МІНІСТЕРСТВО ОБОРОНИ УКРАЇНИ
Військовий інститут телекомунікацій та інформатизації
імені Героїв Крут

MINISTRY OF DEFENCE OF UKRAINE
Military Institute of Telecommunications and Informatization Technologies
named after Heroes of Kruty

Системи і технології зв'язку, інформатизації та кібербезпеки
№ 5

Communication, informatization and cybersecurity systems and technologies
№ 5

У збірнику викладено статті наукових та науково-педагогічних працівників, докторантів, ад'юнктів (аспірантів), курсантів, здобувачів інституту та інших установ (організацій) за наступними науковими напрямами:

перспективи розвитку телекомунікаційних систем, комплексів та засобів спеціального призначення;

захист інформації в спеціальних інформаційно-комунікаційних системах;

стан і розвиток автоматизованих систем управління військами та зброєю;

інформаційні системи та мережі, системи підтримки прийняття рішень спеціального призначення;

бойове застосування систем зв'язку та автоматизації Збройних сил України;

теорія і практика кібербезпеки та інформаційної боротьби в комп'ютеризованих системах і мережах.

Запрошуємо до співробітництва всі зацікавлені установи та організації, які проводять наукові дослідження та науково-технічні розробки за даними напрямами.

The book containes articles of scientific and teaching staff, post graduate students, adjuncts, institute applicants and other institutions (organizations) applicants in the following fields:

prospects of telecommunications systems, development, facilities and means of special purpose;

in special information protection and communication systems;

automated systems state and development of army weapons;

information systems and networks, decision support systems for special purposes;

combat use of communications systems and automation of Armed Forces of Ukraine;

theory and practice of cyber security and information warfare in computerized systems and networks.

All interested institutions and organizations, who conduct research and development in the directions state, are invited for cooperation.

Київ – 2024 – Kyiv

УДК 621
ББК Ц4 (4Укр) 39
З – 415

Системи і технології зв'язку, інформатизації та кібербезпеки: збірник наукових праць / за заг. ред. Г. Д. Радзівілова. Київ: Військовий інститут телекомунікацій та інформатизації імені Героїв Крут. 2024. № 5. 240 с. DOI: 10.58254/viti.5.2024.

ISSN 2786-6610

Статті, розміщені у збірнику, затверджені Вченою радою Військового інституту телекомунікацій та інформатизації імені Героїв Крут (протокол засідання № 16 від 28.05.2024 року).

Науковий профіль видання:
125 Кібербезпека;
126 Інформаційні системи та технології;
255 Озброєння та військова техніка

Засновник – Військовий інститут телекомунікацій та інформатизації імені Героїв Крут
(код за ЄДРПОУ 24978555).
Свідоцтво про державну реєстрацію видання: КВ № 25184-15124 Р від 20.07.2022.
Адреса редакції: 01011, м. Київ, вул. Князів Острозьких, 45/1. Тел. 256-22-73.
Електронна адреса: naukaviti@gmail.com

З М І С Т

C O N T E N T S

DOI: 10.58254/viti.5.2024.01.05

УДК 004.75

д-р філософії Бабарика А. О. ORCID: 0000-0001-8534-7764 (НАДПСУ)
канд. техн. наук Басараб О. К. ORCID: 0000-0002-2852-9534 (НАДПСУ)
Площик А. С. ORCID: 0000-0002-7034-1930 (НАДПСУ)
Табенський С. М. ORCID: 0000-0001-8771-5671 (НАДПСУ)

ВИБІР АРХІТЕКТУРИ ПОБУДОВИ СИСТЕМ ВІДЕОСПОСТЕРЕЖЕННЯ ДЕРЖАВНОЇ ПРИКОРДОННОЇ СЛУЖБИ УКРАЇНИ НА ОСНОВІ ХМАРНИХ ТА ТУМАННИХ ТЕХНОЛОГІЙ

Сучасні системи відеоспостереження еволюціонували від замкнутих систем телебачення до рівня складних систем, які функціонують у складі складних комплексних систем та вирішують задачі не лише фіксації подій в секторах огляду камер відеоспостереження, але і проводять аналіз отриманої відеоінформації. Збільшення обсягів інформації, що циркулює в сучасних системах, вимагає створення нових концепцій. В останні роки увагу науковців привернули технології розподіленої обробки інформації. Концепція хмарних обчислень (Cloud computing) стрімко розвивається, і основні ідеї, які були закладені при її побудові, вже не можуть вирішувати проблемні питання, які постають перед технологією. Це призвело до появи так званих постхмарних архітектур, які розширюють та доповнюють можливості хмарних обчислень. До таких архітектур відносять Mist, Edge, Fog тощо.

Метою статті є дослідження систем відеоспостереження, побудованих за концепцією хмарних і туманних обчислень, з метою подальшої реалізації у складі інформаційно-комунікаційних систем Державної прикордонної служби України. Під час проведення дослідження використано методи аналізу та узагальнення, моделювання, експеримент. Зазначений підхід дав можливість провести порівняльне дослідження систем відеоспостереження, побудованих за двома архітектурами. Отримані результати вказують на те, що технологія туманних обчислень має перевагу у зменшенні часу затримки, мінімізуючи потребу в повторних запитах до хмари, виконуючи обчислення на межі хмари. В моделі, побудованій за концепцією хмарних обчислень, застосовуються сервіси, які використовують ресурси хмари та призводять до збільшення навантаження на мережу. Разом з тим технологія туманних обчислень дозволяє розвантажити мережеве навантаження, виконуючи частину обчислень вузлами туману.

Результати експериментального дослідження показують переваги туманних обчислень для мереж, які чутливі до затримок. Проте, якщо розглянути систему відеоспостереження з основним завданням фіксації подій в секторах огляду камер відеоспостереження та можливості перегляду відеоінформації в режимі реального часу, то очевидно, що вузли туману не забезпечуватимуть тривалого зберігання відеоінформації, а затримки будуть некритичними. У випадку використання системи відеоспостереження з функціями відеоаналітики, вузли туману зможуть виконувати частину алгоритмів відеоаналітики, розвантажуючи водночас хмару. Тому актуальною задачею є дослідження ефективності побудови систем відеоспостереження з функціями відеоаналітики на основі туманної архітектури.

Ключові слова: *хмарні обчислення, туманні обчислення, відеоспостереження, моделювання, критерій, ефективність, iFogSim.*

A. Babaryka, O. Basarab, A. Ploshchyk, S. Tabenskyi Choice of architecture of video surveillance systems of State Border Guard of Ukraine based on cloud and fog technologies.

Modern video surveillance systems have evolved from closed-circuit television systems to the level of complex systems that operate as part of complex integrated systems and solve the tasks of not only recording events in the sectors of view of video surveillance cameras but also analysing the received video information. The increase in the amount of information circulating in modern systems requires the creation of new concepts. In recent years, the attention of scientists has been drawn to distributed information processing technologies. The concept of cloud computing is developing rapidly, and the basic ideas that were laid down in its construction can no longer solve the problems faced by the technology. This has led to the emergence of so-called post-cloud architectures that extend and complement the capabilities of cloud computing. These architectures include Mist, Edge, Fog, etc.

The purpose of the article is to study video surveillance systems based on the concept of cloud and fog computing and criteria for assessing their effectiveness. The research was conducted using the methods of analysis and generalisation, modelling, and experimentation. This approach made it possible to conduct a comparative study of video surveillance systems built on two architectures. The results obtained indicate that fog computing technology has the advantage of reducing latency, minimising the need for repeated requests to the cloud by performing calculations at the cloud edge. In a model based on the cloud computing concept, services that use cloud resources lead to an increase in

network load. At the same time, fog computing technology allows you to relieve the network load by performing part of the computation by fog nodes.

The results of the experimental study show the advantages of fog computing for networks that are sensitive to delays. However, if we consider a video surveillance system with the main task of recording events in the sectors of view of CCTV cameras and the ability to view video information in real time, it is obvious that fog nodes will not provide long-term storage of video information, and delays will not be critical. In the case of a video surveillance system with video analytics functions, fog nodes will be able to perform part of the video analytics algorithms, thus unloading the cloud. Therefore, an urgent task is to study the effectiveness of building video surveillance systems with video analytics functions based on fog architecture.

Keywords: *cloud computing, fog computing, video surveillance, modelling, criterion, efficiency, iFogSim.*

Постановка проблеми у загальному вигляді. Сучасні виклики та загрози вимагають нових підходів щодо їх протидії. У сфері забезпечення безпеки Державного кордону України особливої ваги набирають сучасні системи відеоспостереження, що розгортаються як на протяжних ділянках кордону, так і на військових об'єктах та територіях з обмеженим доступом.

Територіальна розподіленість систем відеоспостереження, а також можливість використовувати у їхньому складі мобільні складові на основі роботизованих систем, наприклад, БпЛА, вимагає від дослідників шукати нових підходів під час побудови таких систем. При цьому, за своєю функціональністю можуть використовуватись технології Інтернету речей (*англ.* IoT, Internet of Things).

В останні роки, зважаючи на зростання обсягів даних, які потребують обробки, увагу науковців привертають технології хмарних обчислень. Проте, притаманні традиційній клієнт-серверній архітектури недоліки, що пов'язані з проблемами масштабованості та надійності, можуть призвести до перевантаження серверів. Ці недоліки певною мірою вирішують постхмарні технології, такі як Mist, Edge, Fog тощо.

Для розробки моделей систем на основі хмарних та постхмарних концепцій, застосовуються різноманітні симулятори. Для оцінки ефективності функціонування таких систем використовують ряд метрик. Проте вказані метрики не завжди можуть адекватно відображати якісні характеристики систем, оскільки не враховують особливостей їхньої побудови.

Аналіз останніх публікацій. У роботі [1] проведено розширене дослідження критеріїв оцінки ефективності використання систем, побудованих на основі Cloud, Mist, Edge та Fog-архітектур. Також у вищевказаному дослідженні описано симулятори, що використовуються для моделювання систем за архітектурами Cloud, Mist, Edge та Fog.

Концепція туманних обчислень (Fog) за принципами побудови достатньо часто використовується під час побудови складних систем, таких як «Розумне місто».

Для вивчення систем, побудованих за Fog-архітектурою, дослідниками було розроблено ряд програмних симуляторів, таких як iFogSim [2–4], Edge-Fog [5], FogNetSim++ [6], YAFS [7], MobFogSim [8] тощо.

Описані авторами Harshit Gupta, Amir Vahid Dastjerdi, Soumya Kanti Ghosh та Rajkumar Buyya [4] можливості симулятора iFogSim, дозволяють провести порівняльне дослідження критеріїв оцінки ефективності функціонування систем відеоспостереження, побудованих за концепцією туманних та хмарних обчислень (Fog та Cloud).

Метою статті є дослідження систем відеоспостереження побудованих за концепцією хмарних і туманних обчислень, з метою подальшої реалізації у складі інформаційно-комунікаційних систем Державної прикордонної служби України.

Виклад основного матеріалу дослідження.

Основним завданням пристроїв в парадигмі IoT є періодичний збір даних з реального світу. [9; 10]. Пристроєм IoT може бути БпЛА з камерою відеоспостереження на борту.

Далі інформація з камери відеоспостереження може передаватися до місця відображення та/або зберігання.

У разі застосування технологій відеоаналітики, інформація може оброблятися обчислювальним блоком відразу на борту БпЛА, при цьому в мережу буде передаватись як відеопотік, так й інформація з результатами відеоаналітики. В іншому разі, при варіанті відеоаналітики на серверному обладнанні, інформація передається до місця зберігання, після чого автоматично обробляється, та результати роботи алгоритмів відеоаналітики відображаються користувачам.

Технології хмарних обчислень надають такі моделі послуг [11]:

- хмарне програмне забезпечення як послуга (Cloud Software as a Service, SaaS);
- хмарна платформа як послуга (Cloud Platform as a Service, PaaS);
- хмарна платформа як послуга (Cloud Infrastructure as a Service, IaaS).

SaaS – це готове програмне забезпечення, призначене для користувачів хмари, яке усуває потребу у використанні клієнтських програм, зберіганні даних, підтримці та адмініструванні програмних продуктів.

IaaS – відповідає за розміщення даних, конфігурацію мережі та обчислювальних потужностей. IaaS – це послуга хмарної системи, яка надається у формі платформи у віртуальному середовищі. Клієнтам не потрібно купувати сервери, центри обробки даних, мережеве обладнання або простір. Прикладом є Amazon EC2.

PaaS – відповідає за розвиток платформи з пристроями та компонентами для створення, тестування та запуску застосунків користувачів хмари [11]. PaaS є послугою хмарної системи, яка підтримує повний життєвий цикл програмного забезпечення та дозволяє користувачам розробляти хмарні додатки та сервіси. Програмісти та розробники не повинні купувати власне обладнання; замість цього вони використовують посередницьке обладнання та надсилають розроблені додатки клієнтам через інтернет. Google App Engine, Azure services platform від Google, реляційні служби баз даних Amazon (RDS) є прикладами послуги PaaS.

В останні роки дослідники розглядають також засновану на віртуалізації контейнерів модель послуг CaaS (Container as a Service). CaaS виник як хмарна модель для вирішення проблем розробки додатків у середовищі PaaS. Модель хмарного обчислення CaaS спрямована на розширення можливостей додатків, роблячи їх незалежними від специфікацій середовища PaaS. Прикладами моделі є CaaS Amazon EC2 Container Service (ECS) та Google Container Engine.

Хмарні архітектури можна класифікувати за рівнем належності та розміщенням елементів:

1. Публічна хмара;
2. Приватна хмара;
3. Хмара спільноти;
4. Гібридна хмара.

Публічна архітектура є найпоширенішою. Її використовують великі компанії, адміністративні установи й інші організації. Приватні хмари призначені для внутрішнього використання організаціями; водночас обладнання може бути фізично розташоване як у власника, так і бути під контролем третьої сторони. Громадські хмари призначені виключно для обчислень групами користувачів, які спільно виконують загальні завдання. Гібридні хмари складаються з двох або більше хмарних інфраструктур, які є окремими сутностями, але переносимість даних між ними забезпечується за допомогою стандартизованих технологій.

Централізована архітектура хмарних обчислень викликає проблеми у безпеці та затримках між пристроями та хмарою. Значною мірою нова парадигма туманних обчислень вирішує ці проблеми, доповнюючи хмарні послуги. Туманні обчислення розширюють хмарні

обчислення, включаючи кілька граничних вузлів, безпосередньо підключених до фізичних пристроїв.

Існує ряд переваг, пов'язаних із туманними обчисленнями, які забезпечують їхній успіх. Першою перевагою є зменшення мережевого трафіку, оскільки неконтрольоване збільшення мережевого трафіку може призвести до перевантажень і, як наслідок, до збільшення затримок. Туманні обчислення забезпечують платформу для фільтрації та аналізу даних, що генеруються датчиками, використовуючи ресурси периферійних пристроїв. Це значно зменшує трафік, що надсилається до хмари.

Значне зменшення затримки передачі даних є наступною важливою перевагою використання парадигми туманних обчислень, особливо для критично важливих додатків, які вимагають обробки даних у реальному часі. Одними з найкращих прикладів таких додатків є організація систем відеоспостереження з використанням мобільних роботизованих датчиків (БпЛА з камерами відеоспостереження). Саме тут на допомогу приходять туманні обчислення, які виконують обробку даних системи управління дуже близько до БпЛА, що робить можливим реагування в реальному часі.

Система на основі туманної архітектури являє собою невеликі обчислювальні вузли (Fog nodes), розташовані на граничному сегменті мережі. Головною метою технології є підтримка ресурсоємних інтернет-застосунків, які не потребують спеціальних вимог до часу затримки.

Основна різниця між туманними та хмарними обчисленнями полягає в тому, що хмара – це централізована система, а туман – розподілена децентралізована інфраструктура.

Туманні обчислення виступають посередником між обладнанням і віддаленими серверами. У туманному середовищі визначається, яка інформація буде відправлена на сервер, а яку можна редагувати локально. Так, туман можна розглядати як інтелектуальний шлюз, який розвантажує хмару, забезпечуючи більш ефективну роботу та аналіз даних.

Важливо зазначити, що туманне середовище не має окремої архітектури і не замінює хмарні обчислення, але доповнює їх, максимально наближаючись до джерела інформації [11]. Модель архітектури туманних обчислень показано на рисунку 1.

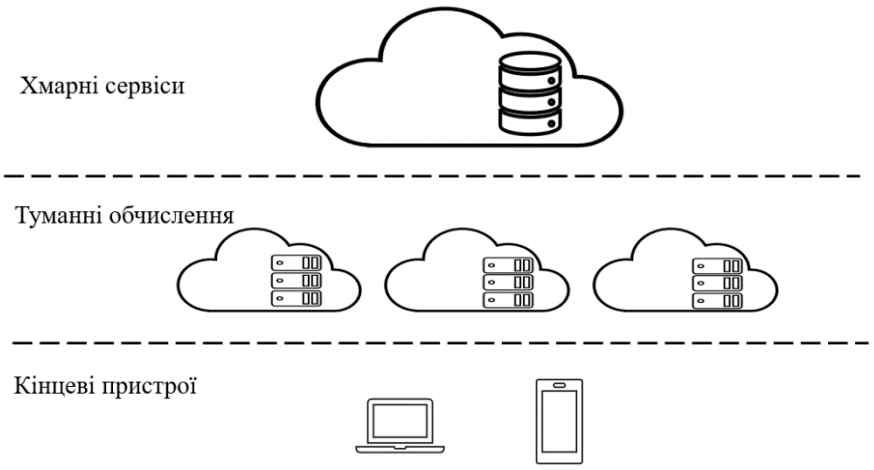


Рис. 1. Архітектура хмарних обчислень

Для дослідження різних методів управління ресурсами та планування, а також зв'язування, міграції та консолідації завдань моделей систем на основі хмарних архітектур необхідно мати інструментарій з можливістю проведення аналітики в режимі реального часу. Хоча реальне середовище Інтернету речей є бажаним тестовим майданчиком, у багатьох випадках воно є занадто дорогим та не забезпечує повторюваність і контрольованість

середовища. Для усунення цього недоліку дослідниками запропоновано симулятор під назвою iFogSim, який дозволяє моделювати управління ресурсами та політику планування додатків на периферійних і хмарних ресурсах за різних сценаріїв і умов.

Авторами Harshit Gupta, Amir Vahid Dastjerdi, Soumya Kanti Ghosh та Rajkumar Buyya в роботі [4], для реалізації функцій туманної архітектури використано базові функції симуляції подій з функціоналу існуючого симулятора CloudSim [13].

Основні класи iFogSim показано на рисунку 2 [4].

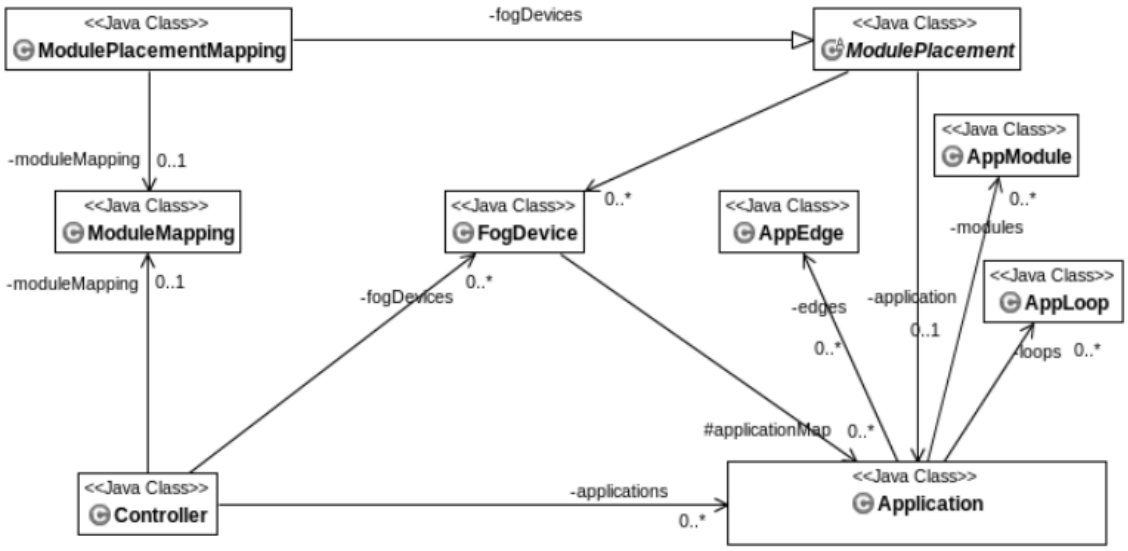


Рис. 2. Основні класи iFogSim

Реалізація iFogSim складається з імітованих сутностей та сервісів. Коротко опишемо основні поняття iFogSim, реалізовані у фізичній топології класів (рис. 3).

FogDevice – клас, який визначає апаратні характеристики Fog-пристроїв, їх підключення до інших Fog-пристроїв, датчиків та виконавчих механізмів. Основними атрибутами класу FogDevice є доступна пам’ять, процесор, розмір сховища, пропускна здатність лінії зв’язку. Методи цього класу визначають, як ресурси Fog-пристрою розподіляються між запущеними на ньому прикладними модулями, а також як модулі розгортаються і виводяться з експлуатації. Перевизначення цих методів дозволяє розробникам підключати власні політики для вищезазначених функцій.

Sensor – це сутності, які діють як датчики IoT. Клас містить атрибути, що являють собою характеристики датчика, починаючи від його підключення і закінчуючи вихідними атрибутами. Клас містить атрибут посилання на шлюзовий пристрій Fog, до якого підключено датчик, та затримку з’єднання між ними.

Tuple – це кортежі, які є основною одиницею зв’язку між сутностями в туманній архітектурі. Характеризується типом, модулями джерела та призначення. Атрибути класу визначають вимоги до обробки та довжину даних, інкапсульованих у кортежі.

Actuator – це клас, що моделює пристрій, визначаючи властивості його мережевого з’єднання. Атрибут у класі посилається на шлюз, до якого підключено Actuator, та затримку цього з’єднання. У класі визначено метод, що виконує дію при надходженні кортежу (Tuple) з модуля додатку (Application).

Application – додаток, який моделюється у вигляді орієнтованого графа, вершини якого є модулями, що виконують обробку вхідних даних, а ребра позначають залежності даних між модулями. Ці сутності реалізуються за допомогою класів AppModule, AppEdge, AppLoop.

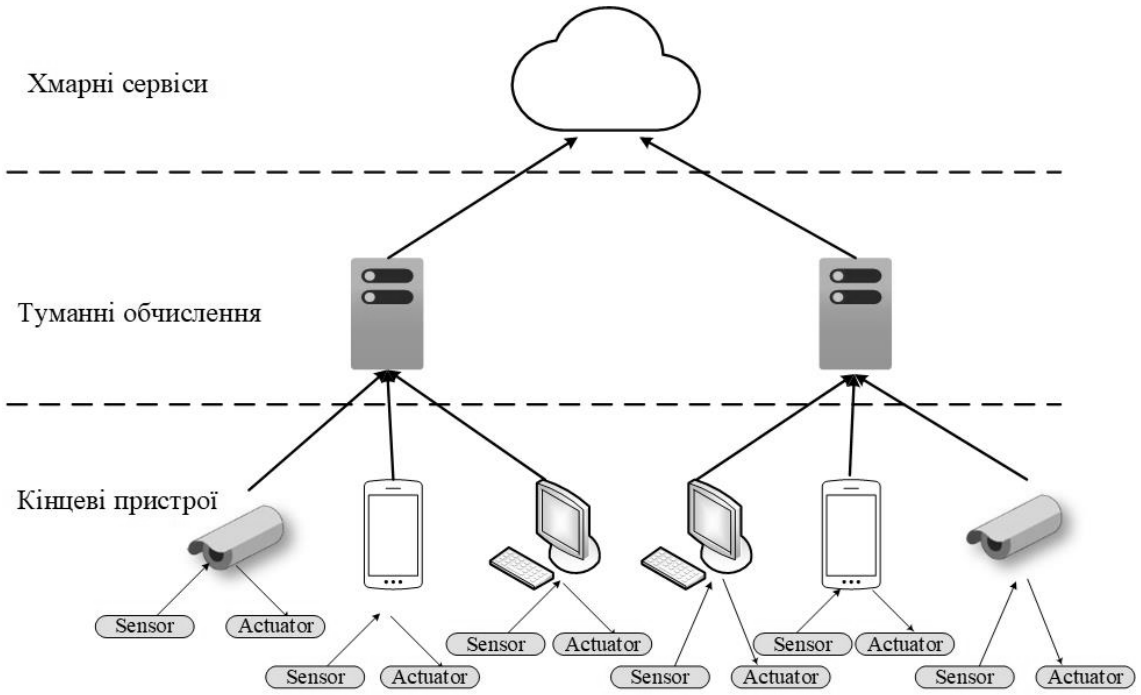


Рис. 3. Фізична топологія iFogSim

Після моделювання постає актуальне питання оцінки ефективності змодельованої системи. На основі проведеного в роботі [1] аналізу, авторами роботи виокремлено ряд критеріїв ефективності функціонування систем, побудованих за хмарними та постхмарними архітектурами (у т. ч. і за туманною). У своїй роботі автори наводять класифікацію на основі концепції MAPE-K, яка була запропонована компанією IBM [13; 14]. Відповідно до вказаної концепції критерії категорійовано за класами: моніторинг, аналіз, планування та виконання. Враховуючи особливості туманної архітектури, визначено критерії, за якими доцільно оцінювати ефективність систем, побудованих на основі туманної архітектури:

- використання ресурсів;
- навантаження на ресурси;
- термін використання ресурсу;
- максимальне навантаження на контейнер;
- час відгуку;
- час затримки/очікування;
- резервованість контейнерів;
- енергоспоживання;
- температура;
- надійність;
- безпека.

Використання ресурсів (*англ.* Resource Utilization) – визначається як відсоток ресурсів, який споживається вхідним навантаженням. Іншими словами, цей критерій показує наскільки завантажений процесор. Ресурсами можуть бути процесор, оперативна пам’ять, пам’ять, пропускна здатність тощо.

Ресурсне навантаження (*англ.* Resource Load) – визначається як відсоток навантаження на фізичний хост, віртуальну машину, контейнер тощо. Це вимір кількості завдань, що очікують на виконання в черзі на процесор, а також тих, що виконуються у цей момент.

Термін використання ресурсу (*англ.* Resource lifetime) – це критерій для обчислення часу обслуговування віртуальною машиною у хвилинах, годинах тощо. Тобто час, протягом якого ресурс віртуальної машини був орендований обчислювальним процесом [15].

Час відгуку (*англ.* Response time) – вказує на час від надходження завдання на вхід навантаження до повернення відповіді користувачеві. Час відгуку також називають часом виконання – це час, необхідний для повного виконання робочого обчислювального процесу. Він також відомий як час завершення – термін, який необхідний для конкретних хмар або завдань для завершення роботи [14]. Цей критерій є більш ефективним для оцінки продуктивності додатків та графічно орієнтованих робочих навантажень [16].

Час очікування/затримки (*англ.* Delay/latency) – це проміжок часу, протягом якого завдання очікує на виконання. Наприклад, якщо очікуваний час виконання завдання становить 2 секунди, а фактичний – 3 секунди, то час затримки для цього завдання становить 1 секунду. Цей критерій використовується для ситуації, коли фактичний час відгуку для такої задачі вже виміряний [17].

Енергоспоживання (*англ.* Energy Consumption) – визначається як кількість енергії, що споживається ресурсом для завершення виконання робочого обчислювального процесу. Цей критерій є ефективним для оцінки робочих навантажень при тестуванні продуктивності систем [17].

Температура (*англ.* Temperature) – визначається як кількість тепла, що генерується в обчислювальному середовищі базовою інфраструктурою, такою як центр обробки даних, при виконанні завдань. Ідеться про тепловиділення в центрі обробки даних, коли завдання виконуються на базовій інфраструктурі, що здебільшого стосується хмарних інфраструктур, які мають потужні обчислювальні ресурси, а не обмежені в енергоспоживанні пристрої Інтернету речей або периферійні вузли. Детальний опис критеріїв для хмарних обчислень з урахуванням температури можна знайти в [18].

Надійність (*англ.* Reliability) – визначається як здатність системи зберігати у часі в установлених межах значення всіх параметрів, які характеризують здатність виконувати потрібні функції в заданих режимах та умовах застосування. Надійність вузлів у віртуальному середовищі змінюється адаптивно. Вузол є надійним лише тоді, коли він може працювати в невизначених ситуаціях, таких як збій, інакше вузол не є надійним [17]. Цей критерій є найбільш ефективним для оцінки вебсайтів, сервісів зберігання та резервного копіювання, а також мобільних обчислювальних сервісів.

Змоделюємо систему відеоспостереження з використанням iFogSim за фізичною топологією на основі концепції Fog (рис. 4) та Cloud (рис. 5).

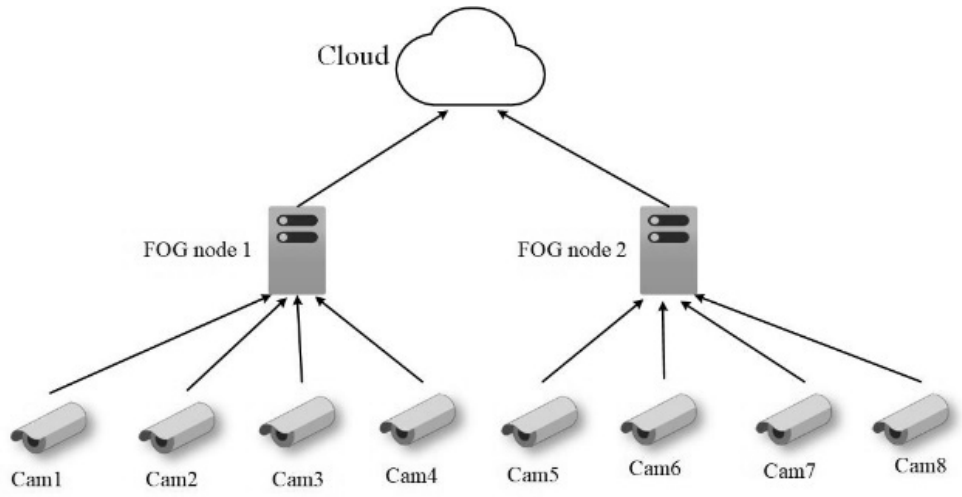


Рис. 4. Модель системи відеоспостереження за концепцією Fog

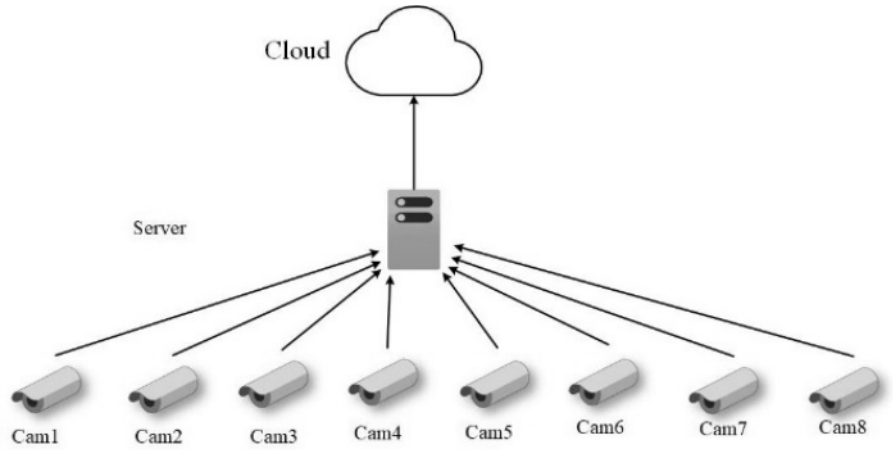


Рис. 5. Модель системи відеоспостереження за концепцією Cloud

Характеристики структурних елементів для моделі Cloud показано в таблиці 1.

Таблиця 1

Характеристики структурних елементів Cloud

Параметр	Cloud	Server	Cam X
CPU length (MIPS)	50000	4000	1000
RAM(MB)	40000	8000	1000
Downlink Bandwidth	10000	10000	
Uplink Bandwidth	100	10000	10000

Характеристики структурних елементів для моделі Fog показано в таблиці 2.

Таблиця 2

Характеристики структурних елементів Fog

Параметр	Cloud	Fog node X	Cam X
CPU length (MIPS)	50000	4000	1000
RAM(MB)	40000	8000	1000
Downlink Bandwidth	10000	10000	
Uplink Bandwidth	100	10000	10000

Використавши при моделюванні параметри, визначені в таблиці 1 та таблиці 2, скомпілюємо створений Java-файл, використовуючи різні змінні static int *numOfCamerasPerArea*: 4, 8, 16, 32, 64.

В результаті проведеного експерименту отримано дані, на основі яких сформовано таблицю 3 та таблицю 4. У вказаних таблицях результати експериментального дослідження наведено за критеріями *Latency* та *Network usage.*

Таблиця 3

Результати експериментального дослідження для моделі Fog

Кількість камер	Latency	Network usage
4	4.27	796.0
8	5.62	3518.08
16	8.31	3925.0
32	9.92	6425.20
64	12.6	12736.0

Таблиця 4

Результати експериментального дослідження для моделі Cloud

Кількість камер	Latency	Network usage
4	4.31	1512.0
8	6.11	3518.08
16	10.12	29125.2
32	302.54	63258.2
64	3221.45	73258.1

Критерій $Latency$ характеризується часом, який потрібний для передачі даних або виконання обчислень між кінцевим пристроєм та вузлом туману Fog Node, який розташований ближче до кінцевого пристрою, ніж обчислювальні ресурси хмари. Це важливий критерій ефективності, що впливає на загальну реакцію системи. Зменшення $Latency$ (1) в туманних обчисленнях допомагає покращити реакцію системи на події в реальному часі та забезпечити швидку і ефективну обробку даних:

$$Latency = \alpha + \mu + \varphi, \quad (1)$$

де α – затримка обчислювального пристрою Cam під час отримання та обробки зображення;

μ – час завантаження зображень на вузол туману та у хмару для обробки і зберігання;

φ –час, необхідний для відображення обробленої інформації на дисплеї монітору оператора системи відеоспостереження.

Критерій *Network usage* характеризує навантаження на мережу, яке виникає внаслідок збільшення трафіку.

Висновки і перспективи подальших досліджень.

Аналіз отриманих результатів (табл. 3, табл. 4) вказує на те, що в системах, побудованих за архітектурами Cloud та Fog, час виконання обчислювального процесу та навантаження на мережу зростає зі збільшенням кількості камер відеоспостереження. Проте, кількісні показники вказують на те, що технологія Fog має перевагу у зменшенні часу затримки, мінімізуючи потребу в повторних запитах до хмари, виконуючи обчислення на межі хмари. В моделі, побудованій за концепцією Cloud, порівняно з Fog, сервіси, які використовують ресурси хмари, призводять до збільшення навантаження на мережу. Проте технологія Fog дозволяє розвантажити мережеве середовище, виконуючи частину обчислень вузлами Fog nodes.

Результати експериментального дослідження показують переваги архітектури Fog для мереж, які чутливі до затримок. Але якщо розглянути систему відеоспостереження з основним завданням фіксації подій в секторах огляду камер відеоспостереження та можливості перегляду відеоінформації в режимі реального часу, то очевидно, що Fog nodes не забезпечуватимуть тривалого зберігання відеоінформації, а затримки будуть некритичними.

У разі використання системи відеоспостереження з функціями відеоаналітики, що є особливо актуальним у прикордонних підрозділах, вузли Fog nodes зможуть виконувати частину алгоритмів відеоаналітики, розвантажуючи при цьому хмару.

Тому, актуальною задачею є **подальші дослідження** ефективності побудови систем відеоспостереження з функціями відеоаналітики на основі туманної архітектури.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Performance evaluation metrics for cloud, Fog and edge computing: A review, taxonomy, benchmarks and standards for future research. *Internet Things* / Aslanpour Mohammad Sadegh et al. 2020. Vol. 12. P. 20–31.

2. Rajkumar B., Satish Narayana S. Modeling and Simulation of Fog and Edge Computing Environments Using iFogSim Toolkit. *Fog and Edge Computing: Principles and Paradigms*. 2019. P. 433–465.

3. iThermoFog: IoT-Fog based Automatic Thermal Profile Creation for Cloud Data Centers using Artificial Intelligence Techniques. *Internet Technology Letters* / Tuli Shreshth et al. 2020. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/itl2.198 (date of access: 20.02.2024).

4. iFogSim: A toolkit for modeling and simulation of resource management techniques in the Internet of Things, Edge and Fog computing environments Gupta. *Software: Practice and Experience* / Harshit Gupta, Amir Vahid Dastjerdi, Soumya Kanti Ghosh and Rajkumar Buyya. 2016. Vol. 47. P. 1275–1296.

5. Mohan N., Kangasharju J. Edge-Fog cloud: A distributed cloud for Internet of Things computations. *Cloudification of the Internet of Things (CIoT)*. Paris, France. 2016. P. 1–6.

6. FogNetSim++: A Toolkit for Modeling and Simulation of Distributed Fog Environment. *IEEE Access* / Qayyum T. et al. 2018. Vol. 6. P. 63570–63583.

7. Lera I., Guerrero C., Juiz C. YAFS: A simulator for IoT scenarios in fog computing. *IEEE Access.* 2019. Vol. 7. P. 91745–91758.

8. Tuli S., Mahmud R., Buyya R. Fogbus: A blockchain-based lightweight framework for edge and fog computing. *J. Syst. Softw*. 2019. Vol. 154. P. 22–36.

9. Evans D. The internet of things: How the next evolution of the internet is changing everything. 2011. 11 p. URL: https://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf (date of access: 20.02.2024).

10. Mattern F., Floerkemeier C. From the Internet of Computers to the Internet of Things. *From Active Data Management to Event-Based Systems and More*. 2010. P. 242–259.

11. Заковоротний О., Орлова Т. Порівняльний аналіз хмарних та туманних середовищ інтернету речей. *Системи управління, навігації та зв'язку*. Полтава: ПНТУ, 2023. Т. 3 (73). С. 152–154.

12. Cloudsim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience* / Calheiros R. et al. 2011. Vol. 41 (1). P. 23–50.

13. LARPA: A learning automata-based resource provisioning approach for massively multiplayer online games in cloud environments. *International Journal of Communication Systems* / Aslanpour M. S. et al. 2019. URL: https://onlinelibrary.wiley.com/doi/10.1002/dac.4090 (date of access: 20.02.2024).

14. CHOPPER: an intelligent QoS-aware autonomic resource management approach for cloud computing. *Cluster Computing /* Gill Sukhpal Singh et al. 2018. P. 1–39.

15. Calheiros R. N., Ranjan R., Buyya R. Virtual machine provisioning based on analytical performance and QoS in cloud computing environments. *Parallel processing (ICPP).* 2011. P. 295–304.

16. Gill S. S., Garraghan P., Buyya R. ROUTER: Fog enabled cloud based intelligent resource management approach for smart home IoT devices. *Journal of Systems and Software*. 2019. Vol. 154. P. 125–138.

17. Madni S. H. H., Latiff M. S. A., Coulibaly Y. Recent advancements in resource allocation techniques for cloud computing environment: a systematic review. *Cluster Computing*. 2017. Vol. 20. No. 3. P. 2489–2533.

18. ThermoSim: Deep learning-based framework for modeling and simulation of thermal-aware resource management for cloud computing environments. *Journal of Systems and Software* / Gill S. S. et al. 2020.

DOI: 10.58254/viti.5.2024.02.15

УДК 621.396:004.89

канд. техн. наук Бєляков Р. О. ORCID: 0000-0001-9882-3088 (ВІТІ ім. Героїв Крут)
д-р філософії Фесенко О. Д. ORCID: 0000-0002-2114-5327 (ВІТІ ім. Героїв Крут)

ОЦІНКА ЕФЕКТИВНОСТІ МЕТОДОЛОГІЧНИХ ПІДХОДІВ ІЄРАРХІЧНОГО ІНТЕЛЕКТУАЛЬНОГО УПРАВЛІННЯ НАЗЕМНО-ПОВІТРЯНОЮ КОМУНІКАЦІЙНОЮ AD-HOC МЕРЕЖЕЮ

У статті показано особливості процесів інтелектуального управління нового типу наземно-повітряних комунікаційних мереж, що стрімко впроваджуються. Динамічна природа умов функціонування та поведінки комунікаційних вузлів, як наземних, так і повітряних, обумовлює стрімке збільшення обсягів службової інформації, необхідної для забезпечення безперервного й адаптивного управління в режимі реального часу. Одним зі шляхів вирішення цієї проблеми є перерозподіл завдань управління на різних етапах циклу управління, який класично розподілений на етап планування, розгортання та оперативного управління. З одного боку збільшення ентропії на етапі планування ускладнює цей процес, однак такий підхід дозволить збільшити ймовірність прийняття «правильних» управлінських рішень з позиції управління якістю (мережевими метриками).

У роботі досліджується нова архітектура ієрархічної інтелектуальної системи управління наземно-повітряною комунікаційною мережею на основі безмодельного алгоритму Reinforcement learning, у якості мережевого агента Q-навчання та алгоритмів онлайн-послідовного екстремального машинного навчання FA-OSELM – агентів вузлового рівня. Представлено модель інтелектуальної системи управління, перевірено її адекватність, показано процес її навчання на етапі планування на різних моделях мобільності. Важливою особливістю процесу навчання інтелектуальної системи управління є застосування розробленої моделі мобільності, розкритої в статті, що на більш глибокому рівні описує процеси взаємодії комунікаційних вузлів. В роботі проведено дослідження репрезентативності навчальної вибірки, отриманої із використанням розробленої моделі мобільності щодо наявних, та визначено, що незважаючи на менший обсяг популяції вихідних даних вдалося забезпечити кращу якість управління ресурсами.

Ключові слова: *наземно-повітряна комунікаційна мережа, MANET, FANET, FA-OSELM, Q-навчання, інтелектуальна система управління, управління ресурсами, БпЛА, прогнозування, алгоритм машинного навчання, нейромережа.*

R. Bieliakov, O. Fesenko Evaluation of the effectiveness of methodological approaches to hierarchical intelligent control of ground-air Ad-Hoc communication network.

The article shows the features of intelligent management processes of a new type of ground-air communication networks that are rapidly being implemented. The dynamic nature of the conditions of operation and behavior of communication nodes, both ground and air, causes a rapid increase in the volume of service information necessary to ensure continuous and adaptive management in real time. One of the ways to solve this problem is the redistribution of management tasks at different stages of the management cycle, which is classically divided into the stages of planning, deployment and operational management. On the one hand, the increase in entropy at the planning stage complicates this process, but this approach will increase the probability of making "correct" management decisions from the standpoint of quality management (network metrics).

The work investigates a new architecture of a hierarchical intelligent ground-air communication network control system based on the model-free Reinforcement learning algorithm as a Q-learning network agent and FA-OSELM online sequential extreme machine learning algorithms as node-level agents.

The ICS model is presented, its adequacy is checked, and the process of its learning at the planning stage on various mobility models is shown. An important feature of the ICS training process is the application of the developed mobility model, disclosed in the article, which describes the interaction processes of communication nodes at a deeper level. The work conducted a study of the representativeness of the training sample, obtained using the developed mobility model, that was carried out relative to the existing ones, and it was determined that despite the smaller volume of the population of the initial data, it was possible to ensure a better quality of resource management.

Keywords: *ground-air communication network, MANET, FANET, FA-OSELM, Q-learning, intelligent control system, resource management, UAV, forecasting, machine learning algorithm, neural network.*

Вступ. Мобільні мережі наземного рівня (MANET) та літаючі мережі верхнього рівня (FANET) є особливими типами мереж, які не потребують фіксованої топології та здатні до самоорганізації, адаптації до змін відносно умов середовища. Мережі MANET та FANET

мають широкий спектр застосувань, такі як військові операції, рятувальні місії, сенсорні мережі, інтернет речей, автономні транспортні системи тощо. До ключових факторів, що впливають на продуктивність та надійність цих мереж, відноситься мобільність вузлів, яка визначає динамічність топології мережі, характеристики каналу зв'язку, доступність ресурсів та якість обслуговування. Тому важливо враховувати сучасні вимоги рівня представлення моделі мобільності, які відображають реальну поведінку комунікаційних вузлів у мережах, для аналізу та оцінки різних аспектів мережевої діяльності, таких як маршрутизація, безпека, енергоефективність тощо.

На сьогодні існує значна кількість наукових статей, присвячених розробці наукових методів та методик, моделей управління мережами, зокрема моделюванню процесів мобільності вузлів у мережах MANET та FANET [1–5]. Однак у більшості з цих робіт застосовується апроксимація на високому рівні, тобто сценарії руху мобільних користувачів (комунікаційних вузлів) описуються наближеною моделлю – як точка в довільній площині. Крім того, більшість існуючих моделей мобільності не враховують вплив рельєфу місцевості на опис руху повітряних вузлів (комунікаційних аероплатформ на БпЛА відносно мобільних користувачів наземної підмережі), що є важливим фактором для відображення реальних сценаріїв функціонування та інтеграції FANET до наземної підмережі, та управління наземно-повітряною мережею (НПМ) для забезпечення інформаційного обміну із заданим рівнем якості.

В наукових працях [6–15] окреслено принципи, методи, моделі та підходи інтелектуального управління наземно-повітряними комунікаційними мережами з використанням нейронних мереж та машинного навчання із підкріпленням та визначено, що для підвищення ефективності цього процесу необхідно збільшити глибину опису функціональної взаємодії мобільних комунікаційних вузлів наземної і повітряної підмережі шляхом моделювання на етапі планування мережі, тим самим підвищуючи умовну ймовірність прийняття «правильних» рішень на етапах розгортання та оперативного управління.

Таким чином, існує потреба в розробці та впровадженні нових методологічних підходів ієрархічного інтелектуального управління вузловими та мережевими ресурсами з урахуванням особливостей середовища функціонування на різних етапах циклу управління.

Аналіз літератури. Моделі мобільності можна класифікувати за різними критеріями, такими як рівень абстракції, тип мережі, тип вузла, тип руху, тип середовища тощо.

Так, найбільш популярною моделлю мобільності для дослідження мереж MANET [16] є Random Waypoint Model (RWP). Суть RWP складається з наступних допущень: вузол вибирає випадкову точку у заданій області, рухається до неї з випадковою швидкістю та зупиняється на довільний час, потім повторює процес. Модель RWP дозволяє моделювати різні сценарії мобільності, але має деякі недоліки, такі як нерівномірний розподіл вузлів, нереалістичність різких змін напрямку та швидкості руху [16].

Модель групового руху (Group Mobility Model, GMM). В основу GMM покладено умови, що вузли рухаються у групах, які мають спільну ціль, лідера, швидкість та напрямок руху. Кожна група може мати свою власну модель мобільності, яка визначає рух вузлів у межах групи. Модель GMM дозволяє моделювати більш реалістичніші сценарії мобільності. Однак ця модель має недоліки, такі як складність визначення параметрів груп, відсутність взаємодії між групами, низька гнучкість та адаптивність до змін середовища [17].

Модель City Section Mobility Model (CSM) [18]. Модель CSM визначає, що вузли рухаються за міськими дорогами, які задаються сіткою прямокутників. Вузли можуть змінювати напрямок руху на перехрестях, але не можуть виходити за межі сітки. Ця модель дозволяє моделювати реалістичні сценарії мобільності для міських мереж MANET, але має такі недоліки, як відсутність врахування рельєфу місцевості, обмеженість сітки, низька

гнучкість та адаптивність до змін середовища, та недоцільність застосування для моделювання дій підрозділів в умовах реальних бойових зіткнень.

Модель Roadway Mobility Model (RMM) [19]. Логіка моделі будується на тому, що вузли рухаються по трасах, які задаються картою доріг. Вузли можуть змінювати напрямок руху на розвилках, але не можуть виходити за межі карти. Модель RMM дозволяє моделювати реалістичні сценарії мобільності для міжміських мереж MANET, але має такі недоліки, як відсутність врахування рельєфу місцевості, обмеженість карти, низька гнучкість та адаптивність до змін середовища.

Модель Trajectory Mobility Model (TMM) [20]. Ця модель припускає, що вузли рухаються за заданими траєкторіями, які визначаються набором точок у просторі, які вузол повинен досягти. Вузли можуть змінювати напрямок руху на кожній точці, але не можуть відхилятися від траєкторії. Ця модель дозволяє моделювати реалістичні сценарії мобільності для мереж FANET, але має такі недоліки, як складність визначення траєкторій, відсутність взаємодії між вузлами, низька гнучкість та адаптивність до змін середовища.

Модель мобільності на основі карт (Map-based Mobility Model, MBM) враховує місцевість, представлену у вигляді цифрової карти. Вузли переміщуються за дорогами, стежками або іншими шляхами, визначеними на карті. MBM забезпечує більш реалістичне моделювання сценаріїв мобільності для мереж MANET та FANET, але має недоліки, такі як необхідність наявності детальних карт, обмеженість руху вузлів картою та низька адаптивність до динамічних змін середовища [21].

Модель мобільності на основі перешкод (Obstacle Mobility Model, OMM) враховує вплив перешкод, таких як будівлі, дерева або пагорби, на рух вузлів. Вузли обходять перешкоди, змінюючи напрямок або швидкість руху. OMM дозволяє моделювати реалістичні сценарії мобільності для мереж MANET та FANET в умовах складного рельєфу місцевості, але має недоліки, такі як складність визначення параметрів перешкод, відсутність взаємодії між вузлами та низька гнучкість [22].

Модель мобільності на основі потенціальних полів (Potential Field Mobility Model, PFMM) використовує концепцію потенціальних полів для моделювання руху вузлів. Кожен вузол має свій потенціал, який визначає його параметр тяжіння або відштовхування від інших вузлів або точок інтересу. PFMM дозволяє моделювати реалістичні сценарії мобільності для мереж MANET та FANET з урахуванням взаємодії між вузлами, але має недоліки, такі як складність визначення параметрів потенціальних полів та низька масштабованість [23].

Модель мобільності на основі машинного навчання (Machine Learning-based Mobility Model, MLMM) використовує алгоритми машинного навчання для прогнозування руху вузлів на основі історичних даних про їх переміщення. MLMM дозволяє моделювати реалістичні сценарії мобільності для мереж MANET та FANET з урахуванням індивідуальних особливостей поведінки вузлів, але має недоліки, такі як необхідність великого обсягу навчальних даних, складність налаштування параметрів алгоритмів та низька інтерпретованість результатів [24].

Для збору статистичних даних мобільності вузлів у мережах MANET та FANET в статті було вибрано три моделі: Random Waypoint (RWP), Reference Point Group Mobility (RPGM) та Random Direction (RD). Ці моделі вважаються найбільш поширеними для дослідження характеристик мобільності в різних сценаріях. Комбінація вище зазначених моделей дозволяє охопити широкий спектр сценаріїв мобільності та отримати статистичні дані про різні аспекти руху вузлів, такі як швидкість, напрямок, розподіл вузлів у просторі, час зупинки тощо. Вибір моделей RWP, RPGM та RD для збору статистичних даних про мобільність вузлів є обґрунтованим та доцільним для дослідження характеристик мереж MANET та FANET.

Таким чином, аналіз літератури показує, що існує дуже широкий спектр моделей мобільності для дослідження мереж MANET та FANET, кожна з яких має свої переваги та

недоліки. Вибір відповідної моделі залежить від конкретних вимог та умов застосування мережі. Однак більшість існуючих моделей не враховують вплив рельєфу місцевості на рух вузлів, що є важливим фактором для реалістичного моделювання сценаріїв FANET з використанням БпЛА. Тому розробка нової адаптованої моделі мобільності до функціонування в наближено реальних умовах, які б поєднували переваги існуючих підходів, для збору даних і навчання нейромережевих агентів вузлового та мережевого рівня на етапі планування є актуальним напрямком досліджень.

Метою статті є аналіз статистичних даних розробленої моделі мобільності мереж FANET-MANET на етапі планування, визначення репрезентативності навчальної вибірки порівняно з існуючими підходами, оцінка ефективності інтелектуального управління наземно-повітряною комунікаційною мережею з використанням алгоритму Reinforcement Learning – FA-OSELM.

Основні завдання дослідження полягають у наступному:

формалізація математичної моделі мобільності мережі FANET-MANET з урахуванням параметрів руху вузлів, таких як: швидкість, напрямок, висота рельєфу місцевості та зависання комунікаційних аероплатформ, потужність передачі, метод модуляції тощо;

опис імітаційної моделі мобільності мереж FANET-MANET, основна суть якої полягає в процесі генерації різних сценарії руху вузлів з урахуванням рельєфу місцевості, використовуючи тривимірні карти рельєфу;

математичний опис взаємодії агентів вузлового та мережевого рівнів;

збір та аналіз статистичних даних імітаційної моделі функціонування наземно-повітряної комунікаційної мережі Ad-Hoc (FANET-MANET), таких як: середня швидкість, середня відстань, середній час обміну повідомленнями, час переривання зв'язку, середня кількість змін напрямку, кількість сусідів, кількість стрибків, кількість переходів групи тощо;

здійснення оцінки продуктивності управляючих рішень агента мережевого рівня Reinforcement learning – FA-OSELM інтелектуальної системи управління (ІСУ) НПМ порівняно з існуючими моделями мобільності, такими як Random Waypoint, Random Direction, Reference Point Group Mobility, використовуючи метрики інформаційно-комунікаційного обміну пропускної спроможності, затримки та втрати пакетів.

Виклад основного матеріалу.

Модель мережі. Розглядається мережа MANET-FANET, яка складається з трьох рівнів: мобільних користувачів, мобільних базових станцій, елементів повітряної мережі комунікаційних аероплатформ на БпЛА.

Кількість вузлів наземно-повітряної мережі $N_{\text{НПМ}} \leq 80$, місцевість функціонування НПМ визначена завчасно (фізичні перешкоди та рельєф місцевості відомі); вузли наземної комунікаційної мережі (НКМ) розподілені за рангами $\text{P}_i^{\text{НМ}}, i = \overline{1,7}$, кожен ранг при цьому має свою обмежену ділянку функціонування на початковому етапі. В подальшому розташування вузлів НКМ може змінюватись випадковим чином та займати на деякий час іншу ділянку. Модель мережі НКМ проілюстрована на рисунку 1 [6].

Мережеву модель можна охарактеризувати наступними параметрами: $N_{\text{і}}$ – загальна кількість вузлів на ψ-му рівні, де $\psi = \overline{1,3}$, (x_i, y_i, z_i) – координати i-го вузла в тривимірному просторі, та $N = \sum_{\psi=1}^{3} N_{\text{і}}$, (v_i, a_i, θ_i) – швидкість, прискорення та напрямок i-го вузла, R_i – радіус радіодоступності i-го вузла, D_i – коефіцієнт діаграми направленості антени i-го вузла, E_i – енергія батареї i-го вузла, T_i – початкова таблиця маршрутизації найкоротших шляхів i-го вузла, λ_i – інтенсивність пакетних потоків i-го вузла, P_i – протокол доступу до каналу i-го вузла, де $i = \overline{1,N}$.

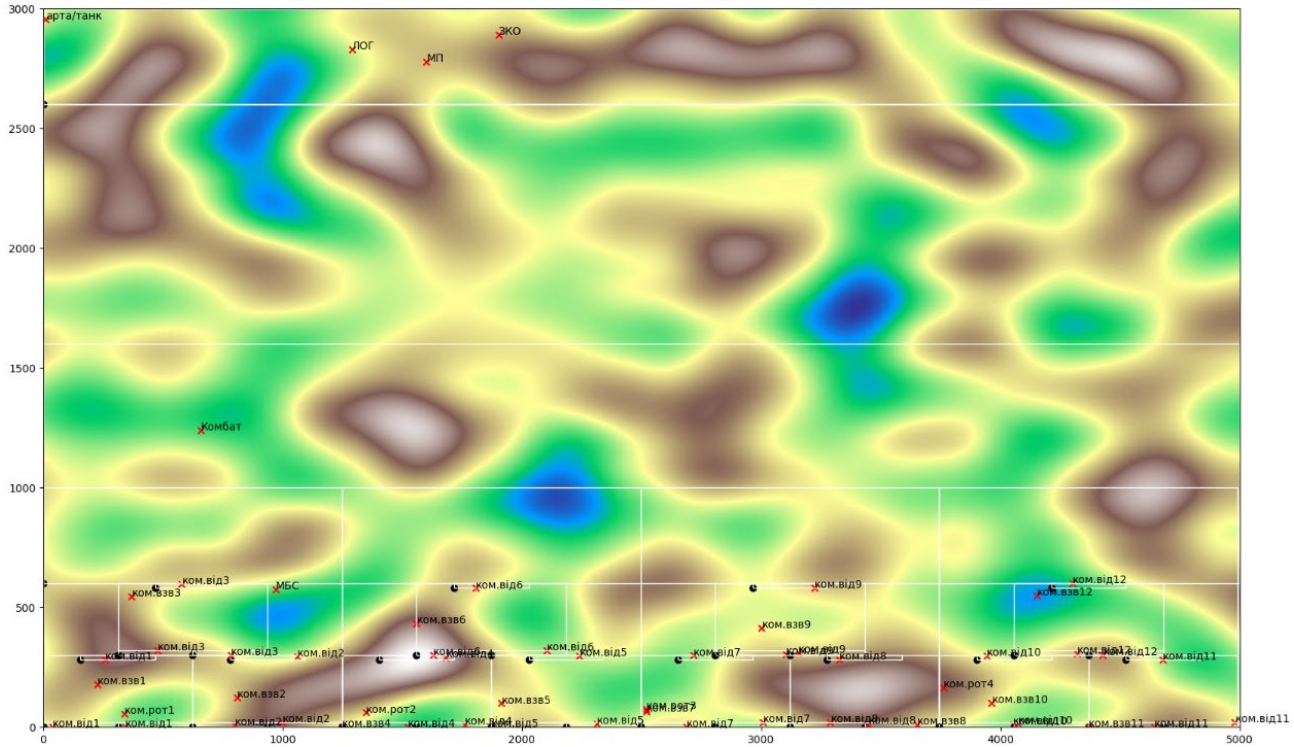


Рис. 1. Варіант початкового розміщення мобільних користувачів наземної компоненти НПМ

Початкове положення зависання комунікаційних аероплатформ мережі повітряного рівня визначено за критерієм рівномірного розподілу покриття зон функціонування мобільних користувачів, де кожна мобільна базова станція (МБС) управляє підгрупою комунікаційних аероплатформ (КА) за загальною методикою управління переміщенням, описаною в [13].

h – висота зависання БпЛА КА становить 700 метрів; d – відстань між КА становить до 800 метрів.

Врахування рельєфу місцевості. Для моделювання тривимірної карти рельєфу місцевості руху абонентів враховано вплив рельєфу місцевості на рух вузлів.

Висота рельєфу місцевості в точці з координатами (x, y) задається функцією $h(x, y)$, яка отримана з тримірної карти (висот) рельєфу.

Обмеження на висоту повітряного комунікаційного вузла в точці з координатами (x, y) задається функцією $g(x, y) = h(x, y) + H_{max}$, де H_{max} – максимальна відстань від поверхні землі, на якій вузол може підніматись.

Умова на висоту вузла в точці з координатами (x, y) задається нерівністю $h(x, y) \leq z \leq g(x, y)$, де z – висота вузла. Якщо повітряний вузол порушує цю умову, тоді він змінює свій кут зеніту ϕ так, щоб повернутися в допустимий діапазон висоти.

Врахування моделі мобільності вузлів мережі для процесу навчання нейромережевих агентів вузлового рівня.

Переміщення вузлів мережі описується наступним чином:

1. Координати вузла в момент часу t задаються вектором $x_t = (x_t, y_t, z_t)^T$, де x_t, y_t, z_t – географічна широта, довгота та висота вузла відповідно.
2. Швидкість вузла в момент часу t задається вектором $v_t = (v_t, \theta_t, \phi_t)^T$, де v_t – модуль швидкості, θ_t – азимутальний кут, ϕ_t – зенітний кут вузла відповідно.
3. Прискорення вузла в момент часу t задається вектором $a_t = (a_t, \alpha_t, \beta_t)^T$, де a_t – модуль прискорення, α_t – азимутальний кут, β_t – зенітний кут вузла відповідно.
4. Рух вузла в момент часу t описується наступними рівняннями:

$$x_t = x_{t-1} + v_{t-1}\Delta t + \frac{1}{2}a_{t-1}\Delta t^2,$$

$$v_t = v_{t-1} + a_{t-1}\Delta t,$$
$$a_t = f(x_t, v_t, a_{t-1}, U, Q_t),$$

де Δt – часовий крок; f – функція, що визначає прискорення вузла на основі його поточного стану, вагових коефіцієнтів U функції значення-дії Q_t.

Процес оцінки поточного стану вузла виражається функцією значення-дії нейромережевого агенту $Q_t(s,a)$ і відображає, наскільки задовольняє умова для переходу вузла в стан s для виконання дії a в момент часу t. Необхідно зазначити, що процес оновлення параметрів функції значення-дії на основі розробленого алгоритму Reinforcement Learning – FA-OSELM, який використовує негайний сигнал підкріплення r_t як дохід (суму підкріплень), отриманим вузлом від мережі, та коефіцієнту дисконтування γ, впливає на майбутній стан вузла в момент прийняття рішення. Основна ітераційна формула алгоритму Reinforcement learning – FA-OSELM (1):

$$Q_{t+1}(s,a) = (1-\eta)Q_t(s,a) + \eta\left[r_t + \gamma(U)\max_{a_{t+1}} Q_t\,(s_{t+1}, a_{t+1})\right], \quad (1)$$

де η – коефіцієнт навчання, який визначає швидкість навчання на інтервалі $(0,1)$.

Агент вузлового рівня складається з підмереж агентів FA-OSELM, кожен з яких є нейронною мережею (НМ) прямого зв'язку з одним прихованим рівнем, яка може навчатися реальному часі. Кожна НМ має K прихованих вузлів з сигмоїдною функцією активації, обраною для спрощення навчання при нелінійних вхідних даних та один вихідний вузол з функцією лінійної активації для полегшення процесу тренування нейронної мережі, оскільки вона не вносить додаткову нелінійність. Вхідний рівень кожної нейромережі (Full Adaptive Online Sequential Extreme Learning Machine) FA-OSELM приймає вектор стану агента мережевого рівня FANET-MANET як вхід, а вихідний рівень – як метрику маршрутизації.

Стан мережі – це вектор, який містить наступну інформацію: ідентифікатор вузла-відправника; ідентифікатор вузла призначення; місцезнаходження; швидкість; напрямок переміщення; кількість вузлів-сусідів; кількість переходів (стрибків) до вузла призначення і залишкову ємність батареї.

Метрика маршрутизації – це скалярна величина, яка визначає рівень готовності поточного вузла для ретрансляції відповідно до заданого алгоритму маршрутизації. Чим вище значення метрики маршрутизації, тим краще поточний вузол відповідає умовному критерію готовності до інформаційно-комунікаційного обміну. Кожна підмережа OSELM формує різну метрику маршрутизації на виході. Таким чином, процес ініціалізації нейромережі FA-OSELM відбувається випадковим чином відповідно до загального стану мережі FANET-MANET на основі зворотного зв'язку.

Зворотний зв'язок у процесі навчання агентів вузлового рівня представлено у вигляді функції винагороди, яка залежить від продуктивності мережі і формується, використовуючи такі показники, як коефіцієнт доставки пакетів, затримка і енергоспоживання вузлами мережі.

Навчання агента *Q*-навчання мережевого рівня, що являє собою безмодельний алгоритм RL, полягає у вивченні станів та формуванні політики вибору найкращого рішення агентів вузлового рівня FA-OSELM (кожного вузла) на основі стану мережі та функції винагороди. Агент *Q*-навчання має *Q*-таблицю, в якій зберігаються значення для кожної пари стан – дія, де стан є вектором стану мережі, а дія – індексом рішення підмереж (агентів) вузлового рівня FA-OSELM.

Q-значення – це скаляр, який вказує на очікувану майбутню винагороду за дію в певному стані. Чим вище значення *Q*, тим відповідно вищий пріоритет вибору дії в заданих умовах.

Q-таблиця на початковому етапі ініціалізується випадковим чином і оновлюється онлайн в реальному часі відносно зміни стану мережі та отриманого зворотного зв'язку. Процес оновлення Q-таблиці ґрунтується на формулі (2):

$$Q_t(s,a) = Q_{t-1}(s,a) + \alpha * \Big(R + \gamma * \max_{(a')}\big(Q_{t-1}(s',a')\big) - Q_{t-1}(s,a)\Big). \quad (2)$$

Оновлення Q-значення стан – дія (s, a) в момент часу t, де s – поточний стан, a – поточна дія, R – поточна винагорода, s' – наступний стан, a' – наступна дія, з метою (3)

$$a_t = \text{argmax}_{(a)} Q_t(s,a), \quad (3)$$

вибору дії з найбільшим Q-значенням в поточному стані на кроці t та є цільовою функцією нейромережевого агента мережевого рівня.

Моделювання процесу інтелектуального управління НПМ.

Комунікаційні аероплатформи підтримують однакові протоколи інформаційного обміну, мають обмежені дальність радіозв'язку і швидкість обміну. Кожен вузол мережі має власну систему управління, діє в кооперації з іншими вузлами мережі і КА. КА обладнані інерціальною навігаційною системою та приймачем GPS-сигналів глобальної навігаційної супутникової системи, антенною системою MIMO 4×4. Основні характеристики обладнаної антени MIMO: вихідна потужність в межах 8–20 Вт; підсилення сигналу 5–6 дцБ з функцією Effective Isotropic Radiated Power (EIRP); швидкість передачі даних складає до 100 мбіт/с.

Наземні абоненти обладнано: 2x2 MIMO, з вихідною потужністю передавача 1 мВт – 1 Вт (до 2 Вт у режимі TX Beamforming); смуга пропускання – 1,25 МГц, 2,5 МГц, 5 МГц; швидкість передачі даних – до 20 Мбіт/с; робочі частоти – 1350–1440 МГц, 2200–2500 МГц; 4400–4940 МГц.

Для забезпечення адаптивних процесів на мережевому рівні вузлів MANET та FANET в моделі мобільності застосовується вдосконалений протокол маршрутизації на основі MAODV [16].

Основні характеристики антен MIMO мобільних базових станцій, мобільних користувачів та вузлових елементів повітряної мережі зведено до таблиці 1, де

P_{out} – вихідна потужність, яку випромінює антена, Вт;

G – коефіцієнт підсилення, дБ;

η – ККД, який є відношенням потужності, що випромінює антена, до потужності, яка подається на антену, %;

P – вид поляризації, яка є орієнтацією вектора електричного поля електромагнітної хвилі, яку випромінює антена;

D – коефіцієнт направленої дії, дБ;

B – ширина променя діаграми направленості B, що є кутовою шириною основного пелюстка діаграми направленості антени, (°).

Таблиця 1

Параметри антени MIMO відповідно до типу вузла

Тип вузла	$\boldsymbol{P_{out}}$, W	$\boldsymbol{G}$, дБ	$\boldsymbol{\eta}$, %	$\boldsymbol{P}$	$\boldsymbol{D}$, дБ	$\boldsymbol{B}$, (°)
Мобільний користувач	0,1–1	0–10	50–90	Лінійна	0–10	15–30
Мобільна базова станція	0,5–5	10–20	70–95	Кругова	10–20	30–60
Комунікаційна аероплатформа	1–10	20–30	80–98	Еліптична	20–30	60–120

Мобільність вузлів моделюється на основі запропонованої моделі мобільності [6] на основі математики алгоритму градієнтного спуску з моментумом (Gradient Descent with Momentum), який є розширенням стандартного градієнтного спуску. Він допомагає подолати деякі з проблем, властиві класичному градієнтному спуску, таких як повільна збіжність та осцилююча поведінка і застосовується для апроксимації реалістичності руху вузлів.

Рухливість вузла характеризується наступними параметрами: мінімальна швидкість; максимальна швидкість; прискорення; напрямок та час паузи. Характеристики мобільності вузлів змінюються залежно від типу носія комунікаційного обладнання, як показано в таблиці 2.

Таблиця 2

Характеристики мобільності вузлів НПМ

Тип вузла	**Мінімальна швидкість,** м/с	**Максимальна швидкість,** м/с	**Прискорення,** м/с^2	**Час паузи, чи зависання БпЛА,** с
Мобільна базова станція	0	10	0,5	10–20
Мобільний користувач	0	2	1	5–15
Комунікаційна аероплатформа	10	45	1,5	5–10

Енергія вузла характеризується наступними параметрами: початкова енергія; потужність передачі; потужність прийому; витрата потужності в режимі очікування. Параметри енергії вузла змінюються залежно від типу вузла, як показано в таблиці 3.

