнього полотна, початкової швидкості гальмування та маси автомобіля. Таким чином регенерація енергії при максимально ефективному сценарії доходить до 78% у випадку низької автоматизації [24].

ADAS відкривають нові можливості для більш ефективного використання регенераційних гальм. На ранньому етапі транспортний засіб (водій або автоматизована система водіння) має бути проінформований про майбутні події чи ситуації. У дослідженні в [25] була введена система енергетично оптимального уповільнення для посилення рекуперативних гальм. Він показав 33% покращення ефективності рекуперативного гальмування для CAV порівняно з людьми-водіями.

Іншим технологічним рішенням виступає електромагнітна активна підвіска, може значно покращити динамічні характеристики автомобіля з точки зору комфорту їзди та стабільності керування. Також технологія відновлення енергії вібрації може зменшити споживання енергії транспортним засобом на 7% – 10% [26], компенсувати споживання енергії активною підвіскою та підвищити ефективність використання джерела живлення автомобіля для досягнення мети енергозбереження.

Процес рекуперації енергії існуючих активних підвісок повітряного та гідравлічного типу відносно складний: енергію вібрації необхідно перетворити у внутрішню, а потім за допомогою двигуна перетворити в електричну. Обидва потребують додавання допоміжних установ, що ускладнює контроль і знижує ефективність перетворення енергії [27].

Електромагнітна підвіска використовує струм як робоче середовище для полегшення лінійного керування. У той же час, відповідно до характеристик

самого двигуна, енергія вібрації може бути безпосередньо перетворена в електричну енергію, яка має характеристики компактною структури та можливості регенерації енергії. Таким чином, система рекуперації електромагнітної енергії підвіски, яка враховує характеристики активної динаміки та регенерації енергії, відповідає потребам розробки майбутніх шасі та поступово стає гарячою точкою для вивчення вчених з різних країн.

### Висновки

У підсумку запропоновані підходи мають потенціал для розвитку і є важливими в контексті сучасних інтелектуальних транспортних систем. Кожен з методів має свої переваги та недоліки, хоча не потрібно робити вибір одного, тому майже всі методи можуть працювати сумісно у рамках однієї інтелектуальної системи. Це стосується і напрямку розвитку. Для деяких методів економія спожитої енергії є доповненням до основного напрямку роботи. Однак поєднання інтелектуальних транспортних засобів та відповідних засобів організації дорожнього руху може сприяти подальшій реалізації транспортних переваг інтелектуальних транспортних засобів. Результати огляду методів можна побачити в табл. 1. Оцінка ризиків і питання безпеки є одними з найважливіших питань у розробці автономного транспорту. Варто зазначити, що кожен з методів повинен враховувати безпеку як важливе питання. Автоматизація транспорту та сприяння використанню відновлюваної енергії не є самоціллю, а частиною необхідної відповіді – з точки зору стратегії та скоординованих інструментів для побудови стійкої мобільності. Окрім безпеки, ADAS також відіграє важливу роль в оптимізації ефективності електромобілів, що є ключовим моментом у широкому впровадженні стійкої мобільності. Завдяки динамічному регулюванню енергоспоживання ADAS максимізує ефективність електромобілів, усуваючи проблеми із запасом ходу та максимізуючи відстань, яку можна подолати на одному заряді. Вкрай важливо підвищити швидкість встановлення модулів V2X на транспортні засоби та прискорити інтелектуальну модернізацію транспортної інфраструктури.

Таблиця 1

Енергетичний вплив переліку методів

Категорія	Вплив	Опис	Енергетичний вплив
Дорожня карта і оптимізація маршруту	Оптимізація доставки останньої милі	Методи перегляду варіантів та оцінка компромісів	Параметр економії має високу залежність від типу і розміру доставки тому має лише загальні рекомендації до оптимізації
	Система контролю і відстеження траєкторії руху	Точне відстеження транспортним засобом траєкторії є корисним для економії електроенергії та збільшення запасу ходу акумулятора	Біля 1% на 1 милію

Категорія	Вплив	Опис	Енергетичний вплив
Часовий трафік і гармонізація дорожнього руху	Формування автоколонни	Використання колони до 10 машин з використанням модельного прогнозованого управління	До 50%
Слідування за автомобілем	Адаптивний круїз контроль	Ефективність залежить від інтенсивності руху	Від 2% до 36%
	Стратегії регулювання маршруту	Вищий рівень автоматизації та більша кількість автономних транспортних засобів можуть зменшити споживання енергії.	До 10%
«Ідеальне перехрестя»	Оптимізація методом змішаного цілочисельного лінійного програмування	MILP контролер та з'єднання V2V надає місцезнаходження кожного транспортного засобу і будує алгоритм ефективного проходження перехрестя	До 20%
Формування корисних навичок	Еко-водіння	Дотримання правил еко-водіння	Від 10% до 21%
Внутрішня комплектація	Рекуперативна гальмівна система	Накопичення енергії отримуючи енергію при гальмуванні	До 78% відновлення енергії
	Електромагнітна гальмівна підвіска	Коливання автомобіля під час руху регульованих можна перетворити	Від 7% до 10%

**Список літератури:**

- [1] Z. Liu, X. Liu, and F. Zhao, "Research on NEV Platform Development Strategies for Automotive Companies," *World Electric Vehicle Journal*, vol. 12, no. 4, p. 201, Dec. 2021, [Online]: <https://doi.org/10.3390/wevj12040201>.
- [2] Rolando Campos Canales, and Gabriel Pérez. "Autonomous Vehicles and Alternative Energies for Logistics in the Wake of the Pandemic." *Facilitation of Transport and Trade in Latin America and the Caribbean*, vol. 4, no. 338, 8 Nov. 2021. ISSN: 1564-4243.
- [3] Yadlapalli, Ravindranath Tagore, et al. "A Review on Energy Efficient Technologies for Electric Vehicle Applications." *Journal of Energy Storage*, vol. 50, June 2022, <https://doi.org/10.1016/j.est.2022.104212>
- [4] Shih-Chieh Lin et al. 2018. The Architectural Implications of Autonomous Driving: Constraints and Acceleration. *SIGPLAN Not.* 53, 2 (February 2018), 751–766. <https://doi.org/10.1145/3296957.3173191>
- [5] Sarnago, H.; Lucía, Ó. Popa, I.O.; Burdío, J.M. Constant-Current Gate Driver for GaN HEMTs Applied to Resonant Power Conversion. *Energies* 2021, 14, 2377. <https://doi.org/10.3390/>
- [6] VAN DER AALST, Wil. Six Levels of Autonomous Process Execution Management (APEM). arXiv preprint arXiv:2204.11328, 2022. <https://doi.org/10.48550/arXiv.2204.11328>
- [7] D. Katare, D. Perino, J. Nurmi, M. Warnier, M. Janssen and A. Y. Ding, "A Survey on Approximate Edge AI for Energy Efficient Autonomous Driving Services," in *IEEE Communications Surveys & Tutorials*, vol. 25, no. 4, pp. 2714–2754, Fourthquarter 2023, doi: 10.1109/COMST.2023.3302474.
- [8] Hamed Faghihian, and Arman Sargolzaei. "Energy Efficiency of Connected Autonomous Vehicles: A Review." *Electronics*, vol. 12, no. 19, 29 Sept. 2023. [Online]: <https://doi.org/10.3390/electronics12194086>.
- [9] Jiang Y, Zhao B, Li M, Yao Z. A Two-Level Model for Traffic Signal Timing and Trajectories Planning of Multiple CAVs in a Random Environment. *Journal of Advanced Transportation*. 2021 Apr 26;2021:1–13. [Online]: <https://doi.org/10.1155/2021/9945398>
- [10] Gee, K. Faust, and M. Webber, "A framework for determining energy use in rural food delivery services: capturing system interdependencies through an agent-based discrete-event approach," *Environmental Research: Infrastructure and Sustainability*, Sep. 2021, doi: <https://doi.org/10.1088/2634-4505/ac2b10>.
- [11] Á. Halldórsson and J. Wehner, "Last-mile logistics fulfilment: A framework for energy efficiency," *Research in Transportation Business & Management*, vol. 37, p. 100481, Apr. 2020, doi: <https://doi.org/10.1016/j.rtbm.2020.100481>.
- [12] Ma, Hao, et al. "Battery Energy Consumption Analysis of Automated Vehicles Based on MPC Trajectory Tracking Control." *Electrochem*, vol. 3, no. 3, 1 Sept. 2022, pp. 337–346, [www.mdpi.com/2673-3293/3/3/23](http://www.mdpi.com/2673-3293/3/3/23), <https://doi.org/10.3390/electrochem3030023>. Accessed 3 Dec. 2023.
- [13] Dollar, R.A.; Vahidi, A. Quantifying the impact of limited information and control robustness on connected automated platoons. In *Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, Yokohama, Japan, 16–19 October 2017; pp. 1–7. doi: 10.1109/ITSC.2017.8317604.
- [14] Lee, Jooyong, and Kara Kockelman. *Energy Implications of Self-Driving Vehicles*. The University of Texas at Austin, Jan. 2019. [Online]: [https://www.cae.utexas.edu/prof/kockelman/public\\_html/TRB19EnergyAndEmissions.pdf](https://www.cae.utexas.edu/prof/kockelman/public_html/TRB19EnergyAndEmissions.pdf)
- [15] Song H, Zhao F, Zhu G, Liu Z. Impacts of Connected and Autonomous Vehicles with Level 2 Automation on Traffic Efficiency and Energy Consumption. *Journal of Advanced Transportation*. Apr. 2023;2023:e6348778. [Online]: <https://www.hindawi.com/journals/jat/2023/6348778/> <https://doi.org/10.1155/2023/6348778>

- [16] "SAE: Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems", SAE On-road Automated Vehicles Standards Committee and others, 2021. [Online]. Available: [https://www.sae.org/standards/content/j3016\\_202104/](https://www.sae.org/standards/content/j3016_202104/)
- [17] Chen B, Chen Y, Wu Y, Xiu Y, Fu X, Zhang K. The Effects of Autonomous Vehicles on Traffic Efficiency and Energy Consumption. *Systems*. 2023 Jul 1;11(7):347. [Online]: <https://www.mdpi.com/2079-8954/11/7/347> doi.org/10.3390/systems11070347
- [18] Qu X, Yu Y, Zhou M, Lin CT, Wang X. Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: A reinforcement learning based approach. *Applied Energy*. 2020 Jan;257:114030. [Online]: <https://doi.org/10.1016/j.apenergy.2019.114030>
- [19] S. A. Fayazi and A. Vahidi, "Mixed-Integer Linear Programming for Optimal Scheduling of Autonomous Vehicle Intersection Crossing," in *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 3, pp. 287-299, Sept. 2018, doi: 10.1109/TIV.2018.2843163.
- [20] He, Xiaoyi, et al. "Energy Consumption Simulation for Connected and Automated Vehicles: Eco-Driving Benefits versus Automation Loads." *SAE International Journal of Connected and Automated Vehicles*, vol. 6, no. 1, 9 May 2022.[Online]: <https://doi.org/10.4271/12-06-01-0002>.
- [21] U.S. Environment Protection Agency (EPA), "Compliance and Fuel Economy Data for Vehicles and Engines," accessed November 12, 2019, [Online]: <https://www.epa.gov/compliance-and-fuel-economy-data/data-cars-used-testing-fuel-economy>.
- [22] R. Ma, X. He, Y.-L. Zheng, B. Zhou, S.-N. Lu, and Y. Wu, "Real-world driving cycles and energy consumption informed by large-sized vehicle trajectory data," *Journal of Cleaner Production*, vol. 223, pp. 564–574, Jun. 2019, doi: <https://doi.org/10.1016/j.jclepro.2019.03.002>.
- [23] Zhao, W.; Wu, G.; Wang, C.; Yu, L.; Li, Y. Energy transfer and utilization efficiency of regenerative braking with hybrid energy storage system. *J. Power Sources* 2019, 427, 174–183.
- [24] J. Valladolid, M. Calle, and A. Guiracocha, "Analysis of regenerative braking efficiency in an electric vehicle through experimental tests," *Ingenius*, no. 29, pp. 24–31, Jan. 2023, doi: <https://doi.org/10.17163/ings.n29.2023.02>.
- [25] Kim, D.; Eo, J.S.; Kim, Y.; Guanetti, J.; Miller, R.; Borrelli, F. Energy-Optimal Deceleration Planning System for Regenerative Braking of Electrified Vehicles with Connectivity and Automation; Technical Report; SAE:Warrendale, PA, USA, 2020.
- [26] Abdelkareem, M., Lin, X., Ahmed, A., Ahmed, E., Jia, M. et al., "Vibration energy harvesting in automotive suspension system: A detailed review," *Applied Energy*, vol. 229, pp. 672–699, Nov. 2018, doi: <https://doi.org/10.1016/j.apenergy.2018.08.030>.
- [27] R. Zhang, X. Wang, and S. John, "A Comprehensive Review of the Techniques on Regenerative Shock Absorber Systems," *Energies*, vol. 11, no. 5, p. 1167, May 2018, doi: <https://doi.org/10.3390/en11051167>

*Надійшла до редакції 10.04.2024*



Г. Ю. Терещенко<sup>1</sup>, І. В. Кириченко<sup>2</sup>, К. С. Смеляков<sup>3</sup>, А. Є. Олійник<sup>4</sup>

<sup>1</sup>ХНУРЕ, м. Харків, Україна, hlib.tereshchenko@nure.ua, ORCID iD: 0000-0001-8731-2135

<sup>2</sup>ХНУРЕ, м. Харків, Україна, iryna.kyrychenko@nure.ua, ORCID iD: 0000-0002-7686-6439

<sup>3</sup>ХНУРЕ, м. Харків, Україна, kyrylo.smelyakov@nure.ua, ORCID iD: 0000-0001-9938-5489

<sup>4</sup>ХНУРЕ, м. Харків, Україна, artem.oliinyk@nure.ua, ORCID iD: 0009-0008-8613-964X

## АНАЛІЗ МЕТОДІВ СТИСНЕННЯ ЗОБРАЖЕНЬ ДЛЯ ЗБЕРІГАННЯ У ДЕЦЕНТРАЛІЗОВАНИХ БЛОКЧЕЙН СХОВИЩАХ

Розглянуто методи стиснення зображень для зберігання у децентралізованих блокчейн-сховищах, зокрема в системі IPFS. Пропонується методологія оцінки ефективності різних алгоритмів стиснення зображень, що включає як стиснення без втрат (PNG, TIFF, GIF), так і стиснення з втратами (JPEG, WEBP). Виконується порівняння показників розміру файлів, якості зображень (за допомогою метрики PSNR), часу завантаження в IPFS та пропускної здатності для кожного методу стиснення. Для кожного зображення вимірюється хеш, який зберігається у смарт-контракті в локальній блокчейн-мережі, розгорнутій за допомогою Ganache. Використання таких методів дозволяє оптимізувати зберігання та передачу зображень у блокчейн-сховища, забезпечуючи баланс між якістю зображень, розміром файлів та швидкістю доступу.

БЛОКЧЕЙН, СТИСНЕННЯ ЗОБРАЖЕНЬ, СХОВИЩА ДАНИХ, IPFS, GANACHE, ГРАФІКА.

**G.Yu. Tereshchenko, I.V. Kyrychenko, K.S. Smelyakov, A.E. Oliynyk. Analysis of image compression methods for storage in decentralized blockchain repositories.** The methods of image compression for storage in decentralized blockchain repositories, specifically in the IPFS system, are examined. A methodology for evaluating the efficiency of different image compression algorithms is proposed, incorporating both lossless compression (PNG, TIFF, GIF) and lossy compression (JPEG, WEBP). The comparison includes metrics such as file size, image quality (measured by PSNR), upload time to IPFS, and bandwidth for each compression method. Each image is hashed, and the hash is stored in a smart contract within a local blockchain network deployed using Ganache. Such methods enable the optimization of image storage and transfer within blockchain repositories, ensuring a balance between image quality, file size, and access speed.

BLOCKCHAIN, IMAGE COMPRESSION, DATA STORAGE, IPFS, GANACHE, GRAPHICS.

### Вступ

Сучасний науковий прогрес активно просуває технології, випереджаючи їхній розвиток і впровадження нововведень. Раніше для зберігання великого обсягу зображень доводилося витратити значний час і ресурси. Тепер майже кожна компанія спрямована на поліпшення процесу зберігання шляхом інтеграції систем стиснення зображень у свої продукти.

Системи стиснення здатні працювати з різними типами даних, такими як зображення, звук, відео, текст тощо. Вони відрізняються одна від одної своєю архітектурою, типом пам'яті, тощо. Актуальність даної теми полягає в тому, що за останні роки в галузі зберігання даних зростає потреба у надійних і безпечних методах забезпечення конфіденційності та цілісності інформації. В контексті розвитку технологій зображень, виникає велика потреба у збереженні та обміні цими даними в безпечних та надійних умовах. Таким чином, вивчення методів стиснення зображень для їх зберігання у блокчейн-сховищі має велике практичне значення у сучасному інформаційному середовищі. У цьому дослідженні вирішено дослідити методи стиснення зображень для їх ефективного зберігання у блокчейн-сховищі оскільки технологія блокчейну [1] пропонує ряд переваг для зберігання

даних, таких як підвищена безпека, надійність та масштабованість.

На сьогоднішній день існує багато доступних систем стиснення зображень, які відрізняються за якістю в залежності від алгоритму. Якість стиснення прямо пропорційна якості алгоритму, який, у свою чергу, залежить від розміру, типу та обсягу даних, а також від типу системи, в якій він застосовується.

Метою роботи є визначення найбільш ефективного методу стиснення зображень для зберігання в блокчейн сховищах з точки зору розміру файлу, якості зображення та продуктивності передачі даних.

Об'єктом дослідження є ефективність методів стиснення зображень, які використовуються для зберігання у блокчейн сховищах. Предметом дослідження є процес стиснення зображень та його вплив на розмір файлів та якість зображення.

Методами дослідження є проведення експериментів зі стисненням зображень різними методами та аналіз результатів за допомогою спеціалізованих програмних інструментів.

Отримані результати можуть бути використані для покращення процесу зберігання зображень у блокчейн сховищах, а також можуть послужити основою для подальших досліджень у галузі стиснення даних у децентралізованих системах.

### 1. Технології збереження зображень у блокчейн сховищах

Зображення, які використовуються в сучасних інформаційних системах, зазвичай мають великий обсяг даних і потребують ефективних методів зберігання та передачі. У зв'язку з цим, блокчейн сховища активно досліджуються для зберігання зображень, оскільки вони пропонують децентралізовану, безпечну та прозору інфраструктуру для зберігання даних [2]. В даному дослідженні розглядається найпопулярніше блокчейн сховище — IPFS.

IPFS активно використовується як протокол комунікації та розподілу даних, побудований на основі блокчейн технології. Це блокчейн сховище, що пропонує децентралізовану архітектуру для зберігання та обміну даними [3].

При виборі IPFS для зберігання зображень у блокчейн сховищах враховуються кілька важливих чинників. IPFS вирізняється своєю децентралізованістю та масштабованістю, що робить його привабливим варіантом для зберігання даних. Ця характеристика широко поєднується з основними принципами блокчейн технології, забезпечуючи надійність та безпеку даних.

Архітектура IPFS базується на концепції розподіленої мережі з додатковим шаром блокчейн технології. Кожен вузол мережі може бути як клієнтом, що зберігає дані, так і постачальником контенту [4]. Коли клієнт запитує доступ до певного файлу за його хешем, мережа автоматично знаходить найближчий вузол, який має цей файл, та передає його клієнту через блокчейн (див. рис. 1).

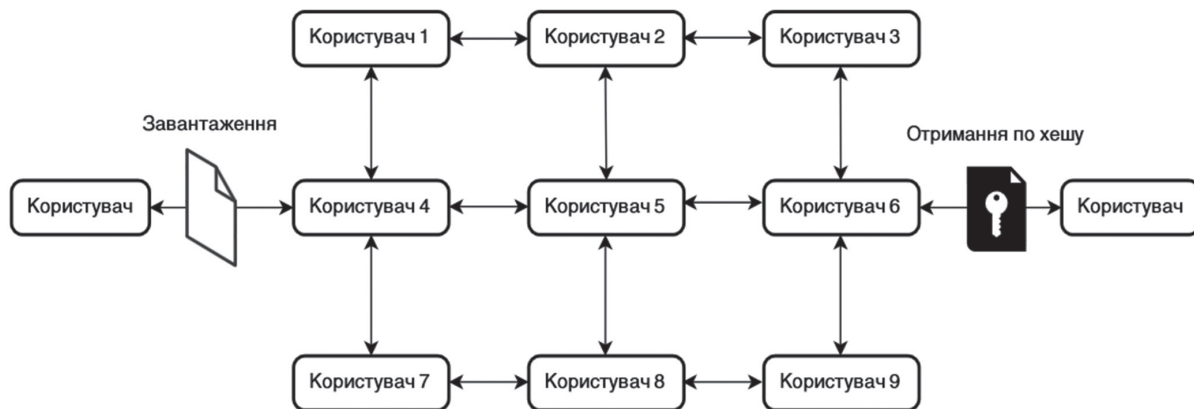


Рис. 1. Архітектура сучасних блокчейн сховищ

Додатково, в сучасному світі IPFS активно використовують в інтеграції з блокчейн технологією, що дозволяє забезпечити прозорість та цілісність збережених даних. Оскільки кожен файл отримує

унікальний хеш [5], він може бути легко записаний в блокчейн, забезпечуючи незмінність даних та перевірку їхньої цілісності (див. рис. 2).



Рис. 2. Взаємодія IPFS з блокчейн мережею

Враховуючи постійний розвиток IPFS та активні дослідження в галузі блокчейн технологій, обрання цього блокчейн сховища для зберігання зображень відкриває широкі перспективи для подальшого розвитку та вдосконалення цієї інфраструктури.

Незважаючи на переваги, використання IPFS як блокчейн сховища також має свої виклики, такі як обмежена пропускна здатність мережі. Це обов'язково враховується при виборі методів стиснення (див. табл.1).



Таблиця 1

Актуальні переваги та недоліки зберігання зображень в децентралізованих сховищах

Характеристика	Плюси	Мінуси
Децентралізованість	Забезпечує високий рівень безпеки та надійності даних, оскільки дані розподіляються по різних вузлах мережі	Обмежена пропускна здатність, особливо в глобальній мережі
Масштабованість	Здатний ефективно обробляти великі обсяги даних та взаємодіяти з великою кількістю користувачів	Проблеми з пропускною здатністю при великих обсягах даних
Унікальність хешу	Кожен файл отримує унікальний хеш, що гарантує цілісність та безпеку даних	Можливість втрати доступу до даних у разі втрати хешу
Прозорість даних	Забезпечує прозорість та відкритий доступ до даних	Проблеми з конфіденційністю даних та приватністю
Вартість зберігання	Вартість зберігання даних зазвичай нижча порівняно з централізованими рішеннями	Витрати на мережеву активність та маніпуляції даними

Отже, вибір відповідних методів стиснення є ключовим для забезпечення ефективного, економічного та безпечного зберігання зображень у блокчейн сховищах.

**2. Технології стиснення зображень**

Технології стиснення зображень можна розділити на два стиснення з втратами та стиснення без втрат.

Стиснення з втратати - метод, який дозволяє значно зменшити розмір файлу за рахунок втрати якості зображення. Цей метод дозволяє досягти значного зменшення розміру файлу без помітних втрат якості;

Стиснення без втрат - метод, який дозволяє зберегти початкову якість зображення. Однак, цей метод, як правило, не дозволяє досягти такого значного зменшення розміру файлу, як стиснення з втратами.

До плюсів втратного стиснення можна віднести високу ефективність. Втратне стиснення може забезпечити значно більший ступінь стиснення, ніж безвтратне. Це може призвести до значної економії місця на диску, підвищення пропускної здатності мережі та покращення ефективності обробки зображень.

Тим часом до мінусів втратного стиснення зазвичай відносять втрату якості. Втратне стиснення передбачає видалення деякої інформації з зображення. Це може призвести до втрати деталей, зміни кольору

або контрасту. Також серед мінусів розглядають неможливість відновлення даних. Втративши інформацію, яку було видалено при стисканні, її неможливо відновити.

Втратне стиснення широко застосовується для обробки зображень, де незначна втрата деталізації є допустимою і не впливає суттєво на їх сприйняття. Використання методів втратного стиснення, дозволяє значно зменшити розмір файлу за рахунок часткового зменшення якості зображення, що є прийнятним компромісом у багатьох випадках. Ці методи базуються на алгоритмах, що відкидають або узагальнюють деякі деталі зображення, які менш помітні для людського ока, зберігаючи при цьому загальну візуальну якість.

Тим часом до плюсів безвтратного стиснення відносять збереження якості. Безвтратне стиснення не призводить до втрати інформації з зображення. Це означає, що стиснене зображення буде повністю відповідати оригінальному.

Але основний мінус безвтратного стиснення це низька ефективність. Безвтратне стиснення зазвичай забезпечує нижчий ступінь стиснення, ніж втратне. Це означає, що для зберігання стиснених зображень може знадобитися більше місця на диску або пропускної здатності мережі.

Безвтратне стиснення використовується в різних областях, де збереження якості зображень є вирішальним аспектом. Наприклад, у медичній сфері або в обробці зображень для друку, важливо зберегти всі важливі деталі та точність, тому безвтратні методи стиснення є невід'ємною частиною процесу стиснення зображень. Використання безвтратного стиснення дозволяє зменшити обсяг даних, зберігаючи при цьому високий рівень якості, що робить його незамінним інструментом у сучасній обробці зображень.

Наразі вибір методу стиснення залежить від конкретного застосування. Для зображень з високою деталізацією, для яких втрата якості не є допустимою, слід використовувати безвтратне стиснення. Для зображень, для яких важлива пропускна здатність або розмір файлу, можна використовувати втратне стиснення

Після того, як ми зрозуміли актуальність дослідження методів стиснення зображень у контексті їх зберігання в блокчейн-сховищах, наступним кроком є аналіз доступних алгоритмів стиснення. Розглянемо основні методи, які широко використовуються для зменшення розміру зображень без втрати якості та з втратами, щоб визначити їх ефективність у контексті наших досліджень.

У дослідженні розглядаються алгоритми втратного стиснення, такі як JPG, WEBP та алгоритми стиснення без втрат, такі як PNG, TIFF, GIF.

### 3. Метод стиснення PNG

Метод стиснення PNG є форматом зображень, який активно використовується для зберігання растрових зображень [6]. Однією з ключових особливостей PNG є можливість стиснення без втрат якості. Для цього реалізований алгоритм стиснення Deflate.

Deflate – це алгоритм стиснення з втратами, що використовується в форматі PNG для зменшення розміру файлу без втрати якості. Він базується на комбінації алгоритму LZ77 та кодування Хаффмана [7].

Алгоритм LZ77 використовується для заміни повторюваних послідовностей байтів на зворотні посилання на знайдені раніше послідовності в тексті. Цей процес ґрунтується на вікні ковзання, яке зберігає копії останніх даних. Замість зберігання повної копії, використовується посилання на позицію та довжину послідовності.

Після застосування LZ77 до даних, отримані символні послідовності кодуються за допомогою кодування Хаффмана, яке стискає часто повторювані символи у менші біти та рідко повторювані символи у більші.

Початковий етап полягає у визначенні ймовірностей появи кожного символу у зображенні. У контексті стиснення зображень ці символи можуть бути представлені, наприклад, значеннями пікселів або іншими кодами, що представляють кольори.

Для обчислення ймовірностей використовується формула 1:

$$p_i = \frac{n_i}{N} \quad (1)$$

де  $p_i$  – ймовірність появи символу  $i$ ,  $n_i$  – кількість входжень символу  $i$  у зображенні,  $N$  – загальна кількість символів у зображенні.

Після обчислення ймовірностей будується дерево Хаффмана. Для кожного символу будується листок дерева з вагою, що відповідає ймовірності появи символу. Внутрішні вузли мають вагу, яка є сумою ваг дітей. Після побудови дерева вузлам надаються ваги, відповідні їх ймовірностям.

Для призначення кодів Хаффмана кожному символу призначається унікальний код, що визначається шляхом від кореня дерева до відповідного листка. Код складається з послідовності бітів, де 0 вказує на ліве відгалуження, а 1 – на праве.

Після цього зображення кодується, замінюючи кожен символ його відповідним кодом. Це призводить до створення бітової послідовності, яка представляє стиснене зображення.

Наприклад для побудови дерева Хаффмана для даних «AAAABBBBBBCCCCDDDD» після аналізу частоти, буде отримано наступні значення:

- A: частота 4 рази, ймовірність появи 4/16;

- B: частота 5 разів, ймовірність появи 5/16;
- C: частота 4 рази, ймовірність появи 4/16;
- D: частота 3 рази, ймовірність появи 3/16.

Після цього визначається вага кожного листка і будується дерево Хаффмана (див. рис. 3).

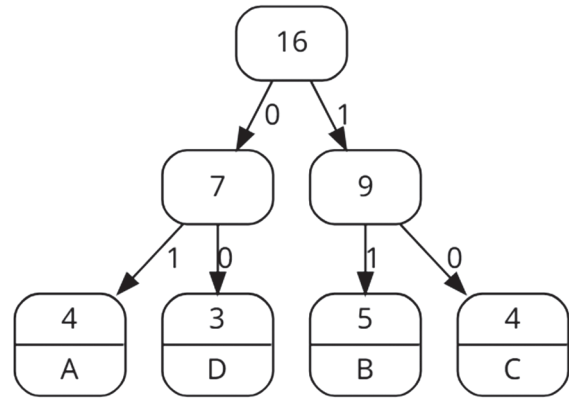


Рис. 3. Приклад діаграми Хаффмана

Таким чином, метод стиснення PNG забезпечує зменшення розміру файлу без втрати якості. Це дозволяє ефективно зберігати та передавати зображення, зберігаючи всі деталі та точність. В сучасному світі використання цього методу є критичним для сфер медичної діагностики, наукових досліджень, поліграфії та інших сфер де важлива висока якість зображення.

### 4. Метод стиснення TIFF

Метод стиснення TIFF є форматом зображень, який використовується для зберігання растрових зображень. Однією з ключових особливостей TIFF є можливість стиснення без втрат якості. Для цього використовується алгоритм стиснення LZW.

Алгоритм LZW – це алгоритм стиснення без втрат, що використовується в форматі TIFF для зменшення розміру файлу без втрати якості. Він базується на використанні словника фіксованого розміру, який динамічно оновлюється під час процесу стиснення [8].

Першим етапом є створення словника, який містить всі можливі символи вхідних даних. Під час процесу стиснення, алгоритм шукає найдовшу послідовність символів, яка вже є в словнику, і замінює її на відповідний код словника. Ця послідовність символів потім розширюється на один символ, і нова послідовність додається до словника.

Наприклад, розглянемо стиснення рядка «АВАВАВА» за допомогою LZW. Спочатку словник містить лише окремі символи «А» з кодом «1» та «В» з кодом «2».

Під час стиснення, алгоритм спочатку знаходить «А» в словнику і виводить відповідний код. Потім він розширює послідовність до «АВ», якої ще немає в словнику, тому «АВ» додається до словника з кодом «3», а код для «А» виводиться.

Процес продовжується, поки весь рядок не буде стиснутий, і в результаті отримується послідовність кодів словника. Таким чином, вихідний код для рядка «АВАВАВА» за допомогою алгоритму LZW буде 1, 2, 3, 1, 3 (див. табл. 1.2).

**Таблиця 2**  
**Покроковий принцип роботи алгоритму LZW**

Крок	Словник (символ : код)	Вивід алгоритму
1	A : 1, B : 2, AB : 3	1
2	A : 1, B : 2, AB : 3, BA : 4	1, 2
3	A : 1, B : 2, AB : 3, BA : 4, ABA : 5	1, 2, 3
4	A : 1, B : 2, AB : 3, BA : 4, ABA : 5	1, 2, 3, 1
5	A : 1, B : 2, AB : 3, BA : 4, ABA : 5	1, 2, 3, 1, 3

Цей алгоритм широко використовується для стиснення даних, які містять повторювані послідовності символів, і дозволяє стискати зображення без втрат. Важливо зазначити, що, хоча LZW є алгоритмом стиснення без втрат, він може бути неефективним для стиснення деяких типів даних. Наприклад, він може не зменшувати розмір файлу, якщо дані містять високий рівень шуму або якщо повторювані послідовності символів є дуже короткими або рідкісними.

**5. Метод стиснення GIF**

Метод стиснення GIF є форматом зображень, який використовується для зберігання растрових зображень. Однією з ключових особливостей GIF є можливість стиснення без втрат якості. Для цього використовується алгоритм стиснення LZW, аналогічний до того, що використовується в TIFF. Однак, варто зазначити, що GIF відрізняється від TIFF декількома нюансами.

По-перше, GIF підтримує анімацію, що дозволяє зберігати кілька зображень в одному файлі.

По-друге, GIF використовує палітру кольорів, що обмежує кількість кольорів, які можуть бути використані в зображенні, до 256 [9]. Саме через цю особливість, за останні роки використання цього формату стали менш привабливими, але незважаючи на це, GIF досі продовжує використовуватись для зображень з обмеженою палітрою кольорів, таких як текст, логотипи або прості ілюстрації, і не дивлячись на його недоліки, є досить ефективним.

Алгоритм LZW в GIF працює так само, як і в TIFF. Він створює словник, який динамічно оновлюється під час процесу стиснення, замінюючи найдовші послідовності символів, які вже є в словнику, на відповідні коди словника. Це дозволяє GIF стискати зображення без втрати якості. Але також варто зазначити, що, цей метод може бути неефективним, якщо повторювані послідовності символів є дуже короткими або дані містять високий рівень шуму.

**6. Метод стиснення JPEG**

Метод стиснення JPEG є форматом зображень, який використовується для зберігання растрових зображень. Однією з ключових особливостей JPEG є можливість стиснення з втратами якості. Для цього використовується алгоритм стиснення, що базується на дискретному косинусному перетворенні (DCT) [10]. DCT є методом стиснення, що використовується в форматі JPEG для зменшення розміру файлу за рахунок втрати якості. Він базується на перетворенні просторових даних зображення в частотні дані, що дозволяє відкинути частотні компоненти, які менш важливі для сприйняття людським оком. Процес стиснення JPEG можна розбити на п'ять основних кроків.

Першим кроком зображення розбивається на блоки розміром 8x8 пікселів, цей процес можна описати формулою 2:

$$N = \frac{W \cdot H}{64} \tag{2}$$

де  $N$  – загальна кількість блоків,  $W$  – ширина зображення,  $H$  – висота зображення.

Далі кольоровий простір зображення перетворюється з RGB в YCbCr, що відокремлює яскравість (Y) від хроматичності (Cb та Cr) [11].

Третім кроком кожен блок 8x8 пікселів перетворює просторові дані координати  $x, y$  в частотні дані (компоненти низьких і високих частот) за допомогою DCT. Матриця результатів DCT розраховується за формулою 3:

$$F = T \cdot B \cdot T^T \tag{3}$$

де  $F$  – матриця результатів DCT,  $B$  – матриця 8x8 пікселів зображення,  $T$  – матриця косинусних перетворень розміром 8x8,  $T^T$  – транспонована матриця перетворень.

Матриця косинусних перетворень розраховується за формулою 4:

$$T_{[i,j]} = \alpha(i) \cdot \cos\left(\frac{(2j+1) \cdot i \cdot \pi}{16}\right) \tag{4}$$

де  $i$  – рядок матриці,  $j$  – стовпець матриці,  $\alpha(i)$  – визначається за формулою 5:

$$\alpha(i) = \begin{cases} \frac{1}{\sqrt{8}}, & \text{якщо } i = 0 \\ \frac{1}{2}, & \text{якщо } i > 0 \end{cases} \tag{5}$$

Четвертим кроком частотні координати кожного блоку нормалізуються за допомогою таблиці квантування, що зменшує точність височастотних компонентів. На цьому етапі відбувається втрата інформації. Коефіцієнти ДКП квантуються, тобто округляються до заданих значень. Матриця квантування розраховується за формулою 6:

$$B' = \left\lfloor \frac{F}{Q} \right\rfloor \quad (6)$$

де  $B'$  – квантизовані коефіцієнти DCT,  $F$  – матриця результатів DCT,  $Q$  – матриця квантування,  $\lfloor \cdot \rfloor$  – операція взяття підлоги (округлення вниз),

Матриця  $Q$  може бути визначена за допомогою таблиць квантування, які містять значення квантування для різних частотних компонентів та рівнів стиснення [6].

Останнім кроком нормалізовані блоки кодується за допомогою кодування Хаффмана для подальшого зменшення розміру файлу і збираються разом із заголовком JPEG.

Важливо зазначити, що, хоча JPEG є алгоритмом стиснення з втратами, він дозволяє вибрати баланс між розміром файлу та якістю зображення за допомогою параметра якості. Високі значення якості приводять до меншого стиснення та вищої якості зображення, тоді як низькі значення якості приводять до більшого стиснення та нижчої якості зображення.

Оскільки JPEG є одним з найефективніших методів стиснення, переважно для фотографій, які зазвичай містять велику кількість деталей та кольорів, і при цьому зміна якості зображень майже не видна людському оку, він є найпопулярнішим методом стиснення зображень в усьому світі.

### 7. Метод стиснення WEBP

WEBP є форматом зображень, розробленим Google, який активно використовується у веб-застосунках. Однією з ключових особливостей цього формату є можливість стиснення з втратами якості. Для цього використовується алгоритм стиснення VP8.

VP8 є алгоритмом стиснення з втратами, що використовується в форматі WEBP, для зменшення розміру файлу. Він базується на використанні дискретного косинусного перетворення (DCT), аналогічного до того, що використовується в JPEG. DCT перетворює зображення з просторового (піксельного) представлення до частотного. Це дозволяє видалити високо-частотні компоненти, які людське око менш чутливе, зменшуючи розмір файлу без значного зниження якості зображення.

На відміну від JPEG, у форматі WEBP, який використовує алгоритм стиснення VP8, розбивка на блоки може бути більш гнучкою. VP8 може використовувати блоки різного розміру, включаючи 4×4, 8×8, 16×16 та інші. Це дозволяє алгоритму оптимізувати стиснення для різних частин зображення.

Також, варто зазначити, що VP8 відрізняється від JPEG ще декількома нюансами. По-перше, VP8 використовує більш ефективний метод кодування, який включає в себе прогнозування між кадрами та внутрішньо-кадрове прогнозування. По-друге, VP8

включає в себе вбудований механізм корекції помилок, що дозволяє відновити зображення навіть при втраті даних під час передачі. По-третє, VP8 підтримує альфа-канал, що дозволяє створювати прозорі зображення.

Важливо зазначити, що, хоча VP8 є алгоритмом стиснення з втратами, він може бути неефективним для стиснення деяких типів даних. Наприклад він може бути менш ефективним для стиснення зображень з високою деталлю, оскільки він може призвести до втрати дрібних деталей під час стиснення.

Незважаючи на недоліки та переваги кожного з методів стиснення, важливо враховувати багато факторів, таких як якість зображення, розмір файлу, підтримка прозорості та інші вимоги до стиснення.

### 8. Пропонований метод аналізу стиснення зображень для зберігання в децентралізованих блокчейн сховищах

Для вирішення проблеми ефективного зберігання зображень у блокчейн сховищах, пропонується метод аналізу, що ґрунтується на поєднанні технологій стиснення зображень та децентралізованої системи зберігання IPFS (InterPlanetary File System). Це було обрано через актуальність проблеми збереження великих обсягів даних у децентралізованих мережах. Цей метод дозволяє оцінити вплив різних методів стиснення на ефективність зберігання зображень, включаючи основні показники: розмір файлу, якість зображення, час завантаження та пропускну здатність.

З метою реалізації цього методу було розроблено який забезпечує користувачів можливістю імпортувати зображення та вибирати методи стиснення. Після імпорту зображення стискаються за допомогою різних алгоритмів, таких як PNG (deflate), TIFF (lzw), GIF (lzw-like) для стиснення без втрат, та JPEG (DCT), WEBP (VP8) для стиснення з втратами. Цей набір методів було обрано для забезпечення широкого спектру можливих сценаріїв використання.

Якість зображення, оцінюється за допомогою метрики PSNR (Peak Signal-to-Noise Ratio) і зберігається для подальшого аналізу. Використання метрики PSNR було обрано через її здатність об'єктивно оцінювати якість зображення після стиснення, що є критично важливим для цього дослідження.

Стиснені зображення завантажуються в IPFS через сервіс Pinata, який виступає посередником для доступу до всесвітньої мережі. Використовуючи технологію IPFS, кожне завантажене зображення отримує унікальний хеш, що забезпечує його однозначну ідентифікацію та можливість доступу до нього в будь-який момент часу. Це гарантує безпеку та незмінність даних, що є важливим аспектом у контексті децентралізованих мереж.

Після завантаження зображень в IPFS, отримані хеші зберігаються у смарт контракті, розгорнутому в локальній блокчейн мережі, що імітує Ethereum. Для цього використовуються технології Ganache для розгортання локальної мережі та Truffle для управління смарт контрактами, написаними на мові Solidity. Використання цих технологій забезпечує високу надійність та масштабованість системи.

Фронтенд частина веб-застосунку, реалізована за допомогою технології ASP.NET RazorPages, забезпечує інтерактивний і зручний у використанні інтерфейс. Користувачі можуть завантажити зображення зі свого пристрою, вибрати методи стиснення за допомогою чекбоксів, переглянути зібрану інформацію про стиснені зображення, порівняльні графіки, а також відкрити та переглянути всі стиснені зображення для візуального порівняння якості. Інтерфейс користувача розроблений з урахуванням принципів юзабіліті, що забезпечує інтуїтивно зрозуміле взаємодію та мінімізацію помилок з боку користувача. Це було обрано, оскільки зручність використання та інтуїтивність інтерфейсу є критичними для залучення широкої аудиторії користувачів.

Бекенд частина, реалізована на основі .NET 8 та C# 12.0, відповідає за обробку даних, взаємодію з IPFS та блокчейном, а також за виконання основних бізнес-логік застосунку. Використовуючи бібліотеку ImageMagick, бекенд частина здійснює стиснення

зображень за обраними методами, обчислюючи розмір файлу та метрику якості PSNR. Це рішення було обрано через високу продуктивність та гнучкість бібліотеки ImageMagick, яка дозволяє ефективно працювати з різними форматами зображень.

За допомогою HttpClient відбувається завантаження стиснених зображень до IPFS і отримання унікальних хешів для кожного зображення. Для збирання показників часу завантаження в IPFS та пропускної здатності використовуються інструменти вимірювання часу, такі як Stopwatch, для обчислення тривалості завантаження та розміру файлу після стиснення. Це забезпечує точний та надійний збір даних для подальшого аналізу.

Блокчейн частина системи відповідає за збереження хешів зображень у децентралізованій мережі, що гарантує їх безпеку, незмінність та прозорість. Локальна блокчейн мережа розгортається за допомогою Ganache, що дозволяє швидко створити середовище для тестування, а Truffle використовується для управління смарт контрактами, написаними на мові Solidity. Цей підхід був обраний через його надійність та широке визнання в спільноті розробників блокчейн-рішень.

Нижче наведено малюнок який візуалізує взаємодію вищезазначених модулів розробленої системи (див. рис. 4).

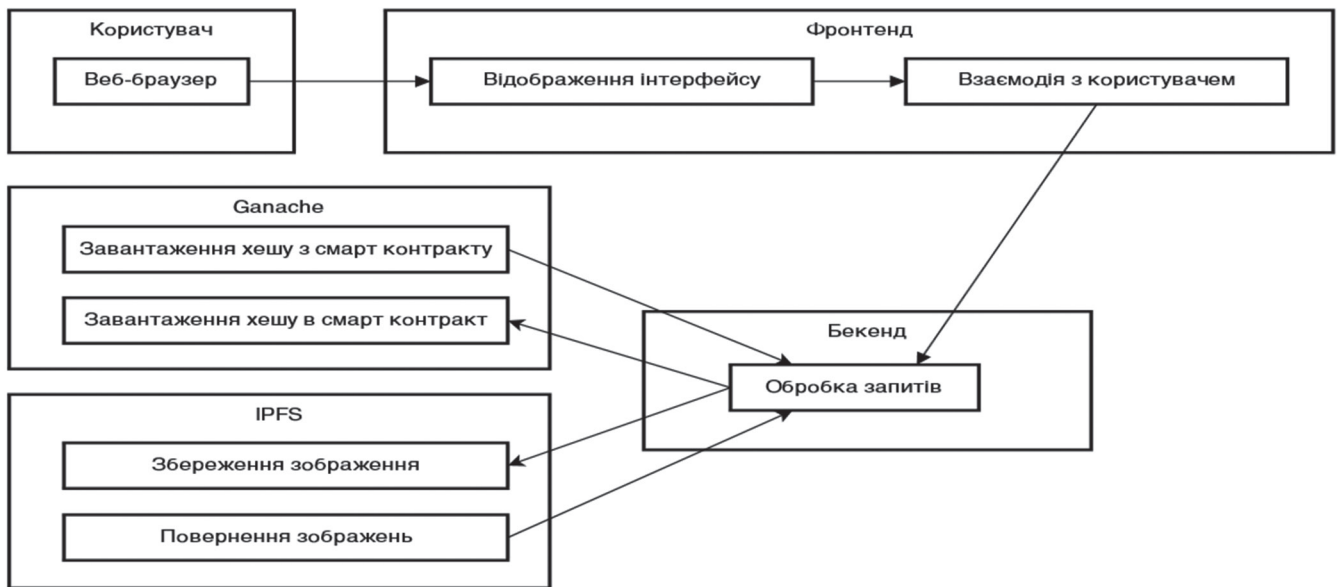


Рис. 4. Архітектура пропонованого застосунку для аналізу.

Таким чином, розроблений метод дозволяє користувачам проводити комплексний аналіз ефективності різних методів стиснення зображень у контексті зберігання в децентралізованих блокчейн сховищах. Це забезпечує можливість знайти оптимальний баланс між розміром файлу та якістю зображення, що

є критично важливим для багатьох сучасних застосунків. Завдяки використанню передових технологій та інтеграції з блокчейн, пропонований метод забезпечує високу надійність, прозорість та масштабованість, що робить його інноваційним рішенням у сфері управління цифровими даними.

## 9. Експеримент

Для дослідження ефективності різних методів стиснення зображень у контексті їх зберігання та передачі в IPFS було обрано різноманітний набір зображень. Зображення представляють типові сценарії використання у веб-застосунках і включають фотографії, медичні зображення, графіку та логотипи. Всі зображення були отримані у форматі BMP, що забезпечує високу якість без втрат і є оптимальним для подальшого порівняльного аналізу методів стиснення.

Для кожного зображення було застосовано п'ять популярних методів стиснення, такі як: JPEG, PNG, WEBP, GIF та TIFF.

JPEG використовується для фотографій завдяки своєму ефективному алгоритму стиснення з втратами, який добре працює з деталями та градієнтами кольору. PNG забезпечує стиснення без втрат, що робить його ідеальним для графіки та ілюстрацій. WEBP поєднує в собі переваги як стиснення з втратами, так і без втрат, пропонуючи високий ступінь компресії. GIF часто використовується для простих графічних зображень і анімацій, тоді як TIFF забезпечує високоякісне стиснення без втрат для збереження максимальної якості.

Набір досліджуваних даних включає в себе 20 фотографій, які відрізняються такими характеристиками як: наявність монохромності, кольорова гамма та температура, контрастність, насиченість та різна щільність пікселів.

Також набір даних містить 15 логотипів, які можна розділити на 3 групи по 5 логотипів: кольорові логотипи, чорно-білі логотипи та чорно-білі логотипи без заливки кольором.

Окрім цього датасет включає в себе 5 графічних зображень, серед яких є чорно-білі зображення. Графічні зображення відрізняються за розміром, кольором, насиченістю та іншими характеристиками.

Також для всебічного аналізу, в досліджувані дані було додано чорно-білі медичні зображення, такі як знімки МРТ, зображення клітин та інші зображення з медичної сфери.

Кожне зображення з набору даних було ретельно підібране, щоб включати широкий спектр характеристик. Першою з основних характеристик при відборі зображень була роздільна здатність, тобто кількість пікселів по горизонталі та вертикалі, яка впливає на розмір файлу та якість зображення. Також пріоритетною характеристикою була щільність пікселів, що визначається кількістю пікселів на дюйм і впливає на чіткість зображення. При відборі також була важлива різноманітність кольорової гамми, яка відображає діапазон кольорів, що наявні на зображенні і контрастність та насиченість.

Після того як був сформований набір даних, стало можливим провести аналіз досліджуваних зображень,

для того щоб визначити оптимальний метод стиснення для зберігання у блокчейн сховищі. Для використання показники що описані нижче.

Оскільки одною з основних цілей під час зберігання файлів є зниження витрат на зберігання, то було визначено один з головних показників, який прямо впливає на вартість зберігання зображень в IPFS – розмір завантаженого файлу. Відсоткове співвідношення розміру стисненого файлу до початкового файлу обчислювалось за формулою 7:

$$S = \frac{C}{O} \cdot 100 \quad (7)$$

де  $S$  – відсоткове співвідношення розмірів,  $C$  – розмір стиснутого файлу,  $O$  – розмір оригінального файлу.

Аналогічно до розміру файлу, на вартість зберігання також впливає витрачена пропускна здатність каналу (мережі) який був наданий кінцевому користувачу. Пропускна здатність, вимірювана у кілобайтах на секунду обчислювалась за формулою 8:

$$Tp = \frac{C}{t} \quad (8)$$

де  $Tp$  – пропускна здатність,  $C$  – розмір стиснутого файлу у кілобайтах,  $t$  – час, витрачений на завантаження в IPFS, вимірюваний в секундах.

Також, важливою вимогою для клієнта, під час процесу стиснення є максимально можливе збереження якості файлу після форматування. Якість вимірюється шляхом калькуляції піксельної різниці сигналу до шуму (PSNR) розраховувалась за формулою 9:

$$psnr = 10 \cdot \log_{10} \left( \frac{255^2}{mse} \right) \quad (9)$$

де  $psnr$  – пікове відношення сигналу до шуму,  $mse$  – значення середньоквадратичної похибки.

Ці показники були обрані для об'єктивної оцінки ефективності кожного методу стиснення з точки зору розміру файлу, швидкості завантаження, та якості зображення.

В першу чергу, було виміряно та проаналізовано розміри фотографій після стиснення за допомогою різних методів (див. рис. 5).

Як видно з діаграми, найбільший розмір файлу після стиснення спостерігався для методу TIFF (алгоритм LZW). Це можна пояснити тим, що методи стиснення без втрат зазвичай зберігають більше інформації, ніж методи з втратами. З іншого боку, метод WEBP (алгоритм VP8) показав найкращий результат з точки зору зменшення розміру файлу.

Другим важливим показником було вимірювання якості зображення після стиснення за допомогою метрики PSNR (Peak Signal-to-Noise Ratio) (див. рис. 6).

Як показує діаграма, методи стиснення без втрат, такі як PNG і TIFF, повністю зберігають якість зображення, що підтверджується максимальними

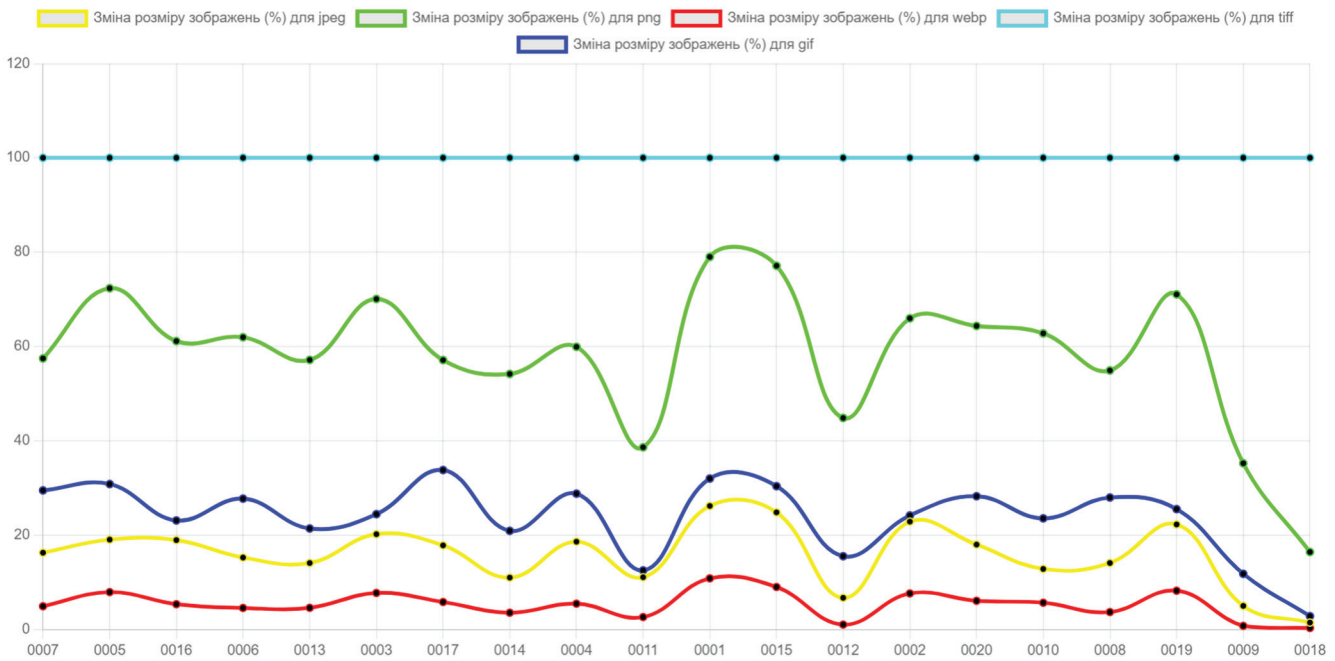


Рис. 5. Графік зміни розміру для фотографій

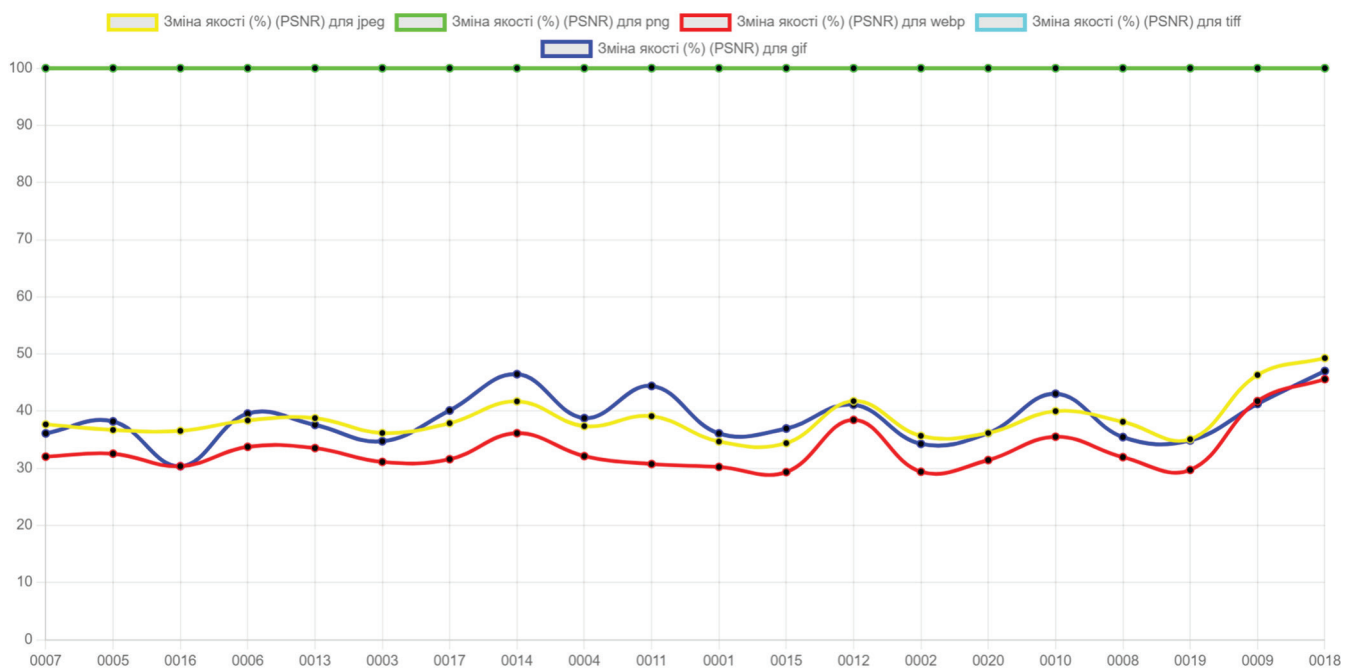


Рис. 6. Графік зміни якості для фотографій

значеннями PSNR. Методи стиснення з втратами, такі як JPEG, WEBP і GIF, демонструють нижчі значення PSNR, що є очікуваним результатом через втрату частини інформації при стисненні.

Наступним важливим показником була вимірювана пропускна здатність при завантаженні зображень в IPFS (див. рис. 7).

Як показує діаграма, найбільш ефективні з точки зору пропускної здатності, тобто витрачають її менше, є методи стиснення з втратами, такі як WEBP, JPEG та GIF. Це можна пояснити тим, що менші розміри файлів дозволяють швидше передавати дані через мережу.

Аналіз показав, що для фотографій, які зазвичай містять велику кількість деталей та кольорів, методи стиснення з втратами, такі як JPEG і WEBP, є більш ефективними для зменшення розміру файлу. На основі результатів ми робимо висновок, що метод JPEG забезпечує оптимальний баланс між зменшенням розміру файлу та якістю зображення для стиснення фотографій.

Аналогічно фотографіям, було виміряно та проаналізовано розміри медичних зображень після стиснення за допомогою різних методів (див. рис. 8).

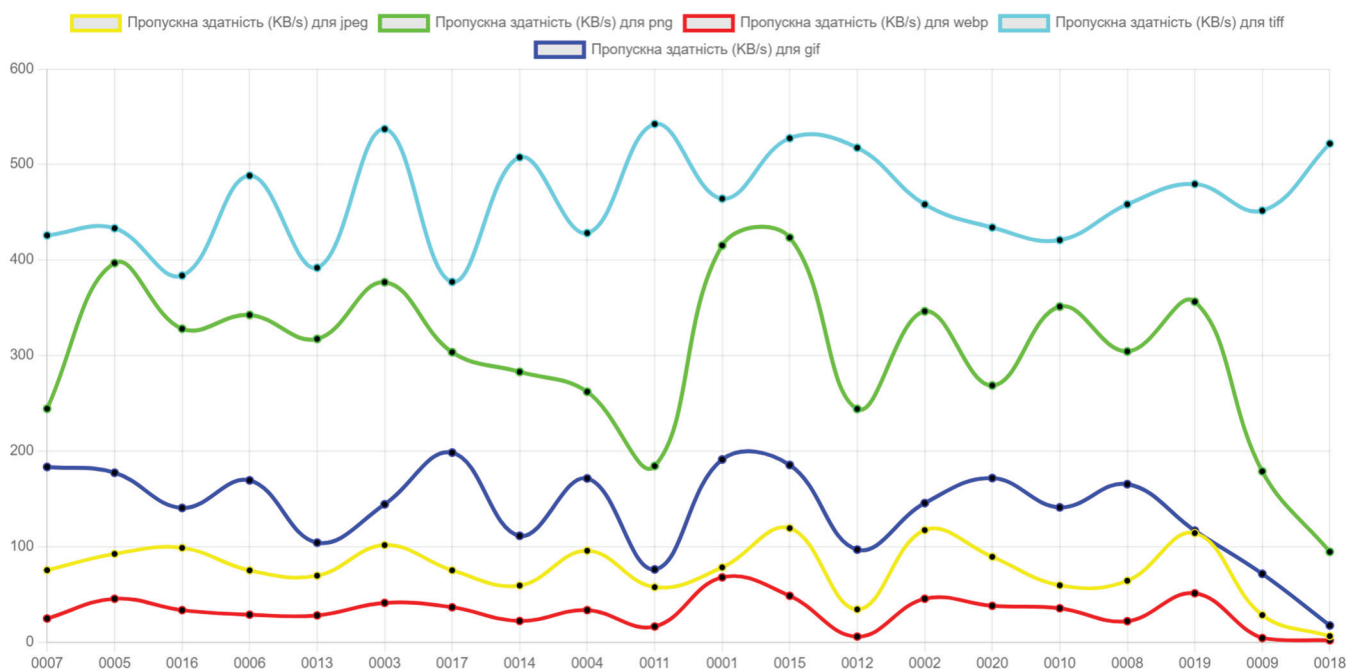


Рис. 7. Графік пропускної здатності для фотографій

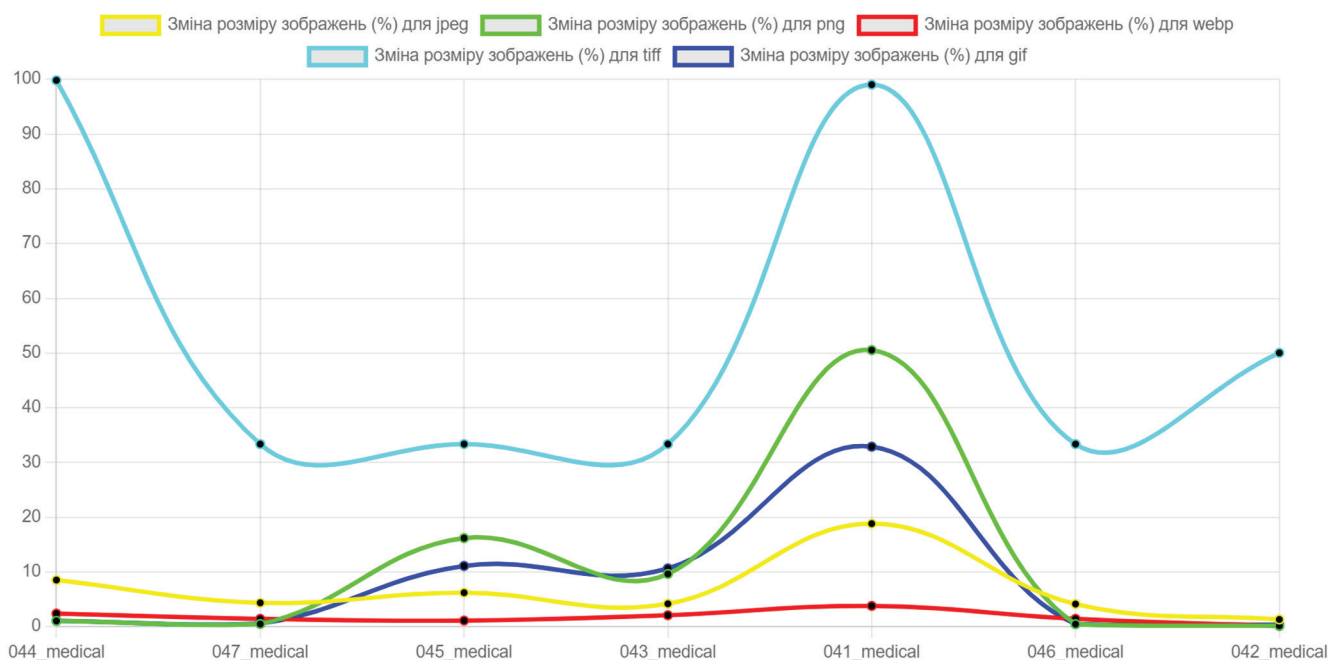


Рис. 8. Графік зміни розміру для медичних зображень

Як видно з діаграми, методи стиснення з втратами, такі як JPEG та WEBP, значно зменшили розмір файлу, пропонуючи ефективніші результати з точки зору обсягу даних. Також можна побачити що метод PNG, який стискає зображення без втрат, показав найкращі результати для більшості даних, в порівнянні з іншими методами стиснення, навіть з тими, що нехтують якістю під час форматування. Другим важливим показником було вимірювання якості медичних зображень після стиснення за допомогою метрики PSNR (див. рис. 9).

Як показує діаграма, методи стиснення без втрат, такі як PNG і TIFF, повністю зберігають якість зображення. Методи стиснення з втратами, такі як JPEG, WEBP та для більшості вибірки GIF, демонструють нижчі значення PSNR, що може бути критичним для медичних зображень, де важливе збереження максимальної кількості деталей.

Наступним важливим показником була виміряна пропускна здатність при завантаженні медичних зображень в IPFS (див. рис. 10).



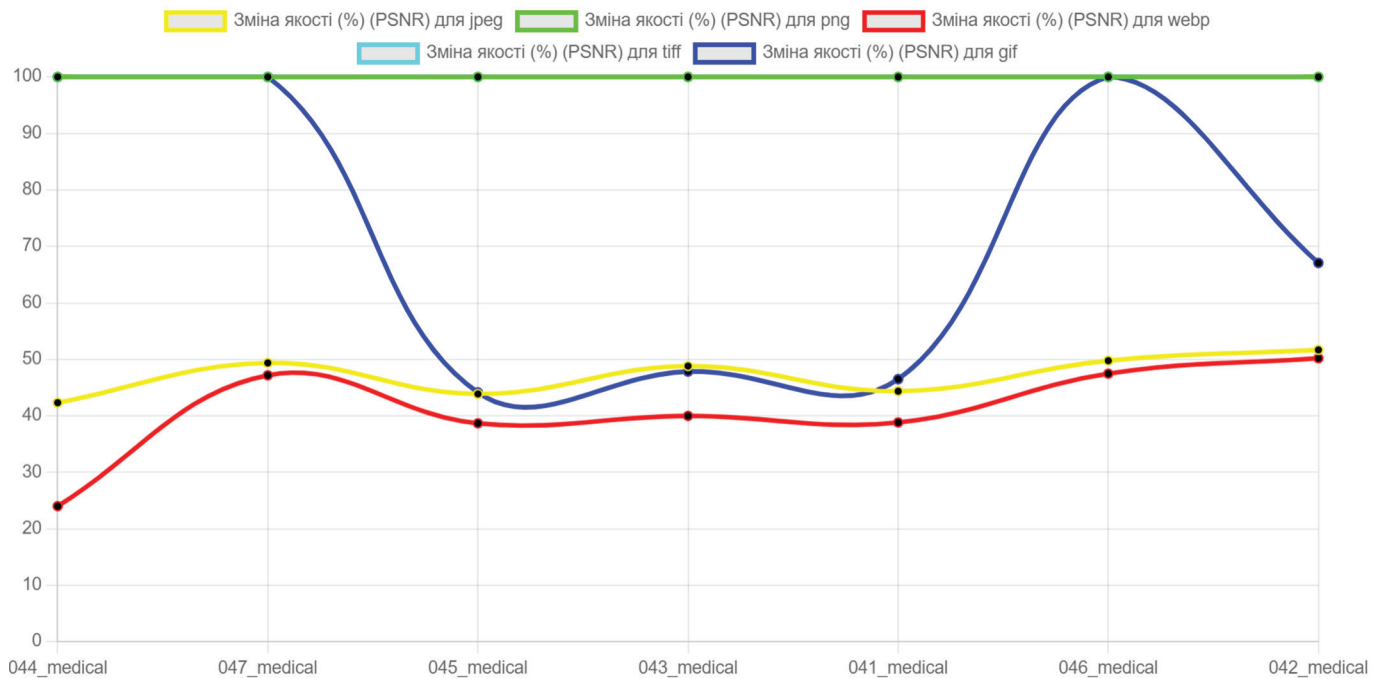


Рис. 9. Графік зміни якості для медичних зображень

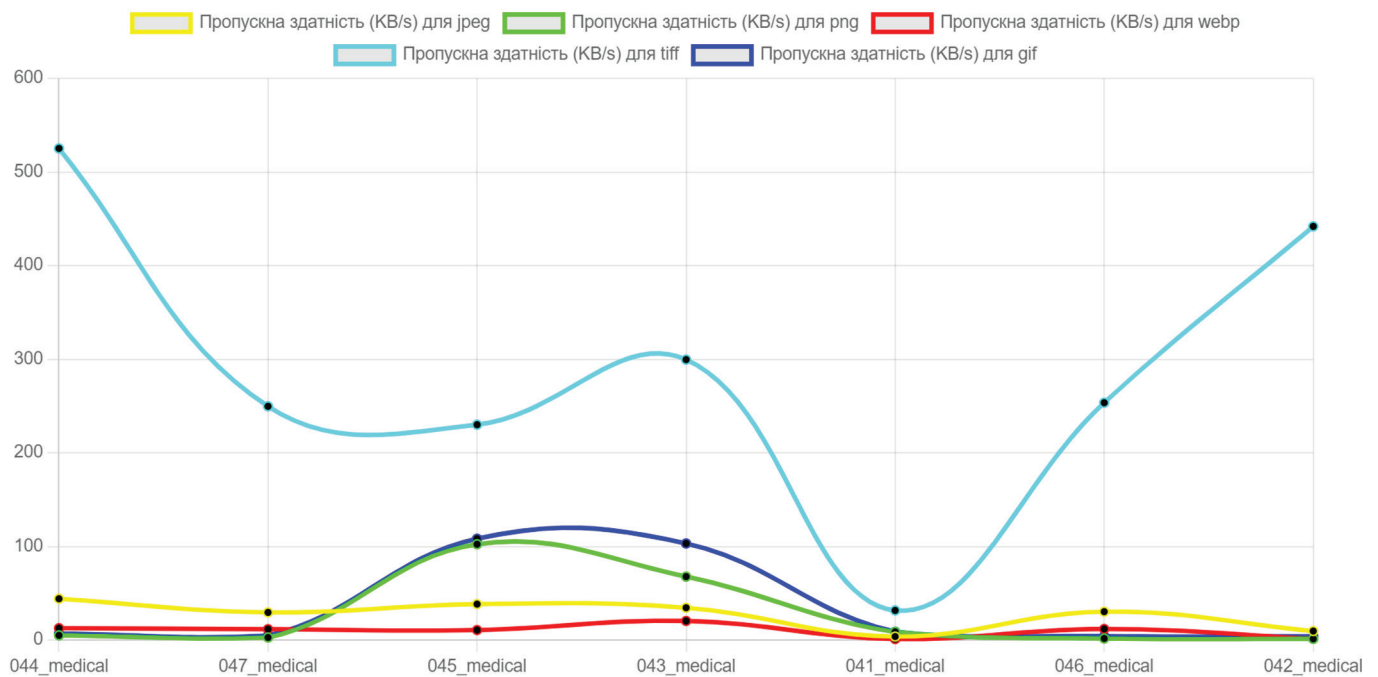


Рис. 10. Графік пропускної здатності для медичних зображень

Як показує діаграма, методи стиснення з втратами, такі як WEBP та JPEG, є найбільш ефективними з точки зору пропускної здатності, але в той же час, PNG показав хоч і не для всієї вибірки, але для більшості, найкращі показники.

Аналіз показав, що для медичних зображень критично важливим є збереження максимальної якості, тому метод PNG є більш придатним за інші, оскільки стискає зображення без втрати якості і показує найкращі результати для більшості досліджуваних файлів за такими характеристиками як розмір файлу

та пропускна здатність під час завантаження в блокчейн сховище. Проте, якщо необхідно надійно зменшити розмір файлу в декілька разів для усієї вибірки, враховуючи можливу втрату важливих деталей, може бути використаний метод стиснення з втратами WEBP, оскільки показав найкращі результати для цього показника.

Далі було виміряно та проаналізовано розміри графіки та логотипів після стиснення за допомогою різних методів (див. рис. 11).

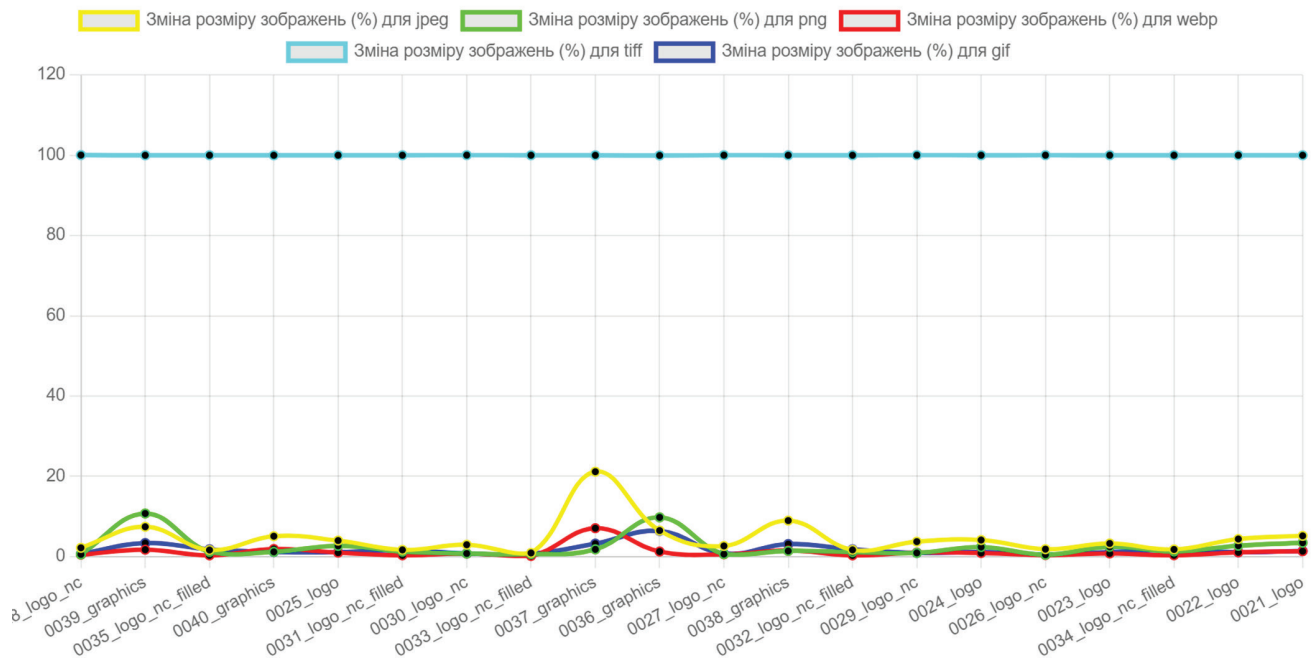


Рис. 11. Графік зміни розміру графічних зображень і логотипів

Як видно з діаграми, метод стиснення без втрат TIFF, має більші розміри файлів у порівнянні з методами з втратами, такими як JPEG, WEBP та GIF.

У той же час метод стиснення без втрат PNG показує кращі результати стиснення за JPEG та GIF (див. рис. 12.)

Зображення 1					
Type:	png, Quality: 100%,	Size: 7.51 KB	Time: 1.22s,	Throughput: 6.18 KB/s	<a href="#">Open Image</a>
Type:	webp, Quality: 42.74%,	Size: 8.17 KB	Time: 0.93s,	Throughput: 8.78 KB/s	<a href="#">Open Image</a>
Зображення 2					
Type:	png, Quality: 100%,	Size: 35.35 KB	Time: 1.27s,	Throughput: 27.76 KB/s	<a href="#">Open Image</a>
Type:	webp, Quality: 30.72%,	Size: 38.54 KB	Time: 1.23s,	Throughput: 31.32 KB/s	<a href="#">Open Image</a>
Зображення 3					
Type:	png, Quality: 100%,	Size: 180.47 KB	Time: 1.33s,	Throughput: 135.24 KB/s	<a href="#">Open Image</a>
Type:	webp, Quality: 25.3%,	Size: 706.89 KB	Time: 1.47s,	Throughput: 481.33 KB/s	<a href="#">Open Image</a>

Рис. 12. Порівняння розмірів деяких файлів для WEBP та PNG

Попередній аналіз показав, що пропускна здатність прямо залежить від розміру файлу, тому проводити додатковий аналіз діаграми пропускної здатності для графічних зображень та логотипів не є необхідним. Також оскільки для цих зображень є доречним максимальне збереження, тому метод PNG є оптимальнішим за інші, оскільки проводить безвтратне стискання і показує гарні результати стиснення файлу. Проте, як альтернативу для методу PNG при збереженні графічних файлів можна відмітити WEBP, оскільки цей метод стиснення показує чи найкращі показники співвідношення якості та розміру файлу серед методів стиснення з втратами. Проте в такому випадку буде мати місце значна втрата якості.

## Висновки

В ході проведення аналізу було досліджено методи стиснення зображень для зберігання у децентралізованих блокчейн сховищах на прикладі IPFS.

Для цього спочатку було розглянуто особливості зберігання даних у блокчейн сховищах, зокрема в IPFS і проаналізовано методи стиснення зображень, які можуть бути застосовані для зменшення розміру файлів перед зберіганням в IPFS.

Наступним кроком було визначено показники ефективності для кожного з методів стиснення (розмір файлу, якість зображення, пропускна здатність при завантаженні, час завантаження).

Після цього для проведення дослідження було розроблено бекенд на .NET 8 та фронтенд на Razor Pages для імпорту та стиснення зображень. Для імітації роботи реальної блокчейн мережі було розгорнуто локальну блокчейн мережу Ethereum за допомогою Ganache й написано смарт контракт на мові Solidity [12] який був пізніше задеплоєний в локальну блокчейн мережу за допомогою Truffle.

Після реалізації системи, яка дозволить провести аналіз, було проведено експеримент зі стискання зображень різними методами та збереженням їх в IPFS із фіксацією отриманих хешів у смарт-контрактах на локальній блокчейн мережі. За результатами експерименту зібрано та проаналізовано дані про розмір файлів, якість зображень (за метрикою PSNR), час завантаження в IPFS та пропускну здатність для кожного методу стиснення;

Для повноти картини було візуалізовано результати дослідження у вигляді графіків та таблиць.

За результатами дослідження, по-перше, було визначено, що методи стиснення без втрат, такі як PNG і TIFF, дозволяють зберегти високу якість зображення. Проте, це призводить до значного збільшення розміру файлу. Така особливість може бути критично важливою для типів зображень, де максимальна деталізація є необхідною, як-от медичні зображення.

По-друге, дослідження показало, що методи стиснення з втратами, такі як JPEG та WEBP, дозволяють значно зменшити розмір файлу. Це покращує ефективність зберігання та передачі даних у блокчейн сховищах. Зокрема, вони є оптимальними для використання з фотографіями та графікою, де допускається деяке зниження якості без значних втрат візуальної інформації.

Крім того, було встановлено, що метод WEBP демонструє найкращі результати з точки зору зменшення розміру файлу при збереженні достатньої якості зображення. Це робить його оптимальним для використання у випадках, коли необхідно знайти баланс між розміром файлу та якістю.