Таблиця 3

Енергетичні параметри комунікаційних вузлів

Тип вузла	**Початкова енергія,** Дж	**Потужність передачі,** Вт	**Потужність прийому,** Вт	**Потужність в режимі очікування,** Вт
Мобільний користувач	1000	0,1–1	0,05	0,01
Мобільна базова станція	2000	0,5–5	0,1	0,02
Комунікаційна аероплатформа	3000	1–10	0,2	0,04

Радіодоступність вузла представлена моделлю логарифмічного нормального затінення [25] (Lognormal Shadowing Model) і є часто поширеною моделлю для опису радіодоступності вузлів в бездротових мережах. Ця модель враховує випадкові флуктуації потужності сигналу, викликані затіненням перешкодами. Радіодоступність вузла характеризується наступними параметрами:

дальність передачі – показує типовий діапазон відстаней, на яких може здійснюватися передача сигналу між мобільним користувачем та базовою станцією або іншим пристроєм. Для мобільного користувача цей діапазон становить від 100 м до 400 м;

експоненціальна втрата шляху – характеризує швидкість загасання сигналу зі збільшенням відстані. Чим більше значення експоненти, тим швидше сигнал слабшає при віддаленні від передавача;

відхилення затінення – показує стандартне відхилення втрат сигналу через ефекти затінення, такі як перешкоди на шляху поширення (будівлі, дерева тощо);

коефіцієнт спрямованості – цей параметр характеризує здатність антени фокусувати випромінювання в певному напрямку.

Параметри радіодоступності вузла варіюються залежно від типу вузла і типу каналу, як показано в таблиці 4.

Таблиця 4

Параметри радіодоступності комунікаційних вузлів НПМ

Тип вузла	**Дальність передачі,** м	**Експонента втрати шляху**	**Відхилення затінення,** дБ	**Коефіцієнт спрямованості**
Мобільний користувач	100–400	3	6	2
Мобільна базова станція	400–800	3,5	8	4
Комунікаційна аероплатформа	800–1600	4	10	8

Трафік вузлів представлено моделлю трафіку з постійною бітовою швидкістю Constant Bit Rate [26], яка може генерувати постійні та періодичні пакети даних.

У якості протоколу канального доступу вузла застосовується IEEE 802.11 – протокол безпроводової локальної мережі. Протокол характеризується наступними параметрами: пропускною здатністю каналу; методом модуляції; швидкістю кодування; порогом чутливості. Параметри протоколу доступу змінюються залежно від типу вузла та типу зв’язку, як показано в таблиці 5.

Таблиця 5

Сигнальні параметри комунікаційних вузлів НПМ

Тип вузла	**Смуга пропускання каналу,** МГц	**Метод модуляції**	**Швидкість кодування**	**Поріг чутливості,** дБм
Мобільний користувач	1,25; 2,5; 5	BPSK, QPSK, 16-QAM	1/2, 2/3, 3/4	-82, -79, -77
Базова станція	5, 10 ,20	QPSK, 16-QAM, 64-QAM	1/2, 2/3, 3/4	-77, -74, -72
БПЛА	10, 20 ,40	16-QAM, 64-QAM, 256-QAM	1/2, 2/3, 3/4	-72, -69, -66

На рисунку 2 зображено основі етапи процесу навчання алгоритму машинного навчання з підкріпленням для управління мережею FANET-MANET, які складаються з чотирьох блоків.

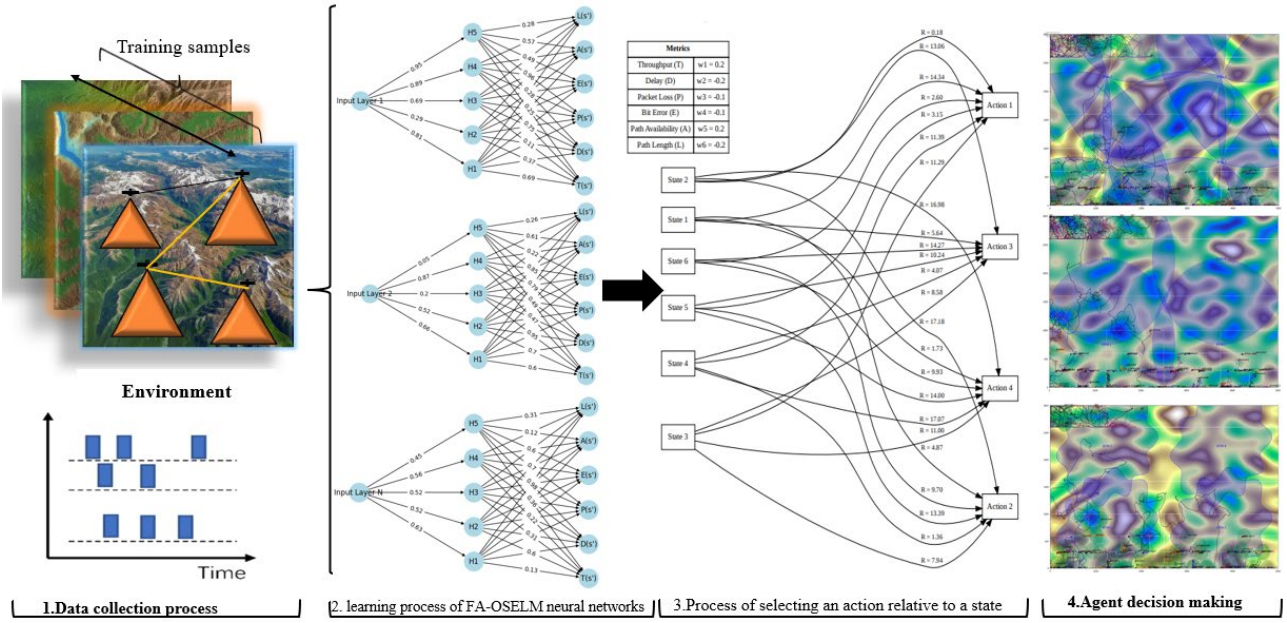


Рис. 2. Основні етапи процесу машинного навчання з підкріпленням ІА для інтелектуального управління мережею FANET-MANET

Блок 1 ілюструє процес збору статистичних даних поведінки мобільних користувачів мережі нижнього рівня MANET та верхнього рівня FANET з урахуванням заданих особливостей функціонування.

Блок 2 показує процес параметризації та навчання нейромережевої архітектури агентів вузлового рівня на основі алгоритмів FA-OSELM, який описується наступними рівняннями:

1. $w_k = \text{randn}(K, 1)$ – ініціалізація вектора вагових коефіцієнтів k-го рівня OSELM;
2. $b_k = \text{randn}(K, 1)$ – ініціалізація вектора зміщення b_k;
3. $\beta_k = \text{randn}(K, 1)$ – ініціалізація вектора вагового коефіцієнта β;
4. $H_k = \text{zeros}(K, K)$ – ініціалізація вихідної матриці прихованого k-го рівня OSELM;

5. $P_k = \text{eye}(K, c)$ – ініціалізація оберненої матриці вихідного *k*-го рівня OSELM, де *c* – параметр регуляризації;

6. $x_t = [s, a, d, c, L, v, \theta, |N|, h, b]$ – вектор стану мережі в момент часу *t*;

7. $y_t = R$ – функція винагороди в момент часу *t*;

8. $h_{(k,t)} = \frac{1}{1+e^{(-w_k * x_t - b_k)}}$ – вектор вихідного рівня OSELM в момент часу *t*;

9. $o_{k,t} = \beta_k * h_{k,t}$ – результат обчислення вихідного *k*-го рівня нейромережі в момент часу *t*;

10. $e_{k,t} = y_t - o_{k,t}$ – оцінка прогнозу, тобто похибка в реальному часі;

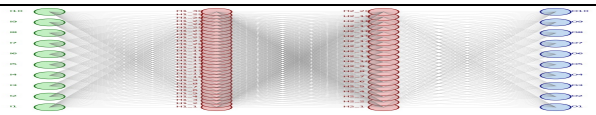
11. $g_{k,t} = \frac{P_k \cdot h_{k,t}}{1+h_{k,t} \cdot P_k \cdot h_{k,t}}$ – вектор коефіцієнта *k*-го OSELM в момент часу *t*;

12. $\beta_k = \beta_k + g_{k,t} \cdot e_{k,t}$ – процес оновлення вектору вихідних вагових коефіцієнтів *k*-го рівня;

13. $P_k = P_k - g_{k,t} \cdot h_{k,t} \cdot P_k$ – процес оновлення матриці *k*-го вихідного рівня OSELM.

Таблиця 6

Параметри нейромереж FA-OSELM комунікаційних вузлів НПМ

Основні параметри нейромережевого алгоритму	**Значення**
Коефіцієнт навчання (швидкість навчання)	0.1
Коефіцієнт дисконтування	0.9
Вагові коефіцієнти	U = 0.2,0.3,0.4,0.1,0.1,0.1,0.2,0.2,0.2,0.1,0.1,0.1,0.1,0.1,0.1
Q-learning parameters	α = 0.1, γ = 0.9
Функція винагороди	α = 0.5, β = 0.3, γ = 0.2
Первина архітектура	 50, 100 , 80, 50
Функції активації нейронів	$sigmoid, tanh, relu, softmax, linear$

Блок 3 та 4 описує процес управління параметрами алгоритму машинного навчання з підкріпленням (функція винагороди, штрафу) та процес прийняття рішення для управління мережею НПМ (FANET-MANET).

Функція винагороди агента $R(s, a, s')$ визначена як сума вагових коефіцієнтів та метрик, які оцінюють продуктивність мережі FANET-MANET, що наведено нижче у наступному виразі:

$$R(s, a, s') = w_1 \cdot T(s') + w_2 \cdot D(s') + w_3 \cdot P(s') + w_4 \cdot E(s') + w_5 \cdot A(s') + w_6 \cdot L(s'),$$

де (s) – поточний стан мережі; (a) – дія, вибрана агентом; (s') – наступний стан мережі після виконання дії; (w_i) – вагові коефіцієнти для кожної метрики; $(L(s'))$ – довжина маршруту в наступному стані, яка враховує потенційні довгострокові наслідки.

Кожна метрика визначена як:

- пропускна спроможність мережі $\big(T(s')\big)$: $\left[T(s') = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} b_{i,j}\right]$;
- затримка передачі пакетів даних $\big(D(s')\big)$: $\left[D(s') = \frac{1}{M}\sum_{k=1}^{M} d_k\right]$;
- втрата пакетів $\big(P(s')\big)$: $\left[P(s') = \frac{1}{M}\sum_{k=1}^{M} p_k\right]$;

- бітова помилка $\left(E(s')\right)$: $\left[E(s') = \frac{1}{M}\sum_{k=1}^{M} e_k\right]$;
- доступність маршруту $\left(A(s')\right)$: $\left[A(s') = \frac{1}{M}\sum_{k=1}^{M} a_k\right]$.

Функція винагороди агента мережевого рівня керується наступною логікою:

- винагорода R за забезпечення заданої пропускної спроможності $(T(s'))$ мережі;
- штрафом Pen за високу затримку передачі пакетів $(D(s'))$;
- втрата пакетів $(P(s'))$ зі штрафом за перевищену втрату пакетів, що призводить до відсутності зв'язку між мобільними користувачами;
- бітова помилка $(E(s')$ не менше ніж до заданої;
- винагорода R за доступність маршруту $(A(s'))$ від відправника до кінцевого користувача.
- довжина маршруту $(L(s'))$ зі штрафом Pen за довгі маршрути, а саме вибір оптимального маршруту з найменшим числом стрибків із урахуванням чутливості приймача, залишкової енергії та матриці навантаження.

Для кожної метрики визначаються порогові значення, чи отримує агент винагороду або штраф.

Наприклад, якщо затримка $(D(s'))$ перевищує певний поріг, агент отримує штраф, що відображається у від'ємному значенні вагового коефіцієнта (w_2).

Функція винагороди повинна бути максимізована, щоб підвищити ефективність мережі. Вагові коефіцієнти (w_i) визначають важливість кожної метрики в контексті загальної цілі мережі. Наприклад, якщо пропускна спроможність є критичною для мережі, (w_1) може бути вищим, ніж інші вагові коефіцієнти.

Математично процес параметризації агента з числовими штрафами для субоптимального рішення у контексті машинного навчання з підкріпленням представлений у вигляді матриці винагороди (R) та матриці штрафів (P), де кожен елемент матриці відображає винагороду або штраф за певну дію агента відносно оцінки стану.

Нехай є (M) можливих станів мережі та (A) можливих дій, які може вибрати агент. Тоді матриця винагород (R) та матриця штрафів (P) буде мати розмірність $(M \times A)$:

$$R = \begin{bmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,A} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,A} \\ \vdots & \vdots & \ddots & \vdots \\ r_{M,1} & r_{M,2} & \cdots & r_{M,A} \end{bmatrix},$$

де $\left(r_{m,a}\right)$ – винагорода за виконання дії (a) у стані (m);

$$Pen = \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,A} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,A} \\ \vdots & \vdots & \ddots & \vdots \\ p_{M,1} & p_{M,2} & \cdots & p_{M,A} \end{bmatrix},$$

де $\left(pen_{m,a}\right)$ – штраф за виконання дії (a) у стані (m).

Якщо дія (a) призводить до покращення пропускної спроможності, то відповідний елемент $\left(r_{m,a}\right)$ у матриці (R) буде мати позитивне значення. Якщо дія (a) призводить до збільшення затримки, то відповідний елемент $\left(pen_{m,a}\right)$ у матриці (Pen) буде мати негативне значення.

Агент у процесі навчання намагається максимізувати сумарну винагороду, отриману протягом часу, вибираючи дії, які забезпечують позитивні значення у матриці (R) та мінімізувати негативні значення у матриці (Pen).

Для субоптимального рішення для управління мережею FANET-MANET, функцію винагороди враховує не тільки поточні метрики продуктивності, але й потенційні довгострокові наслідки дій агента.

Архітектура імітаційної моделі. Нижче наведена архітектура програмної імітаційної моделі, яка складається з наступних модулів (рис. 3).

Рис. 3. Архітектура програмної імітаційної моделі збору статистичних даних інтелектуального управління НПМ

Модуль генерації тривимірної карти рельєфу. Цей модуль відповідає за генерацію тривимірної карти рельєфу для заданої місцевості, використовуючи дані із доступних географічних інформаційних систем (ГІС), таких як Google Earth, NASA World Wind, OpenStreetMap з використанням бібліотеки в програмному середовищі Python GeoPandas. Модуль генерації тривимірної карти рельєфу використовує алгоритми інтерполяції та апроксимації для створення гладкої та реалістичної поверхні землі, яка враховує різні типи рельєфу, такі як гори, долини, рівнини, озера, річки тощо.

Модуль генерації сценаріїв, суть якого полягає у забезпечені процесу генерації сценаріїв руху вузлів у мережах FANET-MANET. Модуль генерації сценаріїв руху вузлів використовує тривимірну карту рельєфу для визначення допустимого діапазону висоти для кожного вузла, а також враховує різні параметри та обмеження, кількість вузлів, час симуляції, максимальну швидкість, максимальне прискорення, відстань, кількість сусідів, максимальну кількість стрибків, кількість переходів групи тощо.

Далі відбувається агрегація результатів даних у файл сценарію руху вузлів, який містить координати, швидкість, прискорення, кут, параметри потужності передачі, метод модуляції та іншу інформацію про кожен вузол на кожному часовому кроці.

Модуль симуляції мережі. Цей модуль відповідає за симуляцію мережі FANET-MANET, використовуючи файл сценарію руху вузлів.

Модуль симуляції мережі використовує програмне забезпечення на основі мови програмування Python та враховує протокол маршрутизації, безпеки, енерговитрати, вимоги із якості обслуговування тощо. Далі відбувається процес агрегації вихідного файлу статистичних даних моделі мобільності FANET-MANET, формуються метрики та показники ефективності управління НПМ, такі як: пропускна здатність, затримка, втрата пакетів, бітова помилка, кількість підмереж, розмір кожної підмережі, зв'язність кожної підмережі та мережі в цілому, довжина маршрутів, якість та доступність маршрутів тощо.

Збір та аналіз статистичних даних процесу інтелектуального управління НПМ.

Методика збору даних.

Процес визначення репрезентативності навчальної вибірки важливий з точки зору вирішення протиріччя між обсягом службового трафіку та якості управління. В основу дослідження покладено процес визначення цільової популяції вхідних даних за визначеним критерієм відбору мінімально-допустимої репрезентативної вибірки.

Процес пошуку розмірності вибірки відбувається на основі довірчого інтервалу та за допомогою методів кластерної та систематичної вибірки.

Таким чином, процес визначення репрезентативності вибірки відбувається внаслідок порівняння вихідних результатів, отриманих на основі популяції вихідних даних з отриманими середнім значенням, медіаною, дисперсією, з різною частотою оновлення тощо. Для перевірки гіпотез про відповідність аргументів цільової функції управляючих рішень агенту мережевого рівня всієї популяції та підібраної вибірки в роботі пропонується застосувати існуючі методи математичної статистики, такі як t-тест та F-тест ANOVA.

Основні параметри для симуляції мережі:

кількість вузлів – 80;
час симуляції – 5000 с;
максимальна швидкість вузлів – 20 м/с;
мінімальна швидкість вузлів – 5 м/с;
максимальна швидкість – 20 м/с;
максимальне прискорення – 2 м/с^2;
максимальний час паузи вузлів – 10 с;
мінімальний час паузи вузлів – 2 с.;
розмір області симуляції – 5000 м × 3000 м;
тип мережі – FANET-MANET;
тип мобільності – RZRVPM/RD/RVM;
тип протоколу маршрутизації: запропонований на основі MAODV ;
тип протоколу MAC – IEEE 802.11n;
тип антени – MIMO
розмір пакета – 512 байт;
інтервал відправки пакетів – 0,1 с;
максимальна відстань – 1000 м.

Результати аналізу даних. Результати аналізу статистичних даних моделі мобільності FANET-MANET були згенеровані модулем симуляції мережевої діяльності і описуються наступними виразами:

середня швидкість вузлів в момент часу t

$$\bar{v}_t = \frac{1}{N}\sum i = 1^N v_{i,t},$$

де N – кількість вузлів, $v_{i,t}$ – швидкість i-го вузла в момент часу t;

середня відстань між вузлами в момент часу t

$$\bar{d}t = \frac{2}{N(N-1)}\sum i = 1^{N-1}\sum_{j=i+1}^{N} d_{i,j,t},$$

де $d_{i,j,t}$ – відстань між i-м та j-м вузлами в момент часу t;

середній час зв'язку між вузлами задається формулою

$$\bar{T}c = \frac{1}{M}\sum k = 1^M T_{c,k},$$

де M – кількість пар вузлів, які мають зв'язок, $T_{c,k}$ – час зв'язку між k-ю парою вузлів;

середній час переривання зв'язку між вузлами задається виразом

$$\bar{T}b = \frac{1}{M}\sum k = 1^M T_{b,k},$$

де $T_{b,k}$ – час переривання зв'язку між k-ю парою вузлів;

середня кількість змін напрямку вузлів в момент часу t

$$\overline{N_\theta} = \frac{1}{N}\sum_{i=1}^{N} N_{\theta,i,t},$$

де $N_{\theta,i,t}$ – кількість змін напрямку -го вузла в часовий момент

середня кількість сусідів вузлів в момент часу t

$$\overline{N_n} = \frac{1}{N}\sum_{i=1}^{N} N_{n,i,t},$$

де $N_{n,i,t}$ – кількість сусідів -го вузла в часовий момент t;

середня кількість стрибків між вузлами в момент часу t

$$\overline{N_h} = \frac{1}{M}\sum_{k=1}^{M} N_{h,k,t},$$

де $N_{h,k,t}$ – кількість стрибків між k-ю парою вузлів, які обмінюються пакетами даними, в момент часу t;

середня кількість переходів групи між вузлами

$$\overline{N_g} = \frac{1}{M}\sum_{k=1}^{M} N_{g,k,t},$$

де $N_{g,k,t}$ – кількість переходів групи між вузлами, які мають зв'язок, в часовий момент .

Для дослідження впливу фізичного середовища переміщення абонентів наземної та повітряної комунікаційної мережі було вибрано три типи рельєфу для симуляції: рівнинний, гірський та річковий.

У процесі моделювання було зібрано та збережено у файл сценарію руху вузлів такі дані: координати, швидкість, прискорення, кут, потужність передачі, метод модуляції та іншу інформацію про кожен вузол у кожен дискретний відлік часу. Також було збережено дані про пропускну здатність, затримку, втрату пакетів, бітову помилку, кількість підмереж, розмір кожної підмережі, зв'язок кожної підмережі, довжину, якість та доступність шляхів тощо. Обробка даних відбувалась за допомогою мови програмування Python, що дозволяло виконувати різні статистичні та графічні операції.

Далі відбувається процес генерації сценаріїв мобільності вузлів, використовуючи наступні моделі мобільності для порівняння з запропонованою моделлю мобільності [6]:

Random Waypoint: модель мобільності використовується для допущення, що вузли вибирають випадкову точку призначення в межах області симуляції та рухаються до неї з випадковою швидкістю. Після досягнення точки вузли зупиняються на випадковий час паузи та повторюють процес. Основні параметри цієї моделі мобільності є мінімальна та максимальна швидкості, мінімальний та максимальний час паузи.

Random Direction: у параметрах моделі мобільності є допущення, що вузли вибирають випадковий напрямок та рухаються в ньому з випадковою швидкістю доти, поки не досягнуть межі області симуляції. Далі вузли змінюють напрямок на випадковий кут та продовжують рух. Основні параметри цієї моделі мобільності є мінімальна та максимальна швидкості, мінімальний та максимальний кути повороту.

Reference Point Group Mobility: у параметрах моделі мобільності є допущення, що вузли розбиті на групи, кожна з яких має лідера, який визначає траєкторію руху групи. Вузли в межах групи рухаються відносно свого лідера з деяким відхиленням. Основні параметри цієї моделі мобільності є кількість груп, кількість вузлів в групі, мінімальна та максимальна швидкість лідера, мінімальний та максимальний час паузи лідера, мінімальне та максимальне відхилення вузлів від лідера.

Статистичний метод ANOVA [27] для обчислення F-статистики ґрунтується на основі наступних виразів:

$$F = \frac{MS_{between}}{MS_{within}},$$

де $MS_{between}$ – середній квадрат між групами, а MS_{within} – середній квадрат в межах груп:

$$MS_{between} = \frac{SS_{between}}{df_{between}},$$

$$MS_{within} = \frac{SS_{within}}{df_{within}},$$

де $SS_{between}$ – сума квадратів між групами, SS_{within} – сума квадратів в межах груп, $df_{between}$ – ступені свободи між групами, а df_{within} – ступені свободи в межах груп, що наведено в наступних виразах:

$$SS_{between} = \sum_{i=1}^{k} n_i(\bar{x}_i - \bar{x})^2,$$
$$SS_{within} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \left(x_{ij} - \bar{x}_i\right)^2,$$
$$df_{between} = k - 1,$$
$$df_{within} = N - k,$$

де k – кількість груп, n_i – кількість спостережень в *i*-тій групі, $\bar{x}_i$ – середнє значення *i*-тої групи, $\bar{x}$ – загальне середнє значення, x_{ij} – *j*-те спостереження в *i*-тій групі, а N – загальна кількість спостережень.

Для застосування t-тесту обчислення t-статистики відбувається на основі виразу:

$$t = \frac{\overline{x_1} - \overline{x_2}}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

де $\overline{x_1}$ – середнє значення першої групи, $\overline{x_2}$ – середнє значення другої групи, s_p – оцінка стандартного відхилення в обох групах, n_1 – кількість спостережень в першій групі, а n_2 – кількість спостережень в другій групі. Величина s_p обчислюється за допомогою наступної формули:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}},$$

де s_1 – стандартне відхилення першої групи, а s_2 – стандартне відхилення другої групи.

Необхідно визначити розмір вибірки, який забезпечить статистичну значущість та репрезентативність вибірки – достатню точність та надійність оцінок параметрів популяції, враховуючи довірчій інтервал та похибку. Наприклад, якщо оцінка середньої якості зв'язку Q в популяції з довірчим інтервалом 95 % та похибкою 0,01, тоді для вибірки із достатньо великої популяції для забезпечення заданого довірчого інтервалу використовуємо:

$$n = \frac{z^2\sigma^2}{E^2},$$

де *n* – розмір вибірки, *z* – критичне значення нормального розподілу для заданого рівня довіри (для 95 % це 1,96), σ – стандартне відхилення Q в популяції (якщо не відоме, можна використати оцінку з попередніх досліджень), *E* – допустима похибка (0,01). Якщо популяція мала, то потрібно використати наступний вираз з поправкою на розмір популяції *N*:

$$n = \frac{Nz^2\sigma^2}{E^2(N-1) + z^2\sigma^2}.$$

Наступним кроком є вибір методу визначення розмірності вибірки: випадкова вибірка, стратифікована вибірка, кластерна вибірка, систематична вибірка. Необхідно зазначити, що у дослідженні статистична вибірка симуляції моделі мобільності відображає комбінації популяцій за типами рельєфу, тоді оптимальним варіантом є метод стратифікованої вибірки, де є необхідність декомпозиції популяції на групи (страти) за типами територій, такими як рівнинний, ліси, водойми (табл. 7).

Таблиця 7

Середні значення вихідних даних розробленої моделі функціонування вузлів НПМ

Тип рельєфу	**Середня швидкість,** м/с	**Середня відстань,** м	**Середній час зв'язку,** с	**Середній час радіомовчання,** с	**Середня кількість змін напрямку**	**Середня кількість сусідів**	**Середня кількість стрибків**	**Середня кількість переходів групи**
Рівнинний	10,5	500	800	200	5	8	3	2
Гірський	8,7	400	600	400	7	6	4	3
Річковий	9,1	430	650	350	6	7	4	3

Далі необхідно вибрати випадкову підвибірку з кожної групи, пропорційну її розміру в загальній популяції (табл. 8). Це дозволить отримати більш точні оцінки параметрів статистичної популяції на відміну від методів, які базуються лише на випадкових розподілах.

Таблиця 8

Стратифікація популяції за типами рельєфу

Тип рельєфу	**Розмір групи в популяції**	**Розмір підвибірки**
Рівнинний	1000	200
Ліси	1500	300
Водойми	500	100

Процес перевірки репрезентативності вибірки полягає у порівнянні характеристик популяції, таких як середні, медіана, дисперсія, частота, гістограми, кореляції тощо. В цьому випадку можливо застосовувати статистичні тести для перевірки гіпотез про рівність параметрів вибірки та популяції, такі як t-тест та F-тест ANOVA. Якщо виявлено значні відмінності, тоді необхідно коректувати вибірку, виключивши або додавши елементи та змінивши метод вибору. Наприклад, якщо необхідно перевірити репрезентативність вибірки за середньою якістю зв'язку Q, тоді можливо застосувати t-тест для порівняння середніх вибірки та популяції:

$$t = \frac{\overline{Q_s} - \overline{Q_p}}{s/\sqrt{n}},$$

де t – статистика t-тесту, $\overline{Q_s}$ – середня вибірки, $\overline{Q_p}$ – середня Q популяції, s – стандартне відхилення вибірки, n – розмір вибірки. Якщо t перевищує критичне значення для заданого рівня довіри (1,96 для 95 %), тоді відповідно робиться висновок про відхилення нульової гіпотези про рівність середніх і приймається альтернативна гіпотеза про нерівність середніх. Це означає, що отримана вибірка не репрезентативна для забезпечення заданої якості зв'язку.

Нижче наведено результати статистичного аналізу, які враховують різні типи рельєфу для моделей Random Waypoint, Random Direction та Reference Point Group Mobility, середні значення вихідних даних зведено до таблиць 9–14.

Таблиця 9

Середні значення вихідних даних моделі Random Waypoint

Тип рельєфу	**Середня швидкість,** м/с	**Середня відстань,** м	**Середній час зв'язку,** с	**Середній час радіо-мовчання,** с	**Середня кількість змін напрямку**	**Середня кількість сусідів**	**Середня кількість стрибків**	**Середня кількість переходів групи**
Рівнинний	12,3	750	1200	300	7	12	5	4
Гірський	10,1	600	900	600	9	9	6	5
Водойми	10,8	645	975	525	8	11	6	5

Таблиця 10

Стратифікація популяції за типами рельєфу для моделі Random Waypoint

Тип рельєфу	**Розмір групи в популяції**	**Розмір підвибірки**
Рівнинний	5000	1000
Ліси	7500	1500
Водойми	2500	500

Таблиця 11

Середні значення вихідних даних моделі Random Direction

Тип рельєфу	**Середня швидкість,** м/с	**Середня відстань,** м	**Середній час зв’язку,** с	**Середній час радіо-мовчання,** с	**Середня кількість змін напрямку**	**Середня кількість сусідів**	**Середня кількість стрибків**	**Середня кількість переходів групи**
Рівнинний	11,7	700	1100	280	6	11	4	3
Ліси	9,6	570	855	570	8	8	5	4
Водойми	10,3	615	920	500	7	10	5	4

Таблиця 12

Стратифікація популяції за типами рельєфу для моделі Random Direction

Тип рельєфу	**Розмір групи в популяції**	**Розмір підвибірки**
Рівнинний	6000	1200
Ліси	9000	1800
Водойми	3000	600

Таблиця 13

Середні значення вихідних даних моделі Reference Point Group Mobility

Тип рельєфу	**Середня швидкість,** м/с	**Середня відстань,** м	**Середній час зв’язку,** с	**Середній час радіо-мовчання,** с	**Середня кількість змін напрямку**	**Середня кількість сусідів**	**Середня кількість стрибків**	**Середня кількість переходів групи**
Рівнинний	13,1	800	1300	320	8	13	6	5
Ліси	10,9	640	960	640	10	10	7	6
Водойми	11,5	690	1040	560	9	12	7	6

Таблиця 14

Стратифікація популяції за типами рельєфу для моделі Reference Point Group Mobility

Тип рельєфу	**Розмір групи в популяції**	**Розмір підвибірки**
Рівнинний	8000	1600
Ліси	12000	2400
Водойми	4000	800

Популяція вихідних даних для альтернативних моделей значно більша порівняно з розробленою моделлю, що суттєво впливало на обчислювальну складність в процесі навчання.

Порівнюючи показники моделей Random Waypoint, Random Direction та Reference Point Group Mobility з розробленою, можливо зробити висновок про перевагу в ефективності та точності моделювання ІСУ НПМ, незважаючи на меншу популяцію вихідних даних.

Результати застосування методу ANOVA для розрахунку метрик продуктивності та надійності мережі, яких вдалося досягти із використанням інтелектуальної системи управління НПМ, наведено нижче в таблиці 15.

Таблиця 15

Результати розрахунку статистичних показників навчальних даних

Метрика	p-значення t-тесту (ANOVA) розробленої моделі	p-значення t-тесту (ANOVA) моделі RWP	p-значення t-тесту (ANOVA) моделі RD	p-значення t-тесту (ANOVA) моделі RPGM
Пропускна спроможність	0,0001	0,0003	0,0002	0,0004
Затримка передачі	0,0002	0,0005	0,0004	0,0006
Втрата пакетів	0,0001	0,0004	0,0003	0,0005
Коефіцієнт доставки пакетів	0,0001	0,0003	0,0002	0,0004

Розраховане значення F-статистики для всіх метрик перевищує 18, що вказує на високу статистичну значущість управляючих рішень агентів вузлового рівня для формування управляючих рішень агента мережевого рівня.

Для всіх моделей мобільності, використаних для формування управляючих рішень вузлового та мережевого рівнів, p-значення є меншими за 0,0006, що підтверджує надійність виявлених відмінностей на рівні значущості 0,05, крім того, додатково підтверджує адекватність ієрархічного послідовного підходу інтелектуального управління за загальною схемою нейромережі FA-OSELM (агенти вузлового рівня) – Q-RL з підкріпленням (агент мережевого рівня) для оптимізації параметрів мережі в реальному часі, адаптуючись до змін у середовищі функціонування.

Аналіз продуктивності НПМ з різними моделями мобільності.

Для оцінки продуктивності НПМ застосовувались наступні метрики та показники:

пропускна спроможність B – дозволяє оцінити кількість даних, які передаються в одиницю часу в мережі і визначається виразом:

$$B = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} b_{i,j},$$

де T – час симуляції, N – кількість вузлів, $b_{i,j}$ – кількість байтів, які передаються від i-го до j-го вузла;

метрика затримки D – дозволяє оцінити час, який потрібен для передачі повідомлень від вузла відправника до вузла-адресанта в мережі:

$$D = \frac{1}{M}\sum_{k=1}^{M} d_k,$$

де M – кількість пар вузлів, які мають зв'язок, d_k – затримка між k-ю парою вузлів;

метрика втрати пакетів P – дозволяє оцінити відсоток пакетів даних, які не досягають вузла-адресанта в мережі:

$$P = \frac{1}{M}\sum_{k=1}^{M} p_k,$$

де p_k – втрата пакетів між k-ю парою вузлів;

бітова помилка E – дозволяє отримати відсоток бітів, які передаються неправильно (з помилкою) в мережі:

$$E = \frac{1}{M}\sum_{k=1}^{M} e_k,$$

де e_k – бітова помилка між k-ю парою вузлів;

кількість підмереж S – ця метрика визначає кількість підмереж, на які розбивається мережа внаслідок руху вузлів:

$$S = \max_{t} s_t,$$

де s_t – кількість підмереж в в момент часу t;

розмір підмережі R – ця метрика визначає кількість вузлів, що належать до кожної підмережі:

$$R_i = \sum_{j=1}^{N} r_{i,j},$$

де i – індекс підмережі, $r_{i,j}$ – бінарний індикатор, який показує, чи належить j-й вузол до i-ї підмережі;

зв'язність кожної підмережі С – ця метрика дозволяє оцінити ступінь радіозв'язності кожної підмережі:

$$C_i = \frac{2}{R_i(R_i-1)} \sum_{j=1}^{R_i-1} \sum_{k=j+1}^{R_i} c_{i,j,k},$$

де $c_{i,j,k}$ – бінарний індикатор, який показує, чи існує зв'язок між j-м та k-м вузлами в i-й підмережі;

L – довжина шляху (маршруту) – метрика, яка відображає кількість стрибків, що потрібні для передачі пакета даних від адресата до адресанта в мережі:

$$L = \frac{1}{M} \sum_{k=1}^{M} l_k,$$

де l_k – довжина шляху між k-ю парою вузлів, які мають зв'язок.

Статистична оцінка продуктивності процесу інтелектуального управління НПМ

На рисунку 4 наведено гістограми середніх значень кількості групових переходів, середнього часу переривання зв'язку та середньої кількості сусідів у процесі управління метриками маршрутизації НПМ при різних моделях мобільності.

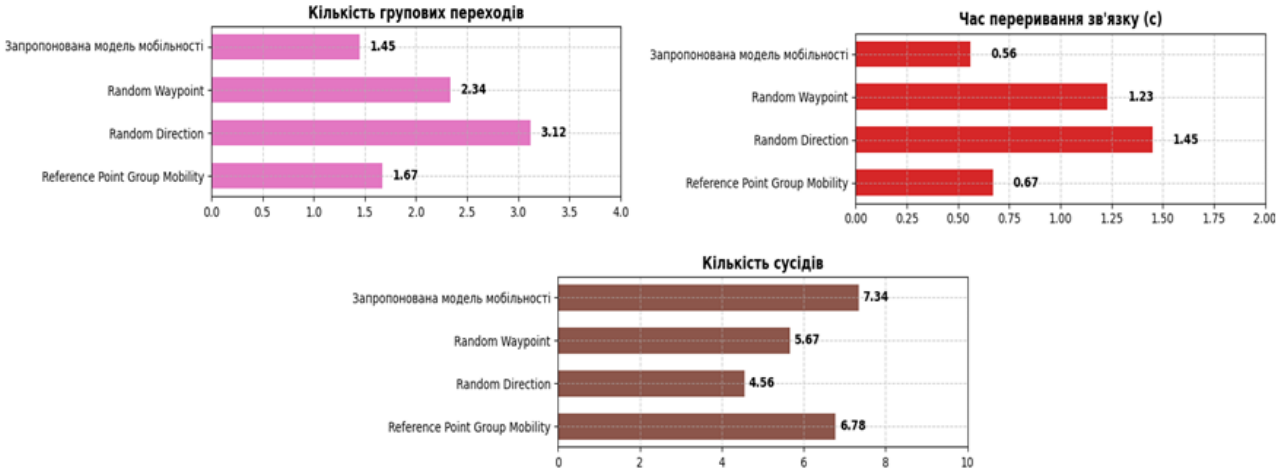


Рис. 4. Гістограми метрик маршрутизації процесі інтелектуального управління НПМ при різних моделях мобільності

Таблиця 16

Часові показники доставки повідомлень в симуляційній моделі оперативного управління НПМ

Модель мобільності	**Середній час доставки повідомлень**	**Час переривання зв'язку, с**	**Кількість групових переходів**
Запропонована ММ	2,34	0,56	1,45
Random Waypoint	3,21	1,23	2,34
Random Direction	4,12	1,45	3,12
Reference Point Group Mobility	2,89	0,67	1,67

Із отриманих результатів (табл. 16, рис. 4), можна зробити висновок, що ІСУ RL – FA-OSELM, навчена із використанням розробленої моделі мобільності, показує кращі результати порівняно з іншими розглянутими моделями. Зокрема, середній час доставки повідомлень за

однакової матриці вхідного навантаження становить 2,34 с, а час переривання зв’язку є мінімальним (0,56 с), що підкреслює високу стабільність інформаційно-комунікаційного обміну з використанням запропонованих рішень. На рисунку 5 показано результати оцінки ефективності процесу оперативного управління з використанням запропонованих рішень в ІСУ НПМ після навчання на різних моделях мобільності відносно розробленої.

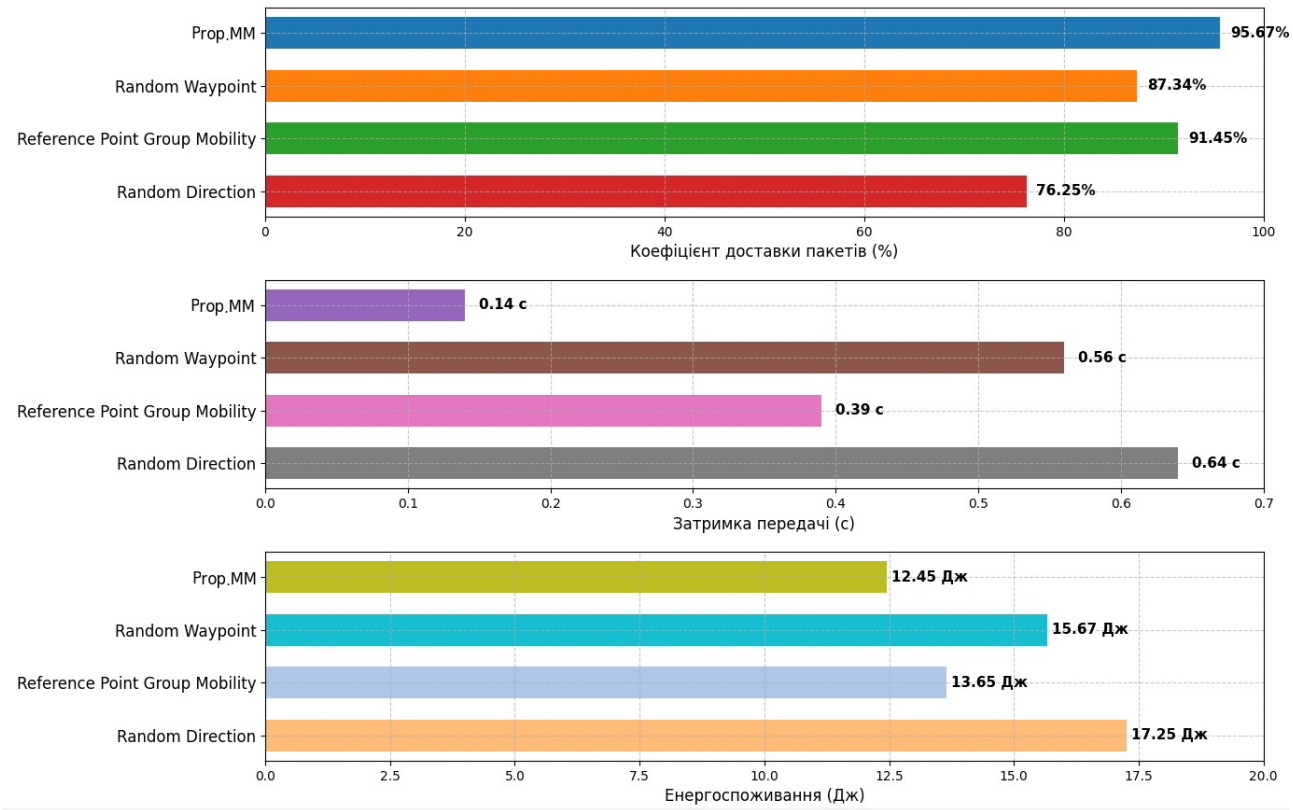


Рис. 5. Результати оцінки ефективності процесу оперативного управління з використанням запропонованих рішень в ІСУ НПМ

Аналізуючи представлені дані про ефективність процесу інтелектуального управління НПМ у контексті коефіцієнту доставки пакетів, затримки передачі та енергоспоживання вузлів мережі, ієрархічна інтелектуальна система управління НПМ на основі RL – FA-OSELM за умови застосування розробленої моделі мобільності на етапі планування мережі показує високий коефіцієнт доставки пакетів, і складає 95,67 %; затримка передачі становить 0,14 с; та найменше енергоспоживання на рівні 12,45 Дж.

Висновки. У статті було проведено комплексне дослідження ефективності методологічних підходів ієрархічного інтелектуального управління наземно-повітряною комунікаційною Ad-Hoc мережею спеціального призначення.

Результати статистичного аналізу показали, що запропонована модель ІСУ НПМ на основі RL – FA-OSELM адекватна.

Ключовим аспектом застосування запропонованої ІСУ є її навчання на етапі планування та донавчання в процесі розгортання та оперативного управління.

Процес навчання запропонованої ІСУ на розробленій моделі мобільності здійснюється на репрезентативній вибірці вихідних даних порівняно меншої розмірності, ніж при існуючих альтернативних моделях.

Таким чином, результати дослідження демонструють ефективність запропонованих підходів побудови та застосування моделі RL – FA-OSELM для ієрархічного інтелектуального

управління наземно-повітряною комунікаційною Ad-Hoc мережею спеціального призначення. Ця модель забезпечує високу продуктивність, надійність та адаптивність мережі в динамічних умовах, що є критично важливим для успішного функціонування таких мереж у реальних сценаріях застосування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. A 3D Smooth Random Walk Mobility Model for FANETs / N. Lin et al. *2019 IEEE 21st International Conference on High Performance Computing and Communications; IEEE 17th International Conference on Smart City; IEEE 5th International Conference on Data Science and Systems (HPCC/SmartCity/DSS)*, Zhangjiajie, China, 10–12 August 2019. URL: https://doi.org/10.1109/hpcc/smartcity/dss.2019.00075.

2. Bani M., Alhuda. Flying Ad-Hoc Networks: Routing Protocols, Mobility Models, Issues. *International Journal of Advanced Computer Science and Applications*. 2016. Vol. 7, no. 6. URL: https://doi.org/10.14569/ijacsa.2016.070621.

3. A mobility model for UAV ad hoc network / O. Bouachir et al. *2014 International Conference on Unmanned Aircraft Systems (ICUAS)*, Orlando, FL, USA, 27–30 May 2014. URL: https://doi.org/10.1109/icuas.2014.6842277.

4. Flying ad-hoc network application scenarios and mobility models / A. Bujari et al. *International Journal of Distributed Sensor Networks*. 2017. Vol. 13, no. 10. P. 155014771773819. URL: https://doi.org/10.1177/1550147717738192.

5. Bujari A., Palazzi C. E., Ronzani D. FANET Application Scenarios and Mobility Models. *the 3rd Workshop*, Niagara Falls, New York, USA, 23 June 2017. New York, New York, USA, 2017. URL: https://doi.org/10.1145/3086439.3086440.

6. Bieliakov R., Fesenko O. Модель мобільності наземної комунікаційної мережі спеціального призначення. *Computer-Integrated Technologies: Education, Science, Production*. 2023. № 51. С. 130–138. URL: https://doi.org/10.36910/6775-2524-0560-2023-51-17 (дата звернення: 25.12.2023).

7. Romaniuk V. A., Bieliakov R. O. Objective control functions of FANET communication nodes of land-air network. *Computer-Integrated Technologies: Education, Science, Production*. 2023. No. 50. P. 125–130. URL: https://doi.org/10.36910/6775-2524-0560-2023-50-19 (date of access: 25.12.2023).

8. Методологія синтезу інтелектуальних систем управління вузлами перспективних мобільних радіомереж з динамічною топологією / В. А. Романюк та ін. *Збірник наукових праць ХУПС*. 2012. № 4 (33). С. 112–116.

9. Бєляков Р., Фесенко О. Модель інтелектуального управління ресурсами наземної комунікаційної мережі класу manet. *Information technology and society*. 2023. № 3 (9). С. 6–14. URL: https://doi.org/10.32689/maup.it.2023.3.1 (дата звернення: 05.05.2024).

10. Bieliakov R. Ієрархічна модель інтелектуального управління наземно-повітряної комунікаційної мережі спеціального призначення. *Computer-integrated technologies: education, science, production*. 2024. № 54. С. 225–235. URL: https://doi.org/10.36910/6775-2524-0560-2024-54-28 (дата звернення: 05.05.2024).

11. Bieliakov R. Проблема інтеграції повітряної мережі класу FANET в мобільну комунікаційну мережу спеціального призначення. *Computer-integrated technologies: education, science, production*. 2023. № 53. С. 263–276. URL: https://doi.org/10.36910/6775-2524-0560-2023-53-40 (дата звернення: 05.05.2024).

12. Бєляков Р. О., Фесенко О. Д. Концептуальна модель управління наземно-повітряною мережею manet і fanet класів спеціального призначення. *Вісник Херсонського національного технічного університету*. 2024. № 1 (88). С. 203–210. URL: https://doi.org/10.35546/kntu2078-4481.2024.1.28 (дата звернення: 05.05.2024).

13. Bieliakov R., Fesenko O. FANET management process simulation at the deployment and operation stage. *Technology audit and production reserves*. 2023. Vol. 5, no. 2(73). P. 40–47. URL: https://doi.org/10.15587/2706-5448.2023.290033 (date of access: 07.05.2024).

14. Bieliakov R. O., Fesenko O. D. Improved model of intelligent management of node resources of the terrestrial communication network of the MANET class. *Scientific notes of Taurida National V. I. Vernadsky University. Series: Technical Sciences*. 2023. No. 5. P. 93–98. URL: https://doi.org/10.32782/2663-5941/2023.5/16 (date of access: 05.05.2024).

15. Bieliakov R. O. Development of a routing method for ground-air Ad-Hoc network of special purpose. *Technology audit and production reserves*. 2024. Vol. 2, no. 2(76). P. 44–51. URL: https://doi.org/10.15587/2706-5448.2024.302394 (date of access: 07.05.2024).

16. Camp T., Boleng J., Davies V. A survey of mobility models for ad hoc network research. *Wireless Communications and Mobile Computing*. 2002. Vol. 2, no. 5. P. 483–502. URL: https://doi.org/10.1002/wcm.72.

17. A group mobility model for ad hoc wireless networks / X. Hong et al. *the 2nd ACM international workshop*, Seattle, Washington, United States, 20 August 1999. New York, New York, USA, 1999. URL: https://doi.org/10.1145/313237.313248.

18. Yujin Li, Ming Zhao, Wenye Wang. Internode Mobility Correlation for Group Detection and Analysis in VANETs. *IEEE Transactions on Vehicular Technology*. 2013. Vol. 62, no. 9. P. 4590–4601. URL: https://doi.org/10.1109/tvt.2013.2264689.

19. Choffnes D. R., Bustamante F. E. An integrated mobility and traffic model for vehicular wireless networks. *the 2nd ACM international workshop*, Cologne, Germany, 2 September 2005. New York, New York, USA, 2005. URL: https://doi.org/10.1145/1080754.1080765.

20. A Smooth-Turn Mobility Model for Airborne Networks / Y. Wan et al. *IEEE Transactions on Vehicular Technology*. 2013. Vol. 62, no. 7. P. 3359–3370. URL: https://doi.org/10.1109/tvt.2013.2251686.

21. Medina A., Matta I., Byers J. On the origin of power laws in Internet topologies. *ACM SIGCOMM Computer Communication Review*. 2000. Vol. 30, no. 2. P. 18–28. URL: https://doi.org/10.1145/505680.505683.

22. Towards realistic mobility models for mobile ad hoc networks / A. Jardosh et al. *the 9th annual international conference*, San Diego, CA, USA, 14–19 September 2003. New York, New York, USA, 2003. URL: https://doi.org/10.1145/938985.939008.

23. Biao Zhou, Kaixin Xu, Gerla M. Group and swarm mobility models for ad hoc network scenarios using virtual tracks. *IEEE MILCOM 2004. Military Communications Conference,* 2004. Monterey, CA, USA. URL: https://doi.org/10.1109/milcom.2004.1493283.

24. Kraaier J., Killat U. The random waypoint city model. *the 3rd ACM international workshop*, Cologne, Germany, 2 September 2005. New York, New York, USA, 2005. URL: https://doi.org/10.1145/1080730.1080749.

25. Stuedi P., Alonso G. Log-normal shadowing meets SINR: A numerical study of Capacity in Wireless Networks. *2007 4th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, 18–21 June 2007. 2007. URL: https://doi.org/10.1109/sahcn.2007.4292867.

26. Doci A., Xhafa F. WIT: A Wireless Integrated Traffic Model. *Mobile Information Systems*. 2008. Vol. 4, no. 3. P. 219–235. URL: https://doi.org/10.1155/2008/341498.

27. Introduction to Probability and Statistics for Engineers and Scientists. Elsevier, 2014. URL: https://doi.org/10.1016/c2013-0-19397-x.

DOI: 10.58254/viti.5.2024.03.37

УДК 621.396

Бернацький А. П. ORCID: 0000-0003-0379-075X (ВІТІ ім. Героїв Крут)

ВДОСКОНАЛЕНИЙ МЕТОД ПЛАНУВАННЯ ШЛЯХУ АВТОНОМНОГО НАЗЕМНОГО РОБОТА З ВИКОРИСТАННЯМ АЛГОРИТМУ MBD-RRT*FFT

*В дослідженні проведено аналіз проблем, пов'язаних із плануванням шляху переміщення роботів і підвищення точності та надійності їх наведення у режимі реального часу. Задля вирішення цієї проблеми досліджено, проведено вдосконалення модифікації асимптотично оптимального алгоритму BD-RRT*FT. Розроблений алгоритм MBD-RRT*FFT при застосуванні в динамічних середовищах завдяки використанню багатопотокового обчислення має кращі можливості динамічного планування. Використання Fitch's-алгоритму вибору оптимального результату пошуку надає спрямованості вибору оптимального прямолінійного шляху.*

*Задля перевірки ефективності запропонованого методу вдосконалення алгоритму проведено імітаційне моделювання з використанням власної програми моделювання. Проведено порівняння алгоритму MBD-RRT*FFT з іншими алгоритмами за трьома картами з оцінюванням показників продуктивності.*

З метою оцінки поведінки програми, а також ідентифікації вузьких місць виконання алгоритму на всіх етапах перевірки проводився контроль навантаження на пам'ять системи прийняття рішення. Це дозволило побачити підвищення споживання ресурсів пам'яті на 12–15 %, але враховуючи отримані результати щодо оптимальності та швидкості обчислень, зроблено висновки щодо мінімальності впливу. Це надало додаткове розуміння необхідності уважного вибору апаратної складової систем прийняття рішень.

*Наукова новизна методу полягає у застосуванні науково-методичного апарату з вдосконалення алгоритму BD-RRT*FT, що надало подальший розвиток щодо розширення та доповнення відомих даних про алгоритми з плануванням шляху робота і підвищення точності та надійності його наведення у режимі реального часу. Метод охоплює способи дослідження, систематизацію, коригування нових і отриманих раніше знань.*

Ключові слова: *алгоритм, багатопотоковий, швидке дослідження, випадкове дерево, планування шляху в реальному часі, жадібний пошук, відновлення шляху, асимптотичний, Fitch's-оптимізація, БАНЗ, робот.*

A. Bernatskyi An improved method planning path of an autonomous ground robot with using the MBD-RRT*FFT algorithm.

*The study analyzed the problems associated with planning the path of robots and increasing the accuracy and reliability of their guidance in real time. To solve this problem, a practical study was conducted, based on the obtained analysis, a modification of the asymptotically optimal BD-RRT*FT algorithm and was proposed and improved the MBD-RRT*FFT algorithm. was developed. Which, when applied in dynamic environments due to the use of multi-threaded computations, has better dynamic scheduling capabilities, and the use of Fitch's optimal search result selection algorithm provides a general tendency to choose the optimal straight-line path.*

*To verify the effectiveness of the proposed algorithm improvement method, simulations were carried out using our own simulation program. A comparison of the MBD-RRT*FFT algorithm with other algorithms was carried out on three maps with an evaluation of performance indicators.*

In order to evaluate the behavior of the program, as well as to identify bottlenecks in the execution of the algorithm, the memory load of the decision-making system was monitored at all stages of the verification. This made it possible to see an increase in the consumption of memory resources by 12-15%, but taking into account the obtained results regarding the optimality and speed of calculations, conclusions were made regarding the minimal impact. Additional understanding of the need for careful selection of the hardware component of decision-making systems is provided.

*The scientific novelty of the method consists in the application of a scientific and methodological apparatus for improving the BD-RRT*FT algorithm, which ensured further development in terms of expanding and supplementing known data on algorithms with robot path planning and increasing accuracy and reliability in real time. The method includes methods of research, systematization and adjustment of new and previously acquired knowledge.*

Keywords**:** *algorithm, multiprocessing, fast exploration, random tree, real-time path planning, greedy search, path recovery, asymptotic optimality, fitch's optimization, robot, UGV.*

Постановка завдання. В розрізі парадигми Стратегії воєнної безпеки України «Воєнна безпека – всеохоплююча оборона» в короткостроковій перспективі приділяється особлива

увага до «розроблення, виробництва й оснащення сил оборони сучасним озброєнням, військовою та спеціальною технікою, забезпечення засобами ураження, у тому числі безпілотними і роботизованими...» [6].

Завдання, що покладаються на безпілотні автономні наземні засобі (БАНЗ) під час проведення бойових дій в урбанізованому просторі в умовах щільної забудови міста з постійною зміною ландшафту, мають загальну тенденцію до постійного збільшення вимог до складності в режимі реального часу, а сценарії використання БАНЗ демонструють диверсифіковану тенденцію розвитку інтелектуальних здібностей [5], що вимагає розробки військових мобільних робототехнічних систем із системами прийняття рішень, здатними на самостійне (автономне) вирішення проблем пошуку шляху.

Враховуючи вище зазначене, сучасною актуальною науковою задачею [1–6; 22] постає розробка ефективних алгоритмів пошуку маршруту переміщення БАНЗ без зіткнень в умовах вузьких ділянок урбанізованого середовища та оперативного простору з перешкодами, що динамічно змінюються [1].

Таким чином, стаття призначена опису методу модифікації алгоритму BD-RRT*FN, який можна застосовувати в динамічних середовищах з метою покращення часу планування алгоритму, довжини рішення шляху, форми шляху, швидкості конвергенції, задля задоволення вимог планування шляху в реальному часі, дозволяючи БАНЗ швидко отримати оптимальний шлях без зіткнень у динамічних середовищах у режимі реального часу.

Аналіз останніх публікацій. Відомо багато прикладів досліджень проблематики планування шляху роботів [1; 2; 7; 10–12; 14–17; 19–21]. У роботі [1] Бернацьким А. П. проведено поглиблений аналіз і запропоновано метод планування шляху автономного наземного робота з використанням модифікації динамічного двонаправленого алгоритму RRT*FN. Під час експериментальної перевірки була доведена ефективність алгоритму BD-RRT*FT, що задовольняє вимогам планування шляху в реальному часі, дозволяючи БАНЗ швидко отримати оптимальний шлях без зіткнень у динамічних середовищах.

У запропонованому методі під час модифікації акцентована увага на загальну проблематику вибору фіксованої кількості вузлів відбору точок вибірки. Таке рішення є відносно розумним завдяки використанню порівняльного вибору, мінімального за кількістю точок вибірки з використанням випадкового $p_{rand}(x)$, вибраного алгоритмом RRT*, що дозволяє уникнути впливу неправильного вибору. Але на рівні початкового припущення мінімальної оцінки на вибірковому етапі ми не розуміємо кількість майбутніх ітерацій загального пошуку довжини шляху через циклічність виконання пошуку зіткнення або закінчення виконання завдання. Це спонукає до того, що початковий шлях пошуку найменшої вибірки може бути хибним і становити «максимально велике» фінальне значення. Це також призведе до сповільнення швидкості роботи внаслідок взаємозв'язку, при якому чим більший розмір дерева, тим більше навантаження на динамічний алгоритм при оновленні в реальному часі, коли алгоритм динамічного пошуку оновлює інформацію про навколишнє середовище, кількість точок відбору проб поступово збільшуватиметься з часом.

Метою статті є модифікація алгоритму пошуку шляху BD-RRT*FT в умовах динамічного урбанізованого середовища.

Виклад основного матеріалу. В роботі [1] запропонована модифікація алгоритму RRT*FN із впровадженням модифікації алгоритму та застосуванням принципу динамічного оновлення та відновлення шляху. А саме під час кожної ітерації алгоритм оновлюватиме інформацію про середовище та виконуватиме оновлення планування на основі вихідних вузлів. Крім того, також проводиться оновлення інформації про розташування БАНЗ. Щоб зменшити розмір дерева під час процесу планування, непотрібні вузли і допоміжні гілки, через які проходить БАНЗ, маркуються як *непотрібні*, використовуючи поточне розташування БАНЗ як новий кореневий вузол. Якщо запланований шлях знищено перешкодами, алгоритм

відкине вузли на дереві, покриті перешкодами, і використає простір для перебудови дерева з повними вузлами й ефективного відновлення шляху.

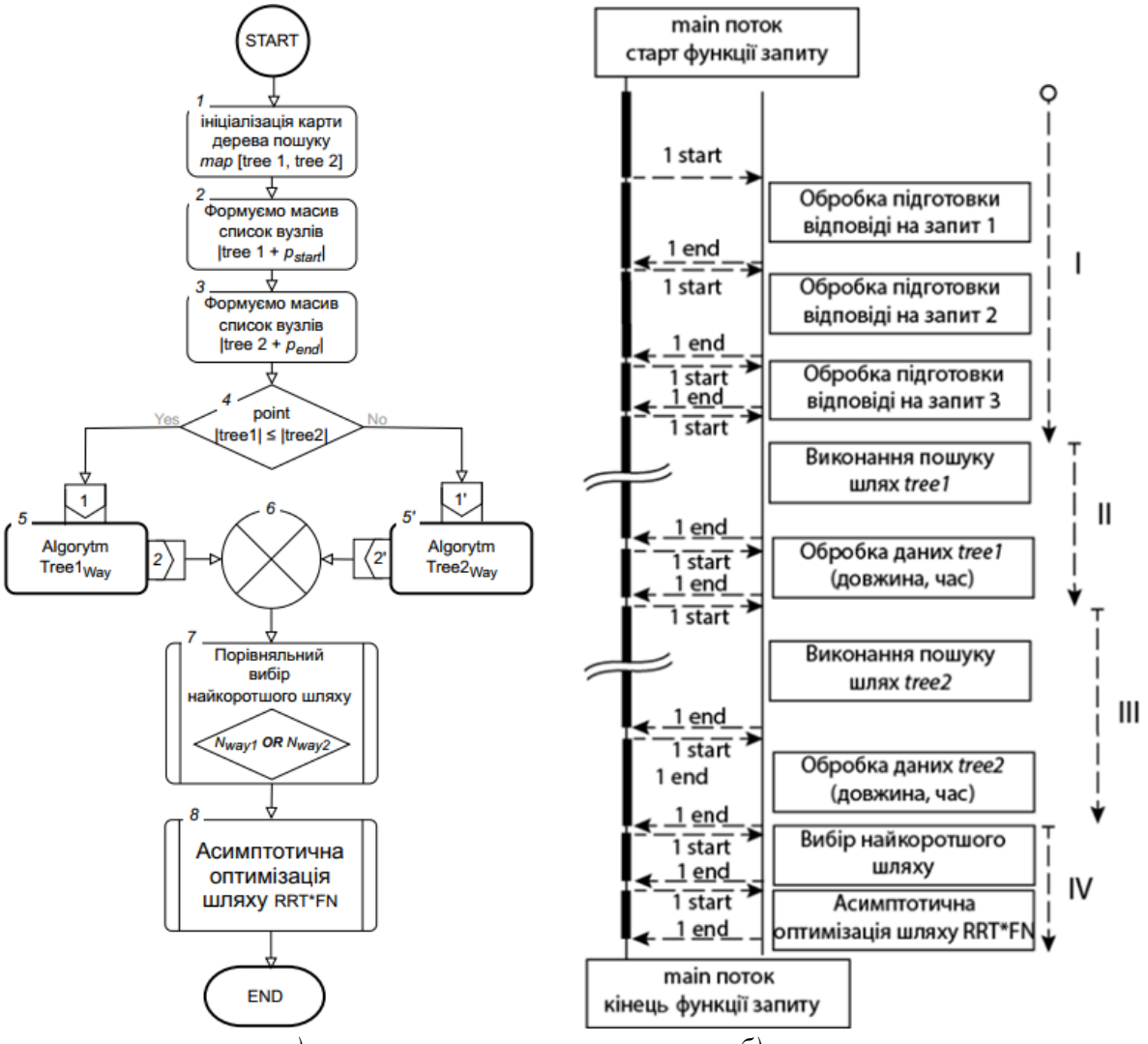
Було визначено, що алгоритм BD-RRT*FT має ймовірнісну повноту та асимптотичну оптимальність. Хоча алгоритм має ймовірнісну повноту, у діапазоні m⩽n<1, якщо кількість реперних вузлів встановлена надто малою, ймовірність отримання шляхового рішення майже дорівнює 0.

$$\begin{cases} L\{p_1, p_2, \dots, p_n\} = \sigma^*, \\ \forall p \in \sigma^*, p_{rand}(x) = p, x \in (0, iter). \end{cases} \quad (1)$$

Але якщо значення реперного вузла встановлено занадто великим, воно займатиме забагато місця в пам'яті. Зазвичай кількість реперних вузлів вибирається на основі емпіричних методів. Але для уникнення людського фактору модифікований алгоритм автоматично перевіряє кілька груп за умови фіксованого простору, станів і розміру кроку.

Через випадковість точок вибірки, вибраних алгоритмом RRT, метод вибору фіксованої кількості вузлів може бути не оптимальним, але цей метод відбору є відносно розумним і дозволяє уникнути впливу неправильного вибору фіксованої кількості вузлів для експерименту.

Загальна реалізація алгоритму BD-RRT*FT показана на рисунку 1.



а) *б)*

Рис. 1. Алгоритм BD-RRT*FT [1]:
а – блок-схема алгоритму; *б* – функціональний часовий потоковий розріз виконання алгоритму

Алгоритм BD-RRT*FT має явно виражену асинхронну властивість, яка є запозиченою з батьківського алгоритму RRT*. Така батьківська властивість вимагає послідовних однопотокових обчислень. На рівні початкового припущення мінімальної оцінки на вибірковому етапі кількість майбутніх ітерацій загального пошуку довжини шляху не відома внаслідок циклічності виконання пошуку зіткнення або закінчення виконання завдання (рис. 2). Через це початковий шлях пошуку найменшої вибірки може бути хибним і становити «максимально велике» фінальне значення, що призведе до сповільнення швидкості роботи внаслідок взаємозв'язку, при якому чим більший розмір дерева, тим більше навантаження на динамічний алгоритм при оновленні в реальному часі.

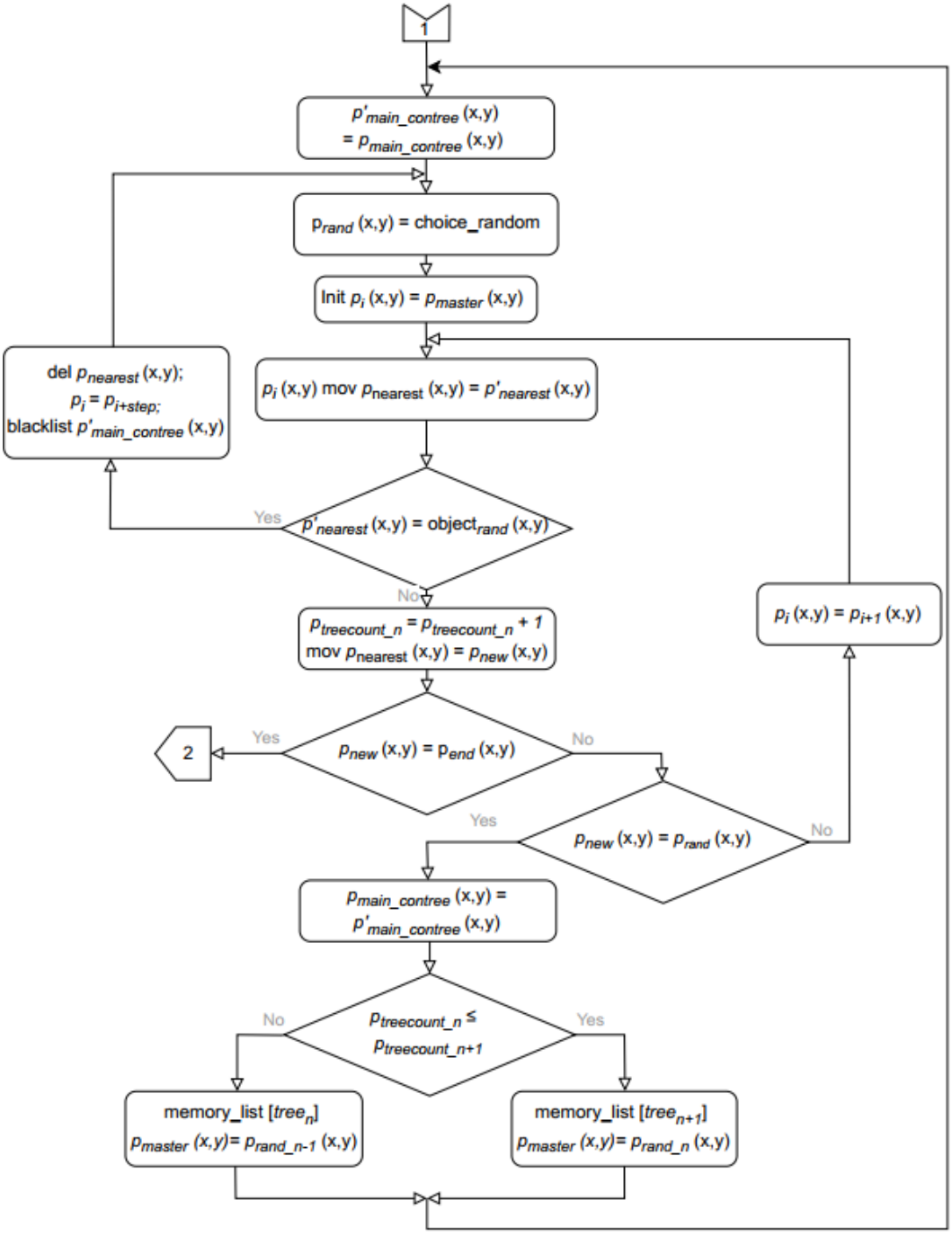
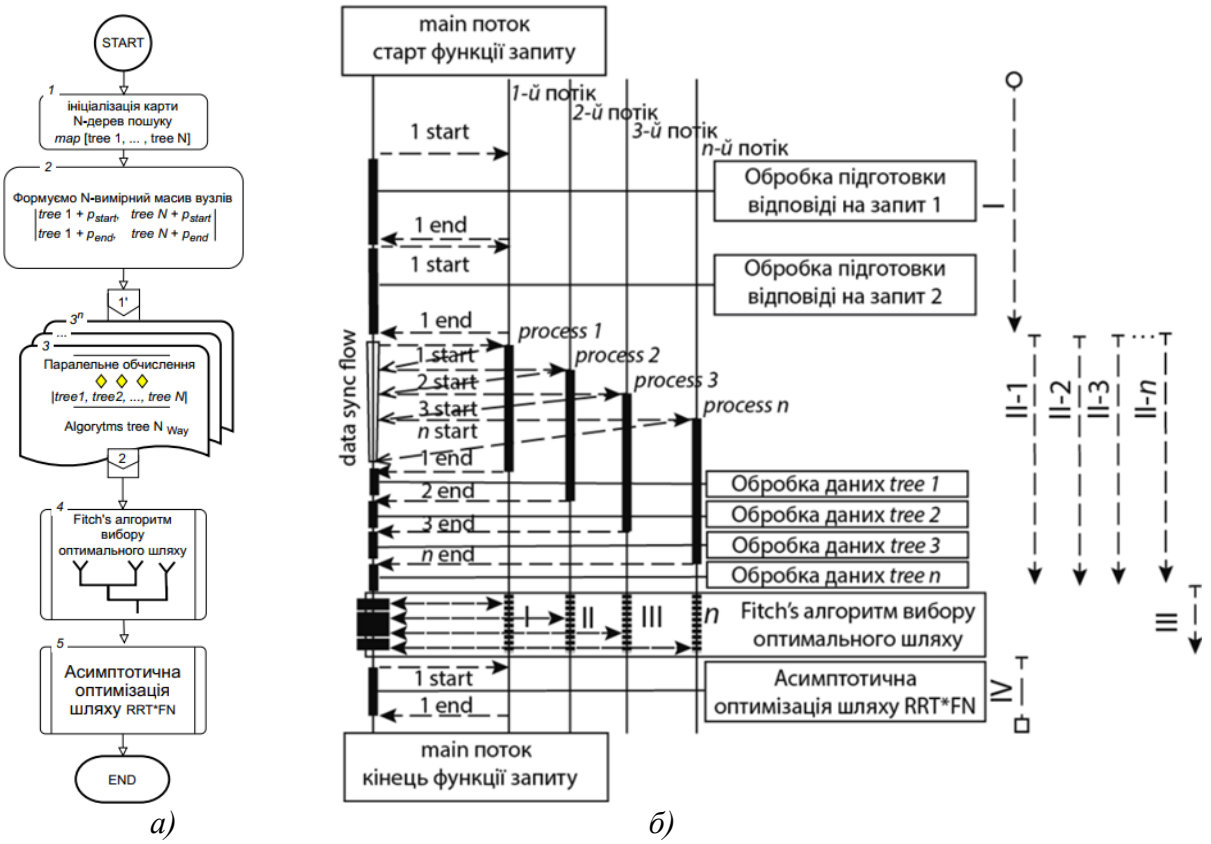


Рис. 2. Представлення у вигляді блок-схеми виконавчого блоку алгоритму пошуку шляху алгоритму BD-RRT*FT [1]

Асинхронна алгоритмізація та багатопоточність часто використовуються для досягнення схожих цілей, але між ними є ключові відмінності. Асинхронні алгоритми дозволяють продовжувати виконання операцій, не чекаючи завершення тривалої операції, що покращує чуйність. Але щоб вирішити вищеозначену проблему, ми можемо використовувати багатопотоковість, яка дозволяє алгоритму створювати кілька потоків для одночасної обробки кількох завдань пошуку шляхів.