Дослідження також виявило, що пропускну здатність при завантаженні в IPFS є найвищою у методів стиснення з втратами, що зумовлено їх меншим розміром файлів. Це сприяє швидшій передачі даних у мережі, що є важливим показником ефективності.

На основі отриманих результатів можна зробити висновок, що для випадків, де критично важлива висока якість зображень, найкращим вибором є метод стиснення без втрат PNG. Для більш ефективного використання ресурсів та зменшення часу завантаження, методи стиснення з втратами, такі як JPEG та WEBP, є більш доцільними. Проведене дослідження

показало, що метод JPEG забезпечує оптимальний баланс між зменшенням розміру файлу та якістю зображення.

Для підтвердження отриманих результатів та зроблених висновків необхідні подальші дослідження. Вони мають охоплювати різноманітні типи зображень та включати більш складні сценарії використання, що дозволить отримати більш точні дані щодо ефективності кожного методу стиснення у різних умовах.

#### Список літератури:

- [1] Blockchain Technology Explained: Powering Bitcoin. URL: <https://www.ibm.com/topics/what-is-blockchain> (date of access: 29.03.2023)
- [2] Терещенко Г.Ю., Груздо І.В. Застосування симетричних алгоритмів в блокчейні. Біоніка інтелекту. – Харків : ХНУРЕ. – 2020. – № 1 (94). – С. 33-39
- [3] Huckle S. Internet of Things, Blockchain and Shared Economy Applications / S. Huckle, R. Bhattacharya, M. White, N. Beloff // Procedia Comput. Science. – Oct. 2016. – Vol. 98. – P. 461–466.
- [4] W. Pennebaker, J. Mitchell, JPEG: Still Image Data Compression Standard. – Springer New York, NY, 1992. – 638p. – ISBN: 978-0-442-01272-4.
- [5] 6. N. Bilous, G. Tereshchenko Copyright protection using blockchain. Біоніка інтелекту – Харків : ХНУРЕ. – 2019. – № 1 (92). – С. 52-58.
- [6] GIF File Format Summary. URL: <https://www.fileformat.info/format/gif/egff.htm> (date of access: 07.04.2023).
- [7] LZW compression. URL: <https://www.techtarget.com/whatis/definition/LZW-compression> (date of access: 07.04.2023).
- [8] Білоконенко В.М , Ревенчук І.А. Алгоритми сегментації зображень на базі побудови матриць збігів. Східно-Європейський журнал передових технологій.- №2(62), том 2.-2013.-С.43-45
- [9] Специфікація алгоритму стискання даних формату DEFLATE версія 1.3 (RFC1951). URL : <http://www.kytk.org.ua/post/cpetsyifikatsiya-alhorytmu-styskannya-danykh-formatu-deflate-versiya-1-3-rfc1951> (date of access: 07.04.2023).
- [10] J. Solem. Programming Computer Vision with Python: Tools and algorithms for analyzing images. O'Reilly Media, Inc., 2012. 264p
- [11] JPEG YCbCr Support. URL: <https://learn.microsoft.com/en-us/windows/win32/wic/jpeg-ycbcr-support> (date of access: 08.04.2023).
- [12] Шаронова Н.В., Терещенко Г.Ю. Проблеми і перспективи практичного застосування інформаційної технології blockchain в smart-контрактах. Інтелектуальні системи та інформаційні технології (ISIT-2019). – Матеріали Міжн. Наук.-практ. Конф. – Одеса, 19–24 серпня 2019 р. – С. 214–219.

*Надійшла до редколегії 7.02.2024*



Д. І. Гольдінер<sup>1</sup>, О. І. Матвієнко<sup>2</sup>

<sup>1</sup>Аспірант, Харківський національний університет радіоелектроніки, Харків, Україна;  
e-mail: denys.holdiner@nure.ua; ORCID ID: 0000-0002-1456-1867

<sup>2</sup>Канд. техн. наук, доцент, доцент кафедри прикладної математики,  
Харківський національний університет радіоелектроніки, Харків, Україна;  
e-mail: olha.matviienko@nure.ua; ORCID ID: 0000-0001-7492-7616

## ЗМЕНШЕННЯ ЙМОВІРНОСТІ ВІДМОВИ В СИСТЕМАХ МАСОВОГО ОБСЛУГОВУВАННЯ З ОБМЕЖЕНОЮ ЧЕРГОЮ ІЗ ЗАСТОСУВАННЯМ ПРІОРИТЕЗАЦІЇ ЗА РОЗМІРОМ ТА ШТУЧНОГО ІНТЕЛЕКТУ

Стаття присвячена збільшенню ефективності обробки заявок у багатоканальних системах масового обслуговування з обмеженою чергою та відмовами у випадку її переповнення. Предметом даної статті є: методи та підходи до оптимізації обробки потоків заявок. Метою роботи є: запропонувати новий підхід до пріоритизації та балансування категорій вимог задля зменшення ймовірності відмови. Завдання статті полягає у: формулюванні досліджуваної системи масового обслуговування; визначенні джерела оптимізації; описі методу розбиття загального потоку вимог на категорії; переліку підходів до оцінки розмірів заявки; визначенні алгоритму обробки заявок із застосуванням пріоритизації за часом обробки; пропозиції рішення проблем оцінки навантаження та балансування пріоритетів із застосуванням штучного інтелекту. Використовуються такі методи: теорія масового обслуговування, UML діаграми, штучний інтелект. Було здобуто наступні результати: запропоновано підхід до зменшення ймовірності відмови в багатоканальних системах масового обслуговування з обмеженою чергою, за рахунок збільшення пріоритетів менших заявок; запропоновано методи оцінки складності вимог та розбиття загального потоку заявок на категорії згідно з їхнім розміром; Запропоновано підхід до балансування пріоритетів за допомогою штучного інтелекту.

МАСОВЕ ОБСЛУГОВУВАННЯ, ПРІОРИТЕТИ, ШТУЧНИЙ ІНТЕЛЕКТ, НАВЧАННЯ, ШТУЧНА НЕЙРОННА МЕРЕЖА, ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ, ПРИЙНЯТТЯ РІШЕНЬ, МОДЕЛЬ ПОВЕДІНКИ, АДАПТАЦІЯ, БАЛАНСУВАННЯ НАБОРІВ ДАНИХ, ЗАДАЧА КАТЕГОРИЗАЦІЇ

**D. Goldiner, O. Matviienko. Rejection probability reduction in queueing systems with limited queue using size based prioritization and artificial intelligence.** The article is dedicated to enhancing the efficiency of request processing in multi-channel mass service systems with limited queue and refusals in case of overflow. The subject of this article is the methods and approaches to optimizing the processing of request flows. The aim of the work is to propose a new approach to prioritizing and balancing request categories to reduce the probability of refusals. The tasks of the article include: formulating the studied mass service system; determining the source of optimization; describing the method of dividing the overall request flow into categories; listing approaches to determining the request sizes; defining the algorithm for processing requests using size-based prioritization; proposing solutions to load assessment and priority balancing problems using artificial intelligence. The following methods are used: mass service theory, UML diagrams, artificial intelligence. The following results were obtained: a method to reduce the probability of refusals in multi-channel mass service systems with a limited queue by increasing the priority of smaller requests was proposed; methods for assessing request complexity and dividing the overall request flow into categories according to their complexity were proposed; an approach to balancing priorities using artificial intelligence was proposed.

MASS SERVICE SYSTEM, PRIORITIES, ARTIFICIAL INTELLIGENCE, TRAINING, ARTIFICIAL NEURAL NETWORK, APPLICATION OF NEURAL NETWORKS, DECISION MAKING, BEHAVIOR MODEL, ADAPTATION, DATA BALANCING, CATEGORIZATION TASK

### Вступ

У сучасному світі, навколо є безліч різноманітних процесів, що включають в себе обробку потоків вхідних даних. Значне поширення мають прикладні задачі, що можуть бути описані та оптимізовані за допомогою теорії масового обслуговування, наприклад:

- управління запасами;
- людськими ресурсами;
- оптимізація використання обладнання на виробництвах;
- керування взаємодією з клієнтами;
- обробка черг замовлень в торгівлі;
- побудова автоматизації виробництв;
- масштабування систем надання послуг;

– забезпечення збільшення пропускної здатності трафіка в цифрових мережах.

До математичних моделей можуть бути зведені задачі, що виникають в сферах: інформаційно-керуючих систем, системах контролю та управління процесами, телекомунікаційних технологіях та мережах, а також, у менеджменті. До кожної задачі можна підібрати відповідний тип системи масового обслуговування (СМО), який найкраще описуватиме специфіку фактичного процесу [1].

Зазвичай, прикладні задачі, що описуються за допомогою систем масового обслуговування, накладають обмеження, пов'язані з фактичними можливостями масштабування. Більшість задач, що мають

черги, не можуть дозволити собі зняти обмеження на кількість заявок, що перебувають в очікуванні. Також, зазвичай, не є можливим збільшувати чергу кожен раз, як вона переповнюється. Отже, набувають актуальності моделі з обмеженими чергами та відмовами у випадку переповнення. Такі СМО надають математичне обґрунтування та рекомендації до оптимізації процесів задля отримання кращої пропускної здатності. Тобто головна задача поставлена наступним чином: необхідно зменшити ймовірність відмови черговій заявці, що надходить до СМО, не змінюючи при цьому її структуру та основні характеристики.

Серед технологій, які яскраво про себе заявили в останнє десятиріччя, мають місце машинне навчання та штучний інтелект (ШІ).

Після спаду інтересу до штучного інтелекту в 70-х роках, у 80-х ця галузь отримала новий поштовх завдяки розробці експертних систем. Однак, технологія залишалась дуже вузькоспеціалізованою. Галузь машинного навчання значно розвинулася в 90-х роках, коли почали широко застосовуватись нейронні мережі, завдяки алгоритму зворотного поширення помилки. Що, у свою чергу, покращило ефективність моделей та їхню здатність до навчання. Головним стримуючим фактором залишалось обмеження обчислювальних потужностей і доступу до великих об'ємів даних [2]. У 2000-х роках ситуація почала значно покращуватись. На додачу до цього, глибокі нейронні мережі, що складаються з багатьох шарів, продемонстрували видатні результати у задачах розпізнавання образів, обробки мови та гри у складні ігри, як шахи або го. Починаючи з 2010 року ШІ стає все більш доступним як для професійного, так і побутового використання. У повсякденному житті дедалі частіше зустрічаються додатки та сервіси, які застосовують машинне навчання для збільшення ефективності або якості надання послуг. Останніми роками такі технології почали підтримуватись навіть у мобільних пристроях. Кожен наступний етап розвитку штучного інтелекту відкривав нові сфери його застосування, розширюючи можливості для покращення життя.

Це дослідження присвячено пошуку нових підходів до збільшення продуктивності багатоканальних СМО з обмеженою чергою та відмовами.

### 1. Аналіз поточного стану справ

Як правило, складні системи з високим навантаженням оптимізують за рахунок додавання пріоритетів. Розрізняють наступні види систем з пріоритетом [3, 4]:

– Абсолютний пріоритет – передбачає, що заявки з вищим пріоритетом завжди обслуговуються першими, з можливістю переривання обробки вимог з нижчим пріоритетом;

– Відносний пріоритет – заявки з вищим пріоритетом мають перевагу, але не можуть переривати виконання інших вимог;

– Пріоритет за класами обслуговування – передбачає класифікацію заявок на категорії, кожна з яких має свою чергу та пріоритет. Заявки одного типу обслуговуються за принципом FIFO (перший прийшов перший пішов), але між собою категорії взаємодіють згідно пріоритетів;

– Динамічний пріоритет – у цьому випадку пріоритет заявок може змінюватись з часом в залежності від певних умов та стану системи;

– Випадковий пріоритет – передбачає встановлення пріоритету заявці випадковим чином;

– Комбіновані методи – дозволяють довільну комбінацію раніше згаданих типів пріоритетів задля найкращого забезпечення потреб реальної задачі.

Такі підходи добре діють, але залишається проблема їхнього застосування для випадку, коли задача не регулює пріоритети на вимоги за замовченням. Для таких випадків необхідно напрацювати механізм переходу від звичайних СМО до таких, що використовують систему пріоритетів.

### 2. Постановка задачі

Будемо розглядати задачу про багатоканальну систему масового обслуговування із обмеженою чергою вимог та відмовою у випадку переповнення черги. У якій до  $n$  – однакових каналів обслуговування надходить сімейство з  $r$  пуассонівських потоків заявок інтенсивності  $\lambda_j$ , де  $j=1,2,\dots,r$ . Отже, сукупний потік є пуассонівським з інтенсивністю (1):

$$\lambda_T = \sum_{j=1}^r \lambda_j \quad (1)$$

Якщо на момент приходу нової заявки є хоча б один вільний канал, він негайно починає обробку. У випадку, коли усі канали зайняті, вимога стає останньою до загальної черги ємності  $k$ . Заявки покидають чергу для подальшого обслуговування у тій самій послідовності у якій вони надходили на очікування. Канал, що звільнюється від виконання, одразу починає обробку першої в черзі вимоги. При цьому кожна вимога обслуговується тільки одним каналом, і кожен канал може обслуговувати не більше однієї вимоги одночасно. У випадку, якщо вільними є декілька каналів обслуговування, для обробки буде обрано канал випадковим чином [4]. Час, необхідний на обробку однієї вимоги, є випадковою величиною з експоненціальним законом розподілу ймовірностей, та, загалом кажучи, відрізняється для різних потоків вимог (2):

$$F_j(x) = 1 - e^{-v_j x}, v_j > 0, j = 1, 2, \dots, r. \quad (2)$$

В даному випадку  $v_j$  є коефіцієнтом складності задачі. Вимоги, що мають менше значення, будуть обслуговуватись швидше, а з більшим – довше. Існує

безліч поширених практичних застосувань, що вписуються у дану модель, та яким притаманно мати задачі різної тривалості. Отже, задана система підпадає під умовне визначення:  $M / M / n / m$ , де:

– перша  $M$  вказує на Марківський вхідний потік вимог;

– друга  $M$  вказує на те що процес обробки вимог також є Марківським;

–  $n$  – визначає, що система є багатоканальною, і задає кількість каналів;

–  $m$  – описує обмежену ємність системи та визначає кількість місць для очікування  $m > 0$ .

Кожен вхідний потік вимог має задовольняти наступні вимоги [4]:

- стаціонарність потоку;
- відсутність післядії;
- ординарність.

Функція розподілу загального часу обслуговування вимог потоків всіх розмірів має вигляд (3) [3]:

$$F(x) = \frac{1}{\lambda} \sum_{j=1}^r \lambda_j F_j(x), \quad (3)$$

де  $F_j(x)$  – функція розподілу часу обслуговування вимоги  $j$ -го розміру, що має середній час обслуговування  $\frac{1}{\mu_j}$ . Переривання обслуговування в досліджуваній системі не допускається. На даному етапі потоки заявок усіх розмірів мають рівний пріоритет, і поєднуються в загальний потік вимог.

Ймовірність надходження нової вимоги довільного розміру до системи за проміжок часу  $t$  може бути визначений наступним чином (4):

$$P_i(t) = \frac{(t\lambda_{\bar{r}})^i}{i!} e^{-t\lambda_{\bar{r}}}, \quad (4)$$

де  $\lambda_{\bar{r}}$  – сукупна інтенсивність прибуття вимог усіх розмірів, що надходять до системи на одиницю часу, а  $i$  – кількість вимог, що присутні у системі, включно з наступною, в момент часу  $t$ .

Враховуючи обмеження на кількість вимог, що можуть очікувати на обслуговування у черзі, розраховуємо ймовірність відмови у виконанні вимоги через переповнення черги. Оскільки майбутнє протікання обслуговування, у контексті теорії ймовірностей, не залежить від того, що відбувалось до моменту часу  $t_0$  в силу особливостей ймовірнісного розподілу [3, 4].

Інтенсивність навантаження  $j$ -ї категорії знайдемо через рівняння:  $\rho_j = \frac{\lambda_j}{n\nu_j}$ , де  $j=1,2,\dots,r$ . Тоді загальне навантаження системи можемо розрахувати так (5):

$$R_{zah} = \sum_{j=1}^r \rho_j, R_0 = 0, j=1,2,\dots,r, \quad (5)$$

Для описаної системи ймовірність відмови дорівнює ймовірності знаходження у системі рівно  $i$  вимог на момент надходження чергової заявки довільного розміру. Маємо наступний вираз (6), (7):

$$\begin{cases} 1 \leq i \leq n, & P_i = \frac{n^i R_{zah}^i}{i!} P_0, \\ n < i \leq n+m, & P_i = \frac{n^n R_{zah}^i}{n!} P_0, \end{cases} \quad (6)$$

$$P_0 = \left[ \sum_{i=0}^{n-1} \frac{n^i R_{zah}^i}{i!} + \frac{n^n}{n!} \cdot \frac{R_{zah}^n (1 - R_{zah}^{m+1})}{1 - R_{zah}} \right]^{-1}. \quad (7)$$

Розглядатимемо випадки, коли  $R_{zah} > n$ , оскільки за такої умови забезпечується наповнення черги і відмова через брак місця для очікування. Отже, ймовірність відмови через переповнення черги може бути розрахована виразом (8), (9):

$$P_{vidm} = P_{n+m} = \frac{n^n R_{zah}^{n+m}}{n!} P_0 \quad (8)$$

$$P_{vidm} = \frac{n^n R_{zah}^{n+m}}{n!} \left[ \sum_{i=0}^{n-1} \frac{n^i R_{zah}^i}{i!} + \frac{n^n}{n!} \cdot \frac{R_{zah}^n (1 - R_{zah}^{m+1})}{1 - R_{zah}} \right]^{-1} \quad (9)$$

Даний вираз означає, що запит отримає відмову щодо обслуговування, якщо всі канали й місця очікування будуть зайняті. При цьому розбиття загального потоку заявок на категорії за розміром, саме по собі, не змінює загальну пропускну здатність системи.

Алгоритм дії даної системи зображено на рис.1.

Перед нами постає задача про збільшення пропускну здатності описаної системи задля зменшення загальної ймовірності відмови. При цьому, з точки зору результатів, для нас всі заявки мають однакове значення, незалежно від їхнього розміру та часу перебування в системі.

### 3. Вирішення проблеми

В даній статті будемо розглядати рішення поставленої задачі за рахунок розбиття загального потоку заявок на сімейство потоків за розміром. З подальшим наданням меншим задачам більшого пріоритету. Суть підходу полягає в тому, що ми від простої СМО переходимо до системи з комбінованими пріоритетами [5]. Цільова схема має гібрид пріоритетів за класами вимог з динамічними, відносними пріоритетами. Це означає, що ми визначимо фіксовану кількість категорій заявок. І сортуватимемо вимоги на вході в систему. При цьому пріоритет буде гнучко підлаштовуватись під стан системи [6]. Надання пріоритету меншим задачам дозволить зменшити загальну ймовірність відмови через переповнення черги в системі. Такий вибір ґрунтується на тому, що причина відмови залежить від кількості задач, а не їх складності. На відміну від продуктивності системи, яка у свою чергу, залежить від часу, необхідного на обробку заявки, і не залежить від кількості заявок в черзі. Відповідно, надаючи перевагу меншим задачам, ми можемо збільшити кількість опрацьованих задач за відрізок часу. І таким чином мінімізувати кількість відмов [5, 6].

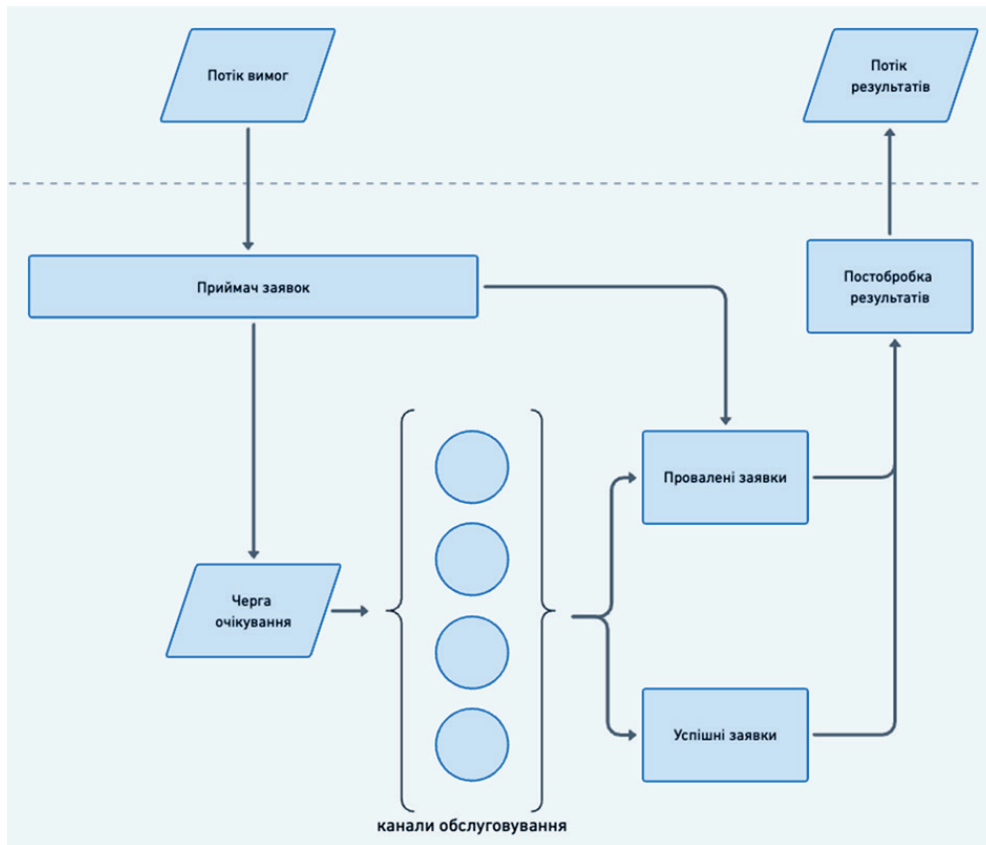


Рис. 1. Алгоритм обробки заявок

Встановлення пріоритетів для вимог, що очікують – один із ефективних способів керування розмірами черги та часом перебування в ній [5]. Для системи з пріоритетами всі вимоги поділяються на категорії, а заявки більш високої категорії при обслуговуванні мають певні переваги перед вимогами з нижчим пріоритетом. Оскільки ми не маємо на меті припинити обслуговування довгих вимог, введемо кілька додаткових правил. По перше – ми не перериватимемо обробку вимоги, навіть якщо вона має нижчий пріоритет (застосування відносного пріоритету). По друге – наявність заявок вищої категорії в черзі не забороняє взяття в роботу менш важливої задачі, а тільки надають додаткових преференцій при виборі. Для кількісної оцінки ефективності системи масового обслуговування з пріоритетами використовують ті самі характеристики, як і для простої системи. Але, на додачу до загальних показників, це робиться для кожної категорії окремо.

Оскільки початковий потік вимог є не ранжованим, постає потреба у визначенні способу його розбиття на окремі категорії заявок. Цю задачу буде вирішувати модуль пріоритетизації. В його обов'язки входить:

- оцінка складності заявки та призначення певної категорії;
- оцінка поточного завантаження системи та ведення статистики щодо допущення задач різної складності;

– прийняття рішення про надання вимоги на виконання.

### 3.1. Оцінка складності

За критерій розбиття ми обираємо час, необхідний для обробки вимоги. Оскільки саме ця ознака повинна надати нам підстави для ефективнішого планування роботи. Для оцінки складності заявки можна застосувати один з наступних підходів або їхню комбінацію:

- самоідентифікація заявки;
- особливості функції ймовірнісного розподілу часу обслуговування загального потоку заявок;
- застосування навченої моделі штучного інтелекту.

Найпростішим рішенням може бути делегація зобов'язання за визначення розміру на саму заявку. Тобто, в момент класифікації вимога має надати чисельне значення еквівалентне рівню складності її обробки. Таким чином, алгоритм категоризації може бути суттєво спрощений. Однак, постають кілька недоліків, а саме: заявка може надати недостовірну інформацію; вимоги різного походження можуть використовувати відмінні відносні шкали оцінки складності. Для усунення першої проблеми ми можемо скомбінувати даний підхід з більш достовірними опціями. У другому випадку нам знадобиться напрацювати алгоритм узгодження оцінок.

При застосуванні функції розподілу початкового вхідного потоку вимог нам буде необхідно зібрати певну статистику залежності часу виконання від параметрів. Наступним етапом буде визначення того, за яких вхідних даних час обслуговування потрапляє в зону математичного очікування, а за яких — відхиляється в менший або більший бік. Завдяки цьому, для подальших надходжень матимемо алгоритм категоризації за розміром та призначення пріоритету.

На відміну від попереднього підходу, застосування штучного інтелекту (ШІ) вимагає заздалегідь виконаного навчання моделі під потреби конкретної задачі [7]. Однак, якість визначення категорії буде вищою за рахунок більш адаптивного принципу дії ШІ та машинного навчання. Додатково можна продовжувати вдосконалення моделі прямо під час обробки потоків заявок. Тим самим гнучко підлаштовуватись під зміни у вхідних даних.

Застосування двох принципів одночасно може підвищити якість передбачення складності задачі та покращити ефективність оптимізації [8]. В такому разі аналіз функції розподілу може використовуватись для побудови початкових припущень, які в подальшому будуть уточнюватись за допомогою штучного інтелекту.

### 3.2. Оцінка поточного завантаження

Важливим етапом оптимізації є контроль поточно-го завантаження системи. При цьому нас не цікавить кількість одночасно працюючих каналів обслуговування. Оскільки, з точки зору модуля пріоритезації, всі сценарії, за яких система працює з недобором роботи, означають відсутність необхідності розділяти вхідний потік [3]. Для правильного функціонування необхідно збирати інформацію щодо поточного складу черги. Таким чином, ми будемо розуміти відносно чого ми пріоритезуємо задачі. Відповідно, ширина часових діапазонів (складності вимог) для категорій задач та загальна кількість класів заявок залежить від розміру черги.

Унаслідок використання пріоритету за класами обслуговування під кожен клас виділяється окрема черга [8, 9]. Це забезпечить збереження послідовності обслуговування заявок, що були розподілені до однієї категорії. Розміри цих черг залежать від функції ймовірнісного розподілу, що регулює час обробки заявок загального потоку вимог та визначаються перед початком роботи системи. Разом з тим, наповненість кожної з черг є важливою ознакою. За такої будови наша система буде давати відмову тільки тим заявкам, які неможливо додати в чергу згідно з визначеною категорією [7]. Відповідно, ми маємо неперервно слідкувати за наповненістю кожної з черг категорій задля розуміння завантаженості СМО. Для цього використовуватимемо

метрику — кількість вільних місць на певний клас заявок.

### 3.3. Балансування пріоритетів

Головним завданням, яке вирішує модуль пріоритезації є прийняття рішення про обрання класу заявок, що використовуватиметься як джерело для постанови в чергу, з якої беруть собі роботу канали обслуговування. З технічної точки зору це означає, що даний елемент системи, окрім черги, яка використовується для подачі вимог на виконання, взаємодіє з окремими чергами для задач кожного з класів.

Таким чином, на рис. 2 можемо побачити оновлений алгоритм для заданої СМО. Зміни торкнулися передусім механізму прийому заявок в систему, оскільки черга виконання залишилась тою самою. Але тепер їй передедує аналіз та сортування заявок по чергах відповідних розмірів. Після чого, враховуючи поточний склад черги виконання, а також наповненість черг по категорії заявок, відбувається прийняття рішення про постанову вимоги певного розміру до черги виконання.

Застосування такого рішення дозволяє уникнути додаткової категорії відмов та збільшити пропускну здатність системи. Оскільки кількість місць очікування в системі збільшилась, маємо зміни у формулі, яка описує загальну ймовірність відмови в системі. Для найгіршого сценарію, коли всі заявки, що надходять до системи мають один розмір отримаємо наступний вираз (10):

$$P_{\text{відм}} = P_{n+m+j} = \frac{n^n R_{\text{zah}}^{n+m+j}}{n!} P_0, \quad (10)$$

де  $m$  — кількість місць очікування в черзі з якої отримують заявки канали обслуговування, та  $m_j$  — розмір черги для  $j$ -ї категорії,  $j = 1, 2, \dots, r$ .

Сортувальник заявок має повідомляти відправнику заявок про переповнення черги певного розміру шляхом обміну сигналами. Це необхідно робити з метою зміни пріоритетів та уникання подальших відмов через переповнення черги категорії. Окрім того, ми маємо пропускати великі задачі до системи, щоб не припинити їх обслуговування [10]. Про це має дбати модуль пріоритезації. За таких обставин наш алгоритм не гарантує співпадіння послідовності надходження заявок до системи з послідовністю надходження на виконання. Але ми зберігаємо цю послідовність в рамках кожної окремої категорії задач. Послідовність надходження результатів обслуговування при кількості каналів більше за один — не має обмежень щодо відповідності послідовності через асинхронну та випадкову природу процесу обслуговування.



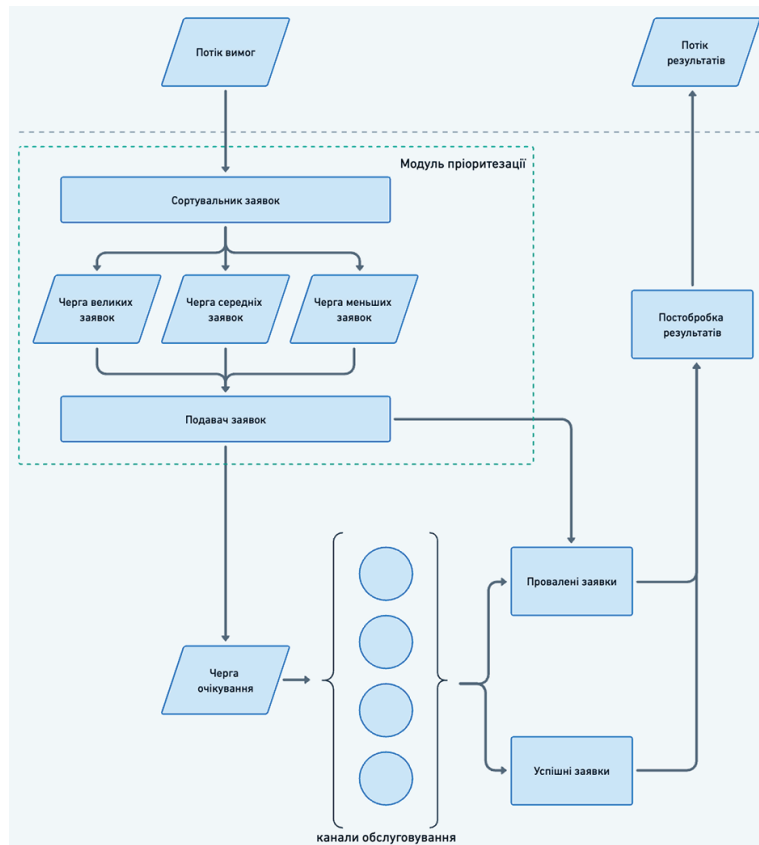


Рис. 2. Алгоритм обробки заявок з розбиттям

#### 4. Результати дослідження

Розбиття загального вхідного потоку заявок на класи за їх розміром – є необхідною передумовою для подальших удосконалень моделі системи. Запропоновано декілька підходів до визначення складності вимоги, серед яких:

- самоідентифікація заявки;
- застосування функції ймовірнісного розподілу;
- застосування штучного інтелекту.

Було розглянуто переваги та недоліки для кожного з вище наведених рішень. Задля підвищення якості категоризації можливе застосування довільних комбінацій даних підходів.

Для оцінки результатів дослідження пропонується використовувати декілька чисельних метрик, які мають якнайкраще характеризувати наслідки застосування оптимізації. Головною метрикою буде загальна ймовірність відмови новоприбулій довільній заявці через переповнення черги очікування. Для глибшого розуміння природи оптимізації розраховуватимемо показники ймовірності відмови за умови повної пріоритетизації по кожному з розмірів вимог. На рис. 3 зображено три графіки кожен з яких відповідає різним стратегіям:

- пріоритет великим задачам (пунктирна лінія);
- без застосування пріоритетизації (цілісна лінія);
- пріоритет меншим задачам (штрих-пунктирна лінія).

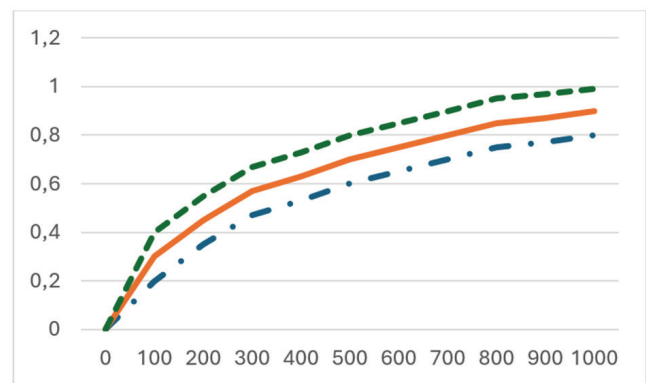


Рис. 3. Ймовірність відмови за пріоритетних категорій

Як бачимо, за умов коли інтенсивність надходження задач суттєво перевищує пропускну здатність системи, пріоритетизація певної категорії заявок призводить до суттєвих змін пропускну здатності системи. Це пояснюється витісненням непріоритетних задач, а також збільшенням відсотку відмов для такої категорії. Відповідно, коли ми надаємо пріоритет меншим задачам, ми майже повністю перестаємо обслуговування великих задач, і навпаки.

Оскільки ми зацікавлені у збереженні балансу розміру виконаних задач, щоб оптимізація не призвела до повного усунення всіх заявок, розподілених до класу більших, будемо розглядати балансні політики. Вони призводять до потрапляння в проміжок між лініями графіків пріоритету менших задач та без пріоритету [11]. Це дозволить збільшити загальну

пропускну здатність системи, надаючи при цьому можливість обробляти для більших задач.

Наступним кроком буде розробка алгоритму, який даватиме стабільне потрапляння в зону балансу. Задля чисельного вимірювання успішності механізму пріоритетизації, введемо метрику, яка буде відображати пропорційність добору задач. Вона представлятиме собою співвідношення відмов великим заявкам до загальної кількості відмов за проміжок часу  $t$  (11).

$$V(t) = \frac{V_L}{V_{zah}}, \quad (11)$$

де  $V_L$  – це відсоток відмов для великих задач, а  $V_{zah}$  – загальна кількість відмов у системі за проміжок часу  $t$ . Визначивши розмір проміжку часу, за який вестимуться спостереження, ми отримуємо періоди коригування балансу [8]. Після завершення збору аналітики по відмовах у системі, настає корекція коефіцієнту пріоритетизації, який змінить певним чином пропорцію великих, середніх та маленьких заявок що перебувають у черзі.

Для апроксимації коефіцієнта пріоритетизації пропонується застосовувати штучний інтелект. Модель якого попередньо має бути натренована на широкій вибірці даних. Важливо, щоб функція ймовірнісного розподілу вхідного потоку співпадала з функцією, що буде використовуватись при навчанні моделі штучного інтелекту. Із поглибленням інтеграції та довчанням моделі під час її застосування, будуть розширюватись можливості щодо виявлення певних патернів (послідовностей) поведінки в балансуванні навантаження [9]. Це дозволить передбачати наперед надмірне завантаження системи та завчасно адаптуватись за рахунок зміни пріоритетів. Дані переваги, у свою чергу, в повній мірі виправдовують додаткову складність, пов'язану з використанням штучного інтелекту для вирішення проблеми збільшення пропускну здатності СМО.

### Висновки

Метод розбиття загального потоку заявок багатоканальної системи масового обслуговування з обмеженою чергою та відмовами на категорії за часом необхідним для обробки може бути дієвим способом зменшити загальну ймовірність відмови через переповнення черги. Такий підхід дозволяє більш ефективно задіяти вже наявні ресурси, за умови відсутності можливості збільшити кількість каналів обслуговування. Задля підвищення ефективності балансування пріоритетів, бажано на перших етапах використовувати особливості функції ймовірнісного розподілу інтенсивності надходження заявок до системи. В подальшому рекомендується застосування навченої моделі штучного інтелекту до оцінки складності задач, а також балансування пріоритетів між більшими та меншими вимогами. Такий метод

збільшення продуктивності системи може бути інтегрований в архітектуру застосунку, призначеного для програмного моделювання СМО.

Було вперше запропоновано застосування оптимізації обробки заявок шляхом розбиття вхідного потоку вимог на класи згідно з розмірами та поєднання із використанням штучного інтелекту для оцінки складності заявок і подальшого балансування пріоритетів.

В подальшому дані підходи мають бути застосовані та реалізовані у програмному забезпеченні, призначеного для моделювання СМО. Такий продукт, у свою чергу, надасть можливість напрацювати певні балансні стратегії та порівняти їх ефективність на практиці.

### Список літератури

- [1] Newell G. Applications of queuing theory. Second edition. – 1982. – Chapman and Hall – P. 1-302.
- [2] Haenlein M., Kaplan A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence // California Management Review. – 2019. – №. 61(4). – P. 5-14.
- [3] Ложковський А. Теорія масового обслуговування в телекомунікаціях. – 2010. – Одеса: ОНАЗ ім. О. С. Попова – 112 с.
- [4] Литвинов А. Теорія систем масового обслуговування. – 2018. – Харків: ХНУМГ ім. О. М. Бекетова – 141 с.
- [5] Alotaibi F., Ullah, I., Ahmad S. Modeling and Performance Evaluation of Multi-Class Queuing System with QoS and Priority Constraints // Electronics. – 2021. – №. 10(4). – P. 500.
- [6] Beshley M., Kryvinska N., Beshley H., Yaremko O.; Pyrih J. Virtual Router Design and Modeling for Future Networks with QoS Guarantees // Electronics. – 2021. – № 10(10). – P. 1139.
- [7] Vercellino C., Scionti A., Varavallo G., Viviani P., Vitali G., Terzo O. A Machine Learning Approach for an HPC Use Case: the Jobs Queuing Time Prediction // Future Generation Computer Systems. – 2023. – №. 143. – P. 215-230.
- [8] Malik S., Gupta K., Gupta D., Singh A., Ibrahim M., Ortega-Mansilla A., Goyal N., Hamam H. Intelligent Load-Balancing Framework for Fog-Enabled Communication in Healthcare // Electronics. – 2022. – №. 11(4). – P. 566.
- [9] Malik N., Sardaraz M., Tahir M., Shah B., Ali G., Moreira F. Energy-Efficient Load Balancing Algorithm for Workflow Scheduling in Cloud Data Centers Using Queuing and Thresholds // Applied Sciences. – 2021. – №. 11(13). – P. 5849.
- [10] Apachidi X., Katsman Yu. Development of a Queuing System with Dynamic Priorities // Key Engineering Materials. – 2016. – №. 685. – P. 934-938.
- [11] Chikriy A., Gubarev V., Kondratenko Y., Turovyerova N. Multi-channel Queuing Systems with the Dynamic Priority // Journal of Automation and Information Sciences. – 2009. – №. 41(8). – P. 49-54.