Пропонується вдосконалений алгоритм, що використовує стратегію алгоритму BD-RRT*FT двостороннього жадібного пошуку шляху з принципом динамічного оновлення та відновлення шляху з додаванням блока паралельних багатопотокових обчислень.

Реалізація багатопокового алгоритму MBD-RRT*FFT зображена на рисунку 3.



а) *б)*

Рис. 3. Алгоритм MBD-RRT*FFT:
а – блок-схема алгоритму; *б* – функціональний часовий потоковий розріз виконання алгоритму

Опис алгоритму MBD-RRT*FFT:

Крок 1. Ініціалізуємо карту та дерева пошуку $|tree_x|$. Враховуючи, що подальший пошук планується виконувати багатопотоково в паралельних площинах, кількість дерев буде дорівнювати кількості потоків обчислення, але не менше 2, тобто $N_{tree} \vDash N_{flow} \geq 2$. Додаємо p_{start} до списку вузлів $|tree_1 ... tree_{N-1}|$ і p_{end} до списку вузлів $|tree_2 ... tree_N|$.

Крок 2. Випадковим чином вибираємо точку відбору проб p_{rand} у просторі, завдяки чому формуємо масив проб для початкової ініціалізації потоків обчислень.

Крок 3...3^n. В кожному окремому потоці вибираємо найближчий вузол $p_{nearest\_n}$ й починаємо рухатися до p_{rand_n}, зростаючи з певним кроком, і отримуємо p_{new_n}, де n – номер потоку, визначений на етапі *Крок 1*. Після чого виконується послідовність виконавчого блоку алгоритму пошуку шляху алгоритму BD-RRT*FT (див. рис. 2) [1].

Крок 4. В зв'язку з тим, що в алгоритмі BD-RRT*FT використовується асинхронна послідовна структура обчислень, для аналізу залучаються не більше 2 попередньо відібраних

вибірок шляхів, а *блок 7* (див. рис. 1) використовує метод порівняльного аналізу з вибіркою найкоротшого шляху. Але враховуючи, що модифікований алгоритм MBD-RRT*FFT на етапі находження оптимального найкоротшого шляху (*блок 4*, рис. 3) задіює багатопотокові паралельні обчислення, то використання послідовного порівняльного аналізу є недоречним і тому для аналізу *пропонується додаткова модифікація* завдяки використанню *Fitch's*-алгоритму вибору оптимального шляху із залученням в якості базового методу аналізу метод фільтрації (*filter methods*).

Крок 5. Виконуємо асимптотичну оптимізацію шляху відповідно до кроків 1–3 алгоритму RRT*FN [1; 7; 10; 14; 20].

Важливим доповненням є те, що оновлення карти після кожного завершення зростання/повторного підключення, яке ми виконували на етапі 9 алгоритму BD-RRT*FT, в цій модифікації виконується під час виконання *Кроку 3…3*^{n}. А вузли, що покрити перешкодами й вузли, лінії яких перетинаються з перешкодами, не видаляються, а маркуються як *black_list*.

Враховуючи те, що модифікація алгоритму стосується тільки методів обчислення та прикінцевого вибору оптимального шляху, а описаний в роботі [1] виконавчий елемент пошуку залишається без змін, імовірнісна повнота та асимптотична оптимальність алгоритму буде наслідувати батьківську властивість [8; 9; 18; 21; 23], а саме [1; 2]:

$$\forall p \in \sigma, p_{rand}(x) = p, x \in (0, iter). \tag{2}$$

Кількість ітерацій внаслідок певної циклічності достатньо велика, рівняння (3) залишається вірним:

$$\begin{cases} L\{p_1, p_2, \dots, p_n\} = \sigma^*, \\ \forall p \in \sigma^*, p_{rand}(x) = p, x \in (0, iter). \end{cases} \tag{3}$$

Враховуючи зазначене, додаткову перевірку та повторне доведення ймовірнісної повноти та асимптотичної оптимальності алгоритму RRT* [1; 2] в розрізі цього дослідження автор описувати не буде.

На думку автора, цікавим елементом модифікації є додавання *Fitch's*-алгоритму вибору оптимального шляху із залученням категорії *Метод фільтрації* (*filter methods*) в якості базового методу аналізу на етапі пошуку оптимального найкоротшого шляху БАНЗ (*Блок 4,* рис. 3). Доведено [1; 19; 20], що якщо фінальних вибірок дуже багато, маємо суттєве збільшення часу роботи класифікатора. Тому якщо необхідно протестувати кілька класифікаторів з метою вибору кращого, то час, необхідний для обчислення, може стати просто величезним. Для подолання цієї проблематики пропонується використати *Fitch's*-алгоритм.

Загально відомо [9; 23; 24], що метод відбору *Fitch's* поділений на три категорії: *filter methods, wrapper methods* й *embedded methods.*

В зв'язку з тим, що *метод фільтрації (filter methods)* ґрунтується на статистичних методах, і, як правило, розглядають кожну вибірку незалежно, що дозволяє оцінити і ранжувати вибірки за значимістю, за яку приймається ступінь кореляції цієї вибірки з цільової змінної, задля наших цілей будемо використовувати саме цю категорію методу.

Формула ентропії точки вибірки $tree_{rand}$ для n-мірної мапи визначається як:

$$H(tree(X)) = -\sum_{x_i \in X} p(tree(x_i)) \times \log_2(p(tree(x_i)), \tag{4}$$

де $p(tree(x_i))$ – ймовірність того, що змінна $tree(X)$ прийме значення $tree(x_i)$. Така можливість розглядається на *Кроці 3…3*^{n} з урахування того, що $tree(X) = tree(x_i)$ і розділене на загальну кількість умовно випадкових кроків.

Для розрахунку кореляції між змінними потрібно визначити специфічну умовну ентропію (*specific conditional entropy*), відносну ентропію (*conditional entropy*) і значення інформаційного приросту (*information gain*).

$$H(tree(X_{new}) \mid tree(X) = tree(x_i)), \quad (5)$$

де ентропія $H(Tree(X_{new}))$ визначає лише ті записи, в котрих $tree(X) = tree(x_i)$.

Відносна ентропія (*conditional entropy*) вважається як:

$$H(tree(X_{new}) \mid tree(X)) = \sum_{tree(x_i) \in tree(X)} p(tree(x_i)) \times H(tree(X_{new}) | tree(X) = tree(x_i)), \quad (6)$$

У досліджені ця величина цікавить не як окреме значення, а саме як її різниця із звичайною ентропією вибірки $tree(X_{new})$.

Нам є важливим, наскільки більш упорядкованою стає змінна $tree(X_{new})$, якщо ми знаємо значення *X*. Або, простіше кажучи, чи існує кореляція між значеннями *X* і X_{new}, і наскільки вона велика. Про це говорить величина значення інформаційного приросту (*information gain) P_{IG}*:

$$P_{IG}(X_{new}|X) = H(X_{new}) - H(X_{new}|X), \quad (7)$$

Чим більший параметр P_{IG}, тим сильніша кореляція. Таким чином, ми легко можемо обчислити величину *значення інформаційного приросту* для всіх вибірок і відкинути ті, що слабко впливають на цільову змінну. Завдяки чому планується скоротити час розрахунку шляху.

Експериментальна перевірка алгоритму. Імітаційне моделювання було здійснено завдяки розробленій програмі моделювання з використанням мови програмування Python 3+ та бібліотек PyGame, NumPy і ***Multiprocessing***. В якості системи обчислювання використовувалась система 2-сокетна Workstation із системними параметрами Intel® Xeon® Processor E5-2689 v4 (25M Cache, 3.10 GHz) =2/ 4*64 Gb DDR4 2133/ SSD 250Gb/ VA Asus GeForce GTX 1650 GDDR6 4096Mb.

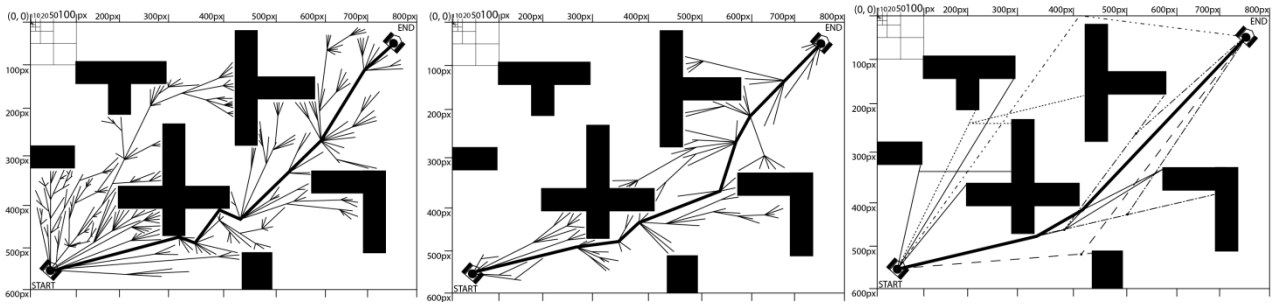
Щоб реалізувати необхідну багатопоточність в Python задля перевірки запропонованого алгоритму, було обрано бібліотеку *multiprocessing*. Цей вибір пов'язано з тим, що модуль *multiprocessing* дозволяє створювати, керувати та виконувати процеси окремо у власному просторі пам'яті. Це означає, що кожен процес може виконуватися як незалежно, так і мати залежність від результатів виконання інших процесів.

Подібно до роботи [1], задля перевірки ефективності запропонованого алгоритму проведено порівняння алгоритму MBD-RRT*FFT з іншими алгоритмами за трьома картами з оцінюванням показників продуктивності. Щоб полегшити аналіз симуляції та врахувати раціональність, перешкоди на мапах використані як блоки піксельної графіки, БАНЗ розглядається як точка піксельного розміру 4х4 рх, а інші нерелевантні змінні, окрім алгоритму, контролюються, щоб бути узгодженими. Оскільки алгоритм RRT базується на випадковій вибірці, у процесі моделювання є непередбачуваність, і кожен результат вимірювання може мати відмінності, щоб усунути вплив випадковості, під час експерименту проводилось 100 незалежних експериментів для кожної ситуації та формувалися результати для порівняльного аналізу.

Перевірка роботи алгоритмів на карті 1 (Карта зі статичними перешкодами)

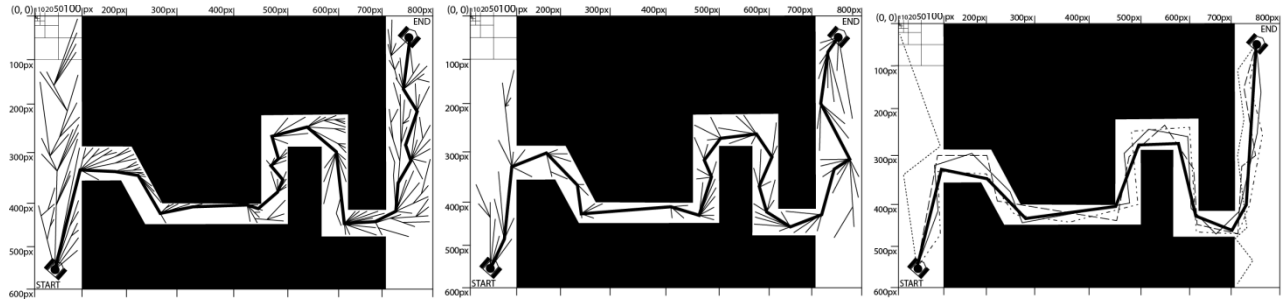
Використовуємо карту 1, розміром 800 рх × 600 рх, як звичайну карту статичних перешкод (рис. 4, *а*), яка використовується для перевірки швидкості реакції та індексу шляху алгоритму в нормальному середовищі перешкод. На цій карті верхній лівий кут встановлюється як початок координат [0, 0], а карта розташована в 4-му квадранті, початкова

та кінцева позиція БАНЗ [20, 580], [780, 20] позначена значками БАНЗ та надписами START, END відповідно.



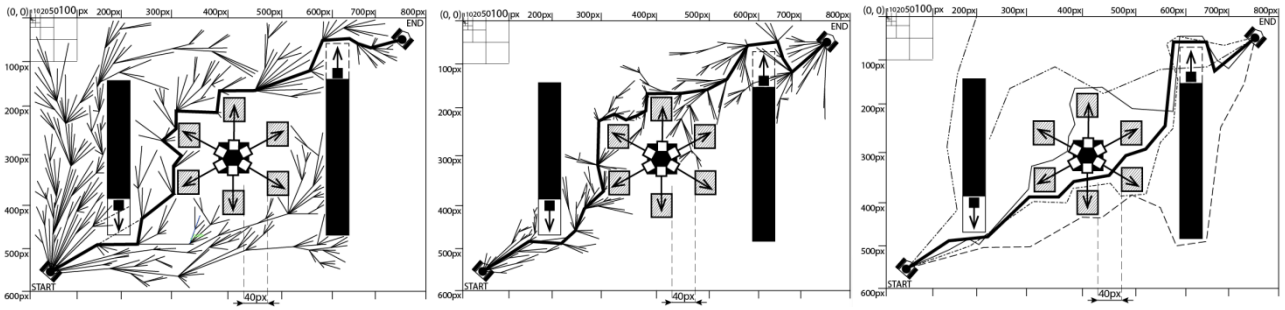
*RRT*FN* *BD-RRT*FT* *MBD-RRT*FFT*

а) візуалізація результату роботи алгоритмів на карті 1



*RRT*FN* *BD-RRT*FT* *MBD-RRT*FFT*

б) візуалізація результату роботи алгоритмів на карті 2



*RRT*DWA* *BD-RRT*FT* *MBD-RRT*FFT*

в) візуалізація результату роботи алгоритмів на карті 2

Рис. 4. Візуалізація результату роботи алгоритмів планування шляху за трьома типами карт:
а – карта 1: середовище статичних перешкод; *б* – карта 2: простір вузьких тунелів;
в – карта 3: середовище динамічних перешкод

Перевіряються алгоритми RRT*FN, BD-RRT*FT та MBD-RRT*FFT. Результати планування шляху показано на рисунку 4, *а*. Товста лінія позначає отримане рішення шляху. Інші дерева є ітераційними обчисленнями.

Як і в досліджені [1], алгоритм RRT*FN, незважаючи на двосторонність, має занадто багато надлишкових точок вибірки в процесі планування. Для інтуїтивного порівняння, середня довжина шляху рішення отримана зі 100 експериментів.

Алгоритм RRT*FN, алгоритм BD-RRT*FT і алгоритм MBD-RRT*FFT належать до алгоритмів оптимального рішення, за результатами в таблиці 1 бачимо, що багатопотоковий двонаправлений алгоритм MBD-RRT*FFT має певні переваги у швидкому пошуку рішення.

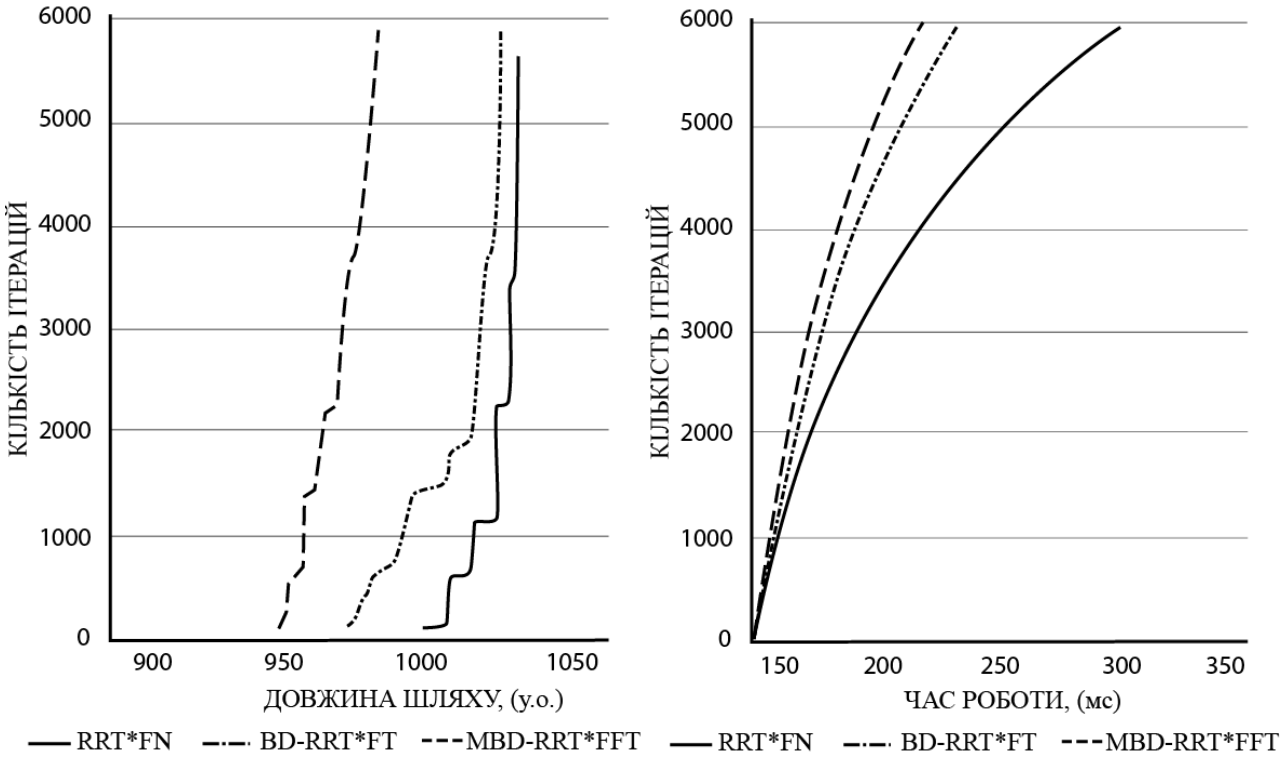
Враховуючи, що вдосконалений алгоритм використовує стратегію *Fitch’s*-вибору, загальна вартість шляху суттєво зменшується і набуває ще більш прямолінійного характеру.

На рисунку 5, *а, б* показано графічний порівняльний аналіз продуктивності кожного алгоритму на карті 1.

Таблиця 1

Дані порівняння продуктивності алгоритму на карті 1

Алгоритм	RRT*FN	BD-RRT*FT	MBD-RRT*FFT
Середня довжина шляху, *у. о.*	1032,02	1003,43	962,825
Середній час роботи, *мс*	299,4	236,6	226,9
Споживана пам’ять, *Мб*	386,53	348,51	397, 81



а) *б)*

Рис. 5. Порівняльний аналіз продуктивності кожного алгоритму на карті 1:
а – зв’язок між кількістю ітерацій та довжиною шляху;
б – зв’язок між кількістю ітерацій та часом виконання

Перевірка роботи алгоритмів на карті 2 (карта з вузьким коридором).

Використовуємо карту 2 розміром 800 px × 600 px, як карту з вузьким коридором (див. рис. 4, *б*), що імітує середовище, близьке до міського тунелю або печер. Загальні налаштування карти 2, за винятком конфігурації розташування перешкод на карті, такі самі, як і для карти 1.

Візуалізація результатів планування шляху алгоритмами RRT*FN, BD-RRT*FT та MBD-RRT*FFT наведена на рисунку 4, *б*. Як і у попередньому експерименті товста лінія позначає отримане рішення шляху. Інші дерева є ітераційними обчисленнями.

Для карт із вузьким коридором простору, через сліпоту зростання й завдяки односторонньому зростанню, RRT*FN алгоритму важко отримати точки вибірки у вільному просторі в межах вузького каналу, в результаті чого кількість ітерацій і час виконання набагато вищі, ніж у алгоритмів з двостороннім характером BD-RRT*FT та MBD-RRT*FFT, а точки

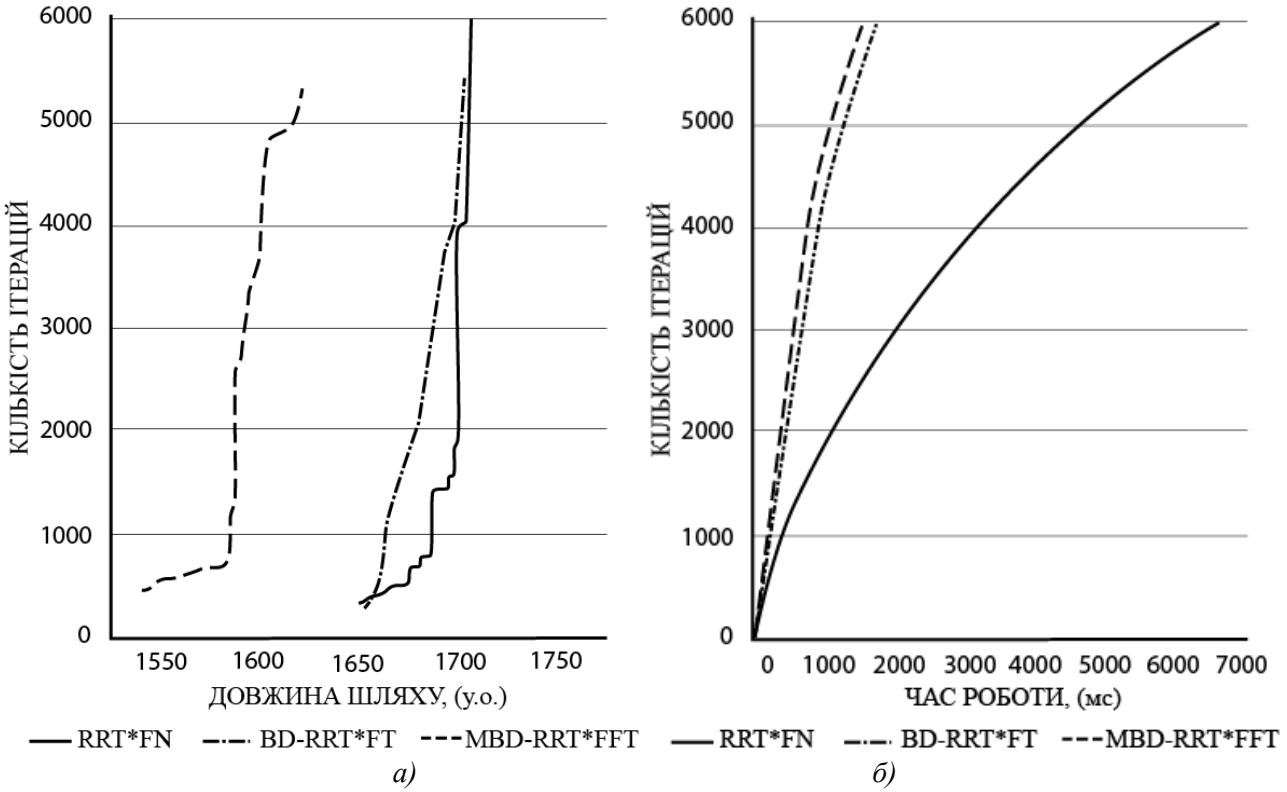
вибірки здебільшого зосереджені в лівому вільному просторі, як і в [1], що підтверджує правильність напрямку модифікації алгоритму.

В таблиці 2 наведено отримані під час експерименту результати дослідження пошуку шляху в умовах вузького середовища.

Таблиця 2

Порівняння показників ефективності алгоритму на карті 2

Алгоритм	RRT*FN	BD-RRT*FT	MBD-RRT*FFT
Середня довжина шляху, *у. о.*	1712,95	1711,99	1591,67
Середній час роботи, *мс*	6797	1955	1817
Споживана пам’ять, *Мб*	387,84	339,52	356,75

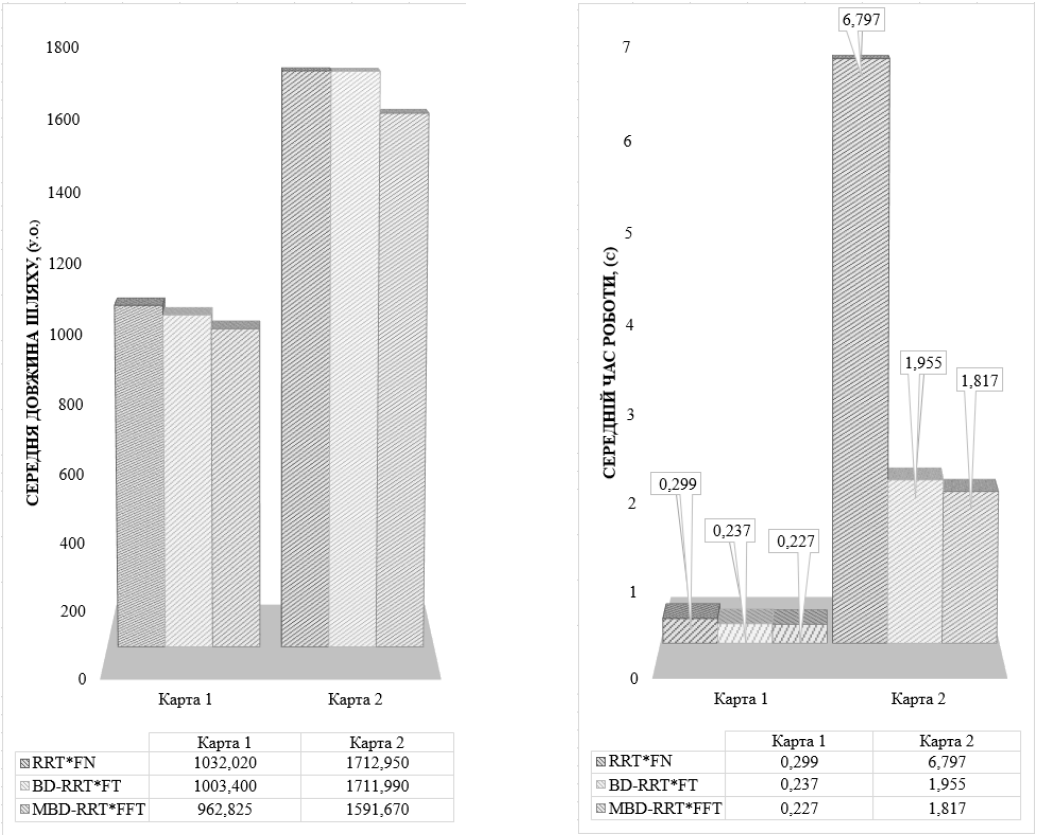


а) *б)*

Рис. 6. Порівняльний аналіз продуктивності кожного алгоритму на карті 2:
а – зв’язок між кількістю ітерацій та довжиною шляху;
б – зв’язок між кількістю ітерацій та часом виконання

Порівняння продуктивності алгоритмів пошуку оптимального рішення шляху для карти 2 та результати ітераційного процесу з урахування виконання асимптотичної оптимізації в 3 алгоритмах показано на рисунку 6, *а, б*. З чого видно, що довжина та швидкість пошуку шляху, отриманого алгоритмом MBD-RRT*FFT, кращі, ніж у алгоритмів RRT*FN і BD-RRT*FT.

При перевірці роботи методу вдосконалення алгоритму, застосовуючи карти зі статично встановленими перешкодами (карта 1 і карта 2), модифікований алгоритм MBD-RRT*FFT, маючи спадкові зв’язки з алгоритмом BD-RRT*FT, показав високу продуктивність порівняно з іншими алгоритмами (рис. 7), через що для подальшої перевірки роботи вдосконаленого алгоритму в режимі наявності динамічних об’єктів було прийнято рішення на часткову заміну набору перевірочних алгоритмів для карти 3.



а) *б)*

Рис. 7. Порівняльні діаграми продуктивності алгоритмів на карті 1 і карті 2:
а – зв’язок між кількістю ітерацій та довжиною шляху;
б – зв’язок між кількістю ітерацій та часом виконання

Перевірка роботи алгоритмів на карті 3 (карта з динамічними об’єктами).

Для подальшого вивчення продуктивності алгоритму в динамічному середовищі створено динамічну карту розміром 600 px × 800 px, що містить три рухомі об’єкти-перешкоди. Так як результати в дослідженні [1] були занадто «приємними», було прийнято рішення карту 3 зробити більш складною. Тому перший і третій об’єкти залишились подібними до карти 3 дослідження [1].

Перший X_1:[180] і третій об’єкти-перешкоди X_3:[580] візуально позначені як чорні прямокутники розміром 40 px × 200 px, що рухаються вперед-назад уздовж осі Y, а діапазон руху становить Y_1:[110–490] px і Y_2:[50–500] px відповідно. Другий об’єкт-перешкода має складний характер, а саме блок-перешкода розміром 40 px × 40 px (центр X_{20}:[400], Y_{20}:[300]) розпадається на шість незалежних рівновіддалених об’єктів з радіусом центрів 100 px.

В якості порівняльних алгоритмів вибрано алгоритм BD-RRT*FT та алгоритм RRT*DWA. Оскільки в динамічному середовищі довжина рішення шляху змінюється нерегулярно, то задля спрощення аналізу приймемо, що лише довжина шляху переміщення динамічної перешкоди буде впливати на глобальний шлях БАНЗ. Середній час виконання побудови шляху та довжина наведені в таблиці 3, графічна візуалізація результатів планування шляху показані на рисунку 4, *в*, а порівняння продуктивності алгоритмів пошуку оптимального рішення шляху для карти 3 та результати ітераційного процесу з урахування виконання асимптотичної оптимізації в 3 алгоритмах показано на рисунку 8, *а, б*.

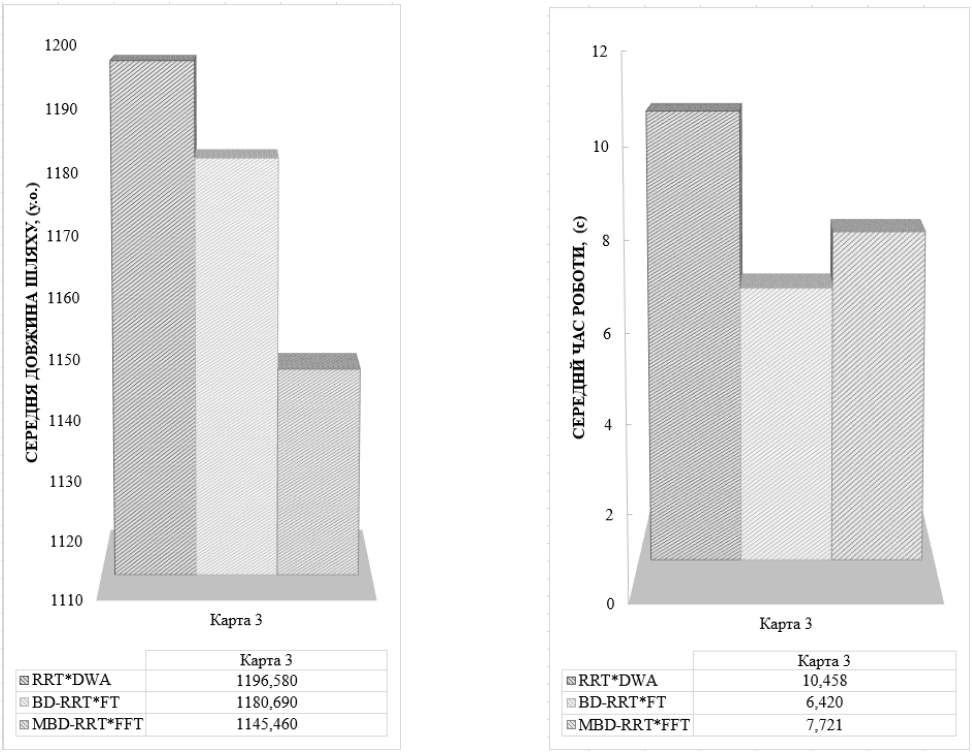
Таблиця 3

Порівняння показників ефективності алгоритму на карті 3

Алгоритм	RRT*DWA	BD-RRT*FT	MBD-RRT*FFT
Середня довжина шляху, *у. о.*	1196,58	1180,69	1145,46
Середній час роботи, *мс*	10458,4	6420,2	7720,9
Споживана пам’ять, *Мб*	470,66	387,04	433,14

На рисунку 4, *в* явно помітна наявна властивість всіх алгоритмів можливості планування шляху в реальному часі та з корекцією шляху задля уникнення динамічних (рухомих) об’єктів-перешкод.

Відомо [7; 8; 10; 16; 21], що алгоритм RRT*DWA має початкову спадкову властивість планування шляху від алгоритмів RRT і RRT*, що саме й використовують для глобального планування шляху на перших етапах, і тільки потім, під час руху БАНЗ, використовують алгоритм DWA.



а) *б)*

Рис. 8. Порівняльна діаграма продуктивності алгоритмів з динамічними об’єктами на карті 3:
а – зв’язок між кількістю ітерацій та довжиною шляху;
б – зв’язок між кількістю ітерацій та часом виконання

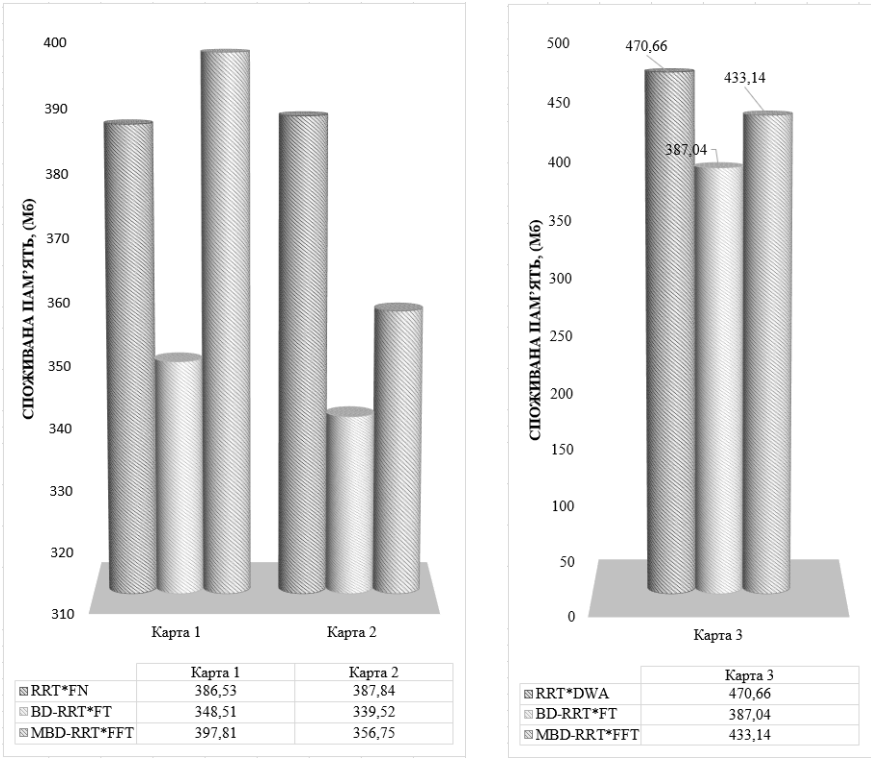
Локальне планування шляху використовується для досягнення динамічного уникнення перешкод, але через це легко впасти в локальну оптимальність, а середня довжина шляху та глобальний час планування шляху довші, ніж запропонований автором модифікований алгоритм BD-RRT*FT в роботі [1]. Вдосконалений алгоритм MBD-RRT*FFT, маючи загальну спадковість пошуку від алгоритму BD-RRT*FT, має загальну тенденцію до того, що запланований шлях може бути тимчасово поганим через те, що деякі вузли вибірки відкидаються й маркуються в *black_list*, після того, як шлях знищено перешкодами. Але завдяки використанню багатопотокового обчислення експеримент реально показав, що вдосконалений алгоритм MBD-RRT*FFT має кращі можливості динамічного планування,

а використання *Fitch’s*-алгоритму вибору оптимального результату пошуку надає загальній тенденції прагнення до *вибору більш прямолінійного оптимального шляху*.

Як і в дослідженні [1] під час проведення експериментальної частини, на всіх етапах перевірки проводився контроль навантаження на пам’ять системи прийняття рішення.

Середні результати споживання при роботі всіх алгоритмів показані в таблицях 1–3.

На рисунку 9. представлено порівняльні діаграми споживання об’єму пам’яті алгоритмами для всіх карт, що досліджувалися. Ми бачимо, що для вдосконаленого алгоритму присутнє *підвищене споживання пам’яті*, що свідчить про використання *багатопоточності* при виконанні обчислень, але враховуючи отримані результати щодо оптимальності та швидкості обчислень, ми розуміємо, що, порівнюючи з батьківським алгоритмом, приріст споживання в 12–15 % можна *вважати незначним*, але необхідно враховувати при виборі апаратної складової керуючого контролера та систем прийняття рішень.



	Карта 1	Карта 2
RRT*FN	386,53	387,84
BD-RRT*FT	348,51	339,52
MBD-RRT*FFT	397,81	356,75

	Карта 3
RRT*DWA	470,66
BD-RRT*FT	387,04
MBD-RRT*FFT	433,14

а) *б)*

Рис. 9. Порівняльна діаграма ефективності завантаження пам’яті системи прийняття рішень при виконанні пошуку шляху алгоритмами:
а – робота алгоритмів на картах 1 і 2; *б* – робота алгоритмів на картах 3

Висновки

У задачах навігації рухомих об’єктів одним із головних напрямів є вирішення проблеми планування шляху переміщення роботів і підвищення точності та надійності їх наведення у режимі реального часу. Розуміючи складність завдань під час проведення бойових дій в урбанізованому просторі щільної забудови міста з постійною зміною ландшафту, такі завдання, що покладаються на БАНЗ, мають тенденцію до постійного ускладнення.

Задля вирішення цієї важливої практичної проблеми з урахуванням проведеного дослідження [1] на основі отриманої аналітики було запропоновано і проведено вдосконалення модифікації асимптотично оптимального алгоритму BD-RRT*FT та розроблено алгоритм MBD-RRT*FFT, який при застосуванні в динамічних середовищах, не дивлячись на загальну спадковість алгоритму BD-RRT*FT, має загальну тенденцію до того,

що запланований шлях може бути тимчасово поганим, але завдяки використанню багатопотокового обчислення має кращі можливості динамічного планування, а використання *Fitch's*-алгоритму вибору оптимального результату пошуку надає загальну тенденцію до вибору оптимального прямолінійного шляху.

Експериментально доведено ефективність алгоритму MBD-RRT*FFT задля задоволення вимог планування шляху в реальному часі, дозволяючи БАНЗ швидко отримати оптимальний шлях без зіткнень у *складних динамічних* середовищах у режимі реального часу.

Важливо зауважити, що завдяки використанню в алгоритмі MBD-RRT*FFT багатопоточності з'являється підвищене споживання пам'яті в 12–15 %, але враховуючи отримані результати щодо оптимальності та швидкості обчислень, ми розуміємо, що, порівнюючи з батьківським алгоритмом, такій приріст можна вважати незначним, але необхідним для розуміння при виборі апаратної складової керуючого контролера та систем прийняття рішень

Напрями подальших досліджень

Враховуючи отримані результати дослідження, в подальшому планується розробити й описати *методику* застосування вдосконаленого модифікованого алгоритму MBD-RRT*FFT для застосування в реальних умовах середовища з динамічними перешкодами.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Бернацький А. П. Метод планування шляху автономного наземного робота з використанням модифікації динамічного двонаправленого RRT-алгоритму // Системи і технології зв'язку, інформатизації та кібербезпеки. 2023. № 4. С. 16–31. DOI: 10.58254/viti.4.2023.02.16.

2. Бернацький А. П. Основи робототехніки військового призначення / А. П. Бернацький, І. В. Панченко, О. І. Восколович. Київ: ВІТІ, 2021. 496 с.

3. Бернацький А. П., Панченко І. В., Восколович О. І. Розширена математична модель руху автономного наземного робота розвідника в умовах бойових дій в урбанізованому просторі // Озброєння та військова техніка. 2021. № 30 (2). С. 121–129.

4. Бідюк П. І. Системи і методи підтримки прийняття рішень / П. І. Бідюк, О. Л. Тимощук. Київ: КПІ ім. Ігоря Сікорського, 2022.

5. Про нарощування спроможностей сил оборони: Указ Президента України від 06.02.2024 № 51/2024. URL: https://www.president.gov.ua/documents/512024-49625.

6. Про рішення Ради національної безпеки і оборони України від 25 березня 2021 року «Про Стратегію воєнної безпеки України»: Указ Президента України від 21.03.2021 № 121.

7. Adiyatov, Olzhas; Varol, Huseyin Atakan. A novel RRT-based algorithm for motion planning in Dynamic environments. In Mechatronics and Automation (ICMA), 2017 IEEE International Conference on, 2017. P. 1416–1421. URL: https://doi:10.1109/ICMA.2017.8016024.

8. Black, Paul E. Greedy algorithm. Dictionary of Algorithms and Data Structures, US National Institute of Standards and Technology* PE Black. 2010. URL: http://www.nist. gov/dads/HTML/greedyalgo.html.

9. Cormen T., Leiserson C., Rivest R., Stein C. Introduction to Algorithms 3rd ed. Cambridge, London: The MIT Press, 2009. 1312 p.

10. Gammell J. D., Srinivasa S. S., Barfoot T. D. Informed RRT*: Optimal sampling-based path planning focused via direct sampling of an admissible ellipsoidal heuristic // Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems; Chicago, IL, USA. 14–18 September. 2014. URL: https://doi.org/10.1109/IROS.2014.6942976.

11. Jeong I. B., Lee S. J., Kim J. H. Quick-RRT*: Triangular inequality-based implementation of RRT* with improved initial solution and convergence rate. Expert Syst. Appl. – 2019; 123. P. 82–90. URL: https://doi:10.1016/j.eswa.2019.01.032.

12. Kadry S., Alferov G., Fedorov V. D-Star Algorithm Modification // International Journal of Online and Biomedical Engineering (iJOE). 2020. Vol. 16, No. 8. P. 108–113. URL: https://doi.org/10.3991/ijoe.v16i08.14243.

13. Kagan E., Ben-Gal I. A Group-Testing Algorithm with Online Informational Learning // IIE Transactions (Institute of Industrial Engineers). No. 46 (2). P. 164–184. URL: https://doi.org/10.1080/0740817X.2013.803639.

14. Klemm S., Oberländer J., Hermann A., Roennau A., Schamm T., Zollner J. M., Dillmann R. RT-Connect: Faster, asymptotically optimal motion planning // 2015 IEEE International Conference on Robotics and Biomimetics (ROBIO 2015) At: Zhuhai, China. Volume 12. URL: https://doi.org/10.1109/ ROBIO.2015.7419012.

15. Li B., Chen B. An Adaptive Rapidly-Exploring Random Tree // IEEE/CAA Journal of Automatica Sinica. 2021. URL: https://doi.org/10.1109/JAS.2021.1004252.

16. Majeed A, Hwang SO. Path planning method for UAVs based on constrained polygonal space and an extremely sparse waypoint graph // Applied Sciences. 2021. No. 11 (12). P. 5340. URL: https://doi.org/10.3390/app11125340.

17. Maw A. A., Tyan M., Nguyen T. A. et al. iADA*-RL: Anytime graph-based path planning with deep reinforcement learning for an autonomous UAV // Applied Sciences. 2021. No. 11 (9). P. 1–18. URL: https://doi.org/10.3390/app11093948.

18. Schapire R., Freund Y. Boosting: Foundations and Algorithms. MIT, 2012.

19. Spanogianopoulos, Sotirios and Sirlantzis, Konstantinos Non-holonomic Path Planning of Car-like Robot using RRT*FN // In: 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI). 2015 IEEE, pp. 53–57. URL: https://doi:10.1109/URAI.2015.7358927.

20. Taheri E., Ferdowsi M. H., Danesh M. Fuzzy greedy RRT path planning algorithm in a complex configuration space // Int. J. Control. Autom. Syst. 2018. No. 16. P. 3026–3035. URL: https://doi:10.1007/ s12555-018-0037-6.

21. Thomas H. Corman, Charles I. Leiserson, Ronald L. Rivest, Clifford Stein. Algorithms: construction and analysis. Introduction to Algorithms. Michigan U.: Williams, 2006.

22. Winnfield J. A. Unmanned Systems Integrated Roadmap FY2011-2036 / Winnfield J. A. Jr., Kendall F. Washington, DC: U.S. Department of Defense, March 9, 2012.

23. Wirth N. Algorithms + Data Structures = Programs (Prentice-Hall Series in Automatic Computation) / N. Wirth. Hoboken: Prentice Hall, 1976. 366 p.

24. Zhang L., Shen J., Yang J., Li G. Analyzing the fitch method for reconstructing ancestral states on ultrametric phylogenetic trees // Bulletin of Mathematical Biology. 2010. P. 1760–1782. DOI: 10.1007/s11538-010-9505-8.

DOI: 10.58254/viti.5.2024.04.52

УДК 621.391.372

Бондаренко Л. О. ORCID: 0000-0003-1850-0508 (ВІТІ ім. Героїв Крут)
канд. техн. наук Масесов М. О. ORCID: 0000-0003-4537-4295 (ВІТІ ім. Героїв Крут)
канд. техн. наук Коваленко І. Г. ORCID: 0000-0002-6827-5196 (ВІТІ ім. Героїв Крут)
Руденко В. І. ORCID: 0000-0003-3563-548X (ВІТІ ім. Героїв Крут)
Шугалій О. О. ORCID: 0000-0002-6587-0096 (ВІТІ ім. Героїв Крут)

УЗАГАЛЬНЕНА МЕТОДИКА ОЦІНКИ ЗАСТОСУВАННЯ СУПУТНИКОВИХ РЕТРАНСЛЯТОРІВ ЗВ'ЯЗКУ ЗА ЦІЛЬОВИМ ПРИЗНАЧЕННЯМ

Забезпечення електронними комунікаційними послугами військових формувань, які захищають державу в умовах відбиття повномасштабного воєнного вторгнення рф, вимагає створення системи зв'язку (СЗ) Збройних Сил України (ЗС України), яка володіє необхідними властивостями щодо виконання поставлених завдань за призначенням [1; 2]. Згідно зі стандартом AJP-3 «Об'єднана доктрина ведення операцій НАТО» [3], нарощування бойових спроможностей об'єднаних сил та підтримка операцій (бойових дій) здійснюється в тому числі завдяки використанню супутникового зв'язку.

Потенційні можливості супутникового зв'язку, такі як широкомовність, висока надійність та якість каналів зв'язку, незалежність від відстані, фізико-географічних умов та висока мобільність дозволяють забезпечити зв'язком абонентів системи електронних комунікацій збройних сил, яка є бойовою платформою, що дозволяє системі управління отримувати мережеві інформаційні послуги в реальному масштабі часу. Сучасні супутникові ретранслятори (СР) можуть мати до кількох десятків транспондерів кількох частотних діапазонів [4–6]. Супутникові системи зв'язку (ССЗ), що включають кілька СР, можуть забезпечити високу пропускну спроможність радіоканалів та можливість передачі ними великих обсягів інформації і дозволяють вирішити проблему створення глобальних інформаційно-телекомунікаційних систем та мереж.

Система супутникового зв'язку ЗС України повинна мати можливість зміни конфігурації для визначеного району обслуговування залежно від потреб у певних типах супутникових мереж, а отже ССЗ ЗС України має базуватися в тому числі на СР зв'язку комерційного призначення.

У статті запропонована узагальнена методика оцінки супутникових ретрансляторів зв'язку ССЗ ЗС України, новизною якої є її загальна спрямованість для ССЗ, що побудовані за різними архітектурними принципами, використовують різні частотно-орбітальні ресурси та надають різні послуги із застосуванням методологічного підходу, який базується на положеннях кваліметрії та розрахунку інтегрального та відносного показників ефективності, що враховує характерні показники окремих СР.

Ключові слова: *супутниковий зв'язок, супутниковий ретранслятор, критерій ефективність.*

L. Bondarenko, M. Masesov, I. Kovalenko, V. Rudenko, O. Shugaliy Generalized methodology of valuate using the satellite repeater connection for targeted purpose.

Providing electronic communication services to military formations that protect the state in the conditions of repelling a full-scale military invasion of the Russian Federation, requires the creation of a communication system (SZ) of the Armed Forces (AF) of Ukraine, which has the necessary properties to perform the assigned tasks as intended [1; 2]. According to the standard AJP-3 «Joint doctrine of the conduct of NATO operations» [3], increasing the combat capabilities of the combined forces and supporting operations (combat operations) is carried out, including through the use of satellite communications.

The potential capabilities of satellite communication, such as broadcasting, high reliability and quality of communication channels, independence from distance, physical and geographical conditions, and high mobility, allow to ensure the communication of subscribers of the electronic communications system of the armed forces, which is a combat platform that allows the system management to receive network information services in real time. Modern satellite repeaters (SR) can have up to several dozen transponders of several frequency bands [4–6]. Satellite communication systems, which include several SRs, can provide high bandwidth of radio channels and the possibility of transmitting large volumes of information through them and allow solving the problem of creating global information and telecommunication systems and networks.

The satellite communication system of the AF of Ukraine should be able to change the configuration for a specific service area depending on the needs of certain types of satellite networks, and therefore the satellite communication system of the AF of Ukraine should be based, including on SR communication of commercial appointment.

The article proposes a generalized methodology for evaluating satellite communication relays of the SSR of the AF of Ukraine, the novelty of which is its general orientation for satellite communication systems built according to different architectural principles, using different frequency-orbital resources and providing different services using a

methodological approach, which is based on the provisions of qualimetry and the calculation of integral and relative efficiency indicators, which takes into account the characteristic indicators of individual satellite repeaters.

Keywords: *satellite communication, satellite repeaters, efficiency criteria.*

Постановка завдання в загальному вигляді

Планування електронних комунікацій в системі планування операції (бою) є найбільш складним, трудомістким та відповідальним періодом діяльності посадових осіб органів управління зв’язком, спрямованим на вироблення обґрунтованих рішень щодо організації зв’язку та розроблення відповідних документів.

Супутниковий зв’язок – це радіозв’язок, що здійснюється між земними станціями за допомогою ретрансляції радіосигналів через СР [1; 2], основними перевагами якого, порівняно з іншими родами зв’язку, є:

значне розширення території, в межах якої можливий зв’язок між абонентами, та одночасне використання ними СР;

висока пропускна спроможність радіоканалів супутникового зв’язку і можливість передачі по них великих обсягів інформації;

можливість забезпечення зв’язком важкодоступних та віддалених районів за практичної незалежності вартості супутникового каналу від відстані між абонентами.

Система супутникового зв’язку ЗС України являє собою складну цілеспрямовану технічну систему, що складається з сукупності елементів (різної природи), які знаходяться у відносинах та зв’язках один з одним [6; 7]. Оскільки ССЗ призначена для надання послуг електронних комунікацій кінцевим споживачам із заданою якістю, то планування застосування системи має проводитися при комплексному обліку різних за своєю природою факторів, різноманіття відносин та зв’язків між ними, зовнішніх умов тощо, які впливають на показники якості послуг.

Оцінка ССЗ проводиться з метою вирішення наступних завдань:

прийняття рішення щодо допустимості практичного використання способу дій, що оцінюється, у тій чи іншій ситуації;

виявлення вкладів (ефектів) різних факторів у загальну ефективність ССЗ, впливу взаємодій факторів на ефективність;

визначення шляхів підвищення ефективності ССЗ (виявлення резервів ефективності);

визначення функціональних можливостей технічних засобів, що використовуються;

зіставлення (порівняння) кількох альтернативних варіантів дій або технічних засобів, їх ранжування за рівнями ефективності (встановлення відношення переваги на множині можливих варіантів).

Методологічною основою підготовки та обґрунтування рішень при вирішенні проблем технічного характеру є системний аналіз. Прийоми та методи системного аналізу спрямовані на висування альтернативних варіантів вирішення проблеми, виявлення масштабів невизначеності (невідповідності) щодо кожного варіанту та зіставлення варіантів щодо їхньої ефективності.

Планування побудови ССЗ, яка задовольняє заданим показникам якості, повинні ґрунтуватися на методах порівняння показників ефективності за обраними, обґрунтованими критеріями.

Актуальність викладеного матеріалу полягає в тому, що відсутність єдиних підходів, а відтак і єдиних методів та методик оцінки ССЗ, з урахуванням особливостей її побудови, ускладнює, а часом і перешкоджає проведенню оперативних розрахунків під час планування побудови ефективної архітектури ССЗ ЗС України, яка задовольняє за показниками якості системи управління, на користь якої вона функціонуватиме. Вирішення зазначеної проблеми полягає в подальшому вивченні глибинної суті проблеми, її опису та формуванні пропозицій щодо методів та узагальненої методики вирішення цього актуального завдання.

Аналіз останніх публікацій

Класичні роботи в області ССЗ, наприклад, [4–6; 8], присвячено побудові сучасних супутників зв’язку і земних станцій на основі досліджень взаємозалежності характеристик пропускної спроможності та завадостійкості ліній супутникового зв’язку, захисту ліній від зовнішніх і взаємних завад при багатостанційному доступі, розподілу енерго-просторових ресурсів у багатоствольних та багатопроменевих системах, орбітальної побудови багатосупутникових систем.

У зазначених роботах розгляд питань ефективності ССЗ проводиться з позицій техніко-економічної ефективності, вартісних функцій систем з урахуванням космічного сегмента, оптимізації вартості ліній супутникового зв’язку, питомих витрат за одиницю пропускної спроможності та питомої вартості обслуговування систем.

У монографії [6] пропонується та досліджується багаторівнева структурно-параметрична («статична») модель ССЗ як сукупності та взаємозалежності компонентів із множиною параметрів та показників, що підлягають оптимізації.

У статті [9] розглядається залежність стратегії розвитку супутникового зв’язку на основі нових технологій шляхом застосування багатопараметричних оцінок та інтегральних моделей ефективності. Сформовано систему приватних та узагальнюючих показників, розроблено модель інтегральної оцінки ефективності, процедуру експертного оцінювання складу, значущість окремих показників інтегральної моделі та безпосередню оцінку ефективності. Запропоновано оцінку достовірності вибору ефективних варіантів побудови ССЗ, що здійснюється на основі коефіцієнта узгодженості думок експертів.

У статті [10] пропонується методика оцінки ефективності ССЗ із СР на різних типах орбіт, яка зводить оцінку ефективності до розрахунку собівартості одиниці передачі інформації. Однак запропонований підхід не дозволяє порівняти ССЗ із СР на GEO-, LEO- та MEO-орбітах, параметри та послуги яких суттєво відрізняються.

У роботі [11] досліджуються питання оцінки та підвищення ефективності каналів зв’язку сучасних ССЗ, реалізованих на базі технології VSAT, з використанням технологій TDM/TDMA та комутацією пакетів з урахуванням особливостей Ка-діапазону. Запропоновано узагальнені методики та моделі для оцінки та вибору варіантів, що забезпечують підвищення техніко-економічної ефективності каналів у мережах зв’язку на базі сучасних технологій VSAT з урахуванням особливостей Ка-діапазону.

Таким чином, аналіз літератури та інтернет-ресурсів у галузі ССЗ [4–6; 9–11] показує, що нині відсутня загальна методика оцінки ефективності ССЗ, які побудовані за різними архітектурними принципами, використовують різні частотно-орбітальні ресурси та надають різні послуги.

Мета статті

Метою статті є формування методологічного апарату оцінки ССЗ та розробка узагальненої методики оцінки застосування супутникових ретрансляторів зв’язку за цільовим призначенням в ССЗ ЗС України. Така методологія може використовуватися під час розробки вихідних даних при моделюванні процесів у ССЗ ЗС України на користь науково обґрунтованого вибору принципів організації мереж зв’язку у них.

Новизною представленої в роботі методики є її загальна спрямованість для ССЗ, що побудовані за різними архітектурними принципами, використовують різні частотно-орбітальні ресурси та надають різні послуги із застосуванням методологічного підходу, який базується на положеннях кваліметрії та розрахунку інтегрального та відносного показників ефективності, що враховує характерні показники СР ССЗ ЗС України.

Виклад основного матеріалу

Ефективність ССЗ виступає як важлива категорія теорії та практики застосування системи за цільовим призначенням, що відображає результат прояву всієї сукупності факторів

та умов, які визначають хід і результат забезпечення зв’язку, а значить й організації управління [12]. Ефективність складних цілеспрямованих технічних систем, таких як ССЗ, визначається безліччю різних за своєю природою факторів. Під фактором розумітимемо рушійну силу будь-якого процесу (яви) чи умов, що впливає на той чи інший процес (явище) [7].

Ефективність функціонування ССЗ сприймається як міра узагальненої (інтегральної) властивості, характеризує рівень її відповідності своєму призначенню [7]. Ефективність відображає якість функціонування ССЗ у конкретних умовах обстановки та показує, наскільки успішно вирішуються (можуть вирішуватись) завдання надання встановленого набору послуг електронних комунікацій. Оцінка ефективності ССЗ може здійснюватися на різних етапах її синтезу (побудови) та функціонування.

Для оцінки цільової ефективності ССЗ визначимо алгоритм дій, який складається з:

1) визначення цілей та завдань оцінки ефективності;

2) вибору показників та визначення критеріїв ефективності;

3) вибору (побудови) моделей для розрахунку показників ефективності;

4) підготовки вихідних даних для розрахунку показників ефективності;

5) проведення розрахунків для оцінки ефективності, аналізу результатів розрахунків та прийняття рішення [7; 12].

При дослідженні цільової ефективності ССЗ виділимо три групи факторів, наведених на рисунку 1.

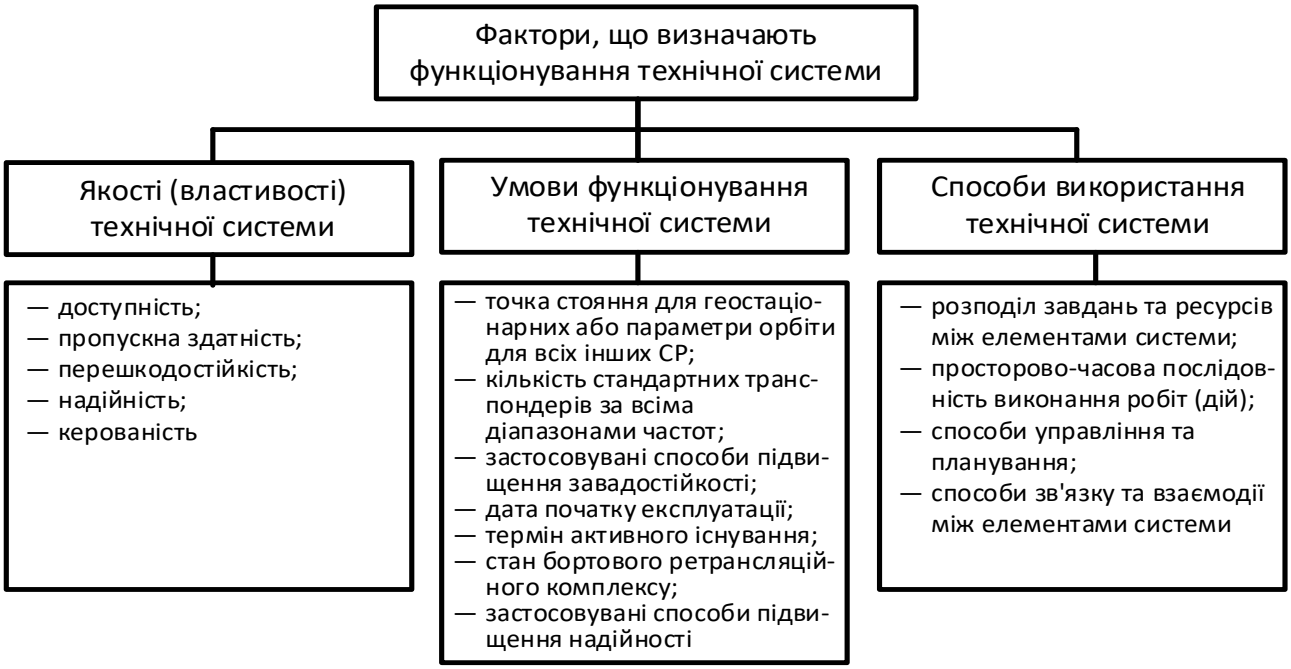


Рис. 1. Фактори функціонування ССЗ як технічної системи

Для оцінювання цільової ефективності ССЗ необхідно визначити склад початкових і вихідних даних, а також показників і критеріїв, за якими буде здійснюватися їхнє порівняння між собою.

Оцінку якості ССЗ з позицій виконання завдань за призначенням пропонується визначати за цільовими характеристиками СЗ, такими як [1; 2]:

– доступність;

– пропускна здатність;

– завадозахищеність;

– надійність;

– керованість.

Розглянемо найбільш значущі характеристики СР, визначимо показники та критерії їхньої ефективності, а також способи розрахунку.

Показник доступності СР та критерій оцінювання його ефективності

Доступність СР – здатність ретранслятора забезпечувати абонентам супутникової мережі зв’язку доступ до радіоресурсу за збереження призначених пріоритетів та способів встановлення зв’язку [6].

Одним із найбільш часто обчислюваних показників доступності СР є ймовірність санкціонованого доступу абонента незалежно від його розташування щодо СР. Очевидно, що ймовірність санкціонованого доступу залежить від взаємного розташування абонента супутникової мережі зв’язку та зони радіопокриття СР, що фактично визначається ймовірністю знаходження абонента в зоні обслуговування ретранслятора та залежить від виконання умови взаємної радіовидимості супутникового абонентського терміналу та СР. Взаємозв’язок зон покриття, обслуговування та радіовидимості СР показано на рисунку 2.

За наявності рухомих абонентів розрахунок показника доступності СР проводиться з урахуванням перетину зони радіопокриття та району розташування абонентів, а також часу знаходження терміналу абонентів у зоні взаємної радіовидимості [6].

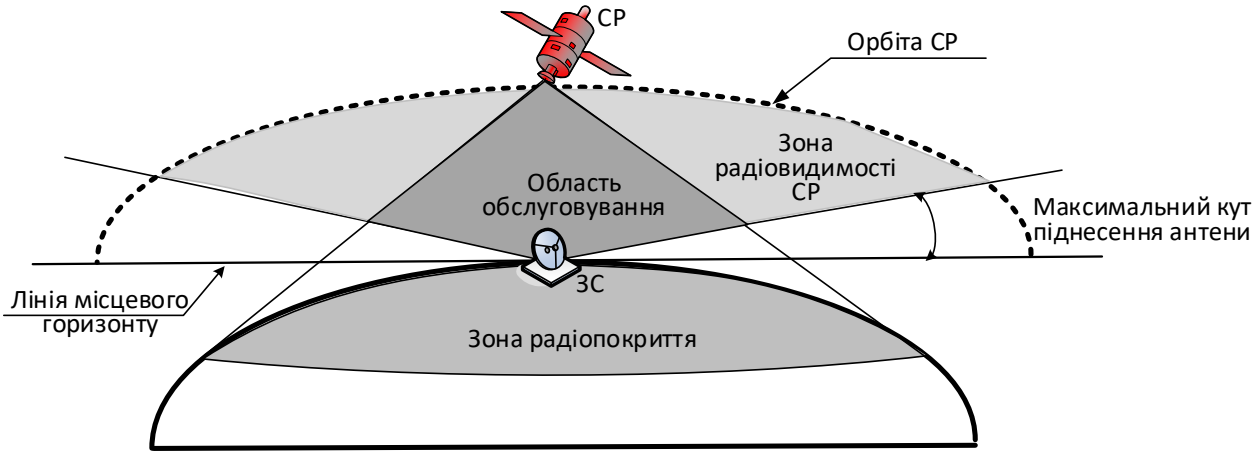


Рис. 2. Взаємозв’язок зон покриття, обслуговування та радіовидимості СР

Прийнявши припущення про те, що всі абоненти рівномірно розподілені по зоні їх розташування, розрахунок показника доступності СР ($P_Д$) виконується за формулою (1) [6]:

$$P_Д = \frac{S_П T_a}{S_a T_o}, \quad (1)$$

де $P_Д$ – показник доступності СР;

$S_П$ – площа перетину району розташування абонентів із зоною радіопокриття СР;

S_a – площа району розташування абонентів:

T_a – період часу циклу обігу СР навколо Землі, коли абонент знаходиться в зоні радіопокриття СР;

T_o – період часу обігу СР навколо Землі.

Очевидно, що окремим випадком при розміщенні СР на геостаціонарній орбіті відношення дорівнюватиме 1, тому що абоненти завжди знаходяться в одному і тому самому положенні щодо апарата. У разі, коли районів розташування абонентів декілька, розрахунок ведеться по перетині всіх районів, що розглядаються.

Показник пропускної спроможності СР та критерій оцінювання ефективності

Пропускна здатність СР – здатність ретранслятора, що визначає сукупний обсяг інформації, яка може передаватись і оброблятись в утворених ним радіоканалах, за одиницю часу і може бути кількісно оцінена матрицею пропускних здатностей напрямів зв'язку. У цьому випадку її прийнято розраховувати через сумарну кількість еквівалентних (стандартних) транспондерів у заданих діапазонах частот [6; 8]. В якості величини еквівалентного транспондера СР прийнята смуга частот, що дорівнює 36 МГц. У цьому випадку нормований показник пропускної спроможності СР ($P_{\text{ПС}}$) обчислюється за формулою (2) [6; 8]:

$$P_{\text{ПС}} = \frac{\sum_{i-1}^{d} N_T C_T}{C_{\text{ET}} N_{max}}, \tag{2}$$

де d – кількість діапазонів частот, в яких може функціонувати СР;

N_T – кількість транспондерів СР у заданому діапазоні частот;

C_T – смуга частот транспондерів СР у заданому діапазоні;

C_{ET} – смуга частот еквівалентного транспондера СР;

N_{max} – максимальна кількість транспондерів СР в множині СР, що розглядається (або в орбітальному угрупованні).