Надійшла до редколегії 14.05.2024



С. М. Неронов<sup>1</sup>, Г. А. Плехова<sup>2</sup>, М. В. Костікова<sup>3</sup>

<sup>1</sup>ХНАДУ, м. Харків, Україна, sernikner@gmail.com, ORCID iD: 0000-0003-2381-1271

<sup>2</sup>ХНАДУ, м. Харків, Україна, plehovaanna11@gmail.com, ORCID iD: 0000-0002-6912-6520

<sup>3</sup>ХНАДУ, м. Харків, Україна, kmv\_topaz@ukr.net, ORCID iD: 0000-0001-5197-7389

## ВІРТУАЛЬНЕ УПРАВЛІННЯ ПЕРЕВІЗНИМИ ПРОЦЕСАМИ НА АВТОМОБІЛЬНОМУ ТРАНСПОРТІ

У зв'язку з доцільністю перенесення віртуального управління автотранспортом в хмарну середу в роботі розглянута апаратна віртуалізація, Web-рішення та інтерфейс користувача Neuronet автотрансферу. Цей матеріал буде корисним при розгортанні клієнт-сервісної технології рухомого складу автомобільного трансферу. АВТОТРАНСФЕР, NEURONET, WEB-КОМПЛЕКС, OPENGTS, МАРШРУТИЗАТОР

**S. M. Neronov, G. A. Pliekhova, M. V. Kostikova. Virtual management of transport processes in road transport.** In connection with the expediency of transferring virtual vehicle management to a cloud environment, hardware virtualization, Web solutions and the Neuronet user interface of auto transfer are considered in the paper. This material will be useful when deploying the customer service technology of rolling stock of car transfer. AUTO TRANSFER, NEURONET, WEB-COMPLEX, OPENGTS, ROUTER

### Вступ

Завдяки віртуалізації ІТ адміністратор або аналітик вимог може поділити фізичний сервер на будь-яку кількість віртуальних серверів, кожен з яких працює під управлінням власної операційної системи і кожному з яких виділяються такі ресурси, як пам'ять, CPU, ділянки дискового простору. Деякі технології віртуалізації навіть дозволяють переміщати запущені екземпляри віртуальних серверів з одного фізичного сервера на інший. З точки зору користувача для додатків, які працюють на віртуальному сервері, не існує ніяких можливостей визначити, чи є сервер, на якому вони працюють, віртуальним або фізичним. Ряд віртуалізаційних технологій, доступних на ринку, використовують різні підходи до проблеми віртуалізації. Гіпервизор створює рівень апаратних абстракцій, який дозволяє операційним системам спільно використовувати ресурси фізичного сервера, при цьому гостьові операційні системи не мають фізичного доступу до цих ресурсів, як «своїм», так і «чужим».

Загальновідоме заперечення проти віртуалізації, особливо з боку тих, хто працював з емуляторами на настільних комп'ютерах під управлінням настільних операційних систем полягає в тому, що віртуалізовані системи істотно втрачають в продуктивності. Але в хмарі це загальноприйняте упередження проти віртуалізації не має під собою підстав з наступних причин: навіть знижена продуктивність обладнання вашого постачальника хмарних послуг, цілком ймовірно, виявиться краще, ніж максимальна продуктивність вашого споживчого сервера, і це за тієї умови, що він оптимальним чином налаштований; технології віртуалізації, призначені для великих підприємств. Це дозволяє їм досягти майже такої ж продуктивності, як і при роботі програм на «живому залізі». Розглянемо також і особливості створення хмарних програм.

### Виклад основного матеріалу

#### 1. Апаратна віртуалізація

Абстрагування від апаратних засобів в хмарі здійснюється не тільки завдяки заміні фізичних серверів віртуальними. Віртуалізації підлягають і системи фізичного зберігання даних. Хмарне сховище дозволяє перекидати дані в хмару, і при цьому не турбуватися про те, як саме вони зберігаються, і не замислюватися про їх резервне копіювання. Коли дані, переміщені в хмару, знадобляться знову, досить буде просто звернутися в хмару і отримати свої дані. При цьому ви можете не знати, як зберігаються ці дані, де вони зберігаються, чи того, що відбувається з тим або іншим обладнанням, коли ви періодично переміщуєте дані в хмару і витягаєте їх звідти.

Як і у випадку з іншими елементами хмарних обчислень, на ринку ІТ-індустрії пропонується кілька підходів до хмарного сховища. Говорячи загальними словами, вони пов'язані з розбивкою ваших даних на невеликі ланцюжки і зберігання їх на безлічі серверів. Ланцюжки даних забезпечуються індивідуально обчисленими контрольними сумами так, щоб дані можна було швидко відновити, незалежно від того, що могло б статися протягом часу зберігання з накопичувачами, фізично зберігаючи дані, і скомпрометувати хмару.

Бувають випадки, коли люди, початківці працюючи з хмарою, намагалися поводитися з хмарним сховищем так, як якщо б воно представляло собою мережевий накопичувач. З точки зору принципу роботи хмарне сховище принципово відрізняється від традиційних мережевих накопичувачів, і служить принципово іншим цілям. Хмарне сховище має тенденцію працювати набагато повільніше і має набагато більшу ступінь структурованості, внаслідок чого його використання в якості оперативного сховища даних

непрактично, незалежно від того, чи працює програма, що використовує ці дані, в хмарі або десь ще.

Cloud сховище не підходить для оперативного використання хмарними програмами на базі транзакцій. Поки що про хмарне сховище можна думати приблизно так само, як про аналог резервної копії на стрічковому носії, але, на відміну від системи резервного копіювання з стрічковим приводом, при роботі в хмарі не потрібні ні привід, ні стрічки.

Компанія Amazon виступила з комерційною пропозицією, яка називається Amazon CloudFront. Amazon CloudFront використовує Amazon S3 як засіб для реалізації мережі, призначеної для розподілу інформаційного вмісту. Ідея, на якій ґрунтується Amazon CloudFront, полягає в реплікації вашого інформаційного вмісту по периферії мережі. В майбутньому воно обіцяє стати ключовим компонентом систем швидкого розподілу статичного контенту [1].

## 2. Web-рішення та інтерфейс користувача Neuronet автотрансферу

Розглянемо проблему визначення інтерфейсу користувача Neuronet автотрансферу. Для автотрансферу основним є обрання найбільш раціонального маршруту доставки вантажів. Формально є чотири завдання. По-перше, це завдання поставки вантажів від декількох постачальників одному споживачеві. По-друге, це зворотне першому – завдання поставки вантажів від одного постачальника декільком споживачам. По-третє – це об'їзд кількох пунктів призначення одним транспортним засобом. По-четверте – це знаходження найбільш раціонального маршруту проїзду з вихідного пункту в пункт призначення. Транспортна система в такій постановці розглядається як мережа, що описується зваженим графом, в якому ребрам відповідають ваги. У транспортних системах цим вагам відповідають вартість транспортного процесу, надійність його реалізації, а в найпростішому випадку відстань. Тоді відповідною оцінкою транспортної системи буде матриця ваг  $W = [w_{ij}]$ .

Практично всі відповідні математичні задачі: і транспортна задача лінійного програмування, і задача комівояжера, і просто задача знаходження найкоротшого шляху вирішуються за допомогою цієї оцінки  $W = [w_{ij}]$  транспортної системи. Для нього поряд із загальними методами вичерпного пошуку і решета застосовують як динамічне програмування, так і порівняно універсальний відомий алгоритм Дейкстри. Для його реалізації частина карти з мережею автомобільних доріг, на якій вибирають раціональний маршрут руху, представляють у вигляді матриці суміжності  $W$ , в якій:  $w(i, j)$  довжина ребра між вершинами  $i$  і  $j$ ;  $w(i, j) = +\infty$ , якщо між  $i$  і  $j$  немає ребра;  $w(i, i) = 0$  для всіх  $i = 1, 2, \dots, m$ , де  $m$  – кількість вузлів.

Власне пошук оптимального маршруту зводиться до ітераційної процедури аналізу результату додавання-видалення відповідного пункту-вершини для поновлення варіантів маршруту. Найкоротший шлях до даної вершини відшукується в явному вигляді простеження шляху по такому своєрідному графу дереву з кореня – початкового пункту до що цікавить нас вершини-пункту призначення.

Близькою до задачі знаходження найкоротшого шляху є відома задача комівояжера, яка дозволяє вибрати найменший за витратами маршрут з послідовності пунктів призначення  $i$ , що не повторюються. Для вирішення цієї задачі, попередньої і практично будь-якої подібної задачі використовуються як розглянутий алгоритм, так і універсальні методи динамічного програмування, найбільш узагальнені методи гілок і меж, решета і пошуку з поверненням. В окремому випадку можливий пошук оптимального рішення методами лінійного програмування. Так в задачах маршрутизації широко використовується відома транспортна задача організації перевезень вантажів.

Веб-інтерфейс порталу повинен мати можливість відбиття картографічних даних, інформації у вигляді таблиць та графічних даних. Він повинен враховувати сучасні тенденції подання даних та бути інтуїтивно-зрозумілим для користувачів. Транспортний Інтернет-портал повинен містити готовий набір Web 2.0 сервісів в єдиній компонентній архітектурі. Сервіси повинні мати, як серверні компоненти, так і готові графічні інтерфейсні компоненти для кінцевого користувача, реалізовані в єдиному стилі. Сервер транспортного порталу повинен відповідати вимогам необхідності і достатності для виконання відповідних завдань з урахуванням зростання обсягів обчислень в міру збільшення потоку інформації.

Однією з основних функцій транспортного порталу є відбиття місця розташування рухомого об'єкту на карті місцевості. Для рішення цієї задачі доцільним є вибір системи просторово-часової орієнтації OpenGTS (Open GPS Tracking System). Це проект з відкритим вихідним кодом, призначений спеціально для забезпечення веб-GPS спостереження за транспортними засобами.

Зараз OpenGTS був завантажений та введений в експлуатацію більше ніж в 95 країнах світу для відстеження транспортних засобів. Ця система пристосована для моніторингу різних типів транспортних засобів: автомобілів, спеціальних транспортних засобів із гусеничним ходом, таксі, автофургонів, вантажівок, сільськогосподарської техніки, приватних транспортних засобів. Також OpenGTS застосовується для обслуговування транспортних засобів, контейнерів, квадроциклів, персональних систем спостереження, визначення та моніторингу місця знаходження стільникових телефонів та інших засобів.

Хоча OpenGTS розроблена, щоб задовольнити потреби початкового рівня систем стеження, ця система гнучка у конфігуруванні та має засоби масштабування на рівні обслуговування великого підприємства (<http://www.opengts.org/>). OpenGTS розповсюджується за ліцензією Apache Software, версія 2. Відповідно до умов цієї ліцензії, будь-хто може

вільно завантажити і поширювати відповідні інструментальні засоби. Для комерційних додатків пропонується платна версія – GTS Enterprise.

OpenGTS не залежить від базової операційної системи тому, що її реалізовано на мові Java із використанням технології Apache Tomcat для розгортання веб-служб, а також MySQL для сховища даних (рис. 1).

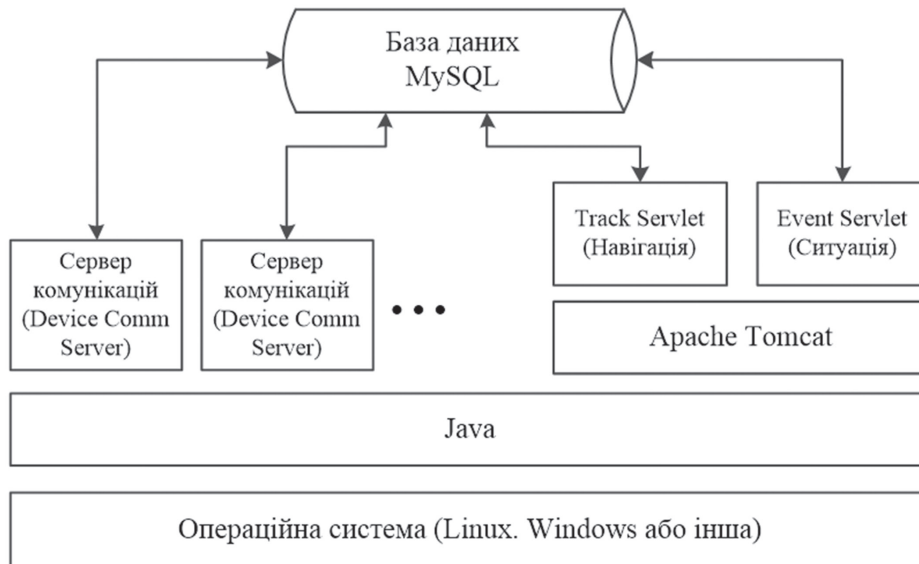


Рис. 1. Просторово-часова орієнтація (OpenGTS)

Таким чином, OpenGTS працюватиме на будь-якій системі, яка підтримує ці технології (включаючи Linux, MacOS X, FreeBSD, Solaris, Windows та багато ін.). OpenGTS підтримує не тільки збір та зберігання GPS-даних та телеметричної інформації з віддалених пристроїв, а також включає до себе набір функцій:

- Web-аутентифікація (кожен обліковий запис може підтримувати декілька користувачів, де кожен користувач має свій власний пароль, логін й контрольований доступ до розділів;

- GPS-пристрої (підтримуються різні апаратні платформи, наприклад: Aspicore GSM Tracker на телефонах Nokia, Samsung, Sony Ericsson, персональні трекери Sanav (протокол на базі HTTP), GPSReader для комп'ютерів на базі операційної системи Windows, стільникові телефони та WindowsCE-комунікатори (на базі протоколу OpenDMTP), смартфони із операційною системою Android (застосунок GPS2OpenGTS), засоби на базі TAIP (Trimble ASCII Interface Protocol), пристрій TrackStick та ін.;

- настройка зовнішнього вигляду веб-сторінок (зовнішній вигляд та інтерфейс веб-сайту можна налаштувати відповідно прийнятому стилю в організації, що впроваджує систему);

- OpenGTS поставляється з підтримкою OpenLayers/OpenStreetMap додатково до підтримки Google Maps та Mapstraction. У рамках OpenGTS можна застосовувати й інших постачальників послуг, відповідний функціонал може бути легко інтегровано до системи;

- використання внутрішньої XML-розмітки надає можливості налаштувати вигляд вихідних звітів;
- підтримуються різні мови інтерфейсу, наприклад, англійська, французька, німецька, російська та ін.

Для швидкого розгортання та знайомства із системою OpenGTS доцільно обрати LiveCD, що забезпечує запуск повністю сконфігурованої системи (<http://www.nwitechsupport.com/2011/11/15/opengts-livecd/>). Після запуску OpenGTS LiveCD віддалений користувач може вести обмін даними із сервером.

Для взаємодії з OpenGTS віддалений пристрій повинен сформувати HTTP-запит у форматі «[http://192.168.1.4:8080/gprmc/Data?acct=admin&dev=test01&code=61472&gprmc=\\$GPRMC,\\*,\\*,\\*,\\*3В](http://192.168.1.4:8080/gprmc/Data?acct=admin&dev=test01&code=61472&gprmc=$GPRMC,*,*,*,*3В)», де 192.168.1.4:8080 – адреса серверу й відповідний порт сервісу OpenGTS; acct – обліковий запис користувача; dev – обліковий запис пристрою чи (або) автотранспортного засобу, позиція якого відстежується; code – код стану системи, позиція якої відстежується; gprmc – строчка «\$GPRMC», що надається GPS-приймачем та є рекомендованим мінімумом даних за стандартом NMEA.

Відповідно, транспортний портал може використовувати систему OpenGTS для відбиття даних про стан рухомих одиниць, та надавати інші сервісні функції користувачам та учасникам єдиного простору транспортних організацій (рис. 2).

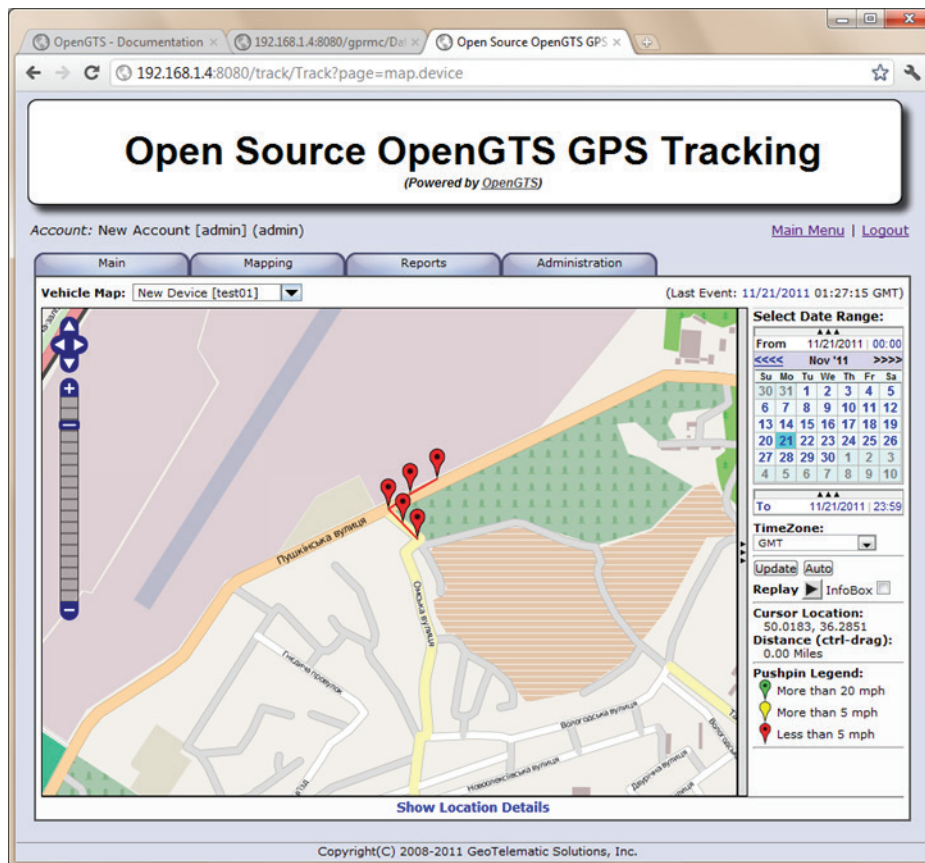


Рис. 2. Відбиття даних системою OpenGTS

У складі OpenGTS LiveCD, який у свою чергу оснований на Linux-дистрибутиві CentOS 6, є система phpMyAdmin, яка дозволяє досліджувати стан бази даних. Наявність мови php у LiveCD надає можливості розробки на її базі свого сервісу, що буде надавати дані про стан транспортного засобу.

Web-технологія руху наземного транспорту великих міст забезпечує синергетичне об'єднання комп'ютерних ресурсів усіх учасників дорожнього руху – від окремої транспортної машини до корпоративного рівня транспортної організації. Вона призначена для моніторингу транспортної системи міста та (або) регіону й повинна визначати місцезнаходження транспортного засобу, стан середовища руху і забезпечувати учасників дорожнього руху (водії транспортних засобів, транспортні організації) даними про стан транспортної мережі.

Інформаційно-комунікаційна технологія забезпечує реєстрацію, обробку та представлення даних учасникам руху в режимі реального часу, як для оперативного прийняття рішень з аналізу транспортних ситуацій, так і для накопичення даних. Інформаційні функції рішення завдань безперервного моніторингу міської транспортної мережі сприяють скороченню витрат на удосконалення існуючих транспортних систем. Відповідно, покращується якість прийняття рішень щодо управління транспортним обслуговуванням великих міст та регіонів України.

Вхідними даними для Web-комплексу є показники давачів прискорень транспортного засобу, його поточні координати в просторі та швидкість. Ці дані оброблюються та з них формується пакет, який із застосуванням засобів безпроводного зв'язку передається до транспортного порталу. Дані передаються та зберігаються у відкритому текстовому форматі (ASCII). Однак, сам пакет даних формується відповідно до прийнятих протоколів TCP/IP мережі Інтернет. Додатковою інформацією у системі є відеодані про стан середовища руху. Ця інформація потребує підвищеної пропускної здатності комунікаційних каналів.

Програмно-апаратне забезпечення Web наземного транспорту (Web-комплекс) дозволить проводити:

- відеоспостереження за станом середовища руху транспортних систем;
- екологічний моніторинг (реєстрація забруднення повітря);
- оцінку стану покриття (рівності та зчепних якостей) доріг;
- прив'язку даних про стан автомобільних доріг до певної точки траси вулично-дорожньої мережі із урахуванням зміни швидкості автомобіля під час вимірювань;
- управління процесом вимірювань та перегляду ретроспективної інформації;
- формування електронної бази даних моніторингу стану середовища руху;

– обробку результатів вимірювань та накопичення даних про стан транспортної системи.

Для роботи із Web-комплексом користувач може задіяти безпроводний зв'язок до Інтернет із свого смартфона або комп'ютера. У разі відсутності технічної можливості встановити обладнання інформаційно-комунікаційного комплексу у транспортний засіб можливим є застосування смартфона, планшетного комп'ютера або ноутбука із підключенням до приймача GPS-сигналів у якості своєрідного «давача» системи

На рис. 3 наведено функціональну схему програмно-апаратного комплексу Web-технології маршрутизації. Функціональна схема програмно-апаратного комплексу визначає, що для розробки інформаційно-комунікаційної технології руху наземного транспорту великих міст передбачено застосування одного (позначення: «1») транспортного порталу та багатьох фактично однотипних інформаційно-комунікаційних комплексів (позначення: «...»), що встановлюються на автотранспортний засіб.

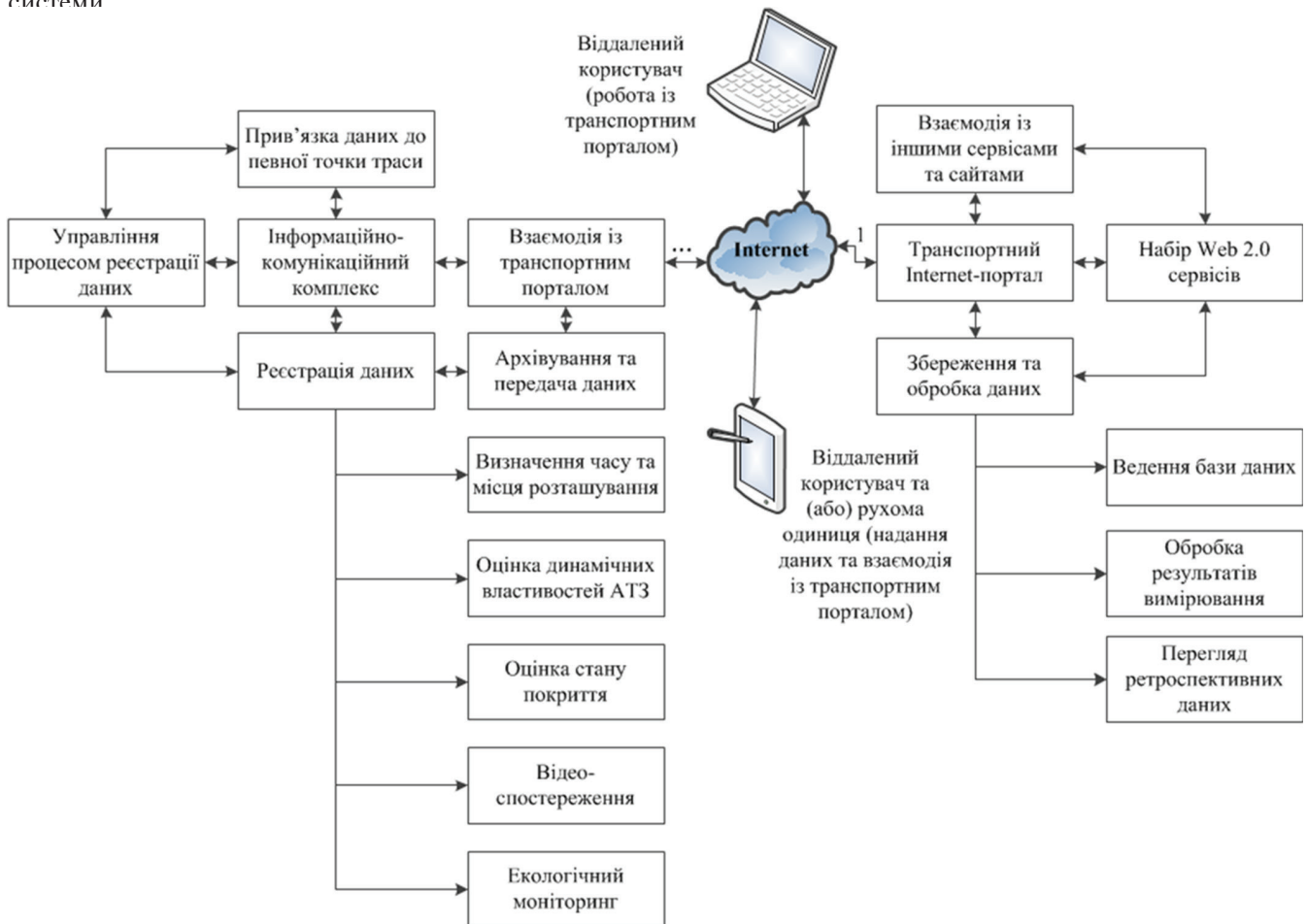


Рис. 3. Функціональна схема програмної платформи Web-порталу

Особливістю системи, що пропонується, є розміщення обладнання на борту транспортного засобу без значних конструкційних змін, як транспортного засобу, так й інформаційно-комунікаційного комплексу. Зараз для моніторингу у русі транспортних засобів на території України можна застосовувати безпроводний зв'язок за технологією 2G, 3G, WiMAX та Wi-Fi.

До стандартів передачі даних за стільниковою технологією другого покоління, або 2G, можна віднести технології GPRS та EDGE. Загальний сервіс пакетної радіопередачі (англ. General Packet Radio Service, GPRS) – стандарт, який використовує не зайняту голосовим зв'язком смугу частот для передачі інформації. Використовується в мобільних пристроях для передачі MMS, WAP-серфінгу та повноцінного

з'єднання з Інтернетом. EDGE (англ. Enhanced Data Rates for GSM Evolution) – сучасна технологія, що забезпечує передачу великих обсягів інформації в мережі мобільного зв'язку. Технологія EDGE підтримує у середньому в три рази вищу швидкість передачі даних, ніж GPRS, крім того, забезпечується більш ефективне використання частотних ресурсів і поліпшення покриття мережі в порівнянні зі звичайною мережею GSM. Максимально досяжна швидкість передачі інформації в мережі EDGE – 474,6 кбіт/с.

3G (англ. 3rd Generation) – «третє покоління», набір послуг, котрий включає до себе як високошвидкісний мобільний доступ до послуг мережі Інтернет, так і технологію радіозв'язку. Мережі третього покоління 3G працюють на частотах дециметрового діапазону

(близько 2 ГГц), швидкість передачі даних становить понад 2 Мбіт/с. Такі мережі надають можливість організувати відеозв'язок, дивитись на мобільному телефоні фільми й телепрограми та ін. В світі існує два стандарти 3G: UMTS (чи W-CDMA) та CDMA-2000. UMTS більш розповсюджений в основному в Європі, CDMA2000 – в Азії та США

Термін 3G використовується для опису сервісів мобільного зв'язку стандартів наступного покоління, які забезпечують більш високу якість звуку, а також високошвидкісний Інтернет-зв'язок та мультимедійні сервіси. Мобільні мережі третього покоління (3G) відрізняються від мереж другого покоління (2G), таких як наприклад цифровий стандарт мобільного зв'язку GSM, зв'язок перехідного покоління (2.5G) GPRS, набагато більшою швидкістю передачі даних, а також більш широким набором і високою якістю послуг, що надаються.

В Україні підприємство Укртелеком має мережу мобільного зв'язку 3G, що працює в стандарті UMTS 2100 з надбудовою HSDPA (3,5G). Якщо в базовій версії UMTS забезпечує пікові швидкості від 2 Мбіт за секунду для статичних об'єктів поблизу соти, та 384 Кб/с для мобільних абонентів, то для пристроїв, що підтримують HSDPA швидкості в теорії можуть досягати 14,4 Мбіт/с. На практиці ж, реальні швидкості не перевищують 3 Мбіти, а в умовах високих, щільних забудов та завантаженості мережі, ще менше. Зв'язок дає можливість здійснювати відеодзвінки, ширококутний доступ в Інтернет, а також переглядати потокове відео (<http://wikipedia.org>).

Також на території України компанія «Телесистеми України» експлуатує мережу 3G під торговою маркою «PEOPLEnet» стандарту CDMA2000 1xEV-DO rev0 (800 мГц) та EV-DO revA, а також МТС Україна надає послуги за технологією CDMA2000 EV-DO revA (450 мГц).

Відповідні стандарти застосовує компанія Інтертелеком з мережею CDMA2000 EV-DO revA (800 мГц) та EVDO Rev.B, яка дозволяє отримати абонентам можливості, які раніше були доступні лише при використанні проводного інтернету. 14,7 Мбіт/с – максимальна швидкість, з допомогою якої можливо обмінюватися фотографіями, дивитися відео, слухати музику у реальному часі.

Технології Wi-Fi та WiMAX також поширені на території України. Відповідні послуги надають Інтернет-провайдери. Ці технології відрізняються від стільникових та потребують наявності спеціального обладнання. Перевагою такого обладнання (переважно маршрутизатори та точки доступу) є можливість підключення клієнтського обладнання за провідною технологією Ethernet.

Технологія Wi-Fi (від англ. Wireless Fidelity) – торгова марка, що належить Wi-Fi Alliance.

Загальноживана назва для стандарту бездротового (радіо) зв'язку передачі даних, який об'єднує декілька протоколів та ґрунтується на сімействі стандартів IEEE 802.11 (Institute of Electrical and Electronic Engineers – міжнародна організація, що займається розробкою стандартів у сфері електронних технологій). Найвідомішим і найпоширенішим на сьогодні є протокол IEEE 802.11g, що визначає функціонування бездротових мереж. Установка Wireless LAN рекомендується там, де розгортання кабельної системи було неможливо або економічно недоцільно. Нині в багатьох організаціях використовується Wi-Fi, оскільки при визначених умовах швидкість роботи мережі вже перевищує 100 Мбіт/с. Користувачі можуть переміщатись між точками доступу по території покриття мережі Wi-Fi. Мобільні пристрої (смартфони, планшетні комп'ютери або ноутбуки), оснащені клієнтськими Wi-Fi прийомо-передаючими пристроями, можуть підключатися до локальної мережі і отримувати доступ в Інтернет через точки доступу.

WiMAX (від англ. Worldwide Interoperability for Microwave Access) – стандарт IEEE 802.16 безпроводного зв'язку, що забезпечує ширококутний зв'язок на значні відстані зі швидкістю, порівняною з кабельними з'єднаннями. Назву «WiMAX» було створено WiMAX Forum – організацією, яку засновано в червні 2001 року з метою просування і розвитку WiMAX. Форум описує WiMAX як «засновану на стандарті технологію, яка надає високошвидкісний бездротовий доступ до мережі, альтернативний виділенням ліній і DSL».

Виконав порівняльний аналіз різних технологій безпроводного доступу до Інтернет, можна стверджувати, що технології WiMAX та Wi-Fi є дуже перспективними, однак покриття відповідних операторів зв'язку не зовсім повністю охоплює територію великих міст України. Ця технологія ідеально підходить для передачі даних на зупинках транспортних засобів й може бути використана у інформаційно-комунікаційному комплексі, що розроблюється, для передачі значних обсягів даних, наприклад, відеофайлів.

Для розробки експериментального зразка інформаційно-комунікаційного комплексу було обрано рішення на базі технології 3G. Це обумовлено наявністю широкого покриття території України комунікаційним обладнанням операторів стільникового зв'язку. У якості найбільш універсального пристрою зв'язку обрано USB-модем Huawei E1550. Це USB-пристрій за стандартом UMTS 2100, GSM 850/900/1800/1900, який надає можливості передавати дані за технологіями: HSDPA, EDGE class 12, GPRS клас 12. Максимальна швидкість завантаження даних – 3,6 Мбіт/с (HSDPA), а передачі – 384 Кбіт/с. Пристрій має вбудовану антену та рознімання для підключення SD-карти пам'яті, що надає можливості



його використання у якості накопичувача чи зчитувача flash-карт пам'яті (обмеження – 2 Гбайта).

Для налаштування маршрутизатора та обладнання, яке можна до нього підключити по інтерфейсу USB, застосовується веб-інтерфейс.

Обраний маршрутизатор створено на базі мікропроцесору BCM4704/BCM4780 з частотою 264 MHz та має 32 Мбайт оперативної пам'яті та 8 Мбайт постійної пам'яті типу flash. Це дозволяє використовувати оновлену альтернативну прошивку на базі ядра Linux 2.6.22.

Оскільки операційна система Linux 2.6 має підтримку на рівні ядра драйверів підключення Web-камери із інтерфейсом USB, то до обраного маршрутизатора можливе підключення сучасної Web-камери із протоколом передачі даних UVC (USB video device class). Також, практично до будь-якого маршрутизатору можливо підключити спеціальну відеокамеру із інтерфейсом Ethernet або Wi-Fi.

Завдяки порівняно невеликій швидкості передачі даних за технологією 3G та невеликій продуктивності процесору камери та (або) маршрутизатора (у разі USB-підключення дискретизація картинки складає порядку 1 с) застосування таких апаратних засобів можливо у окремих випадках. Наприклад, у разі зйомки статичних об'єктів при зупинці транспортного засобу. Тому інформаційно-комунікаційний комплекс доцільно комплектувати автомобільним реєстратором, або цифровим фотоапаратом чи професійною відеокамерою.

Визначені пристрої дозволяють отримати зображення у русі транспортного засобу та виконують збереження даних на свій внутрішній носій. Зареєстровані візуальні дані можна завантажувати на сервер транспортного порталу для подальшої обробки та прив'язки до телеметричних даних.

Система, що розроблюється є універсальною. Тому для підтримки спеціального устаткування, наприклад, дозиметрів, пірометрів, систем оцінки забруднення повітря, що не мають цифрового інтерфейсу, передбачається пакетне завантаження файлів із даними, які зареєстровані, на Інтернет-сервер транспортного порталу.

Відповідні дані, як правило, збираються не досить часто (не більше декількох замірів на 1 км), що надає

можливості не розробляти дороге обладнання для автоматизації вводу цих специфічних даних.

Розглянемо Web-технології та системи Neuronet автотранферу на прикладі найбільш простої її реалізації Cloud – GRID. Розвиток технологій GRID знаходиться на стадії реалізації проектів, спрямованих на розгортання великих інфраструктур. GRID являє собою інфраструктуру, що складається із розподілених ресурсів, телекомунікацій (мережеві ресурси) та взаємного проміжного (middleware) програмного забезпечення, що підтримує виконання дистанційних операцій, а також виконуючої функції контролю й керування операційним середовищем. GRID створюється власниками ресурсів, що виділяють їх у загальне користування. Власники й споживачі, що діють на підставі певних правил надання/споживання ресурсів, утворюють віртуальну організацію.

### Висновки

Головне в удосконаленні перевізного процесу для ланцюга виробника, промисловості, перевізника, отримувача є задача надання учасникам перевізного процесу, особам, що приймають рішення з віртуального управління транспортними та дорожніми організаціями, інформацію про дорожні ситуації.

Рішення має інструментальний засіб – Internet-сайт, когнітивний комп'ютерної технології Web прийняття рішень щодо раціональної організації автотранспорту (будь якого пересування пасажирів або вантажу у просторово-часовому просторі перевізних процесів) з урахуванням стану дорожнього середовища. На відміну від існуючого стану логістики, основних законів, правил та принципів розвитку IT-індустрії передбачається інтерактивний моніторинг як автомобілю, так і учасників перевізного процесу, саме дороги.

Тому особлива увага у даній статті була спрямована на розгляд апаратної віртуалізації, Web-рішення та інтерфейсу користувача Neuronet автотранферу, надані поради щодо практичного застосування розробленого експериментального зразка інформаційно-комунікаційного комплексу.

### Список літератури:

- [1] Риз Дж. Облачные вычисления: Пер. с англ. – СПб.: БХВ-Петербург, 2011. – 288 с.

*Надійшла до редколегії 21.02.2024*

УДК 004.056

DOI 10.30837/bi.2024.1(100).07

Г. А. Плехова<sup>1</sup>, С. М. Неронов<sup>2</sup>, М. В. Костікова<sup>3</sup>, С. О. Кашкевич<sup>4</sup><sup>1</sup>ХНАДУ, м. Харків, Україна, plehovaanna11@gmail.com, ORCID iD: 0000-0002-6912-6520<sup>2</sup>ХНАДУ, м. Харків, Україна, sernikner@gmail.com, ORCID iD: 0000-0003-2381-1271<sup>3</sup>ХНАДУ, м. Харків, Україна, kmv\_topaz@ukr.net, ORCID iD: 0000-0001-5197-7389<sup>4</sup>НАУ, м. Харків, Україна, svitlana.kashkevych@npp.nau.edu.ua, ORCID iD: 0000-0002-4448-3839

## УДОСКОНАЛЕННЯ МОДЕЛІ БЕЗПЕЧНОЇ МАРШРУТИЗАЦІЇ В ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖАХ

Наразі розгортання таких мережних архітектур, як програмно-конфігуровані мережі (Software-Defined Networking, SDN), стикається з новими загрозами кібербезпеці, які вимагають розробку та дослідження нових спеціалізованих рішень щодо підвищення рівня мережної безпеки. Незважаючи на високу відкритість і можливості програмованості, архітектура SDN замінює традиційну мережу, проте збільшує кількість потенційних мережних атак, що призводить до нових проблем безпеки.

Зростаючий інтерес до SDN та широкому розгортанню програмно-конфігурованих мереж різних типів дозволяють виявляти їхні недоліки в процесі боротьби із загрозами кібербезпеці. Очевидно, що питання безпеки тісно пов'язані з характеристиками самих SDN мереж. Крім того, проблеми безпеки в SDN можна розділити на основі трьох рівнів: площини даних, площини управління та площини застосунків.

Водночас серед об'єктів атак можуть бути пристрої різних рівнів SDN. Отже, відповідно до багаторівневої архітектури SDN можна класифікувати загрози безпеці на рівнях передачі даних, управління та застосунків. Зі свого боку, площина даних складається з комутаторів та інших мережних пристроїв і головним чином відповідає за обробку даних, їх пересилання, відкидання, а також збір статистики. Функціонування площини даних відбувається на основі правил потоків, що надаються контролером мережі. Тоді як основними причинами проблем безпеки є власне архітектура SDN, зовнішні шкідливі атаки, недостатність контролю доступу та засобів шифрування.

На сьогоднішній день важливе місце у комплексі засобів підвищення мережної безпеки, у тому числі мереж SDN, відводиться протоколам маршрутизації, які потребують системної та скоординованої взаємодії одночасно множини мережних елементів – SDN-комутаторів, і контролерів мережі під час формування (розрахунку) шляхів і правил потоків, вздовж яких має забезпечуватися необхідний рівень безпеки за обраними показниками або критеріям.

В роботі проведена аналіз того як модифікувати маршрутні метрики таким чином, щоб отримувана модель набула властивостей безпечної QoS-маршрутизації. Показано що удосконалення моделі та вибір маршруту потрібно обирати з урахуванням базових метрик критичності вразливостей, і пропускної здатності каналів зв'язку, що складають цей маршрут.

МОДЕЛЬ, МАРШРУТИЗАЦІЯ, ВРАЗЛИВІСТЬ, БЕЗПЕКА, МЕТРИКА, ДЕЦЕНТРАЛІЗОВАНІ СХОВИЩА ДАНИХ, СТИСНЕННЯ ЗОБРАЖЕНЬ

**G. A. Pliekhova, S. M. Neronov, M. V. Kostikova, S. O. Kashkevich. Improvement of the secure routing model in software-configured networks.** Currently, the deployment of such network architectures as Software-Defined Networking (SDN) is facing new cyber security threats that require the development and research of new specialized solutions to increase the level of network security. Despite its high openness and programmability, the SDN architecture replaces the traditional network, but it increases the number of potential network attacks, which leads to new security problems.

The growing interest in SDN and the widespread deployment of software-configured networks of various types allow identifying their shortcomings in the process of combating cyber security threats. Obviously, security issues are closely related to the characteristics of SDN networks themselves. Furthermore, security issues in SDN can be divided based on three layers: data plane, control plane, and application plane.

At the same time, devices of different SDN levels can be among the objects of attacks. Therefore, according to the multilayer architecture of SDN, security threats can be classified at the data transmission, management and application layers. For its part, the data plane consists of switches and other network devices and is mainly responsible for data processing, forwarding, discarding, and collecting statistics. The data plane functions on the basis of flow rules provided by the network controller. While the main causes of security problems are the SDN architecture itself, external malicious attacks, insufficient access control and encryption tools.

Today, an important place in the complex of means of increasing network security, including SDN networks, is given to routing protocols, which require the systematic and coordinated interaction of a number of network elements at the same time – SDN switches and network controllers during the formation (calculation) of paths and flow rules, along which the required level of security must be ensured according to selected indicators or criteria.

The paper analyzes how to modify route metrics in such a way that the resulting model acquires the properties of secure QoS routing. It is shown that the improvement of the model and the choice of the route should be chosen taking into account the basic metrics of the criticality of vulnerabilities and the bandwidth of the communication channels that make up this route.

MODEL, ROUTING, VULNERABILITY, SECURITY, METRICS, DECENTRALIZED DATA STORAGE.

### Вступ

Головною метою використання NFV (Network function virtualization) є необхідність використовувати стандартні технології віртуалізації для консолідації

апаратного забезпечення та віртуалізації мережних функцій у блоки, які можна об'єднувати для створення наскрізних комунікаційних послуг. Це може бути реалізовано для будь-якої функції площини

управління або площини даних у середовищі як про-  
водових, так і безпроводових мереж. Національна  
база даних вразливостей та загальна система оцінки  
вразливостей.

Розглянемо більш детально сферу технічних враз-  
ливостей. Видатним ресурсом є національна база  
даних про вразливості NIST – National Vulnerability  
Database (NVD), і відповідна загальна система оцінки  
вразливостей – Common Vulnerability Scoring System  
(CVSS), описана в NISTIR 7946, Посібник із впро-  
вадження CVSS. NVD – це вичерпний список відо-  
мих технічних вразливостей систем, апаратного та  
програмного забезпечення. CVSS забезпечує відкри-  
ту структуру для передачі характеристик вразливос-  
тей. CVSS визначає вразливість як помилку, недолік,  
слабкість або відкритість програми, системного при-  
строю чи сервісу, що може призвести до збою конфі-  
денційності, цілісності чи доступності.

Отже, модель CVSS намагається забезпечити по-  
вторювані та точні вимірювання, одночасно дозволя-  
ючи користувачам переглядати базові характери-  
стики вразливості, які використовуються для створення  
числових оцінок. CVSS надає загальну систему ви-  
мірювання для галузей промисловості, організацій  
та урядів, які вимагають точних і послідовних оцінок  
використання вразливостей та їхнього впливу.

Розуміння CVSS дозволяє оцінити широкий  
спектр вразливостей, які впливають на системи. Крім  
того, систематизована схема для оцінки вразливостей  
у CVSS є корисною для розробки подібного систем-  
ного підходу до інших вразливостей, таких як ті, що  
пов'язані з організаційними питаннями, політикою  
та процедурами, а також фізичною інфраструктурою.  
На сьогоднішній день CVSS широко прийнятий і ви-  
користовуваний підхід. Наприклад, використання  
CVSS рекомендується для кількісного розрахунку  
рівня вразливості мережного обладнання.

Кожен запис NVD містить наступну інформацію:

- унікальний словниковий ідентифікатор враз-  
ливостей і ризиків – Common Vulnerabilities and  
Exposure (CVE);
- опис вразливості;
- посилання на веб-сайти та інші посилання з  
інформацією, пов'язаною з вразливістю;
- метрики CVSS.

### 1. Мета та постановка задачі

Метою роботи є розв'язання задачі, яка спрямо-  
вана на розгляд категорії вразливостей, використан-  
ня національної бази даних вразливостей та загаль-  
ної системи оцінки вразливостей, метрики загальної  
системи оцінки вразливостей виходячи з проведеного  
аналізу стандарту CVSS щодо кількісного розрахун-  
ку рівня вразливості мережного обладнання. Також  
приділено увагу використанню існуючої методики  
розрахунку метрик маршрутизації на основі базових  
метрик критичності вразливостей.

Таким чином бажаючи удосконалити математич-  
ну модель безпечної маршрутизації з урахуванням

базових метрик критичності вразливостей в роботі [1]  
розроблено та досліджено модель безпечної марш-  
рутизації з балансуванням навантаження в мережах  
на основі SD-WAN. Технологічне завдання безпеч-  
ної маршрутизації з балансуванням навантаження  
було сформульовано у формі оптимізаційної задачі з  
квадратичним критерієм оптимальності. Така форма  
критерію дозволяє збалансувати частки потоків, що  
передаються в мережі.

Перспективним в моделі [1] представляється те,  
що використовується комбінована метрика під час  
розрахунку мультишляху:

$$f_{i,j}^{\text{комб}} = f_{i,j}^{\text{OSPF}} + f_{i,j}^{\text{SEC}}, \quad (1)$$

де  $f_{i,j}^{\text{OSPF}}$  – метрика, обрана за аналогією з протоколом  
OSPF, яка відповідає за те, щоб розраховуваний марш-  
рут містив у собі найбільш продуктивні канали зв'язку

$$f_{i,j}^{\text{OSPF}} = \frac{10^8}{c_{i,j}}, \quad (2)$$

де  $c_{i,j}$  – пропускна здатність відповідного каналу  
зв'язку;  $f_{i,j}^{\text{SEC}}$  – метрика, заснована на параметрі ме-  
режної безпеки – ймовірності компрометації каналу  
зв'язку

$$f_{i,j}^{\text{SEC}} = \frac{10^8}{R} p_{i,j}. \quad (3)$$

При цьому  $R$  є співвідношенням між ваговими  
коефіцієнтами метрик продуктивності та мережної  
безпеки:

$$R = \frac{w^{\text{OSPF}}}{w^{\text{SEC}}}, \quad w^{\text{SEC}} = \frac{w^{\text{OSPF}}}{R} = \frac{10^8}{R}. \quad (4)$$

Результати моделювання показали, що досліджу-  
ване навантаження на канал зв'язку (а саме частка  
потoku, що передається, оскільки використовувалась  
стратегія багатошляхової маршрутизації) зменшуєть-  
ся із збільшенням ймовірності компрометації каналу.  
Аналіз результатів дослідження також виявив зна-  
чення співвідношення  $R$  метрик мережної безпеки  
та продуктивності, коли модель найбільш чутлива  
до погіршення ймовірності компрометації каналів  
зв'язку та мережі загалом.

### 2. Аналіз публікацій

Представлена в [1] модель безпечної маршрутиза-  
ції з балансуванням навантаження з адитивною ком-  
бінованою метрикою враховує продуктивність і без-  
пеку мережі, дозволяє ефективніше використовувати  
наявні мережні ресурси, але також враховує ймовір-  
ність компрометації каналів зв'язку під час прийня-  
ття маршрутних рішень. Отже, в роботі [2] розроблено  
проактивне рішення щодо забезпечення мережної  
безпеки, а саме метод безпечної маршрутизації пові-  
домлень шляхами, що перетинаються. Новизна цього  
методу полягає в тому, що допускається використан-  
ня шляхів, які перетинаються. Вони також станов-  
лять основу композитних шляхів і містять мереж-  
ні фрагменти з послідовним та (або) паралельним  
з'єднанням каналів зв'язку. Крім того, основу мето-  
ду складає оптимізаційний процес вибору множини

композитних шляхів і балансування ними частин повідомлення, що передається, із забезпеченням допустимих значень його ймовірності компрометації. Проведені числові дослідження довели працездатність та ефективність запропонованого методу, тобто його використання в межах наведених розрахункових прикладів дозволило знизити ймовірність компрометації повідомлень, що передаються, в середньому від 5 – 10% до 25 – 50% з огляду на можливості використання композитних шляхів [2].

Другий метод, розроблений і запропонований в [2], пов'язаний з необхідністю оперативного розрахунку та зміни множини шляхів, що використовуються для передачі частин конфіденційних повідомлень, у разі зміни стану мережі, викликаній порушенням рівня безпеки. Тому підходи та механізми щодо швидкої перемаршрутизації з локальним чи глобальним захистом елементів мережі можуть розглядатися як реалізація реактивного підходу щодо забезпечення безпечної маршрутизації [2]. Отже, в [2] було розроблено та досліджено метод безпечної швидкої перемаршрутизації повідомлень у мережі, який орієнтує на реалізацію як проактивної, так і реактивної безпечної маршрутизації конфіденційних повідомлень.

Сутність методу безпечної швидкої перемаршрутизації (Secure Fast ReRoute, S-FRR) полягає в тому, що порушення вимог мережної безпеки, викликаного підвищенням ймовірності компрометації одного або множини композитних шляхів, тобто основного мультишляху, призводить до використання попередньо розрахованою множиною резервних композитних шляхів для багатошляхової передачі частин конфіденційного повідомлення із забезпеченням заданих значень ймовірності його компрометації. Крім того, метод передбачає реалізацію захисту або основного мультишляху загалом, або одного чи декількох задалегідь заданих композитних шляхів, що містить основний мультишлях [2].

Застосування методу S-FRR дозволило в режимі реального часу забезпечити задані значення показника мережної безпеки, а саме ймовірності компрометації повідомлень, що передаються, навіть в умовах динамічної зміни стану мережі (ймовірності компрометації каналів і шляхів) на підставі розрахунку й оперативного переходу на використання резервних композитних шляхів за умови багатошляхової передачі частин конфіденційного повідомлення [2].

Підсумовуючи, можна зробити висновок, що розроблені в [2] методи безпечної маршрутизації можуть рекомендуватися до використання як основа нових мережних протоколів безпечної маршрутизації та безпечної швидкої перемаршрутизації для багатошляхової передачі частин конфіденційного повідомлення із заданими вимогами щодо граничної ймовірності його компрометації в мережі.

Алгоритми маршрутизації з урахуванням параметрів ризику інформаційної безпеки розглянуті у роботах [3] пропонуються під час вибору маршруту в ІКМ враховувати ризики інформаційної безпеки.

Це забезпечується шляхом відповідного формування маршрутних метрик, коли в них сумісно з показниками якості обслуговування враховується показник ризику інформаційної безпеки елементів мережі. Цей підхід дозволяє динамічно вибрати найбільш безпечний маршрут потоків, що передаються в умовах активних атак і в разі пасивного аналізу ризиків у системі маршрутизації.

Так, у [3] запропоновано використовувати параметр ризику інформаційної безпеки у формулі розрахунку метрики протоколу EIGRP для маршрутизації трафіка найбільш безпечними шляхами в мережі. Метод пропонує розраховувати ризик на основі двох параметрів: ризик, який розраховується на основі стандарту NIST CVSS, і ризик, який розраховується на основі формули для ступеня вразливості вузла з теорії живучості інформаційних систем. Це дозволяє враховувати інформаційну безпеку маршрутизованих пакетів і структурну цілісність мережі. Також у [3] розроблено модифікований алгоритм балансування навантаження між шляхами, який дозволяє розвантажити найефективніший вузол маршрутизації під час атаки «відмова в обслуговуванні» (DoS). Результати дослідження довели, що запропонований підхід ефективний під час запобігання порушенню інформаційної безпеки маршрутизованих пакетів і забезпечення безпеки найефективніших вузлів маршрутизації в мережі, які дозволяють ефективно маршрутизувати довірений трафік, коли мережа піддається DoS-атаці або бракує критичних системних ресурсів.

Потокові моделі маршрутизації на основі базових метрик критичності вразливостей у роботах [4, 5] пропонуються потокові моделі маршрутизації з урахуванням ризиків інформаційної безпеки за допомогою базових метрик критичності вразливостей. Основу моделі складають умови реалізації одно- та багатошляхової маршрутизації, збереження потоку та запобігання перевантаженню каналів зв'язку мережі, а задача безпечної маршрутизації також сформульована як оптимізаційна. У моделі [4] для розрахунку маршрутних метрик використовуються вирази, які характеризують ризик інформаційної безпеки в каналах зв'язку мережі та відповідно до рекомендацій NIST враховують збитки від порушення конфіденційності та цілісності інформації, доступності мережного ресурсу у випадку використання наявних вразливостей; показники складності використання вразливостей на вузлах мережі та отримання доступу до мережних елементів та мережі загалом внаслідок використання зазначених вразливостей. Запропонований авторами підхід до формування маршрутних метрик може бути використаний під час комплексного врахування в процесі розв'язання задач маршрутизації як показників мережної безпеки, так і показників якості обслуговування.

Безпечна маршрутизація на основі SPREAD включає відомий підхід щодо застосування механізмів SPREAD (Secure Protocol for Reliable dAta Delivery) та H-SPREAD (Hybrid Secure Protocol for

Reliable dAta Delivery) та їх розвиток запропоновано у роботах [2, 3], що націлені на посилення безпечної передачі та маршрутизації секретних повідомлень у MANET (рис. 1). Основною ідеєю є те, що конфіденційне повідомлення розділяється на декілька фрагментів – частин, які потім передаються від відправника до отримувача множиною шляхів, що не перетинаються. Таким чином, що навіть якщо деяка кількість фрагментів повідомлення буде

скомпрометовано, то секретне повідомлення загалом залишиться нескомпрометованим [2, 6]. Отже, в [2, 6] було запропоновано загальну математичну модель для створення та реконструкції фрагментів секретного повідомлення, оптимальний розподіл його частин за декількома шляхами з урахуванням параметра мережної безпеки, а також підходи щодо розрахунку мультишляху в безпроводових мережах MANET.

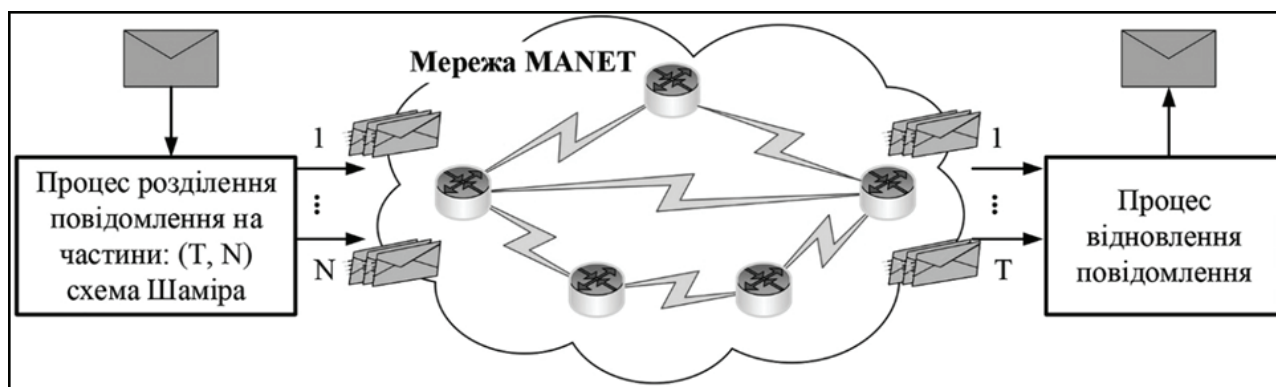


Рис. 1. Загальна архітектура роботи механізму SPREAD

Відмічається, що у порівнянні з проводовими мережами забезпечення безпеки в MANET пов'язано з виявленням і запобіганням множини наявних вразливостей та атак [2]. Перш за все, радіоканали більш сприйнятливі до атак пасивного прослуховування й активного втручання в сигнали та здійснення завад. Крім того, більшість протоколів маршрутизації MANET припускають, що взаємодія між вузлами для здійснення передачі пакетів є довірчою. Таким чином, передача даних стає більш вразливою щодо несанкціонованого доступу, підміни даних, а також атак типу «відмова в обслуговуванні» (Denial of Service, DoS). Також відсутність фіксованої інфраструктури та централізованого управління ускладнює застосування більшості традиційних рішень щодо забезпечення мережної безпеки.

Численні дослідження застосування механізму SPREAD та його модифікацій дозволяє знизити ймовірність компрометації секретного повідомлення, що передається, завдяки ускладненню завдання зловмисника: йому необхідно скомпрометувати не один маршрут, яким передається нерозділене повідомлення, а всі шляхи, якими передаються його фрагменти [2, 6]. Згідно з [2] під компрометацією повідомлення розуміється подія, пов'язана з несанкційним доступом до його вмісту.

Для забезпечення безпечної маршрутизації повідомлення в мережі під час застосування механізму SPREAD необхідно вирішити такі завдання [3, 6]:

1. Розрахувати множини маршрутів, що не перетинаються, між заданими вузлами «відправник» і «отримувач».

2. Розділити конфіденційне повідомлення, що передається, на множину частин відповідно до обраної схеми Шаміра.

3. Розподілити множини фрагментів повідомлення між множиною маршрутів, визначених на першому кроці.

Отже, ймовірність компрометації маршруту (шляху) багато в чому залежить від кількості вузлів і каналів зв'язку, що його складають, і від параметрів їх безпеки, тобто кожен елемент мережі (вузол, канал) шляху може бути скомпрометований з певною ймовірністю [2, 6]. У загальному випадку шляхи для передачі частин розділеного відповідно до схеми Шаміра [2, 6] повідомлення, можуть мати різні рівні безпеки, що визначаються ймовірністю компрометації цих шляхів [2]. Крім того, подібні рішення орієнтовані на використання лише шляхів, що не перетинаються. Це негативно впливає на ефективність використання доступного мережного ресурсу.

Проактивні та реактивні методи безпечної маршрутизації повідомлень шляхами, що перетинаються також розглядаються у [2] запропоновано використання особливого класу шляхів, що перетинаються, які зі свого боку можуть формувати композитні шляхи та містять відповідно мережні фрагменти з послідовним та (або) паралельним з'єднанням каналів зв'язку мережі. Такий підхід щодо вибору шляхів орієнтує на зниження ймовірності компрометації конфіденційних (секретних) повідомлень, які передаються в комунікаційній мережі.

Отже, в роботі [2] розроблено проактивне рішення щодо забезпечення мережної безпеки, а саме метод безпечної маршрутизації повідомлень шляхами, що перетинаються. Новизна цього методу полягає в тому, що допускається використання шляхів, які

перетинаються. Вони також становлять основу композитних шляхів і містять мережні фрагменти з послідовним та (або) паралельним з'єднанням каналів зв'язку. Крім того, основу методу складає оптимізаційний процес вибору множини композитних шляхів і балансування ними частин повідомлення, що передається, із забезпеченням допустимих значень його ймовірності компрометації. Проведені числові дослідження довели працездатність та ефективність запропонованого методу, тобто його використання в межах наведених розрахункових прикладів дозволило знизити ймовірність компрометації повідомлень, що передаються, в середньому від 5 – 10% до 25 – 50% з огляду на можливість використання композитних шляхів [2].

Другий метод, розроблений і запропонований в [2], пов'язаний з необхідністю оперативного розрахунку та зміни множини шляхів, що використовуються для передачі частин конфіденційних повідомлень, у разі зміни стану мережі, викликаній порушенням рівня безпеки. Тому підходи та механізми щодо швидкої перемаршрутизації з локальним чи глобальним захистом елементів мережі можуть розглядатися як реалізація реактивного підходу щодо забезпечення безпечної маршрутизації [2]. Отже, в [2] було розроблено та досліджено метод безпечної швидкої перемаршрутизації повідомлень у мережі, який орієнтує на реалізацію як проактивної, так і реактивної безпечної маршрутизації конфіденційних повідомлень.

Сутність методу безпечної швидкої перемаршрутизації (Secure Fast ReRoute, S-FRR) полягає в тому, що порушення вимог мережної безпеки, викликаного підвищенням ймовірності компрометації одного або множини композитних шляхів, тобто основного мультишляху, призводить до використання попередньо розрахованою множиною резервних композитних шляхів для багатошляхової передачі частин конфіденційного повідомлення із забезпеченням заданих значень ймовірності його компрометації. Крім того, метод передбачає реалізацію захисту або основного мультишляху загалом, або одного чи декількох задалегідь заданих композитних шляхів, що містить основний мультишлях [2].

Застосування методу S-FRR дозволило в режимі реального часу забезпечити задані значення показника мережної безпеки, а саме ймовірності компрометації повідомлень, що передаються, навіть в умовах динамічної зміни стану мережі (ймовірності компрометації каналів і шляхів) на підставі розрахунку й оперативного переходу на використання резервних композитних шляхів за умови багатошляхової передачі частин конфіденційного повідомлення [2].

Підсумовуючи, можна зробити висновок, що розроблені в [2] методи безпечної маршрутизації можуть рекомендуватися до використання як основа нових мережних протоколів безпечної маршрутизації та безпечної швидкої перемаршрутизації для багатошляхової передачі частин конфіденційного повідомлення із заданими вимогами щодо граничної ймовірності його компрометації в мережі.

Алгоритми маршрутизації з урахуванням параметрів ризику інформаційної безпеки розглянуті у роботах [3] пропонується під час вибору маршруту в ІКМ враховувати ризики інформаційної безпеки. Це забезпечується шляхом відповідного формування маршрутних метрик, коли в них сумісно з показниками якості обслуговування враховується показник ризику інформаційної безпеки елементів мережі. Цей підхід дозволяє динамічно вибрати найбільш безпечний маршрут потоків, що передаються в умовах активних атак і в разі пасивного аналізу ризиків у системі маршрутизації.

Запропоновано використовувати параметр ризику інформаційної безпеки у формулі розрахунку метрики протоколу EIGRP для маршрутизації трафіка найбільш безпечними шляхами в мережі. Метод пропонує розраховувати ризик на основі двох параметрів: ризик, який розраховується на основі стандарту NIST CVSS, і ризик, який розраховується на основі формули для ступеня вразливості вузла з теорії живучості інформаційних систем. Це дозволяє враховувати інформаційну безпеку маршрутизованих пакетів і структурну цілісність мережі. Розроблено модифікований алгоритм балансування навантаження між шляхами, який дозволяє розвантажити найефективніший вузол маршрутизації під час атаки «відмова в обслуговуванні» (DoS). Результати дослідження довели, що запропонований підхід ефективний під час запобігання порушенню інформаційної безпеки маршрутизованих пакетів і забезпечення безпеки найефективніших вузлів маршрутизації в мережі, які дозволяють ефективно маршрутизувати довірений трафік, коли мережа піддається DoS-атаці або бракує критичних системних ресурсів.

Потокові моделі маршрутизації на основі базових метрик критичності вразливостей у роботах [4, 5] пропонуються потокові моделі маршрутизації з урахуванням ризиків інформаційної безпеки за допомогою базових метрик критичності вразливостей. Основу моделі складають умови реалізації одно- та багатошляхової маршрутизації, збереження потоку та запобігання перевантаженню каналів зв'язку мережі, а задача безпечної маршрутизації також сформульована як оптимізаційна. У моделі [4] для розрахунку маршрутних метрик використовуються вирази, які характеризують ризик інформаційної безпеки в каналах зв'язку мережі та відповідно до рекомендацій NIST враховують збитки від порушення конфіденційності та цілісності інформації, доступності мережного ресурсу у випадку використання наявних вразливостей; показники складності використання вразливостей на вузлах мережі та отримання доступу до мережних елементів та мережі загалом внаслідок використання зазначених вразливостей. Запропонований авторами підхід до формування маршрутних метрик може бути використаний під час комплексного врахування в процесі розв'язання задач маршрутизації як показників мережної безпеки, так і показників якості обслуговування.

### 3. Виклад основного матеріалу

Таким чином пропонуємо використання методи- ки розрахунку метрик маршрутизації на основі базо- вих метрик критичності вразливостей відповідно до результатів, отриманих у роботах [4, 5, 7–9], вияв- лено, що одним з ефективних засобів забезпечення захисту мереж є попереднє оцінювання ризику ін- формційної безпеки (РІБ). Процес оцінювання РІБ спрямований на запобігання використанню відомих вразливостей, потенційно наявних у мережі, що за- хищається [9]. Зі свого боку РІБ може розраховува- тися за допомогою використання зазначених у ре- комендації NIST CVSS v3 [8, 10] метрик критичності вразливостей: базових, часових і метрик середовища користувачів.

У межах запропонованого у [4] та удосконалено- го у [9] рішення для розрахунку вагових коефіцієн- тів  $w_{i,j}$ , що використовувались у процесі отримання маршрутних рішень, обрано базові метрики, які ха- рактеризують наявні вразливості елементів мережі та дозволяють оцінити ризик інформаційної безпеки мережі загалом, а не для окремих випадків компро- метації мережних елементів, на відміну від часових метрик і метрик середовища.

В роботах [4, 9] введено наступні позначення:

- $U = \{U_i^q; q = \overline{1, Q}, i = \overline{1, m}\}$  – множина вразливос- тей, виявлених на вузлах (маршрутизаторах) мережі;
- $U_i^q$  –  $q$ -та вразливість на  $i$ -му вузлі мережі;
- $U_i^* \subset U$  – множина вразливостей на  $i$ -му вузлі мережі;
- $BS_i^q$  – показник критичності  $q$ -ї вразливості на  $i$ -му вузлі мережі, що розраховується за допо- могою базових метрик системи оцінювання вразливос- тей відповідно до NIST CVSS v3 [8, 11] та характери- зує умовні збитки від використання вразливості  $U_i^q$  зловмисником;
- $P_i^q$  – ймовірність використання  $q$ -ї вразли- вості зловмисником на  $i$ -му вузлі мережі, що за фі- зичним змістом є ймовірністю компрометації.

У [12] зазначено наступний вираз для розрахунку

ризикі інформаційної безпеки від використання на- явних вразливостей на  $i$ -му вузлі мережі:

$$R^i = \sum_{U_i^q \in U_i^*} BS_i^q \cdot P_i^q. \quad (5)$$

Згідно з рекомендацією NIST [8, 10], збитки щодо базових метрик вразливостей на вузлах мережі розра- ховуються як

$$BS_i^q = (0,6 \cdot \text{Imp}_i^q + 0,4 \cdot \text{Ex}_i^q - 1,5) \cdot f(\text{Imp}_i^q), \quad (6)$$

де  $\text{Imp}_i^q$  – потенційний збиток від використання  $q$ -ї вразливості зловмисником на  $i$ -му вузлі мережі;  $\text{Ex}_i^q$  – складність використання  $q$ -ї вразливості зло- вмисником на  $i$ -му вузлі мережі;  $f(\text{Imp}_i^q)$  – функція від потенційного збитку в разі використання  $q$ -ї враз- ливості зловмисником на  $i$ -му вузлі мережі.

Тоді як потенційний збиток від використання вразливості можна отримати таким чином [7, 9]:

$$\text{Imp}_i^q = 10,41 \left[ 1 - (1 - \text{Conf}_i^q) \cdot (1 - \text{Int}_i^q) \cdot (1 - v_i^q) \right], \quad (7)$$

де  $\text{Conf}_i^q$  – збитки від порушення конфіденційності інформації, що передається мережею та не може бути отримана неавторизованим, наприклад, зовнішнім, користувачем (зловмисником);  $\text{Int}_i^q$  – збитки від порушення цілісності мережі, що характеризуються модифікацією, зміною та руйнуванням інформації неавторизованим користувачем (зловмисником);  $\text{Av}_i^q$  – збитки від порушення доступності мережного ресурсу у випадку використання  $q$ -ї вразливості на  $i$ -му вузлі мережі.

Три метрики базової групи  $\text{Conf}_i^q$ ,  $\text{Int}_i^q$  та  $\text{Av}_i^q$  ви- значають можливі наслідки використання зловмис- ником  $q$ -ї вразливості на  $i$ -му вузлі мережі.

Відповідно до [8, 10], у кожній із цих метрик збит- ки від використання вразливості можуть бути:

- відсутніми із значенням 0;
- частковими із значенням 0,275;
- повними із значенням 0,66.

Значення метрик  $\text{Conf}_i^q$ ,  $\text{Int}_i^q$  та  $\text{Av}_i^q$  наведено в табл. 1.

Таблиця 1

Значення показників для розрахунку базових метрик вразливостей елементів мережі [4, 9]

Значення	Опис	Числова характеристика
Збиток конфіденційності $\text{Conf}_i^q$		
Відсутній (В)	Можливість порушення конфіденційності інформації відсутня	0,000
Частковий (Ч)	Існує значне, однак обмежене розголошення конфіденційної інформації	0,275
Повний (П)	Існує повне розкриття конфіденційної інформації	0,660
Збиток цілісності $\text{Int}_i^q$		
Відсутній (В)	Можливість порушення цілісності інформації відсутня	0,000
Частковий (Ч)	Існує можливість часткової модифікації даних або системних файлів	0,275

Значення	Опис	Числова характеристика
Повний (П)	Існує можливість модифікації будь-яких даних вузла	0,660
Збиток доступності $Av_i^q$		
Відсутній (В)	Можливість порушення доступності ресурсу відсутня	0,000
Частковий (Ч)	Існує можливість зниження продуктивності або виведення з ладу деяких функцій вузла	0,275
Повний (П)	Існує можливість повного виведення вузла з ладу	0,660

Крім того, складність використання вразливості отримується як [6, 8]:

$$Ex_i^q = 20 \cdot Ac_i^q \cdot Au_i^q \cdot AcV_i^q, \quad (8)$$

де  $Ac_i^q$  – показник системи оцінки вразливості, що характеризує складність отримання доступу (вектор доступу);  $Au_i^q$  – показник системи оцінки вразливості, що відповідає за вимоги до автентифікації;  $AcV_i^q$  – показник системи оцінки вразливості, який відображає спосіб використання  $q$ -ї вразливості на  $i$ -му вузлі

мережі, що за фізичним змістом характеризується «віддаленістю» зловмисника, тобто кількістю пристроїв та/або обмежень доступу, через які зловмисник може досягнути  $i$ -го вузла мережі для здійснення атаки.

Зазначені показники є базовими метриками [8], які характеризують загальну складність реалізації атаки у використанні тієї чи іншої вразливості на  $i$ -му вузлі мережі (табл. 2).

Таблиця 2

Значення показників системи оцінки вразливості, які характеризують складність використання вразливостей [4, 9]

Значення	Опис	Числова характеристика
Вектор доступу $Ac_i^q$		
Потрібен локальний доступ (Л)	Зловмисникові потрібен безпосередній фізичний доступ до об'єкта, на якому розташована вразливість	0,395
Можливий доступ із суміжної мережі (СММ)	Зловмисникові потрібен доступ у межах однієї локальної мережі (одного ширококомовного домену) до вразливого об'єкта	0,646
Можливий доступ із будь-якої мережі (М)	Зловмисник може використовувати вразливість віддалено з будь-якої ділянки мережі, зокрема з допомогою інтернету	1,000
Вимоги до автентифікації $Au_i^q$		
Множинна (М)	Зловмисник повинен зробити більше ніж одну процедуру автентифікації для експлуатації вразливості вузла	0,450
Одинична (О)	Зловмиснику досить один раз автентифікуватися для експлуатації вразливості вузла	0,560
Відсутня (В)	Зловмисникові не потрібно проходити процедуру автентифікації для експлуатації вразливості вузла	0,704
Збиток доступності $Av_i^q$		
Складна (Ск)	Існує низка жорстких обмежень доступу до вузла. Наприклад, експлуатація вразливості вузла можлива тільки в дуже короткий проміжок часу або вимагає застосування соціальної інженерії, коли зловмисника може бути опізнано	0,350
Середня (Ср)	Існують деякі обмеження доступу до вузла. Наприклад, підключення до вразливого пристрою можливе тільки з певних вузлів або вразливий пристрій має функціонувати з нестандартними налаштуваннями	0,610
Легка (Л)	Немає особливих умов доступу до вразливості вузла. Наприклад, коли система доступна багатьом користувачам одночасно або коли вразлива конфігурація працює на множині вузлів мережі	0,710

Згідно з [4, 9] функція від потенційного збитку  $f(\text{Imp}_i^q)$  приймає значення 0, якщо збитку немає, тобто  $f(\text{Imp}_i^q) = 0$ . Тоді розглядатиметься випадок, коли потенційний збиток наявний ( $\text{Imp}_i^q \neq 0$ ), і в подальших розрахунках використовувалось

$$f(\text{Imp}_i^q) = 1,176 [3, 7].$$

Отже, для розрахунку ризику інформаційної безпеки за умови компрометації каналу зв'язку  $E_{i,j} \in E$ , що виходить з  $i$ -го вузла, використовується наступний вираз [9, 12]:



$$R_{i,j} = w_{i,j} \cdot \ln \sum_{U_i^q \in U_i^*} e^{BS_i^q}, \quad (9)$$

де  $w_{i,j}$  – вагові коефіцієнти (вага компрометації) оцінювання ризику, створюваного використанням вразливостей на  $i$ -му вузлі мережі. Фактично коефіцієнти  $w_{i,j}$  кількісно характеризують потенційний збиток у разі застосування наявних на  $i$ -му вузлі мережі вразливостей.

Якщо компрометація каналу зв'язку  $E_{i,j} \in E$  відбувається тільки через використання вразливостей на  $i$ -му вузлі, то ризики інформаційної безпеки вузла та каналу зв'язку тотожно рівні [9]:

$$\sum_{U_i^q \in U_i^*} BS_i^q \cdot P_i^q = w_{i,j} \cdot \ln \sum_{U_i^q \in U_i^*} e^{BS_i^q}. \quad (10)$$

Більше того, розрахунок вагових коефіцієнтів  $w_{i,j}$  ґрунтується на тому, що компрометація каналу зв'язку  $E_{i,j} \in E$  відбуватиметься внаслідок компрометації  $i$ -го вузла мережі.

Далі відповідно до виразів (9) – (10) значення вагових коефіцієнтів  $w_{i,j}$ , що можуть у подальшому слугувати метриками маршрутизації, обчислюються наступним чином [4, 9]:

$$w_{i,j} = \frac{\sum_{U_i^q \in U_i^*} BS_i^q \cdot P_i^q}{\ln \sum_{U_i^q \in U_i^*} e^{BS_i^q}}. \quad (11)$$

Таким чином можливо удосконалення існуючої моделі безпечної маршрутизації з урахуванням базових метрик критичності вразливостей.

### Висновки

В роботі проведена аналіз того як модифікувати маршрутні метрики таким чином, щоб отримувана модель набула властивостей безпечної QoS-маршрутизації. Показано що удосконалення моделі та вибір маршруту потрібно обирати з урахуванням базових метрик критичності вразливостей, і пропускної здатності каналів зв'язку, що складають цей маршрут. В подальших роботах планується провести удосконалення математичної моделі безпечної маршрутизації з урахуванням базових метрик критичності вразливостей та розв'язати технологічну задачу, яку буде сформульовано як оптимізаційну з квадратичною цільовою функцією, коли оптимальний маршрут обирався відповідно до комбінованої метрики на основі базових метрик критичності вразливостей і пропускної здатності каналів зв'язку, що складають цей маршрут.

### Список літератури:

[1] Yeremenko O., Persikov M., Lemeshko V., Altaki B. Research and development of the secure routing flow-based model with load balancing. Проблеми телекомунікацій. 2021.

№ 2(29). С. 3–14. URL: [https://pt.nure.ua/wp-content/uploads/2021/12/212\\_yeremenko\\_secure.pdf](https://pt.nure.ua/wp-content/uploads/2021/12/212_yeremenko_secure.pdf).

- [2] Лемешко О. В., Єременко О. С., Невзорова О. С. Потоківі моделі та методи маршрутизації в інфокомунікаційних мережах: відмовостійкість, безпека, масштабованість. Харків: ХНУРЕ, 2020. 308 с. DOI: <https://doi.org/10.30837/978-966-659-282-1>.
- [3] Lou W., Liu W., Fang Y. SPREAD: Enhancing Data Confidentiality in Mobile Ad Hoc Networks. INFOCOM 2004: Proceedings of the Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies. Hong Kong, China, 7–11 March, 2004. IEEE, 2004. P. 2404–2413. DOI: <https://doi.org/10.1109/INFCOM.2004.1354662>.
- [4] Snihurov A., Chakrian V. Improvement of EIGRP Protocol Routing Algorithm with the Consideration of Information Security Risk Parameters. Scholars Journal of Engineering and Technology. 2015. Vol. 3, No. 8. С. 707–714.
- [5] Євдокименко М. О., Шаповалова А. С., Шаповал М. М. Потоківі модель маршрутизації із врахуванням ризиків інформаційної безпеки за допомогою базових метрик критичності вразливостей. Проблеми телекомунікацій. 2020. № 1(26). С. 48–62. URL: [http://pt.nure.ua/wp-content/uploads/2021/03/201\\_yevdokimenko\\_security.pdf](http://pt.nure.ua/wp-content/uploads/2021/03/201_yevdokimenko_security.pdf).
- [6] Lou W., Kwon Y. H-SPREAD: A Hybrid Multipath Scheme for Secure and Reliable Data Collection in Wireless Sensor Networks. IEEE Transactions on Vehicular Technology. 2006. Vol. 55, No. 4. P. 1320–1330. DOI: <https://doi.org/10.1109/TVT.2006.877707>.
- [7] Yevdokymenko M., Yeremenko O., Shapovalova A., Shapoval M., Porokhniak V., Rogovaya N. Investigation of the Secure Paths Set Calculation Approach Based on Vulnerability Assessment. Workshop Proceedings of the MoMLeT+DS 2021: 3rd International Workshop on Modern Machine Learning Technologies and Data Science, June 5, 2021, Lviv-Shatsk, Ukraine. P. 207–217. URL: <http://ceur-ws.org/Vol-2917/paper19.pdf>.
- [8] Stallings W. Effective Cybersecurity: A Guide to Using Best Practices and Standards. Addison-Wesley Professional, 2018. 800 p.
- [9] CVSS v3.1 Specification Document. FIRST – Forum of Incident Response and Security Teams. URL: <https://www.first.org/cvss/v3.1/specification-document>.
- [10] Моделювання та оптимізація процесів безпечної та відмовостійкої маршрутизації в телекомунікаційних мережах: монографія / О. В. Лемешко, О. С. Єременко, М. О. Євдокименко та ін. Харків: ХНУРЕ, 2022. 198 с. DOI: <https://doi.org/10.30837/978-966-659-378-1>.
- [11] Abedin M., Nessa S., Al-Shaer E., Khan L. Vulnerability analysis For evaluating quality of protection of security policies. Quality of Protection (QoP): Proceedings of the 2nd ACM Workshop, 2006. P. 49–52. DOI: <https://doi.org/10.1145/1179494.1179505>.
- [12] CVSS v3.0 User Guide. FIRST – Forum of Incident Response and Security Teams. URL: <https://www.first.org/cvss/v3.0/user-guide>.