Показник завадостійкості СР та критерії оцінювання ефективності

Завадостійкість СР – здатність ретранслятора виконувати завдання за призначенням в умовах впливу всіх видів завад [1; 2].

Виконати спрощене оцінювання завадостійкості СР можливо через порівняння сучасних стандартів супутникового зв'язку, що реалізуються в СР. Завадостійкість СЗ знаходиться в жорсткому співвідношенні зі швидкістю передачі інформації і безпосередньо пов'язана з питаннями оптимального використання спектрального ресурсу супутникового каналу. Сучасні стандарти супутникового зв'язку передбачають можливості адаптивної зміни балансу між завадостійкістю та швидкістю передачі інформації завдяки вибору сигнально-кодових конструкцій. У сучасних СР найчастіше застосовуються стандарти DVB-RCS (Digital Video Broadcasting – Return Channel via Satellite), DVB-S/S2 (Digital Video Broadcasting – Sputnik), а також режим псевдовипадкової перебудови робочої частоти (ППРЧ) та використання фазоманіпульованих шумоподібних сигналів (ФМ ШПС), порівняння яких виконано в роботі [13]. При проведенні розрахунків повинні враховуватися можливості застосування інших способів забезпечення стійкості до завад: обробка сигналу на борту, використання багатопроменевих антен, управління променями антени, формування «нулів» діаграми спрямованості антени. Обчислення показника стійкості до завад СР проводиться за формулою, що дозволяє визначити співвідношення наявності та ефекту технологій забезпечення завадостійкості СР [6] за формулою (3):

$$P_{\text{ЗС}} = \sum_{i=1}^{m} K_i, \tag{3}$$

де m – кількість технологій, що впливають на завадостійкість СР;

K_i – показник ефекту технологій забезпечення завадостійкості СР, який визначається експертно за таблицею 1 [13].

Таблиця 1

Технології забезпечення завадостійкості СР

№ з/п	Технології забезпечення завадостійкості	K_i
1	Стандарт DVB-S	0,07
	Стандарти DVB-S2/ DVB-RCS	0,09
2	Використання ФМ ШПС	0,18
3	Режим ППРЧ	0,22

№ з/п	Технології забезпечення завадостійкості	K_i
4	Багатопроменеві антени	0,10
5	Режим обробки сигналів на борту	0,14
6	Управління променями антени	0,11
7	Формування «нулів» діаграми спрямованості антени	0,16

Показник надійності СР та критерії оцінювання ефективності

Надійність – здатність системи військового зв'язку та автоматизації виконувати завдання за призначенням, зберігаючи в часі значення експлуатаційних показників у межах, передбачених експлуатаційною документацією [1; 2].

СР належать до складних технічних систем, оскільки вони мають слабо передбачувану поведінку зі здатністю ухвалення рішення [7]. З метою досягнення заданого рівня надійності СР застосовують конструкторські рішення, що забезпечують відновлюваність виробу, наприклад: структурне, функціональне, алгоритмічне, часове та інформаційне резервування.

На етапі конструювання СР використовуються комплексні узагальнені показники надійності, які характеризують систему на всіх етапах її життєвого циклу.

Для оцінювання СР за показником надійності доцільно використовувати узагальнені комплексні показники [14]: коефіцієнт готовності ($K_{\text{г}}$), коефіцієнт оперативної готовності ($K_{\text{ог}}$), коефіцієнт технічного використання ($K_{\text{тв}}$), виходячи з яких розраховується підсумковий показник надійності. Розрахунок узагальненого показника надійності пропонується проводити за такою формулою (4) [14]:

$$P_{\text{н}} = K_{\text{г}} * K_{\text{ог}} * K_{\text{тв}}, \tag{4}$$

Значення комплексних показників конкретного типу СР задаються у тактико-технічному завданні на проєктування та містяться у технічних умовах на виріб. Зазвичай значення зазначених характеристик варіюється в межах 0,9÷0,995.

Показник керованості СР та критерії оцінювання ефективності

Під керованістю розуміють здатність системи переходити за кінцевий (заданий) час з одного стану в інший (необхідний) під дією керуючих впливів [7].

У загальному сенсі керованість є здатністю системи виконувати команди управління та оперативно реагувати на них. Очевидно, керувати можна лише завадостійкою системою. Якщо сигнали керування спотворюються в недопустимих межах при їх передачі до об'єктів управління, то управління позбавляється сенсу. Керованість забезпечується насамперед наявністю прямого та зворотного зв'язків, які служать для передачі до керованої системи команд (сигналів) управління, отримання від неї повідомлень про неузгодженість реального (фактичного) і необхідного станів керованого об'єкта та виконання команд управління. Керованість поєднує такі властивості системи, як гнучкість управління, його оперативність, точність, швидкодію, інерційність та ін. Для складних систем керованість включає також здатність вироблення рішень, на основі яких формуються керуючі дії [7].

На взаємодію бортової системи управління СР із наземним автоматизованим комплексом управління та цільовою бортовою апаратурою зв'язку чималий вплив має тип орбіти. Збільшення відстані від наземного автоматизованого комплексу управління до СР тягне за собою погіршення якості радіосигналу в прямому та зворотному напрямку (на лінії «вгору» і на лінії «вниз»), що веде до затримок в управлінні, зменшення швидкості передачі інформації, отже, і до зниження обсягів телеметричної інформації від цільової бортової апаратури [6].

Оперативне, наближене оцінювання керованості ($P_{\text{к}}$) СР визначається типом орбіти за умови відсутності впливу зовнішніх дестабілізуючих факторів штучного походження. Складність розрахунку цього показника в режимі масштабу часу, наближеного до реального,

визначає вибір на користь експертного способу оцінювання [9]. У таблиці 2, створеній виходячи з аналізу публікацій [4; 6; 9], представлено показники керованості СР залежно від типу орбіти [10].

Таблиця 2

Взаємозв'язок показника керованості СР та типу його орбіти

№ з/п	Тип орбіти СР	$P_{\text{у}}$
1	Низька кругова	1,00
2	Середня кругова	0,95
3	Геостаціонарна	0,90
4	Високоеліптична	0,85

Розрахунок показника цільової ефективності СР

Перелічені вище показники ефективності застосування СР дозволяють всебічно дослідити ефективність використання СР, а саме з погляду доступності – $P_{\text{Д}}$, пропускної спроможності – $P_{\text{ПС}}$, завадостійкості – $P_{\text{ЗС}}$, надійності – $P_{\text{Н}}$ та керованості – $P_{\text{К}}$. Тож постає актуальною задача визначення узагальненого показника цільової ефективності СР.

У спрощеному вигляді для узагальнення отриманих характеристичних показників СР можна використовувати розрахунок узагальнюючого інтегрального показника цільової ефективності СР за формулою (5):

$$P_i = (P_{\text{Д}} * P_{\text{ПС}} * P_{\text{ЗС}} * P_{\text{Н}} * P_{\text{К}})_i, \tag{5}$$

де i – умовний порядковий номер СР у множини оцінюваних або у складі аналізованого орбітального угруповання (ОУ). У загальному випадку оцінюваний СР може і не належати до складу ОУ.

Для врахування ваги окремих показників ефективності необхідно застосувати відповідні вагові коефіцієнти згідно з виразом (6):

$$P_i = (P_{\text{Д}} * q_{\text{Д}} + P_{\text{ПС}} * q_{\text{ПС}} + P_{\text{ЗС}} * q_{\text{ЗС}} + P_{\text{Н}} * q_{\text{Н}} + P_{\text{К}} * q_{\text{К}})_i, \tag{6}$$
$$q_{\text{Д}} + q_{\text{ПС}} + q_{\text{ЗС}} + q_{\text{Н}} + q_{\text{К}} = 1,$$

де $q_{\text{Д}}, q_{\text{ПС}}, q_{\text{ЗС}}, q_{\text{Н}}, q_{\text{К}}$ – вагові коефіцієнти відповідних показників ефективності.

Зазвичай, для визначення вагових коефіцієнтів найчастіше використовують методи експертних оцінок (експертний метод), метод використання показника варіації (варіаційний метод) та метод визначення закону розподілу [15]. Останні два можуть застосовуватися при наявності певної кількісної, статистичної інформації. У випадку, коли статистична інформація про досліджувані показники відсутня або її не можна кількісно виразити, використовується метод експертних оцінок.

Для визначення вагових коефіцієнтів пропонується використовувати метод експертних оцінок, який полягає в обробці інформації, отриманої шляхом опитування експертів. Використання експертів як джерел інформації про майбутній розвиток досліджуваного процесу (явища, об'єкта) ґрунтується на гіпотезі наявності у провідних спеціалістів визначеної галузі глибоких і достатніх знань про шляхи розв'язання досліджуваних проблем [16].

В основі більшості застосовуваних на практиці методів лежить опитування експертів з математичною обробкою їхніх суджень. Найчастіше використовують такі математичні методи [17]:

Пряма розстановка. Експертам пропонується розставити коефіцієнти k_i при відповідних факторах, виходячи з умови $\sum_{i=1}^{n} k_i = 1$, тобто розв'язати задачу безпосередньо.

Ранжування факторів. Ранжування дозволяє упорядкувати фактори за ступенем збільшення або зменшення їхнього впливу на досліджуваний феномен, що цікавить дослідника. Результати ранжування n факторів m експертів можна представити у вигляді матриці:

$$X_{11}, X_{21}, \dots, X_{n1}$$
$$X_{12}, X_{22}, \dots, X_{n2}$$
$$\dots, \dots, \dots, \dots$$
$$X_{1m}, X_{2m}, \dots, X_{nm.}$$

Зведені оцінки вагових коефіцієнтів можна отримати в результаті усереднення приватних рангів.

Парне порівняння. В цьому методі експертам пропонується порівняти фактори попарно. Інформація від кожного експерта надходить у формі булевої матриці парних порівнянь:

$$y_i = (y_{ik,j}),$$

де i, k = 1, …, n – фактори (показники), що порівнюються; j = 1, …, m – експерт, що виконує порівняння; $y_{ik,j}$ – результат парного порівняння j-м експертом факторів X_i і X_k може виражатися одиницею або нулем із наступного правила.

Якщо експерт порівнює фактори між собою, то 1, якщо на думку j-го експерта параметр X_i впливає сильніше ніж X_k, в іншому випадку 0.

Проаналізувавши різні методи експертних оцінок, для визначення вагових коефіцієнтів узагальненого показника цільової ефективності СР доцільно використовувати метод аналітичних записок, який дозволяє мінімізувати кількість експертів, а за математичний апарат – методи прямої розстановки і ранжування, які дозволяють врахувати помилки в оцінках експертів на основі застосування довірчого інтервалу.

Безпосередньо процес експертизи можна розбити на декілька етапів, згідно з методами, розглянутими в [17]:

1. Визначення придатності і компетентності експертів.
2. Проведення опитування.
3. Обробка результатів опитування.
4. Визначення узгодженості думки експертів.

На першому етапі кожному експерту виставляють за певним набором критеріїв деяку оцінку за бальною шкалою. Такими критеріями можуть бути, наприклад, фах за освітою з точки зору галузі дослідження, рівень освіти, загальний стаж роботи в галузі, стаж роботи в досліджуваному питанні, досвід участі в подібних експертизах, рівень ознайомленості з методиками дослідження тощо.

Для визначення компетентності на основі складається анкета, згідно з якою експерти особисто собі виставляють оцінки. Ступінь придатності спеціаліста до експертизи по анкетному опитуванню визначається за коефіцієнтом компетентності за виразом (7):

$$K_a = \frac{\sum V_{ij}}{\sum V_j}, \tag{7}$$

де K_a – коефіцієнт компетентності за анкетним опитуванням;

V_{ij} – вага j-ї градації (обраної експертом, що оцінюється) i-й характеристиці, в балах;

V_j – максимальна вага j-ї характеристики, в балах [16].

Сформована робоча група експертів є компетентною і здатною кваліфіковано вирішувати поставлені перед нею завдання за такої умови:

$$0{,}67 \le M \le 1,$$

де M – рівень компетентності експертної групи, що визначається формулою (8):

$$M = \frac{1}{n}\sum_{i=1}^{n} K_i\,, \tag{8}$$

де n – кількість експертів в групі.

Для визначення вагових коефіцієнтів методом ранжування [16] необхідно присвоїти кожному рангу певну кількість балів за умови, що перший ранг отримує найбільшу кількість балів, а останній – найменшу. Далі необхідно визначити суму балів по кожному критерію:

$$r_i = \sum_{m=1}^{n} r_{mi},$$

де r_i – сума балів по i-му критерію;

r_{mi} – кількість балів по i-му критерію від n-го експерта.

Безпосередньо розрахунок коефіцієнтів виконується за формулою (9):

$$k_i = \frac{r_i}{\sum_{i=1}^{m} r_i}. \quad (9)$$

Окремо проводиться оцінку ступеня узгодженості експертів для перевірки отриманих результатів.

Таким чином, за допомогою отриманих вагових коефіцієнтів та формули (5) можна отримати значення для узагальненого показника цільової ефективності для окремого СР.

Для зручності оцінки доцільно використовувати відносний показник цільової ефективності, який набуває значень у діапазоні $W_i \in [0;1]$ і може бути розрахований на основі узагальненого показника за формулою (10):

$$W_i = \frac{P_i}{\sum_{i=1}^{n} P_i}, \quad (10)$$

де W_i – відносний показник цільової ефективності i-го СР з множини оцінюваних;

i – умовний порядковий номер СР у багатьох оцінюваних $i = \overline{1,n}$;

P_i – значення інтегрованого показника цільової ефективності i-го СР, $P_i \in [0;1]$.

На практиці також зручно використовувати нормований показник цільової ефективності СР $P_{i\,\text{норм}}$, який пропонується розраховувати за формулою (11):

$$P_{i\,\text{норм}} = {P_i}/{P_{max}} \quad (11)$$

де $P_{i\,\text{норм}}$ – нормоване значення інтегрованого показника цільової ефективності i-го СР, $P_i \in [0;1]$;

P_{max} – максимальне значення інтегрованого показника цільової ефективності СР, що входить до складу ОУ, відносно якого буде здійснено порівняння інших ретрансляторів.

Нормований показник ефективності відображає, наскільки застосування оцінюваного СР відповідає цільовому призначенню щодо еталонного (найкращого) апарату, що входить до складу аналізованого орбітального угруповання.

Порівняння отриманих нормованих показників ефективності окремих СР дасть змогу проводити оцінку, вибір та побудову складових ССЗ, в тому числі перспективних супутників зв'язку України. Це дозволить побудувати найбільш ефективну ССЗ ЗС України.

Висновки

Супутникова система зв'язку ЗС України, що може включати кілька СР, в тому числі і комерційного призначення, може забезпечити високу пропускну спроможність радіоканалів та можливість передачі ними великих обсягів інформації і дозволяє вирішити проблему створення глобальних інформаційно-телекомунікаційних систем та мереж. При цьому постає задача розробки єдиних методів та методик оцінки ССЗ ЗС України, з урахуванням особливостей її побудови. Це дозволить проведення оперативних розрахунків під час планування побудови ефективної архітектури ССЗ ЗС України, яка повинна задовольняти за показниками якості системи управління, на користь якої вона функціонуватиме.

У статті вирішується актуальне наукове завдання – формування методологічного апарату оцінки ССЗ та розробка узагальненої методики оцінки застосування супутникових ретрансляторів зв'язку за цільовим призначенням в ССЗ ЗС України. Зазначена методологія може використовуватися під час розробки вихідних даних при моделюванні процесів у ССЗ

ЗС України на користь науково обґрунтованого вибору принципів організації мереж зв’язку у них.

Запропонована узагальнена методика оцінки супутникових ретрансляторів зв’язку ССЗ ЗС України, новизною якої є її загальна спрямованість для ССЗ, що побудовані за різними архітектурними принципами, використовують різні частотно-орбітальні ресурси та надають різні послуги з застосуванням методологічного підходу, який базується на положеннях кваліметрії та розрахунку інтегрального та відносного показників ефективності, що враховує характерні показники окремих СР. В якості основних характеристик, що підлягають оцінюванню, запропоновано вибрати доступність, пропускну здатність, завадостійкість, надійність і керованість супутникових ретрансляторів, на основі яких розраховується узагальнена інтегральна та нормована оцінка.

Представлена в роботі узагальнена методика буде корисною фахівцям для обґрунтування нових технологічних рішень при розробці вітчизняних ССЗ в інтересах ЗС України. Крім того, методологія буде корисна науковцям та здобувачам, які ведуть наукові дослідження в галузі супутникового зв’язку та під час розробки вимог до перспективних супутників зв’язку України.

Напрямки подальших досліджень

У подальших дослідженнях пропонується деталізувати розрахунки для окремих показників ефективності застосування ССЗ, розробити відповідні методики та провести розрахунки оцінки ефективності застосування для ССЗ, які можливо бути використовувати у Збройних Силах та інших силах оборони України.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. ДСТУ В3265 – 95. Зв’язок військовий. Терміни та визначення. [Чинний від 1997-01-01]. Державний стандарт України. 23 с.

2. Військова система стандартизації. Військовий стандарт 01.112.001. Військовий зв’язок. Терміни та визначення. К.: Міністерство оборони України, 2006. 25 с. Реєстраційний номер А2187/000020.

3. NATO STANDARD AJP-3ALLIED JOINT DOCTRINE FOR THE CONDUCT OF OPERATIONS. Edition C Version 1, 2019 // NATO MILMED COE. URL: https://www.coemed.org/files/stanags/01_AJP/AJP-3_EDC_V1_E_2490.pdf.

4. Ільченко М. Ю., Наритник Т. М., Присяжний В. І., Капштик С. В., Матвієнко С. А. Низькоорбітальна супутникова система інтернету речей на базі розподіленого супутника // Космічна наука і технологія. 2020. 26, № 4 (125). С. 57–85. URL: https://doi.org/10.15407/knit2020.04.057.

5. Ніколаєнко Б. А., Пелешок Є. В. Сучасні супутникові системи зв’язку: навч. посібник. К.: ІСЗЗІ КПІ ім. Ігоря Сікорського, 2022. 146 с.

6. Конин В. В. Спутниковые системы связи, навигации, наблюдения. К.: Кафедра АНС, 2007. 350 с.

7. Василевський О., Ігнатенко О. Нормування показників надійності технічних засобів: навч. посіб. Вінниця: ВНТУ, 2013. 160 с.

8. Беркман Л. Н., Стеклов В. К., Кільчицький Є. В. Оптимізація та моделювання пристроїв і систем зв’язку. Київ: Техніка, 2004. 576 с.

9. Фролов В. Ф., Кирпач Л. А., Ільїн О. Ю. Використання узагальнених критеріїв при проектуванні супутникових систем // Зв’язок. 2018. № 1. С. 9–11.

10. Кирпач Л. А., Власенко Г. М., Срібна І. М. Методи оцінювання критеріїв ефективності супутникових систем // Зв’язок. 2018. № 4. URL: file:///C:/Users/User/Downloads/2114-%D0%A2%D0%B5%D0%BA%D1%81%D1%82%20%D1%81%D1%82%D0%B0%D1%82%D1%82%D1%96-7478-1-10-20190513.pdf.

11. Галаган В. І., Турейчук А. М., Бондарчук С. В., Прокопенко О. С., Панадій К. В. Розгляд проблемних питань оптимізації супутникового зв’язку // Збірник наукових праць Центру воєнно-

стратегічних досліджень Національного університету оборони України імені Івана Черняховського. 2015. № 2. С. 30–37. URL: http://nbuv.gov.ua/UJRN/ Znpcvsd_2015_2_7.

12. Zinchenko A. Методичний підхід з оцінки ефективності системи зв’язку спеціального призначення / A. Zinchenko, R. Pikul, K. Zinchenko, K. Liubarchuk, E. Yusukhno, A. Maksymenko, Yu. Sokol, Yu. Rindin // Системи управління, навігації та зв’язку. Полтава: ПНТУ, 2020. Т. 1 (59). С. 132–136. DOI: https://doi.org/10.26906/SUNZ.2020.1.132.

13. Шульга О., Сокіріна В. Вимірювання завадостійкості супутникових радіонавігаційних систем методом теорії ігор в умовах невизначних завад // Measuring And Computing Devices In Technological Processes. № 4. С. 142–148. URL: https://doi.org/10.31891/2219-9365-2022-72-4-20.

14. ДСТУ 2860-94. Надійність техніки. Терміни та визначення. [Чинний від 1996-01-01]. К., 1995. 75 с.

15. Гуляр Р. Е. Методи визначення вагових коефіцієнтів при розрахунку таксономічних показників / Р. Е. Гуляр // Харківська національна академія міського господарства. 3 с. URL: http://eprints.kname.edu.ua/29737/1/44.pdf.

16. Грабовецький Б. Є. Методи експертних оцінок: теорія, методологія, напрямки використання: монографія / Б. Є. Грабовецький. Вінниця: ВНТУ, 2010. 171 с.

17. Сахно В. П. До аналізу методів визначення періодичності виконання технічних впливів / В. П. Сахно, О. М. Іванушко // Вісник Національного транспортного університету. Серія «Технічні науки»: науково-технічний збірник. К.: НТУ, 2017. Вип. 3 (39). С. 53–65.

DOI: 10.58254/viti.5.2024.05.64

УДК : 355.004

Гроздов А. А. ORCID: 0009-0001-9892-6738 (КСП ЗСУ)
Зінченко І. А. ORCID: 0000-0002-6516-9558 (ВІТІ ім. Героїв Крут)
Громлюк М. М. ORCID: 0009-0005-9490-8435 (в/ч А0515);
Білий О. А. ORCID: 0000-0003-3056-8652 (ВІТІ ім. Героїв Крут)
Івченко М. М. ORCID: 0000-0002-0039-2812 (ВІТІ ім. Героїв Крут)
Цимбал І. В. ORCID: 0000-0001-7294-3794 (ВІТІ ім. Героїв Крут)

МЕТОДИКА ОЦІНЮВАННЯ СТІЙКОСТІ ФУНКЦІОНУВАННЯ СИСТЕМИ ЗВ’ЯЗКУ НА ОСНОВІ БОЙОВИХ СПРОМОЖНОСТЕЙ ВІЙСЬК

Проведений авторами за досвідом ведення бойових дій на сході країни аналіз обчислення значення показника стійкості системи військового зв’язку засвідчив, що використання класичного підходу посадовими особами органу управління зв’язком не можливо. Він потребує проведення численних спеціалізованих розрахунків окремо за напрямками зв’язку з подальшим їхнім узагальненням за систему військового зв’язку у цілому в умовах невизначеності ведення активних бойових дій. Крім того, стійкість є інтегральним показником від згортки імовірнісних часткових показників живучості, надійності, завадозахищеності та кіберзахищеності. Тому авторами було запропоновано розробити нову методику обчислення значення стійкості функціонування системи військового зв’язку у сучасних умовах ведення бойових дій. На відміну від класичного підходу, вона повинна бути заснована на використанні інших показників, які дозволять спростити проведення розрахунків. Авторами була розроблена графічна модель функціонування системи військового зв’язку у сучасних умовах ведення активних бойових дій.

Для опису функціонування запропонованої моделі авторами були використанні значення бойових спроможностей військ (сил). Методика враховує прогнозований вплив засобів вогневого ураження, радіоелектронної боротьби противника на нашу систему військового зв’язку та виконання підрозділами зв’язку заходів бойового забезпечення і радіоелектронного захисту. Для формалізації методики авторами було одержано ряд компактних аналітичних виразів, які встановлюють імовірності співвідношення між значеннями бойових спроможностей наших військ і військ противника. Методика легко реалізується у вигляді інформаційно-розрахункової задачі і значно зменшує час проведення оперативних розрахунків для прийняття органом управління військовим зв’язком обґрунтованих управлінських рішень.

Ключові слова: *система військового зв’язку, стійкість, вогневе ураження, радіоелектронна боротьба, угруповання військ.*

A. Hrozdov, I. Zinchenko, M. Hromliuk, O. Bilyi, M. Ivchenko, I. Tsymbal The method of assessing the sustainability of the functioning system based on the combat capabilities of the armies

The authors' analysis of the calculation of the value of the stability indicator of the military communication system, based on the experience of conducting military operations in the east of the country, proved by the use of the classic approach by the officials of the communication management staff is not possible. It requires carrying out numerous specialized calculations separately for the directions of communication with their further generalization for the military communication system as a whole in the conditions of uncertainty of conducting active hostilities. In addition, resilience is an integral indicator from the convolution of probabilistic partial indicators of survivability, reliability, immunity and cyber security.

Therefore, the authors proposed to develop a new methodology for calculating the value of the stability of the operation of the military communication system in the modern conditions of conducting hostilities. Unlike the classical approach, it should be based on the use of other indicators that will simplify calculations. The authors developed a graphic model of the operation of the military communication system in modern conditions of active hostilities. To describe the operation of the proposed model, the authors used the values of the combat capabilities of troops (forces). The methodology takes into account the predicted impact of means of fire damage, radio-electronic warfare of the enemy on our military communication system and the implementation of combat support and radio-electronic protection measures by the communication units. To formalize the methodology, the authors obtained a number of compact analytical expressions that establish the probabilities of the ratio between the values of the combat capabilities of our troops and the enemy's troops. The methodology is easily implemented in the form of an information and calculation task and significantly reduces the time of operational calculations for the military communications management staff to make sound management decisions.

Keywords: *communication system, stability, fire damage, electronic warfare, grouping of troops.*

Постановка задачі.

Аналіз досвіду ведення збройної боротьби з 2014 року свідчить, що противник зазвичай зосереджує свої зусилля на дезорганізації управління нашими військами (силами) засобами вогневого ураження (ВУ) та радіоелектронного придушення (РЕП) шляхом цілеспрямованого впливу на системи управління. Такі дії противника підтверджують одну з основних тенденцій ведення збройної боротьби у сучасних умовах – цілеспрямовану боротьбу із системами управління. Особливо з технічною основою системи управління – системою зв'язку (СЗ). Тому важливого значення набуває завчасне прогнозування стійкості СЗ угруповань військ (УВ), які ведуть бойові дій (операції).

Під стійкістю розуміється спроможність СЗ виконувати завдання в умовах впливу різноманітних факторів. Стійкість СЗ – інтегральний показник від чотирьох часткових показників: живучість, надійність, завадозахищеність та кіберзахищеність [1–3]. У свою чергу, часткові показники також є складними властивостями. Існуючі методики обчислення стійкості передбачають визначення стійкості окремих напрямків зв'язку з подальшим їх узагальненням за СЗ у цілому. Кількісним показником оцінювання стійкості є ймовірність того, що напрямок (канал, лінія) зв'язку буде працездатним в будь-який випадково обраний момент часу. Тоді значення стійкості СЗ відповідно [1–3] буде обчислюватися за наступним виразом (1):

$$P_{\text{с}} = P_{\text{ж}} \times P_{\text{н}} \times P_{\text{зз}} \times P_{\text{кз}}, \quad (1)$$

де $P_{\text{ж}}$ – показник живучості;

$P_{\text{н}}$ – показник надійності;

$P_{\text{зз}}$ – показник завадозахищеності;

$P_{\text{кз}}$ – показник кіберзахищеності.

Наведений науковий підхід є занадто складним. Він потребує проведення численних спеціалізованих розрахунків окремо за напрямками зв'язку з подальшим їх узагальненням за СЗ у цілому в умовах невизначеності ведення активних бойових дій. Тому існує реальна потреба у розробці методики, яка враховувала б потреби військ і ґрунтувалася на інших часткових показниках. Такий варіант можливо реалізувати на основі використання значень бойових спроможностей військ.

Аналіз останніх досліджень і публікацій.

Оцінюванню функціональних показників системи військового зв'язку завжди приділяється значна увага в дослідженнях. Найбільш розповсюджені методики, що знаходять застосування і досі, були відпрацьовані задовго до початку формування нових концепцій та підходів до управління військами, але варто зазначити, що лише частково відповідають умовам сьогодення [1; 7].

Вивчення та аналіз більш сучасної літератури показав, що в роботах [2; 6; 8; 13] визначено показники та критерії оцінки за основним показником – стійкістю, що більш відповідають вимогам сьогодення та враховують набутий досвід під час антитерористичної операції.

Роботи [11; 12] розглядають сучасні СЗ сил оборони України та їх застосування під час війни, що важливо для врахування у процесі дослідження для оцінки бойових спроможностей військ.

Метою статті є розроблення методики оцінювання стійкості функціонування СЗ в сучасних умовах ведення бойових дій на основі бойових спроможностей військ.

Виклад основного матеріалу.

При проведенні дослідження обмежимося врахуванням деструктивного впливу на СЗ засобів ВУ і РЕП. Для забезпечення формалізації дослідження наведемо спрощену графічну модель функціонування СЗ в сучасних умовах ведення бойових дій (рис. 1).

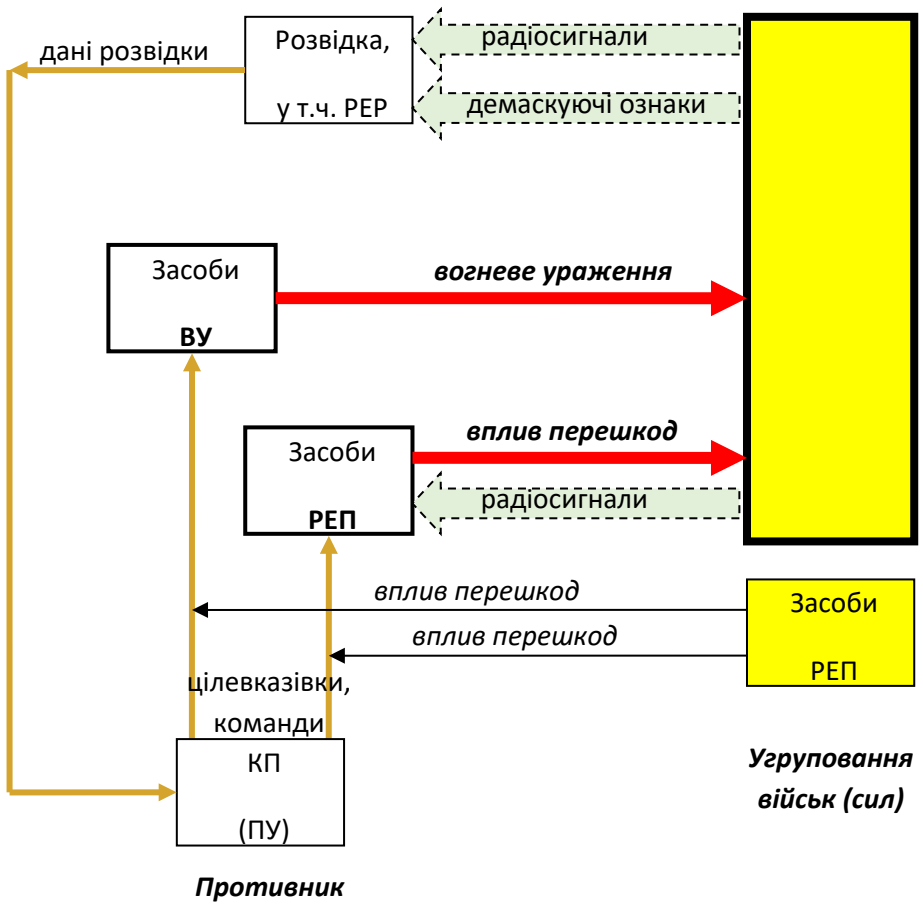


Рис. 1. Графічна модель функціонування СЗ в сучасних умовах ведення бойових дій

Вихідними даними для розрахунку стійкості функціонування СЗ в умовах радіоперешкод є: склад УВ, що бере участь в операції; база даних щодо кількості засобів зв'язку різного функціонального призначення (різних діапазонів), що знаходяться на озброєнні військових частин (підрозділів) зі складу угруповання; типовий розрахунок кількості радіоелектронних засобів (РЕЗ) у складі пунктів управління з'єднань (військових частин, підрозділів); розрахунок кількості радіомереж, що можуть бути організовані на пунктах управління з використанням визначених РЕЗ; склад засобів радіоелектронної боротьби (РЕБ) противника та їхні можливості з РЕП наших систем управління військами та зброєю; прогнозований ступінь ВУ, який може бути досягнутий противником; прогнозована навченість наших розрахунків вузлів зв'язку щодо виконання основних заходів радіоелектронного захисту – резервування каналів зв'язку, використання дротових мереж, виконання інших організаційних і технічних заходів (РЕЗт).

Розрахунок кількості РЕЗ, що входять до складу УВ, здійснюється завчасно, з метою створення бази даних та використання її під час аналізу можливого впливу засобів РЕБ противника на СЗ. Кількість РЕЗт розраховується як сума засобів зв'язку відповідних діапазонів хвиль, що знаходяться на озброєнні військових частин (підрозділів), які входять до складу УВ. Наприклад, кількість РЕЗт короткохвильового (КХ) діапазону розраховується наступним чином згідно з виразом (2):

$$K_{\text{зас.КХ}} = K_{\text{кх1}} + K_{\text{кх2}} + \cdots + K_{\text{кх}V}, \qquad (2)$$

де $K_{кхV}$ – кількість засобів зв'язку КХ діапазону різних типів, що знаходяться на озброєнні військових частин (підрозділів), які входять до складу УВ;

$v = 1,...,V$ – кількість типів РЕЗ КХ діапазону.

Розрахунок кількості РЕЗ інших діапазонів хвиль здійснюється аналогічним способом. За окрему військову частину та УВ розрахунок РЕЗ здійснюється методом підсумовування з оформленням відповідних значень у вигляді таблиць (окремо за кожну військову частину та УВ у цілому).

Розрахунок кількості радіомереж (радіоліній) управління УВ здійснюється з урахуванням аналізу складу підрозділів військових частин, кількості пунктів управління, що розгортаються, а також засобів зв'язку вузлів зв'язку, що знаходяться на цих пунктах управління. Кількість військових частин (пунктів управління (ПУ), радіомереж управління) залежить від кількості військових частин (підрозділів), що входять до складу певного УВ, і розраховується як сума згідно з виразами (3)–(5):

$$K_{в/ч} = K_{омбр} \times n_{омбр} + K_{омпбр} \times n_{омпбр} + \cdots + K_{оабр} \times n_{оабр}, \quad (3)$$

$$K_{ПУ} = K_{ПУомбр} \times n_{омбр} + K_{ПУомпбр} \times n_{омпбр} + \cdots + K_{ПУоабр} \times n_{оабр}, \quad (4)$$

$$K_{РМ} = K_{РМомбр} \times n_{омбр} + K_{РМомпбр} \times n_{омпбр} + \cdots + K_{РМоабр} \times n_{оабр}, \quad (5)$$

де n – кількість окремих військових частин кожного типу, які входять до складу УВ;

$K_{ПУомбр}$ – кількість розгорнутих ПУ для окремих механізованих бригад;

$K_{ПУомпбр}$ – кількість розгорнутих ПУ для окремих мотопіхотних бригад;

$K_{ПУоабр}$ – кількість розгорнутих ПУ для окремих артилерійських бригад;

$K_{РМомбр}$ – кількість розгорнутих радіомереж управління для окремих механізованих бригад;

$K_{РМомпбр}$ – кількість розгорнутих радіомереж управління для окремих мотопіхотних бригад;

$K_{РМоабр}$ – кількість розгорнутих радіомереж управління для окремих артилерійських бригад.

При проведенні розрахунків можливо використовувати декілька підходів. Обчислювати окремо кількість радіомереж або проводити розрахунки за кількістю типових пунктів управління (військових частин). У разі застосування військових частин не у повному складі вводяться відповідні коефіцієнти.

Розрахунок бойових можливостей сил і засобів РЕБ противника здійснюється з урахуванням висновків з оцінки радіоелектронної обстановки та залежить від складу військових частин (підрозділів) РЕБ противника, прогнозованих напрямків їх застосування, а також наявності відповідних засобів РЕБ. Оцінювання бойових можливостей військових частин (підрозділів) РЕБ противника обчислюється як сума можливостей з дезорганізації певних видів зв'язку (відповідних діапазонів хвиль).

Наступним кроком у методиці, що пропонується, є розрахунок ступеню дезорганізації нашої системи управління. Як відомо, існує три ступені дезорганізації управління: **"**зрив управління**"**, **"**порушення управління**"**, **"**утруднення управління**"**. Кожен ступінь дезорганізації залежить від можливості здійснювати одночасний радіоелектронний вплив перешкодами на РЕЗ для досягнення певної мети дезорганізації військами та зброєю. У загальному вигляді ефективність створення перешкод ($E_{РН}$) обчислюється як співвідношення бойових можливостей засобів РЕБ ($Б_{можл}$) зі здійснення одночасного радіоелектронного подавлення виявленої кількості радіомереж управління ($K_{РМ}$), а саме вирази (6), (7):

$$E_{\text{РП}} = \frac{Б_{\text{можл}}}{К_{\text{РМ}}}, \quad (6)$$

або у відсотках

$$E_{\text{РН}} = \frac{Б_{\text{можл}}}{К_{\text{РМ}}} \times 100\% . \quad (7)$$

Разом з наведеним для визначення ступеню дезорганізації управління слід враховувати ще ряд показників, які безпосередньо впливають на його значення. До таких показників слід віднести ефективність ВУ елементів СЗ противником та ефективність виконання заходів РЕЗт військовими частинами і підрозділами зв'язку.

Тому, для подальшої формалізації методики оцінювання стійкості функціонування СЗ в сучасних умовах доцільно перед використанням виразів (6), (7) обчислити кількість пунктів управління (елементів СЗ) і радіомереж, які ймовірно будуть знищені противником відповідно до виразів (8), (9):

$$К_{\text{зн.ПУ}} = К_{\text{ПУ}} \times С_{\text{ВУ}}, \quad (8)$$

$$К_{\text{зн.РМ}} = К_{\text{РМ}} \times С_{\text{ВУ}}, \quad (9)$$

де $С_{\text{ВУ}}$ – обраний ступінь ВУ;

$К_{\text{ПУ}}$ – кількість розгорнутих ПУ;

$К_{\text{РМ}}$ – кількість розгорнутих радіомереж.

При цьому, кількість ПУ, радіомереж і РЕЗ, що збереглися, розраховується як різниця між тими, що були спочатку, та тими, що виведені з ладу відповідно до виразів (10), (11):

$$К_{\text{збер.ПУ}} = К_{\text{ПУ}} - К_{\text{зн.ПУ}}, \quad (10)$$

$$К_{\text{збер.ПУ}} = К_{\text{ПУ}} - К_{\text{зн.ПУ}}, \quad (11)$$

Наведені вирази доцільно застосовувати для кожного із діапазонів хвиль окремо. В подальшому здійснюється визначення ефективності впливу перешкодами противника на нашу СЗ згідно з виразом (12):

$$E_{\text{РП}} = \frac{Б_{\text{можл}}}{К_{\text{збер.РМ}}}, \quad (12)$$

а також ступінь дезорганізації нашої системи управління – вираз (13):

$$С_{\text{тДУ}} = \frac{Б_{\text{можл}}}{К_{\text{збер.РМ}}} \times 100\%. \quad (13)$$

Залежно від показника впливу перешкодами, з урахуванням ВУ, визначається ступінь дезорганізації управління (порушено, зірвано або утруднено).

При цьому стійкість функціонування СЗ у відсотках оцінюється за наступним виразом (14):

$$С_{\text{тСЗ}} = 100 - С_{\text{тДУ}} = 100 - \frac{Б_{\text{можл}}}{К_{\text{збер.РМ}}} \times 100\%. \quad (14)$$

Наступним кроком потрібно врахувати виконання заходів РЕЗт. Підрозділи зв’язку будуть виконувати заходи резервування мереж зв’язку, комплексне використання засобів зв’язку на інформаційних напрямках, організаційні і технічні заходи РЕЗт, чим буде суттєво підвищено стійкість функціонування СЗ. Врахуємо підвищення стійкості СЗ у вигляді $C_{\text{додатСЗ}}$. З урахуванням наведеного вирази (13) і (14) перепишуться до наступного вигляду (15), (16):

$$C_{\text{тДУ}} = \frac{Б_{\text{можл}}}{(1+C_{\text{додатСЗ}})K_{\text{збер.РМ}}} \times 100\%, \quad (15)$$

$$C_{\text{тСЗ}} = 100 - \frac{Б_{\text{можл}}}{(1+C_{\text{додатСЗ}})K_{\text{збер.РМ}}} \times 100\%. \quad (16)$$

Висновок. На основі викладеної методики авторами було розроблено інформаційно-розрахункову задачу. Використання задачі (розробленої в Excel) передбачає введення лише даних за свої війська та засоби перешкод противника, а також можливий ступінь ВУ, визначений противником. Введення даних не перевищує за часом 5 хвилин і дозволяє протягом зазначеного терміну отримати інформацію щодо стійкості функціонування нашої СЗ. Наведене надасть змогу начальникам органів планування зв’язку і РЕБ підготувати командувачу обґрунтовані пропозиції для прийняття рішення. Крім того, ця методика дозволяє провести відповідний аналіз та надати командувачу обґрунтовані пропозиції щодо поліпшення функціонування своєї СЗ.

Шляхами майбутніх досліджень авторами вважається подальше удосконалення запропонованої методики завдяки введенню нових зовнішніх та внутрішніх факторів, що впливають на функціонування системи військового зв’язку, і практична перевірка методики в бойових умовах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Боговик А., Игнатов В. Эффективность систем военной связи и методы ее оценки. СПб: ВАС, 2006. 152 с.

2. Макаров С. А., Бєляк С. П., Висоцький О. В. Оцінка стійкості системи зв’язку та радіотехнічного забезпечення. *Системи озброєння і військова техніка*. 2020. № 1 (61). С. 44–50. URL: https://doi.org/10.30748/soivt.2020.61.05.

3. Методика оцінки стійкості системи військового зв’язку / М. Масесов та ін. *Збірник наукових праць Військового інституту телекомунікацій та інформатизації*. 2016. № 1. С. 94–102.

4. Васюта К., Чекунова О., Макаров С. Застосування досвіду АТО та особливостей проведення ООС для підготовки фахівців зв’язку, РТЗ, А та ІС. Харків: ХНУПС, 2018. 292 с.

5. Шолудько В. Г., Єсаулов М. Ю., Вакуленко О. В., Гурський Т. Г., Фомін М. М.. Організація військового зв’язку: навч. посіб. К.: ВІТІ, 2017. 282 с.

6. Шевченко В. О. Системний підхід до розроблення методологічних основ дослідження телекомунікаційних мереж військового призначення. *Наука і оборона*. 2004. № 4. С. 42–46.

7. Эффективность и электронная защита военных систем связи / А. Н. Авсюкевич и др. Л.: ВОЛКАС, 1980. 32 с.

8. Кучеренко Ю. Ф. Оцінка ефективності автоматизованих систем управління міжвидових угруповань військ. *Наука і техніка Повітряних Сил Збройних Сил України ХУПС*. 2013. № 2 (11). С. 49–51.

9. Управління проектами зі створення зразків озброєння та військової техніки в умовах прояву факторів невизначеності та ризику / Б. О. Демідов та ін. *Озброєння та військова техніка*. К.: ЦНДІ ОВТ ЗС України, 2016. № 2 (10). С. 15–19.

10. Modern means of communication and information communication technologies in of the Armed Forces of Ukraine and the National Guard of Ukraine: present and prospects of application / O. Iokhov et al.

The scientific journal of the National Academy of National Guard «Honor and Law». 2022. Vol. 4, no. 83. P. 111–119. URL: https://doi.org/10.33405/2078-7480/2022/4/83/272323.

11. Лаврут О. О., Іохов О. Ю. Сучасні засоби зв’язку та інфокомунікаційні технології у Збройних Силах України: сьогодення та перспективи застосування. Застосування Сухопутних військ Збройних Сил України у конфліктах сучасності (за досвідом забезпечення національної безпеки складовими сектору безпеки і оборони у російсько-українській війні в 2022 році): тези доп. наук.-практ. конф., м. Львів, 17 листоп. 2022 р. Львів: НАСВ ім. гетьмана Петра Сагайдачного, 2022. С. 11.

12. Лаврут О. О., Лаврут Т. В., Климович О. К., Здоренко Ю. М. Новітні технології та засоби зв’язку у Збройних Силах України: шлях трансформації та перспективи розвитку. *Наука і техніка Повітряних Сил Збройних Сил України*. 2019. № 1 (34). С. 91–101. URL: https://doi.org/10.30748/nitps.2019.34.13.

13. Методика оцінювання стійкості системи зв’язку, радіотехнічного забезпечення та автоматизації управління повітряного командування / М. В. Кас’яненко та ін. *Системи озброєння і військова техніка*. 2020. № 1 (61). С. 13–22. URL: https://doi.org/10.30748/soivt.2020.61.02.

DOI: 10.58254/viti.5.2024.06.71

УДК 681.61

канд. техн. наук Ільницький А. І. ORCID: 0000-0001-5817-4917
(ННІ ТС НТУУ «КПІ ім. Ігоря Сікорського»)
Захарчук Л. В. ORCID: 0009-0004-3171-6332 (ННІ ТС НТУУ «КПІ ім. Ігоря Сікорського»)

МАТЕМАТИЧНА МОДЕЛЬ ОДНОКАНАЛЬНОГО ПОШУКУ СИГНАЛІВ ЗАСОБІВ МОБІЛЬНОГО РАДІОЗВ'ЯЗКУ З ОДНИМ СТУПЕНЕМ ВИЯВЛЕННЯ В ЧАСТОТНО-ЧАСОВІЙ ОБЛАСТІ

В умовах повномасштабного вторгнення і ведення бойових дій за незалежність України питання радіомоніторингу джерел радіовипромінювання засобів телекомунікаційних мереж і систем набувають особливої уваги та вимагають принципово нових підходів до рівня інформатизації, технічної реалізації та підвищення їхньої ефективності.

На сьогодні технічні засоби радіоперехоплення, моніторингу й радіопеленгації в мережах систем мобільного радіозв'язку реалізуються у вигляді програмно-апаратних комплексів, найважливішими показниками ефективності яких вважається швидкодія, точність визначення та ймовірність розпізнавання засобів рухомого радіозв'язку з їхнім інформаційним наповненням. Водночас ці питання і дотепер залишаються проблематичними і потребують подальшого розвитку методів і способів пошуку та виявлення сигналів засобів рухомого радіозв'язку як у частотних, так і в часових телекомунікаційних каналах, та їх інформаційної обробки.

Для вирішення вказаного питання авторами запропоновано математичну модель пошуку сигналів мобільного радіозв'язку у частотно-часовій області з одним ступенем виявлення. Опис математичної моделі визначення спектральних складових сигналів у частотній і часовій областях здійснюється за допомогою основних положень теорії напрямлених імовірнісних графів, геометричних імовірностей з її відомою задачею про зустріч у заданий проміжок часу та теоретичних основ статистичної обробки інформації.

Аналіз запропонованої моделі свідчить, що ймовірність успішного завершення пошуку сигналу з одним ступенем виявлення в частотно-часовій області за заданий час залежить від значення тривалості сигналу, кількості аналізованих частотних каналів, часу аналізу сигналу в кожному каналі та ймовірностей похибок першого і другого роду. У статті наведено результати розрахунків значення ймовірностей успішного завершення пошуку сигналу від часу пошуку при різних значеннях часу аналізу. Показано, що застосування математичної моделі одноканального пошуку сигналів засобів мобільного радіозв'язку з одним ступенем виявлення в частотно-часовій області на практиці дасть можливість визначати якість процесу радіомоніторингу і підвищувати ефективність оцінювання спектральних складових радіосигналів за показниками швидкодії, точності визначення та ймовірності розпізнавання засобів мобільного радіозв'язку з подальшим визначенням їхнього інформаційного наповнення.

Ключові слова: *радіомоніторинг, телекомунікаційні мережі, системи мобільного зв'язку, показники ефективності, швидкодія, точність визначення, імовірність розпізнавання.*

A. Ilnytskyi, L. Zakharchuk Mathematical model of single-channel search for signals of mobile radio communications with one degree of detection in the frequency-time domain.

In the context of a full-scale invasion and military operations for Ukraine's independence, the issues of radio monitoring of radio emission sources of telecommunication networks and systems are gaining special attention and require fundamentally new approaches to the level of informatization, technical implementation and increase of their efficiency.

Today, the technical means of radio interception, monitoring and direction finding in the networks of mobile radio communication systems are realized in the form of software and hardware complexes, the most important performance indicators of which are considered to be speed, accuracy of detection and probability of recognition of mobile radio communication means with their information content. At the same time, these issues still remain problematic and require further development of methods and techniques for searching and detecting signals of mobile radio communications in both frequency and time telecommunication channels and their information processing.

To solve this issue, the authors propose a mathematical model for searching for mobile radio signals in the frequency-time domain with one degree of detection. The description of the mathematical model for determining the spectral components of signals in the frequency and time domains is carried out using the basic provisions of the theory of directed probability graphs, geometric probabilities with its well-known problem of meeting in a given time interval, and the theoretical foundations of statistical information processing.

The analysis of the proposed model shows that the probability of successful completion of the search for a signal with one degree of detection in the frequency-time domain for a given time depends on the value of the signal duration,

the number of analyzed frequency channels, the time of signal analysis in each channel, and the probabilities of errors of the first and second kinds The article presents the results of calculations of the value of the probability of successful completion of the signal search as a function of the search time at different values of the analysis time. It is shown that the application of the mathematical model of single-channel search for signals of mobile radio communication means with one degree of detection in the frequency-time domain in practice will make it possible to determine the quality of the radio monitoring process and increase the efficiency of evaluating the spectral components of radio signals in terms of speed, detection accuracy and probability of recognition of mobile radio communication means with further determination of their information content.

Keywords: *radio monitoring, telecommunication networks, mobile communication systems, efficiency indicators, speed, detection accuracy, recognition probability.*

Постановка проблеми. В умовах повномасштабного вторгнення і ведення бойових дій за незалежність України питання радіомоніторингу (РМ) джерел радіовипромінювання (ДРВп) засобів телекомунікаційних мереж і систем (ТКМС) набувають особливої уваги. На сьогодні технічні засоби радіоперехоплення, моніторингу й радіопеленгації в мережах безпроводових систем мобільного радіозв’язку (МРЗ) реалізуються у вигляді програмно-апаратних комплексів (ПАК) різних модифікацій, які працюють в умовах часткової або повної невизначеності параметрів вхідних сигналів у реальному масштабі часу на тлі величезної кількості і різноманітності ДРВп з найсучаснішими протоколами передачі даних. При цьому, найважливішими показниками ефективності ПАК при моніторингу безпроводових систем МЗР вважають швидкодію, точність визначення та ймовірність розпізнавання ДРВп рухомого радіозв’язку з їхнім інформаційним наповненням. Однак ці питання і дотепер залишаються проблематичними, що потребує подальшого розвитку ефективних методик, методів та способів пошуку і виявлення тональних, мовних і спеціальних сигналів МРЗ у частотно-часових телекомунікаційних каналах та їх інформаційну обробку [1; 2]. Також швидкий розвиток сучасних телекомунікаційних і комп’ютерних технологій, обчислювальної техніки, радіоелектроніки тощо призводить до появи нових технічних засобів різноманітного призначення, в тому числі й подвійного значення, які мають і оперують вже з новими «нетрадиційними» інформаційними ознаками. Вказані обставини потребують вдосконалення і розвитку систем РМ та протидії, поперед всього з позицій та методології високоефективних інтелектуальних систем, що забезпечують збір, обробку, накопичення і використання отриманої моніторингової інформації для прийняття рішення в масштабі реального часу.

Аналіз останніх досліджень і публікацій. У структурі ПАК моніторингу ДРВп засобів МРЗ основне місце має підсистема радіоприймальних пристроїв (РПП), яка вирішує завдання виявлення сигналів у частотно-часовій області за циклічними процедурами з одним або декількома ступенями виявлення, що реалізують методи послідовно-паралельного пошуку й виявлення сигналів. Як приклад можна навести мобільну станцію радіомоніторингу MMS-02, яка має такі технічні характеристики [3]: смуга радіочастот – 30–3000 МГц; чутливість – не нижче 0 дБмкВ/м (1 мкВ/м); динамічний діапазон – не гірше 75 дБ; максимальна швидкість панорамного огляду – не менше 300 МГц/с; невизначеність вимірювання рівня – не більше 3 дБ; точність вимірювання частоти – не менше 10^{-6}.

Під час пошуку сигналів МРЗ у частотній області з одним ступенем виявлення РПП налаштовується на частоту певного каналу (1, 2,…, N), де протягом заданого часу аналізу $t_a = \text{const}$ приймається рішення про наявність або відсутність сигналу. У разі відсутності сигналу приймач переналаштовується на частоту наступного каналу й т. ін. З досягненням останнього каналу РПП налаштовується на частоту першого каналу й пошук починається знову.

Цю процедуру пошуку можна надати у вигляді напрямленого ймовірнісного графу [4] (рис. 1) і, спираючись на теорію напрямлених імовірнісних графів [5], описати математично.

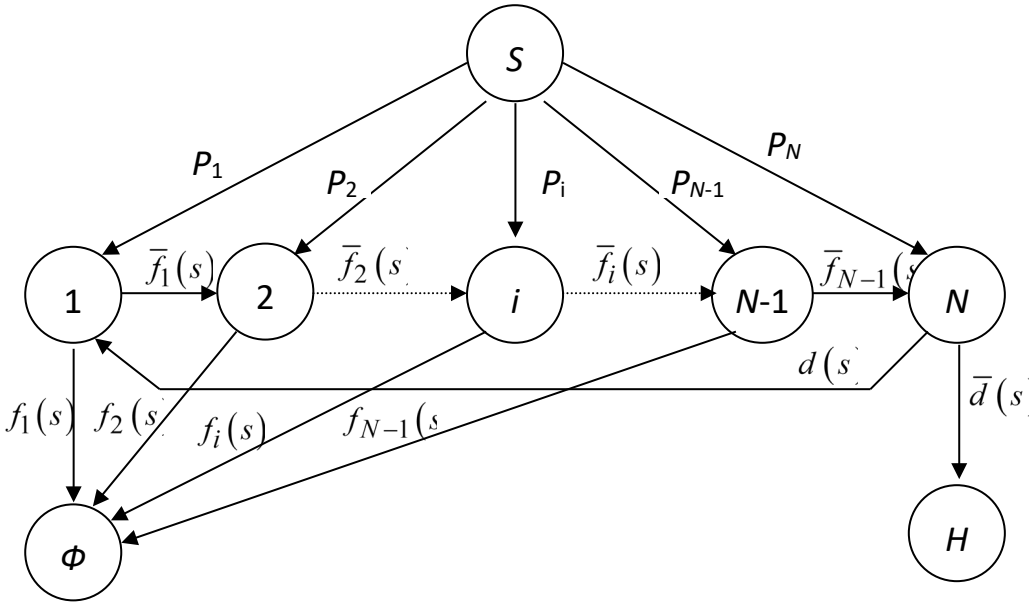


Рис. 1. Граф циклічної процедури одноканального пошуку з одним ступенем виявлення

Вузли графа відображають можливі стани підсистеми пошуку: S – початковий стан; $H(s)$ – поглинальний стан, що відповідає успішному завершенню пошуку сигналу (правильне розпізнавання); $\Phi(s)$ – поглинальний стан, що відповідає помилковому розпізнаванню сигналу. Ребра графа відображають можливі переходи системи з одного стану в інший. При цьому кожне ребро графа має ваги, які визначаються за (1) і враховують імовірність переходу за один крок зі стану, з якого ребро виходить, у стан, в яке воно входить, і час перебування у відповідному стані:

$$d(s) = De^{-t_a s}, \quad \overline{d}(s) = (1-D)e^{-t_a s};$$
$$f_q(s) = F_q e^{-t_a s}, \quad \overline{f}_q(s) = (1-F_q)e^{-t_a s}, \tag{1}$$

де F_q – імовірність помилкового виявлення (похибка першого роду) під час аналізу осередку з номером q (q = 1, 2,…, N-1), що не утримує сигнал;

D – імовірність пропуску сигналу (похибка другого роду) під час аналізу осередку з номером N.

За правилами перетворення графів [6] можна одержати такі вирази для функцій переходу графа з початкового стану S у поглинальні:

стан успішного завершення пошуку (правильне розпізнавання):

$$H(s) = \frac{\overline{d}(s)\sum_{i=1}^{N} P_i \prod_{q=i}^{N-1} \overline{f}_q(s)}{1 - d(s)\prod_{q=1}^{N-1} \overline{f}_q(s)}; \tag{2}$$

стан помилкового розпізнавання сигналу:

$$\Phi(s) = \sum_{i=1}^{N} P_i \frac{\sum_{q=i}^{N-1} f_q(s) \prod_{l=i}^{q-1} \overline{f}_l(s)}{1 - d(s)\prod_{q=1}^{N-1} \overline{f}_q(s)} + \frac{d(s)\prod_{q=i}^{N-1} \overline{f}_q(s)}{1 - d(s)\sum_{q=1}^{N-1} f_q(s)} \sum_{j=1}^{i-1} f_j(s) \prod_{t=1}^{j-1} \overline{f}_t(s), \tag{3}$$

де P_i – імовірність початку перегляду з i-го осередку (i = 1, 2,…, N).

Імовірність успішного завершення пошуку за заданий час визначається як зворотне перетворення Лапласа [5; 6] $P^f_{\Pi}(t) = L^{-1}\{H(s)/s\}$ і після подання $H(s)$ у вигляді геометричного ряду отримаємо, що:

$$P_{\Pi}^{f}(t) = \frac{1-D}{N}\sum_{i=1}^{N}\sum_{n=0}^{\infty} D^{n}(1-F)^{n(N-1)+N-i} h\left[t-(Nn+N-i+1)\ t_a\right], \quad (4)$$

де $h(x)$ – одинична функція;

t_a – час аналізу одного осередку;

$F_1 = F_2 = \ldots = F_{N-1} = F$ – ймовірності помилкового виявлення в кожному каналі, що не містить сигнал;

$P_i = 1/N$ – рівномірний розподіл моментів появи сигналу усередині області пошуку.

Для обчислення ймовірності успішного завершення пошуку за час t необхідно в (4) просумувати всі складові з індексами i, n, що задовольняють нерівності:

$$(Nn+N-i+1)\ t_a \le t. \quad (5)$$

Отже, наведений вираз (4) можна вважати математичним описом (моделлю) успішного завершення пошуку (правильне розпізнавання) сигналу за час t_a аналізу N частотних каналів.

Однак, як показує практика, за час t_a аналізу N каналів сигнал може зникнути або з'явитися при його виявленні. При цьому моменти його появи або зникнення можуть не збігатися з моментами початку або закінчення РМ цього каналу. Крім того, сигнал у каналі може з'явитися й на попередньому циклі перегляду РПП N каналів. Тобто виникає необхідність проаналізувати стан й математично описати підсистеми пошуку сигналу не тільки в частотній, але й у часовій області.

Метою роботи є опис математичної моделі одноканального пошуку сигналів засобів мобільного радіозв'язку з одним ступенем виявлення в частотно-часовій області. Тобто об'єктом дослідження в статті є процес радіомоніторингу джерел радіовипромінювання засобів телекомунікаційних мереж і систем, а предметом дослідження – математична модель одноканального пошуку сигналів засобів мобільного радіозв'язку з одним ступенем виявлення в часовій області. Застосування математичної моделі на практиці дасть можливість підвищити ефективність оцінювання спектральних складових радіосигналів за показниками швидкодії, точності визначення та ймовірності розпізнавання засобів мобільного радіозв'язку з подальшим визначенням їхнього інформаційного наповнення.

Виклад основного матеріалу дослідження. Для опису одноканальної підсистеми пошуку сигналів МРЗ у часовій області можна скористатись теорією геометричних імовірностей та її відомою задачею про зустріч у заданий проміжок часу [5; 6] та розрахувати ймовірність зустрічі РПП й сигналу за час t_a аналізу N каналів, тобто протягом одного циклу перегляду T за умови, що тривалість шуканого сигналу τ_c не менша часу аналізу одного каналу $\tau_c \ge t_a$. При цьому тривалість одного циклу перегляду визначається як $T = t_a N$.

Розрахунок зазначеної ймовірності проведемо за допомогою рисунка 2, де координата x відображає момент надходження РПП в певний частотний канал, а y – момент появи в ньому сигналу. Як видно, на рисунку 2 виділено дві області: область M точок [$T \ge x \ge 0;\ T \ge y \ge -\tau_c$] стану всієї процедури пошуку і область A точок [$x \ge y;\ y \ge x - (\tau_c - t_a)$], що задовольняють умовам пошуку.

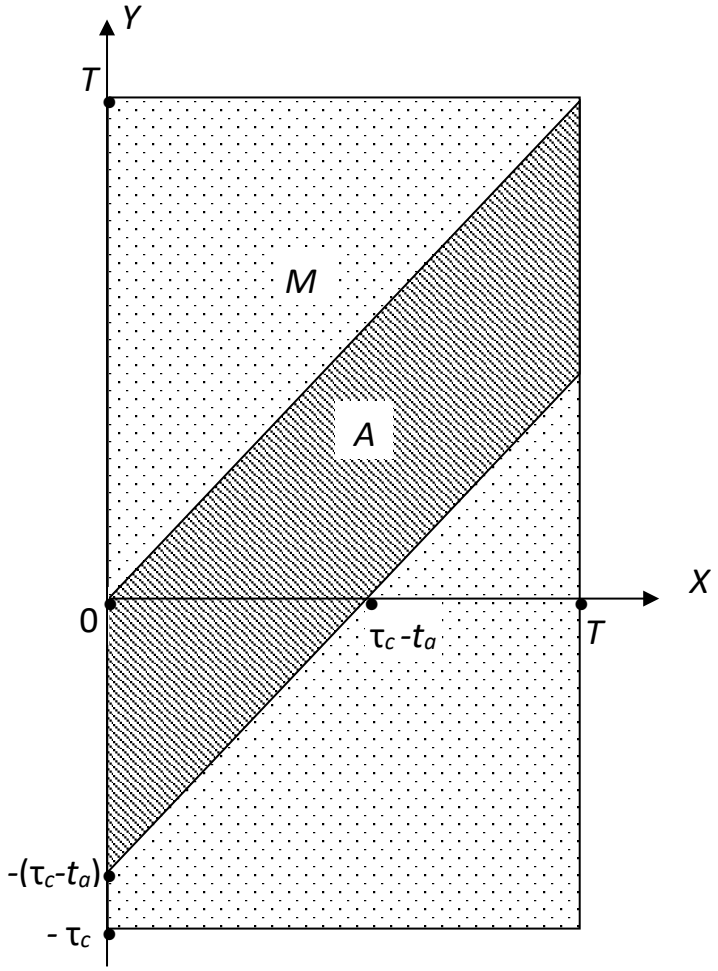


Рис. 2. Геометричне зображення областей пошуку сигналу

За визначенням геометричної ймовірності треба, щоб ймовірність $P^t_П$ попадання точки в область A дорівнювала відношенню площі області S_A до загальної площі області S_M:

$$P^t_П = S_A / S_M.$$

Після обчислення S_A і S_M, з урахуванням рисунка 2, одержимо формулу для розрахунку цієї ймовірності:

$$P^t_П = \frac{\tau_c - t_a}{\tau_c + Nt_a}. \quad (6)$$

Отже, наведений вираз (6) можна вважати математичним описом (моделлю) успішного завершення пошуку (правильне розпізнавання) сигналу τ_c за час t_a аналізу N частотних каналів з одним ступенем виявлення в часовій області.

Отримані математичні вирази (4)–(6) дають підставу розглянути особливості пошуку сигналів з одним ступенем виявлення в частотно-часовій області, де факт одночасного виявлення сигналу в частотній і часовій областях можна розглядати як незалежні випадкові події. При цьому згідно з [5; 6] і врахуванням того, що ймовірність успішного завершення пошуку сигналу в часовій області визначалася за один цикл перегляду РПП всіх каналів, імовірність пропуску сигналу $D_П$ (похибка другого роду) під час пошуку в частотно-часовій області збільшується й визначається за формулою:

$$D_П = 1 - (1 - D) P^t_П \quad (7)$$

З урахуванням (4), (6) і (7) імовірність успішного завершення пошуку за заданий час з одним ступенем виявлення в частотно-часовій області визначається як

$$P_{\text{П}}(t)=\frac{(\tau_c-t_a)(1-D)}{N(\tau_c+Nt_a)}\sum_{i=1}^{N}\sum_{n=0}^{\infty}\left(1-\frac{(\tau_c-t_a)(1-D)}{\tau_c+Nt_a}\right)^n\times$$
$$\times(1-F)^{n(N-1)+N-i}\,h\left[t-(Nn+N-i+1)\;t_a\right]. \qquad (8)$$

З наведеного виразу (8) видно, що ймовірність успішного завершення пошуку сигналу з одним ступенем виявлення в частотно-часовій області за заданий час залежить від значення тривалості сигналу τ_c, кількості аналізованих частотних каналів N, часу аналізу сигналу в кожному каналі t_a та ймовірностей похибок першого F і другого роду D. Із цих величин реально можна змінювати тільки час аналізу сигналу в кожному каналі. Розраховані за виразом (8) значення імовірностей $P_{\text{П}}$ успішного завершення пошуку сигналу від часу пошуку при різних значеннях t_a = 5 мс, 10 мс, 20 мс наведено в таблиці 1 і зображено графічно на рисунку 3, за умови, що τ_c = 100 мс, F = 0,001, N = 10, D = 0,1.

Таблиця 1

Імовірності $P_{\text{П}}$ успішного завершення пошуку сигналу від часу пошуку

t, *мс*	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0	1,1
t_a = 5 мс	**0,81**	**0,95**	**0,98**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**
t_a = 10 мс	**0,4**	**0,64**	**0,78**	**0,86**	**0,91**	**O,94**	**0,96**	**0,97**	**0,98**	**0,98**	**0,98**
t_a = 20 мс	**0,12**	**0,24**	**0,33**	**0,42**	**0,49**	**0,55**	**0,61**	**0,66**	**0,69**	**0,73**	**0,76**

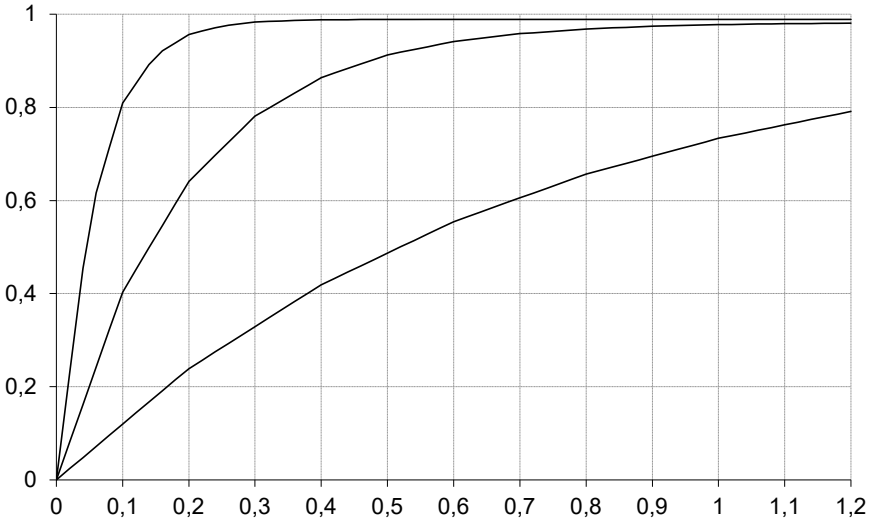


Рис. 3. Залежність імовірності $P_{\text{П}}$ успішного завершення пошуку сигналу з одним ступенем виявлення від часу пошуку пошуку при різних значеннях t_a = 5 мс, 10 мс, 20 мс

Аналізуючи отримані залежності, можна зробити висновок, що для збільшення ймовірності успішного завершення пошуку сигналу й зменшення його часу потрібно зменшувати час аналізу сигналу t_a в кожному каналі. Для цього можна запропонувати використання циклічної процедури пошуку сигналів ДРВп з декількома ступенями виявлення й різним часом аналізу на кожному ступені. Однак ці питання виходять за межі цієї статті і можуть бути розглянуті окремо.

Висновки.

1. Розроблена математична модель одноканального пошуку сигналів засобів мобільного радіозв'язку з одним ступенем виявлення в частотно-часовій області (8) дає можливість

розрахувати ймовірність зустрічі РПП й сигналу за час аналізу N каналів, тобто протягом одного циклу перегляду $T = t_a N$ за умови, що тривалість шуканого сигналу τ_c не менша часу аналізу одного каналу $\tau_c \geq t_a$.

2. Ймовірність успішного завершення пошуку сигналу з одним ступенем виявлення в частотно-часовій області за заданий час залежить від значення тривалості сигналу τ_c, кількості аналізованих частотних каналів N, часу аналізу сигналу в кожному каналі t_a та ймовірностей похибок першого F і другого роду D.

3. Застосування математичної моделі одноканального пошуку сигналів засобів мобільного радіозв'язку з одним ступенем виявлення в частотно-часовій області на практиці дасть можливість підвищити ефективність оцінювання спектральних складових радіосигналів за показниками швидкодії, точності визначення та імовірності розпізнавання засобів мобільного радіозв'язку (див. рис. 3) з подальшим визначенням їхнього інформаційного наповнення.

4. Достовірність результатів досліджень, що наведені у статті, забезпечується несуперечністю отриманих науково-практичних результатів загальновідомим фізичним та математичним положенням теорії цієї предметної області науки і техніки та добрим погодженням із результатами, що наведені в відомих літературних джерелах [7–10].

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Ільченко М. Ю., Кравчук С. О. Телекомунікаційні системи: монографія. Київ, 2017. С. 730.
2. Ільницький А. І., Захарчук Л. В. Проблемні питання радіомоніторингу джерел і об'єктів телекомунікаційних мереж і систем та загально-науковий концептуальний підхід до їх розв'язання. *Системи і технології зв'язку, інформації та кібербезпеки: актуальні питання і тенденції розвитку:* матеріали ІІІ Міжнародної наук.-техн. конф., м. Київ, 30 листопада 2023 р. Київ, ВІТІ, 2023. С. 167. URL: https://drive.google.com/file/d/1nTFvkmgQSrIvdQgFvzSDyz1aStDJpQOC/view (дата звернення: 03.02.2024).
3. Слободянюк П. В., Благодарний В. Г., Ступак В. С. Довідник з радіомоніторингу: посібник. Ніжин, 2008. С. 341. URL: https://duikt.edu.ua/uploads/l_730_62703296.pdf (дата звернення: 15.01.2024).
4. Ільницький А. І., Стретович О. В. Алгоритми декодування тональних сигналів у комплексах розвідки систем рухомого радіозв'язку. *Труди академії.* 2005. № 65. С. 70–75.
5. Гнєденко Б. В. Курс теорії ймовірностей: підручник. Київ, 2010. С. 464. URL: https://probability.knu.ua/userfiles/yamnenko/Gnedenko.pdf (дата звернення: 28.12.2023).
6. Grimmett G. R., Stirzaker D. R. Probability and Random Processes. Oxford University Press Inc., New York. 2001. P. 596. URL: https://www.academia.edu/31210737/Probability_and_random_process_by_geoffrey_grimmett_and_david_stirzaker (дата звернення: 14.02.2024).
7. Ільницький А. І., Ільяшов О. А. Інформативність розвідувальних ознак і сигнатур та міра їх невизначеності при розпізнаванні джерел і об'єктів розвідки. *Збірник наукових праць ЦНДІ ЗСУ*. 2009. № 25. С. 25–40.
8. Ільницький А. І., Бурба О. І., Пасічник О. О. Статистичні характеристики інформаційних ознак джерел випромінювання при радіомоніторингу телекомунікаційних мереж. *УСиМ*. 2016. № 6. С. 48–58. URL: http://usim.org.ua/arch/2016/6/7.pdf (дата звернення: 03.04.2024).
9. Ільницький А. І., Бурба О. І. Статистичні оцінки незміщеності та ефективності параметрів джерел радіовипромінювання при радіомоніторингу телекомунікаційних мереж. *Системи управління навігації та зв'язку.* 2018. № 2 (48). С. 149–153. URL: http://nbuv.gov.ua/UJRN/suntz_2018_2_31 (дата звернення: 03.04.2024).
10. Стретович О. В., Федірко О. М., Рубальський П. С. Математична модель підсистеми пошуку радіосигналів систем рухомого радіозв'язку. *Збірник наукових праць ЦНДІ ЗСУ*. 2005. № 14. С. 36–43 (дата звернення: 03.04.2024).

DOI: 10.58254/viti.5.2024.07.78

УДК 621.317

д-р техн. наук, професор Міночкін А. І. ORCID: 0000-0001-5723-5052 (ВІТІ ім. Героїв Крут)
д-р техн. наук, професор Кузавков В. В. ORCID: 0000-0002-0655-9759 (ВІТІ ім. Героїв Крут)
канд. техн. наук Клімович С. О. ORCID: 0000-0001-7209-2176 (ВІТІ ім. Героїв Крут)

ПЕРСПЕКТИВИ СТВОРЕННЯ ТА РОЗВИТКУ СИСТЕМ ДІАГНОСТУВАННЯ РАДІОЕЛЕКТРОННИХ ЗАСОБІВ ІЗ ВБУДОВАНИМ ПРОГРАМНИМ ЗАБЕЗПЕЧЕННЯМ

У статті пов'язані між собою питання стану військово-політичних відносин України (стійкий напрямок до інтеграції зі стандартами НАТО та реальні бойові дії на сході країни) зі створенням ефективної системи контролю технічного стану озброєння та військової техніки (іноземних виробників різних країн та років випуску) в місцях їхнього безпосереднього застосування.

Зміни, які відбуваються у світі, призводять до необхідності проведення відповідних змін і у середині України. Значний процент сучасної зброї з боку супротивника, а також стрімке зростання кількості зброї іноземного виробництва в нашій країні спонукає до перегляду поглядів на форми та способи ведення бойових дій, форми логістичного забезпечення, стану та організації ефективної системи технічного забезпечення та ремонту військової техніки й озброєння.

Під ефективною системою контролю технічного стану (діагностування) радіоелектронного озброєння розуміємо програмно-апаратні засоби визначення технічного стану (на базі комп'ютерно-вимірювальних систем), які забезпечують підвищення показників надійності й ефективності функціонування озброєння при оптимальному співвідношенні витрати на систему – ефект від застосування. При цьому, прийняття рішення про фактичний технічний стан об'єкта контролю здійснюється шляхом розв'язання задачі аналізу значень фізичних величин, виміряних безконтактним методом.

На сьогодні, при зростаючій складності сучасного радіоелектронного озброєння, збільшенні числа контрольованих параметрів, мініатюризації розмірів елементів, а також низького рівня оснащення системами, які виконують функції прогнозування та діагностування, актуальними є завдання щодо прогнозування відмов, зменшення часу відновлення, економії ресурсів сил та засобів, необхідних для вирішення задач діагностування, особливо на місцях експлуатації радіоелектронного озброєння.

Ключові слова: *контроль технічного стану, радіоелектронне озброєння, вбудоване програмне забезпечення, комп'ютерно-вимірювальна система, надійність, ефективність.*

A. Minochkin, V. Kuzavkov, S. Klimovych Prospects for the creation and development of radio electronic devices diagnostic systems with built-in software.

In the article, the issues of the current state of military-political relations of Ukraine (stable direction towards integration with NATO standards and real combat operations in the east of the country) are connected with the creation of an effective system of monitoring the technical condition of weapons and military equipment (foreign manufacturers of different countries and years of manufacture) in places of their direct application.

Changes in military and political relations that are taking place in the world lead to changes both at the global level and in the middle of Ukraine. The level of the enemy's equipment with modern weapons, as well as the rapid development of scientific and technical progress, prompts a review of views on the forms and methods of conducting hostilities, the form of logistical support, the state and organization of an effective system of technical support and repair of military equipment and weapons.

By an effective system for monitoring the technical condition (diagnosis) of radio-electronic weapons, we mean a set of hardware and software tools (based on computer-measurement systems) that are designed to solve the problems of technical diagnosis and ensure an increase in the reliability and efficiency of the weapon with an optimal cost-effectiveness ratio on the system - the effect of application.

At the moment, with the increasing complexity of modern radio-electronic weapons, the increase in the number of controlled parameters, the miniaturization of the size of elements, as well as the low level of equipment with systems that perform the functions of forecasting and diagnosis, the tasks of predicting failures, reducing recovery time, saving resources of forces and means are urgent necessary for solving diagnostic problems, especially at the sites of operation of radio-electronic weapons.

Keywords: *technical condition control, radio-electronic weapons, embedded software, computer-measurement system, reliability, efficiency.*

Постановка проблеми. Система діагностування технічного стану радіоелектронного озброєння (далі – РЕО) є важливим інструментом забезпечення надійності технічних систем та безпеки її застосування. Ця система повинна здійснювати моніторинг та контролювати основні параметри устаткування, оперативно сигналізувати про можливі несправності, запобігати їх виникненню, а у випадку виявлення відхилень від заданих нормативів – автоматично генерувати попередження або сигнали тривоги оператору.

Для об’єктів РЕО із вбудованим програмним забезпеченням (далі – ВПЗ) система діагностування може бути інтегрована до основного устаткування на різних рівнях (на апаратному або програмному), або, як перспективна альтернатива, виконуватися у вигляді автономних уніфікованих систем контролю зі складними алгоритмами обробки інформації та доступом до глобальних інформаційних баз. На рівні апаратного забезпечення система оснащується датчиками та сенсорами безперервного моніторингу основних параметрів об’єкта контролю (далі – ОК). На рівні програмного забезпечення (далі – ПЗ) система повинна використовувати алгоритми та моделі для аналізу та обробки отриманих діагностичних даних.

Аналіз основних досліджень і публікацій. Сучасні системи діагностування (далі – СД) РЕО можуть бути вбудованими в апаратуру або надаватись їй. Фізично такі системи виконуються на програмно-логічних інтегральних схемах чи з використанням мікропроцесорної техніки. Перспективним вважається метод автоматичного самоконтролю [1; 2]. Відповідно до основного принципу вказаного методу, контрольне обладнання повинно проєктуватися одночасно з основним устаткуванням. Збільшення кількості територіально рознесених пристроїв із провідним та бездротовим інтерфейсом, до яких належать не тільки засоби зв’язку, а й комунікаційні пристрої, об’єднані в мережі, їхня мініатюризація, також сприяють підвищенню вимог до СД РЕО.

У нашій країні та за кордоном ведеться значна кількість робіт із дослідження, розробки, виробництва та експлуатації СД РЕО. Ці системи різні за призначенням, складністю, вартістю, принципами управління та обробки інформації, а також технічним рівнем їх виконання. Зарубіжний досвід показує, що застосування CALS-технологій (Continuous Acquisition and Life-Cycle Support – інформаційна підтримка життєвого циклу РЕО на всіх його етапах), заснована на використанні єдиного інформаційного простору та дозволяє підвищити достовірність, ефективність й інші показники якості СД [3].

Питання надійності апаратної складової діагностичної системи (як і будь-якої технічної системи) досліджено та описано достатньо розгорнуто. Низка праць представляє ПЗ безвідмовним. У таких роботах оцінюється тільки надійність апаратних засобів (АЗ) [4; 5]. У випадку, коли оцінка надійності складної системи включає в себе обидві складові (програмної й апаратної), загальні значення показників надійності можуть змінюватися суттєво. Навіть якщо програмна складова функціонує відповідно до заданих параметрів, ефективність всієї системи залежатиме від відмов апаратної складової [6]. І навпаки, бездоганно налаштована апаратна складова втрачає свою ефективність у випадку збоїв програмної складової [7].

Наявність різних підходів підтверджує складність оцінки надійності програмно-апаратних засобів та необхідність аналізу конкретної ситуації з урахуванням можливих впливів для отримання об’єктивних результатів.

Під час оцінювання надійності засобів, які функціонують під керівництвом ПЗ, необхідно враховувати особливості функціонування ПЗ на всіх етапах життєвого циклу технічних засобів. Для цього застосовується модель, в якій після процедури оновлення ПЗ надійність зростає стрибкоподібно. Однак навіть у випадку використання такої моделі надійності ПЗ, ймовірність безвідмовної роботи системи буде заниженою, оскільки не враховуються появи відмов і збоїв АЗ та ПЗ. Тому, комплексний аналіз надійності (як складової ефективності) програмно-апаратних засобів має неабияке наукове значення.

При цьому, існуючі варіанти побудови систем контролю мають низку недоліків, які іноді не дозволяють забезпечити високу достовірність та ефективність контролю. Згадані недоліки обумовлені відсутністю єдиного методологічного підходу до створення систем контролю, а також необхідністю інтегрування зразків цивільного виробництва в існуючу систему технічного обслуговування ЗСУ. Основними недоліками сучасних СД слід вважати [8]: вартість, складність, низьку надійність, недостатню завадостійкість, відсутність можливості уніфікації та самонавчання.

У зв'язку з цим доцільно проводити дослідження з можливості спрощення систем контролю, зменшення вартості діагностичного обладнання з одночасним підвищення їхньої надійності та точності вимірювань. Спрощення системи контролю обумовлене необхідністю проведення процедур перевірки (визначення технічного стану РЕО) безпосередньо на місцях експлуатації силами обслуговуючого персоналу (та засобами, які є в наявності). Одним зі шляхів досягнення цієї мети є створення уніфікованих автономних автоматизованих систем контролю (діагностування) РЕО.

Виклад основного матеріалу дослідження. Створення, утримання та розвиток зразків озброєння та військової техніки (далі – ОВТ) супроводжується значними витратами бюджетних коштів в оборонній сфері. При цьому величина витрат на утримання зразків ОВТ (та ефективність використання технічних засобів) не рівномірна на різних стадіях життєвого циклу (далі – ЖЦ). Це обумовлено нерівномірністю в рівнях наукоємності та часових витратах від початку обґрунтування необхідності створення системи до обґрунтованого етапу припинення використання певного зразку озброєння.

Для оцінки стану та готовності складових оборони, країни Північноатлантичного альянсу регулярно проводять огляд результатів виконання заходів 4-річного циклу оборонного планування (NATO Defense Planning Process). Цей цикл включає в себе наступні етапи: аналіз політичних вказівок; визначення потреб; розподіл потреб за цілями; сприйняття реалізації плану; оцінювання отриманих результатів. Загальною метою цього процесу є узгодження національних оборонних програм із цілями та завданнями НАТО щодо досягнення визначеного рівня колективної безпеки та оборони. [9].

Відповідно до нормативних документів Альянсу, ЖЦ зразків ОВТ представляється як еволюція системи, виробу, послуг, проєктів тощо в часі, починаючи від його задуму на створення зразку до його вилучення з обігу (утилізації) [10]. Типовий ЖЦ виробу ОВТ включає в себе декілька стадій і етапів (рис. 1).

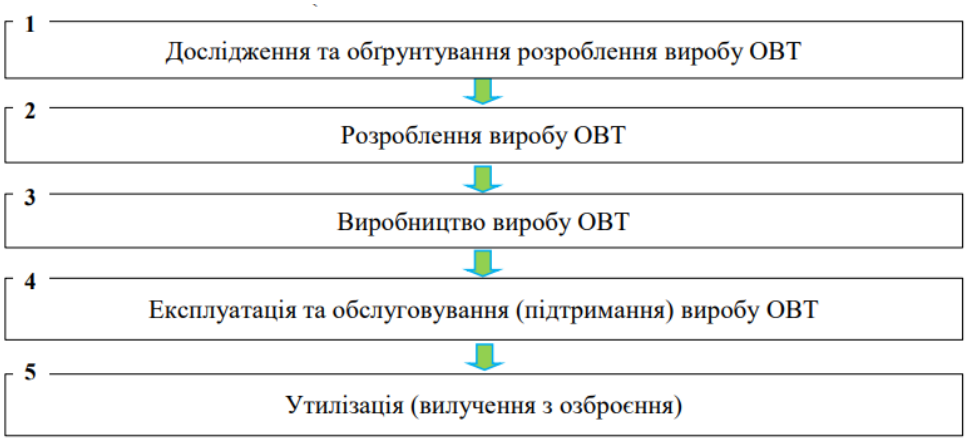


Рис. 1. Типові стадії ЖЦ виробу ОВТ

В нашій державі (на відміну від набору стадій ЖЦ зразків ОВТ, прийнятого в країнах-членах НАТО) відсутня стадія «перед-концепція» (preconception stage) [11]. Цю стадію використовують як сполучення питань планування оборонного бюджету та програм створення

зразків ОВТ. При цьому, слід розуміти, що на утримання, розвиток і створення ОВТ витрачають близько 60 % від загальної сукупності оборонного бюджету.

Як зазначалось раніше, система технічного контролю (діагностування) спрямована на підвищення ефективності використання систем ОВТ, в першу чергу шляхом зменшення витрат на підтримку ЖЦ технічних засобів (від початку етапу експлуатації виробу). Саме на цьому етапі відбувається зниження ефективності зразку, пов’язане з додатковими витратами на планові та позапланові ремонтно-відновлювальні роботи, які спрямовані на підтримання зразка у боєздатному стані (шляхом відновлення його ресурсу до показників, визначених у технологічній документації).

Вбудовані системи контролю не уніфіковані та дуже обмежені за своїми можливостями. Для виробників технічних систем загального призначення створення ефективної системи контролю є додатковим фінансовим навантаженням. В умовах, коли на озброєння надходять системи різних країн (різного рівня технологічної складності), напрямок створення уніфікованої системи контролю та діагностування набуває ще більшої важливості.

Отже, ЖЦ є складний процес, кожний етап якого потребує часових, фінансових, технологічних та інших ресурсних витрат, які впливають на показники ефективності технічного засобу. Час, необхідний для створення нового зразка озброєння, становить приблизно 10–12 років. Враховуючи час, необхідний для налагодження серійного виробництва, цей термін може сягати 15–20 років. Тому термін ефективного використання зразку озброєння також повинен становити не менше 15 років.

В якості критерія ефективності зразків ОВТ може виступати показник (співвідношення) отриманого ефекту з матеріальними витратами. Візуально це можливо представити у вигляді графіку, який пов’язує стадії ЖЦ РЕО (ОВТ) з ефектом/вартістю (рис. 2, який є баченням авторського колективу).

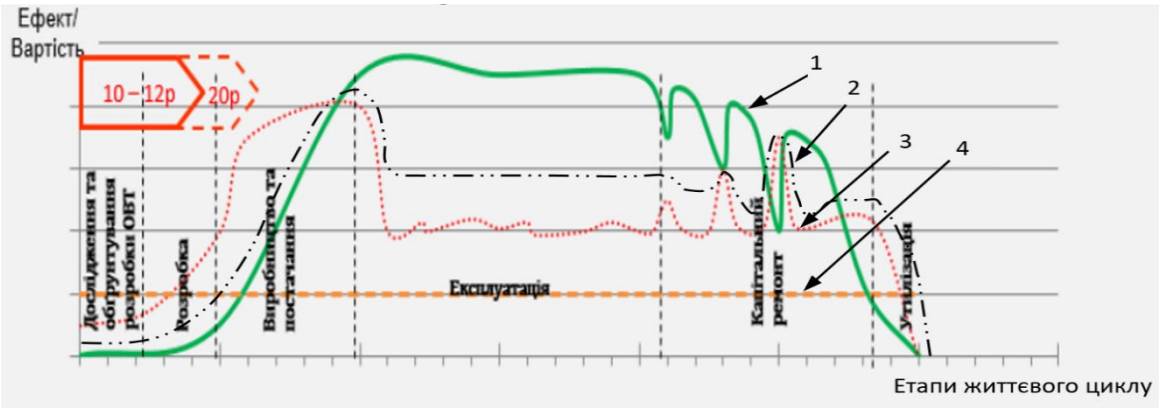


Рис. 2. Залежність ефекту/вартості від стадії ЖЦ виробу ОВТ:
1 – витрати фінансових ресурсів на реалізацію стадій ЖЦ; *2* – ефект від функціонування зразку ОВТ із вбудованою системою контролю (діагностування); *3* – ефект від функціонування зразку ОВТ із зовнішньою системою контролю; *4* – гранично допустимий рівень експлуатаційних витрат

Найбільш тривалий період ЖЦ РЕО пов’язаний саме з етапом експлуатації. Головне завдання системи технічного обслуговування (системи технічного діагностування) полягає в максимальному подовженні цього періоду, бажано завдяки скороченню (уникненню) етапу капітального ремонту. Досягнення цієї мети цілком можливе за принциповими умовами:

перехід на сучасні форми технічного обслуговування (за фактичним технічним станом);

наявність ефективної автономної, уніфікованої системи контролю технічного стану (діагностування) РЕО.

Автономні системи контролю (за зразком інших сучасних технічних систем) також будуються за схемою: апаратна складова – програмна складова в єдиній системі. Ефективність такої системи залежить від надійності обох складових, а відмови можуть бути спричинені наступними причинами: випадкові відмови та дефекти апаратної складової; помилки та відмови ПЗ; збої внаслідок зовнішнього впливу.

Модель поведінки програмно-апаратних засобів зазвичай можливо описати Марковською моделлю, в якій визначаються стани, пов’язані з надійністю ПЗ, та стани, які не залежать від надійності ПЗ. Однак така модель не враховує всі можливі види відмов та помилок ПЗ, їхні наслідки та способи їх виправлення.

Тому для визначення показників надійності ПЗ програмно-апаратних засобів доцільно використовувати моделі на основі підрахунку відмов, які дозволяють враховувати різноманітні можливі сценарії відмов та їхні наслідки.

Апаратні засоби контролю є додатковим обладнанням для здійснення перевірки правильності функціонування ОК (бажано без впливу на функціональність та швидкодію основного устаткування). Вони є більш ефективними порівняно з програмними засобами через постійне вдосконалення апаратної складової та зниження фінансової вартості.

Працездатність складної системи можливо встановити шляхом аналізу результатів виконання спеціальної програми (перевірного тесту), для якої відомий правильний результат. Висновок про технічний стан об’єкта контролю робиться за результатами автоматизованого порівняння результатів контролю з еталонними результатами [12].

Залежно від призначення виділяються налагоджувальні тести (перевірка працездатності окремих елементів програмної системи на початковому етапі наладки), перевірочні тести (періодична перевірка працездатності системи та виявлення несправностей на етапі експлуатації) і діагностичні тести (використовуються для виявлення помилок і локалізації місця несправності).

Доведено, що системи діагностики, які базуються на комп’ютерно-вимірювальних засобах, забезпечують підвищення показників надійності й ефективності функціонування [13].

Оскільки оцінюється ефективність використання автоматизованої системи контролю технічного стану (яка є програмно-апаратним комплексом), оцінку надійності та ефективності слід розглядати окремо для програмної та апаратної складових. Тренд зміни показників надійності ПЗ й апаратної складових істотно різниться залежно від вхідних даних і часу функціонування системи. Інтенсивна обробка даних спричиняє додаткове навантаження на апаратну частину (а надмірне навантаження процесора призводить до зростання споживаної потужності та додаткового нагріву, що може спричинити підвищення частоти збоїв).

Типові криві надійності програмно-апаратних засобів наведено на рис. 3.

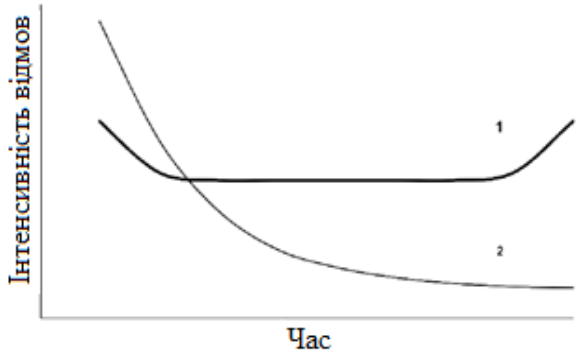


Рис. 3. Залежність інтенсивності відмов від часу експлуатації апаратних (*1*) і програмних (*2*) засобів

Інтенсивність відмов апаратури (крива 1 на рис. 3) змінюється з часом експлуатації. У початковому періоді функціонування, на етапі виявлення та виправлення помилок проєктування та виробничих дефектів, інтенсивність відмов апаратури зменшується з часом. Потім, протягом більшої частини терміну служби, вона залишається стабільною. Наприкінці терміну служби інтенсивність відмов значно збільшується внаслідок зношування апаратури.

Залежність надійності ПЗ (крива 2 на рис. 3) від часу експлуатації має непрямий тренд. Частота виявлення помилок визначається лише вхідними даними. Зменшення інтенсивності відмов ПЗ з часом є результатом виявлення та усунення прихованих помилок у процесі експлуатації.

Оцінку ефективності функціонування програмної складової системи з ВПЗ можна проводити на основі різних моделей. Це може бути модель з дискретно-знижувальною частотою (інтенсивністю) виявлення помилок – аналітичні моделі надійності спеціалізованого ПЗ або моделі з дискретним збільшенням часу до відмови.

Окремо стоять експоненціальні моделі, або моделі по імені авторів (Джелінського – Моранди, Шика – Уолвертона, Липова тощо) [4]. Найбільший інтерес викликає експоненціальна модель, яка добре узгоджується з фізико-хімічними процесами старіння напівпровідникових структур і характером зміни у часі числа помилок у програмі. Такі моделі надають можливість досліджувати закономірності виникнення помилок, а також прогнозувати надійність ПЗ на всіх етапах життєвого циклу РЕО.

У цьому випадку використовуються наступні характеристики надійності:

ймовірність того, що помилки програми не виявляться на інтервалі часу від 0 до t, позначається як функція надійності $P(t)$ і визначає час її безвідмовної роботи протягом часу t;

ймовірність того, що на інтервалі часу від 0 до t виникне відмова програмної складової, позначається як функція ненадійності $Q(t)$:

$$Q(t) = -P(t).$$

Інтенсивність відмов λ(t) і функція надійності $P(t)$ пов'язані між собою виразом [14]:

$$P(t) = exp[-n\lambda(t)dt].$$

Число помилок, що залишилися $m_0(\tau)$ після тестування, можливо визначити виразом:

$$m_0(\tau) = M - m(\tau),$$

де τ – інтервал тестування (час тестування); n – число виявлених відмов протягом певного інтервалу часу; M – кількість помилок у програмі перед фазою тестування; $m(\tau)$ – кінцева кількість виправлених помилок.

Середній час напрацювання до відмови (T_0) (за припущенням, що в процесі коригування нові помилки не з'являються):

$$T_0 = T_0{}' exp\left(\frac{\lambda(\tau)}{MT_0{}'}\right).$$

Інтенсивність відмов ($\lambda(\tau)$):

$$\lambda(\tau) = \frac{1}{T_0{}'},$$

де $T_0{}'$ – вихідне значення середнього часу напрацювання перед тестуванням.

При оцінці ефективності функціонування апаратної складової системи контролю необхідно враховувати надійність та оперативність функціонування цієї системи за показником $\Delta P_З$ (приріст надійності прогнозу завдяки впровадженню підсистеми контролю) [15]:

$$\Delta P_З = P_M - P_{бс},$$

де P_M – ймовірність досягнення мети; $P_{бс}$ – ймовірність досягнення мети без системи контролю.

Приріст оперативності прогнозу при наявності системи контролю ($\Delta T_{ОП}$):

$$\Delta T_{ОП} = M(T_M) - M(T_{бс}),$$

де $M(T_M)$ – математичне очікування часу досягнення мети; $M(T_{бс})$ – математичне очікування часу досягнення мети без системи контролю.

Через відсутність достовірних відомостей про надійність та час проведення робіт у випадку відсутності системи контролю розрахувати ці два показники практично неможливо. Тому, для оцінки ефективності функціонування системи контролю будемо використовувати інші параметри:

$P_{ВН}$ – ймовірність виявлення несправності системою контролю;

ΔP – програш у безвідмовності контрольованого пристрою під час використання вбудованої системи контролю;

ΔD – виграш у достовірності під час використання системи контролю.

Для оцінки ефективності застосування системи контролю можливо використовувати показник «програш у безвідмовності контрольованого пристрою з вбудованою системою контролю».

Значення показника ΔP визначається виразом:

$$\Delta P = |P_{ВИХ}P_K - P_{ВИХ}|,$$

де $P_{ВИХ}$ – ймовірність безвідмовної роботи початкової схеми без системи контролю; P_K – ймовірність безвідмовної роботи початкової схеми за наявності системи контролю.

Імовірність безвідмовної роботи початкової схеми $P_{ВИХ}$:

$$P_{ВИХ} = \frac{1}{1+\frac{\lambda_0}{\theta}},$$

де λ_0 – параметр потоку відмов ОК; θ – інтенсивність відновлення ОК.

Виграш у достовірності при використанні системи контролю:

$$\Delta D = P_{ВИХ} - D, \quad (1)$$

де D – достовірність функціонування ОК у процесі перевірки:

$$D = P_{ВН} + P_{ВИХ} \times P_K - P_{ВИХ}P_K P_{ВН}. \quad (2)$$

Підставивши вираз (2) у вираз (1), отримаємо:

$$\Delta D = P_{ВН} - P_{ВИХ}(1 - P_K) - P_{ВИХ}P_K P_{ВН}.$$

Графіки залежності ΔD та ΔP від кількості діагностичної інформації δ % будуються для різних значень $P_{ВН}$ і ймовірності безвідмовної роботи вихідної системи $P_{ВИХ}$ (рис. 4, 5).

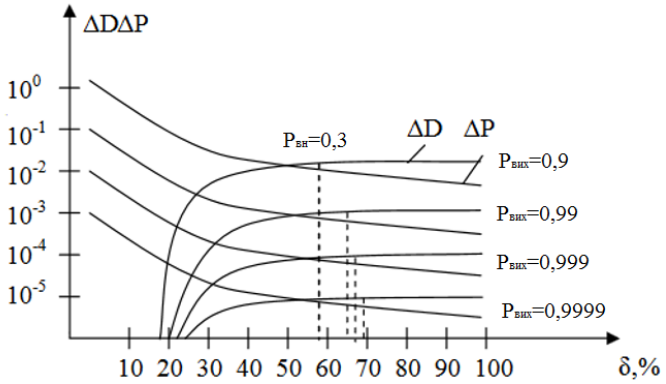


Рис. 4. Залежності ΔP ΔD від δ % при $P_{ВН} = 0{,}3$ та різних значеннях імовірності безвідмовної роботи вихідної схеми $P_{ВИХ}$

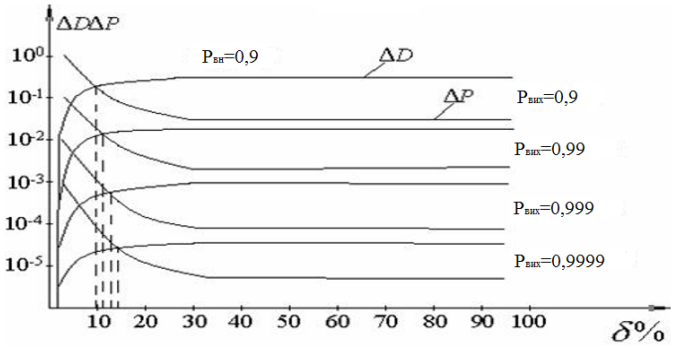


Рис. 5. Залежності ΔP ΔD від δ % при $P_{ВН} = 0{,}9$ та різних значеннях імовірності безвідмовної роботи вихідної схеми $P_{ВИХ}$

Чисельні значення параметру δ реєструються в точках перетину кривих ΔP та ΔD як компроміс між виграшом у достовірності та програшом у безвідмовності системи контролю.

Порівняння показника ймовірності виявлення дефектів для ОК, який містить в собі вбудовану систему контролю, показує можливий приріст у достовірності виявлення дефектів на рівні від 10 % до 30 % залежно від імовірності безвідмовної роботи вихідної схеми.

Збільшення обсягу діагностичної інформації призводить до незначного зростання ймовірності виявлення дефектів. При цьому, ускладнення системи контролю завдяки збільшенню кількості діагностичних показників є економічно неефективним.

Висновки. У перспективі розвитку технологій системи автоматичного контролю повинні стати не лише складовою частиною, а й важливим елементом систем, що мають здатність до самоорганізації. Це означає, що такі системи будуть здатні не лише реагувати на виникнення проблем та несправностей, а й автоматично адаптуватися до змінних умов і вирішувати проблеми безпосередньо на місці, забезпечуючи тим самим безперебійну роботу і високий рівень надійності.

У цьому контексті важливим аспектом є забезпечення систем діагностики технічного стану неруйнівними методами контролю, які дозволяють здійснювати моніторинг стану об'єктів без пошкодження їхньої структури. Крім того, використання принципів штучних нейронних мереж у розробці систем контролю дозволить розширити їхні логічні можливості та надати їм здатність до самонавчання.

Такий підхід до розвитку систем контролю технічного стану відкриває широкі перспективи для покращення їхньої ефективності та надійності. Впровадження інноваційних технологій у цю галузь дозволить забезпечити більш високий рівень безпеки та стабільності функціонування технічних систем у різних галузях промисловості та виробництва.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Воровичб Б. О., Наливайко А. Д., Поляєв А. І. Проблемні питання щодо удосконалення системи оборонного планування в силах оборони України. *Збірник наукових праць Центру воєнно-стратегічних досліджень Національного університету оборони України імені Івана Черняховського*. Київ, 2017. № 2 (60). С. 30–33. DOI: https://doi.org/10.33099/2304-2745/2017-2-60/30-33.

2. Слюсар В. І., Кулагін К. К. Особливості процесу оборонного планування НАТО. *Наука і техніка Повітряних Сил Збройних Сил України*. Харків, 2019. № 3 (36). С. 47–59. DOI: https://doi.org/10.30748/ nitps.2019.36.06.

3. Дєнєжкін М. М., Крикун П. М., Руснак І. С. Проблеми проведення комплексного огляду сектору безпеки та оборони України: погляди на його організацію та вирішення завдань. *Наука і оборона*. Дніпропетровськ. 2014. № 4. С. 3–10.

4. Koren I., Mani C. Krishna Fault-Tolerant Systems. Elsevier Science, 2010. 400 p.

5. Rausand M., Barros A., Hoyland A. System Reliability Theory: Models, Statistical Methods, and Applications. John Wiley & Sons, 2021. 864 p.

6. Yusupbekov N. R., Gulyamov Sh. M., Mirzaev D. A., Kuziyev Z. J. Analysis of the current status of the theory and practice of assessing the reliability of software of automated information and control systems. *Galaxy international interdisciplinary research journal*. 2022. Vol. 10, Issue 10. P. 418–4025.

7. Pham H. Handbook of Reliability Engineering. London:Springer, 2003. 696 p.

8. Про затвердження Порядку організації та здійснення оборонного планування в Міністерстві оборони України, Збройних Силах України та інших складових сил оборони: Закон України від 22.12.2020 № 484. URL: https://zakon.rada.gov.ua/laws/show/z0196-21#Text (дата звернення: 20.02.2024).

9. Наливайко А. Д., Поляєв А. І., Сівоха І. М. Генезис та розвиток оборонного планування в Україні. *Сучасні інформаційні технології у сфері безпеки та оборони*. Київ, 2017. № 2 (29). С. 138–143. DOI: https://doi.org/10.33099/2311-7249/2017-29-2-138-143.

10. ДСТУ В 15.004:2022 Система розроблення і поставлення на виробництво озброєння та військової техніки. Стадії життєвого циклу озброєння та військової техніки.

11. AAP-20:2015 NATO Programme Management Framework (NATO Life Cycle Model).

12. Sipser M. Introduction to the Theory of Computation. PWS, 2012. 452 p.

13. Кузавков В. В., Хусаінов П. В. Прогнозування технічного стану однотипних програмно-апаратних засобів. *Інформатика та математичні методи в моделюванні*. 2018. № 1 С. 57–68.

14. Василевський О. М., Поджаренко В. О. Нормування показників надійності технічних засобів. Вінниця: ВНТУ, 2010. 130 с.

15. Кузавков В. В., Янковський О. Г., Болотюк Ю. В. Обґрунтування вибору показників оцінки ефективності функціонування автоматизованої системи контролю. *Сучасні інформаційні технології у сфері безпеки та оборони.* Київ. 2022. Том 2 (44). С. 21–27. DOI: 10.33099/2311-7249/2022-44-2-21-27.

DOI: 10.58254/viti.5.2024.08.87

УДК 621.39

д-р техн. наук, професор Міночкін А. І. ORCID: 0000-0001-5723-5052 (ВІТІ ім. Героїв Крут)
канд. техн. наук Масесов М. О. ORCID: 0000-0003-4537-4295 (ВІТІ ім. Героїв Крут)
канд. техн. наук Шаціло П. В. ORCID: 0009-0008-7932-2294 (ВІТІ ім. Героїв Крут)

ПРИКЛАДНЕ МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ УПРАВЛІННЯ ВИТРАТАМИ ЕНЕРГОРЕСУРСУ ВУЗЛІВ СЕНСОРНОЇ МЕРЕЖІ ВІЙСЬКОВОГО (СПЕЦІАЛЬНОГО) ПРИЗНАЧЕННЯ

Аналіз досвіду проведення військових операцій у війні російської федерації проти України засвідчив, що використання у Збройних силах передових технологій дозволяє своєчасно викрити нарощування військ ворога на кордонах між державами, які беруть участь у військовому конфлікті, виявити опорні пункти противника, передислокацію військ, їхню підготовку до нанесення ракетних ударів та ін. Саме впровадження таких технологій в автоматизованих системах управління військами і зброєю дозволяє своєчасно реагувати на вплив різноманітних зовнішніх факторів та дає змогу командуванню і військам на полі бою швидко пристосовуватись до мінливої ситуації на всьому просторі проведення бойових дій.

До таких передових технологій військові науковці армій передових країн світу відносять технології безпроводових сенсорних мереж. Такі характеристики безпроводових сенсорних мереж, як: 1 – здатність до швидкого розгортання; 2 – самоорганізація; 3 – висока відмовостійкість; 4 – здатність застосування в різних середовищах (повітряні, надземні; підземні; морські), – роблять їх надійним засобом забезпечення ефективного управління виконанням військами бойових завдань.

У статті розглядаються результати дослідження архітектури сенсорної мережі спеціального (військового) призначення, її основних характеристик, принципів функціонування та методів управління енергоспоживанням вузла мережі (як базової компоненти прикладного математичного забезпечення управління витратами енергоресурсу) з точки зору критерію найменшого енергоспоживання. Для впровадження методів управління витратами енергоресурсу в систему оперативного управління сенсорною мережею та розробки методики узгодження дій її компонентів для вирішення задачі енергозбереження розроблена її функціональна модель на платформі методології IDEF (Integrated Definition – цілісна точність).

Ключові слова: *алгоритм, критерій, метод, модель, метрика, сенсор, сенсорний вузол, сенсорна мережа.*

A. Minochkin, M. Masesov, P. Shatsilo Applied mathematical providing of management of energy of resource of knots of sensory network for the (special) military setting charges

The analysis of experience of realization of soldiery operations in war of the Russian federation against Ukraine witnessed, that the use in the armed forces of frontrank technologies the increase of troops of enemy allows in good time to disrobe on borders between the states, that take part in a military conflict, to reduce the strong points of opponent, redeployment of troops, their preparation to the rocket inflicting blows and other. Self introduction of such technologies in CASS of management troops and weapon allows in good time to react on influence of various external factors and gives an opportunity to the command and troops on a battlefield quickly to adapt to the changeable situation on all space spacious realizations of battle actions.

To such frontrank technologies the soldiery scientists of armies of frontrank countries of the world take technologies of wireless sensory networks (WSN). Such descriptions of WSN as: 1 – is a capacity for rapid development; 2 – is independently organization; 3 – is high fault tolerance; 4 – is ability of application in different environment.

The results of research of architecture of sensory network of the special (military) setting are examined in the article, her basic descriptions, principles of functioning and methods of management of knot of network (as a base component of the applied mathematical providing of management the charges of resource) an energy consumption from the point of view of criterion of the least energy consumption. For introduction of methods of management of charges in the system of operative management a sensory network worked out her functional model on the platform of methodology of IDEF (Integrated Definition is integral exactness).

Keywords: *algorithm, criterion, method, model, birth-certificate, touch control, sensory knot, sensory.*

Актуальність дослідження

Рішенням Ради національної безпеки і оборони України від 20.08.2021 року № 0063525-21 введено в дію стратегічний оборонний бюлетень, в якому одним із напрямків реалізації оборонної політики України визначена «…побудова системи об'єднаного керівництва силами оборони та військового управління у Збройних Силах України, яка має

здійснюватися відповідно до передового досвіду, принципів і стандартів держав-членів НАТО».

З метою впровадження цього напрямку інноваційному розвитку в Збройних силах України (ЗСУ) підлягають, в першу чергу: принципи, форми, методи і технології застосування сил і засобів отримання інформації про стан військ і бойових систем противника, а також процеси добування, обробки, аналізу та доведення інформації про стан об'єктів противника до органів військового управління й компонентів озброєння та військової техніки.

Ці форми, методи, процеси і технології реалізуються та здійснюються в автоматизованій системі **C4ISR** (*англ.* **C**ommand and **C**ontrol, **C**ommunications, **C**omputers, **I**ntelligence, **S**urveillance, Reconnaissance – командування та управління, зв'язок, комп'ютери, розвідка).

У цьому напрямку реалізації оборонної політики України на сьогодні в закладах військової освіти і науки ЗСУ активно проводяться наукові дослідження з метою розробки всіх видів забезпечення мобільних сенсорних мереж військового (спеціального) призначення, які реалізують функції збору, обробки і передачі інформації про стан елементів військ і бойових систем противника, тим самим забезпечуючи вирішення задач в автоматизованій системі C4ISR.

Враховуючи зазначене, обраний нами напрям наукового дослідження є достатньо актуальним науковим завданням, виконання якого сприятиме інноваційному розвитку сенсорних мереж військового (спеціального) призначення.

Аналіз останніх публікацій

[1]. Описано принципи, особливості побудови і функціонування мобільних радіомереж типу Ad Hoc і MANET. Викладено теоретичні основи і приклади вирішення задач управління радіомережами на канальному, мережевому і прикладному рівнях моделі взаємодії відкритих систем OSI (The **Open Systems Interconnection** model).

[11]. Запропоновано метод збільшення «часу життя» безпроводової сенсорної мережі (БСМ) з надлишковою кількістю вузлів, який під час реалізації маршрутизації враховує енергію вузлів і надійність безпроводових з'єднань.

[10]. Наукова стаття присвячена вирішенню задачі розробки нових методів енергозбереження та методик їх застосування в сенсорних радіомережах спеціального призначення, які повинні забезпечувати зберігання енергетичних ресурсів вузлів за умови заданої якості передачі інформаційних потоків та адаптуватись до умов функціонування мережі.

[12]. У науковій статті розглянута задача розробки нових методів побудови БСМ, які дозволять дистанційно та оперативно здійснювати моніторинг стану природного середовища заданого регіону та оцінювати загрози й ризики виникнення надзвичайних ситуацій.

[13]. Запропоновано новий адаптивний енергозберігаючий алгоритм передачі даних при прийнятті рішень про стан сенсорної мережі на основі показань із сенсорів.

Мета статті: забезпечити ефективність функціонування підсистеми управління енергоресурсом вузлів сенсорної мережі військового (спеціального) призначення шляхом інновацій її прикладного математичного забезпечення (ПМЗ) (методів, моделей, методик і алгоритмів).

Наукове завдання статті: удосконалення схеми вирішення задач оперативного управління сенсорною мережею військового (спеціального) призначення в частині «*методи – моделі прикладного математичного забезпечення*» для заощадження енергоресурсу вузла мережі військового (спеціального) призначення.

Основна частина

Професор електромеханіки Крістофер Пістер (Каліфорнія, США) розробив концепцію **«smart dust»** (*англ.* smart dust – розумний пил). Згідно з цією концепцією він запропонував застосування систем, які складаються з довільної множини електромеханічних (електронних)

сенсорів (різного функціонального призначення), які здатні обмінюватись інформацією в будь-якому середовищі, де здійснюється моніторинг об'єктів, за якими здійснюється спостереження.

Проєкція концепції **«smart dust»** на такі середовища, як театр проведення військових операцій, морські комунікації, частини і підрозділи збройних сил, згенерувала появу БСМ, які складаються з множини сенсорних вузлів (СВ) з інтегрованими функціями:

– моніторингу навколишнього середовища;
– обробки даних;
– передачі даних та ін.

Нині БСМ називають просто сенсорними мережами (СМ). На рисунку 1 представлено узагальнену структуру СМ.

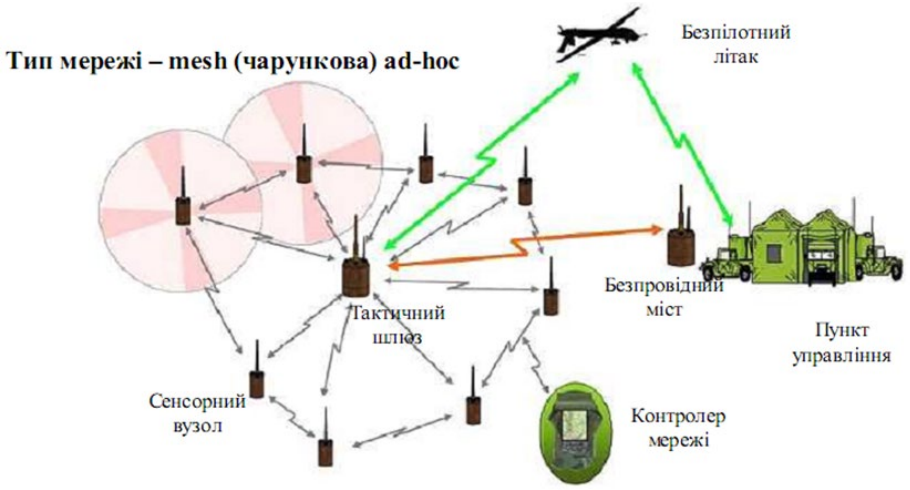


Рис. 1. Узагальнена структура СМ

Базова архітектура СМ і її основні інтерфейси були визначені науково-дослідницькою групою SGSN (Study Group on Sensor Networks) технічного комітету № 1 ISO/IEC JTC1, що дозволило виділити:

1. Види забезпечення СВ СМ:
– апаратне забезпечення;
– базове (системне) програмне забезпечення (комплекс програм, що забезпечують інфраструктуру СВ, на якій можуть працювати прикладні програми).

Головним компонентом базового програмного забезпечення є операційна система (комплекс взаємопов'язаних програм, призначених для управління ресурсами СВ та організації взаємодії з прикладними програмами, з апаратним обладнанням усіх підсистем СВ та користувачами):

– ***прикладне математичне забезпечення*** (реалізоване в прикладному програмному забезпеченні), яке являє собою сукупність сучасних математичних методів, моделей, алгоритмів, що забезпечують розв'язання функціональних задач у системі управління СВ;
– прикладне програмне забезпечення.

2. Базові інтерфейси СВ:
– програмний інтерфейс (інтерфейс між базовим і прикладним програмним забезпеченням СВ);
– інтерфейс користувача;
– інтерфейси між вузлами СМ;

– інтерфейс між СМ і зовнішнім середовищем (користувачі – органи військового управління, інформаційно-аналітична система ЗСУ).

Сенсорний вузол (*англ*. Sensornode) – це пристрій, який складається з одного або декількох взаємодіючих між собою сенсорів (рис. 2).

Основні вимоги до СВ:

а) забезпечення роботи у СМ;

б) ***низьке енергоспоживання***;

в) реалізація функції виділення й обробки сигналів на вході і виході вузла;

г) можливість перетворення сигналів сенсорів (таких як сенсори фізичних, оптичних, електрохімічних і т. ін.) у зручну форму для подальшої обробки і передачі даних.

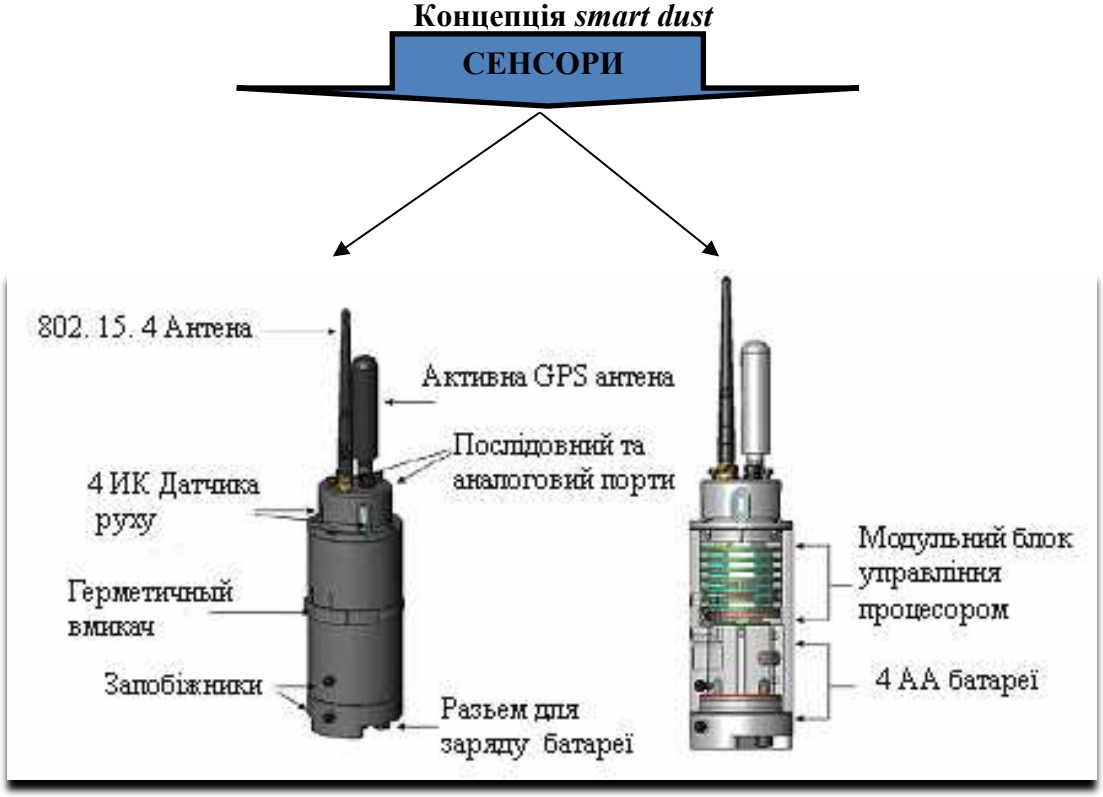


Рис. 2. Сенсори у складі вузла СМ

Групою SGSN визначено, що СВ (як система) складається з таких підсистем:

А) *обчислювальна підсистема* – здійснює обробку даних і забезпечує виконання всіх функцій СВ. Головним елементом цієї підсистеми є мікроконтролер MCU (*англ*. Micro Controller Unit).

До складу мікроконтролера ***MCU*** входять:

– процесор;
– оперативна пам’ять;
– незалежна пам’ять;
– флеш-пам’ять;
– аналого-цифровий перетворювач (ADC);
– таймер;
– порти введення-виведення даних;

Б) *комунікаційна підсистема* – система для управління процесами зв’язку, взаємодії та обміну даними між СВ. Центральним компонентом комунікаційної підсистеми є радіо – трансивер (прийомопередавач);

В) *сенсорна підсистема* – складається із сенсорів, які являють собою датчики (первинні перетворювачі), які сприймають об'єкт контролю (танк, літак, ракета, підводний човен, гармата й ін.). Одночасно сенсор є пристроєм у вигляді конструктивної сукупності одного або декількох вимірювальних перетворювачів величини, яку вимірюють і контролюють, та котрий виробляє вихідний сигнал, зручний для дистанційного передавання, зберігання та використання в системі **C4ISR.**

Г) ***підсистема електроживлення*** – забезпечує енергетичне постачання всіх елементів СВ.

Основною метою застосування СМ у ЗСУ є контроль і моніторинг вимірюваних параметрів об'єктів противника.

Процес збору даних і управління в СВ представлено на рисунку 3.

На сьогодні СМ знаходяться на початковому етапі впровадження в сектор безпеки і оборони України, включно з моніторингом навколишнього середовища, контролем руху військових об'єктів у просторі та ін.

Розглянуті характеристики СМ (відповідно до стандарту IEEE 802.15.4) дозволяють стверджувати, що технологія СМ є:

1 – стійкою у налагодженні;

2 – має високі експлуатаційні параметри.

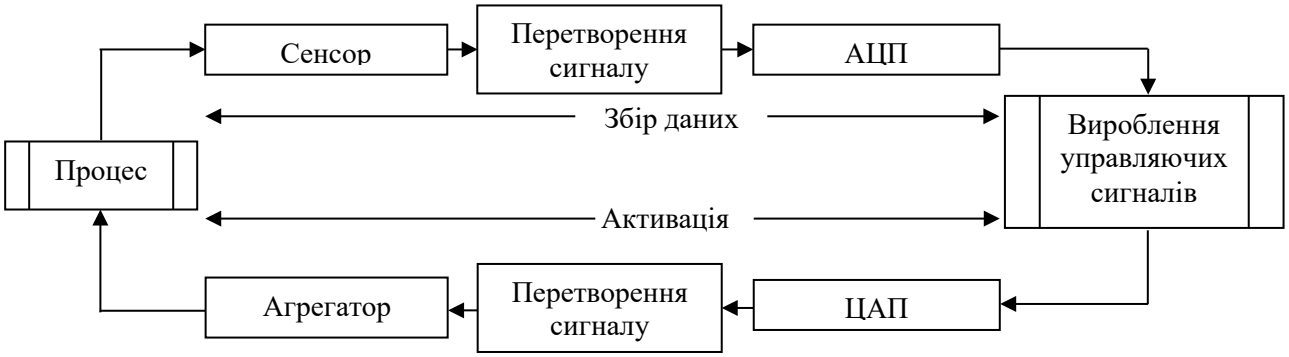


Рис. 3. Збір даних і управління в СВ

Згідно зі стандартом IEEE 802.15.4:

а) в СМ повинна бути узгоджена дальність передачі/прийому інформації;

б) в СМ повинна бути реалізована процедура отримання підтвердження про успішну доставку повідомлень;

в) *енергоспоживання повинне бути якомога меншим* завдяки низькій швидкості передачі даних і відповідати вимогам до «часу життя» СВ.

У стандарті 802.15.4 затверджена робота СМ двох рівнів моделі OSI взаємодії відкритих систем (*фізичний і канальний* рівні).

Загальна класифікація сенсорних мереж спеціального призначення (СМСП):

а) за мобільністю: стаціонарні, рухомі (роботи), гібридні;

б) за середовищем моніторингу: надземні, підземні, повітряні, морські;

в) за організацією: децентралізовані, ієрархічні, гібридні;

г) за видом моніторингу: акустичні, оптичні, сейсмічні тощо.

Враховуючи основні відмінності та особливості системи управління СМСП та основні вимоги до неї, визначено основний принцип її функціонування – **принцип функціональності управління.**

Сутність *принципу функціональності управління* полягає в об'єднанні функцій системи управління СМСП у відносно незалежні групи.

Впровадження методу аналізу в досліджені СМСП та цього принципу дозволяє здійснити декомпозицію управління мережею на підсистеми (що значно спрощує задачу розробки ПМЗ) та виділити **підсистему управління енергоресурсом вузлів СМСП.**

Підсистеми управління:

– маршрутизацією;
– топологією;
– навантаженням;
– безпекою;
– **енергоресурсом вузлів СМСП;**
– радіоресурсом;
– якістю обслуговування Qos;
– збору та зберігання інформації;
– та підсистеми реалізації рішень і моніторингу.

Метою функціонування *підсистеми управління енергоресурсом вузлів СМСП* є мінімізація витрат енергоресурсу СВ та забезпечення вимог до «часу життя» мережі.

Необхідно зазначити, що функціональні підсистеми $СВ_i$ витрачають різну кількість енергії батарей. Тому управління витратами енергоресурсу вузлів СМСП може бути реалізовано:

1) за різними функціями управління і рівнями еталонної моделі OSI взаємодії відкритих систем (фізичному, канальному, мережному);

2) методами збереження витрат електроенергії комунікаційною підсистемою вузла СМСП (управління потужністю прийому-передачі);

3) методами збереження енергії батарей електроживлення СВ;

4) методами мінімізації споживання електроживлення підсистемами (сенсорної, обчислювальної, комутаційної) СВ СМСП.

Значна увага в наукових виданнях [2; 3; 5; 10] приділена методам збереження витрат електроенергії комунікаційною підсистемою вузла СМСП:

– між сусідніми вузлами СМСП (фізичний рівень еталонної моделі);
- при управлінні топологією (канальний і мережевий рівні еталонної моделі);
- при маршрутизації (мережевий рівень еталонної моделі);
- при застосуванні технологій спрямованих антен чи позиціювання (канальний і мережевий рівень еталонної моделі);
- тимчасове виключення прийомопередавачів – трансиверів (між сусідніми вузлами, в масштабах СМСП або її зони);
- зменшення добових передач;
- застосування енергозалежних метрик вибору маршрутів;
- зменшення службового трафіку;
- застосування R-зонової маршрутизації.

Менше уваги науковці приділяють методам збереження витрат електроенергії обчислювальною та сенсорною підсистемою вузла СМСП, а також методам узгодження дій окремих підсистем СВ СМСП при вирішенні задачі комплексної оптимізації витрат енергоресурсу СВ.

ПМЗ управління витратами енергоресурсу вузла СМСП складається із методів, алгоритмів та моделей функціонування **підсистеми управління енергоресурсом вузла СМСП.**

Сутність процесу оперативного управління СМСП полягає в забезпеченні підтримки значень показників її функціонування на рівні заданих шляхом доведення до її елементів управляючих впливів, які є результатом рішень задач управління.

Узагальнена схема вирішення задач оперативного управління СМСП представлена на рисунку 4.

Розв'язанню задачі Z_{ijk}^{r} (де $i=\overline{1,I}$ – рівень СМСП; $j=\overline{1,J}$ – функціональна підсистема СВ СМСП; r – рівень оперативного управління (організаційного або технологічного); $k=\overline{1,K}$ – рівень моделі OSI) передує орієнтація (визначення) щодо рівнів оперативного управління СМСП, рівня моделі OSI і функціональної підсистеми СВ СМСП, а також визначення початкових даних для розв'язання задачі на базі моделі організації інформаційних ресурсів.

Вибір методу, алгоритму розв'язання задачі, моделі елемента застосування задачі і оптимізації її розв'язання здійснюється на основі використання компонентів ПМЗ.

Для виконання свого цільового призначення ПМЗ оперативного управління СМСП повинно характеризуватися властивостями, які визначаються:

1) системою управління СМСП в цілому;
2) функціональними підсистемами СВ СМСП;
3) базовими положеннями з побудови ПМЗ.

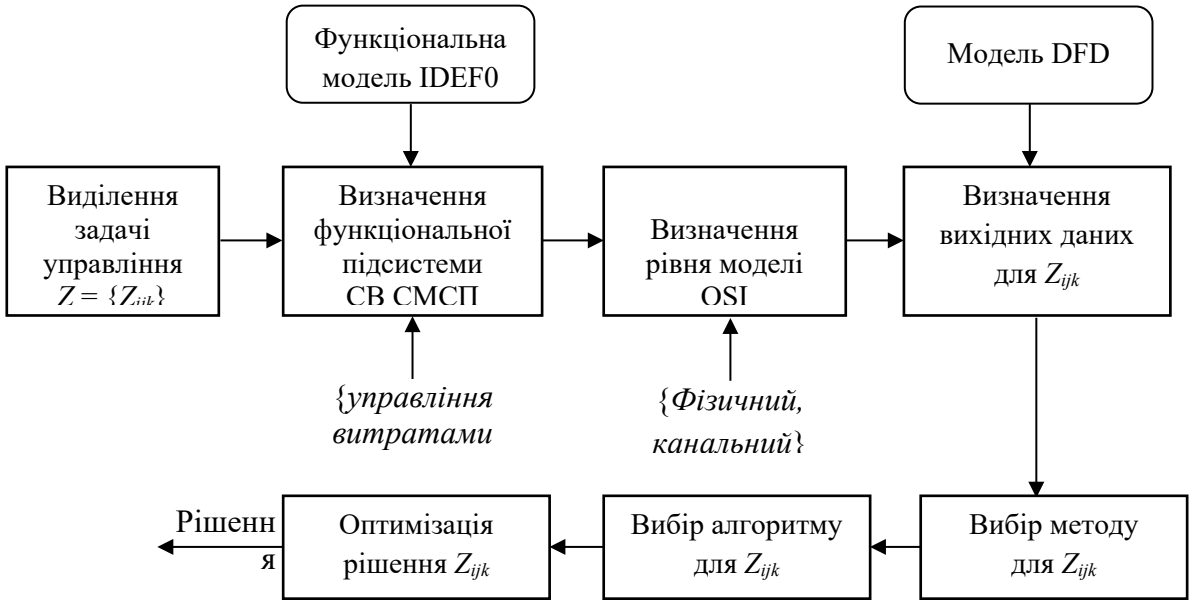


Рис. 4. Узагальнена схема вирішення задач управління СМСП

Перша група властивостей визначає здатність ПМЗ задовольнити потреби СВ СМСП у виконанні свого функціонального призначення. До цієї групи властивостей віднесемо забезпечення обґрунтованості управління і забезпечення оперативності управління (табл. 1).

До другої групи відносяться властивості, які проявляються у взаємодії з функціональними підсистемами СВ СМСП і які окреслюють умови їхнього функціонування. Такою властивістю є якість ПМЗ по k -й підсистемі.

Отже, за головну властивість ПМЗ управління витратами енергоресурсу СВ можна прийняти функціональну якість ПМЗ.

Ця властивість може бути оцінена узагальненим показником – коефіцієнтом якості ПМЗ для *k*-ї підсистеми:

$$K_{\text{їіç}}^{k} = F\left(V_{\text{їіç}}^{k}, C_{\text{їіç}}^{k}, R_{\text{їіç}}^{k}\right),$$

де $V_{\text{їіç}}^{k}$ – показник повноти ПМЗ в k -й функціональній підсистемі СВ МК;

Таблиця 1

Властивості ПМЗ і показники для їхньої оцінки

Властивості ПМЗ		*Показники оцінки*
Група	*Назва*	
1	Забезпечення обґрунтованості управління	Коефіцієнт ступеня забезпечення відповідності сформованого рішення меті операції $K_{\text{а}}(t) = \sum_{i \in I^x} \left[1 - \left(\left\|x_i^{\text{р}}(t) - x_i^{\text{і}}(t)\right\| / \left\|\max\left(x_i^{\text{р}}(t), x_i^{\text{і}}(t)\right)\right\|\right)\right] / \|I\|$
	Забезпечення оперативності управління	Час формування рішення $T_i = T_{\text{і}} = f(\dot{A})$
2	Функціональна якість ПМЗ	Узагальнений показник якості ПМЗ для k-ї підсистеми: $K_{\text{ЇЇÇ}}^k = F\left(V_{\text{ЇЇÇ}}^k, C_{\text{ЇЇÇ}}^k, R_{\text{ЇЇÇ}}^k\right)$, де $V_{\text{ЇЇÇ}}^k = \frac{W_{z^*}^k}{W_{z^*}^k}$ – показник повноти ПМЗ; $\left\|\overline{a} - a\right\| \le C_{\text{ЇЇÇ}}^k$ – показник точності ПМЗ; $R_{\text{ЇЇÇ}}^k = P\left(\left\|\overline{a} - a\right\| \le C_{\text{ЇЇÇ}}^k\right)$ – показник достовірності ПМЗ
3	Модульність	Коефіцієнт сумісності компонентів ПМЗ $K_c = f\left(A = \left\{A^k\right\}, \varphi = \{\varphi_i\}, Z = \left\{Z_{ijk}\right\}\right)$
	Управляємість	Ймовірність своєчасного створення діючої конфігурації ПМЗ $P_{\text{СК}} \ge P_{\text{СК}}^{\text{Н}}$
	Реалізованість	Коефіцієнт реалізованості $K_{\text{р}} = f\left(K_{1,} K_2, K_3, K_4, K_5\right)$
	Інформаційна визначеність	Коефіцієнт невизначеності вибору необхідних даних для розв'язання задачі Z_{ijk}. $E_{\text{В}} = f(J, J_{\text{Н}}) \le 0{,}5$

$V_{\text{ЇЇÇ}}^k = \frac{W_{z*}^k}{W_z^k}$, W_{z*}^k – кількість задач k-ї підсистеми СВ, яка забезпечена засобами ПМЗ;

W_z^k – загальна кількість задач k-ї підсистеми СВ СМСП;

$C_{\text{ЇЇÇ}}^k$ – показник точності ПМЗ в k-й функціональній підсистемі СВ СМСП, який може бути визначений із виразу $\left\|\overline{a} - a\right\| \le C_{\text{ЇЇÇ}}^k$ (a – передбачуваний результат забезпечення роботи k-ї функціональної підсистеми, $\overline{a}$ – статистична оцінка цього результату);

$R_{\text{ЇЇÇ}}^k$ – показник достовірності ПМЗ в k-й функціональній підсистемі СВ СМСП, який може бути визначений як $R_{\text{ЇЇÇ}}^k = P\left(\left\|\overline{a} - a\right\| \le \tilde{N}_{\text{ЇЇÇ}}^k\right)$.

Проведемо аналіз методів управління витратами енергоресурсу в СМ:

Методи *зменшення потужностей передач між сусідніми вузлами СМСП зі збереженням необхідних параметрів радіоканалу, таких як*:

- необхідного рівня відношення сигнал/завада, нормованого на біт SINR. Мінімальне значення SINR (*англ.* Signal to Interference + Noise Ratio) для стабільної роботи в СМСП становить 10 дб;
- коефіцієнта бітових помилок BER (*англ. Bit Error Ratio*) й ін.

Встановимо такі початкові дані:

а) СМСП являє собою N випадково розташованих СВ;

б) позначимо максимальну потужність передавача і-го СВ через P_i^{max}.

Тоді для успішної доставки повідомлення від СВ$_i$ до СВ$_j$ необхідно, щоб представлення рівня відношення сигнал/завада, нормованого на біт SINR, було більше певного порогу β, що визначає коефіцієнт бітових помилок BER (1):

$$\text{SINR} = \frac{P_{ij} G_{ij}}{\sum\limits_{k \neq j} P_{kj} G_{kj} + \eta_j} \geq \beta, \tag{1}$$

де $i, j, k \in$ N;

P_{ij} – потужність передачі трансивера i-го СВ;

G_{ij} (величина загасання сигналу між вузлами ij) визначається як $G_{ij} = kaFdij$ – α (ka – коефіцієнт, що характеризує антени, що використовуються); F – коефіцієнт ослаблення сигналу антени; dij – відстань між СВ; α – ступінь втрати потужності (яка задається $\alpha = 2 \div 4$);

ηj – спектральна потужність шуму приймача трансивера j-го СВ.

У випадку коли рівень сигналу передавача трансивера СВ перевищує заданий поріг, здійснюється зменшення потужності передачі (і відповідно зменшуються витрати енергоресурсу СВ).

Якщо час очікування повідомлення для протоколу канального рівня перевищено, то передача здійснюється з більшою потужністю передавача P_{ij} (і відповідно збільшуються витрати електроенергії).

Управління потужністю передачі трансивера СВ в каналі (ij) здійснюється за наступним ітераційним правилом (2):

$$P_{ij}(n+1) = \min(P^{\max}, \frac{\beta}{\text{SINR}_{ij}(n)} P_{ij}(n)), \tag{2}$$

де n – номер ітерації.

Результати моделювання цього методу показали, що регулювання потужності передачі трансивера СВ дозволяє зменшити витрату енергії на 10–20 % при збільшенні пропускної здатності до 15 % [10].

Методи *управління топологією* СМСП передбачають перерозподіл потужностей передач трансиверів СВ P_i та/або зміну спрямованості їхніх антен [6].

Збільшення потужності передачі трансивера СВ призводить до більшої витрати енергії батарей.

Зменшення потужності передачі трансивера СВ дозволяє зменшити витрати енергії батарей СВ, але призводить до збільшення часу доставки повідомлення та збільшення обсягу службового трафіка.

У наукових виданнях [7; 8] описано методи децентралізованого управління топологією СМ:

1) Метод *BSP* (*англ. BPS* – Planar Power Efficient Structure) використовує систему позиціонування, будує планарний граф (**Планарний граф** – граф, який може бути зображений на площині без перетину ребер), що має мінімальну сумарну потужність передачі по всіх маршрутах СМ (що призводить до зменшення витрати енергоресурсу СВ).