Надійшла до редколегії 7.05.2024



С. Ф. Чалий<sup>1</sup>, В. О. Лещинський<sup>2</sup>

<sup>1</sup>професор кафедри інформаційних управляючих систем,  
Харківський національний університет радіоелектроніки, Україна,  
serhii.chalyi@nure.ua

<sup>2</sup>доцент кафедри програмної інженерії,  
Харківський національний університет радіоелектроніки, Україна

## ПОБУДОВА ПОЯСНЕНЬ НА ЛОКАЛЬНОМУ РІВНІ ПРЕДСТАВЛЕННЯ ПРОЦЕСУ ФОРМУВАННЯ РІШЕНЬ ДЛЯ ВНУТРІШНІХ КОРИСТУВАЧІВ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ

Предметом вивчення в статті є процеси побудови пояснень щодо отриманих результатів та послідовності дій з прийняття рішення в інтелектуальній інформаційній системі.

Метою є розробка підходу до побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної системи на основі побудови темпоральних та каузальних залежностей, що створює умови для виявлення «вузьких місць» даного процесу при представленні інтелектуальної системи як «сірої» скриньки. Завдання: розробка моделі пояснення на локальному рівні представлення для внутрішніх користувачів інтелектуальної системи; розробка методу побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної системи. Використовуваними підходами є: методи побудови пояснень, підходи до побудови темпоральних знань. Висновки. Наукова новизна отриманих результатів полягає в наступному. Розроблено модель пояснення на локальному рівні представлення процесу формування рішення для внутрішніх користувачів інтелектуальної інформаційної системи. Модель використовує поєднання темпоральних залежностей, що визначають лінійну послідовність дій процесу, та каузальних правил-обмежень, які задають обов'язкові дії, що виконуються для всіх варіантів процесу формування рішення. Запропонована модель створює умови для виявлення дій, які обмежують нормальне виконання процесу формування рішення та призводять до результатів, що не відповідають вимогам зовнішніх користувачів. Розроблено метод побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної системи. Метод передбачає виділення темпоральних залежностей між парами послідовних станів поточного процесу формування рішення, відбір темпоральних залежностей, які виконуються для всіх варіантів процесу побудови рішення та формування на їх основі каузальних залежностей. Використання методу дає можливість представити як пояснення обмеження щодо формування рішення, що створює умови для виявлення вузьких місць у послідовності дій з отримання результату в інтелектуальній системі.

ІНТЕЛЕКТУАЛЬНА ІНФОРМАЦІЙНА СИСТЕМА, СИСТЕМА ШТУЧНОГО ІНТЕЛЕКТУ, ПОЯСНЕННЯ, ПРОЦЕС ПРИЙНЯТТЯ РІШЕНЬ, ПРОЦЕСНИЙ ПІДХІД ПРИЧИННО-НАСЛІДКОВИЙ ЗВ'ЯЗОК, МОЖЛИВІСТЬ, КАУЗАЛЬНІСТЬ

**S. Chalyi, V. Leshchynskyi. Construction of explanations at the local level of decision-making process representation for internal users of the intelligent system.** The subject of study in the article is the processes of constructing explanations for the obtained results and the sequence of decision-making actions in an intelligent information system. The goal is to develop an approach to constructing explanations at the local level of representation of the decision-making process of an intelligent system based on building temporal and causal dependencies, which creates conditions for identifying "bottlenecks" of this process when representing the intelligent system as a "gray" box. Tasks: development of an explanation model at the local level of representation for internal users of an intelligent system; development of a method for constructing explanations at the local level of representation of the decision-making process of an intelligent system. The approaches used are: methods of constructing explanations, approaches to constructing temporal knowledge. Conclusions. The scientific novelty of the obtained results is as follows. An explanation model at the local level of representation of the decision-making process for internal users of an intelligent information system has been developed. The model uses a combination of temporal dependencies that determine the linear sequence of process actions and causal rules-constraints that specify mandatory actions performed for all variants of the decision-making process. The proposed model creates conditions for identifying actions that limit the normal execution of the decision-making process and lead to results that do not meet the requirements of external users. A method for constructing explanations at the local level of representation of the decision-making process of an intelligent system has been developed. The method involves identifying temporal dependencies between pairs of consecutive states of the current decision-making process, selecting temporal dependencies that are performed for all variants of the decision-making process, and forming causal dependencies based on them. Using the method makes it possible to present decision-making constraints as an explanation, which creates conditions for identifying bottlenecks in the sequence of actions for obtaining a result in an intelligent system.

INTELLIGENT INFORMATION SYSTEM, ARTIFICIAL INTELLIGENCE SYSTEM, EXPLANATION, DECISION-MAKING PROCESS, PROCESS APPROACH, CAUSAL RELATIONSHIP, POSSIBILITY, CAUSALITY

## Вступ

Сучасні інтелектуальні інформаційні системи (ІС) використовуються для вирішення складних задач у сферах медицини, фінансів, промисловості тощо. ІС орієнтовані на обробку великих обсягів даних, виявлення закономірностей в цих даних та подальше прийняття рішень на основі алгоритмів машинного навчання. Однак, внаслідок використання таких алгоритмів, сучасні інтелектуальні системи, як правило, є «непрозорими» з точки зору користувачів. Тобто внутрішня логіка та процес формування рішень можуть бути незрозумілими для користувачів [1, 2].

Непрозорість функціонування ІС може призвести до недовіри з боку користувачів, особливо у критичних сферах, де рішення системи можуть мати значний вплив на життя та здоров'я людей. Крім того, відсутність розуміння того, як система прийшла до певного рішення, ускладнює процес виявлення та виправлення потенційних помилок у процесі формування рішень в ІС [3, 4].

Для вирішення проблеми непрозорості ІС та підвищення довіри до результатів роботи інтелектуальних систем доцільно надавати користувачам зрозумілі та інформативні пояснення щодо процесу прийняття рішень [5, 6]. Ці пояснення мають відповідати потребам як внутрішніх, так і зовнішніх користувачів.

Зовнішні користувачі є кінцевими споживачами рішень інтелектуальної системи. Вони зазвичай не володіють суттєвими знаннями про її внутрішню структуру та алгоритми роботи. Такі користувачі мають отримати зрозуміле обґрунтування рішень ІС на концептуальному рівні.

Внутрішні користувачі, зокрема розробники, аналітики, інженери з аналізу даних та експерти з предметної області, забезпечують конфігурування та налагодження роботи інтелектуальної системи й відповідають за її розробку, налаштування та підтримку. Для цих користувачів важливо в рамках пояснення отримати опис ключових послідовностей дій та ключових залежностей процесу формування рішень, щоб ефективно виявляти та усувати потенційні «вузькі місця», оптимізувати роботу системи та забезпечувати її відповідність вимогам кінцевого користувача. Зазначене свідчить про актуальність використання процесного підходу до побудови пояснень.

Процесний підхід базується на ідеї представлення процесу прийняття рішень у вигляді послідовності взаємопов'язаних дій та станів, що дозволяє відслідковувати логіку та причинно-наслідкові зв'язки в роботі інтелектуальної системи [7, 8]. Процесний підхід широко використовується в управлінні бізнес-процесами для опису, аналізу та оптимізації складних процесів в організаціях [9, 10].

Застосування процесного підходу до побудови пояснень в інтелектуальних системах створює умови для

побудови структурованого опису процесу прийняття рішень для внутрішніх користувачів. Маючи такі пояснення, розробники та аналітики можуть краще зрозуміти поведінку системи, виявити потенційні помилки та вдосконалити алгоритми прийняття рішень [11].

Дослідження у сфері пояснювального штучного інтелекту (ХАІ) активно розвиваються протягом останніх років. В рамках даного напрямку вирішуються задачі розробки методів побудови пояснень, а також створення ментальних моделей користувачів. Перша задача орієнтована на забезпечення прозорості роботи інтелектуальних систем. Друга задача орієнтована на підтримку розуміння пояснень користувачами [5, 12, 13].

При побудові пояснень суттєва увага приділяється розробці методів, які дозволяють генерувати пояснення на основі аналізу внутрішньої структури та процесу прийняття рішень в інтелектуальних системах. Такі методи, як LIME [14], SHAP [15] та Grad-CAM [16], дозволяють визначити важливість окремих ознак або компонентів моделі для прийняття конкретного рішення.

Інший напрямок досліджень пов'язаний побудовою пояснень з використанням темпоральних та каузальних залежностей. Каузальні моделі дозволяють описати причинно-наслідкові зв'язки між змінними та подіями в процесі прийняття рішень [17, 18]. Використання каузальних залежностей дозволяє створювати більш інтерпретовані та зрозумілі пояснення, які відображають логіку роботи інтелектуальної системи [19, 20]. В роботах [21–23] запропоновано підхід до побудови каузальних моделей пояснень для інтелектуальних систем, який дозволяє враховувати темпоральні та причинно-наслідкові зв'язки між подіями та станами системи.

В роботах [24, 25] запропоновано підхід до побудови пояснень з урахуванням темпорального аспекту.

В роботах [26, 27] запропоновано підходи до опису темпоральних залежностей між подіями та станами в процесі прийняття рішень, а також підхід до побудови процесу прийняття рішень на основі темпоральних залежностей [28]. Ці залежності дозволяють відобразити динаміку зміни станів системи та причинно-наслідкові зв'язки між ними.

Однак в наведених дослідженнях розглядаються як пояснення окремі ключові елементи процесу формування рішень в інтелектуальних системах. Представленню пояснення для внутрішніх користувачів шляхом опису процесу на основі темпоральних та каузальних залежностей не приділяється достатньо уваги, що свідчить про важливість виділення рівнів процесного представлення пояснення для внутрішніх користувачів та розробки відповідних методів побудови пояснень в інтелектуальних системах в темпоральному та каузальному аспектах. Темпоральні

залежності дозволяють описати послідовність станів та подій в процесі прийняття рішень, тоді як каузальні залежності відображають причинно-наслідкові зв'язки між цими станами та подіями. Поєднання цих двох типів залежностей дає можливість створювати детальний опис процесу формування пояснень, які відображають логіку роботи інтелектуальної системи для внутрішніх користувачів.

### 1. Постановка задачі

Метою статті є розробка підходу до побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної системи на основі побудови темпоральних та каузальних залежностей, що створює умови для виявлення «вузьких місць» даного процесу при представленні інтелектуальної системи як «сірої» скриньки.

Для досягнення поставленої мети вирішуються такі задачі:

- розробка моделі пояснення на локальному рівні представлення для внутрішніх користувачів інтелектуальної системи;
- розробка методу побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної системи.

### 2. Модель пояснення на локальному рівні представлення процесу формування рішення

Загальну схему представлення пояснень для зовнішніх та внутрішніх користувачів наведено на рис. 1. Пояснення містить локальний, процесний та глобальний рівні. Перші два рівні відображають процес формування рішення з різним ступенем деталізації та орієнтовані на внутрішнього користувача. Дані два рівня відображають процедурні знання і тому явно враховують темпоральний аспект. Третій рівень орієнтований на зовнішнього користувача та відображає причини отриманого в інтелектуальній системі результату. На даному рівні темпоральний аспект враховується неявно, оскільки залежність між вхідними даними та отриманим результатом передбачає, що був реалізований процес формування рішення. Глобальний рівень представляє процес формування рішення в інтелектуальній інформаційній системі як одну глобальну дію, відображаючи базові зв'язки між вхідними даними та результатом роботи системи.

Пояснення на локальному рівні представлення процесу формування рішення в ІС базується на використанні окремих темпоральних та каузальних залежностей між парами станів процесу або парами подій, що відображають виникнення відповідних станів.

Темпоральні залежності  $f_{n-1,n}^{(i)}$  визначають пари упорядкованих у часі послідовних станів (подій виникнення станів)  $e_{n-1}^{(i)}$  та  $e_n^{(i)}$   $i$ -реалізації процесу

формування рішення. Перехід між цими станами є наслідком виконання дії  $d_{n-1}^{(i)}$ :

$$f_{n-1,n}^{(i)} : e_{n-1}^{(i)} \xrightarrow{d_{n-1}^{(i)}} e_n^{(i)}. \tag{1}$$

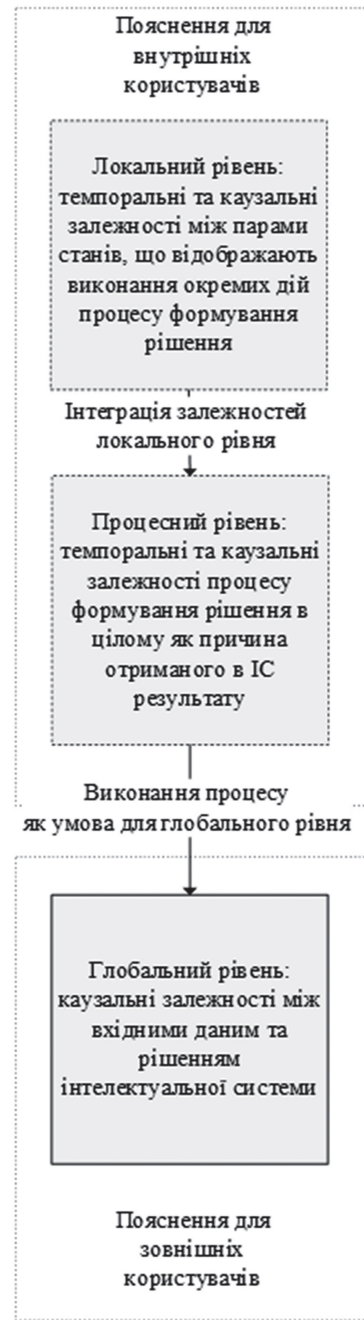


Рис. 1. Ієрархічне представлення пояснення для внутрішніх та зовнішніх користувачів

Тобто поточний процес формування рішення в ІС відображається в поясненні у вигляді лінійної послідовності упорядкованих у часі подій, причому дана лінійна послідовність однозначно задається темпоральними залежностями між парами послідовних подій.

Загальну схему локального рівня процесного опису пояснення для  $i$ -реалізації процесу формування рішення в інтелектуальній системі представлено на рис. 2.

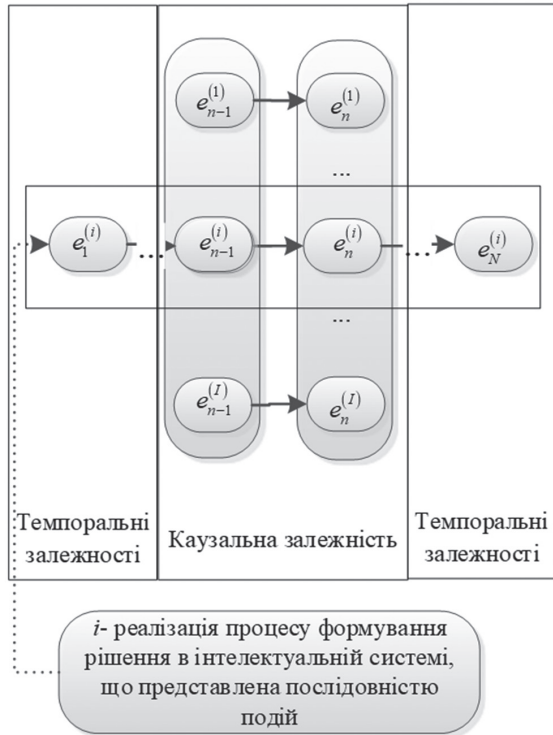


Рис. 2. Локальний рівень процесного опису пояснення

Дана схема містить послідовність подій виникнення станів  $e_1^{(i)}, \dots, e_n^{(i)}, \dots$  для даної реалізації процесу, а також темпоральні та каузальні залежності між парами станів.

Темпоральні залежності мають, наприклад, вигляд  $d_1^{(i)} \rightarrow e_2^{(i)}$  для першої дії процесу формування рішення  $d_1^{(i)}$ , або  $e_{l-1}^{(i)} \rightarrow e_l^{(i)}$  для останньої дії даного процесу.

В тому випадку, якщо одна й та ж сама дія  $d_{n-1}^{(i)}$  в одному й тому ж стані  $e_{n-1}^{(i)}$  виконується для всіх можливих варіантів реалізації процесу прийняття рішення, то відповідне темпоральне правило розглядається як каузальна залежність:

$$e_{n-1}^{(i)} \Rightarrow e_n^{(i)} \mid (\forall i) \exists d_{n-1}^{(i)}. \quad (2)$$

Згідно (2), темпоральна залежність  $e_{l-1}^{(i)} \rightarrow e_l^{(i)}$  не має альтернативи для всіх відомих варіантів процесу формування рішення в ІС. Тому вираз (2) для каузальної залежності приймає вигляд каузального обмеження:

$$c_n^{(n-1)} : e_{n-1} \Rightarrow e_n. \quad (3)$$

Оскільки дане обмеження не залежить від конкретної реалізації процесу формування рішення, то його можна розглядати як детерміноване правило-обмеження.

Подія  $e_n$  фіксує завершення дії  $d_{n-1}$ , тому дану дію можна розглядати як один із атрибутів події  $e_n$ . Сукупність атрибутів даної події фіксує не лише

завершення  $d_{n-1}$ , а і властивості стану після завершення цієї дії.

Тому каузальне обмеження доцільно представити у більш загальному вигляді:

$$c_n^{(n-1)} : e_{n-1} \Rightarrow e_n. \quad (4)$$

Вираз (4) показує, що умовою події  $e_n$  є наявність попередньої події  $e_{n-1}$ , а причина визначається через атрибути цих подій, які були змінені при переході від  $e_{n-1}$  до  $e_n$ .

У випадку визначення підмножини атрибутів  $A$  вираз (4) приймає вигляд:

$$c_n^{(n-1)} : e_{n-1} \stackrel{A}{\Rightarrow} e_n. \quad (5)$$

Ключовим атрибутом зазвичай є дія, тому обмеження (4) визначає обов'язковість виконання дії  $d_{n-1}$  у стані  $e_{n-1}$ .

Відповідно, представлення цього правила як елемента пояснення на локальному рівні дає можливість внутрішньому користувачеві побачити всі обов'язкові дії процесу формування рішення і обов'язкові зміни станів, що є наслідком цих дій. У випадку, якщо такі дії накладають додаткові (необов'язкові) обмеження на процес формування рішення, то це може свідчити про «вузькі місця» у даному процесі.

Коригування наведених обмежень дає можливість усунути ці проблемні місця.

### 3. Метод побудови пояснень на локальному рівні представлення процесу формування рішень інтелектуальної інформаційної системи

На локальному рівні пояснення формується на основ каузального обмеження, що відображає допустимі зміни атрибутів пара послідовних станів процесу формування рішення в інтелектуальній системі. Такі пояснення фактично визначають детерміновані правила побудови рішення ІС.

Метод побудови пояснень на локальному рівні містить такі етапи.

Етап 1. Виділення темпоральних залежностей між парами послідовних станів поточного процесу формування рішення згідно (1).

Результатом даного етапу є множина  $F^{(i)}$  пар подій  $e_{n-1}^{(i)}, e_n^{(i)}$  щодо виникнення станів процесу, який відображає послідовність виконання  $i$ -процесу побудови рішення у часі:

$$F^{(i)} = \left\{ e_{n-1}^{(i)} \xrightarrow{d_{l-1}^{(i)}} e_n^{(i)} \right\}. \quad (6)$$

Етап 2. Відбір підмножини  $A$  атрибутів подій виникнення станів поточної  $i$ -реалізації процесу, що відображають причини переходу між станами.

Кожен стан процесу формування рішення характеризується множиною значень змінних, які

відображають властивості сутностей предметної області. Ці властивості можуть бути статичними або динамічними. Статичні властивості характеризують об'єкт в цілому. Динамічні властивості змінюють свої значення з часом, по мірі виконання процесу формування рішення. Такі властивості дають можливість відрізнити поточний стан сутності від інших її станів. Тому на другому етапі методу відбираються динамічні атрибути станів тобто суттєві властивості, зміна яких відображає перехід між станами процесу. Відбір атрибутів базується на вхідній інформації щодо групування вхідних даних по сутностям предметної області. Фактично, відібрані атрибути мають бути найбільш інформативними для пояснення переходу між станами.

Етап 3. Виділення темпоральних залежностей, які виконуються для всіх варіантів процесу формування рішення в інтелектуальній системі.

На даному етапі перевіряється умова (2) виконання темпоральних залежностей з урахуванням відібраних атрибутів станів для всіх відомих варіантів процесу формування рішення. Це дозволяє виявити детерміновані послідовності зміни станів процесу формування рішення.

Етап 4. Формування каузальних залежностей.

На даному етапі формуються темпоральні обмеження згідно (4) або (5).

Результатом методу є упорядкована за послідовністю дій процесу множини каузальних обмежень, що представляють собою пояснення локального рівня представлення процесу формування рішення в ПС.

Розглянемо приклад використання розробленого методу побудови пояснення на основі побудови каузальних обмежень при формуванні персоналізованого переліку товарів та послуг в рекомендаційній системі.

Вхідними даними є лог процесу формування рекомендацій в онлайн-кінотеатрі. Лог складається з трас. Кожна траса складається із послідовності подій, та є записом про одну реалізацію процесу побудови рекомендацій. Події траси упорядковані за часом виникнення. Атрибути подій містять інформацію про дії з формування персоналізованого переліку фільмів, що відображають

Розглянемо випадок, коли рекомендаційна система надає користувачеві рекомендації щодо фільмів на основі його вподобань, що відображені в історії перегляду.

Процес формування рекомендацій включає наступний узагальнений перелік дій  $D^{(1)} = \langle d_1^{(1)}, \dots, d_7^{(1)} \rangle$ :

– збір даних про користувача ( $d_1^{(1)}$ ): вік, стать, місце проживання тощо.

– аналіз історії перегляду користувача ( $d_2^{(1)}$ ): фільми, що переглядалися, оцінки, що були виставлені.

– визначення вподобань користувача ( $d_3^{(1)}$ ): жанри, актори, режисери тощо.

– пошук схожих користувачів ( $d_4^{(1)}$ ): виконується на основі вподобань та історії перегляду.

– формування списку рекомендованих фільмів ( $d_5^{(1)}$ ): виконується на основі вподобань схожих користувачів.

– ранжування рекомендованих фільмів ( $d_6^{(1)}$ ): виконується за відповідністю до вподобань користувача.

– надання користувачеві топ-N рекомендацій ( $d_7^{(1)}$ ).

Кожна з цих дій відповідає певній події щодо виникнення стану процесу формування рекомендацій. Позначимо ці події як  $e_1^{(1)}, e_2^{(1)}, \dots, e_7^{(1)}$ .

Реалізація методу із наведеними вхідними даними має такий вигляд.

Етап 1. Виділення темпоральних залежностей між парами послідовних станів.

Виділяємо наступні темпоральні залежності між парами послідовних у часі подій:

$$e_1^{(1)} \rightarrow e_2^{(1)}, e_2^{(1)} \rightarrow e_3^{(1)}, e_3^{(1)} \rightarrow e_4^{(1)}, \\ e_4^{(1)} \rightarrow e_5^{(1)}, e_5^{(1)} \rightarrow e_6^{(1)}, e_6^{(1)} \rightarrow e_7^{(1)}.$$

Етап 2. Виділення дій, які привели до переходів між станами.

Для кожної темпоральної залежності визначаємо відповідну дію:

$$e_1^{(1)} \xrightarrow{d_2^{(1)}} e_2^{(1)} \text{ – аналіз історії перегляду користувача;} \\ e_2^{(1)} \xrightarrow{d_3^{(1)}} e_3^{(1)} \text{ – визначення вподобань користувача;} \\ e_3^{(1)} \xrightarrow{d_4^{(1)}} e_4^{(1)} \text{ – пошук схожих користувачів;} \\ e_4^{(1)} \xrightarrow{d_5^{(1)}} e_5^{(1)} \text{ – формування списку рекомендованих фільмів;} \\ e_5^{(1)} \xrightarrow{d_6^{(1)}} e_6^{(1)} \text{ – ранжування рекомендованих фільмів;} \\ e_6^{(1)} \xrightarrow{d_7^{(1)}} e_7^{(1)} \text{ – надання користувачеві топ-N рекомендацій.}$$

Етап 3. Узагальнення темпоральних залежностей та відповідних дій.

На даному етапі виконується аналіз та порівняння з  $D^{(1)}$  схожих трас побудови рекомендацій  $D = \{D^{(i)} : i \neq 1\}$ .

Якщо в результаті аналізу трас логу рекомендаційної системи виявлено, що послідовність дій  $d_3$  – визначення вподобань користувача,  $d_4$  – пошук схожих користувачів,  $d_5$  – формування списку рекомендацій присутня в усіх реалізаціях процесу формування рекомендацій, то на базі вказаних дій можуть бути сформовані обмеження.

Тобто для відповідних темпоральних залежностей виконується умова:

$$(\forall i) \exists d_n^{(i)} \equiv d_n \mid 2 < n < 6. \quad (7)$$

Відповідні темпоральні залежності є однаковими на всіх трасах процесу та мають вигляд:

$$e_2 \xrightarrow{d_3} e_3, e_3 \xrightarrow{d_4} e_4, e_4 \xrightarrow{d_5} e_5. \quad (8)$$

Етап 4. Формування каузальних обмежень.

На основі виявлених на попередньому етапі темпоральних залежностей та дій формуються наступні каузальні обмеження:

$e_2 \xrightarrow{d_3} e_3$ : визначення вподобань користувача  $d_3$  є причиною переходу від події  $e_2$  (проаналізована історія перегляду) до події  $e_3$  (визначені вподобання);

$e_3 \xrightarrow{d_4} e_4$ : Пошук схожих користувачів  $d_4$  є причиною переходу від події  $e_3$  (визначені вподобання) до події  $e_4$  (знайдені схожі користувачі).

$e_4 \xrightarrow{d_5} e_5$ : Формування списку рекомендованих фільмів  $d_5$  є причиною переходу від стану  $e_4$  (знайдені схожі користувачі) до стану  $e_5$  (сформовано список рекомендацій).

Сформовані таким чином каузальні обмеження є поясненнями на локальному рівні представлення. Ці пояснення показують, що визначення вподобань користувача, пошук схожих користувачів та формування списку рекомендацій на їх основі є ключовими діями, які впливають на кінцевий результат — надані користувачеві рекомендації.

Таке пояснення дає можливість команді підтримки деталізувати логіку роботи системи визначити обов'язкові дії, які вплинули на рекомендацію щодо фільмів цільовому користувачеві. Якщо рекомендації виявилися неточними, то на основі пояснення можна зробити висновок, що причиною цього могли бути некоректно визначені вподобання користувача або невдало підібрані схожі користувачі. Це дозволить зробити відповідні корективи в алгоритми рекомендаційної системи для покращення якості рекомендацій.

### Висновки

Представлено багаторівневе представлення пояснення щодо процесу формування рішення в інтелектуальній системі, яке орієнтовано на потреби внутрішніх та зовнішніх користувачів. Пояснення для внутрішніх користувачів базується на формуванні темпоральних та каузальних залежностей, що відображають процес формування рішення. Пояснення для зовнішніх користувачів відображає залежність між вхідними даними та отриманим рішенням.

Розроблено модель пояснення на локальному рівні представлення процесу формування рішення

для внутрішніх користувачів інтелектуальної інформаційної системи. Модель містить сукупність темпоральних залежностей та каузальних правил-обмежень. Темпоральні залежності визначають лінійну послідовність дій процесу, а каузальні обмеження визначають дії, які виконуються для всіх варіантів процесу формування рішення. Модель дає можливість виявити такі дії, які обмежують нормальне виконання процесу формування рішення, що приводить до результатів, які не відповідають вимогам зовнішніх користувачів.

Розроблено метод побудови пояснень на локальному рівні представлення процесу формування рішення інтелектуальної системи. Метод містить етапи виділення темпоральних залежностей між парами послідовних станів поточного процесу формування рішення, виділення темпоральних залежностей, які виконуються для всіх варіантів процесу побудови рішення, формування каузальних залежностей. Метод дає можливість представити як пояснення обмеження щодо формування рішення, що створює умови для виявлення вузьких місць у послідовності дій з отримання результату у інтелектуальній системі.

### Список літератури:

- [1] Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access*, 6, 52138-52160.
- [2] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM computing surveys (CSUR)*, 51(5), 1-42.
- [3] Baier, L., Kühl, N., & Satzger, G. (2019, July). How to cope with change? Preserving validity of predictive services over time. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*.
- [4] Biran, O., & Cotton, C. (2017, August). Explanation and justification in machine learning: A survey. In *IJCAI-17 workshop on explainable AI (XAI) (Vol. 8, No. 1, pp. 8-13)*.
- [5] Чалий, С., & Лещинська, І. Концептуальна ментальна модель пояснення в системі штучного інтелекту // *Вісник Національного технічного університету «ХПІ»*. Серія: Системний аналіз, управління та інформаційні технології — 2023. — Vol. 1(9). — P. 70–75 <https://doi.org/10.20998/2079-0023.2023.01.11>
- [6] Tintarev, N., & Masthoff, J. (2012). Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*, 22(4-5), 399-439. <https://doi.org/10.1007/s11257-011-9117-5>
- [7] Van Der Aalst, W. (2016). *Process mining: data science in action*. Springer.
- [8] Rehse, J. R., Mehdiyev, N., & Fettke, P. (2019). Towards explainable process predictions for industry 4.0 in the dfki-smart-lego-factory. *KI-Künstliche Intelligenz*, 33(2), 181-187.

- [9] Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). *Fundamentals of business process management* (Vol. 1, p. 2). Heidelberg: Springer.
- [10] Weske, M. (2019). *Business process management: concepts, languages, architectures*. Springer.
- [11] Mehdiyev, N., Evermann, J., & Fettke, P. (2020). A novel business process prediction model using a deep learning method. *Business & information systems engineering*, 62(2), 143-157.
- [12] Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... & Herrera, F. (2019). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.
- [13] Tjoa, E., & Guan, C. (2020). A survey on explainable artificial intelligence (XAI): Toward medical XAI. *IEEE Transactions on Neural Networks and Learning Systems*.
- [14] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135-1144. <https://doi.org/10.1145/2939672.2939778>SHAP:
- [15] Lundberg, S. M., & Lee, S. I. (2017). A Unified Approach to Interpreting Model Predictions. *Advances in Neural Information Processing Systems*, 30, 4765-4774. Grad-CAM:
- [16] Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Proceedings of the IEEE International Conference on Computer Vision*, 618-626. <https://doi.org/10.1109/ICCV.2017.74>
- [17] Pearl, J. (2009). *Causality*. Cambridge university press.
- [18] Spirtes, P., Glymour, C. N., Scheines, R., & Heckerman, D. (2000). *Causation, prediction, and search*. MIT press.
- [19] Holzinger, A., Langs, G., Denk, H., Zatloukal, K., & Müller, H. (2019). Causability and explainability of artificial intelligence in medicine. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(4), e1312.
- [20] Moraffah, R., Karami, M., Guo, R., Raglin, A., & Liu, H. (2020). Causal interpretability for machine learning-problems, methods and evaluation. *ACM SIGKDD Explorations Newsletter*, 22(1), 18-33.
- [21] Chalyi, S., & Leshchynskiy, V. (2023). Probabilistic counterfactual causal model for a single input variable in explainability task. *Advanced Information Systems*, 7(3), 54-59.
- [22] Чалий, С. Ф., & Лещинський, В. О. (2021). Можливісна модель каузального зв'язку по вхідній змінній для побудови пояснення в інтелектуальній системі. *Вісник Національного технічного університету "ХПІ". Серія: Системний аналіз, управління та інформаційні технології*, (2), 97-103. <https://doi.org/10.20998/2079-0023.2021.02.17>
- [23] Chalyi, S., & Leshchynskiy, V. (2020). Method of constructing explanations for recommender systems based on the temporal dynamics of user preferences. *EUREKA: Physics and Engineering*, 3, 43-50. <https://doi.org/10.21303/2461-4262.2020.001228>
- [24] Chalyi, S., & Leshchynskiy, V. (2020). Method of constructing explanations for recommender systems based on the temporal dimension of user choice. *EUREKA: Physics and Engineering*, 3, 43-50.
- [25] Чалий, С. Ф., Лещинський, В. О., & Лещинська, І. О. (2021). Контрфактуальна темпоральна модель причинно-наслідкових зв'язків для побудови пояснень в інтелектуальних системах. *Вісник Національного технічного університету "ХПІ". Сер.: Системний аналіз, управління та інформаційні технології*, (2(6)), 41-46.
- [26] Chala O. (2018) Models of temporal dependencies for a probabilistic knowledge base. *Econtechmod. An International Quarterly Journal*. Vol. 7, No. 3. P. 53 – 58.
- [27] Чала О. В. (2020) Модель узагальненого представлення темпоральних знань для задач підтримки управлінських рішень. *Вісник Національного технічного університету «ХПІ». Системний аналіз, управління та інформаційні технології*. № 1(3). С. 14-18. DOI: 10.20998/2079-0023.2020.01.03.
- [28] Levykin V., Chala O. Development of a method of probabilistic inference of sequences of business process activities to support business process management. *Eastern-European Journal of Enterprise Technologies*. 2018. № 5/3(95). P. 16-24. DOI: 10.15587/1729-4061.2018.142664.

*Надійшла до редколегії 16.05.2024*





В. І. Дубровін<sup>1</sup>, К. С. Паничук<sup>2</sup>, К. О. Роговська<sup>3</sup>, Ю. В. Тарасова<sup>4</sup>

<sup>1</sup>к.т.н., професор кафедри програмних засобів,  
Національний університет «Запорізька політехніка», Запоріжжя, Україна,  
e-mail: vdubrovin@gmail.com, ORCID ID: 0000-0002-0848-8202

<sup>2</sup>студентка групи КНТ-130, Національний університет «Запорізька політехніка»,  
Запоріжжя, Україна, e-mail: panichuck.kat@gmail.com, ORCID ID: 0009-0009-4340-6688

<sup>3</sup>IOS Developer IT-компанії «Roll'n'Code», Запоріжжя, Україна, e-mail: katrin.riano@gmail.com

<sup>4</sup>к.т.н., Web Developer IT-компанії «Freshcode», Запоріжжя, Україна, e-mail: julia.tverdohleb@gmail.com

## ПРОГРАМНА РЕАЛІЗАЦІЯ МЕТОДІВ АУДІОВІЗУАЛЬНОЇ СТИМУЛЯЦІЇ ГОЛОВНОГО МОЗКУ

Аудіовізуальна стимуляція мозку, як сучасний підхід до впливу на нейрофізіологічні процеси, надає великий потенціал для поліпшення психічного та фізичного здоров'я людини, а також діагностики та лікування різних порушень. Використання перетворень Фур'є та вейвлет-перетворень дозволяє аналізувати сигнали електроенцефалограм та електрокардіограм і створювати ефективні світлові та аудіостимули. Розглядаються наявні дослідження, які підтверджують важливість біоелектричної активності людини та методів її обробки у виконанні аудіовізуальної стимуляції для лікування різноманітних захворювань у дорослих та дітей. Наведено підхід до програмної реалізації методів аудіовізуальної стимуляції головного мозку та результати розробки. Тестування програмного забезпечення продемонструвало позитивний вплив світлових та аудіостимулів на більшість учасників дослідження. Це свідчить про важливість і перспективи подальших досліджень у цьому напрямку.

АУДІОВІЗУАЛЬНА СТИМУЛЯЦІЯ, ЕЕГ, ЕКГ, ЦЕНТРАЛЬНА НЕРВОВА СИСТЕМА, ПЕРЕТВОРЕННЯ ФУР'Є, ВЕЙВЛЕТ–ПЕРЕТВОРЕННЯ, MATLAB, ARDUINO

**V. Dubrovin, K. Panychuk, K. Rohovska, Y. Tarasova. Software implementation of audiovisual brain stimulation methods.** Audiovisual brain stimulation, as a modern approach to influencing neurophysiological processes, has great potential for improving mental and physical health, as well as diagnosing and treating various disorders. The use of Fourier transforms and Wavelet transforms allows analyzing signals from electroencephalograms and electrocardiograms and creating effective light and audio stimuli. The available studies confirming the importance of human bioelectrical activity and methods of its processing in the implementation of audiovisual stimulation for the treatment of various diseases in adults and children are considered. The approach to the software implementation of audiovisual brain stimulation methods and the results of its development are presented. Software testing has demonstrated a positive effect of light and audio stimuli on the majority of the study participants. This indicates the importance and prospects for further research in this area.

AUDIOVISUAL STIMULATION, EEG, ECG, CENTRAL NERVOUS SYSTEM, FOURIER TRANSFORM, WAVELET TRANSFORM, MATLAB, ARDUINO

### Вступ

В сучасному світі багато людей страждають від різних психічних і фізичних захворювань. Дефіцит ефективних та нефармакологічних методів лікування стає однією з актуальних проблем сучасної медицини. Важливою є необхідність створення терапій, що взаємодіють з організмом природним способом, активуючи його внутрішні резерви. Методи, які впливають на обмін речовин, регуляторні системи, адаптацію та компенсацію, мають великий потенціал [1]. Один із таких методів, що набуває все більшої популярності в медичній спільноті, – це аудіовізуальна стимуляція (АВС). Вона сприяє мобілізації природних резервів організму за допомогою світлових та акустичних впливів і є гарною альтернативою в боротьбі з численними хворобами та станами, де традиційні методи не завжди дають бажаний результат.

Серед таких хвороб варто зазначити депресію та тривожність, що стали актуальними в сучасному

суспільстві. Аудіовізуальна стимуляція може допомогти пацієнтам знижувати рівень тривожності та відновлювати спокій. Крім того, вона є корисною для пацієнтів із порушеннями сну, такими як безсоння [2]. Регулюючи мозкову активність, аудіовізуальна стимуляція сприяє покращенню якості сну та відновленню внутрішнього біологічного годинника.

Застосування аудіовізуальної стимуляції також розглядається у лікуванні синдрому дефіциту уваги з гіперактивністю (СДУГ), де може бути корисною для поліпшення концентрації та розумової діяльності. Ця методика також вивчається для реабілітації пацієнтів з травмами мозку, а також для підтримання мозкової активності та нейропластичності в лікуванні нейродегенеративних захворювань, таких як хвороба Паркінсона та хвороба Альцгеймера [3].

Отже, аудіовізуальна стимуляція відкриває перспективи для покращення психічного та фізичного здоров'я і дає нові можливості в лікуванні та реабілітації.

### 1. Біологічні ритми людини

Одними з найважливіших біологічних показників людини для виконання аудіовізуальної стимуляції є ритми мозку та серця. Для реєстрації біопотенціалів мозку використовується електроенцефалографія (ЕЕГ), а для реєстрації електричної активності серця – електрокардіографія (ЕКГ).

Електроенцефалограма є важливим інструментом для дослідження патологій центральної нервової системи, пов'язаних із судомою та зміною психічного стану [4]. Вона є додатковим тестом до більш складних візуалізаційних досліджень. ЕЕГ широко використовується для оцінки пацієнтів з епілепсією, зміненим психічним статусом або зміненою свідомістю, парасомніями, енцефалопатіями, вторинними по відношенню до різних метаболічних порушень та інтоксикацій, деменціями та інсультами, що протікають у вигляді судомних нападів [4]. ЕЕГ також корисна для визначення прогнозів у пацієнтів з аноксичним ураженням мозку, черепно-мозковими травмами, визначення смерті мозку та медикаментозної інтоксикації. ЕЕГ є універсальним інструментом для оцінки будь-якої інтервальної активності мозкових хвиль і для кращого розуміння основного прогресу у людини, яка не реагує або перебуває в комі. Вона також корисна для оцінки пацієнтів з поведінковими або психогенними розладами, які схожі на судомні напади.