2) Метод *COMROW* передбачає оптимізацію загальної потужності передач в СМ за допомогою зміни її зв'язності (менша потужність передачі – менша зв'язність).

Методи *застосування спрямованих антен і системи позиціонування.* Ці методи дозволяють знизити рівень взаємних перешкод і збільшити енергетику радіоканалів.

Прямі методи збереження енергії батареї

1-й метод. *Тимчасове відключення прийомопередавача трансивера СВ* (пасивний режим або режим «сон»).

Метод «сон» вже стандартизовано. Він припускає централізоване управління відключенням СВ і тому має обмеження для застосування в СМСП.

Значно більший спектр застосування отримав децентралізований варіант цього методу *PAMAS* (Power Aware Multiple-Access Protocol) [8], який реалізує пасивний режим між сусідніми вузлами.

2-й метод. *Зменшення числа повторних передач, які викликані зіткненнями пакетів* (тобто методи канального доступу, що забезпечують мінімум зіткнень пакетів, а тому зменшують витрати енергії).

3-й метод. *Вибір маршрутів за метриками*, що враховує витрату енергоресурсу СВ при передачі/прийому повідомлень і/або ємність батарей [6].

При застосуванні цього методу використовуються такі типи метрик:

а) *метрика – потужність,* що витрачається на передачу/прийом вузла;

б) *метрика – ємність батарей вузлів*;

в) *багатопараметрична метрика* – потужність передачі вузлів, ємність батарей та ін.

Для впровадження методів управління витратами енергоресурсу в систему оперативного управління СМСП пропонується розробка її функціональної моделі на платформі методології IDEF (Integrated Definition – цілісна точність), яка стала стандартом в США та ряді країн Європи.

На рисунку 5 представлена контекстна діаграма оперативного управління СМСП, а на рисунку 6 – діаграма підсистеми управлінням енергоресурсом СМСП, які розроблені за допомогою програмного забезпечення IDEF – методології AllFusion Process Modeler r.7 (Platinum BPWin).

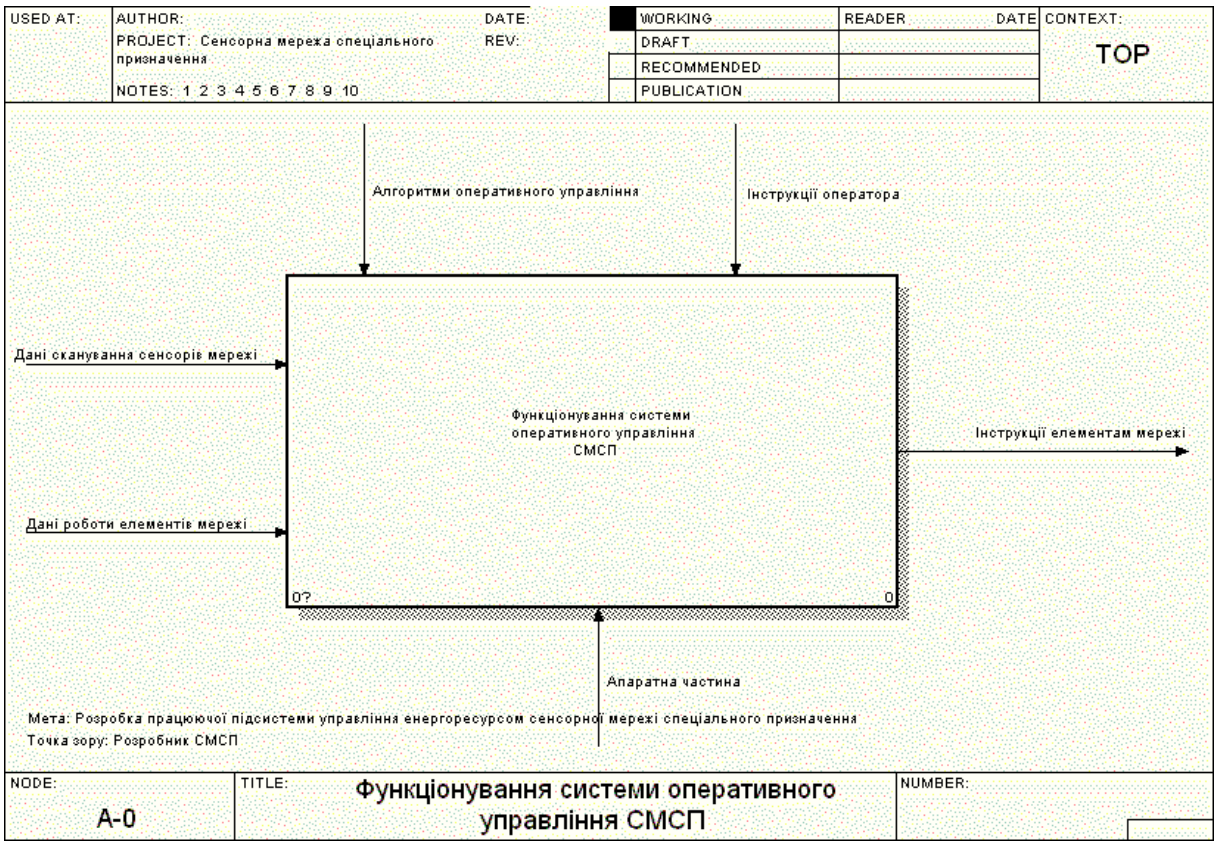


Рис. 5. Контекстна діаграма оперативного управління СМСП

Діаграма А15 розкриває складові підсистеми управління енергоресурсом, поділяючи основні функції за трьома процесами:

– управління міжсистемною взаємодією та формування рішення;
– управління потужністю передачі;
– збереження енергії батареї.

Після прийому йде опрацювання даних з урахуванням алгоритмів оперативного управління енергоресурсом та формування рішення підсистеми управління енергоспоживанням. Опрацьовані дані отримуються у підсистемах управління потужністю передачі та збереження енергії батареї. За допомогою них апаратно змінюються витрати заряду батареї відповідно до вирахуваних раніше показників для збереження поточної ємності батареї або перехід сенсора у режим сну для збереження роботоздатності елемента мережі.

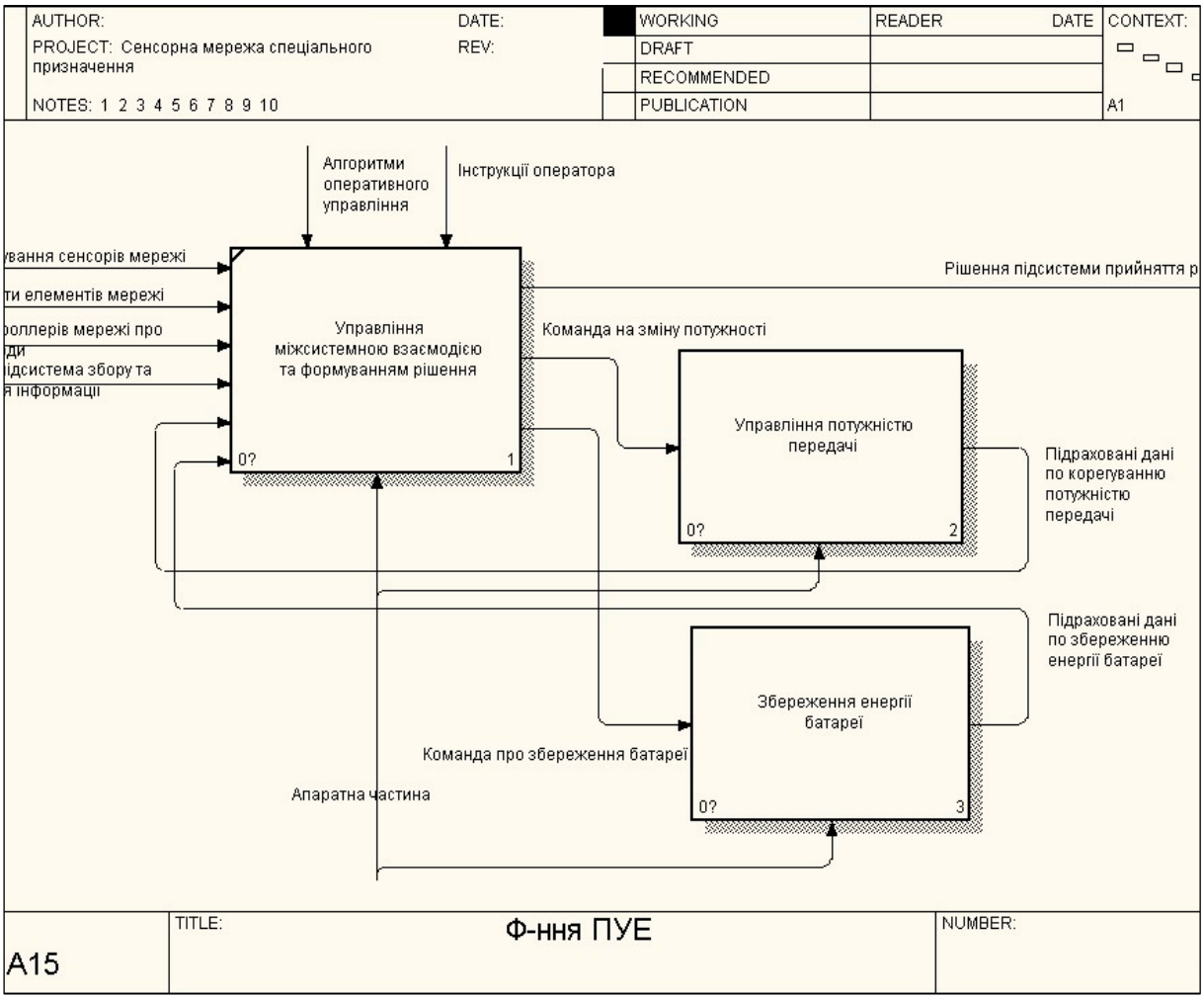


Рис. 6. Діаграма підсистеми управлінням енергоресурсом СМСП

Підсистема управління потужністю передачі використовує метод зменшення потужностей передач між сусідніми вузлами зі збереженням необхідних параметрів радіоканалу, а підсистема збереження енергії батареї застосовує метод тимчасового відключення прийомопередавача (трансивера) СВ.

Застосувавши принцип декомпозиції функціональних блоків для деталізації і уточнення методів управління витратами енергоресурсу в СМСП, отримаємо IDEF Model (як множину діаграм, їхні текстові описи та глосарій), які графічно і вербально відображають й описують функції підсистеми управління витратами енергоресурсу в СМСП.

На рисунку 7 представлено розроблену діаграму управління потужністю передачі, а на рисунку 8 – діаграму управління збереженням енергії батареї сенсора.

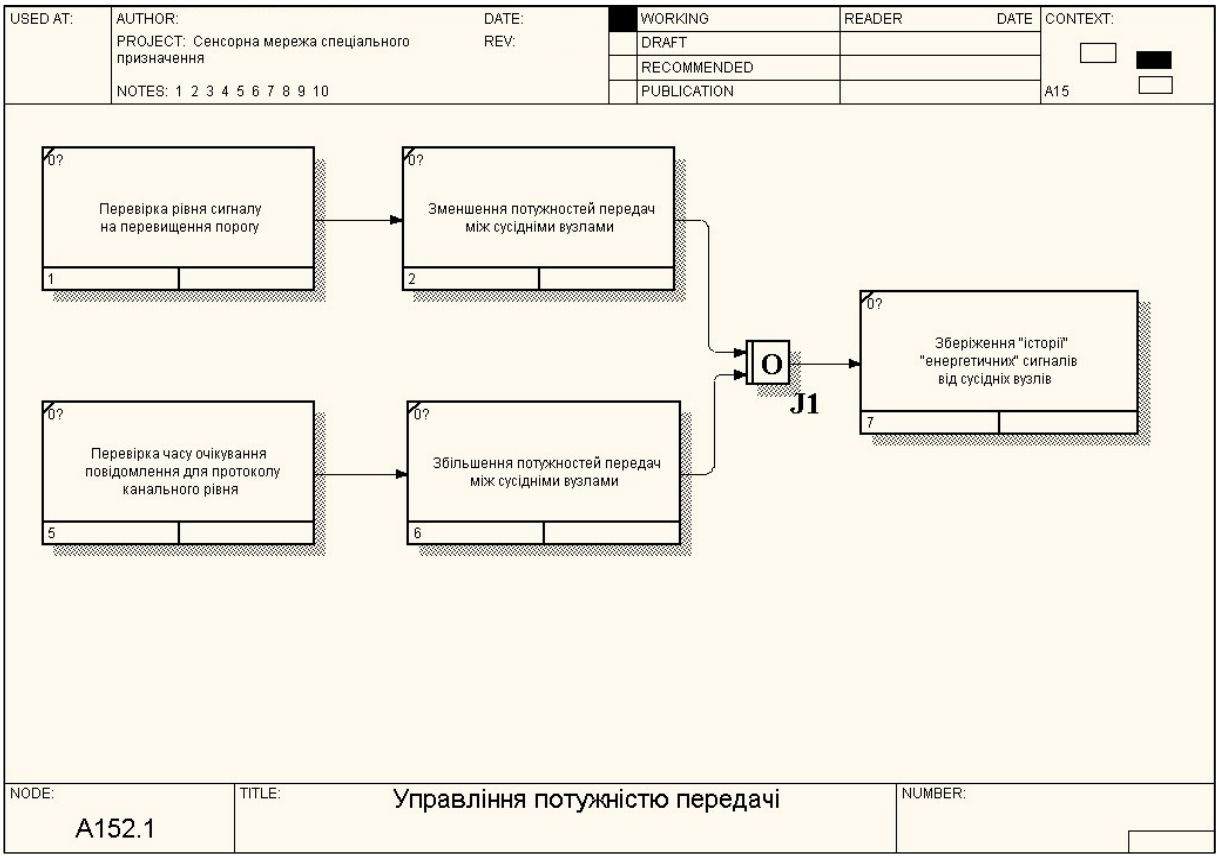


Рис. 7. Діаграма управління потужністю передачі

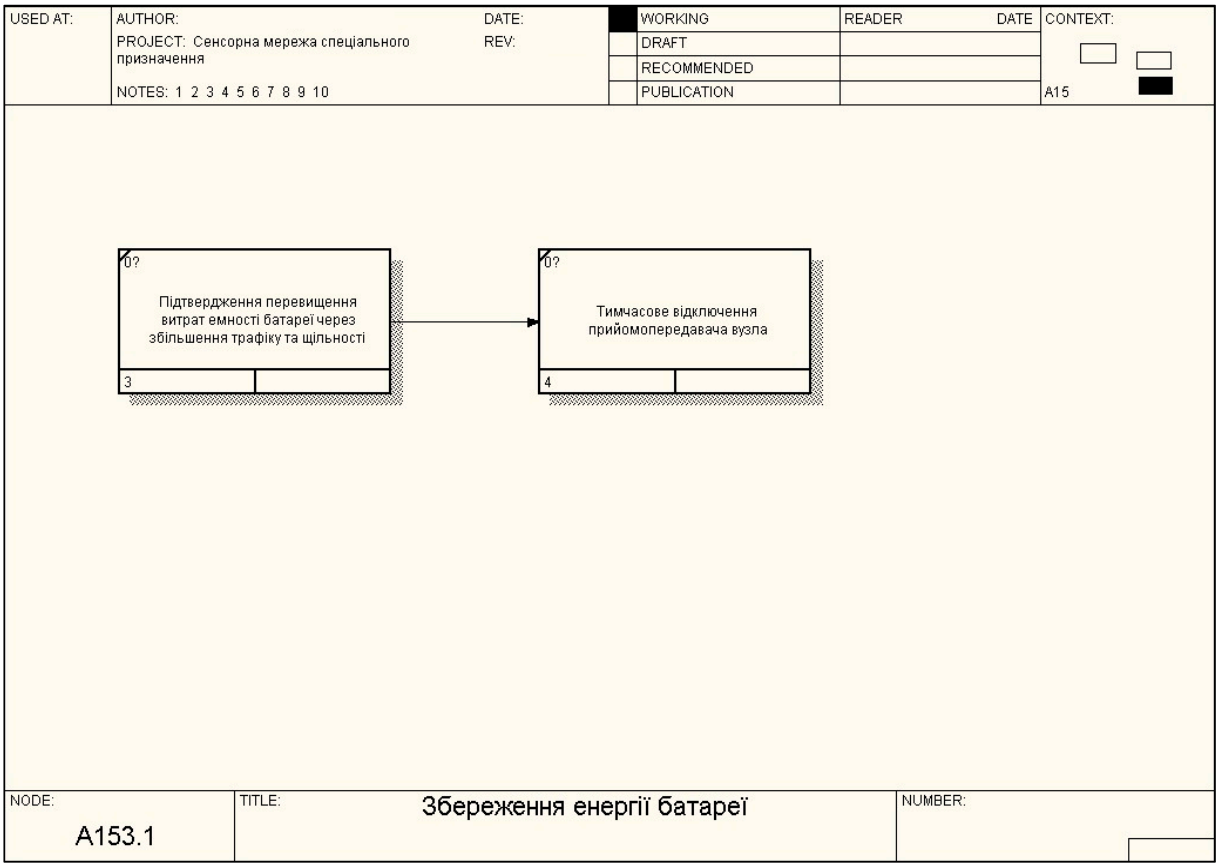


Рис. 8. Діаграма управління збереженням енергії батареї сенсора

Контекстна діаграма А-0 та діаграма підсистеми управління енергоресурсом СМСП А15 розроблені в методології IDEF0 функціонального моделювання, які описують функціональний зміст СМ у вигляді ієрархічної системи взаємопов'язаних функцій.

В свою чергу виконання кожної функції потребує реалізації технологічних процесів. Для опису (документування) технологічних процесів, які реалізуються в системі управління СМ при виконанні її функцій, застосована методологія IDEF3, яка надає інструментарій для візуального моделювання і дослідження сценаріїв технологічних процесів, що реалізуються в системі. Виконання кожного сценарію супроводжується відповідним документообігом.

Модель технологічного процесу в нотації IDEF3 являє собою діаграму PFDD (*англ.* Process Flow Description Diagramm – опис послідовності етапів процесу). Діаграми управління потужністю передачі А152.1 (див. рис. 7) та управління збереженням енергії батареї сенсора А153.1 (див. рис. 8) якраз і є діаграмами PFDD.

Отже, розроблена IDEF-model системи управління СМ дозволить системно (комплексно) перейти до етапу алгоритмізації і програмування функцій підсистеми управлінням витратами енергоресурсу СМСП.

Висновки

1. В інноваційному розвитку БСМ у секторі безпеки та оборони України можна, насамперед, виділити удосконалення енергоефективності СМСП (актуальним є дослідження тривалості функціонування вузлів СМСП («часу життя») через вдосконалення систем живлення вузлів, використання енергоефективних технологій, альтернативних джерел енергії та удосконалення компонентів ПМЗ управління витратами енергоресурсу вузлів мережі).

2. Управління енергоспоживанням вузлів СМСП повинно здійснюватися комплексно за функціями управління мережею на різних рівнях еталонної моделі взаємодії відкритих систем OSI (фізичному, канальному, мережевому).

3. Стандартом IEEE802.15 визначено, що «енергоспоживання в сенсорній мережі повинно бути якомога меншим за рахунок низької швидкості передачі даних і відповідати вимогам до часу життя СВ», але не сформульовано вимоги до енергозбереження в СВ завдяки застосуванню енергоекономних методів за різними функціями управління і рівнями моделі взаємодії відкритих систем (OSI).

4. Проведені дослідження елементів ПМЗ (моделей, методів) управління витратами енергоресурсу СМСП показали, що застосування розглянутих методів дозволяє в середньому збільшити «час життя» СМСП у 1,5–2 рази.

5. Комплексне застосування сукупності методів управління енергоресурсом буде визначатися проєктом конкретної СМСП, предметною областю її застосування і прийнятими рішеннями щодо реалізації інших функцій управління мережею.

Подальші шляхи досліджень

1. Розробка методики побудови ПМЗ управління витратами енергоресурсу СВ СМ, яке буде орієнтоване для потреб всіх підсистем оперативного управління СМСП і яке можливо органічно вбудувати в середовище реалізації процесів управління.

2. Дослідження проблем надійності і функціональної стійкості ПМЗ управління витратами енергоресурсу СМСП.

3. Дослідження науково-методичних аспектів формування складу ПМЗ управління витратами енергоресурсу для архітектури СМ конкретного застосування та специфіки бойового використання СМ.

4. Розробка методичного підходу до вирішення задачі оцінки ефективності методів ПМЗ в контексті управління витратами енергоресурсу СМСП.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Про Стратегічний оборонний бюлетень України: рішення Ради національної безпеки і оборони від 20.08.2021 № 0063525-21 URL: https://zakon.rada.gov.ua/laws/show/n0063525-21.

2. Бунин С. Г., Войтер А. П., Ильченко М. Е., Романюк В. А. Самоорганизующиеся радиосети со сверхширокополосными сигналами. К.: НПП «Издательство наукова думка НАН Украины». 444 с.

3. Шиллер Й. Мобильные коммуникации. Пер. с англ. М.: «Вильямс», 2002. 384 с.

4. Naraynaswamy S., Kawadia V., Sreenivas R. S., Kumar P. R. Power control in ad-hoc networks: Theory. Architecture, algorithm and implementation of the COMROW protocol // In Proceedings of EuroWireless. No. 02. 2002. P. 156–162.

5. Singh S., Raghavendra C. S. PAMAS – pover aware multi – access protocol with signaling for adhoc nttworks // ACM Computer Communications Review, 1998.

6. Xu Y., Heideman J., Estin D. Geography-informed Energy Conservation for Ad Hoc Routing // In Proceedings IEEE MOBICOM. No. 01. 2001.

7. Wang B., Gupta K. S. S-Remit: A Distributed Algorithm for Source-based Energy Efficient Multicast in Wireless Ad Hoc Networks // In Proceedings of GLOBECOM. No. 03. 2003.

8. Sheu J.-P., Chang Y.-C., Tsai H.-P. Power-Balance Broadcast in Wireless Mobile Ad Hoc Networks // In Proceedings of EuroWireless. No. 04. 2004.

9. Шаціло П. В., Цуркан В. В. Методологія аналізу, моделювання та проектування систем і процесів IDEF: навч. посіб. К: Вид-во ІСЗЗІ НТУУ «КПІ», 2012. 148 с.

10. Коваленко І. Г. Аналіз методів енергозбереження в сенсорних радіомережах // Зб. наук. праць ВІТІ НТУУ «КПІ». 2011. № 1. С. 76–84.

11. Новіков В. І. Метод збільшення часу життя безпровідної сенсорної мережі з надлишковою кількістю вузлів під час стеження за цілями моніторингу. К: Вчені записки ТНУ ім. В. І. Вернадського. Серія «Технічні науки». Том 28 (67). 2017. № 2.

12. Лисенко О. І. Функціональна модель системи управління безпроводовою сенсорною мережею із самоорганізацією для моніторингу параметрів навколишнього середовища / О. І. Лисенко, К. С. Козелкова, В. І. Новіков, Т. О. Прищепа, А. В. Романюк // Системи обробки інформації. 2015. Вип. 10 (135). С. 222–226.

13. Гаптельманов А., Міхаль О., Щепка О. Підвищення енергозбереження бездротових сенсорних мереж з використанням методів машинного навчання // Системи управління, навігації та зв’язку. 2022. № 4.

DOI: 10.58254/viti.5.2024.09.101

УДК 621.391

д-р техн. наук Могилевич Д. І. ORCID: 0000-0002-4323-0709 (ВІТІ ім. Героїв Крут)
Долженко Г. В. ORCID: 0009-0005-2940-7087 (ВІТІ ім. Героїв Крут)

АНАЛІЗ СТАНУ СИСТЕМИ ТЕХНІЧНОГО ОБСЛУГОВУВАННЯ ЕЛЕКТРОННОГО КОМУНІКАЦІЙНОГО ОБЛАДНАННЯ МЕРЕЖІ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ

Впровадження великої кількості новітніх засобів зв’язку та експлуатація їх в умовах ведення бойових дій вимагає аналізу стану системи технічного обслуговування та ремонту з урахуванням реальних умов функціонування системи зв’язку. Проведено аналіз останніх публікацій при дослідженні організації технічного обслуговування електронного комунікаційного обладнання мережі спеціального призначення та встановлено, що практично відсутні дослідження оцінки впливу технічного обслуговування на процес функціонування системи зв’язку.

У сучасних умовах на систему зв’язку діють зовнішні та внутрішні чинники негативного впливу. Збереження або автоматичне відновлення можливості виконання всіх або визначених функцій системи зв’язку в умовах впливу потоку експлуатаційних відмов, навмисних пошкоджень, втручання в обмін та обробку інформації, а також помилок обслуговуючого персоналу забезпечується функціональною стійкістю системи. Попередження та зменшення впливу дестабілізуючих внутрішніх факторів, що впливають на систему зв’язку, забезпечується системою технічного обслуговування та ремонту. Існуюча система технічного обслуговування застаріла та не враховує умови експлуатації засобів зв’язку.

Невелика кількість відомих моделей об’єктів технічного обслуговування є недостатньою для дослідження технічного обслуговування існуючих та перспективних зразків техніки. Застосування великої номенклатури сучасного електронного комунікаційного обладнання без належної системи технічного обслуговування може бути причиною зриву функціонування системи зв’язку та втрати управління військами. Подальші дослідження будуть спрямовані на удосконалення методів побудови моделей технічного обслуговування з урахуванням виконання вимог системи зв’язку до функціонування.

Ключові слова: *технічне обслуговування, система зв’язку, функціональна стійкість, надійність.*

D. Mogylevych, H. Dolzhenko Analysis of the technical maintenance system state of electronic communication equipment in a special-purpose network

The implementation of a large number of state-of-the-art communication devices and their operation in combat conditions requires an analysis of the state of technical maintenance and repair systems, taking into account the actual conditions of communication system operation. An analysis of recent publications on the organization of technical maintenance of electronic communication equipment of special-purpose networks has been conducted, revealing a lack of research on the assessment of the impact of technical maintenance on the functioning of communication systems.

In modern conditions, external and internal factors of negative influence affect the communication system. The preservation or automatic restoration of the ability to perform all or certain functions of the communication system under the influence of operational failures, deliberate damage, interference in information exchange and processing, as well as errors by the servicing personnel, is ensured by the functional stability of the system. Prevention and mitigation of the impact of destabilizing internal factors affecting the communication system are provided by the system of technical maintenance and repair. The existing technical maintenance system is outdated and does not take into account the operating conditions of communication devices. The limited number of known models of technical maintenance objects is insufficient for the study of technical maintenance of existing and prospective equipment models. The application of a large range of modern electronic communication equipment without an adequate technical maintenance system can lead to a disruption in the functioning of the communication system and loss of command over the troops. Further research will focus on improving methods of building technical maintenance models, considering the fulfillment of communication system requirements for operation.

Keywords: *technical maintenance, communication system, functional stability, reliability.*

Постановка завдання. Мережа зв’язку спеціального призначення виконує важливу роль у забезпеченні автоматизації управління військами Збройних Сил України, особливо в умовах повномасштабного вторгнення російської федерації в Україну. Зміна потреб військ зв’язку при забезпеченні виконання завдань в умовах ведення бойових дій призвела до потреби впровадження сучасних засобів зв’язку. Водночас існує необхідність дослідження організації системи технічного обслуговування (ТО) електронного комунікаційного обладнання (ЕКО) з урахуванням реальних умов функціонування системи зв’язку.

Аналіз останніх публікацій. Дослідженню проблеми організації ТО військової техніки та озброєння ЗСУ присвячена значна кількість робіт, аналіз основних із них викладено нижче.

В [1] проаналізовано традиційні стратегії ТО, вказано їхні переваги та недоліки. Наведено перспективи застосування прогнозованого ТО ЕКО з використанням машинного навчання, аналізу даних і моніторингу його технічного стану в реальному часі, що дозволить підвищити операційну ефективність, скоротить час простою й оптимізує витрати на ТО. Однак не вказано недоліки цієї стратегії, які потрібно враховувати, а саме: збір та аналіз великого обсягу даних, розробку прогнозних моделей, впровадження систем моніторингу стану засобів комунікаційного обладнання та підготовку обслуговуючого персоналу.

В [2] встановлено позитивний вплив застосування стратегії ТО та ремонту озброєння та військової техніки (ОВТ) за технічним станом на загальну ефективність через зниження питомої вартості; підвищення часу використання за призначенням; зменшення контрольних перевірок завдяки кращій інформованості про стан працездатності ОВТ; зниження витрат на ремонт завдяки запобіганню непередбачених виходів ОВТ з ладу; зменшення кількості випадків по усуненню відмов на місцях виконання завдань ОВТ; зменшення витрат, пов'язаних з наявністю аварій і заміни ОВТ; зниження часу знаходження в обслуговуванні; здатність виявляти й втручатися на ранніх стадіях до виникнення необхідності капітального ремонту двигуна; знижене споживання ЗІП завдяки збільшенню строків використання; зниження вартості управлінських дій завдяки прогнозуванню заявок на поставку ЗІП. З метою досягнення євроатлантичних стандартів та критеріїв, необхідних для набуття членства в НАТО, запропоновано необхідні умови щодо практичного впровадження ТО і ремонту за станом у процесі організації експлуатації ОВТ ЗС України. Однак не досліджено питання проведення ТО за станом для сучасного ЕКО.

В [3] зроблений висновок щодо необхідності вдосконалення чинної системи ТО та приведення її як у відповідність до вимог чинної системи в ЗС країн-членів НАТО, так і до застосування тих видів ТО, які б забезпечували під час обслуговування найкращий показник надійності нової чи вдосконаленої системи з обслуговування, але при дослідженні не враховано реальні умови експлуатації новітніх засобів зв'язку, що може вплинути на організацію системи ТО.

В [4] проведено аналіз наявної системи технічного забезпечення засобів зв'язку та автоматизованої системи управління (АСУ) і системи підготовки фахівців-ремонтників у країнах, що входять до складу ОЗС НАТО. Пропонуються підходи до створення аналогічної системи в Україні, а також до застосування видів ремонту, їх ешелонування, визначення загального критерію оцінки ефективності сил і засобів технічного забезпечення з'єднань з підтримки ОВТ у справному та готовому до бойового застосування стані, але не вказано шляхи розвитку системи ТО ЕКО.

В [5] встановлено, що основними недоліками управління технічним станом військової техніки на теперішній час є недосконалість процесів ТО та відновлення військової техніки, що його забезпечують. Сформульовано основні шляхи усунення вказаних недоліків, але не розглядається, як запропоновані рішення вплинуть на процес експлуатації засобів зв'язку.

В [6] запропоновано методичний підхід, який дозволить спрогнозувати зміни станів засобів зв'язку, що дасть загальне розуміння стану всього комплексу засобів зв'язку та автоматизації АСУ військами пункту управління, але не було запропоновано, що робити з отриманими результатами.

В [7] запропоновано основні напрямки удосконалення системи ТО цифрових засобів зв'язку: визначення головної мети ТО в процесі експлуатації цифрових засобів зв'язку, які надходять на озброєння; визначення сукупності взаємопов'язаних засобів, виконавців і документації з ТО, призначеної для підтримання справного і працездатного стану техніки зв'язку; оптимізація періодичності та видів ТО цифрових засобів зв'язку з визначенням

методик проведення кожного із них, але не було досліджене питання ТО засобів зв'язку подвійного призначення.

В наукових працях, які розглянуто, практично відсутні результати, що стосуються оцінки впливу ТО на функціональну стійкість ЕКО.

Метою статті є аналіз стану системи ТО ЕКО мережі спеціального призначення та її вплив на процес функціонування системи зв'язку.

Виклад основного матеріалу.

Система зв'язку (СЗ) – сукупність взаємозв'язаних, сумісних та узгоджених за завданнями вузлів і ліній військового зв'язку, орендованих каналів передавання і групових трактів та утворених на їхній основі систем (мереж), що призначена для вирішення задач забезпечення управління військами (силами), зброєю в мирний час, під час їх приведення у вищі ступені бойової готовності, підготовки та ведення операцій (бойових дій) [8]. Сьогодні спостерігається стабільна тенденція до загальної модернізації і розвитку СЗ, зокрема, важливим є створення та функціонування багатопараметричної системи, яка інтегрує функції управління військами, зброєю, розвідкою, радіоелектронною боротьбою, а також зв'язком, навігацією та орієнтуванням. Це вимагає відповідності концепції управління військами прийнятій у країнах НАТО та побудови єдиного інформаційного простору.

В цьому контексті відбувається оснащення підрозділів передовими засобами зв'язку та переходу до сучасних цифрових технологій.

Процес функціонування СЗ – це взаємодія компонентів системи, включаючи обладнання та програмне забезпечення, яка забезпечує виконання покладених на них функцій, а саме передачу й обробку інформації між користувачами. У сучасних умовах на СЗ негативно впливають як внутрішні (відмови, збої, помилки персоналу), так і зовнішні (активний або пасивний вплив зовнішнього середовища та противника) фактори. Збереження або автоматичне відновлення можливості виконання всіх або визначених функцій СЗ в умовах впливу потоку експлуатаційних відмов, навмисних пошкоджень, втручання в обмін та обробку інформації, а також помилок обслуговуючого персоналу забезпечується функціональною стійкістю системи.

В роботах [9; 10] сформульована властивість функціональної стійкості та загальна теорія її визначення для складних технічних систем. При цьому, в якості показників функціональної стійкості системи пропонується обрати сімейство $P(F_\tau)$, що визначає ймовірність збереження функціональних властивостей системи для фазової траєкторії $z(t,\alpha)$ на інтервалі часу $I = [0,\tau]$, де $F_\tau = F_\tau\{z(t,\alpha), t \le \tau\}$ – однопараметричне сімейство дійсних функціоналів, $t,\tau \in I, \alpha \in A$, де A – множина структур системи. Після визначення $B^\tau_{A_I}$ як множини припустимих значень F_τ отримано критерій функціональної стійкості системи у вигляді виразу (10) [11]:

$$P\left\{F_\tau[z(t,\alpha), t \le \tau] \in B^\tau_{A_I}\right\}. \quad (1)$$

Місце функціональної стійкості в системі властивостей складних технічних систем, якою є СЗ, визначено у роботі [12]. Зроблено висновок про те, що:

сутність надійності системи полягає в збереженні у встановлених межах значень всіх параметрів, що характеризують здатність виконувати необхідні функції в заданих режимах та умовах застосування, ТО та транспортування під час експлуатаційних відмов, спричинених фізичним старінням, конструктивно-виробничими недоліками; надійність забезпечується застосуванням високонадійної елементної бази та системи ТО;

сутність стійкості до відмов полягає в збереженні працездатності системи в цілому зі встановленою якістю під час експлуатаційних відмов та збоїв у елементах системи; стійкість до відмов забезпечується застосуванням всіх видів резерву, можливістю деградації системи до заданого рівня;

сутність живучості полягає в здатності системи протистояти зовнішнім впливам та бойовим пошкодженням, зберігаючи обмежену працездатність; живучість забезпечується заходами із захисту від протистояння противника та зовнішніх впливів;

сутність адаптивності полягає в збереженні певної частини системи при зміні множини параметрів, що дозволяє оптимальним способом досягти мети, яка була визначена.

Забезпечення функціональної стійкості СЗ досягається застосуванням у складній системі різних, вже існуючих видів надлишковості (апаратурної, структурної, часової, інформаційної, функціональної та ін.) шляхом перерозподілу ресурсів з метою зменшення наслідків позаштатних ситуацій [13].

Попередження та зменшення впливу негативних внутрішніх факторів, що впливають на СЗ та її функціональну стійкість, забезпечується системою ТО та ремонту. Технічне обслуговування являє собою комплекс робіт, які проводяться з метою підтримання техніки зв’язку та автоматизованого управління військами в справному стані під час зберігання, транспортування, підготовки до використання за призначенням та використання за призначенням.

До основних задач ТО належать:

запобігання передчасному зносу і виходу з ладу механічних елементів;

виходу електричних параметрів апаратури за межі встановлених норм;

виявлення і усунення несправностей і причин їх виникнення;

доведення параметрів і характеристик до норм;

продовження міжремонтних ресурсів (термінів) і термінів експлуатації [14].

Система ТО включає в себе наступні завдання:

контроль технічного стану та перевірка ефективності;

технічне обслуговування;

планове поточне обслуговування та капітальний ремонт;

підготовка вимірювальних інструментів для калібрування.

Строки проведення ТО визначені наказом начальника Генерального штабу Збройних Сил України від 11 квітня 2003 року № 22 (зі змінами) [14].

Слід зауважити, що цей наказ розроблявся на основі технічної експлуатації ЕКО, яке знаходилось в той час на озброєнні ЗСУ. На зразках техніки, оснащених комплексами вмонтованого автоматизованого контролю і діагностики, передбачається поточне ТО за станом у необхідному обсязі, але впровадження великої кількості різнотипного обладнання, часто подвійного призначення без відповідного їх метрологічного і діагностичного забезпечення, створює певні труднощі в організації ТО та ремонту – відсутні технологічні картки, принципові схеми та рекомендації до проведення регламентних робіт.

Структура, склад і порядок побудови СЗ визначені в [15]. Згідно з [16] на пункті управління повинен забезпечуватись набір визначених комунікаційних сервісів, які необхідні начальнику для ефективного управління військами (табл. 1)

Таблиця 1

Приклад матриці комунікаційних сервісів

Посадова особа ПУ	Телекомунікаційні сервіси					
	ТА ЗСУ001	ТА ЗСУ002	АРМ ЗК ЗСУ001	АРМ ЗК ЗСУ002	Термінал ВКЗ	АРМ «Віраж»
Начальник ПУ	+	+	+	+	+	+
Заступник НПУ	+	+	+	+		
Начальник штабу ПУ	+	+	+	+		

Надання визначених комунікаційних сервісів забезпечується використанням великої номенклатури ЕКО вітчизняного та закордонного виробництва, зазвичай подвійного

призначення. Вітчизняні виробники розробляють та надають необхідну документацію, що використовується для коректної експлуатації, ТО та ремонту ЕКО.

Розглянемо ЕКО мережі спеціального призначення, що отримало широке застосування при побудові СЗ, а саме радіостанції Harris Falcon II та Falcon III, що забезпечують надійний зв’язок навіть у складних умовах.

Режим псевдовипадкової перебудови робочої частоти робить ці радіостанції стійкими до радіоелектронної протидії противника.

Завдяки широкому діапазону частот радіостанції Harris Falcon II та Falcon III можуть виконувати множинні завдання, такі як:

одноканальний тактичний зв’язок;

вузькосмуговий та широкосмуговий режим роботи;

автоматична ретрансляція та маршрутизація інформації (MANET) [17].

Для радіостанцій виробництва «Harris» основними вважаються два види ТО (профілактичне і позапланове) та три рівні обслуговування. Профілактичне обслуговування – це планові заходи, спрямовані на запобігання поломок обладнання та мінімізацію простоїв.

Профілактичне обслуговування включає в себе:

підтримку обладнання в чистому та сухому стані, без пилу;

щоденний огляд: візуальний огляд, вбудовані тести, перевірка батареї, кабелів, роз’ємів, антени;

щотижневий огляд: більш детальний огляд, включаючи перевірку антени на пошкодження, роз’ємів на корозію, ковпачків;

щорічне обслуговування: комплексна перевірка радіостанції та заміна батареї HUB.

Позапланове ТО проводиться при виявленні несправностей під час профілактичного обслуговування або роботи радіостанції.

Рівні обслуговування включають в себе:

перший рівень – визначення працездатності (виконується оператором без додаткового обладнання);

другий рівень – більш детальна діагностика (виконується оператором з додатковим обладнанням);

третій рівень – діагностика до модуля (виконується сертифікованим персоналом з використанням спеціального обладнання) [8].

При побудові транкінгового зв’язку використовуються абонентські станції (автомобільні й портативні) та ретранслятор (один або декілька) стандарту Mototrbo. Обладнання працює згідно зі стандартом DMR у діапазоні VHF від 136 МГц до 174 МГц (Національна Гвардія України використовує обладнання Mototrbo у діапазоні UHF від 403 МГц до 470 МГц), з чотирьохпозиційною частотною маніпуляцією (4FSK) у цифровому режимі та частотною модуляцією (FM) у аналоговому. Для аналого-цифрового перетворення мови використовується кодек AMBE+2.

Засоби транкінгового зв’язку стандарту Mototrbo виробництва компанії Motorola, які відрізняються високими якістю та функціональними можливостями, підтримують цифровий режим роботи та забезпечують криптографічний захист інформації, на сьогодні широко використовуються у системі військового зв’язку для організації радіозв’язку в тактичній ланці управління і добре зарекомендували себе під час проведення антитерористичної операції.

Технічне обслуговування радіостанцій являє собою комплекс організаційно-технічних заходів планово-запобіжного характеру щодо підтримки їх у стані відповідно до вимог технічної документації протягом усього строку експлуатації. Основні завдання ТО радіостанцій: забезпечення правильного функціонування; контроль технічного стану радіостанцій і визначення придатності до подальшої експлуатації; виявлення й усунення несправностей і причин їх появи; ліквідація або недопущення наслідків впливу несприятливих

кліматичних, експлуатаційних й інших дестабілізуючих факторів; аналіз і узагальнення відомостей результатів виконаних робіт, розробка заходів щодо вдосконалювання форм і методів ТО [18].

ЕКО виробництва компаній «Cisco» та «Grandstream» виконують важливу роль при побудові СЗ. Маршрутизатори, комутатори, голосові шлюзи, а також сервери характеризуються високою надійністю та здатністю до відновлення після збоїв та відмов, часто без участі персоналу, який його експлуатує. Однак на вказане ЕКО відсутні технологічні картки, принципові схеми та рекомендації щодо проведення ТО. Широке застосування ЕКО без належного ТО ставить під загрозу функціонування всієї системи, адже вихід з ладу одного критично важливого елемента буде вимагати або фізичної заміни на працездатне, або довготривалого усунення несправностей. В умовах повномасштабної війни росії проти України це може призвести до фатальних наслідків, таких як зрив управління військами. На цей момент такі критичні ситуації попереджаються наявністю структурного резерву, який в випадку зриву функціонування повністю замінює непрацездатне обладнання.

Несвоєчасне проведення ТО на ЕКО мережі спеціального призначення може призвести до технічних відмов обладнання та викривлення вихідної інформації від зовнішнього середовища та користувачів, стохастичних збоїв й фізичного руйнування елементів і компонентів обладнання, що, в свою чергу, буде причиною зриву функціонування та зниження функціональної безпеки та стійкості.

Дослідження організації оптимального ТО можна класифікувати за двома напрямками:

вивчення конкретних об’єктів ТО;

побудова і дослідження відповідних математичних моделей ТО.

Невелика кількість відомих моделей об’єктів ТО є недостатньою для дослідження ТО наявних та перспективних зразків техніки з урахуванням реальних умов функціонування.

Висновки. Для забезпечення процесу функціонування СЗ необхідно враховувати внутрішні та зовнішні фактори впливу, що можуть негативно вплинути на функціональну стійкість системи. Попередженням та зменшенням негативних чинників слугує організація системи ТО ЕКО мережі спеціального призначення. Застосування великої номенклатури сучасного ЕКО без належної системи ТО може бути причиною зриву функціонування СЗ та втрати управління військами.

Подальші дослідження будуть спрямовані на удосконалення методів побудови моделей ТО з урахуванням виконання вимог СЗ до функціонування.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Anthony Lawrence Paul, Anthony Odu, Joseph Oluwaseyi. Predictive Maintenance: Leveraging Machine Learning for Equipment Health Monitoring. *Ladoke Akintola University of Technology*. 2024. P. 6. URL: https://www.researchgate.net/publication/377411657_Predictive_Maintenance_Leveraging_Machine_Learning_for_Equipment_Health_Monitoring.

2. EXPERIENCE IN THE APPLICATION OF ADAPTIVE STRATEGIES FOR MAINTENANCE AND REPAIR OF MILITARY PRODUCTS IN NATO MEMBER COUNTRIES / P. Openko et al. *Наукові праці Державного науково-дослідного інституту випробувань і сертифікації озброєння та військової техніки*. 2021. No. 8. P. 101–111. URL: https://doi.org/10.37701/dndivsovt.8.2021.11.

3. Напрями вдосконалення системи технічного обслуговування засобів зв’язку та АСУ ЗС України / М. Івченко та ін. *Збірник наукових праць Військового інституту телекомунікацій та інформатизації*. 2019. № 1. С. 18–22. URL: http://nbuv.gov.ua/UJRN/Znpviti_2019_1_4.

4. Желновач О., Ткаченко А. Л., Івченко М. М. Дослідження шляхів розвитку системи технічного обслуговування та ремонту засобів військового зв’язку та АСУ при виконанні завдань логістичного забезпечення силами оборони держави з урахуванням підходів, прийнятих в об’єднаних збройних

силах НАТО. *Збірник наукових праць Національної академії Національної гвардії України*. 2023. № 2 (42). С. 44–51. URL: http://znp.nangu.edu.ua/issue/view/17393/10065.

5. Вплив процесів технічного обслуговування та відновлення на ефективність управління технічним станом військової техніки в провідних країнах світу та Збройних сил України / Ю. Баранов та ін. *Збірник наукових праць Національної академії Державної прикордонної служби України. Серія: військові та технічні науки*. 2020. Т. 81, № 3. С. 302–315. URL: https://doi.org/10.32453/3.v81i3.478.

6. Козачук В. Л. Методичний підхід до визначення параметрів процесу моніторингу озброєння та військової техніки під час експлуатації. *Сучасні інформаційні технології у сфері безпеки та оборони*. 2017. № 2. С. 37–41. URL: http://nbuv.gov.ua/UJRN/sitsbo_2017_2_8.

7. Майборода І. М., Глущенко М., Лазарев В. Д. Методика проведення технічного обслуговування цифрових засобів зв’язку. *Національна академія Національної гвардії України. Системи управління, навігації та зв’язку*. 2021. № 2 (64). С. 154.

8. ВСТ 01.112.001 – 2006. Військовий зв’язок та інформаційні системи. Військовий зв’язок. Терміни та визначення. [Чинний від 2006-03-05]. Вид. офіц. С. 14.

9. Артюшин Л., Машков О. Оптимізація цифрових автоматичних систем, стійких до відмов. К.: КВВАИУ. 1991. С. 89.

10. Теорія автоматичного керування / Л. Артюшин та ін. Львів: УАД, 2004. 272 с.

11. Миколайчук Р. Функціональна стійкість складних технічних систем з динамічною структурою. *Моделювання та інформаційні технології*. 2013. № 70. С. 32–33. URL: http://nbuv.gov.ua/UJRN/Mtit_2013_70_7.

12. Барабаш О. В. Построение функционально устойчивых распределенных информационных систем // К.: НАОУ. 2004. Т. 226. С. 96.

13. Калашник-Рибалко М. А. Методика забезпечення функціональної стійкості пілотажно-навігаційного комплексу літального апарату на окремих режимах польоту. *Збірник наукових праць Харківського національного університету Повітряних Сил*. 2018. № 1 (55). С. 4. URL: https://doi.org/10.30748/zhups.2018.55.09.

14. Керівництво з технічного забезпечення зв’язку та автоматизації управління військами Збройних Сил України: затв. наказом Начальника Генерального штабу Збройних Сил України від 11.04.2003 № 22.

15. Про затвердження Тимчасової настанови з організації зв’язку та інформаційних систем: Наказ Генерального штабу Збройних Сил України від 08.08.2018 № 08. С. 8.

16. Про затвердження Матриці надання мінімально необхідних сервісів (базових та функціональних) оперативному складу органів військового управління на пунктах управління Збройних Сил України: Наказ Головнокомандувача Збройних Сил України від 14.12.2020 № 221. Київ. С. 5.

17. Новітні технології та засоби зв’язку у ЗСУ: шлях трансформації та перспективи розвитку / О. Лаврут та ін. *Наука і техніка Повітряних Сил Збройних Сил України*. 2019. № 1 (34). URL: https://www.ukrmilitary.com/2019/04/signal.html.

18. Сучасні військові засоби радіо та супутникового зв’язку / І. Борисов та ін. Київ: ВІТІ. URL: https://sprotyvg7.com.ua/wp-content/uploads/2023/07/сучасні.pdf.

DOI: 10.58254/viti.5.2024.10.108

УДК 621.396

канд. техн. наук Панченко І. В. ORCID: 0000-0001-5690-3813 (ВІТІ ім. Героїв Крут)
Бондаренко Д. М. (ДержНДІ технологій кібербезпеки)
канд. техн. наук Липський О. А. ORCID: 0009-0007-7355-4433 (ДержНДІ технологій кібербезпеки)
Стефанишин Я. І. ORCID: 0000-0002-8317-4131 (ДержНДІ технологій кібербезпеки)
Ушаков В. Д. (ФОП «Ушаков Віктор Дмитрович»)

АНАЛІЗ ПІДХОДУ ДО ЗАХИСТУ ВІД ЗАВАД СУПУТНИКОВИХ НАВІГАЦІЙНИХ ПРИЙМАЧІВ БПЛА

Розглянуто основні підходи до побудови приймачів сигналів глобальних систем супутникової навігації безпілотних літальних апаратів, захищених від впливу природних чи штучно створених завад. Такі системи оснащуються багатоелементними фазованими антенними решітками та пристроями управління приймальною діаграмою спрямованості антени Controlled Reception Pattern Antenna. У процесі оброблення прийнятого інформаційного сигналу пристроєм Controlled Reception Pattern Antenna забезпечується селекція завад та їх придушення (компенсація), що еквівалентно формуванню провалів у просторовій діаграмі спрямованості антенної решітки на кутах, що відповідають напрямам на джерела завад.

Наведено основні засади реалізації адаптивного алгоритму оброблення прийнятого інформаційного сигналу. Розглянуто конструктивне виконання та склад навігаційного приймача безпілотного літального апарата «Орлан-10», оснащеного пристроєм Controlled Reception Pattern Antenna. Зроблено висновок щодо принципової можливості створення вітчизняного навігаційного приймача з пристроєм Controlled Reception Pattern Antenna.

У перспективі передбачається розгортання робіт щодо створення пристроїв Controlled Reception Pattern Antenna з більшою кількістю антенних елементів та використанням принципів оптимізації кількісного складу операцій адаптації, з метою забезпечення роботоспроможності цих пристроїв при більшій кількості одночасно увімкнених джерел завад та більших рівнях потужності їхніх сигналів.

Ключові слова: *глобальні системи супутникової навігації, безпілотні літальні апарати, пристрої Controlled Reception Pattern Antenna, природні та штучно створені завади.*

I. Panchenko, D. Bondarenko, O. Lypskyi, Ya. Stefanyshyn, V. Ushakov Analysis of the approaches to protection against interference of UAV satellite navigation receivers.

The main approaches to the construction of signal receivers of global satellite navigation systems of unmanned aerial vehicles, protected from the influence of artificially created interference, are considered. Such systems are equipped with multi-element phased antenna arrays and devices of controlled reception pattern antenna. In the process of processing the received information signal by the Controlled Reception Pattern Antenna device. the selection of interference and their suppression (compensation) is ensured, which is equivalent to the formation of dips in the spatial pattern of the antenna array at the corners that correspond to the directions of the interference sources.

The basic principles of the implementation of the adaptive algorithm for processing the received information signal are given. The design and composition of the navigation receiver of the «Orlan-10» unmanned aerial vehicle equipped with a Controlled Reception Pattern Antenna device is considered. A conclusion was made regarding the fundamental possibility of creating a domestic navigation receiver with a Controlled Reception Pattern Antenna device.

In the future, it is envisaged to deploy works on the creation of Controlled Reception Pattern Antenna devices with a greater number of antenna elements and the use of optimization principles of the quantitative composition of adaptation operations, in order to ensure the performance of these devices with a greater number of simultaneously enabled sources of interference and higher levels of their signal power.

Keywords: *global satellite navigation systems, unmanned aerial vehicles, Controlled Reception Pattern Antenna devices, natural and man-made obstacles.*

Постановка проблеми

На поточний час уявити собі навігацію безпілотних літальних апаратів (БпЛА) без використання глобальних навігаційних супутникових систем (GNSS – Global Navigation Satellite Systems) доволі складно. Особливо, якщо мова іде про необхідність досягнення ними кінцевих пунктів маршруту на земній поверхні, що знаходяться на значних відстанях від місць запуску (початкових пунктів маршруту).

У більшості випадків навігація БпЛА забезпечується за допомогою включення до їх складу бортових навігаційних приймачів GNSS, які задовільно функціонують, але тільки за сприятливих зовнішніх умов. Загалом, приймачі GNSS дуже вразливі до перешкод – навмисно створених завад (глушіння – jamming, імітація, перехоплення – spoofing). Тому, без вжиття заходів щодо боротьби із завадами, ефективне спеціальне використання БпЛА (за широкого застосування засобів РЕБ) у більшості випадків неможливе.

На поточний час, що характеризується масовим застосуванням БпЛА на лінії зіткнення, проблема зменшення впливу засобів РЕБ на функціонування GNSS виступає дуже гостро. Особливої актуальності ця проблема набуває при реалізації нагальної потреби налагодження серійного випуску вітчизняних БпЛА та, відповідно, створення сучасної вітчизняної авіоніки для них.

Аналіз останніх досліджень і публікацій

Принцип дії GNSS заснований на розрахунку географічних координат об'єкта на основі прецизійного вимірювання відстані до кількох супутників (мінімально, більше 3), які знаходяться на орбіті. При цьому, просторове положення супутників з великою точністю відоме у будь-який момент часу.

Наявні GNSS тією чи іншою мірою забезпечують навігацію практично в усіх куточках поверхні земної кулі (включаючи моря і океани). Наразі повноцінно функціонують чотири GNSS: GPS, США; GLONASS, рф; GALILEO, Європа; BeiDou, Китай та на різних етапах розгортання ще дві (IRNSS, Індія, та QZSS, Японія). Особливістю всіх GNSS є порівняно невеликі рівні потужностей інформаційних сигналів, які досягають вхідних ланцюгів навігаційних приймачів (орієнтовно «мінус» 125 дБм).

Наразі системи GNSS працюють спільно з так званими системами геостаціонарного доповнення або системами супутникової диференціальної навігації (SBAS – Satellite Based Augmentation System), які дають можливість суттєво зменшити локальну похибку визначення геодезичних координат завдяки використанню додаткового наземного компоненту, покликаного у реальному масштабі часу генерувати відповідні поправки та передавати їх на навколишні приймачі GNSS.

На поточний час у світі активізовано роботи щодо створення і постачання на ринок систем захисту приймачів GNSS від дії природних чи штучно створених завад – так званих пристроїв управління діаграмою спрямованості приймальної антени (CRPA – Controlled Reception Pattern Antenna). За принципом функціонування пристрій CRPA є типовою системою адаптивного оброблення інформації, загальні підходи до алгоритмічного забезпечення яких обґрунтовано у роботі [1]. Зокрема, у цій роботі наведено функціональні схеми оброблення вузькосмугового та широкосмугового сигналів, а також визначено цільовий критерій автоматичного функціонування адаптаційного алгоритму. Авторами роботи [2] представлено графік впливу періоду зникнення сигналу GNSS на похибку визначення координат, що, зважаючи на ітераційний характер адаптаційного процесу, дає можливість оцінити завадозахищеність пристрою CRPA. У роботі [3] підтверджено наявність залежності між кількістю елементів антенної решітки та кількістю джерел завад, які можуть бути усунуті пристроєм CRPA. Авторами роботи [4] розглянуто принципи побудови та функціональну схему аналогового пристрою CRPA. Відомості про технічні характеристики пристроїв CRPA, які постачаються на ринок як окремі прилади, наведено у публікаціях [5–10]. Відомості про конструктивно суміщений з приймачем GNSS пристрій CRPA «Комета-М» представлено у роботі [11].

Метою цієї статті є викладення основних підходів до створення алгоритмів адаптивного оброблення сигналів приладами CRPA (які забезпечують успішне функціонування приймачів сигналів GNSS на борту БпЛА в умовах дії природних та штучно створених завад), а також побіжний аналіз конструктивного виконання навігаційного приймача «Комета–М».

Виклад основного матеріалу

Принципи функціонування адаптивних пристроїв придушення завад. При побудові систем придушення сигналів завад (характеристики яких оперативно змінюються) в інформаційному сигналі навігаційного приймача GNSS наразі використовуються адаптивні системи, теоретичні засади яких висвітлено у роботі [1]. Призначення цих систем полягає у формуванні вихідного сигналу з найбільшим (у тому чи іншому сенсі) наближенням до вхідного неспотвореного сигналу (тобто найменшою похибкою між ними).

Загальна схема придушення завад в адаптивній системі, що має 2 вхідні сигнали, наведена на рисунку 1 [1].

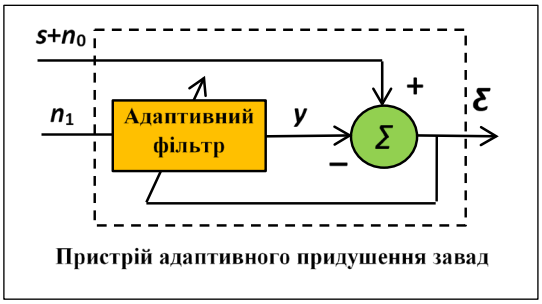


Рис. 1. Функціональна схема адаптивного пристрою придушення завад

Вихідний сигнал адаптивної системи придушення завад, представлений на рисунку 1, записується, як:

$$\varepsilon = s + n_0 - y,$$

де $s + n_0$ – вектор суміші сигналу та завади (вхідний сигнал пристрою адаптивного придушення завад);

n_1 – вектор сигналу завади, некорельованої з сигналом s, але корельованої з сигналом n_0;

y – вектор вихідного сигналу адаптивного фільтра, який приблизно відповідає копії сигналу n_0;

ε – вектор сигналу похибки між істинним та прийнятим сигналом s на виході суматора $\sum$.

Система керується адаптивними алгоритмами, які автоматично формують її вихідний сигнал із найкращим (зокрема, у середньоквадратичному сенсі) наближенням до вхідного сигналу. У процесі функціонування адаптивний фільтр налаштовується відповідно до цього алгоритму так, щоб мінімізувати загальну потужність вихідного сигналу ε, який, у свою чергу, є також сигналом похибки адаптивного процесу.

Враховуючи, що середньоквадратичне відхилення (СКВ) вихідного сигналу системи має тільки один мінімум (локальні мінімуми – відсутні) [1], цей процес легко автоматизується.

Крім того, вибір у якості критерія придушення завади мінімуму СКВ вихідного сигналу дає можливість певним чином максимізувати відношення сигнал/шум на виході системи. Це можна пояснити тим, що повної компенсації завади (шуму) у системі досягти неможливо і мінімізація загальної потужності вихідного сигналу призводить, в основному, до мінімізації його шумової складової.

Адаптивні системи придушення завад, які використовуються у приймальних пристроях GNSS, призначені для оброблення сигналів, прийнятих елементами фазованих антенних решіток, що мають просторове рознесення (орієнтовно, на половину довжини робочої хвилі) і іменуються адаптивними просторовими решітками.

На рисунку 2 наведено функціональну схему одного із варіантів адаптивної просторової решітки – пристрою адаптивного оброблення вузькосмугового сигналу [1]. У цій схемі кожний елемент антенної решітки (1 … k) з'єднаний з перемножувачами на змінні вагові коефіцієнти (w_{11} … w_{k1}), а також з фазообертачами на 90 градусів, які, у свою чергу, з'єднані з

перемножувачами на змінні вагові коефіцієнти (w_{21} … w_{k2}). Таким чином, перебираючи значення вагових коефіцієнтів (за критерієм мінімізації СКВ вихідного сигналу), можна вибрати будь-які фазові значення кута та амплітуди комплексного коефіцієнта передачі вхідного сигналу, що дає можливість принципово отримати компенсацію завади на виході суматора (∑).

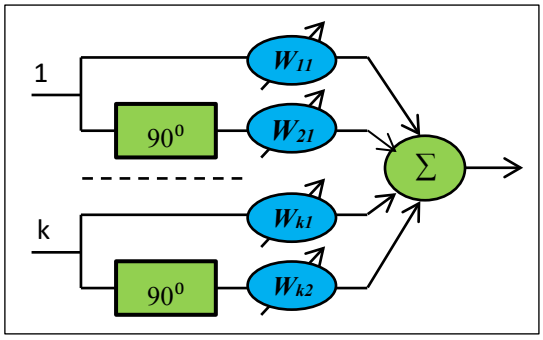


Рис. 2. Функціональна схема пристрою адаптивного оброблення вузькосмугового сигналу

За необхідності забезпечити адаптивне приймання сигналів у певному діапазоні частот, кожний фазообертач, як показано на рисунку 3, замінюється адаптивним трансверсальним фільтром з відводами [1]. Вибір якомога більшої кількості відводів, чи іншими словами менших фазових затримок (**Δ**), дає можливість наблизитися до ідеального фільтра, який дає можливість регулювати амплітуду і фазу сигналу на кожній частоті у смузі пропускання.

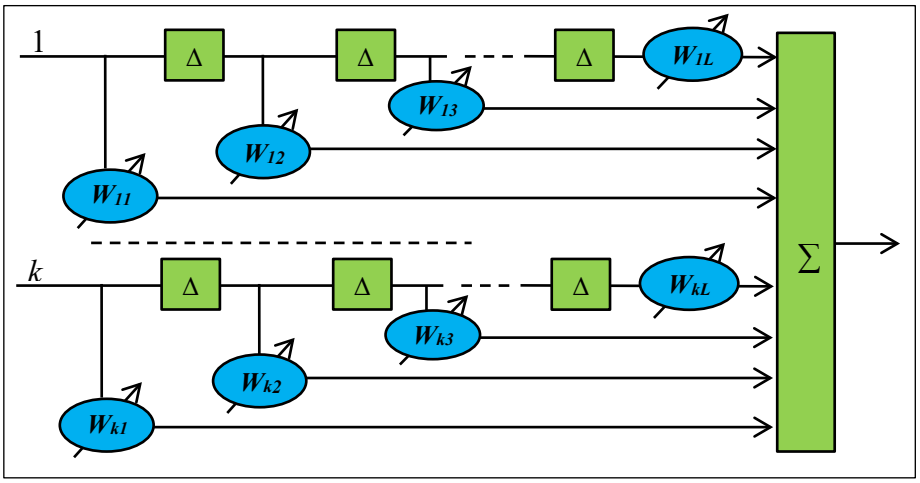


Рис. 3. Функціональна схема пристрою адаптивного оброблення широкосмугового сигналу

Незважаючи на те, що схема, представлена на рисунку 3, принципово не передбачає включення ланцюжка зі зсувом фаз на 90 градусів, інтуїтивно видається доцільним його залишити (хоча б із міркувань більш точного підбору вагових коефіцієнтів). Таке рішення дещо збільшує апаратурні витрати, але при цифровому обробленні сигналу такий недолік можна вважати несуттєвим. Загалом, це питання потребує остаточного вирішення у процесі порівняльних випробувань макетних зразків.

Для реалізації в адаптаційному алгоритмі основної мети підбору вагових коефіцієнтів (забезпечення максимально швидкого досягнення мінімуму СКВ вихідного сигналу) використовуються алгоритми градієнтного спуску, найменших квадратів та інші. Відмінність між алгоритмами полягає у реалізації тих чи інших заходів щодо зменшення загального

терміну адаптації, в основному завдяки збільшенню кроку підбору вагових коефіцієнтів на певних кроках адаптації за певним критерієм (наприклад, відповідно до величини першої похідної на кожному кроці адаптації).

Зазвичай, у складі переважної більшості навігаційних приймачів GNSS використовуються багатоелементні (більше 4) антенні решітки, які дають можливість забезпечити адаптивну компенсацію природних чи штучно створених завад. При цьому селекція завад, яка забезпечується у процесі оброблення прийнятих від багатоелементної антени сигналів, еквівалентна формуванню провалів у просторовій діаграмі спрямованості антенної решітки, на кутах, що відповідають напрямам на джерела завад.

Видається, що детальне практичне відпрацювання алгоритмів адаптації буде досягнуто у процесі проєктування пристрою. Загалом, максимальний цикл адаптації визначається кількістю елементів антенної решітки k, кількістю відводів трансверсального фільтра L, похибкою визначення вагових коефіцієнтів w, проміжною частотою, на якій реалізується адаптивне оброблення сигналу, а також прийнятим алгоритмом адаптації. Наприклад, якщо припустити, що антенна решітка 4-елементна ($k = 4$), трансверсальні фільтри мають $L = 360$ відводів кожний, вагові коефіцієнти змінюються у межах $w = (0 \pm 1)$ і визначаються з похибкою не більше 10^{-4}, кожний антенний елемент має квадратурний канал, а проміжна частота (ПЧ) складає 10 МГц, то будемо мати близько $576 \cdot 10^5$ операцій адаптації і загальний максимальний цикл адаптації – близько 6 с (передбачається, що одна ітерація відбувається за один період робочої частоти). При менших частотах проміжної частоти та виконанні циклу адаптації за один період робочої частоти загальний термін адаптації значно зростає (зокрема, для частоти ПЧ 1 МГц термін адаптації може становити орієнтовно 1 хвилину), що диктує необхідність пошуку певного компромісу (реалізація, за можливості, циклу адаптації не за один період ПЧ, а за частину цього періоду, зменшення кількості відводів трансверсального фільтра, дискретності вагового коефіцієнта чи оптимізації самої кількості циклів адаптації).

Оцінити практичну прийнятність терміну адаптації можна аналізуючи дані, наведені у роботі [2], щодо залежності СКВ похибки визначення координат об'єкта від терміну зникнення сигналу GNSS (за умови комплексування на борту системи GPS та інерційної навігаційної системи), які представлено у таблиці 1. Видається, що для практичного використання, при ураженні БпЛА достатньо протяжних цілей, термін затримки до 10 с, який визначає рівень СКВ величиною 5–6 м, цілком задовільний, а от термін затримки близько 1 хвилини (величина СКВ складає 50–60 м) можна вважати прийнятним тільки при створенні макетного зразка, що демонструє принцип роботи.

Таблиця 1

Похибка визначення координат при зникненні сигналу GNSS

Час зникнення даних GNSS, *с*	**СКВ визначення координат,** *м*
10	5–6
20	7–15
40	20–40
60 (1 хв)	50–60
300 (5 хв)	300

Як наголошується у роботі [3], максимальна кількість джерел завад, яка може бути скомпенсована, не перевищує кількості використовуваних елементів антенної решітки (на практиці, в технічних характеристиках вказується величина на одиницю менша).

Функціональні схеми пристроїв CRPA. Конструктивно пристрої CRPA можуть будуватися як у аналоговому, так і у цифровому вигляді.

Наприклад, у роботі [4] представлена структурна схема 4-канального аналогового пристрою CRPA, яка у спрощеному варіанті наведена на рисунку 4. Вихідний сигнал пристрою CRPA, який не містить сигналу завади взагалі (або інтенсивність завади суттєво зменшена внаслідок реалізації адаптивного алгоритму), подається на антенний вхід штатного навігаційного приймача.

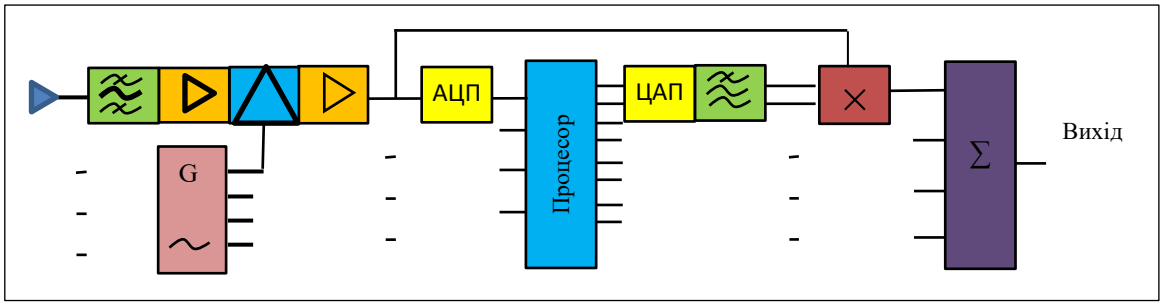


Рис. 4. Функціональна схема аналогового пристрою CRPA

Особливістю наведеної схеми є подача на перемножувачі сигналів кожного із антенних каналів сигналів квадратурних вагових коефіцієнтів, які формуються процесором, що дає можливість не втратити їх комплексний характер і підбирати амплітуду та фазу цих сигналів так, щоб забезпечити точну компенсацію сигналів завад. До недоліків наведеної аналогової схеми побудови CRPA можна віднести використання аналогових перемножувачів, які апаратурно достатньо громіздкі. Крім того, реалізація трансверсального фільтра у аналоговому варіанті (при спробі побудови пристрою адаптивного оброблення широкосмугового сигналу) також становить певні труднощі.

Цифровий пристрій CRPA (що є фактично певною модифікацією аналогового пристрою CRPA) забезпечує цифрове оброблення прийнятого аналогового сигналу на ПЧ.

Функціональна схема цифрового пристрою CRPA представлена на рисунку 5.

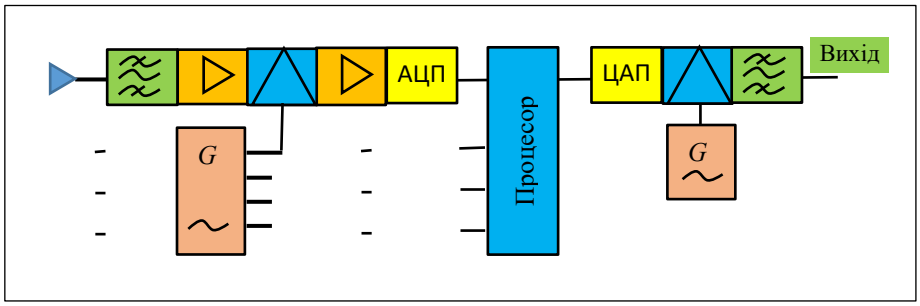


Рис. 5. Функціональна схема цифрового пристрою CRPA

Особливістю цієї схеми побудови CRPA є необхідність зворотнього перенесення обробленого вихідного сигналу з проміжної на робочу частоту перед поданням його на вхід штатного приймача сигналу GNSS. До переваг цифрового пристрою CRPA можна віднести можливості реалізувати адаптивне оброблення широкополосного сигналу з довільною кількістю відводів та оперативно корегувати програмне забезпечення (при необхідності врахування конкретних умов експлуатації).

Крім того, якщо необхідно реалізувати роботу системи CRPA від декількох GNSS, що мають різні робочі частоти, їх розділення достатньо просто забезпечується цифровими фільтрами, реалізованими за допомогою процесора, що загалом дає помітну економію апаратних ресурсів.

Серед недоліків цифрового оброблення сигналів можна назвати дещо більше енергоспоживання, але переваги, які надає мінімізація габаритно-масових показників та пристосованість до серійного виробництва, нівелюють цей недолік.

Приклади пристроїв CRPA, представлених на ринку. Загалом, на поточний час на світовому ринку у модульному виконанні предсталено навігаційні приймачі GNSS, які широко використовуються у приладах і системах різного призначення. При цьому, із комерційних міркувань не кожний навігаційний приймач GNSS потребує компенсації завад і тому не всі вони обладнуються пристроями CRPA.

Зважаючи на те, що пристрій CRPA має функціональну завершеність, він може підключатися до будь-якого штатного навігаційного приймача, суттєво покращуючи його роботу в умовах дії завад.

Враховуючи, що автори не ставили за мету ґрунтовно проаналізувати ринок пристроїв CRPA, для прикладу наводиться тільки декілька найпоширеніших пропозицій.

Турецькою компанією Tualcom пропонується декілька модифікацій функціонально закінчених пристроїв CRPA [5–8], загальний вигляд яких представлено на рисунку 6, а основні технічні характеристики – у таблиці 2.

Tualaj 8200 Tualaj 8100, 8300 Tualaj 8300-D

Рис. 6. Загальний вигляд CRPA пристроїв компанії Tualcom, Туреччина

Таблиця 2

Основні технічні характеристики пристроїв CRPA компанії Tualcom, Туреччина

Найменування параметру	**Тип пристрою CRPA**			
	Tualaj 8100	**Tualaj 8200**	**Tualaj 8300-D**	**Tualaj 8300**
1. Типи сумісних GNSS приймачів	BeiDou (B1C), Galileo (E1), SBAS, GLONASS (L1), GPS (L1)		BeiDou (B1C), Galileo (E1, 5a), SBAS, GLONASS (L1, L2), GPS (L1, L2), QZSS (L1)	
2. Кількість сигналів завад	не більше 7			
3. Чутливість приймача, *дБм*	«мінус» 164			
4. Придушення сигналу завади, *дБ*	не менше 50			
5. Смуга частот, *МГц*	не менше 30			
6. Напруга живлення, *В*	12–28			
7. Потужність споживання, *Вт*	не більше 12		не більше 24	
8. Форм-фактор, *мм*	70 × 139 × 21	90 × 110 × 23	Ø250, h = 98	70 × 139 × 210
9. Маса, *г*	320	380	4150	380
10. Діапазон робочих температур, *градуси*	від «мінус» 40 до 85			

Компанією Height Technologies, Нідерланди [9], пропонується недорогий ($ 50) CRPA-пристрій GPS DOME для навігаційних систем БпЛА, загальний вигляд якого наведено на рисунку 7. Основними особливостями пристрою, окрім підключення додаткової антени, є невеликі габарити (форм-фактор) – 70 мм × 48 мм × 24 мм та маса – 150 г.

Компанією Hexagon Novatel, Канада [10], пропонується декілька варіантів антенних решіток (власне CRPA-пристроїв для GNSS), до складу яких включаються 4- та 7-елементні фазовані антенні решітки. Зокрема, на рисунку 8 представлено загальний вигляд 7-елементної антенної решітки GALT-710ML, а на рисунку 9 – CRPA-пристрою GAJT-AE-N, які призначені для використання у складі колісної та гусеничної техніки.

Антенна решітка GALT-710ML обладнана малошумним підсилювачем з коефіцієнтом підсилення 30 дБ та має діаметр 289 мм, висоту 120 мм і масу 7,5 кг. CRPA-пристрій GAJT-AE-N забезпечує роботу з сигналами наступних GNSS: GPS (L1, L2), QZSS (L1, L2) SBAS (L1), Galileo (E1), споживає 15 Вт електроенергії від джерела живлення 10–32 В, має габаритні розміри 179,5 мм × 155,5 мм × 39 мм та масу не більше 1,2 кг.

Рис. 7. Загальний вигляд CRPA-пристрою GPS DOME компанії Height Technologies, Нідерланди

Рис. 8. Антенна решітка GALT-710ML компанії Hexagon Novatel, Канада

Рис. 9. CRPA-пристрій GAJT-AE-N компанії Hexagon Novatel, Канада

З метою оптимізації масогабаритних характеристик БпЛА його навігаційний приймач GNSS можна виготовляти у вигляді єдиного конструктиву з інтеграцією антенної решітки та CRPA-пристрою. Таке технічне рішення прийняте, зокрема, у системі GPS «Комета–М», якою оснащується більшість воєнної техніки рф (у тому числі і БпЛА «Орлан-10»).

Технічні характеристики системи GPS «Комета–М» наведено у роботі [11]:

кількість антенних елементів – 4;

робоча температура – від «мінус» 40 до 60 градусів;

потужність споживання, не більше – 12 Вт;

кількість супутникових каналів – 32;

підтримувані сигнали GPS/GALILEO/SBAS L1 – 1575,42 МГц, ГЛОНАСС L1 – 1597,5 МГц, 1609,5 МГц;

діапазон навігаційних сигналів, який захищається: ГЛОНАСС L1 – (1593–1610) МГц, GPS L1, GALILEO E1/ SBAS L1 – 1573–1578 МГц;

придушення широкосмугових перешкод – 40–50 дБ;

перешкодозахищеність (завада/сигнал) J/S = 90 дБ;

інтерфейс RS422/RS485, PPS, ВЧ-вихід;

протокол обміну даними – NMEA (за замовчуванням), BINR.

Конструктивно, приймач GPS «Комета–М» складається із 2 плат (антенна 4-елементна решітка та навігаційний модуль з CRPA-пристроєм). Загальний вигляд (з лицьової та тильної сторін) антенної решітки БпЛА «Орлан-10» наведено на рисунку 10, а графік залежності коефіцієнту підсилення її антенного елементу (Patch Antenna Taolas CGGP 18.4.C.02) від частоти, який входить до її складу, – на рисунку 11.

Розмір плати антенної решітки складає 100×100 мм, антенні елементи розташовані по кутах квадрату розміром 85 мм × 85 мм. Враховуючи представлену на рисунку 11 залежність, орієнтовний коефіцієнт підсилення елементу антенної решітки на робочій частоті 1575,42 МГц можна оцінити величиною «мінус» 1,5–2 дБі, що загалом свідчить про його переважальну роль у якості не стільки антени, як частотно- та просторово-селективного засобу.

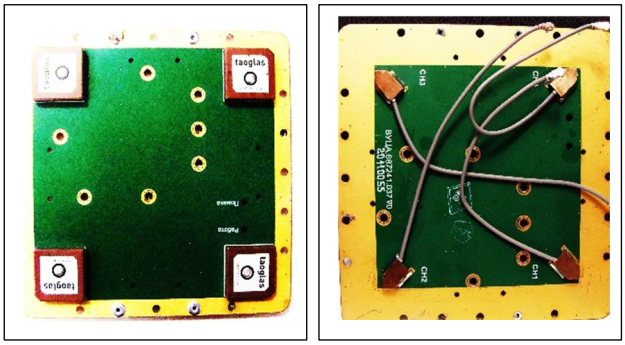


Рис. 10. Загальний вигляд антенної 4-елементної решітки навігаційного приймача БпЛА «Орлан-10», рф

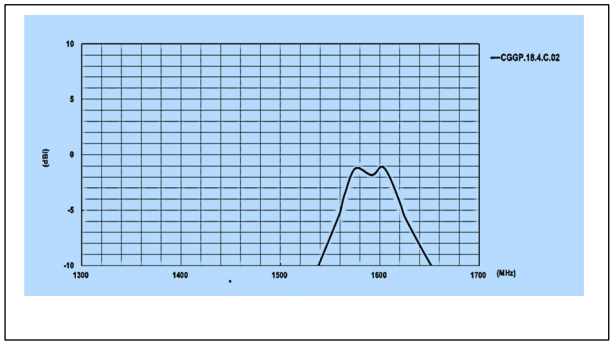
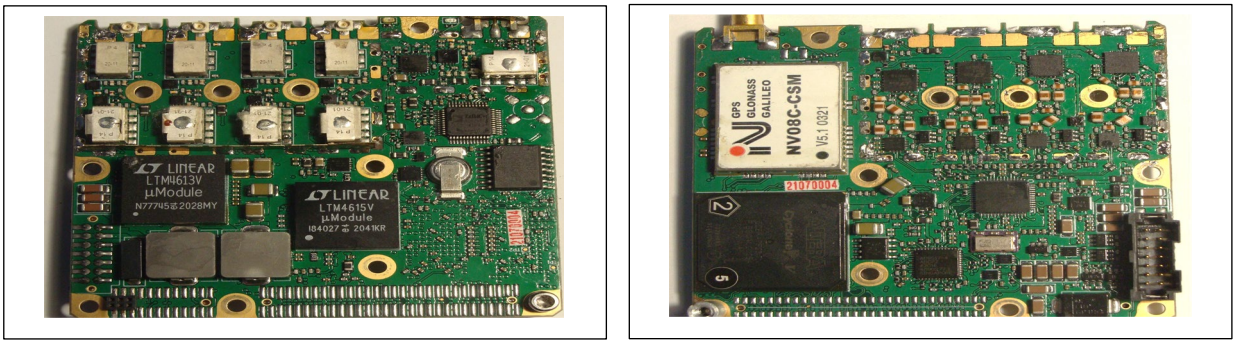


Рис. 11. Залежність коефіцієнта підсилення антенного елементу Patch Antenna Taolas CGGP 18.4.C.02 від частоти

Загальний вигляд (з двох сторін) дещо пошкодженої (відсутність деяких комплектувальних приладів) плати навігаційного модуля з CRPA-пристрою БпЛА «Орлан-10» наведено на рисунку 12.



а) *б)*

Рис. 12. Загальний вигляд плати навігаційного модуля БпЛА «Орлан-10»:
а – вид спереду; *б* – вид з тильної сторони

Аналіз конструктивного виконання плати дає можливість зробити висновок, що приймач GNSS побудований за функціональною схемою з цифровим обробленням інформації, представленою на рисунку 5.

Зокрема, у якості навігаційного приймача використано малогабаритний мультисистемний модуль GPS/GLONASS/GALILEO/COMPASS типу NV08C–CSM [13]. Процесор CRPA, виконаний на базі пристрою Altera Cyclone V компанії Intel [14]. Змішувачі

LTC5541 компанії LT Linear Technology [15] включені у сигнальних ланцюгах кожного із 4 приймальних каналів та у вихідному сигнальному ланцюгу. Гетеродин виконаний на базі 12-вихідного тактового генератора з інтегрованим 1,6 GHz керованим напругою генератором AD9517-4 компанії Analog Devices [16]. Перетворення цифрового сигналу в аналоговий забезпечується 14-розрядним мультиплексованим цифро-аналоговим перетворювачем (ЦАП) AD9759-4 компанії Analog Devices [17]. Зворотне перетворення аналогового сигналу у цифровий (АЦП) забезпечується 14-розрядним аналогово-цифровим 4-канальним перетворювачем LTC2175-14 компанії LT Liner Technology [18].

Таким чином, з урахуванням розглянутих принципів функціонування пристрою CRPA видається цілком можливим створення вітчизняного приймача GNSS для систем БпЛА. Побудова макетного зразка цього приймача (за сприятливих умов, в основному економічного характеру) дала б можливість практично відпрацювати алгоритми оброблення сигналів GNSS, методи ефективного придушення роботи джерел перешкод, і, тим самим, створити передумови його серійного виробництва.

Успішне створення приймача GNSS з пристроєм CRPA можна розглядати як важливий крок у напрямі забезпечення вітчизняною авіонікою створюваних перспективних БпЛА.

Висновки і перспективи подальших досліджень

Селекція природних чи штучно створених завад та їх придушення забезпечується пристроями CRPA, що еквівалентно формуванню провалів у просторовій діаграмі спрямованості антенної решітки, на кутах, що відповідають напрямам на джерела завад.

Оброблення прийнятого сигналу пристроями CRPA забезпечується адаптивними алгоритмами, у якості цільової функції автоматичного функціонування яких виступає досягнення мінімуму СКВ вихідного сигналу. При цьому, окрім придушення завад також максимізується відношення сигнал/шум.

Реалізація адаптивного алгоритму у більшості випадків забезпечується у цифровому вигляді з використанням програмно реалізованих трансверсальних фільтрів.

З високою ймовірністю можна сказати, що саме таким чином побудований приймач GNSS «Комета–М» (до складу якого включено 4-елементну антенну решітку).

Загалом, при наявності зацікавленості і відповідного фінансування, вітчизняний приймач GNSS для БпЛА може бути створений (із налагодженням його серійного виробництва) в осяжні терміни.

У перспективі передбачається розгортання робіт щодо створення пристроїв CRPA з більшою кількістю антенних елементів та використанням принципів оптимізації кількісного складу операцій адаптації, з метою забезпечення роботоспроможності цих пристроїв при більшій кількості одночасно увімкнених джерел завад та більших рівнях потужності їхніх сигналів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Уидроу Б., Стирз С. Адаптивная обработка сигналов. Перевод с англ. М.: Радио и связь, 1989. 440 с.

2. Карпов Д. В., Коробецький О. В., Резніков Ю. В. Рекомендації щодо розробки захищеного від завад приймача глобальних навігаційних супутникових систем для вирішення завдань Збройних сил України // Системи, озброєння і військова техніка. 2020. № 4 (64). URL: https://www.C:/Users/Користувач/Downloads/455-Текст%20статті-814-1-10-20210114.pdf.

3. Теоретические основыі радиолокации / перев. Я. Д. Ширмана. М.: Советское радио, 1970. 560 с.

4. Швець В. А. Структурна схема завадостійкої антенної решітки навігаційних систем GPS, ГЛОНАСС, ГАЛІЛЕО. URL: https://www.core.ac.uk/download/pdf/286628241.pdf.

5. https://www.tualcom/gps-gnss-anti-jam-crpa/tualaj-8200.

6. https://www.tualcom/gps-gnss-anti-jam-crpa/tualaj-8100.
7. https://www.tualcom/gps-gnss-anti-jam-crpa/tualaj-8300.
8. https://www.unmannedsystemstechnology.com/company/tualcom/tualaj-8300-d-gnss-crpa-anti-jam-system.
9. https://gpsdome.nl/wp-content/uploads/2021/02gps-dome_UAS-datasheet.pdf.
10. https://novatel.com/products.
11. Сафронов Тарас. «Комета» – проблема для РЕБ України. URL: https://www.mil.in.ua/uk/articles/kometa-problema-dlya-reb-ukrayiny/ (дата звернення: 01.08.2023).
12. https://allaboutcurcuits.com/electronic-components/datasheet/CGGP.18.4.C.02–Taolas.
13. https://famell/com/datasheets/1792091.pdf, 12/01/2023.
14. https://intel.com/content/www/us/en/docs/programmable/683694/current/cyclon-v-device-overview.html.
15. https://analog.com/media/en/technical-documentation/data-sheets/5541f.pdf.
16. https://analog.com/media/en/technical-documentation/data-sheets/AD9517-4.pdf.
17. https://analog.com/media/en/technical-documentation/data-sheets/AD9755-4.pdf
18. https://analog.com/media/en/technical-documentation/data-sheets/21754314fa.pdf.

DOI: 10.58254/viti.5.2024.11.119

УДК 004.93:004.8

Процюк Ю. О. ORCID: 0000-0001-5193-3669 (ВІТІ ім. Героїв Крут)
Драглюк О. В. ORCID: 0000-0001-8572-7257 (ВІТІ ім. Героїв Крут)
Паламарчук Н. А. ORCID: 0000-0001-8818-7794 (ВІТІ ім. Героїв Крут)
Куцаєв П. В. ORCID: 0000-0002-3235-3316 (ВІТІ ім. Героїв Крут)
канд. техн. наук Овсянніков В. В. ORCID: 0000-0003-0186-6220 (ВІТІ ім. Героїв Крут)

ПІДХОДИ ДО ФОРМУВАННЯ БАЗИ ДАНИХ ЗОБРАЖЕНЬ В СИСТЕМІ ІДЕНТИФІКАЦІЇ ВИБУХОНЕБЕЗПЕЧНИХ ПРЕДМЕТІВ ДЛЯ ПОДАЛЬШОГО НАВЧАННЯ НЕЙРОННОЇ МЕРЕЖІ

Використання безпілотних літальних апаратів для виявлення вибухонебезпечних предметів є перспективним напрямком у галузі розмінування. Наразі значна роль при ідентифікації виявлених предметів залишається за оператором (людиною). Для вирішення задач з ідентифікації зображень вибухонебезпечних предметів, безпілотні літальні апарати або інші роботизовані системи доцільно обладнувати відповідними програмно-технічними засобами із вбудованими елементами штучного інтелекту, які б дозволили з певною ймовірністю розпізнавати виявлені об'єкти за їхніми характерними локальними ознаками (формою/зображенню, матеріалом виготовлення корпусу, температурою корпусу) та класифікувати їх без участі оператора.

Розпізнавання полягає в обчисленні міри релевантності між описами об'єкта та еталона, образи яких задані у вигляді множин векторів-дескрипторів та оптимізації значень цієї міри для бази еталонів.

У статті розглянуто основні алгоритми (методи) ідентифікації зображень та обрано ті, які мають вищу швидкодію та побудовані за бінарним типом, що спрощує процес обчислення та знижує затрати. В цілому, комбінування дескрипторів ORB та BRISK дає кращі результати, оскільки враховується спотворення кольору, погане освітлення, наявність шумів, різних проєкцій.

У статті запропоновано підходи до формування бази даних зображень вибухонебезпечних предметів, яка в подальшому виступатиме вхідним шаром для навчання нейронної мережі для ідентифікації вибухонебезпечних предметів. У процесі роботи був написаний програмний код, який надає можливість перевірити працездатність комбінованих методів (алгоритмів) ідентифікації зображень щодо формування бази даних, її збереження та тестування.

Ключові слова: *безпілотні літальні апарати, вибухонебезпечні предмети, дескриптор, ідентифікація зображення, нейронні мережі, розмінування, штучний інтелект.*

Y. Protsiuk, O. Draglyuk, N. Palamarchuk, P. Kutsaiev, V. Ovsyannikov. Approaches to the formation of an image database in the explosive ordnance identification system for further training of a neural network

The use of unmanned aerial vehicles to detect explosive objects is a promising area of demining. Currently, a significant role in identifying detected objects remains with the operator (human). To solve the problem of identifying images of explosive objects, it is advisable to equip unmanned aerial vehicles or other robotic systems with appropriate software and hardware with built-in elements of artificial intelligence that would allow to recognize detected objects with a certain probability by their characteristic local features (shape/image, body material, body temperature) and classify them without operator participation.

Recognition consists in calculating the relevance measure between the descriptions of the object and the reference, whose images are given as sets of descriptor vectors, and optimizing the values of this measure for the reference database.

The article discusses the main algorithms (methods) for image identification and selects those that have higher performance and are built according to the binary type, which simplifies the calculation process and reduces costs. In general, the combination of ORB and BRISK descriptors gives better results, since it takes into account color distortion, poor lighting, noise, and different projections.

The paper proposes approaches to forming a database of explosive ordnance images, which will further serve as an input layer for training a neural network to identify explosive ordnance. In the course of the work, a program code was written that makes it possible to test the performance of combined methods (algorithms) of image identification in terms of database formation, storage and testing.

Keywords: *unmanned aerial vehicles, explosive ordnance, descriptor, image identification, neural networks, demining, artificial intelligence.*

Постановка завдання. За період ведення бойових дій в Україні забрудненою вибухонебезпечними предметами (далі – ВНП) вважається вся територія, де йшли бої, не враховуючи території, де нині проходять бойові дії. За оцінками Асоціації саперів, перевірці на ВНП потребує п'ята частина території країни, що становить приблизно 139 тис. кв. км, і за оцінками деяких експертів на розмінування потрібно приблизно 750 років. Для порівняння, під час виконання робіт з очищення території України представниками ДСНС України з 2014 року по 2023 рік знешкоджено 1 262 134 од. ВНП, площа очищеної території склала 3 082 кв. км [1–3].

На сьогодні розробляються та використовуються новітні технології розмінування, такі як механізоване розмінування, роботизовані системи (комплекси), безпілотні літальні апарати (далі – БпЛА) з датчиками різних типів для виявлення та знешкодження ВНП [4]. Всі ці технології дозволяють віддалене управління засобами розмінування та убезпечують операторів від випадкового спарцювання ВНП, але не виключають їхньої присутності у роботі цих систем. Для вирішення задач з ідентифікації зображень ВНП, БпЛА або інші роботизовані системи доцільно обладнувати відповідними пристроями на кшталт камери з високою роздільною здатністю, тепловізора, автоматизованої системи зйомки тощо, із вбудованими елементами штучного інтелекту, які б дозволили з певною ймовірністю розпізнавати виявлені об'єкти за їхніми характерними локальними ознаками (формою/зображенню, матеріалом виготовлення корпусу, температурою корпусу) та класифікувати їх без участі оператора (людини). Тому, для проведення ідентифікації ВНП по їхньому зображенню в автоматизованому режимі необхідно розробити методику формування бази даних зображень для подальшого навчання нейронної мережі.

Аналіз останніх досліджень і публікацій. Дослідження, що проводяться в області розпізнавання предметів, зводяться до створення та впровадження сучасних засобів робототехніки з широким використанням елементів штучного інтелекту. Так, у [5–7] був проведений аналіз актуальних, сучасних та перспективних технологій робототехніки, зокрема БпЛА, їх застосування за допомогою штучного інтелекту, які без втручання людини можуть визначати наявність небезпеки шляхом дослідження результатів обстеження територій БпЛА та управління якістю розмінування. У [9] проведено навчання регіонально-згорткової нейронної мережі (Faster R-CNN) щодо ідентифікації ВНП за допомогою мультиспектральних і теплових даних, зібраних автоматизованою системою зйомки з БпЛА.

У [8; 10; 12; 13] авторами проведена робота з розробки та тестування методів обробки даних багатоканального дистанційного зондування на основі нейронних мереж, досліджена тема розпізнавання візуальних об'єктів із використанням класифікації дескрипторів особливих (ключових, контрольних) точок зображення на основі навчання нейронної мережі з описом методів SIFT, SURF, ORB та BRISK.

У [4] запропоновано синтез системи виявлення ВНП на базі БпЛА зі встановленим на ньому тепловізором, в поєднанні з металошукачем та маніпулятором із вибухівкою для дистанційного розмінування. Хоча значного прогресу в цій галузі вже досягнуто, залишається багато нерозв'язаних задач, у тому числі ідентифікація зображень ВНП в автоматизованому режимі. Тому пошук шляхів підвищення ефективності при виявленні та ідентифікації зображень ВНП за допомогою нейронних мереж є актуальним.

Метою статті є розробка підходів до формування бази даних зображень в системі ідентифікації вибухонебезпечних предметів для подальшого навчання нейронної мережі.

Виклад основного матеріалу. Постійний розвиток нових технологій напряму сприяє розвитку основних військових функцій. Технологічні зміни завжди впливали на військові дії та є суттєвим чинником результату військових конфліктів. Наукові розробки у США та країнах Європи спрямовані на використання цифрових технологій для аналізу даних за допомогою штучного інтелекту. До основних напрямків розвитку штучного інтелекту

відносять: глибинне навчання (Deep Learning), машинне навчання (Machine Learning), комп'ютерне бачення, автономні системи, робототехніку тощо.

Одним з найбільш розвинених алгоритмів для глибинного навчання з ідентифікацією зображень та відео є згорткові нейронні мережі (CNN – Convolutional Neural Network), які адаптовані для обробки зображень і спроєктовані з припущенням, що на вхід подається еталонне зображення. На відміну від стандартних нейронних мереж, шари у згорткових нейронних мережах є тривимірними, а нейрони одного шару зв'язані тільки з певною областю наступного шару. Це дозволяє вибудувати певні характеристики в архітектуру, спрямовану на поліпшення класифікації зображень.

Структурні методи розпізнавання об'єктів, що засновані на локальних ознаках зображення, набули популярності та широко застосовуються в сучасних системах комп'ютерного зору. Вони базуються на визначенні множини особливих (контрольних) точок (далі – контрольні точки) та їхньому описові у вигляді числового чи бінарного вектора – дескриптора, який відображає властивості зображення для локальних околиць цих точок. Зазначене дає можливість програмі працювати із візуальними образами подібно людині. Значення дескриптора є незмінним стосовно зміщення, повороту, масштабування об'єктів на зображенні, при цьому кількість утворених дескрипторів, що формують опис, повинна бути достатньою для прийняття рішення відносно ідентифікації. Розпізнавання полягає в обчисленні міри релевантності між описами об'єкта та еталона, образи яких задані у вигляді множин дескрипторів та оптимізації значень цієї міри для бази еталонних зображень.

Ідентифікація (розпізнавання) ВНП, як і будь-яких інших об'єктів, здійснюється за їхніми характерними ознаками. Так, аналіз ВНП показав [4], що найбільш розповсюдженими на поверхні є протипіхотні фугасні міни (ПФМ-1, міна-метелик), міни-пастки, протитанкові міни фугасної дії, протипіхотні міни осколкової дії, протитанкові міни дистанційного мінування. Для фугасних мін характерне використання пластмас, для мін осколкової дії – металу, відповідно основними характерними ознаками для виявлення та ідентифікації ВНП можуть бути їхні форма/зображення (висота, радіус, довжина, ширина), матеріал виготовлення корпусу (пластмаса, поліетилен, сталь, чавун), температура корпусу. За своєю специфікою ВНП можуть бути розкидані в різній місцевості (середовищі) – трава, пісок, сніг, листя та в різних проєкціях, що в подальшому впливатиме на обробку зображень. Ці характерні ознаки є визначальними при виборі способу (засобу) виявлення ВНП, кожний із яких має переваги та недоліки.

Пошук металевих (металовмісних) ВНП здійснюється за допомогою металошукача (під землею або на поверхні землі), пошук пластикових ВНП, що знаходяться на поверхні землі – за допомогою візуального аналізу зображення та за допомогою спеціального програмного забезпечення із залученням тепловізійної чи оптичної камери. Основною функцією металошукача є виявлення металу у міні, а оскільки багато років використовуються міни в неметалевих корпусах, то це значно ускладнює роботу. Також уповільнює процес виявлення наявність великої кількості хибних сигналів від осколків ВНП та інших металів, якими перенасичений ґрунт внаслідок ведення бойових дій [4].

Тепловізійний моніторинг, згідно з [14], найбільш ефективний лише у сонячні дні, в хмарну погоду є недоцільним, а встановлені в тіні зразки не ідентифікуються, як і закопані в ґрунт на глибину 2–5 см. Хоча при тепловізійному моніторингу наявність забруднення металевими частками, на відміну від класичних металошукачів, не становить проблеми. При цьому значна роль при ідентифікації виявлених предметів залишається за оператором (людиною).

Одним із варіантів ідентифікації ВНП, які знаходяться на поверхні землі, може бути їх розпізнавання в автоматизованому режимі з використанням алгоритмів обробки зображень та нейронної мережі. В цій роботі пропонується підхід до розробки методики формування бази

даних зображень ВНП на прикладі протипіхотної фугасної міни (ПФМ-1, міна-метелик) з подальшим використанням її для навчання нейронної мережі. ПФМ-1 обрано для дослідження, оскільки вони є найпоширенішими в Україні, встановлюються методом дистанційного мінування (на поверхні), що дозволяє отримати їхнє зображення, а також можуть маскуватися під місцевість (пісок, листя та ін.), що матиме певні особливості при обробці зображень. Проблема при виявленні цих мін полягає в тому, що вони мають малий розмір, маскуються та зроблені переважно з пластику, відповідно їх досить важко знайти за допомогою металодетекторів, а ефективність використання тепловізорів пов'язана із зазначеними обмеженнями та значною мірою залежить від погодних умов [4; 14].

Запропонована методика містить ряд кроків: від створення відповідного набору даних зображень (на прикладі ПФМ-1), їх попередньої обробки та ін., до тестування сформованої бази даних (рис. 1).

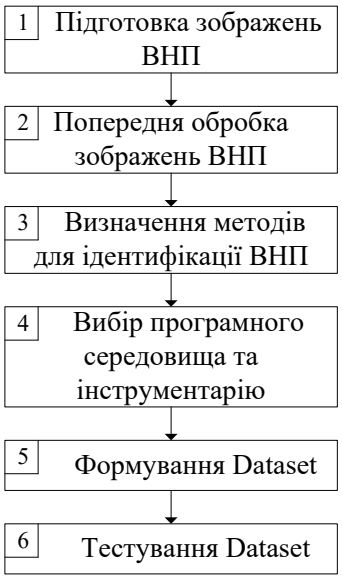


Рис. 1. Методика формування Dataset ВНП

Для якісного навчання нейронної мережі, необхідно на вхідний шар мережі подати сформовану базу даних зображень, еталонів ВНП. База даних зображень (далі – Dataset) повинна бути репрезентативною для всіх можливих варіантів, зображення повинні бути стійкі до спотворення кольору, поганого освітлення, наявності шумів, різних проєкцій, тобто, повинні відповідати вимогам до класифікації зображень.

Розглянемо детальніше кожний із наведених кроків методики. На першому кроці здійснюється підготовка бази зображень ВНП на прикладі ПФМ-1. Для наповнення Dataset можливо використовувати як спеціальне програмне забезпечення для колекціонування зображення з різних ресурсів мережі Інтернет (готові бази даних з різних платформ), так і створити власноруч із зображень, які були отримані за допомогою різних пристроїв – камер, сенсорів, супутників, БпЛА тощо. Для формування Dataset було завантажено 78 зображень ПФМ-1 з різних джерел в мережі Інтернет. На цих зображеннях ПФМ-1 зображені при різному освітленні (штучному та природному), зроблені у різний час доби (вдень та вночі), в різних ракурсах, кутах та позиціях (зверху, знизу та з боку), середовищах та шумах. Далі, на другому кроці, ці зображення були попередньо оброблені з використанням платформи Depositphotos, а саме проведено затирання фону та виділено об'єкти, які обробляються, що дозволило отримати еталонні зображення ПФМ-1 (рис. 2) для формування Dataset.

З метою класифікації зображення (після попередньої обробки) визначаються його локальні особливості, тобто контрольні точки. Контрольна точка – точка на зображенні, що задовольняє ряд властивостей [15]:

визначеність – властивість точки виділятися на фоні серед інших сусідніх точок;

стійкість – властивість, при якій зміна яскравості, контрастності та кольорової гами не повинні здійснювати вплив на місце контрольної точки;

інваріантність – властивість точки бути стійкою до повороту, масштабу, зміни ракурсу;

стабільність – властивість, при якій наявність шумів (сніг, листя тощо) не повинна впливати на роботу детектора;

інтерпретованість – представлення точки в прийнятному для роботи форматі;

кількість – необхідна кількість точок, яка допоможе провести зіставлення зображення.

а) *б)*

Рис. 2. Приклад обробки зображень з використанням платформи Depositphotos:
а – зображення ПФМ-1 на місцевості; *б* – еталонне зображення ПФМ-1 (після обробки)

Для обробки зображень з метою їх ідентифікації використовують ряд алгоритмів (методів), які базуються на комп’ютерному баченні. До основних з них належать [13; 15]:

1. SIFT (Scale Invariant Feature Transform) – алгоритм застосовується для аналізу зображень з подальшим виявленням, описуванням та зіставлянням локальних ознак (точок) у зображеннях та створенні дескриптора. Для побудови простору зображень метод SIFT використовує фільтр Гауса, де обчислюється різниця Гауса, а саме попіксельне віднімання зображень в одній октаві. Октаву формують зображення одного масштабу, розмиті фільтром Гауса з різним радіусом розмиття. На цьому етапі забезпечується інваріантність до масштабування, далі визначаються екстремуми, які записуються в список потенційних контрольних точок.

Переваги: інваріантність щодо рівномірного масштабування спрямування та зміни освітлення; часткова інваріантність щодо афінного спотворення. Недоліки: вимогливість до обчислювальної потужності та ліцензійність.

2. SURF (Speeded up Robust Features) – подібний до алгоритму SIFT, але працює за допомогою матриці Гессе, яка досягає свого екстремуму в точках максимальної зміни градієнту яскравості. Градієнт яскравості обчислюється за допомогою фільтрів Хаара. Метод добре виявляє плями, кути, краї, але не інваріантний відносно масштабу. Після знаходження контрольних точок формуються дескриптори.

Переваги: інваріантність відносно повороту та масштабування, різниці яскравості; можливість виявлення багатьох об’єктів на заданому зображенні. Недоліки: складність реалізації та не дуже швидка робота.

3. FAST (Features from Accelerated Segment Test) – алгоритм детектування контрольних точок на зображеннях, заснований на швидкому алгоритмі перевірки великої кількості пікселів навколо потенційної контрольної точки. Визначає потенційні контрольні точки, шукаючи точки, у яких інтенсивність значно вища або нижча від інтенсивності центрального

пікселя. Після вибору центрального пікселя, алгоритм перевіряє інтенсивності пікселів навколо центрального пікселя в чотирьох напрямках. Якщо знайдено стійкий або більше швидких переходів із темного до світлого або зі світлого до темного, центральний піксель вважається контрольною точкою.

Переваги алгоритму: швидкість детектування контрольних точок; ефективність виявлення контрольних точок на складних зображеннях з великою кількістю деталей; простота реалізації. Недоліки: чутливість до шуму та текстур, до орієнтації; не є інваріантним до масштабу; відсутність можливості генерації дескрипторів.

4. BRIEF (Binary Robust Independent Elementary Features) – алгоритм створення локальних дескрипторів для обробки зображень, який використовується для виявлення та опису контрольних точок на зображеннях. Основна ідея методу полягає в тому, щоб створити короткі, бінарні дескриптори, які можуть швидко порівнюватися для відповідності між зображеннями.

Переваги: швидкість порівняння зображень через використання бінарних дескрипторів та прості порівняння пікселів; ефективність в обробці великих обсягів даних; простота реалізації; малий обсяг пам'яті, що дозволяє ефективно зберігати та обробляти багато дескрипторів одночасно. Недолік: низька робучість; низька точність збігу; чутливість до шуму.

5. ORB (Oriented FAST and Rotated BRIEF) – алгоритм, який обчислює орієнтацію ключової точки зображення та бінарне узгодження з вже отриманих орієнтацій. Є ефективним варіантом SIFT та SURF, який забезпечує порівняну ефективність та швидкість. ORB розроблений на базі детектора контрольних точок і дескриптора.

6. BRISK (Binary Robust Invariant Scalable Keypoints) – продовження алгоритмів SIFT і SURF, але з подальшим удосконаленням складових FAST. BRISK складає дескриптор шляхом об'єднання результатів порівняльних тестів яскравості, враховуючи зазначений масштаб та розташування кожної контрольної точки, щоб забезпечити інваріантність до обертання та одержати адаптовано нормалізовані дескриптори. Дескриптор BRIEF забезпечує ідентифікацію однакових ділянок зображення, отриманих з різних точок зору.

Переваги: вища швидкодія, завдяки спрощенню процесу оброблення та побудови і використання дескриптора бінарного типу; можливість виявлення якісно вищих контрольних точок на зображенні та досягнення стійкості до геометричних перетворень.

Порівняльний аналіз характеристик зазначених алгоритмів наведено в таблиці 1.

Таблиця 1

Порівняльний аналіз характеристик зазначених алгоритмів

Найменування алгоритму / Характеристика	FIST	SURF	FAST	BRIEF	ORB	BRISK
Швидкість	висока	висока	висока	висока	середня	середня
Інваріантність до масштабу	відсутня	відсутня	відсутня	відсутня	наявна	наявна
Інваріантність до обертання	відсутня	відсутня	відсутня	відсутня	наявна	наявна
Витримка до шуму	висока	висока	низька	середня	висока	висока
Діапазон освітленості	середній	широкий	високий	середній	високий	високий
Діапазон детектованого кута	великий	широкий	середній	середній	великий	великий
Робоча пам'ять	середня	висока	низька	низька	середня	середня
Навчальна складність	висока	висока	низька	середня	середня	середня
Простота реалізації	середня	середня	висока	висока	середня	середня
Ресурсоємність	низька	висока	низька	низька	середня	середня
Відносна надійність	висока	висока	висока	середня	висока	висока
Популярність у використанні	середня	висока	висока	середня	висока	висока

Проаналізувавши зазначені типи алгоритмів та результати їх використання при ідентифікації зображень, автори дійшли висновку, що у подальшій роботі для ідентифікації ВНП доцільно використовувати комбінування ORB та BRISK (третій крок методики), оскільки вони мають вищу швидкодію, побудовані по бінарному типу, що спрощує процес обчислення та знижує затрати. В цілому, їх комбінування дає кращі результати, оскільки враховується спотворення кольору, погане освітлення, наявність шумів та різних проєкцій зображень.

Розглянемо математичне представлення застосування ORB та BRISK для класифікації зображень ВНП [16]. Припустимо, що у нас є певна множина $Z = \{Z^j\}_{j=1}^{J}$ ознак бази зображень, яку складають еталони (Z^j – еталон, J – число класів). Під час попередньої обробки множину Z поділяємо на певну кількість k кластерів $M = \{M_i\}_{j=1}^{k}$, так що $M_i \cap M_d = \emptyset, Z = \cup M_i, M = Z$. Кожний з кластерів *Mi* має свій центр m_i.

Кластерування полягає у відображенні еталонного набору контрольних точок в себе Z→Z, де кожна контрольна точка належить до одного з кластерів.

Створена навчальна множина даних $W = \{x | x \in R^n\}$, $W \subseteq R^n$, яка складається з набору контрольних точок дескрипторів. Після цього можемо збудувати кортеж k опорних векторів $M = \{m_i\}_{i=1}^{k}, m_i \in R^n, i = 1, \dots, k$.

Для сформованого набору M апроксимація довільного вектора $x \in W$ конкурентним способом означає визначення номера υ найближчого до нього (у метриці ρ(x, m_i)) вектора $m_\upsilon \in M$ у просторі модельних векторів відповідно до виразу (1):

$$\upsilon = \arg \min_{i=1,\dots,k} \rho(x, m_i). \quad (1)$$

Для традиційного варіанта онлайн-навчання, якщо дескриптори контрольних точок $x[t] \in W$ поступають по черзі, центр m_υ кластера, який став переможцем в (1), на кроці t = 1, 2, ..., s навчання коригується відповідно до виразу (2):

$$m_\upsilon[t+1] = m_\upsilon[t] + \alpha[t](x[t] - m_\upsilon[t]), \quad (2)$$

де $s = cardW$ – загальне число контрольних точок бази зображень навчальної вибірки, $\alpha[t]$ швидкість навчання, яка задається, водночас $\alpha[t] \to 0$ при $t \to s$.

Критерій похибки оцінюється якістю кластерування – усередненої суми квадратів відхилень від центрів сформованих кластерів відповідно до виразу (3):

$$E = \frac{1}{2} \sum_{i=1}^{k} \sum_{\upsilon=1}^{s(i)} \rho^2(x_\upsilon, m_i), \quad (3)$$

де s(i) – потужність *i*-го кластера [3].

У випадку, коли число k кластерів дорівнює числу еталонів бази зображень, кластерування виконує класифікацію на k класів у просторі W. Множина дескрипторів контрольних точок, що описує довільний візуальний об'єкт, може бути класифікована як один із еталонів зображення.

Наряду з класифікацією, на підставі мережі Кохонена, розглянемо евристичний підхід бінарного аналізу, суть якого полягає в наступному.

1. На підставі логічного правила визначимо вектор m_i центра класу для кожного з еталонів множини Z_i дескриптора BRISK, який має бінарний вид відповідно до виразу (4):

$$m_i(b) = \begin{cases} 1, \text{при} \sum_{d=1}^{s(i)} x_d(b) \geq s(i)/2, \\ 0, \text{при} \sum_{d=1}^{s(i)} x_d(b) < s(i)/2, \end{cases} \quad (4)$$

$$x_d \in Z^i \; b = \overline{1,12},$$

де $x_d(b)$ – біт з номером b для дескриптора з номером d в описі еталону зображення.

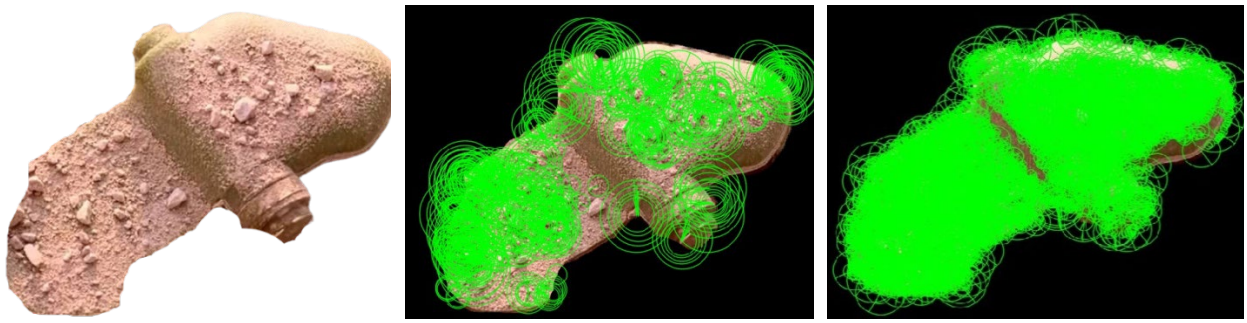
Таким чином, значення центрів для кожного з бітів m_i для i-го класу визначається переважною більшістю значень відповідних бітів всіх дескрипторів контрольних точок, які належать еталону Z^i.

2. Аналізуємо елементи $x \in Z$ спільного вмісту структурних описів бази еталонів (навчальна вибірка) шляхом віднесення їх до відповідного класу з використанням конкурентного способу (1). Тут у якості $\rho(x, m_i)$ застосуємо метрику Хемінга, що визначає кількість розбіжних бітів для двох двійкових послідовностей однакової довжини (детекторів контрольних точок) відповідно до виразу (5):

$$\rho(x, m_i) = \sum_{b=1}^{512} |x(b) - m_i(b)|, \quad (5)$$

У результаті оброблення для кожного еталона отримуємо його кластерне подання $H[Z^i] = (h_1^i, \ldots, h_j^i)$, де h_a^i – цілі числа, які відповідають розподілу елементів множини Z^i за класами еталонів.

Етап попередньої обробки, що ґрунтується на бінарному аналізі, можна розглядати як альтернативний метод хешування для набору дескрипторів контрольних точок еталону. Для аналізу множини дескрипторів ORB можна використати аналогічну обробку. В якості прикладу, на рисунку 3 відображено формування контрольних точок алгоритмами ORB та BRISK для ПФМ-1.



а) *б)* *в)*

Рис. 3. Приклад формування контрольних точок для ПФМ-1:
а – еталон ПФМ-1; *б* – дескриптор ORB; *в* – дескриптор BRISK

Алгоритм ORB обробляє зображення швидше за BRISK, що дає перевагу під час використання в реальному часі. Проте BRISK формує значно більше контрольних точок та аналізує зображення детальніше, що дозволяє забезпечувати більш точніші розпізнавання об'єктів. Відповідно, в процесі комбінування цих методів можливо співставляти контрольні точки ORB та BRISK, що дозволяє забезпечити створення продуктивного методу ідентифікації зображення на основі бітового оброблення та побудову класифікаційних рішень щодо відповідності зображення еталону.

Четвертий крок. В якості програмного середовища для формування Dataset (як зазначалося, бази даних еталонних зображень для ПФМ-1) з використанням зазначених методів ідентифікації ВНП, обрано онлайн-платформу Google Colab, яка дозволяє писати та запускати програмний код мовою Python у хмарі, використовуючи лише браузер. Для цього, імпортуємо необхідні бібліотеки в середовище Google Colab, такі як: OpenCV (для обробки зображень), NumPy (для роботи з масивами даних), Pandas (для роботи з даними у вигляді таблиць) та Matplotlib (для візуалізації даних).

На п'ятому кроці даємо команду на створення об'єктів алгоритмами ORB та BRISK для їх ініціалізації та формування «нульового» Dataset в Google Colab. З метою наповнення Dataset завантажується частина попередньо оброблених зображень (перший та другий крок). Після їх завантаження, об'єкти ORB та BRISK проводять ініціалізацію зображень, позначають на зображеннях контрольні точки та формують дескриптори (у форматі CSV файлу), які і є наповненим Dataset. Якщо Dataset сформовано та збережено коректно, отримуємо повідомлення «Datasets saved successfully», в іншому випадку – код видаватиме помилку з вказівкою місця хибної роботи (рис. 4).

```
def extract_keypoints_and_descriptors(image, detector):
    gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
    keypoints, descriptors = detector.detectAndCompute(gray, None)
    return keypoints, descriptors

def save_dataset_to_csv(dataset, csv_filename):
    with open(csv_filename, mode='w', newline='') as file:
        writer = csv.writer(file)
        writer.writerow(['Image Path', 'Keypoints', 'Descriptors'])
        for data in dataset:
            keypoints = [keypoint.pt for keypoint in data['keypoints']]
            descriptors = data['descriptors']
            writer.writerow([data['image_path'], keypoints, descriptors])
```

```
# Ініціалізація детекторів ORB та BRISK
orb = cv2.ORB_create()
brisk = cv2.BRISK_create()

# Формування датасету для ORB та BRISK
orb_dataset = []
brisk_dataset = []

# Завантаження зображень
uploaded = files.upload()

# Збереження датасету в CSV
save_dataset_to_csv(orb_dataset, 'orb_dataset.csv')
save_dataset_to_csv(brisk_dataset, 'brisk_dataset.csv')

print("Datasets saved successfully.")
```

Рис. 4. Фрагмент програмного коду для ініціалізації детекторів та формування «нульового» Dataset

Для тестування (на шостому кроці) завантажуються вибіркові контрольні зображення (частина зображень, які вже збережені в Dataset та інші, які були попередньо оброблені, але не завантажувалися на попередньому кроці), які порівнюються зі збереженими даними Dataset. Якщо результат ідентифікації зображень є позитивним, виводиться ідентифіковане зображення (за допомогою функції cv2_imshow в середовищі Google Colab) та повідомлення, що зображення є «True». Якщо зображення не ідентифікується, то виводиться повідомлення

«False». Отримані результати підтверджують, що сформований Dataset пройшов тестування успішно та може бути використаний як вхідний шар для навчання нейронної мережі.

Висновки. В статті було розглянуто характерні локальні ознаки ВНП, за якими їх можна ідентифікувати. Аналіз методів ідентифікації зображень показав, що доцільно комбінувати методи ORB та BRISK, це дає більш високу точність при класифікації зображень, оскільки враховується спотворення кольору, погане освітлення, наявність шумів та різних проєкцій.

Запропонована методика містить кроки щодо формування Dataset на прикладі ПФМ-1, починаючи з пошуку, вибору алгоритмів ідентифікації зображень до розроблення програмного коду та тестування сформованого Dataset.

За допомогою платформи Google Colab реалізовано коректну роботу алгоритмів ідентифікації зображень, формування та тестування зазначеного Dataset. Сформований Dataset може бути використаний як вхідний шар для навчання нейронної мережі для ідентифікації ВНП. Подальшим напрямком досліджень є розширення даних Dataset завдяки збільшенню кількості зображень в різних середовищах (з використанням різного програмного забезпечення та порівняння отриманих результатів) та навчанню нейромережі для ідентифікації більшої кількості зображень ВНП.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Мірошниченко Б. Десятки років та мільярди доларів. Коли розмінують українські поля та міста? // Економічна правда. 2023. URL: https://www.epravda.com.ua/publications/2023/03/7/697737/.

2. Аналіз виконання робіт щодо очищення території України від вибухонебезпечних предметів по роках // Державна служба України з надзвичайних ситуацій. 2023. URL: https://dsns.gov.ua/protiminna-diyalnist/gumanitarne-rozminuvannya/analiz-vikonannya-robit-schodo-ochischennya-teritoriyi-ukrayini-vid-vibuhonebezpechnih-predmetiv-po-rokah.

3. Собенко Н. Повне розмінування України може зайняти понад 750 років – WP // Суспільні Новини. 2023. URL: https://suspilne.media/534815-povne-rozminuvanna-ukraini-moze-zajnati-ponad-750-rokiv-wp/.

4. Чередниченко О. Ю., Паламарчук Н. А., Шемендюк О. В., Мартинюк В. В. Синтез системи виявлення вибухонебезпечних предметів на базі безпілотного літального апарата // Системи і технології зв'язку, інформатизації та кібербезпеки. Випуск 3, 2023. С. 163–170. ISSN 2786-6610. DOI: 10.58254/viti.3.2023.

5. Невлюдов І. Ш., Янушкевич Д. А., Толкунов І. О., Попов І. І., Іванець Г. В. Обґрунтування необхідності створення робототехнічних комплексів для гуманітарного розмінування. Problems of Emergency Situations (Проблеми надзвичайних ситуацій). Серія: Civil Security. Вип. 2 (38). С. 17–38. DOI: 10.52363/2524-0226-2023-38-2.

6. Рибалка А. В., Скорлупін О. В., Подорожняк А. О. Аналіз можливості застосування технологій штучного інтелекту для виявлення вибухонебезпечних предметів та подальшого гуманітарного розмінування // Тези доповідей. Інформаційна безпека та інформаційні технології: VI Всеукраїнська науково-практична конференція молодих учених, студентів і курсантів, 30 листопада 2023 року. Львів, ЛДУ БЖД, 2023. С. 404–406.

7. Янушкевич Д. А., Іванов Л. С., Толкунов І. О. Креативні підходи управління якістю у сфері гуманітарного розмінування із застосуванням робототехнічних систем // Тези доповідей. Збірник матеріалів V форуму «Автоматизація, електроніка та робототехніка. Стратегії розвитку та інноваційні технології», AERT-2023. Х.: ХНУРЕ, каф. МТС, 29–30 листопада 2023 р. 149 с. С. 55–58.

8. Єремеєв О., Васильєва І., Макарічев В., Рубель О., Лі Ф., Чернова Г., Коваленко Б., Ворзель Б., Лукін В. Деякі аспекти обробки багатоканальних даних дистанційного зондування з використанням нейромереж // Prospective global scientific trends: Informatics, Architecture, Innovative technology, Ecology: Monographic series «European Science». 2023. Book 19. Part 2. С. 7–60. DOI: 10.30890/2709-2313.2023-19-02-001.

9. Jasper Baur, Gabriel Steinberg, Alex Nikulin, Kenneth Chiu, Timothy S. de Smet. Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines // Remote sensing. 2020. URL: https://www.mdpi.com/2072-4292/12/5/859#.

10. Szeliski R. Computer vision: algorithms and applications. Springer Science & Business Media. 2010.

11. Shnain, Noor Abdalrazak; Hussain, Zahir M.; LU, Song Feng. A feature-based structural measure: an image similarity measure for face recognition. Applied Sciences, 2017, 7.8: 786.

12. Gorokhovatskyi, Volodymyr; Tvoroshenko, Iryna. Image classification based on the Kohonen network and the data space modification. 2020.

13. Leutenegger, Stefan; Chli, Margarita; Siegwart, Roland Y. BRISK: Binary robust invariant scalable keypoints. In: 2011 International conference on computer vision. Ieee, 2011. P. 2548–2555.

14. Опришко О. О., Пасічник Н. О., Шворов С. А., Кіктєв М. О., Дудник А. О., Сович В. І. Дослідження перспектив дистанційного моніторингу мін на полях з використанням тепловізійного знаряддя // Енергетика і автоматика. 2023. № 6. С. 74–89. DOI: http://dx.doi.org/ 10.31548/energiya6(70).2023.074.

15. Тимчишин Р. М., Волков О. Є., Господарчук О. Ю., Богачук Ю. П. Сучасні підходи до розв’язання задач комп’ютерного зору // Управляющие системы и машины. 2018. № 6. С. 46–73. Бібліогр.: 48 назв. – укр.

16. Gorokhovatsky V. O. Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення / V. O. Gorokhovatsky, D. V. Pupchenko, K. G. Solodchenko // Системи управління, навігації та зв’язку. Збірник наукових праць. Полтава: ПНТУ, 2018. Т. 1 (47). С. 93–98. DOI: https://doi.org/10.26906/SUNZ.2018.1.093.

DOI: 10.58254/viti.5.2024.12.130

УДК. 621.391

канд. техн. наук, професор Радзівілов Г. Д. ORCID: 0000-0002-6047-1897 (ВІТІ ім. Героїв Крут)
канд. техн. наук Ільїнов М. Д. ORCID: 0009-0008-6945-3354 (ВІТІ ім. Героїв Крут)
Хоменко П. В. ORCID: 0000-0002-8543-1971 (ВІТІ ім. Героїв Крут)

ОСОБЛИВОСТІ РОЗРАХУНКУ ДІАГРАМИ НАПРАВЛЕНОСТІ КІЛЬЦЕВИХ АНТЕННИХ РЕШІТОК, ЯКІ ВИКОНАНІ НА ПІВХВИЛЬОВИХ СИМЕТРИЧНИХ ВІБРАТОРАХ, РОЗМІЩЕНИХ НАД ЦИЛІНДРИЧНОЮ ПОВЕРХНЕЮ

Застосування антенно-фідерних пристроїв, особливо встановлених на рухомих об'єктах, вказує на необхідність модернізації та розробки антенних пристроїв нового типу для підвищення ефективності функціонування системи радіозв'язку в умовах активного радіоелектронного подавлення. Одним із варіантів забезпечення завадозахисту в каналах зв'язку з рухомими об'єктами є застосування вузьконаправлених Smart антен із керованою діаграмою направленості. Smart антени, які ще мають назву інтелектуальні антени, це один з різновидів фазованих антенних решіток. Smart антени використовують набір випромінюючих елементів, побудованих у формі решіток. Найбільше поширення отримали кільцеві антенні решітки, які відрізняються компактністю та великими функціональними можливостями.

Актуальність завдання по компонуванню Smart антен полягає у зменшенні загальної кількості випромінювачів без погіршення параметрів антени. Це завдання особливо важливе для антен з круговими діаграмами випромінювання в системах зв'язку з рухомими об'єктами.

Щоб зменшити кількість вібраторів у таких антенах, зазвичай знаходять максимально велику відстань між вібраторами, при якій рівномірність β діаграми направленості не знижується нижче заданого рівня. Щоб знайти допустиму відстань, можна розрахувати залежність β від радіуса решітки безпосередньо при розрахунку азимутальної діаграми направленості антени. Для вирішення цього завдання було розглянуто зовнішні характеристики кільцевих антенних решіток, які виконані на півхвильових симетричних вібраторах в ідеальних умовах та вібраторах, розміщених над циліндричною поверхнею. Отримані результати наглядно показують можливості компонування кільцевих антенних решіток залежно від радіуса антенної решітки та радіуса циліндричної опори, та при яких умовах геометричні розміри кільцевих антенних решіток будуть зберігати коефіцієнт рівномірності діаграми направленості.

Ключові слова: *діаграма направленості, Smart антена, фазована антенна решітка, кільцева антенна решітка, коефіцієнт нерівномірності діаграми направленості.*

H. Radzivilov, M. Ilyinov, P. Khomenko Features computation diagram of annular direction antenna arrays what made on half-wave vibrators located above cylindrical surface.

The use of antenna-feeder devices, especially installed on moving objects, indicates the need to modernize and develop new type antenna devices to increase the efficiency of the radio communication system in conditions of active radio electronic suppression. One of the options for providing interference protection in communication channels with moving objects is the use of narrowly directed Smart antennas with a controlled directional pattern. Smart antennas, which are also called intelligent antennas, are one of the varieties of phased antenna arrays. Smart antennas use a set of radiating elements built in the form of grids. The most widespread are the ring antenna arrays, which are characterized by their compactness and great functionality.

The relevance of the task of designing Smart antennas is to reduce the total number of emitters without degrading the antenna parameters. This task is particularly important for antennas with circular radiation patterns in communication systems with moving objects.

The obtained results clearly show the possibilities of ring antenna arrays layout, depending on the radius of the antenna array and the radius of the cylindrical support, and under which conditions the geometric dimensions of the ring antenna arrays will preserve the coefficient of uniformity of the directional pattern. In order to reduce the number of vibrators in such antennas, the maximum distance between the vibrators is usually found, at which the uniformity of β directional diagrams does not decrease below a given level. To find the permissible distance, you can calculate the dependence of β on the radius of the grating directly when calculating the azimuthal directional diagrams of the antenna. To solve this problem, the external characteristics of the ring antenna arrays performed on half-wave symmetric vibrators in ideal conditions and vibrators placed over a cylindrical surface were considered.

Keywords: *directional pattern, Smart antenna, phased antenna array, annular antenna array, coefficient of unevenness of the directional diagram.*

Постановка завдання.

Аналіз досвіду застосування антенно-фідерних пристроїв, особливо встановлених на рухомих об’єктах, вказує на необхідність модернізації та розробки антенних пристроїв нового типу для підвищення ефективності функціонування системи радіозв’язку в умовах активного радіоелектронного подавлення. Одним з варіантів забезпечення завадозахисту в каналах зв’язку з рухомими об’єктами є застосування вузьконаправлених Smart антен з керованою діаграмою направленості (ДН). Smart антени, які ще мають назву інтелектуальні антени, це один з різновидів фазованих антенних решіток (ФАР) [1].

Smart антени – це найбільш перспективніша технологія, що покращує якість прийнятого сигналу завдяки збільшенню ємності системи, зменшує багатопроменевість та інтерференцію між сусідніми каналами. Найчастіше на базі інтелектуальних антен використовують формування вузької діаграми направленості. Вузька діаграма направленості дозволяє сфокусувати енергію сигналу у визначеному напрямку (зазвичай назустріч приймальному пристрою), що збільшує відношення сигнал/шум, або, навпаки, сформувати провал у діаграмі направленості в напрямку завади, що особливо критично для придушення радіоелектронних джерел противника. При вузькому антенному промені зменшуються також вплив завад, збільшується відношення сигнал/шум, і таким чином підвищується ефективність використання спектра частот [2]. Smart антени використовують набір випромінюючих елементів, побудованих у формі решіток [3]. Слід зазначити, що найбільше поширення отримали кільцеві антенні решітки (КАР), які відрізняються компактністю та великими функціональними можливостями.

З метою можливого застосування КАР, виникає завдання в проведенні аналізу електричних характеристик антени залежно від їхніх геометричних розмірів.

Аналіз публікацій за темою дослідження.

У роботі [4] проведено аналіз завдань, які виникають при впровадженні технології адаптивного діаграмоутворення в мобільних радіомережах в умовах активної радіоелектронної протидії, основними з яких є наступні: визначення просторових координат власної станції, кореспондентів мережі та постановника навмисних завад; розрахунок оптимального кута орієнтації основного пелюстка діаграми направленості антени при прийомі сигналу від сусідніх кореспондентів з урахуванням взаємного розташування радіостанцій та постановника завад. Але приклади діаграм направленості кільцевої антенної решітки представлено в ідеальних умовах, без врахування реальної побудови КАР, яке суттєво буде впливати на масо-габаритні та електричні характеристики антени в цілому.

У роботі [5] розкрито особливості розрахунку зовнішніх характеристик кільцевих антенних решіток на основі низькопрофільних випромінювачів над циліндричною поверхнею. Розглянуто вирішення задачі про розрахунок характеристики направленості в азимутальній площині. Але наглядно представлено тільки приклади розрахунку двох- та трьохелементної антенної решітки, що не дає повної оцінки ефективності компонування кільцевих антенних решіток.

У статті [6] представлено удосконалений метод адаптивного прийому сигналів від рухомих джерел, який дозволяє знизити обсяг обчислень порівняно з існуючим, коли ваговий вектор з адаптивними антенними решітками є оптимальним і вибирається як власний вектор сигнальної кореляційної матриці, що визначає його новизну, при цьому можливо забезпечити досить високу ефективність прийому сигналів, однак автори не зосереджують увагу на форми ДН та їхню залежність від компонування антени.

У науковій праці [7] представлено методику адаптивного управління параметрами багатоантенних систем військового радіозв’язку з активними фазованими антенними решітками, проте в роботі не розкрито залежність процесу діаграмоутворення від зміни параметрів управління антени.

У роботі [8] результат наукового дослідження показав, що одним із найбільш перспективних підходів вирішення задачі, що підкреслює актуальність, є підвищення ефективності прийому сигналу при використанні адаптивних антенних решіток і відповідних методів просторової обробки сигналів, проте конфігурація адаптивних антенних решіток передбачається довільною.

Тому, метою статті є дослідження можливих варіантів конфігурації Smart антен на рухомих об’єктах з мінімальними масо-габаритними показниками.

Виклад основного матеріалу.

Актуальність завдання по компонуванню Smart антен полягає у зменшенні загальної кількості випромінювачів без погіршення параметрів антени. Це завдання особливо важливе для антен з круговими діаграмами випромінювання в системах зв’язку з рухомими об’єктами. Часто такі антени виконані у вигляді кругових рівновіддалених вертикальних вібраторах, розташованих навколо висотних опор (місце кріплення антени). Великі відносно робочої довжини хвилі поперечні розміри опор потребують використання великої кількості вібраторів для отримання в горизонтальній площині близької до кругової ДН антени. Щоб зменшити кількість вібраторів у таких антенах, зазвичай знаходять максимально велику відстань між вібраторами, при якій рівномірність β ДН не знижується нижче заданого рівня. Щоб знайти допустиму відстань, можна розрахувати залежність β від радіуса решітки безпосередньо при розрахунку азимутальної ДН антени за формулою β(kr) = |Fmax(kr)| / |Fmin(kr)|). При цьому важливо мінімізувати масо-габаритні показники та зберегти жорсткість конструкції.

Для досягнення поставленої мети необхідно вирішити ряд часткових задач, а саме:

− розрахувати геометричні розміри кільцевої антенної решітки, виконаної на півхвильових симетричних вібраторах (рис. 1), та кільцевої антенної решітки, виконаної на півхвильових симетричних вібраторах, розміщеної над циліндричною поверхнею (рис. 3);

− за допомогою середовища математичного моделювання MathCad 15 визначити характеристики антен (ДН та коефіцієнт нерівномірності ДН).

Розрахунок кільцевої антенної решітки, виконаної на півхвильових симетричних вібраторах

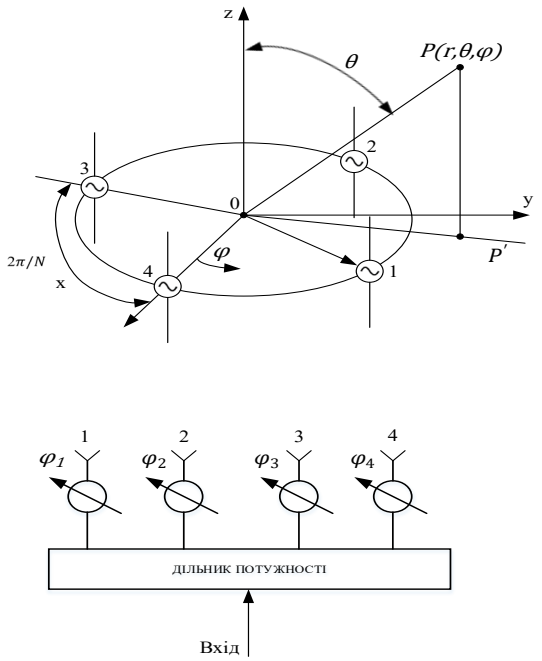


Рис. 1. Кільцева антенна решітка, виконана на півхвильових симетричних вібраторах

Геометрична симетрія в КАР забезпечить рівномірне розміщення елементів антенної решітки по кільцю, з радіусом r, з координатами $\varphi_n = (n-1)\frac{360}{N}$, де N – кількість випромінюючих елементів. Електрична симетрія буде реалізована внаслідок збудження випромінюючих елементів, амплітудно-фазовий розподіл яких ідентичний в секторах, співпадаючий з кутовими відстанями між суміжними випромінювачами $I_{im} = \frac{1}{\sqrt{N}} l^{jm(i-1)\frac{360}{N}}$ де, $m = 0; \pm 1, {}^{N}/_{2}$ – номер моди (типу) розподілу струму, наприклад, $m = 0$ – синфазне збудження елементів решітки, $m = 1$ – квадратурне ($N = 4$).

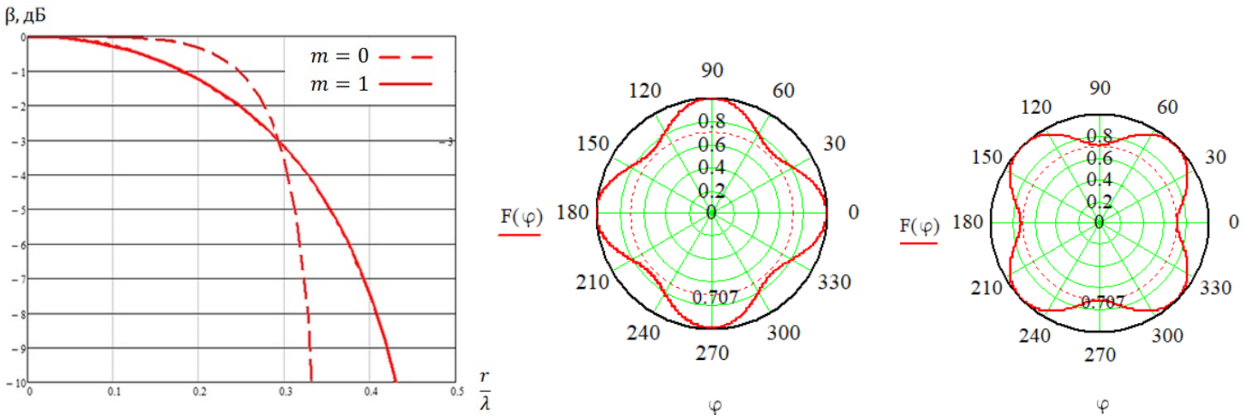
Просторова напруженість електричного поля для КАР з рівномірним розташуванням N елементів по колу з радіусом r [9], визначимо за виразом (1):

$$f(\varphi) := \left| \sum_{i=1}^{N} \left(e^{i\cdot\psi(i)} \cdot F0(\varphi - \phi(i)) \cdot e^{i\cdot kR\cdot\sin(\theta)\cdot\cos(\varphi-\phi(i))} \right) \right|. \quad (1)$$

Коефіцієнт направленої дії (КНД) КАР кількісно визначимо через просторовий розподіл напруженості електричного поля [10] за виразом (2):

$$D := \frac{4\pi}{\int_0^{2\pi} \int_0^{\pi} F(\varphi)^2 \cdot \sin(\theta) \; d\theta \; d\phi}. \quad (2)$$

Проведемо аналіз зовнішніх характеристик КАР з урахуванням $R = 0{,}29\lambda$ (радіус решітки); $l = 0{,}25\lambda$ (довжина половини вібратора). Формування поля випромінювання в азимутальній площині визначається не тільки амплітудно-фазовим розподілом струмів у випромінюючих елементах КАР ($m = 0$, $m = 1$), а й її геометричними розмірами, що наочно показано на рисунку 2.