Основні ритми мозку (рис. 1) – це альфа- (частота від 8 до 13 Гц, амплітуда 15-100 мкВ), бета- (частота від 13 до 35 Гц, амплітуда 3-10 мкВ), дельта- (частота від 0,3 до 4 Гц, амплітуда 20-30 мкВ), тета- (частота від 4 до 8 Гц, амплітуда до 40 мкВ), і гамма-ритм (частота від 35 до 500 Гц, амплітуда зазвичай до 10 мкВ) [5].

ЕЕГ сигнал	Назва груп хвиль	Частота хвилі, Гц	Фізичний стан людини
	Gamma	35-500	Висока розумова активність
	Beta	13-35	Тривога
	Alpha	8-13	Розслаблення
	Theta	4-8	Рання стадія сну
	Delta	0,3-4	Глибокий сон

**Рис. 1. Діапазони частот хвиль мозкової активності, що характеризують фізичний стан людини [2]**

Альфа-ритм максимально виражений в задніх (потиличних) відведеннях при закритих очах. Альфа-хвилі пов'язані з умінням людини розслабитися, здатністю долати стрес, а також засвоювати нову інформацію. У нормі відзначається зниження альфа-ритму на ЕЕГ при розплющенні очей, занепокоєнні, при активній розумовій діяльності, а також під час сну [5].

Бета-ритм має максимальну вираженість в лобово-центральному відведенні. Він посилюється в період сонливості, при засинанні та іноді при пробудженні.

В період глибокого сну амплітуда і вираженість бета-ритму істотно знижується. Посилення активності бета-ритму відзначається при прийомі психоактивних препаратів. Регіональне зниження бета-ритму одночасно зі зниженням альфа-ритму може говорити про структурне ушкодження або дефект кори головного мозку [5].

Максимальна вираженість тета-ритму зустрічається у дітей 4-6 років. Він може реєструватися на ЕЕГ здорової притомної людини під час емоційної активації. Однак існує безліч патологічних станів або змінених станів свідомості (сон, медитація), що супроводжуються розвитком довготривалої і короткочасної тета-активності, більшість з яких вимагає проведення нейровізуалізації [5].

Дельта-ритм реєструється під час глибокого сну, при гіпервентиляції або в стані наркозу. Домінування дельта-активності у підлітків і дорослих в стані неспання – ознака патології. Виявляється у пацієнтів з наявністю енцефалопатій, що супроводжуються змінами рівня свідомості (кома), а також є ознакою важкого структурного мозкового порушення (пухлина, інсульт, травма, абсцес). В поодиноких роботах зустрічається інформація, що коливання в цьому діапазоні можуть бути ідентифіковані в сигналі ЕЕГ у стані спокою при деяких формах стресу [5].

Гамма-ритм спостерігається під час високої розумової активності, такої як навчання, сприймання інформації, а також розв'язання складних завдань. Реєструється в прецентральному, фронтальному, скроневій і тім'яній зонах кори головного мозку [6].

Отримані ритми ЕЕГ фільтруються та аналізуються для подальшого створення звукових доріжок та налаштування світлових стимулів, які використовуються під час аудіовізуальної стимуляції.

ЕКГ – це неінвазивний метод діагностики, який має значний клінічний вплив на дослідження тяжкості серцево-судинних захворювань [7]. ЕКГ все частіше використовується для моніторингу пацієнтів, які приймають антиаритмічні препарати та інші ліки, як невід'ємна частина передопераційної оцінки пацієнтів, яким проводять некардіохірургічні операції, а також для скринінгу осіб, які займаються професіями з підвищеним рівнем ризику та/або спортом. Крім того, ЕКГ слугує інструментом дослідження для спостереження та експериментальних випробувань ліків з визнаними кардіологічними ефектами [7].

Під час реєстрації електричної активності серця на електрокардіограмі присутні такі складові (рис. 2), як Р – зубець, що відображає перебіг поширення збудження передсерддями; інтервал Р-Q – час від початку збудження передсердь до початку збудження шлуночків; інтервал Q-T – час електричної систоли шлуночків, що включає поширення збудження шлуночками серця – комплекс QRS, сегмент ST і зубець

T; хвиля U, яка в нормі спостерігається не завжди; R-R (P-P) – міжцикловий інтервал; T-P – діастолічний інтервал [8].

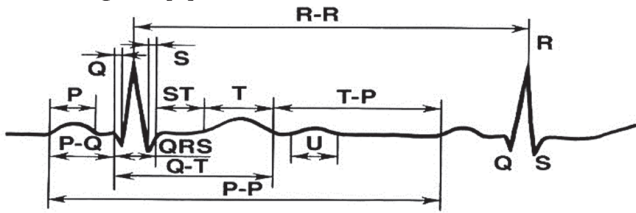


Рис. 2. Схематичне зображення нормальної електрокардіограми [8]

Електрокардіограми також використовуються для створення звукових доріжок, які застосовуються під час АВС.

Використання електроенцефалографії (ЕЕГ) та електрокардіографії (ЕКГ) при аудіовізуальній стимуляції головного мозку дозволяє створювати індивідуалізовані стимули та ефективно оцінювати їх вплив на мозкову і серцеву активність. Завдяки синергії цих методів можна отримати комплексну картину фізіологічних реакцій організму, що сприяє розробці оптимальних програм для релаксації, покращення когнітивних функцій та лікування різних психічних і фізичних розладів.

## 2. Цифрова обробка сигналів

Використання перетворення Фур'є та вейвлет-перетворень стало важливою складовою аналізу електроенцефалограм та електрокардіограм. Перетворення Фур'є дозволяє розкласти сигнали на їхні складові частоти, допомагаючи виявити основні ритми та характеристики активності мозку та серця. Вейвлет-перетворення, з іншого боку, надають додаткову інформацію про часові та частотні властивості сигналів, що дозволяє виявляти короточасні зміни та артефакти.

Перетворення Фур'є – це дослідження функції на наявність у ній гармонічної складової, що має частоту  $\omega$ , та визначення комплексної ваги цієї складової (амплітуди, початкової фази) у вибраній системі координат [9].

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt, \quad (1)$$

Формула перетворення Фур'є (1).

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega, \quad (2)$$

Функцію  $f(t)$ , яка визначається формулою (2), називають зворотним перетворенням Фур'є [9].

Вейвлети – це математичні функції, що дозволяють аналізувати різні частотні компоненти даних. Вейвлети мають істотні переваги в порівнянні з перетворенням Фур'є, тому що вейвлет-перетворення дозволяє робити висновки не тільки про частотний

спектр сигналу, але також про те, в який момент часу з'явилася та чи інша гармоніка [10].

Сигнал аналізується шляхом розкладання за базисними функціями, які було отримано з деякого прототипу шляхом стиснень, розтягувань і зсувів (3). Функція-прототип називається аналізуючим (материнським) вейвлетом [11].

Вейвлет-функція повинна задовольняти двом умовам [11]:

1. Середнє значення (інтеграл усією прямою) дорівнює 0. (Тобто, перші  $m$  моментів рівні 0, причому значення кількості моментів і вважається порядком вейвлета).

2. Функція швидко спадає при  $t \rightarrow \infty$ , де  $t$  – вісь часу.

Для точного вейвлет-аналізу застосовуються функції, що володіють достатньою гладкістю [11].

У загальному випадку вейвлет-перетворення функції  $f(t)$  виглядає так:

$$W(x, s) = \frac{1}{s} \int_{-\infty}^{\infty} \Psi^* \left( \frac{t-x}{s} \right) f(t) dt, \quad (3)$$

де  $\Psi$  – функція-вейвлет,  $t$  – вісь часу,  $x$  – момент часу,  $s$  – параметр, зворотний частоті, \* – комплексно-поєднане [10].

Використання перетворення Фур'є та вейвлет-перетворень допомагає виявляти аномалії, вивчати динаміку певних станів, таких як епілепсія чи аритмія, та вдосконалювати методи діагностики та моніторингу пацієнтів у медичній сфері.

Саме за допомогою результатів цих методів створюються світлові та аудіостимули в АВС.

## 3. Наявні дослідження та їхні результати

У дослідженні [12] було проведено аналіз опублікованих даних, який показав, що майже всі дослідження ефектів АВС, у тому числі [13–17], проводилися без врахування власних біоелектричних потенціалів піддослідного. У той же час у деяких публікаціях розповідається про дуже перспективний, але недостатньо вивчений підхід до організації різних видів стимуляції – методологія, коли сенсорні подразники автоматично підлаштовуються під поточні біоелектричні процеси [18–20].

Порівняння ефектів цих двох підходів до виконання АВС також було виконано у цьому дослідженні [12]. У ньому взяли участь 18 осіб (8 жінок і 10 чоловіків, 48–65 років), які скаржилися на стан психоемоційного напруження та стресу. Було проведено два експерименти.

У першому експерименті добровольцям було продемонстровано АВС, що складалася з світлових та аудіостимулів, які генерувалися трансформацією поточних параметрів ЕЕГ і серцевого ритму піддослідних. У другому експерименті використовували

задалегідь підготовлену композицію світлових та аудіостимулів без використання керуючих сигналів зворотного зв'язку від поточної біоелектричної активності піддослідних.

Було встановлено, що найбільш виражені зрушення об'єктивних та суб'єктивних показників (максимальне зростання потужності альфа-ритму ЕЕГ щодо фону, позитивні емоційні реакції та зрушення функціонального стану організму) відзначаються у випадках, коли управління АВС здійснюється безпосередньо власними електрофізіологічними характеристиками піддослідних [12].

Середні значення та стандартні відхилення параметрів у двох експериментах, рівень значущості  $p$  цих зрушень та достовірність відмінностей ( $p$ ) між зрушеннями в експерименті та в контролі наведені в табл. 1. Жирним шрифтом виділені значення зрушень з рівнем значущості  $p < 0,05$  [12].

Таблиця 1 [12]

Параметр	Експеримент				$p$ контр. /екс- пер.
	контроль		експеримент		
	зрушення	$p$	зрушення	$p$	
$\alpha$ -ритм ЕЕГ, відн. од.	<b>12,3 ± 23,5</b>	<b>0,042</b>	<b>22,9 ± 26,3</b>	<b>0,002</b>	0,272
Тест САН, самопочуття, бали	0,4 ± 4,6	0,725	<b>5,0 ± 3,2</b>	<b>0,001</b>	<b>0,002</b>
Тест САН, активність, бали	0,3 ± 4,1	0,782	2,3 ± 5,3	0,079	0,224
Тест САН, настрої, бали	1,1 ± 4,4	0,301	<b>4,7 ± 2,2</b>	<b>0,001</b>	<b>0,014</b>
Тест РЕД, емоційна дезадаптація, бали	-0,3 ± 1,8	0,523	<b>-0,7 ± 1,2</b>	<b>0,023</b>	0,367
Тест РС, рівень стресу, бали	-0,1 ± 1,5	0,762	<b>-0,9 ± 1,1</b>	<b>0,003</b>	<b>0,039</b>

Ці ефекти є результатом залучення інтероцептивних механізмів у систему механізмів, що відповідають за нормалізацію функціонального стану людини під впливом АВС, тобто механізмів мультисенсорної інтеграції, нейропластичності та резонансних механізмів мозку [12].

У публікації [21] було виконано огляд наявних публікацій (чотирнадцять статей (188 учасників) і два огляди літератури) та представлено узагальнення результатів, яке показує, як АВС впливає на пацієнтів з геміанопсією, а саме на зорову окоорухову функцію, функціональну здатність у повсякденному житті, геміаноптичну дислексію, зорове сканування

та пошукові завдання, підтримання функціональної здатності після тренування та вплив на мультисенсорну інтеграцію мозку за допомогою нейровізуалізації.

Гомогенна геміанопсія (ГГ) — це дефект поля зору, який визначається повною або частковою сліпотою в полях зору з правого або лівого боку обох очей, найчастіше спричинений інфарктом головного мозку [22].

Можливість викликати тривале покращення після АВС у дітей з хронічним ГГ внаслідок набутих уражень головного мозку вивчалася лише в одному дослідженні [23]. У ньому взяли участь троє дітей (один хлопчик і дві дівчинки віком від 9 до 17 років). Тренування тривали півтори години щодня протягом 4 тижнів. Показниками результату були правильна кількість візуальних розпізнавань, здатність до візуального пошуку та швидкість читання. Тест візуального пошуку складався з шести різних підтестів: яблуко, жаба, посмішка, літери Е-Ф, трикутник і числа [23]. Кожного учасника тестували до і після періоду тренування, а також після періоду спостереження протягом 1 місяця, а в одному випадку подальше спостереження було проведено через 12 місяців.

Автори побачили помітне поліпшення у виявленні та часі реакції лише тоді, коли суб'єкти використовували дослідницькі рухи очима, а не фіксували погляд на центральній точці [23]. Це свідчить про те, що покращення зорового сприйняття, викликане тренуванням, опосередковується окооруховою системою, що посилює орієнтацію в сліпому півполі. У всіх тестах основний фактор сесії був значущим, коли розглядався час відповіді [23].

Покращення швидкості читання після тренування спостерігалось для читання одного слова у всіх піддослідних. Результати цього дослідження підтвердили, що АВС може також сприяти активації зорової реакції окоорухової системи у дітей та підлітків з дефектами полів зору, оскільки поведінка зорового пошуку стала більш ефективною та швидкою після лікування.

Автори публікації [23] стверджують, що це свідчить про важливу роль мультисенсорної інтеграції в цьому типі очної компенсації і в пластичності зорової системи при наявності «сліпоти», навіть при повному пошкодженні потиличної кори. Довготривалий ефект лікування був зафіксований як в 1-місячних, так і в 12-місячних контрольних тестах [23].

#### 4. Програмне забезпечення

Програмне забезпечення (ПЗ) повинно проводити зчитування стандартизованого файлу ЕЕГ, відображати основні ритми мозку графічно, дозволяти апроксимувати та фільтрувати сигнал, мати змогу виділяти основні ритми головного мозку, транслювати

їх в обраний аудіофайл та генерувати синхронні світлові імпульси.

Крім того, повинні бути відображені дані з підключення апаратної частини та її готовності до роботи. Пропонована схема аудіовізуальної стимуляції представлена на рис. 3.



Рис. 3. Схема аудіовізуальної синхронізації мозкових хвиль

Оскільки головною метою програмного забезпечення є його універсальність для різних груп користувачів з різними порушеннями центральної нервової системи, в роботі з сигналами ЕЕГ використовуються

моноуральні ритми, які сприймаються незалежно від емоційного та фізичного стану особи. Це значно спрощує апаратну синхронізацію звуку та світла.

Розроблене ПЗ має дві програмовані частини – основну програму на ПК, яку було розроблено за допомогою MATLAB, та апаратну частину, реалізовану з використанням мікроконтролера Arduino Nano (рис. 4).

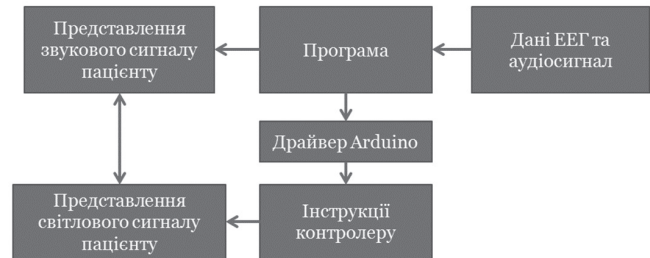


Рис. 4. Взаємодія компонентів

За допомогою лаконічного та мінімалістичного, але зрозумілого інтерфейсу (рис. 5) реалізована наочність характеристик сигналу, що буде оброблюватися.

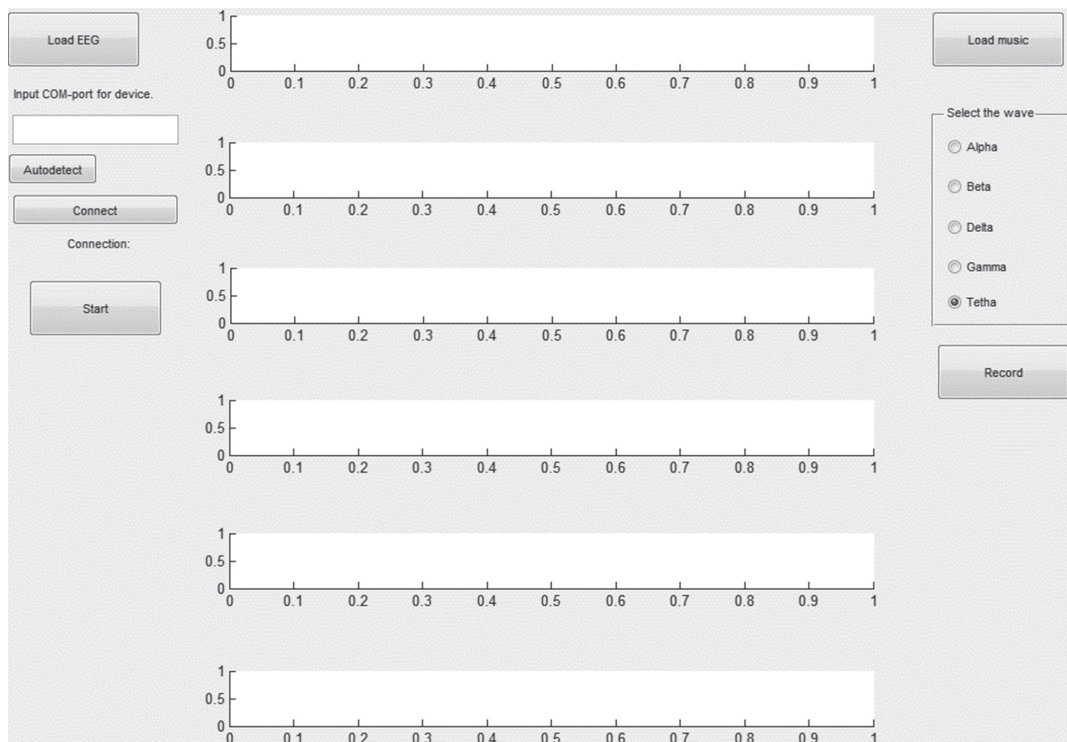


Рис. 5. Інтерфейс програми

Використання в якості носія сигналу моноуральних ритмів спрощує синхронізацію програми на ПК та на мікроконтролері, внутрішня пам'ять якого обмежена 16 Кб.

Перед трансляцією сигнал енцефалограми повинен бути відфільтрований та розділений на необхідні хвилі. Застосування фільтрації дозволить компенсувати спотворення даних, які неминуче виникають при використанні сучасних енцефалографів. Ці спотворення можуть бути обумовлені різними

факторами, від взаємного перекривання потенціалів відведень до мимовільних рухів пацієнта, тому необхідно підібрати гнучкий метод фільтрації, який можна буде використовувати незалежно від причини виникнення шуму.

Саме таким методом на цей час є вейвлет-аналіз. Для виконання фільтрації сигналу обрано сімейство вейвлетів Добеши, загальний вигляд носія яких представлено на рис. 6.

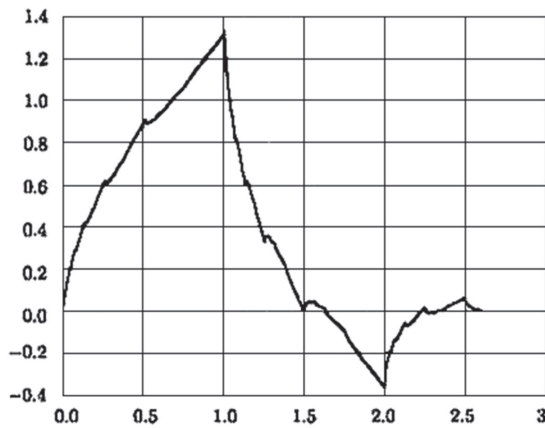


Рис. 6. Вейвлет Добеші 2-го порядку

Існує ряд процедур порогової обробки коефіцієнтів деталізації. Введемо в розгляд лише м'який і жорсткий трешолдинги, запропоновані Д. Донохо і І. Джонстоном (D. Donoho, I. Johnstone) на початку 90-х років.

Використання м'якого трешолдингу означає перерахунок коефіцієнтів деталізації  $\{d_{j,k} | k \in Z\}$  наступним чином:

$$d_{j,k}^s = (|d_{j,k}| - \tau) \text{sign}(d_{j,k}). \quad (4)$$

Тоді як процедура жорсткого трешолдингу не дає ніяких допусків на значення коефіцієнтів:

$$d_{j,k}^h = d_{j,k} I(|d_{j,k}| > \tau). \quad (5)$$

У розробці використано м'який трешолдинг згідно з формулою (4), оскільки деякі відхилення сигналів енцефалограми (що не є шумом, але достатньо малі, щоб бути відфільтрованими жорстким методом порогової обробки) мають значущий потенціал для корегування та синхронізації мозкових хвиль.

Після фільтрації необхідно здійснити виділення ритмів мозку (альфа, бета, гамма, дельта та тета). Для цього використовуються двосмугові фільтри Баттерворта.

Фільтр Баттерворта проектується таким чином, щоб його амплітудно-частотна характеристика була максимально гладкою на частотах смуги пропускання. Аналогічно до інших фільтрів, при розгляданні частотних характеристик використовують фільтр низьких частот, з якого легко можна отримати фільтр високих частот. За допомогою послідовної комбінації декількох таких фільтрів можна створити смуговий або режекторний фільтр.

Амплітудно-частотна характеристика  $G(\omega)$  фільтра Баттерворта  $n$ -го порядку може бути отримана з передавальної функції  $H(s)$ :

$$G^2(\omega) = |H(j\omega)|^2 = \frac{G_0^2}{1 + \left(\frac{\omega}{\omega_c}\right)^{2n}}, \quad (6)$$

де  $n$  — порядок фільтра,  $\omega_c$  — частота пропускання,  $G_0$  — коефіцієнт посилення нульової компоненти.

Таким чином, він здатний виділяти частоти, задані двома граничними значеннями, що дуже зручно для відокремлювання діапазону хвиль (наприклад, альфа-ритми мають основні частоти 8–13 Гц).

Отримані хвилі накладаються на аудіо- та візуальну компоненти стимуляції, що синхронно транслюватимуться користувачу.

В тестуванні роботи створеної системи прийняли участь 30 людей віком від 18 до 55 років, кожному з них було представлено вибір музики для обробки. На обрані звукові файли було накладено певні частоти ритмів мозку, і учасники мали прослуховувати музику протягом 14 днів по 10 хвилин на день, займаючись повсякденними справами чи знаходячись у розслабленому стані. Після завершення вони мали заповнити анкету.

Дані щодо вікового розподілу учасників представлено у табл. 2.

Таблиця 2

Віковий розподіл учасників

Віковий діапазон	Кількість опитаних	Відсоток
18–25	10	33
25–40	10	33
40–50	8	30
більше 50	2	4
<b>Всього</b>	<b>30</b>	<b>100</b>

Більшість учасників відмітили певний позитивний вплив скорегованої музики, що виражалось у підвищенні настрою, зменшенні рівня стресу та покращенні загального самопочуття. Однак частина з них не відчула жодних змін, що може бути пов'язано з індивідуальними особливостями сприйняття музичних стимулів. Інша частина учасників відмітила зміни як негативні. Дані про відчуття та їх розподіл серед учасників представлені в табл. 3.

Таблиця 3

Суб'єктивна оцінка впливу аудіовізуальної корекції

Оцінка змін	Кількість учасників	Відсоток
Позитивна	17	57
Негативна	3	10
Нейтральна	10	33
<b>Всього</b>	<b>30</b>	<b>100</b>

З отриманих результатів можна зробити висновок, що в цілому взаємодія виявилася вдалою, і алгоритми корекції мозкових хвиль, засновані на математичному апараті обробки сигналів, працюють чітко та коректно.

### Висновки

Аудіовізуальна стимуляція мозку на сьогоднішній день є перспективним напрямком в сфері медицини. Вона відкриває нові можливості для поліпшення психічного та фізичного здоров'я людини, а також діагностики та лікування різних порушень.

Було розглянуто біологічні ритми людини, які можуть бути використані для створення персоналізованих стимулів, та методи цифрової обробки отриманих сигналів. Розглянуті наявні дослідження дозволяють підтвердити важливість біоелектричної активності людини та підходів до її обробки у виконанні АВС для лікування різноманітних захворювань у дорослих та дітей.

Розроблене програмне забезпечення відображає характеристики сигналу енцефалограми та транслює енцефалограму у звуковий файл одночасно з виведенням світлового сигналу. В процесі отримання хвиль мозкової активності було застосовано вейвлет-перетворення та режекторні фільтри Баттерворта. Крім того, додатково було розроблене апаратне забезпечення у вигляді керованого з основної програми мікроконтролера для виведення світлового сигналу.

Отримані результати при подальших дослідженнях можуть бути використані для розробки перспективних методів персоналізованого впливу АВС для своєчасного усунення функціональних розладів та відновлення оптимального стану організму. Вони можуть бути корисними для медичних практиків, дослідників та пацієнтів, що шукають ефективні підходи до аудіовізуальної стимуляції мозку.

#### Список літератури:

- [1] *Дубровін, В. І.* Аудіовізуальна стимуляція головного мозку [Текст] / В. І. Дубровін, К. О. Роговська // Комп'ютерні технології: інновації, проблеми, рішення : І всеукраїнська наук.–техн. конф., м. Житомир, 17–18 травня 2015 р.: зб. тез. – Житомир: ЖДТУ, 2015. – С. 125–126.
- [2] *Твердохліб, Ю. В.* Биоакустическая коррекция психоэмоционального состояния человека [Текст] / А. В. Рашавченко, Ю. В. Твердохліб, В. І. Дубровін // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: VII міжнар. наук.–практ. конф., 17–19 вересня 2014 р., м. Запоріжжя : зб. тез. – Запоріжжя: ЗНТУ, 2014. – С. 220–221.
- [3] *Твердохліб, Ю. В.* Корекція роботи центральної нервової системи за допомогою біоакустичної терапії [Текст] / М. М. Захарова, А. В. Рашавченко, Ю. В. Твердохліб // Комп'ютерні технології: інновації, проблеми, рішення : І всеукраїнська наук.–техн. конф., м. Житомир, 17–18 травня 2015 р.: зб. тез. – Житомир: ЖДТУ, 2015. – С. 109–110.
- [4] *Rayi, A.* Electroencephalogram [Text] / A. Rayi, N. I. Murr. – Treasure Island (FL) : StatPearls Publishing, 2022.
- [5] Дослідження сукупного впливу стресового фактору та музичного сигналу на психофізичний стан людини [Текст] / [А. В. Паренюк, Д. В. Паренюк, К. С. Дрозденко та ін.] // Мікросистеми, Електроніка та Акустика. – 2021. – Т. 26, № 1. – С. 71–79.
- [6] *Купін, А. І.* Інтелектуальна система контролю уважності оператора на основі енцефалографа [Текст] / А. І. Купін, О. В. Градовий // Вчені записки ТНУ імені В.І. Вернадського. Серія: Технічні науки – 2021. – Т. 32 (71), № 2. – С. 149–157.
- [7] *Sattar, Y.* Electrocardiogram [Text] / Y. Sattar, L. Chhabra. – Treasure Island (FL) : StatPearls Publishing, 2023.
- [8] Електрокардіографія: клінічні можливості і досягнення в історичному аспекті : (огляд) [Текст] / [В. К. Ташук, О. С. Полянська, П. Р. Іванчук та ін.] // Клініч. анатомія та операт. хірургія. – 2015. – Т. 14, № 3. – С. 123–134.
- [9] *Конончук, Г. Л.* Вступ до Фур'є-оптики : навч. посіб. [Текст] / Г. Л. Конончук, В. М. Прокопеч, В. В. Стукаленко. – Київ : Видавничо-поліграфічний центр «Київський університет», 2009. – 320 с.
- [10] *Куракіна, Т. В.* Метод аналізу ЕКГ-сигналу на основі безперервного вейвлет-перетворення [Текст] / Т. В. Куракіна, Ю. В. Твердохліб // Сучасні інформаційні системи і технології. – 2013. – № 1. – С. 139–140.
- [11] *Дубровін, В. І.* Програмна реалізація методів спектрального аналізу нестационарних сигналів [Текст] / К. О. Роговська, Ю. В. Твердохліб, В. І. Дубровін // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: VII міжнар. наук.–практ. конф., 17–19 вересня 2014 р., м. Запоріжжя : зб. тез. – Запоріжжя: ЗНТУ, 2014. – С. 65–66.
- [12] Effects of Audio–Visual Stimulation Automatically Controlled by the Bioelectric Potentials from Human Brain and Heart [Text] / [A. I. Fedotchev, S. B. Parin, S. A. Polevaya et al.] // Hum. Physiol. – 2019. – Vol. 45, № 5. – P. 523–526.
- [13] Effect of Audiovisual Stimulation on the Psychophysiological Functions in Track-and-Field Athletes [Text] / [M. S. Golovin, N. V. Balioz, R. I. Aizman et al.] // Hum. Physiol. – 2015. – Vol. 41, № 5. – P. 532–538.
- [14] Integration of Functional, Psychophysiological, and Biochemical Processes in Athletes after Audiovisual Stimulation [Text] / [M. S. Golovin, N. V. Balioz, S. G. Krivoschekov et al.] // Hum. Physiol. – 2018. – Vol. 44, № 1. – P. 54–59.
- [15] Effects of Polymodal Rhythmic Sensory Influences on the Human CNS State and Autonomic Functions [Text] / [O. S. Glazachev, S. Ya. Klassina, O. V. Bobyleva et al.] // Hum. Physiol. – 2010. – Vol. 36, № 2. – P. 172–178.
- [16] Open-Loop Neurofeedback Audiovisual Stimulation: A Pilot Study of Its Potential for Sleep Induction in Older Adults [Text] / [H. Y. Tang, M. V. Vitiello, M. Perlis et al.] // Appl. Psychophysiol. Biofeedback. – 2015. – Vol. 40, № 3. – P. 183–188.
- [17] Open-Loop Audio-Visual Stimulation (AVS): A Useful Tool for Management of Insomnia? [Text] / [H. Y. Tang, B. Riegel, S. M. McCurry et al.] // Appl. Psychophysiol. Biofeedback. – 2016. – Vol. 41, № 1. – P. 39–46.
- [18] Closed-loop brain training: the science of neurofeedback [Text] / [R. Sitaram, T. Ros, L. Stoeckel et al.] // Nat. Rev. Neurosci. – 2017. – Vol. 18, № 2. – P. 86–100.
- [19] Clinical, hemispheric, and autonomic changes associated with use of closed-loop, allostatic neurotechnology by a case series of individuals with self-reported symptoms of post-traumatic stress [Text] / [C. H. Tegeler, J. F. Cook, C. L. Tegeler et al.] // BMC Psychiatry. – 2017. – Vol. 17, № 1 : 141.
- [20] Editorial: Closed-Loop Systems for Next-Generation Neuroprostheses [Text] / [T. Levi, P. Bonifazi, P. Massobrio et al.] // Front. Neurosci. – 2018. – Vol. 12 : 26.
- [21] *Alwashmi, K.* Audio-visual stimulation for visual compensatory functions in stroke survivors with visual field defect: a systematic review [Text] / K. Alwashmi, G. Meyer, F. J. Rowe // Neurol Sci. – 2022. – Vol. 43, № 4. – P. 2299–2321.
- [22] Cognitive training in an everyday-like virtual reality enhances visual-spatial memory capacities in stroke survivors with visual field defects [Text] / [L. B. Dehn, M. Piefke, M. Toepper et al.] // Top Stroke Rehabil. – 2020. – Vol. 27, № 6. – P. 442–452.
- [23] *Tinelli, F.* Audio-visual stimulation improves visual search abilities in hemianopia due to childhood acquired brain lesions [Text] / F. Tinelli, G. Purpura, G. Cioni // Multisens Res. – 2015. – Vol. 28, № 1–2. – P. 153–171.



В. В. Кирий<sup>1</sup>, О. В. Калиниченко<sup>2</sup>, О. В. Ляпота<sup>3</sup>

<sup>1</sup>ХНУРЕ, м. Харків, Україна, oleh.liapota@nure.ua

<sup>2</sup>ХНУРЕ, м. Харків, Україна, valentyna.kyryi@nure.ua

<sup>3</sup>ХНУРЕ, м. Харків, Україна, olga.kalynychenko@nure.ua

## ОСНОВНІ ПРИНЦИПИ ПРОЕКТУВАННЯ ТА РОЗРОБКИ СИСТЕМ ГЕНЕРАЦІЇ ЗВІТНОСТІ МЕДИЧНИХ ЗАКЛАДІВ

Розглянуто базові принципи проектування та розробки систем генерації звітності медичних закладів. Знайдено проблеми та запропоновано рішення щодо побудови архітектури систем: використання розподілених сервісів та методів управління ними для покращення продуктивності, застосування черг повідомлень для спілкування між ними. Розглянуто загальні методи збереження даних та використання сховищ даних для спрощення схеми баз даних, що використовуються для ведення звітності медичних закладів. Запропоновано принципи та підходи для створення тих частин систем, що відповідають за кінцеву обробку даних та генерацію звітів з підтримкою різних форматів виведення документів.

АРХІТЕКТУРА СИСТЕМИ, ЗВІТ, МЕДИЧНИЙ ЗАКЛАД, СЕРВІС, СХОВИЩЕ ДАНИХ

**V.V. Kyrii, O.V. Kalynychenko, O.V. Liapota. Basic principles of designing and developing reporting systems for medical facilities.** The basic principles of designing and developing reporting systems for medical facilities are discussed. Problems are identified and solutions are proposed for building the architecture of systems: the use of distributed services and methods of managing them to improve performance, the use of message queues for communication between them. The general methods of data storage and the use of data warehouses to simplify the scheme of databases used for reporting by medical institutions are considered. The principles and approaches for creating those parts of the systems responsible for final data processing and report generation with support for various document output formats are proposed.

SYSTEM ARCHITECTURE, REPORT, MEDICAL FACILITY, SERVICE, DATA WAREHOUSE

### Вступ

Згідно з вимогами сучасної медичної діяльності, необхідність в створенні ефективних систем генерації звітів, які б враховували специфіку роботи медичних закладів, стала першорядною. Збір, аналіз, управління та візуалізація даних — це ті основні завдання, які вирішують ці системи, що забезпечують оперативне та точне інформування персоналу медичних закладів, влади та пацієнтів.

Використання систем генерації звітності найбільш поширене у приватних медичних закладах. Галузь приватних медичних закладів активно розвивається, наприклад, протягом 2019 року розмір приватного сектору медичних закладів в Україні збільшився на 28%. Активний розвиток галузі призводить до актуальності розробки новітніх систем, що використовуються у рамках відповідної діяльності.

Принципи побудови систем генерації звітів полягають в розробці сучасної архітектури системи, що враховує основні операційні процеси роботи медичних закладів; створенні сховища даних, де зберігаються та обробляються великі обсяги інформації; розробці підсистеми генерації звітів, яка включає в себе аналіз даних та формування звітів на їх основі, і це є основною метою використання системи генерації звітності.

Метою даної роботи є обґрунтування використання кращих практик під час проектування та впровадження інформаційних систем генерації звітності, що дозволить медичним закладам більш ефективно використовувати доступні ресурси, поліпшувати

якість обслуговування та підвищувати свою конкурентоспроможність.

### 1. Принципи створення архітектури системи

#### 1.1. Визначення вимог та принципів щодо створення архітектури системи

Архітектура системи генерації звітності медичних закладів має важливу мету — створити структуроване та логічне середовище обробки даних, яке ефективно використовуватиме ресурси та максимально спростуватиме процес генерації звітів. У світі, де кількість даних щодня зростає, правильно побудована архітектура може служити основою для довготривалої та стабільної роботи системи.

Важливість архітектури системи набуває ще більшої актуальності, коли ми говоримо про медичні заклади, де вимоги до точності та достовірності даних є особливо високими. Крім того, звіти таких установ часто мають значну складність та велику кількість сторінок, що вимагає високої продуктивності та оптимізації процесів на всіх рівнях системи.

Архітектура системи генерації звітності медичних закладів має забезпечувати гнучкість та масштабованість, що дозволить адаптуватися до змінюваних вимог та викликів. Вона повинна передбачити інтеграцію з різними джерелами даних, підтримку різних форматів звітів та мати можливість легкого покращення їхнього вигляду з можливістю підтримки стилів, що вказуються користувачами. Також важливим аспектом є обробка великих обсягів даних,



мінімізація помилок при цій обробці та швидкість генерації результуючих звітів.

Саме завдяки правильно побудованій архітектурі можливе ефективне використання комп'ютерних ресурсів, забезпечення швидкої та якісної генерації звітів, що можуть включати велику кількість даних та бути складними з точки зору структури та формату. Це вкрай важливо, адже в медичному середовищі час та точність інформації часто грають вирішальну роль.

Враховуючи це, можна висунути наступне твердження щодо важливості створення архітектури для системи генерації звітності: архітектура системи у такому випадку не лише технічна основа для оптимізації роботи з даними, вона є стратегічним інструментом, який допомагає медичним установам виконувати свої ключові задачі, покращувати ефективність роботи та забезпечувати високу якість послуг.

Тому для створення адекватної архітектури варто дотримуватись наступних принципів:

- виконувати ретельний аналіз потреб медичних закладів у цільовій країні;
- визначати вимоги до звітності;
- розглядати усі вірогідні джерела даних, що можуть бути запропоновані медичними закладами;
- враховувати обсяги та можливості обробки даних;
- брати до уваги динаміку розвитку медичної галузі на цільовому ринку та загальних технологій для роботи з даними.

Останній пункт особливо важливий для досягнення такого рівня системи, на якому вона змогла б без великих витрат адаптуватись до нових викликів і не втрачати своєї актуальності з часом.

### 1.2. Встановлення ключових компонентів системи

Системи, призначені для генерації звітності медичних закладів, загалом можна розділити на декілька окремих блоків, кожен з яких відповідає за окрему частину роботи:

- блок адміністрування: забезпечує контроль, налаштування та управління усією системою загалом; через цей блок можливе управління доступом до даних, налаштування внутрішніх параметрів генерації звітів, моніторинг стану системи та виконання нею усіх задач тощо;

- блок завантаження та первинної обробки даних: його завдання полягає в отриманні вхідних даних з різних джерел, їх валідація, фільтрація та первинна обробка; цей блок також забезпечує зберігання даних у необхідному для системи та зручному для подальшого використання форматі;

- блок генерації звітів: відповідає за обробку даних у відповідності до обраного типу звіту, що може включати агрегацію, аналіз даних та формування інформації для кінцевого звіту; цей блок використовує відповідні дані, отримані від попереднього блоку, та генерує фінальні звіти.

Кожен з трьох блоків може бути представлений у вигляді окремого модулю та розміщуватись на власному вузлі у системі (див. рис. 1), взаємодія з якими відбувається через будь-який користувацький інтерфейс, зокрема це може бути:

- графічний веб-інтерфейс;
- застосунок під будь-яку операційну систему стаціонарних пристроїв;
- застосунок під будь-яку операційну систему мобільних пристроїв;
- інтерфейс командного рядка.

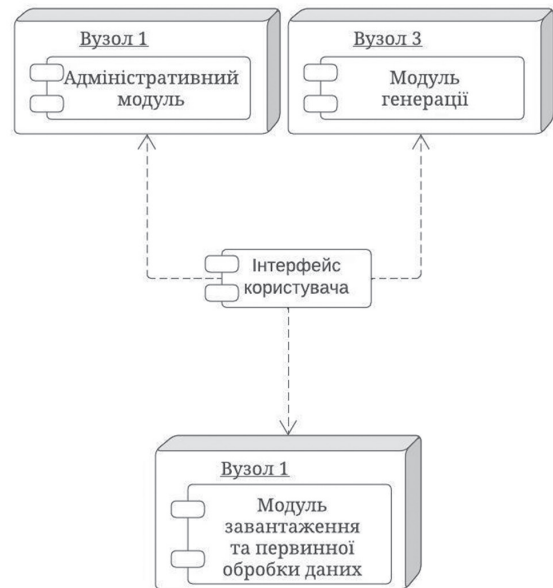


Рис. 1. Модулі на вузлах системи (2 рази вузол 1)

Завдяки використанню окремих вузлів для розміщення модулів системи: це можуть бути як повноцінні сервери, так і віртуальні машини або контейнери, – досягається їхня мінімальна зв'язність та можливість розширення кількості використовуваних вузлів за кожним окремим модулем (див. рис. 2).

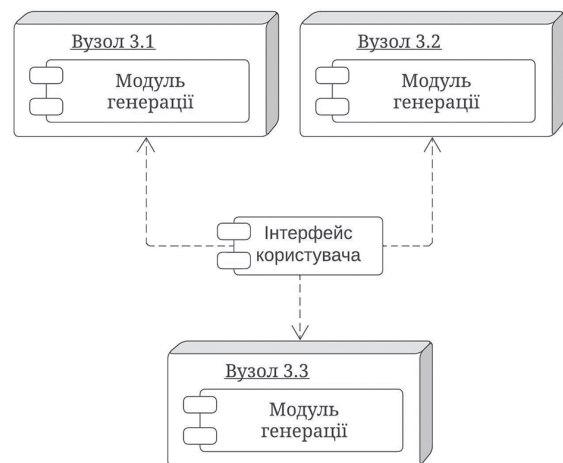


Рис. 2. Розширена кількість вузлів системи

Кожен з основних трьох модулів може складатись з власного набору сервісів, які також можуть розміщуватись на окремих вузлах. Зокрема, модуль генерації звітів може складатись з трьох частин:

- сервіс генерації та перевірки шаблонів для звітів;
- сервіс кінцевої обробки даних для звіту;
- сервіс генерації звітних документів.

Для зовнішньої взаємодії з цими частинами також створюється окремий програмний інтерфейс – Application Programming Interface (API).

Таким чином сервіс генерації звітів перетворюється на окремий повноцінний застосунок у рамках системи, сервіси якого також можуть існувати у декількох екземплярах (див. рис. 3) для покращення продуктивності та підтримки великої кількості користувачів та запитів до системи для генерації звітів.

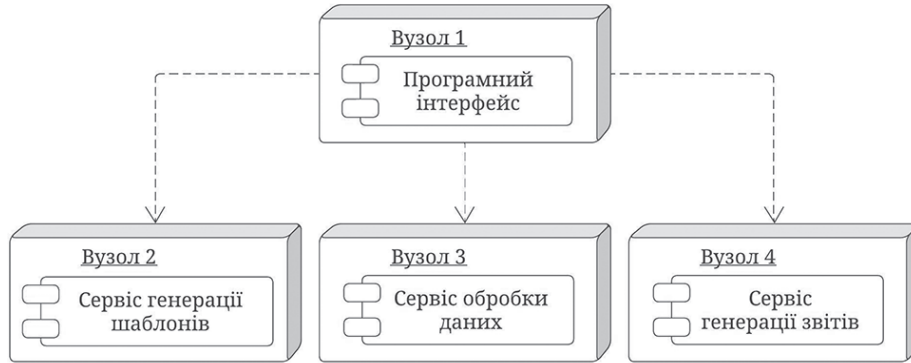


Рис. 3. Сервіси модулю генерації звітів

Таким чином досягається розділення системи на окремі сервіси, завдяки чому таку архітектуру можна вважати сервіс-орієнтованою – Service-Oriented Architecture (SOA) [1].

Наступні переваги [2] сервіс-орієнтованої архітектури реалізуються у системах генерації звітів медичних закладів:

- реалізується концепція слабо зв'язаних сервісів, що дозволяє з легкістю модифікувати або замінювати кожен з них;
- забезпечується економічна ефективність у випадку інтеграції вже існуючих відокремлених систем у нову;
- досягається зменшення бізнес-ризиків завдяки покращенню видимості окремих бізнес-операцій, що розподіляються між сервісами.

**1.3. Визначення зв'язків між компонентами системи**

Взаємодія між сервісами у такій системі, що складається в вище описаних компонентів, може виконуватись за певною низкою протоколів, кожен з яких має власні переваги та недоліки у порівнянні один з одним. Зокрема, найчастіше використовуються наступні:

- Hyper Text Transfer protocol (HTTP): базовий протокол, що використовується для клієнт-серверної взаємодії та виконується у форматі «запит-відповідь»;
- Simple Object Access Protocol (SOAP): протокол для доступу до об'єктів у мережі, що зазвичай використовується поверх HTTP протоколу, та на якому базується технологія веб-сервісів (веб-служб); являє собою XML структуру, що повинна відповідати усім вимогам кінцевого сервісу, до якого надсилається;
- Advanced Message Queuing Protocol (AMQP): протокол для передачі повідомлень між компонентами системи, заснований на ідеї використання брокера повідомлень як проміжної ланки між постачальником та споживачем повідомлень; за такої взаємодії

постачальник повідомлень відправляє їх брокеру на відомий відкритий API, брокер обробляє ці повідомлення та заносить їх у відповідну чергу, повідомлення у якій у свою чергу відстежує та обробляє споживач.

Саме останній із запропонованих протоколів більш доцільно використовувати у модулях, що потребують значного часу на обробку, зокрема з раніше наведених це модуль генерації звітів з його сервісами генерації шаблонів та, власне, звітів.

У такій системі брокером виступає програмне забезпечення RabbitMQ [3], яке реалізує взаємодію за протоколом AMQP.

У результаті отримуємо асинхронну генерацію кожної зі складових, що необхідні у процесі генерації документу зі звітом (див. рис. 4).

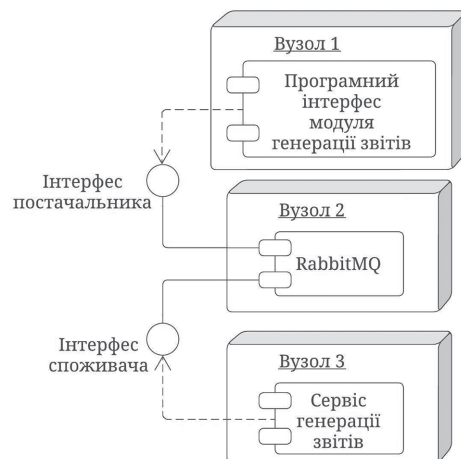


Рис. 4. Брокер повідомлень на схемі розгортання частини модулю генерації звітів

**1.4. Можливості подальшого покращення та розвитку архітектури**

Подальше покращення подібної архітектури можливе за рахунок продовження розділення модулів та

їхніх внутрішніх сервісів на менші сервіси за рахунок виділення окремих задач, які можуть вирішуватись самостійно в обмеженому контексті більш глобальних задач.

Так, модуль додавання та первинної обробки даних можна розділити на окремі сервіси додавання даних, кожен з яких буде підтримувати власне джерело даних, як от, наприклад, вручну введені користувачем данні, додані за допомогою завантаження файлів з даними або ж завантажені через відповідний програмний інтерфейс з іншого застосунку.

Такий розподіл допоможе отримати вже архітектуру, яку можна вважати мікросервісами. Та, відповідно, кожен з модулів загальної системи генерації звітності можливо представити у такому вигляді задля досягнення найменшої сфери застосування кожного з них з подальшою можливістю горизонтального розширення окремих сервісів, що мають найбільше навантаження та найчастіше використовуються користувачами під час роботи із завантаженням даних та генерацією аналітичних та статистичних звітів для покращення ефективності роботи медичних закладів.

## 2. Організація збереження даних

Основний принцип реалізації систем бізнес звітності, або звітності підприємства включає в себе процедури вилучення, перетворення та завантаження даних (ETL), сховище даних (data warehouse) для збереження даних та використання одного, або більше засобів генерації звітності.

### 2.1. Структура сховища даних

Структура сховища даних (див. рис. 5) містить кілька рівнів, кожен з яких відіграє свою роль.



Рис. 5. Фундаментальна структура сховища даних

Робота сховища даних розпочинається, коли відкрито одне чи декілька зовнішніх джерел даних або операційних баз даних. Дані з джерел вилучаються, очищуються та трансформуються, щоб отримати структуру, яка відповідає структурі сховища даних. За допомогою інструменту завантаження зібрані дані імпортуються в сховище даних, після цього сховище

даних оновлюється за допомогою інструмента оновлення, як впливає з його назви [4].

Система генерації звітності медичних закладів передбачає роботу по заповненню сховища даних, під час процесу закриття часового періоду. Як джерело даних використовується низка файлів та операційна база даних, що містить бізнес дані, зібрані у рамках відповідного медичного закладу, та зв'язані з ними системні данні (інформація про медичні заклади, відділи, параметри господарської діяльності відділів, часові періоди).

На етапі вилучення з бази даних отримуються бізнес дані та інформація зі зв'язаних таблиць. На етапі очистки та трансформації зі зв'язаних таблиць отримуються тільки ті дані, що необхідні для суб'єкту, тобто будь-якого зовнішнього інструменту, що у подальшому оперує цими даними.

У системі генерації звітності медичних закладів зовнішній інструмент представляє собою інструмент генерації звітності, інструмент отримує доступ до даних після проходження усіх попередніх етапів.

Інструмент завантаження додає у сховище бізнес дані та системні дані у разі відсутності відповідних. У разі наявності бізнес даних у системі інструмент оновлення модифікує ті дані, що уже наявні у сховищі. Відповідність бізнес даних визначається по набору зв'язаних з ними системних даних.

У системі наявна база, в якій зберігається інформація про файли з даними, які представляють собою зовнішнє джерело, та логи завантажень файлів. Наведена база використовується для моніторингу та адміністрування.

### 2.2. Компоненти архітектури сховища даних

Архітектура сховища даних складається з:

а) сервісу завантаження, який відповідає за збір даних із функціональних структур. Він змінює конфігурацію даних так, щоб залежні компоненти могли їх використовувати. Сервіс містить загальні інструменти та додатки, які повинні накопичувати дані з функціональних структур, налаштувати їх і потім передавати в сховище даних. В системі генерації звітності медичних закладів наявно дві реалізації сервісу завантаження: першого та другого рівня:

1) сервіс завантаження першого рівня відповідає за збір структурованих даних з файлу. Сервіс зчитує дані згідно певної структури, так званої мапи, відповідно до типу мапи менеджер взаємодіє з одною чи декількома таблицями операційної бази даних, завантажуючи рядки, та додаючи зв'язки між таблицями;

2) сервіс завантаження другого рівня відповідає за збір даних з операційної бази даних. Сервіс працює у рамках процесу закриття часового періоду. У ході роботи менеджера, зчитуються дані з таблиць які відносяться до певного часового періоду, зі зчитаних даних обираються орієнтовані на суб'єкт (зовнішній інструмент генерації звітності) данні, які подаються на завантаження до сховища даних.

– сервіс сховища даних організовує дані так, що кожен користувач, або програмний модуль, міг легко проаналізувати їх або знайти необхідну інформацію. Сервіс реалізує підтримку уніфікації та незмінності даних. У системі генерації звітності медичних закладів такий сервіс працює в тандемі з сервісом завантаження другого рівня: отримані у результаті роботи сервісу завантаження дані організовуються у сутності сховища даних та завантажуються у відповідні таблиці;

– сервіс запитів є останньою точкою підключення. Він дозволяє кінцевим клієнтам отримати доступ до даних у сховищі даних [5]. Цей сервіс у рамках системи генерації звітності медичних закладів реалізується шляхом генерації та виконання SQL запитів.

### 2.3. Реалізація процедур вилучення перетворення та завантаження даних

Основним підходом до реалізації процедур вилучення, перетворення та завантаження даних представляє собою:

– вилучення даних: техніка, необхідна для збору даних. Джерела даних включають у себе прості файли (.txt, .csv), або реляційні системи управління базами даних. У рамках системи генерації звітності медичних закладів вилучення даних виконується шляхом читування з файлів рядків, розділення їх на окремі фрагменти за відповідним символом роздільника, та процедури зв'язування відповідних фрагментів з сутностями операційної бази даних згідно з мапою, у разі невідповідності даних мапі витяг не виконується, щоб уникнути втрати даних;

– трансформація даних: для цього кроку розглядається кожна значуща можливість зміни та узгодження даних. Це стандартизує формат розрізнених даних для їх консолідації та об'єднання. На етапі трансформації даних у рамках системи генерації звітності медичних закладів, до даних, отриманих у рамках вилучення додаються дані, що залежать від контексту завантаження (інформація про користувача, клієнта, медичний заклад, часовий період), отримані у результаті структури даних завантажуються в операційну базу даних;

– очищення даних: процес, який реалізує відповідність даних вимогам суб'єкту, що зберігаються в сховищах даних. Точність даних є важливою, оскільки вони використовуються в динамічних циклах обробки даних, таких як генерація звітності. У рамках системи генерації звітності медичних закладів очищення даних виконується після вилучення їх з операційної бази даних. З вилучених даних видаляється інформація, яка не є необхідною для роботи суб'єкту (зовнішнього інструменту генерації звітності), після проходження очистки заповнюються структури, які відповідають вимогам таблиць сховища даних;

– завантаження даних: після виконання стадій витягу, трансформації та очищення даних, вони поміщуються до сховища даних. У рамках системи

генерації звітності медичних закладів завантаження даних реалізується через виконання SQL запитів, до реляційної бази даних, у якій зберігається сховище даних;

– оновлення даних: інструмент, який реалізує оновлення сховища даних в будь-який момент часу. У рамках системи генерації звітності медичних закладів оновлення даних, виконується у разі повторного завантаження даних, при цьому нові дані не замінюють старі, а доповнюють їх.

### 2.4. Розробка схеми сховища даних

Структура конфігурації схеми сховища даних характеризується зв'язками між fact і dimension таблицями.

Схема формату зірки (див. рис. 6) представляє собою найпростіший варіант побудови схеми сховища даних. Схема формату зірки має найбільшу, відносно інших реалізацій схеми сховища даних, швидкість виконання запитів. Dimension таблиці містять властивості певного виміру, корисні для кінцевого користувача дані, які не являються числовими даними. Fact таблиця містить посилання на dimension таблиці, числове значення, на базі якого реалізується робота кінцевого користувача, та міститься в центрі конфігурації.

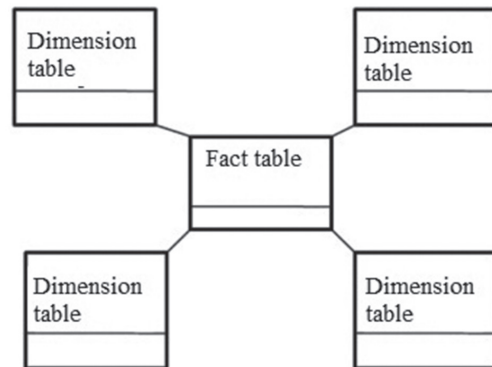


Рис. 6. Схема формату зірки

Для системи генерації звітності медичних закладів використовується сховище даних формату зірки (див. рис. 7).

У fact таблиці містяться посилання на dimension table та числові дані. У dimension таблиці department\_instances знаходиться інформація про відділи: назва стандартного відділу, назва конкретного відділу, коротка назва відділу, назва cost center. У dimension таблиці facilities знаходиться інформація про медичні заклади: назва, аббревіатура, поштовий індекс, основна адреса, допоміжна адреса, місто, контактний телефон, контактна електронна пошта. У dimension таблиці time\_periods наведено інформацію про часовий період: назву часового періоду, коротку назву часового періоду, дату початку, дату завершення, код типу (Q – квартальний, Y – річний). У dimension таблиці measures знаходиться інформація про міри: назву параметру господарської діяльності відділу.

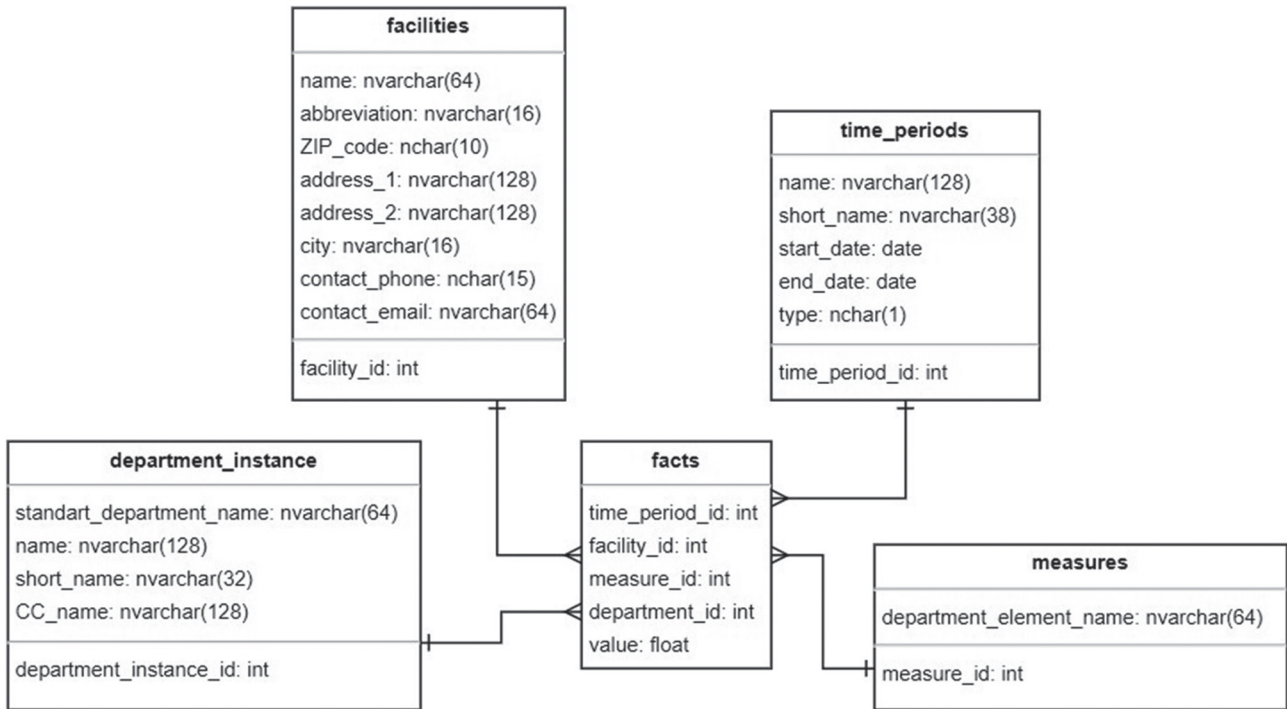


Рис. 7. Схема формату зірки для системи генерації звітності медичних закладів

Зовнішній інструмент генерації звітності може використовувати сховище даних для реалізації свого функціоналу.

### 3. Принципи обробки даних та генерації звітів

#### 3.1. Принципи генерації звітів

Звіти мають враховувати особливості та потреби даних, особливо при обробці великої кількості інформації. Для зменшення обсягу даних, з якими працюємо, можна застосовувати фільтрацію за певними параметрами. Наприклад, якщо ми створюємо звіт з контактною інформацією медичних закладів, ми можемо використовувати фільтри для відбору лише тих закладів, які відповідають заданим критеріям, наприклад, за стандартним відділом або порівняльною групою. Це допомагає зосередитись на суттєвих даних, які відповідають нашим потребам у звіті.

Гнучкість також є важливим аспектом при обробці даних для генерації звітів. Генерація звітів повинна бути гнучкою, щоб користувачі могли настроювати аналіз даних відповідно до своїх потреб та інтересів. Це означає, що ми можемо використовувати вибіркові змінні для аналізу даних, що дозволяє користувачам настроювати аналіз відповідно до їхніх потреб та інтересів. Наприклад, ми можемо вибрати лише заклади з певного міста або з конкретними контактними телефонами або електронною адресою. Такий персоналізований підхід дозволяє отримати більш точні та релевантні результати аналізу, а також враховувати унікальні потреби користувачів при генерації звітів.

Важливою частиною процесу є візуальне представлення потоку даних, яке може бути подібним до використання тегів у системі автоматизованої генерації

звітів (Automated Report Generation System - ASGTR) [6]. Це допомагає проілюструвати, як саме відбувається обробка даних та їх потік з бази даних. Рисунок 8 надає візуальне уявлення про цей потік даних.



Рис. 8. Потік даних в системі генерації звітності медичних закладів

Спочатку аналізатор документу перевіряє кожен з підходів обробки даних на його необхідність, що задана попередньо користувачем. Це дозволяє врахувати особливі потреби користувача та забезпечити виконання лише необхідних підходів. Наприклад, якщо користувач не обрав фільтрацію даних, аналізатор пропускає цей підхід і переходить до наступного етапу обробки документу. Таким чином, користувач отримує інформацію лише згідно обраного шаблону для генерації звіту.

Дані, що оброблюються у процесі генерації звіту, можуть включати різні типи даних, залежно від конкретної задачі та потреб користувача. Основними типами даних, що можуть бути оброблені, це дані про різні відділи, які складаються з різних закладів. Інформація про окремі медичні заклади, включаючи назву, адресу, місто, поштовий індекс, контактні дані (телефон, електронна пошта) та інші релевантні деталі.

Також оброблювальними даними є дані про групу закладів, які використовуються для порівняння. Деякі різні показники, які використовуються для вимірювання та аналізу, такі як активи, зобов'язання, витрати, доходи, обсяги, години роботи тощо. Усі показники даних мають різний часовий період, тобто це розподіл даних за певний період часу, наприклад, за місяць, квартал, рік тощо.

Інші параметри фільтрації та вибору – користувач може обирати додаткові параметри для фільтрації та вибору даних, наприклад, стандартний відділ (якщо є), тип порівняльної групи тощо.

Одним з важливих методів аналізу даних в системі генерації звітів є використання перцентилів. Перцентиль [7] – це статистичний термін, що використовується для визначення значення, нижче якого певний відсоток спостережень знаходиться. Наприклад, перцентиль 75% вказує на значення, яке перевершує 75% спостережень, а 25% спостережень менше за нього.

Перцентилі можуть бути корисними при аналізі даних медичних закладів. Наприклад, можна використовувати перцентилі для визначення медіани (50% перцентиль), що вказує на середнє значення в розподілі даних. Також можна аналізувати верхні та нижні квартилі (25% та 75% перцентилі відповідно), що дозволяють оцінити розмах та розподіл даних.

Наприклад, при аналізі фінансових показників медичних закладів можна використовувати перцентилі для визначення верхньої та нижньої границі для доходів або витрат. Це дозволить ідентифікувати заклади, які виходять за межі звичайного розмаху та потенційно потребують додаткового аналізу.

Перцентилі також можна застосовувати до даних про часові періоди, щоб оцінити розподіл та тренди в часі. Наприклад, використання перцентилів дозволить ідентифікувати певні місяці або квартали, коли показники даних медичних закладів були особливо високими або низькими, що може свідчити про сезонні варіації або зміни тенденцій.

Таким чином, використання перцентилів у системі генерації звітів медичних закладів дозволяє отримувати більш детальний та об'єктивний аналіз даних, встановлювати розмах та розподіл показників, а також ідентифікувати варіації в часі та між різними закладами.

Шаблон документа є важливою складовою процесу генерації звітів. Він представляє собою набір

інструкцій та налаштувань, які використовуються для створення звіту. Шаблон визначає, які дані включати в звіт, який формат і структуру має мати звіт, а також які візуальні елементи, такі як графіки, діаграми та таблиці, використовувати для представлення даних.

Використання шаблонів має декілька переваг. Воно дозволяє стандартизувати формат звіту, що полегшує сприйняття та розуміння інформації користувачем. Крім того, шаблони допомагають забезпечити послідовність та консистентність у генерації звітів, особливо якщо використовуються у командній роботі або для повторюваних завдань.

Один з основних аспектів шаблону документа – це його візуальне оформлення. Графічна візуалізація даних, така як графіки, діаграми та таблиці, є важливим елементом звіту, оскільки допомагає легше сприймати та розуміти дані. Візуальні елементи допомагають встановлювати зв'язки між даними, виділяти тенденції та ілюструвати важливу інформацію. Використання графічної візуалізації сприяє ефективному аналізу результатів, полегшує прийняття рішень та покращує комунікацію з зацікавленими сторонами.

Збереження шаблонів та звітів у базі даних MongoDB є одним із популярних варіантів. MongoDB є нереляційною базою даних, яка зберігає дані у вигляді документів у форматі BSON (Binary JSON). Ця база даних використовує гнучку схему (див. рис. 9), що дозволяє зберігати різноманітні дані без необхідності дотримуватися фіксованої схеми таблиць, як у традиційних реляційних базах даних.



Рис. 9. Схема бази даних MongoDB

MongoDB [8] пропонує API для збереження та отримання документів з бази даних, яке може бути використано для збереження шаблонів документів та звітів. Кожен шаблон документа може бути представлений як окремий документ у колекції шаблонів, яка містить необхідну структуру та налаштування для генерації звіту. Аналогічно, створені звіти можуть бути збережені у колекції звітів з відповідними даними та візуальним оформленням.

Дана база даних також надає можливості для розширеного пошуку та запитів до бази даних, що полегшує виконання пошукових запитів для знаходження певних шаблонів чи звітів за різними критеріями.

При аналізі великої кількості даних важливо мати ефективний спосіб візуалізації результатів. Навіть якщо інформація є розрахунками або статистикою, просте надання її у вигляді чисел або тексту може бути складним для сприйняття та розуміння. Графічна візуалізація даних допомагає представити результати аналізу в зрозумілій та доступній формі. Вона використовує графіки, діаграми, таблиці та інші візуальні елементи, що допомагають ілюструвати тенденції, встановлювати зв'язки та виділяти важливу інформацію. Візуалізація даних сприяє легшому сприйняттю результатів аналізу, полегшує процес прийняття рішень та покращує комунікацію між зацікавленими сторонами.

Крім того, генерація звітів може включати використання попередньо визначених шаблонів або створення власних шаблонів. Це дозволяє включати візуальні елементи в звіт, такі як графіки, діаграми, таблиці та інші візуалізації даних. Використання шаблонів допомагає стандартизувати формат звіту і зробити його більш зрозумілим для користувачів. Крім того, візуальні елементи допомагають ілюструвати важливу інформацію та встановлювати зв'язки між даними. Застосування шаблонів та візуалізація даних сприяють ефективнішому аналізу, зрозумілішому сприйняттю та кращому використанню результатів звіту у прийнятті рішень.

Після завершення обробки всього документу та застосування фільтрації та очищення даних з бази даних за обраними користувачем параметрами, створений документ може бути збережений у різних форматах, таких як текстові файли (txt), документи PDF, розширені таблиці (Excel) або XML-файли. Це дає користувачеві можливість обрати найбільш зручний формат для подальшого використання або розповсюдження звіту.

### Висновки

У роботі були розглянуті особливості роботи систем генерації звітності, зокрема пов'язаних з медичними закладами, та було виділено три основні складові, що є критичними для систем медичної звітності: створена архітектура, яка дозволяє з плином часу не втрачати актуальність з урахуванням сучасних темпів розвитку як медичної галузі, так і технологій загалом; методи ефективного збереження даних для їхнього подальшого використання у звітах; підходи до кінцевої обробки даних та генерації звітних документів.

У контексті побудови архітектури систем генерації звітності було розглянуто: розподіл таких систем на окремі компоненти та сервіси, що вибудовують сервіс-орієнтовану архітектуру; найкращі методи взаємодії між частинами системи, зокрема використання брокера повідомлень RabbitMQ, заснованого

на протоколі AMQP, який надає черги, у які одні сервіси додають завдання, а інші їх отримують та виконують у відповідності до поточної черги; також було вказано на можливість подальшого розділення окремих модулів системи на мікросервісні підсистеми.

Було розглянуто структуру сховища даних, та його компоненти, такі як сервіс завантаження, сервіс сховища даних, сервіс запитів; підходи до реалізації процедур вилучення, перетворення та завантаження даних, а саме витяг, очистка, трансформація, завантаження та оновлення даних; було розглянуто структуру сховища даних та обґрунтовано вибір структури.

Принципи обробки даних було розглянуто такі, що використовують візуалізацію даних: графіки, діаграми та таблиці, для представлення результатів аналізу у зрозумілій та доступній формі. Генерація звітів включає використання попередньо визначених шаблонів або створення власних, що допомагає стандартизувати формат звіту і полегшує сприйняття та використання результатів. Після обробки документа та генерації звіту, його можна зберегти у різних форматах для подальшого використання або розповсюдження.

### Список літератури:

- [1] Faten H. An Overview of Service Composition in Service Oriented Architecture // Modern Applied Science. – 2018. – № 8(12). – P. 172–178.
- [2] Kunjal B. Mankad, Priti S. Sajja. Utilization of Web Services for Service Oriented Architecture // Journal of Global Research in Computer Science. – 2010. – № 1(1). – P. 20–23.
- [3] Amar Ćatović, Nevzudin Buzadija, Samir Lemeš. Microservice development using RabbitMQ message broker // Science, Engineering and Technology. – 2022. – № 1(2). – P. 30–37.
- [4] Chaudhari S., Dayal U. An overview of data warehousing and OLAP technology // ACM sigmoid record – 1997. – №. 26. – P. 65-64.
- [5] Bhatia P. Data mining and data warehousing: principles and practical techniques // Cambridge University Press. – 2019. – P. 388-404.
- [6] Gjorgjevikj, D., Madjarov, G., Chorbev, I., Angelovski, M., Georgiev, M., Dikovski, B. // ASGRT – Automated Report Generation System // ICT Innovations. Communications in Computer and Information Science – 2010. – №1(83) – P. 369-376
- [7] Hyndman, R. J., & Fan, Y. Sample quantiles in statistical packages // The American Statistician. – 2016. – № 50(4). – P. 361-365.
- [8] Shannon Bradshaw. MongoDB: The Definitive Guide // O'Reilly Media – 2013. – № 50(4). – P. 361-365
- [9] Jane Smith. An Overview of Service Composition in Service-Oriented Architecture // Modern Applied Science, – 2018. – №8(12). – P. 172-178

Надійшла до редколегії 21.06.2024

# ПРАВИЛА оформлення рукописів для авторів науково-технічного журналу «БІОНІКА ІНТЕЛЕКТУ»

Науково-технічний журнал «Біоніка інтелекту» приймає до друку написані спеціально для нього оригінальні рукописи, які раніше ніде не друкувались. Структура рукопису повинна бути такою: індекс УДК, відомості про авторів, заголовок, анотації (на трьох мовах), ключові слова, вступ, основний текст статті, висновки, список використаної літератури, резюме.

Відповідно до Постанови ВАК України від 15.01.2003 №7-05/1 (Бюлетень ВАК, №1, 2003, с. 2), стаття повинна мати такі необхідні елементи: постановка проблеми у загальному вигляді та її зв'язок із важливими науковими чи практичними завданнями; аналіз останніх досліджень і публікацій і виділення не вирішених раніше частин загальної проблеми в даній області; формулювання цілей та завдань дослідження; виклад основного матеріалу досліджень з повним обґрунтуванням отриманих наукових результатів; висновки з даного дослідження та перспективи подальших досліджень у даному напрямку.

Статті мають бути виконані в редакторі Microsoft Word. Формат сторінки – А4 (210×297 мм), поля: верхнє – 25 мм, нижнє – 20 мм, ліве, праве – 17 мм. Кількість колонок – 2, з інтервалом між ними 5 мм, основний шрифт Times New Roman, кегль основного тексту – 10 пунктів, міжрядковий інтервал – множник (1,1), абзацний відступ – 6 мм. Обсяг рукопису – від 6 до 12 сторінок (мови: українська, англійська, російська та мовою оригіналу).

УДК друкується з першого рядка, без відступів, вирівнювання по лівому краю.

*ПІБ автора* (-ів), назва статті, назва та адреса учбового закладу необхідно надати повністю російською, українською та англійською мовами.

*Назва статті* друкується прописними літерами; шрифт прямий, напівжирний, кегль 12.

*Назви розділів* нумерують арабськими цифрами, виділяють жирним шрифтом. Відступи для назви статті, ініціалів та прізвищ авторів, відомостей про авторів, назв розділів, вступу та висновків, списку літератури: зверху – 6 пт, знизу – 3 пт.

*Анотації* (мовою статті, абзац 6–12 рядків, кегль 9) розміщують на початку статті, в ній має бути розміщена інформація про очікувані результати описаних досліджень (на трьох мовах).

*Ключові слова* (4–10 слів з тексту статті, які з точки зору інформаційного пошуку несуть змістовне навантаження) наводять мовою рукопису, через кому в називному відмінку, кегль 9.

*Рисунки та таблиці* (чорно-білі, контрастні) розміщуються у тексті після першого посилання у вигляді окремих об'єктів і нумерують арабськими цифрами наскрізною нумерацією за наявності більше ніж одного об'єкта. Невеликі схеми, що складаються з 3–4 елементів виконують, використовуючи вставку об'єкта Рисунок Microsoft Word. Більш складні виконують у графічних редакторах у вигляді чорно-білих графічних файлів форматів .tif, .jpg, .wmf, .cdr із

розділенням 300 dpi. Рисунки мають міститися у текстовому файлі й обов'язково подаватися окремими файлами з відповідними назвами (наприклад, рис1.jpg).

Усі елементи рисунка, включаючи написи, повинні бути згруповані. Усі написи в рисунках і таблицях мають бути виконані шрифтом Times New Roman, кегль у рисунках – 10, у таблицях – 9.

Рисунок повинен мати центрований підпис (поза рисунком), шрифт 9, відступи зверху і знизу по 6 пт. Ширина рисунка має відповідати ширині колонки (або ширині сторінки).

*Формули, символи, змінні* повинні бути набрані в редакторі формул **MathType**. Формули розміщують посередині рядка й нумерують за наявності посилань на них у рукописі. Шрифт – Times New Roman. Висота змінної – 10 пунктів, великих і малих індексів – 8 пт, основний математичний символ – 12 (10) пт. Змінні, позначені латинськими літерами, набирають курсивом, грецькі літери, скорочення російських слів і цифри – прямим написанням. Змінні, які є в тексті, також набирають у редакторі формул.

*Список літератури* вміщує опубліковані джерела, на які є посилання в тексті, укладені у квадратні дужки, друкують без абзацного відступу, кегль 9 пт, відступ зверху – 6 пт.

Після списку літератури з відступом зверху 6 пт зазначають *дату подання статті до редколегії*. Число та місяць задають двозначними числами через крапку. Розмір шрифту – 9 пт, курсив, вирівнювання по правому краю.

*Резюме* (Times New Roman, кегль – 10 пунктів,) подають англійською мовою: обсяг резюме до 2000 знаків (бажаний переклад). *Структура резюме: Background, Materials and methods, Results, Conclusion.*

Разом із рукописом (на аркушах білого паперу формату А4 щільністю 80–90 г/м<sup>2</sup>, надрукований на лазерному принтері) необхідно подати такі документи:

1. Заяву, яку повинні підписати всі автори.
2. Акт експертизи про можливість опублікування матеріалів у відкритому друці (якщо потрібно).
3. Рецензію, підписану доктором чи кандидатом наук.
4. Відомості про авторів.
5. Електронний варіант рукопису, резюме та відомостей про авторів.
6. Зробити оплату публікації.

Необхідно також зазначити один з наступних тематичних розділів, якому відповідає рукопис:

1. Теоретичні основи інформатики та кібернетики. Теорія інтелекту.
2. Математичне моделювання. Системний аналіз. Прийняття рішень.
3. Інтелектуальна обробка інформації. Розпізнавання образів.
4. Інформаційні технології та програмно-технічні комплекси.
5. Структурна, прикладна та математична лінгвістика.
6. Дискусійні повідомлення.



## ЗМІСТ

### **ОБ'ЄКТНЕ МОДЕЛЮВАННЯ ТА НЕЙРОМЕРЕЖЕВІ ТЕХНОЛОГІЇ**

<i>Неронов С. М., Плехова Г. А., Костікова М. В.</i> Удосконалення архітектури Neuronet автотрансферу.....	3
<i>Литвин В. В., Пелещак І. Р., Висоцька В. А., Мотика В. М., Пелещак Р. М.</i> Пасивний підхід до виявлення та класифікації мін із використанням магнітних полів і згорткових нейронних мереж .....	8
<i>Карнішен Б. С.</i> Покращення споживання енергії за допомогою ADAS в автономному водінні комунікаційної системи .....	15

### **ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ. МАШИННЕ НАВЧАННЯ. БАЗИ ДАНИХ**

<i>Терещенко Г. Ю., Кириченко І. В., Смеляков К. С., Олійник А. Є.</i> Аналіз методів стиснення зображень для зберігання у децентралізованих блокчейн сховищах .....	23
<i>Гольдінер Д. І., Матвієнко О. І.</i> Зменшення ймовірності відмови в системах масового обслуговування з обмеженою чергою із застосуванням пріоритизації за розміром та штучного інтелекту .....	36

### **ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ ТА ВІРТУАЛЬНІ МОДЕЛІ УПРАВЛІННЯ**

<i>Неронов С. М., Плехова Г. А., Костікова М. В.</i> Віртуальне управління перевізними процесами на автомобільному транспорті .....	43
<i>Плехова Г. А., Неронов С. М., Костікова М. В., Кашкевич С. О.</i> Удосконалення моделі безпечної маршрутизації в програмно-конфігурованих мережах.....	50
<i>Чалий С. Ф., Лещинський В. О.</i> Побудова пояснень на локальному рівні представлення процесу формування рішень для внутрішніх користувачів інтелектуальної системи .....	58

### **СТРУКТУРНА, ПРИКЛАДНА ТА МАТЕМАТИЧНА ЛІНГВІСТИКА**

<i>Дубровін В. І., Паничук К. С., Роговська К. О., Тарасова Ю. В.</i> Програмна реалізація методів аудіовізуальної стимуляції головного мозку .....	65
<i>Кирій В. В., Калиниченко О. В., Ляпота О. В.</i> Основні принципи проектування та розробки систем генерації звітності медичних закладів .....	72

### **ПРАВИЛА**

оформлення рукописів для авторів науково-технічного журналу «БІОНІКА ІНТЕЛЕКТУ» .....	80
---	----

*Наукове видання*

**БІОНІКА ІНТЕЛЕКТУ**  
**інформація, мова, інтелект**

**Науково-технічний журнал**

**№ 1 (100)**  
**2024**

Головний редактор — *Г. Г. Четвериков*  
Відповідальний редактор — *І. В. Кириченко*

Комп'ютерна верстка — *О. Б. Ісаєва*

Рекомендовано Вченою Радою  
Харківського національного університету радіоелектроніки  
(протокол № 10 от 28.06.2024)

Адреса редакції:  
Україна, 61166, Харків-166, просп. Науки, 14,  
Харківський національний університет радіоелектроніки, к. 127  
тел. 702-14-77, факс 702-10-13,  
e-mail: [bionics@nure.ua](mailto:bionics@nure.ua)

---

Підписано до друку 28.06.2024. Формат 60 × 84 <sup>1</sup>/<sub>8</sub>. Друк ризографічний.  
Папір офсетний. Гарнітура Newton. Умов. друк. арк. 9,5. Обл.-вид. арк. 9,3.  
Тираж 20 прим.

Віддруковано в редакційно-видавничому відділі ХНУРЕ  
61166, Харків, просп. Науки, 14.