а) *б)* *в)*

Рис. 2. ДН при радіусі КАР $R = 0{,}29$:
а – коефіцієнт нерівномірності діаграми направленості; *б* – синфазне збудження випромінюючих елементів мода ($m = 0$); *в* – квадратурне збудження випромінюючих елементів мода ($m = 1$)

Отримані теоретичні результати досліджень зовнішніх характеристик КАР, виконаних на вертикальних симетричних вібраторах, дозволяє зробити наступні висновки:

− незалежно від моди збудження випромінюючих елементів КАР максимальний радіус її побудови дорівнює 0,29 λ;

− при компонуванні КАР, для яких масо-габаритні показники мають важливе значення, необхідно застосовувати синфазне збудження випромінюючих елементів ($m = 0$);

− коефіцієнт нерівномірності ДН в азимутальні площині для моди (m = 0) дорівнює 0 дБ практично до значення радіуса КАР 0,2 λ.

Розрахунок кільцевої антенної решітки, виконаної на півхвильових симетричних вібраторах, розміщених над циліндричною поверхнею

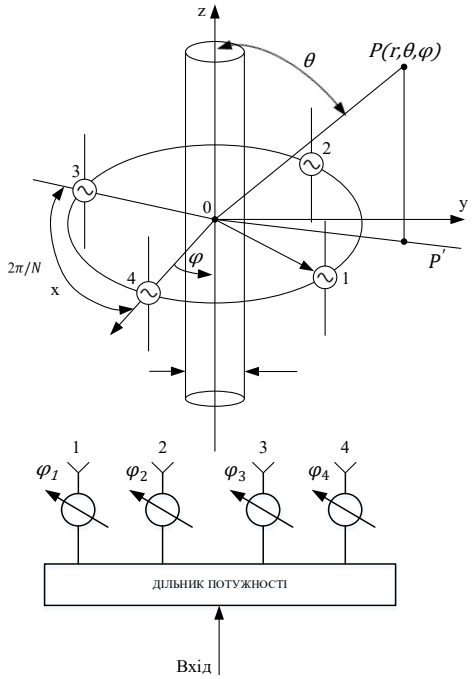


Рис. 3. Кільцева антенна решітка, виконана на півхвильових симетричних вібраторах, розміщених над циліндричною поверхнею

Розглянемо інший варіант побудови КАР у вигляді вертикальних симетричних вібраторів, розташованих біля відбиваючої поверхні.

Така будова КАР дозволить зменшити вплив між випромінюючими елементами, а також розширити функціональні можливості антенної решітки по формуванню поля випромінювання в азимутальні площині завдяки зміні розмірів радіуса відбиваючої поверхні. Характеристики кільцевої антенної решітки з урахуванням дифракції поля випромінюючих елементів, розташованих над циліндричною поверхнею (рис. 3), можна визначити за наступним виразом [11]:

$$(\varphi,\theta) := \left|\frac{\cos(ki\cdot\cos(\theta))-\cos(kI)}{\sin(\theta)}\sum_{i=1}^{N}\sum_{n=0}^{10}\left[\varepsilon(n)i^{n}e^{i\cdot\psi(i)}\left(Jn(n,kr\cdot\sin(\theta))-\frac{Jn(n,ka\cdot\sin(\theta))}{H2(n,ka\cdot\sin(\theta))}H2(n,kr\cdot\sin(\theta))\right)\cos\left[n\left[\varphi-\frac{2\pi}{N}(i-1)\right]\right]\right]\right|, \quad (3)$$

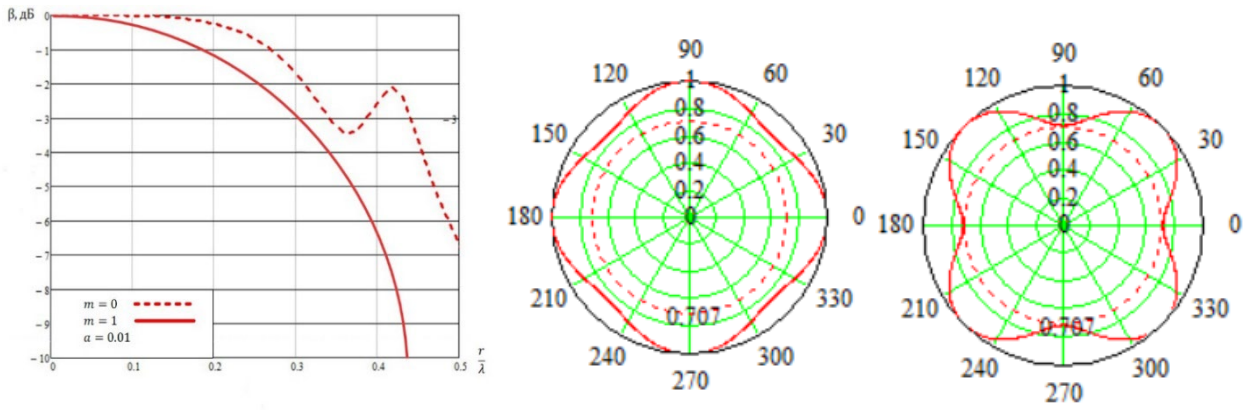
де Jn, H_n^2 – функція Бесселя і Ганкеля n-го порядку відповідно;

N – кількість випромінюючих елементів в КАР;

n – кількість елементів в рядку;

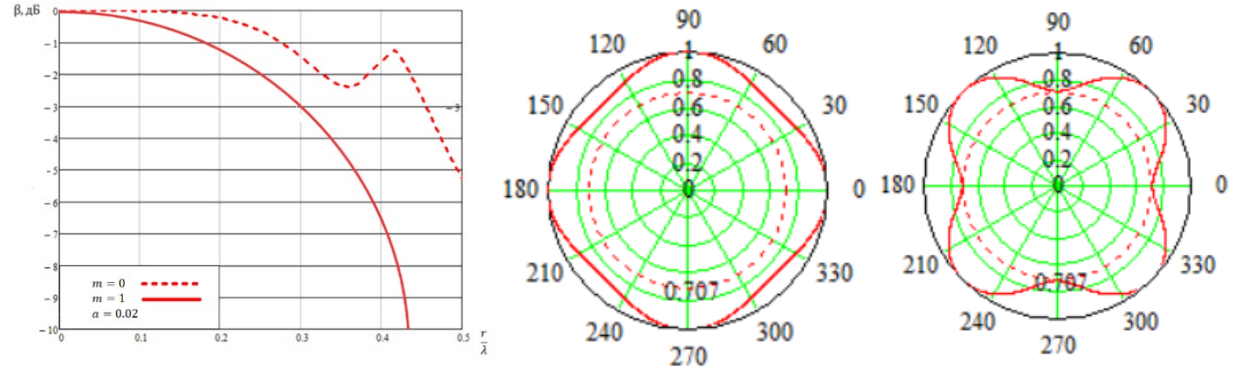
φ, θ – кут розташування n-го вібратора.

Проведемо аналіз зовнішніх характеристик КАР з урахуванням впливу циліндричної поверхні на характеристики випромінювання в азимутальні площині, для різних значень радіуса циліндричної поверхні (a = 0,01λ, рис. 4; a = 0,02λ, рис. 5; a = 0,05λ, рис. 6; a = 0,1λ, рис. 7; R = 0,29λ (радіус решітки); l = 0,25λ (довжина половини вібратора).



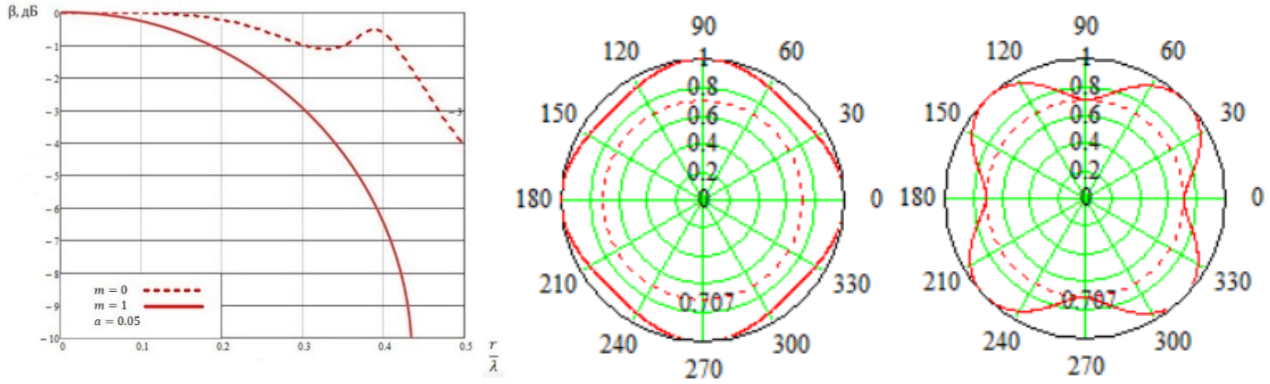
а) *б)* *в)*

Рис. 4. ДН при радіусі КАР R = 0,29, радіус циліндричної поверхні a = 0,01λ:
а – коефіцієнт нерівномірності діаграми направленості; *б* – синфазне збудження випромінюючих елементів мода (m = 0); *в* – квадратурне збудження випромінюючих елементів мода (m = 1)



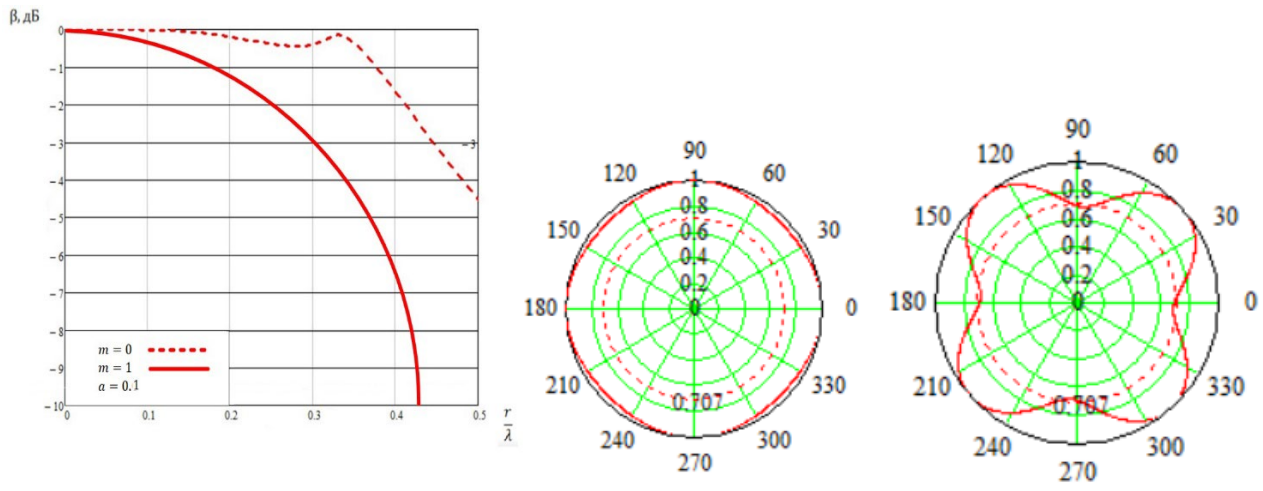
а) *б)* *в)*

Рис. 5. ДН при радіусі КАР R = 0,29, радіус циліндричної поверхні a = 0,02λ:
а – коефіцієнт нерівномірності діаграми направленості; *б* – синфазне збудження випромінюючих елементів мода (m = 0); *в* – квадратурне збудження випромінюючих елементів мода (m = 1)



а) *б)* *в)*

Рис. 6. ДН при радіусі КАР R = 0,29, радіус циліндричної поверхні a = 0,05λ:
а – коефіцієнт нерівномірності діаграми направленості; *б* – синфазне збудження випромінюючих елементів мода (m = 0); *в* – квадратурне збудження випромінюючих елементів мода (m = 1)



а) *б)* *в)*

Рис. 7. ДН при радіусі КАР R = 0,29, радіус циліндричної поверхні а = 0,1λ:
а – коефіцієнт нерівномірності діаграми направленості; *б* – синфазне збудження випромінюючих елементів мода (m = 0); *в* – квадратурне збудження випромінюючих елементів мода (m = 1)

Результати проведеного аналізу зовнішніх характеристик КАР з урахуванням впливу на характеристики випромінювання в азимутальні площині дозволяє зробити наступні висновки:

коефіцієнт нерівномірності ДН в азимутальні площині при квадратурному збудженні випромінюючих елементів (мода, m = 1) зі збільшенням радіуса циліндричної поверхні не змінюється. Максимальний радіус КАР по коефіцієнту нерівномірності більше –3 дБ дорівнює 0,3λ при всіх значеннях радіуса циліндричної відбиваючої поверхні;

при синфазному збуджені випромінюючих елементів КАР (мода, m = 0) коефіцієнт нерівномірності ДН в азимутальні площині змінюється залежно від радіуса циліндричної поверхні. Так, при радіусі циліндричної поверхні а = 0,01λ максимальний радіус КАР складає r = 0,34λ, а при а = 0,05λ, r = 0,47λ.

Отримані результати наглядно показують можливості компонування КАР з ненаправленим випромінюванням в азимутальні площині. Так, при застосуванні таких антенних пристроїв на рухомих об’єктах необхідно врахувати місце кріплення антени, яке буде впливати та формування поля випромінювання азимутальні площині та на геометричні розміри антени.

Висновок. Одним з варіантів забезпечення завадозахисту в каналах зв’язку з рухомими об’єктами є застосування вузьконаправлених Smart антен з керованою діаграмою направленості (ДН).

Вузька діаграма направленості дозволяє сфокусувати енергію сигналу у визначеному напрямку, що збільшує відношення сигнал/шум, або, навпаки, сформувати провал в діаграмі направленості в напрямку завади, що особливо критично для придушення радіоелектронних джерел противника. Smart антени використовують набір випромінюючих елементів, побудованих у формі решіток. Найбільше поширення отримали кільцеві антенні решітки, які відрізняються компактністю та великими функціональними можливостями.

У статті за допомогою середовища математичного моделювання MathCad 15 проведено розрахунки коефіцієнта нерівномірності діаграми направленості 4-елементної кільцевої антенної решітки, виконаної на півхвильових симетричних вібраторах в ідеальних умовах та за умови кріплення антени на опорах з різними радіусами.

Отримані результати наглядно показують можливості компонування КАР з ненаправленим випромінюванням в азимутальні площині залежно від радіуса антенної решітки та радіуса циліндричної опори, та при яких умовах геометричні розміри кільцевих антенних решіток будуть зберігати коефіцієнт рівномірності діаграми направленості.

На основі проведених розрахунків, **напрямки подальших досліджень** будуть направлені на аналіз та вдосконалення системи автоматичного управління діаграмою направленості Smart антен, які встановлені на рухомих об’єктах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Development of a method of increasing the interference immunity of frequency-hopping spread spectrum radio communication devices / O. Kuvshynov et al. *Eastern-European Journal of Enterprise Technologies*. 2019. Vol. 2, no. 9 (98). P. 74–84. URL: https://doi.org/10.15587/1729-4061.2019.160328.

2. Масесов М. О., Субач І. Ю., Руденко Д. М., Станович О. В. Перспективи застосування цифрового діаграмоутворення у станціях тропосферного зв’язку спеціального призначення. *Збірник наукових праць ВІТІ ДУТ*. 2014. Вип. 1. С. 43–48.

3. Булашенко А. В. Принципи формування променя інтелектуальних антен. *Вісник СумДУ*. Серія технічні науки. 2010. № 1. С. 14–19.

4. Increasing the anti-jammingness of mobile radio networks with adaptive beamforming / S. Boholii et al. *Communication, informatization and cybersecurity systems and technologies*. 2022. Vol. 2, no. 2. URL: https://doi.org/10.58254/viti.2.2022.01.5.

5. Ilinov M., Nesterenko I., Iankovskii O. The way of construction compensator interference based on usage multi-input antenna systems. *Communication, informatization and cybersecurity systems and technologies*. 2023. Vol. 3, no. 3. URL: https://doi.org/10.58254/viti.3.2023.02.14.

6. The method of adaptive signal reception with adaptive antenna arrays from moving sources / H. Radzivilov et al. *Communication, informatization and cybersecurity systems and technologies*. 2022. Vol. 1, no. 1. URL: https://doi.org/10.58254/viti.1.2022.09.75.

7. Шишацький А. В., Жук О. Г., Бєляков Р. О. Методика адаптивного управління параметрами MIMO-АФАР. *Системи озброєння і військова техніка*. 2016. № 4. С. 77–82.

8. Частотно неселективний просторовий канал з використанням адаптивних антенних решіток / О. Г. Цатурян та ін. *Вісник ВІТІ. Комунікаційні та інформаційні системи*. 2021. № 1. С. 113–119.

9. Підвищення ефективності функціонування систем радіозв’язку за рахунок використання адаптивних антенних решіток / І. В. Борисов та ін. *Збірник наукових праць ВІТІ*. 2015. С. 16–24.

10. Гриценок К. М., Гурський Т. Г. Методика формування діаграми спрямованості кільцевої антенної решітки радіостанції мобільної радіомережі в умовах навмисних завад. *Збірник наукових праць ВІТІ*. 2018. С. 6–16.

11. Радзівілов Г. Д., Ільїнов М. Д., Хоменко П. В. Фазована кругова антенна решітка на півхвильових симетричних вібраторах. *ІІІ Міжнародна науково-технічна конференція ВІТІ імені Героїв Крут:* Системи і технології зв’язку, інформатизації та кібербезпеки: акт. питання і тенденції розвитку, м. Київ, 19 груд. 2023 р. С. 263–264.

DOI: 10.58254/viti.5.2024.13.138

УДК 621.391

д-р техн. наук Сайко В. Г. ORCID: 0000-0002-3059-6787 (ВІТІ ім. Героїв Крут)
Романов Д. О. ORCID: 0009-0008-5522-7591 (ВІТІ ім. Героїв Крут)
канд. техн. наук, професор Наритник Т. М. ORCID: 0000-0001-8071-6356 (ІЕЗ УАННП)
канд. техн. наук Комаров В. О. ORCID: 0000-0002-4929-4527 (ВІТІ ім. Героїв Крут)
д-р філософії Фомін М. М. ORCID: 0000-0002-6864-4238 (ВІТІ ім. Героїв Крут)

АНАЛІЗ ПЕРСПЕКТИВ ВИКОРИСТАННЯ ТЕРАГЕРЦОВОГО ДІАПАЗОНУ ЧАСТОТ ДЛЯ БЕЗПРОВОДОВИХ МЕРЕЖ ЗВ’ЯЗКУ СПЕЦІАЛЬНОГО ПРИЗНАЧЕННЯ

Частоти від 100 ГГц до 3 ТГц є перспективними діапазонами для нового покоління систем безпроводового зв’язку через широкі смуги спектра. Ці частоти також пропонують потенціал для революційних застосувань, які стануть можливими завдяки новому мисленню та прогресу в пристроях, схемах, програмному забезпеченні, обробці сигналів і системах. Але перехід до реального використання терагерцового діапазону як частотного ресурсу при наданні послуг зв’язку для додатків у мережах безпроводового зв’язку спеціального призначення супроводжується безліччю проблем управління спектром, мережевого аналізу, проєктування та оптимізації архітектури цих мереж.

Новизна цієї роботи обумовлена тим, що, незважаючи на те, що в останні роки було опубліковано ряд публікацій оглядового характеру вітчизняних та зарубіжних авторів, у аспектах розвитку та специфіки можливих застосувань безпроводових мереж зв’язку терагерцового діапазону частот при вирішенні задач силовими структурами та силами спеціальних операцій не розглядалася. Особливу увагу автори акцентували на висвітленні вітчизняних розробок у галузі безпроводових систем доступу терагерцового діапазону.

Метою роботи є аналітичний огляд особливостей переходу до реального використання терагерцового діапазону частот як частотного ресурсу при наданні послуг зв’язку ресурсомісткими додатками безпроводових мереж зв’язку для перспективних застосувань під час вирішення спеціальних задач силовими структурами.

У роботі зроблено систематизований аналіз та класифікацію основних сфер перспективних застосувань безпроводових мереж зв’язку терагерцового діапазону частот при вирішенні задач силовими структурами спеціального призначення та наведено науково-дослідні задачі для переходу до реального використання терагерцового діапазону частот як частотного ресурсу при наданні послуг зв’язку для додатків спеціального призначення.

Розроблені в цій роботі авторські пропозиції для перспективних застосувань терагерцових мереж зв’язку силами спеціальних операцій або спеціального призначення є теоретичною основою створення методичного апарату для обґрунтування рекомендацій щодо вибору перспективних застосунків при використанні терагерцових мереж зв’язку силами спеціальних операцій при вирішенні задач спеціального призначення.

Ключові слова: *терагерцовий діапазон частот, безпроводові мережі зв’язку, мережі зв’язку спеціального призначення, аналітичний огляд, застосування мереж спеціального призначення.*

V. Saiko, D. Romanov, T. Narytnyk, V. Komarov, M. Fomin Analysis of prospects for the use of the terahertz frequency range for special purpose wireless communication networks.

Frequencies from 100 GHz to 3 THz are promising ranges for a new generation of wireless communication systems due to wide spectrum bands. These frequencies also offer the potential for revolutionary applications made possible by new thinking and advances in devices, circuits, software, signal processing and systems. But the transition to the real use of the terahertz range as a frequency resource in the provision of communication services for applications in special purpose wireless communication networks is accompanied by many problems of spectrum management, network analysis, design and optimization of the architecture of these networks.

The novelty of this work is due to the fact that, despite the fact that in recent years a number of review publications by domestic and foreign authors have been published, but in the aspects of development and specifics of possible applications of wireless communication networks of the terahertz frequency range in solving problems with power structures and forces special operations were not considered. The authors focused special attention on highlighted domestic developments in the field of wireless access systems in the terahertz range.

The purpose of the work is an analytical review of the features of the transition to the real use of the terahertz frequency range as a frequency resource in the provision of communication services by resource-intensive applications by wireless communication networks for promising applications in solving special problems by power structures.

In the work, a systematic analysis and classification of the main areas of promising applications of wireless communication networks of the terahertz frequency range in solving problems by special-purpose power structures is made, and scientific research tasks are given for the transition to the real use of the terahertz frequency range as a frequency resource in the provision of communication services for special purpose applications.

The author's proposals developed in this work for promising applications of terahertz communication networks by special operations forces or special purpose forces are the theoretical basis for creating a methodological apparatus for substantiating recommendations for the selection of promising applications when using terahertz communication networks by special operations forces when performing special purpose tasks.

Keywords: *terahertz frequency range, wireless communication networks, special purpose communication networks, analytical review, application of special purpose networks.*

Постановка завдання у загальному вигляді

На сьогодні велика увага приділяється використанню терагерцових сигналів у різних сферах, таких як радіозв’язок, ближня радіолокація та спектроскопія. Це пояснюється тим, що такі сигнали мають численні переваги порівняно з радіосистемами, що використовують сигнали у високочастотному діапазоні до 100 ГГц. Водночас у всьому світі зростає зацікавленість у використанні терагерцового діапазону частот (від 0,1 ТГц до 3 ТГц), який, по-перше, мало використовувався для радіозв’язку та радіолокації, по-друге, має велику інформаційну ємність, і, по-третє, не вимагає ліцензування у більшості країн. Цей діапазон займає значну частину електромагнітного спектра між інфрачервоним і мікрохвильовим діапазонами. Наразі використання терагерцового випромінювання ще досить обмежене і переважно на демонстраційному рівні, незважаючи на його великий потенціал для технічних та біомедичних застосувань.

Перехід до терагерцового діапазону дозволяє вирішити проблеми зі збільшенням швидкості передачі інформації завдяки використанню ширших спектрів частот (декілька десятків ГГц) і забезпечення електромагнітної сумісності радіоелектронних пристроїв, а також зменшує ймовірність радіозавад. Прикладом успішного використання терагерцового діапазону є запуск першого у світі експериментального китайського штучного супутника Землі з бортовим ретранслятором терагерцового діапазону для тестування мереж зв’язку 6G. Порівняно з іншими методами збільшення пропускної здатності, цей метод унікальний, оскільки більшість сучасних методів мають свої обмеження. Терагерцовий діапазон є перспективним для створення компактних завадозахищених високошвидкісних безпроводових пристроїв та систем зв’язку, оскільки він не схильний до впливу існуючих засобів радіоелектронної боротьби. Однак використання терагерцового діапазону також ускладнює роботу мереж зв’язку через швидке ослаблення сигналу у середовищі, ефект молекулярної абсорбції та використання спрямованих антен.

Аналіз публікацій за темою досліджень

Більшість статей щодо застосування терагерцових технологій у телекомунікаційних системах обмежується оглядом окремих вузлів та елементною базою.

В [1] розглянуто телекомунікаційні системи терагерцового діапазону, показано переваги та можливості, яких можна досягнути із використанням терагерцової технології. Також розглянуто експериментальні результати українських науковців у цій галузі, основних виробників радіоелектронних пристроїв та список їхньої продукції, що працює у терагерцовому діапазоні.

У роботах [2; 3] наведено аналітичний огляд теоретичних і експериментальних досліджень у галузі систем зв’язку суб- і терагерцового діапазону. Подальша перспектива застосування телекомунікаційних систем суб- і терагерцового діапазонів вбачається у створенні швидкодіючих широкосмугових радіоелектронних компонентів, включаючи підсилювачі потужності та малошумні підсилювачі і розвиток на цій основі методів побудови систем множинного доступу й інших телекомунікаційних систем суб- і терагерцового діапазонів нового покоління. Наведено результати розробки вітчизняного

експериментального зразка цифрової радіолінії із гігабітною пропускною здатністю, що функціонує в терагерцовому діапазоні і який може бути використаний в надвисокошвидкісних розподільчих мережах мобільного зв’язку нового покоління 5G для забезпечення передавання та приймання цифрової інформації зі швидкістю до 1 Гбіт/с у діапазоні частот 128–134 ГГц на дальності зв’язку в межах 1 км.

У [4] наведено інноваційні технологічні рішення та компоненти для мереж зв’язку терагерцового діапазону частот, шляхи використання телекомунікаційного зв’язку терагерцового діапазону частот, проблемні питання впровадження та напрямки подальших досліджень мереж зв’язку терагерцового діапазону частот.

У роботі [5] наведено огляд досягнень і проблем в області терагерцових комунікацій, орієнтованих на швидкість передавання даних близько 100 Гбіт/с. Подано короткий огляд напівпровідникових приладів із граничною частотою, що знаходиться в терагерцовому діапазоні.

Огляд, присвячений досягненням у галузі розробки та застосування ліній зв’язку терагерцового діапазону частот, а також перспектив розвитку представлених технологій та компонентної бази, наведено у [6].

Створено експериментальні стенди 3D терагерцового ЛЧМ-радіолокатора, проведено дослідження побудованого ЛЧМ-радара з обмеженням у часі частоти биття від трьох шарів відбиття [7; 8].

У роботі [9] наведено методи застосування гібридних технологій для реалізації телекомунікаційних систем міліметрового діапазону.

В [10] наведено результати дослідження ключових особливостей терагерцових систем, аналізу найбільш перспективних наукових програм та можливі відкриті питання дослідження.

Глибокий аналіз особливостей застосування засобів терагерцового діапазону наведено в [11], а також підходи до того, як зменшити енергоспоживання та підвищити продуктивність у кількох проблемних областях. Описано багато технічних проблем і можливостей для безпроводового зв’язку та додатків зондування вище 100 ГГц, а також представлено ряд багатообіцяючих відкриттів, нових підходів і останніх результатів, які допоможуть у розробці та впровадженні безпроводових мереж нового покоління.

Дослідженню можливостей використання терагерцового діапазону в гетерогенній транспортній розподільчій мережі мобільного зв’язку присвячено цикл публікацій [12–18].

Отже, новизна цієї роботи обумовлена тим, що, незважаючи на те, що в останні роки було опубліковано ряд публікацій оглядового характеру вітчизняних та зарубіжних авторів [1–18], в аспектах розвитку та специфіки можливих застосувань безпроводових мереж зв’язку терагерцового діапазону частот при вирішенні задач силовими структурами та силами спеціальних операцій не розглядалася.

Постановка завдання

Не ставлячи за мету охопити весь спектр питань, пов’язаних із перевагами і проблемами терагерцових мереж зв’язку, в цій роботі наводиться аналітичний огляд та авторські пропозиції основних перспективних застосунків і найбільш основних дослідницьких завдань при використанні безпроводових мереж зв’язку для вирішення задач спеціального призначення. Таким чином, представлена робота поєднує попередні дослідження в різних галузях знань, що має спростити входження в таку багатогранну область, як побудова та аналіз терагерцових безпроводових мереж зв’язку спеціального призначення для вирішення професійних задач силами спеціальних операцій.

Виклад основного матеріалу

Перспективні застосування для терагерцових мереж зв’язку спеціального призначення

Поява терагерцових мереж зв’язку уможливлює відразу кілька груп ресурсомістких додатків для застосування під час вирішення професійних задач силовими структурами спеціального призначення. Деякі з них являють собою логічний розвиток поширених додатків з урахуванням особливостей терагерцових мереж зв’язку. Інші ж, навпаки, стали принципово можливими лише з появою безпроводових мереж зв’язку в терагерцовому діапазоні частот [2; 3]. У цьому підрозділі наведено основні напрями з обох категорій, а також запропоновано сценарії використання для терагерцових мереж зв’язку спеціального призначення (рис. 1). Далі зупинимось на деяких із них.

Рис. 1. Основні перспективні сфери використання безпроводових мереж зв’язку терагерцового діапазону для спеціальних задач силових структур

Інформаційний водоспад

Мале територіальне охоплення терагерцових мереж зв’язку, обумовлене швидким загасанням сигналу на великій відстані поширення в середовищі, потребує перегляду концепції «малих стільників». З одного боку, використання терагерцового діапазону для таких стільникових мереж виглядає перспективно через велику потенційну ємність мережі, яку можна створити. З іншого боку, реальний радіус дії стільників за допомогою існуючого обладнання обмежений кількома метрами. Таким чином, повне покриття навіть обмеженої території терагерцовими стільниками виявляється економічно невигідним через велику

кількість необхідних точок доступу і високу вартість окремих точок доступу, особливо при вирішенні спеціальних завдань.

У зв’язку з цим було запропоновано спосіб «інформаційного водоспаду», який передбачає розміщення невеликої кількості терагерцових точок доступу в зонах з найбільш інтенсивним потоком користувачів: на початковій, середній та кінцевій точках маршруту руху колони сил спеціального призначення, у місцях дислокації такої колони, вузьких коридорах або місцях з високою прохідністю тощо.

Отже, незважаючи на те, що повне покриття заданої зони недосяжне, значна частина потоку спеціального призначення періодично опиняється в зоні покриття хоча б однієї терагерцової точки доступу. Наші попередні дослідження показують, що середній час між потраплянням у зону покриття хоча б однієї терагерцової точки доступу становить від декількох десятків секунд до декількох хвилин. Це спостереження лежить в основі способу «інформаційного водоспаду» – точки доступу великої ємності, яка протягом короткого часу, поки користувач перебуває в зоні покриття (приблизно кілька хвилин), організовує передачу великої кількості даних з/на користувальницький термінал. Постійне перенаправлення «важкого» трафіку, який кешується (наприклад, відео), на терагерцові точки доступу дозволяє значно зменшити навантаження на стільникову мережу і звільнити ресурси для тих абонентів, які в конкретний момент не знаходяться в зоні покриття жодної з терагерцових точок доступу.

Високозахищені мережі

Можливість створення високонаправлених діаграм з високою ємністю терагерцових мереж зв’язку дозволяє використовувати ці системи для організації надійних з’єднань. Зазвичай це включає створення дуже вузького променю між рухомими комплексами радіозв’язку (РКР) і пристроями користувачів, такими як мобільні телефони або термінали, які знаходяться в прямій видимості один до одного. Це створює систему, яка може бути прослухана тільки терміналом, розташованим на відрізку між РКР і телефоном/терміналом. Однак цей відрізок є коротким і зазвичай видимим користувачам й камерам РКР. Крім того, за допомогою відповідних кодів можна зробити злам такої лінії зв’язку практично неможливим, забезпечуючи теоретичну стійкість системи. Це означає, що ймовірність успішного дешифрування переданого повідомлення практично дорівнює ймовірності вгадати це повідомлення.

Радіорелейні системи терагерцового діапазону частот

Однією з перспективних сфер застосування терагерцових технологій є системи телекомунікацій [1–18]. Зокрема передбачається створення нових за габаритами, завадозахищеністю та енергоефективністю пристроїв суб- і терагерцового діапазону для транспортних мереж мобільного зв’язку нового покоління, високошвидкісної передачі відеосигналів, радіорелейних систем прямої видимості, транспортних мереж мобільного зв’язку нового покоління та радарів високоточного виявлення і розпізнавання малорозмірних швидкісних цілей, сенсорів для отримання більш точної та детальної оперативної інформації про стан контрольованого об’єкта або місцевості.

Такі розробки ведуться в США під егідою DARPA, в Німеччині, Англії, Китаї для застосування як у військових, так і у цивільних цілях.

В [2; 3] для систем радіозв’язку на основі аналізу наявної електронної арсенід-галієвої елементної бази, теоретичних і експериментальних досліджень проведено проєктування основних вузлів і приймально-передавального модуля телекомунікаційної системи із гігабітною пропускною здатністю до 1200 Мбіт/с в діапазоні частот 130–134 ГГц у складі: частотні перетворювачі (перетворювач вниз та змішувач частоти) із субгармонійною накачкою, гетеродин, що використовує високостабільний задаючий кварцовий генератор з подальшим ланцюгом помножувальних і підсилювальних каскадів, смугопропускний фільтр з використанням тонкої металевої пластини в Е-площині хвилеводного каналу 1,6 мм × 0,8 мм,

кінечна рупорна антена з діелектричним лінзовим концентратором, а також формувачів радіосигналів. Проведено експериментальні дослідження лабораторного зразка цифрової симплексної радіорелейної системи радіозв’язку терагерцового діапазону частот 130–134 ГГц і отримано наступні результати: пропускна канальна здатність до 1200 Мбіт/с; значення ймовірних бітових помилок BER не більше ніж 10^{-6} (вид модуляції – КАМ-64); дальність зв’язку в нормальних умовах в межах 1 км при коефіцієнті підсилення системи на рівні 50 дБ (рис. 2).

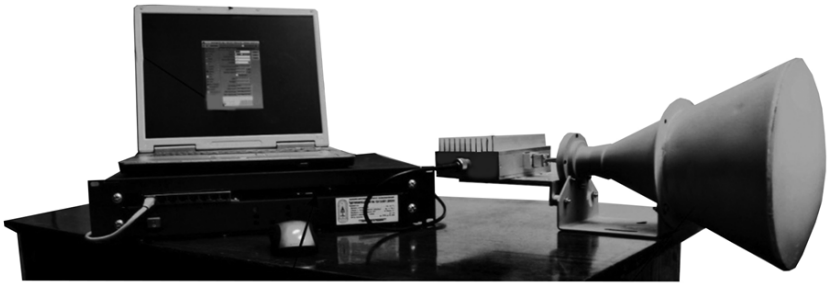


Рис. 2. Зовнішній вигляд лабораторного зразка цифрової симплексної радіорелейної системи терагерцового діапазону

На рисунку 3 наведено нову схему організації цифрового радіорелейного зв’язку для передачі інформації на пункт управління 1 мережею цифрової радіорелейної станції (ЦРС) терагерцового діапазону через вузлову приймально-передавальну ЦРС 4 та приймально-передавальні модулі 2, 3 на 0,13 ТГц. Такий варіант можна використати при розгортанні польового вузла зв’язку спеціального призначення.

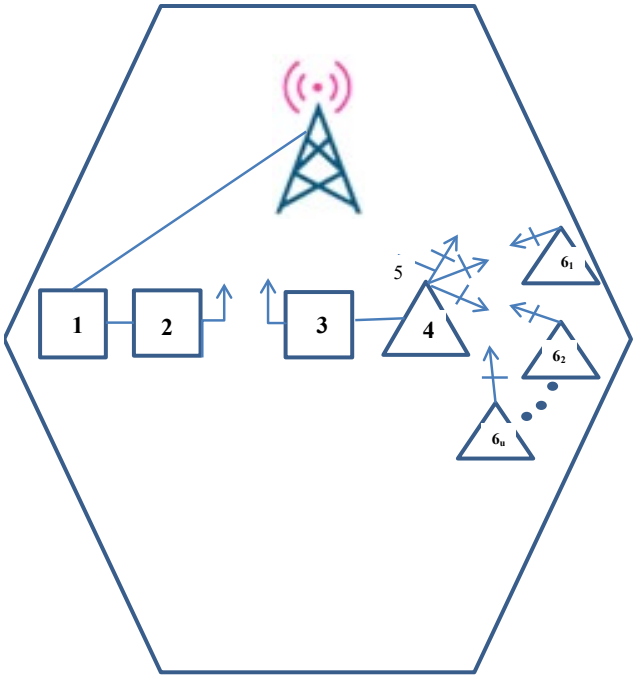


Рис. 3. Схема організації цифрового радіорелейного зв’язку для передачі інформації на пункт управління мережею ЦРС терагерцового діапазону через вузлову ЦРС та приймально-передавальні модулі

Вузлова ЦРС 4 (див. рис. 3) «точка-багатоточка» є пакетною та має на кожному з *N*-напрямків пропускну здатність 2,4 Гбіт/с. Пакетна технологія дозволяє «послідовно»

передавати потоки зі швидкістю 2,4 Гбіт/с на приймально-передавальні модулі терагерцового діапазону частот 2, 3 (див. рис. 3), який весь трафік *N**2,4 Гбіт/с передає на пункт управління мережею 1 ЦРС $6_{1..u}$ терагерцового діапазону, тому передача мультимедійної, графічної й іншої інформації здійснюється в терагерцовому діапазоні. Висока скритність передачі інформації забезпечується особливостями терагерцового діапазону хвиль.

Мережі зв'язку систем безпеки особливо важливих об'єктів спеціального призначення

Відоме інноваційне рішення з використанням терагерцового діапазону частот, яке відноситься до техніки захисту території, що охороняється від несанкціонованого проникнення порушника і може бути використано на особливо важливих військових і державних об'єктах [19].

Спосіб захисту сфокусованим випромінюванням субтерагерцового діапазону зон і об'єктів від несанкціонованого проникнення, в якому визначають місцезнаходження приховано розташованого порушника за допомогою терагерцових датчиків і вплив на нього здійснюють електромагнітним випромінюванням установки субтерагерцового діапазону з урахуванням відбиття електромагнітного випромінювання від різних поверхонь, відрізняється від відомих тим, що робочу частоту випромінювання вибирають в частотному діапазоні «вікна прозорості» 94–96 ГГц, розраховують місце розташування об'єктно-орієнтованої поверхні, яка відбиває промінь та знаходиться на високо піднятій аероплатформі, і кут напрямку впливу на порушника з урахуванням відбиття електромагнітного випромінювання від поверхні на високо піднятій аероплатформі. При цьому спочатку промінь джерела електромагнітного випромінювання субтерагерцового діапазону частот направляється на поверхню, що знаходиться на високо піднятій аероплатформі, а потім сфокусований промінь потрапляє на місце розташування порушника для створення поблизу його знаходження області заданої підвищеної щільності потоку енергії випромінювання.

В основі роботи технічного рішення спрямованої енергії покладено високотемпературний розігрів повітряного простору (понад 900 градусів за шкалою Кельвіна) за допомогою вузьконаправленого радіовипромінювання, частота якого відповідає частоті максимального поглинання енергії молекулами атмосферного кисню і водяної пари.

Принцип дії полягає в наступному: потужний передавач генерує вузький спрямований (шириною до 0,5 градуса) радіопромінь у частотному діапазоні «вікна прозорості» 94–96ГГц, який проникає в шкіру людини на глибину до 0,5 мм, нагріваючи її до 45 $^{\circ}$C і вище. При цьому в людини виникає нестерпний біль. Больова реакція на випромінювання інтенсивністю до 100 Вт/м^2 достатньо вагома, щоб заставити порушника покинути зону, що охороняється.

Реакція на випромінювання виникає протягом 2-3 секунд, стає нестерпною через 5 секунд і зникає після відключення генератора або після того, як порушник залишить зону опромінення. Якщо порушник не покине зону охорони протягом 250 секунд, він отримає опік шкіри (рис. 4). Прогнозована дальність ураження живої сили противника – в межах 1 км і більше.

Технічний результат, на отримання якого спрямоване технічне рішення, – зменшення масо-габаритних характеристик і енергоспоживання компонентів і підсистем завдяки застосуванню твердотільних підсилювальних пристроїв, що дозволяє потенційно зменшити масо-габаритні характеристики системи та енергоспоживання компонентів. Також при цьому спрощується процес технічної експлуатації системи.

В наш час спільно фахівці та науковці Військового інституту телекомунікацій та інформатизації імені Героїв Крут та СП «Інститут електроніки та зв'язку Української академії наук» пропонують до виконання в стислі терміни науково-дослідну роботу «Випромінююча телекомунікаційна система субтерагерцового діапазону для захисту особового складу в бойових операціях». Така випромінююча система, яка базується на використанні наявного

науково-технічного та виробничо-технологічного потенціалу і знаходиться на рівні кращих зарубіжних зразків, призначена для використання в Збройних Силах й інших силових структурах України як високоефективна енергетична зброя та радіоелектронний засіб стримування нового покоління.

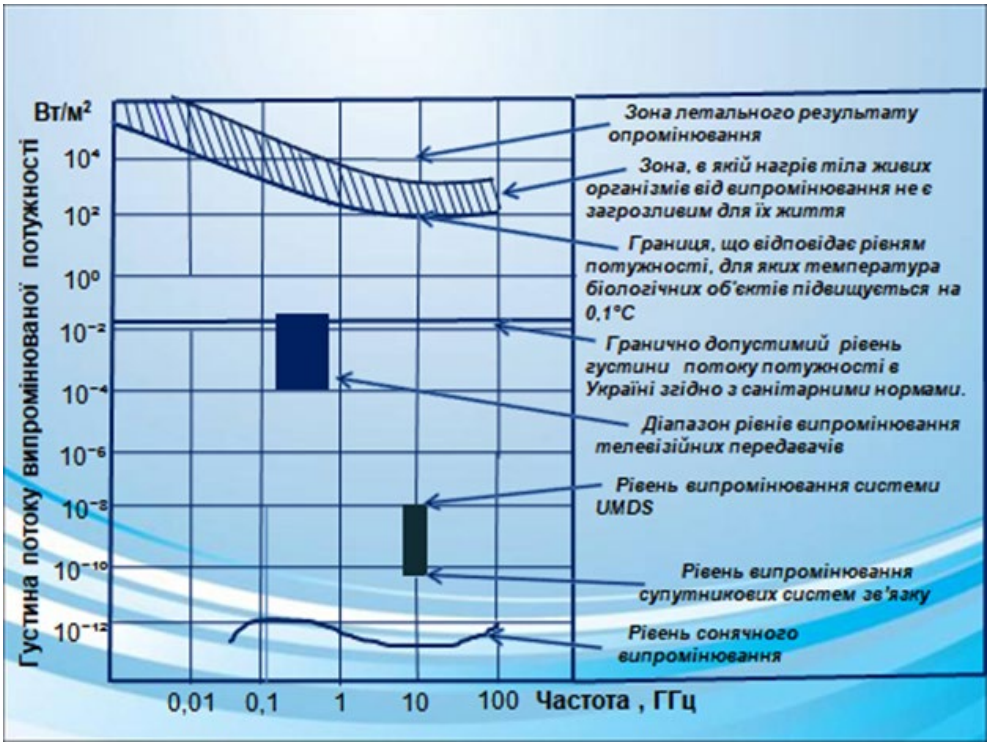


Рис. 4. Діаграма впливу мікрохвильового опромінювання

Системи міжсупутникового зв’язку

Відоме вітчизняне рішення, яке є каналом міжсупутникового зв’язку в терагерцовому діапазоні частот, що включає в себе два геостаціонарних супутники зв’язку – ведучий і ведений, міжсупутниковий двосторонній зв’язок, наземні пункти зв’язку, відрізняється від відомих рішень тим, що в системі міжсупутникового зв’язку двосторонній канал зв’язку базується на каналі передачі в терагерцовому діапазоні частот, а на геостаціонарних супутниках зв’язку розміщені апаратура перенесення стандартного частотного діапазону в терагерцовий частотний діапазон та антени передачі інформації в терагерцовому діапазоні частот [20].

Радіолокатори для високоточних систем наведення і керування

Для високоточних систем наведення та керування розроблено нове апаратно- та схемотехнічне рішення приймально-передавального модуля терагерцового діапазону (рис. 5), яке базується на застосуванні сучасних інтегральних мікросхем терагерцових трансиверів в діапазоні 119–126 ГГц трансивера та плати формування та обробки сигналів (плата синтезатора частоти), що інтегровані між собою відповідними інтерфейсами і сумісні з використаним налагоджувальним набором Sirad Easy компанії Silicon Radar [21–23].

Проведено математичне та імітаційне моделювання електричних схем вузлів приймально-передавального модуля терагерцового діапазону з використанням САПР NI Multisim, а також за допомогою САПР Altium Designer зпроєктовано 2D- та 3D-моделі плат, на основі яких виготовлено та здійснено монтаж електронних компонентів на друковану плату терагерцового трансивера та друковану плату формування та обробки сигналів приймально-передавального модуля терагерцового діапазону. Вперше розроблено макетний зразок

приймально-передавального модуля терагерцового діапазону (119–126 ГГц) для високоточних систем наведення і керування, у якого основні надвисокочастотні вузли (змішувач частот, гетеродин, підсилювач тощо) реалізовані за кремній-германієвою технологією SiGe повністю в межах однієї малогабаритної інтегральної мікросхеми з інтегрованою антеною (габаритні розміри не перевищують по діаметру 35 мм та по довжині 10 мм), що забезпечує при вихідній потужності 3 мВт необхідну (в межах до 100–300 м) дальність дії модуля в складі радара ближньої дії при оптимальних значеннях енергетичного ресурсу радіолінії при практичній відсутності можливості виявлення і несанкціонованого доступу до інформації, що передається в терагерцовому діапазоні частот.

Рис. 5. Пристрої провідних світових виробників на основі IMC компанії Silicon Radar

Стратосферні системи зв’язку

Іншим прикладом корисного використання технології терагерцового діапазону частот є її застосування у стратосферних системах зв’язку. Останнім часом ця концепція здобуває все більшу популярність. Суть стратосферних систем зв’язку полягає у розташуванні приймально-передавального обладнання, тобто базових станцій, на безпілотних стратосферних платформах, таких як повітряні кулі, дрони, дирижаблі й т. ін., які летять на висоті 18–25 км, не перешкоджаючи цивільній авіації. На таких висотах швидкість вітру досить низька, а густина повітря значно менша, ніж на Землі. Застосування стратосферних систем зв’язку має численні переваги, зокрема підвищений кут огляду порівняно з наземними системами. Для забезпечення зв’язку зі стратосферною станцією доцільно використовувати терагерцові системи через їхні переваги, такі як низька завантаженість діапазону, можливість використання широких смуг частот, спрощена процедура виділення частот у багатьох країнах і значне зменшення розмірів антенних систем, які можуть формувати вузькі діаграми напрямленості. Широке використання стратосферних телекомунікаційних систем зв’язку відкриває великі можливості для оперативного зв’язку військових цілей, забезпечення зв’язку у зонах, постраждалих від стихійних лих або військової агресії. Для ефективної роботи систем зв’язку в терагерцовому діапазоні необхідні відповідні засоби, спрямовані на цей діапазон.

Дослідні науково-технічні задачі для терагерцових мереж зв’язку спеціального призначення

Крім описаних вище переваг, безпроводові мережі зв’язку, що працюють у терагерцовому діапазоні частот, мають ряд особливостей, які, у свою чергу, ведуть до набору

дослідницьких завдань, що стоять перед розробниками (рис. 6). Рішення описаних нижче завдань є критично важливим для побудови терагерцових мереж спеціального призначення та підтримки перерахованих категорій застосунків (див. рис. 1).

Рис. 6. Дослідні науково-технічні задачі для мереж зв’язку терагерцового діапазону частот спеціального призначення

Далі зупинимося спочатку на особливостях терагерцового діапазону частот, а потім на деяких науково-дослідницьких задачах.

Особливості терагерцового діапазону частот

Покриття мобільних мереж у терагерцовому діапазоні буде обмежено застосуванням базових станцій з малим та надмалим радіусом дії через ослаблення сигналів в атмосфері Землі та гідрометеорах. Ослаблення терагерцового діапазону хвиль в атмосфері зростає зі збільшенням частоти сигналів, що використовується. Якщо затухання в атмосфері перевищує 10 дБ на відстані 1 км, безпроводовий зв’язок у мережі для радіусів стільників в одиницях кілометрів буде вкрай утруднений.

Субтерагерцовий діапазон має значне загасання сигналів через вплив дощу, що дає майже постійне питоме ослаблення на частотах вище 100 ГГц із втратами від 0,8 дБ/км до 50 дБ/км залежно від інтенсивності дощу відповідно до рекомендації ITU-R P. 838-3 [25].

Вікно прозорості терагерцового діапазону в смузі частот між 252 ГГц і 296 ГГц може забезпечувати дальність зв’язку від базової станції мережі мобільного зв’язку, що не перевищує 2 км, при цьому ослаблення внаслідок впливу атмосфери Землі становить 10 дБ [26].

При використанні в мережах діапазону частот між 192 ГГц і 298 ГГц дальність зв’язку в мережах може становити більше 2 км, а при використанні діапазону частот 121–154 ГГц – понад 8 км. Згасання в атмосфері у терагерцовому діапазоні перевищує 10 дБ/км на всіх частотах вище 351 ГГц. Слід також наголосити, що дальність зв’язку в мережах зменшуватиметься, коли спостерігатимуться опади, наприклад, дощ.

Крім забезпечення обмежених дальностей зв’язку та надзвичайно високих швидкостей передачі даних, квазіоптична природа терагерцових діапазонів частот має багатообіцяючі можливості для функцій зондування, візуалізації та позиціонування (3CLP) у мережах нового покоління, які активно досліджуються.

Таким чином, виходячи з властивостей для вже використовуваного спектра, нижня частина спектра (діапазони частот нижче 6 ГГц та субміліметрового діапазону) забезпечує середні швидкості передачі даних та можливості вимірювання параметрів сигналів з більш високою надійністю та нижчою роздільною здатністю. Між тим викликає зростаючий інтерес верхня частина діапазонів спектра (діапазон субтерагерцовий і терагерцовий), яка забезпечує високу швидкість передачі даних і зондування з високою роздільною здатністю, але зі зменшенням дальності зв’язку і більш переривчастими каналами.

Крім того, використання для більшості послуг субтерагерцового і терагерцового діапазонів частот може допомогти вирішити проблему нестачі спектра для безпроводового зв’язку, яка спостерігається в смугах нижче 6 ГГц. Таким чином, враховуючи особливі переваги, які можуть бути отримані для послуг зв’язку, зондування та формування зображень у терагерцовому діапазоні, спільне використання різних діапазонів забезпечує місцевизначення абонента при наданні послуг зв’язку з високою роздільною здатністю до 1 см у терагерцовому діапазоні. Такі сценарії не тільки забезпечують більш високу ефективність використання спектра, але й відкривають шлях для нових можливостей, що випливають із можливості виконання завдань скоординованого зондування та зв’язку для мереж безпроводового зв’язку спеціального призначення.

Науково-дослідницькі задачі

Побудова детальних моделей для каналів зв’язку терагерцових мереж спеціального призначення

Безліч різноманітних ефектів має бути враховано при побудові детальної моделі терагерцового каналу зв’язку. По-перше, суттєвий вплив частотно-залежного ефекту молекулярної абсорбції призводить до бажання використовувати так звані «вікна прозорості» –- частотні діапазони всередині терагерцового спектра, в яких вплив ефекту молекулярної абсорбції мінімальний. Проте навіть у «вікнах прозорості» вплив молекулярної абсорбції на поширення сигналу не нульовий. Через це, при розгляді навіть найпростіших моделей поширення (наприклад, «Free Space Path Loss») стосовно терагерцових частот, потрібно додавання ще одного компонента (експоненційно залежить від дистанції між передавачем і приймачем) до рівняння [27]. Далі, так як більша частина аналізованих сценаріїв використання терагерцових мереж зв’язку передбачає зв’язок всередині приміщень, всі складні ефекти поширення терагерцового сигналу всередині будівель (такі як віддзеркалювання сигналу від стін та інших об’єктів, ослаблення сигналу при проходженні крізь об’єкти і багато іншого) мають бути враховані. У зв’язку зі складністю цієї задачі, на сьогодні відомі лише кілька аналітичних моделей поширення терагерцового сигналу всередині будівель [28; 29]. Нарешті, володіючи довжиною хвилі в частках міліметра, терагерцовий сигнал розсіюється при контакті практично з будь-яким твердим предметом, як усередині приміщення, так та за його межами. Нарешті, використання спрямованих антен вимагає значно складнішого підходу до аналізу інтерференції в каналі, яка викликана багатопроменевим поширенням корисного сигналу, і впливом сусідніх точок доступу/абонентів. В результаті, існує два основних підходи до побудови моделей каналів для

терагерцових мереж зв’язку. Перший підхід передбачає детальне моделювання простору навколо приймача і передавача з метою побудови всіх можливих (або найімовірніших) траєкторій поширення сигналу. Такий підхід, при правильному застосуванні, дозволяє отримувати результати з високою точністю, проте є обчислювально складним і не може бути використаний для побудови широко застосовних моделей, так як найменша зміна сценарію призводить до необхідності моделювати поширення сигналу ще раз. Другий підхід заснований на «усередненні» множини результатів, отриманих першим підходом, з метою побудови стохастичної моделі каналу, що застосовується для широкого набору сценаріїв, нехай і на шкоду точності результатів, що видаються моделлю у кожному конкретному сценарії.

Застосунки одночасної локалізації та картування (Simultaneous Localization And Mapping, SLAM) у міліметровому або терагерцовому діапазоні дозволяють відновлювати тривимірні карти околиць у невідомих середовищах. SLAM заснований на концепції, згідно з якою пристрій, що володіє сенсорами, переміщаючись в невідомому середовищі, розпізнає навколишні об’єкти (орієнтири) і згодом відновлює 2D- або 3D-карту навколишнього середовища для подальшого підвищення точності локалізації. Поточна точна технологія SLAM (для автоматичного створення внутрішніх та зовнішніх карт) вимагає як високої роздільної здатності при вимірюванні відстані (дальності), так і дуже високих кутових дозволів, які традиційно досягаються за допомогою лазерних та оптичних технологій. Хоча лідари та оптичні камери можуть забезпечувати високу роздільну здатність, вони не можуть працювати в несприятливих погодних умовах (туман, дощ, хмари тощо) та в умовах низького зовнішнього освітлення.

Розробка нових алгоритмів множинного доступу до середовища передачі

Далі не можна не зазначити одну з найбільш складних проблем при розробці терагерцових мереж зв’язку – створення ефективних алгоритмів множинного доступу. Більшість існуючих безпроводових мереж оперують з антенами слабкої спрямованості, що дозволяє передавати сигнальні повідомлення відразу великому числу сусідів. Терагерцові мережі зв’язку, навпаки, характеризуються дуже вузькими променями діаграм спрямованості, що призводить до складнощів при пошуку та підтримці стабільного з’єднання до сусідів, особливо якщо обрані вузли поки що не спілкувалися один з одним і нічого не знають про взаємне розташування.

Також у діапазоні міліметрових і субтерагерцових хвиль мережі повинні використовуватися в поєднанні з масивним MIMO, особливо в застосуваннях з великою пропускною здатністю. Зі збільшенням кількості антенних елементів та переходом у більш високочастотний діапазон сигнальні промені MIMO стануть надзвичайно вузькими. Отже, буде складно генерувати кілька дуже вузьких променів, які точно спрямовані на кількох користувачів одночасно, тому що в багатьох сценаріях буде важко отримати точну інформацію про стан каналу. У таких випадках для надійного формування променю можуть бути розроблені покращені схеми NOMA спільно з розрахованим на багато користувачів попереднім кодуванням. Маючи можливості NOMA, MU-MIMO може використовувати модифіковану схему попереднього кодування. Замість формування дуже вузьких і точних променів, націлених на кожного користувача окремо, прекодер MU-MIMO тепер може генерувати ширші промені для націлювання на групу користувачів, які додатково мультиплексуються за допомогою NOMA. Збільшення ширини променю робить його більш стійким до змін параметрів каналу, наприклад, викликаним мобільністю користувача або затримкою вимірювання та зворотного зв’язку. Міжкористувацькі завади всередині групи додатково придушуватимуться приймачем NOMA за допомогою переданих сигнатур NOMA. Аналогічні ідеї можуть бути додатково застосовані при спільній передачі з кількома базовими станціями. Таким чином, для ефективної роботи терагерцової мережі зв’язку потрібно

створення інноваційних алгоритмів та протоколів встановлення і підтримки з’єднання, а також організації множинного доступу в канал зв’язку.

Підтримка передачі при блокуванні прямої видимості

У реальних умовах найбільші труднощі системам терагерцового діапазону частот у більшості випадків привносять рухливі перешкоди, такі як люди і транспортні засоби, які є блокаторами поширення радіосигналу. У випадку, якщо обладнання тимчасово потрапляє в стан блокування радіосигналу деяким об’єктом, то залежно від середовища розповсюдження сигналу та відстані між мобільним пристроєм і базовою станцією NR BS (англ. New Radio Base Station) цей пристрій може або випасти із зони покриття BS, або знизити свою схему модуляції та кодування так, щоб ймовірність помилки на рівні каналу не перевищувала наперед визначеного цільового значення.

Відомо рішення з одночасним підключенням користувача до кількох базових станцій мережевої інфраструктури в умовах їх щільного розміщення, що враховує блокування каналу прямої видимості перешкод при передачі на вкрай високих частотах [30]. Однак складність його технічної реалізації на існуючій інфраструктурі мобільного зв’язку 5G призводить до додаткових затримок перемикання каналів, оскільки рішення приймаються віддалено, та відповідно знижує загальну ефективність системи. Це пов’язано з тим, що існуюча централізована архітектура інфраструктури мереж мобільного зв’язку на сьогодні є вразливою з точки зору перевантаження обчислювальних ресурсів і тому вона не гарантує безперебійне надання сервісів IoT, у випадку коли у головних серверах виникають збої програмного забезпечення. Тому, зростаюча потреба в різноманітних додатках, що потребують високої пропускної здатності, таких як мобільне потокове відео та обробка великих даних, потребує зміни принципів управління радіоресурсами в мережах мобільного зв’язку, щоб уникнути їхнього дефіциту ресурсів для забезпечення новітніх сервісів для абонентів.

З цієї точки зору представляють інтерес нові технічні рішення, запропоновані авторами в [31; 32], і рішення з інтегрованими інтелектуальними поверхнями, що реконфігуруються (Reconfigurable Intelligent Surface, RIS), які можна використовувати для допомоги у швидкому формуванні променю з використанням точного позиціонування або подолання ефектів блокування завдяки збору даних про канали в системах міліметрових та терагерцових діапазонах хвиль. Інтелектуальні поверхні, що реконфігуруються, відносно недавно стали багатообіцяючою парадигмою проєктування безпроводових мереж і режимів безпроводової передачі [33; 34]. Вони також можуть створювати інтелектуальне радіосередовище (або інтелектуальні радіоканали), тобто поширенням радіохвиль у навколишньому середовищі можна керувати для створення персоналізованого каналу зв’язку. В узагальненій моделі RIS-мережа формується між декількома БС для створення великомасштабних інтелектуальних радіоканалів, що обслуговують кількох користувачів. За відсутності керованого середовища архітектура безпроводової системи та режим передачі можуть бути оптимізовані тільки відповідно до статистичних властивостей фізичних каналів та/або інформацією, що повертається від приймача до передавача. У керованому середовищі RIS-мережі спочатку сприймають дані середовища знаходження і повертають в систему. Виходячи з цих даних, система оптимізує режим передачі та параметри RIS по інтелектуальних радіоканалах на стороні передавача, каналу та приймача. Завдяки підтримці формування променю, пов’язаної з RIS, використання інтелектуальних радіоканалів може значно покращити якість зв’язку, продуктивність системи, покриття стільника та якість зв’язку на межі стільника в безпроводових мережах. Тим не менш, через велику кількість елементів, що відбивають промені на RIS, оцінка каналу виявляється складним завданням. Крім того, потенціал RIS для локалізації отримав лише обмежене висвітлення в літературі, включаючи попередні дослідження, в яких RIS працює в режимі прийому як лінза [33] та в режимі відображення [34].

Висновки

1. Терагерцові мережі зв’язку мають ряд істотних переваг порівняно з сантиметровими та міліметровими мережами, в основному викликаними більшою шириною смуги пропускання каналу. Як проілюстровано в цій роботі, подібні переваги дозволяють терагерцовим мережам більш ефективно підтримувати деякі категорії застосунків при вирішенні задач силовими структурами спеціального призначення, у той час як окремі застосунки та сценарії використання стають можливими тільки при використанні безпроводових мереж, що працюють у терагерцовому діапазоні частот. Водночас, побудова та аналіз терагерцових мереж зв’язку для вирішення задач спеціального призначення вимагає вирішення ряду дослідницьких та інженерних завдань, частина з яких також описана в цій роботі.

2. Розроблені в цій роботі системні пропозиції для перспективних застосувань терагерцових мереж зв’язку силами спеціальних операцій або спеціального призначення є теоретичною основою створення методичного апарату:

- для обґрунтування рекомендацій щодо вибору перспективних застосувань терагерцових мереж зв’язку силами спеціальних операцій при виконані задач спеціального призначення;
- для обґрунтування вимог до випробувальної бази експериментального дослідження.

3. Нині активно досліджуються чотири напрями вирішення критичної проблеми обмеження дальності зв’язку в терагерцовому діапазоні, а саме: проєктування фізичного рівня з урахуванням відстані, використання антен ultra-massive MIMO, рефлекторних решіток, що віддзеркалюють, та інтелектуальних поверхонь RIS, що віддзеркалюють.

Напрямком подальшої роботи є розробка мобільної автономної безпроводової мережі зв’язку терагерцового діапазону частот системи безпеки особливо важливих об’єктів спеціального призначення з підтримкою мобільної мережі нового покоління.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Наритник Т. М. Аналіз терагерцових технологій та їх застосування для створення інноваційних розробок. *Електронне наукове фахове видання – журнал «Проблеми телекомунікацій»*. 2017. № 1 (20). С. 25–30. URL: http://pt.journal.kh.ua (дата звернення: 10.01.2024).

2. Кравчук С. О., Наритник Т. М. Телекомунікаційні системи терагерцового діапазону: монографія. Житомир: ФОП «Євенок О. О.», 2015. 208 с.

3. Сайко В. Г., Наритник Т. М. Безпроводові системи зв’язку терагерцового діапазону: монографія. Німеччина: видавництво LAP Lambert Academic Publishing, 2019. 69 с.

4. Сайко В. Г., Одарченко Р. С., Абакумова А. О., Наритник Т. М., Наконечний В. С., Домрачев В. М., Толюпа С. В., Заблоцький В. Ю., Баховський П. Ф. Мережі мобільного зв’язку нового покоління 4G/5G/6G: монографія. Київ: ТОВ «Про формат», 2021. 200 с.

5. Майборода І. М., Стороженко І. П., Бабенко В. П., Кайдаш М. В. Огляд досягнень в терагерцових комунікаційних системах. *Збірник наукових праць Національної академії Національної гвардії України.* 2016. Вип. 1. С. 45–48.

6. Исаев В. М., Кабанов И. Н., Комаров В. В., Мещанов В. П. Современные радиоэлектронные системы терагерцового диапазона. *Доклады ТУСУРа.* 2014. № 4 (34). С. 5–21.

7. Косовець М. А., Товстенко Л. М. Класичні параметричні методи цифрового спектрального аналізу характеристических функцій 3D терагерцового ЛЧМ радіолокатора. *Зв'язок.* 2016. № 4. С. 53–58.

8. Косовець Н. А., Павлов О. И. Моделирование нелинейных элементов цифрового 3D радара. *Зв'язок.* 2016. № 2. С. 40–47.

9. Кременецька Я. А. Радіофотонні технології та пристрої телекомунікацій: монографія. К.: Друк «ТОВ Три К», 2019. 220 с.

10. Mohammed H. Alsharif. Toward 6G Communication Networks: Terahertz Frequency Challenges and Open Research Issues. *Computers, Materials & Continua.* 2021. vol. 66. no 3. pp. 2831–2839.

11. T. S. Rappaport. 6G and beyond: Terahertz communications and sensing. 2019 Brooklyn 5G Summit Keynote, Apr. 2019. [Online]. Available: URL: https://ieeetv.ieee.org/conference-highlights/keynote-tedrappaport-terahertz-communication-b5gs-2019 (дата звернення 10.01.2024).

12. Сайко В. Г., Наритник Т. М., Грищенко Л. М., Дакова Л. В., Лисенко Д. О., Кравченко В. І. Використання розподілених транспортних радіомереж терагерцового діапазону в рамках побудови мереж мобільного зв’язку нового покоління. *Зв’язок.* 2016. № 6. С. 16–21.

13. Сайко В. Г., Лисенко Д. О., Грищенко Л. М., Дакова Л. В., Кравченко В. І. Метод визначення оптимальних параметрів вікон прозорості в терагерцовому діапазоні. *Телекомунікаційні та інформаційні технології*. 2017. № 1. С. 11–17.

14. V. Saiko, V. Nakonechnyi, S. Toliupa, D. Serhrii. The method for reducing probability of incorrect data reception in radio channels of terahertz frequency range. *14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TSCET) Conference Proceedings.* Lviv-Slavske, Ukraine, February 20–24. 2018. S 11. № 215. # 174. 1043–1046.

15. Сайко В. Г., Наритник Т. М. Радіоканал доступу терагерцового діапазону для безпроводних радіосистем 5-го покоління. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: технічні науки*. 2018. Том 29 (68). № 1, частина 1. С. 51–55.

16. Сайко В. Г., Наритник Т. М., Баховський П. Ф. Надвисокошвидкісний канал радіодоступу терагерцового діапазону для мобільних мереж 4-го та 5-го поколінь. *Технічні вісті*. 2018. № 1 (47). С. 50–55.

17. Сайко В. Г., Нарытник Т. Н., Лутчак А. В. Анализ технических решений создания коммуникационных каналов в нелицензионном частотном диапазоне. *Цифрові технології*. 2016. № 20. С. 30–45.

18. Сайко В. Г., Наконечний В. С., Толюпа С. В., Даков С. Ю. Терагерцовий канал радіодоступу для комплексів безпеки систем виявлення прихованих об’єктів. *Кібербезпека: освіта, наука, техніка*. 2018. № 1(1). С. 17–25.

19. Спосіб захисту сфокусованим випромінюванням субтерагерцового діапазону зон і об’єктів від несанкціонованого проникнення: пат. 138429 Україна: МПК (2006): G08B15/00. № 201905643; заявл. 24.05.2019; опубл. 25.11.2019, Бюл. № 22.

20. Система низькоорбітального супутникового зв’язку із міжсупутниковими каналами зв’язку терагерцового діапазону: пат. 142478 Україна.№ 201911325; заявл. 21.11.2019; опубл. 10.06.2020, Бюл. № 11.

21. Kosovets M. A., Pavlov O. I., Tovstenko L. N. Integral-differential models of characteristic functions of 3D terahertz FMCW radar. *Proceedings of the International Conference on Antenna Theory and Techniques (ICATT’2015)*. 21–24 April. 2015. Kharkiv. Ukraine, 2015. pp. 225–227.

22. M. Kosovets, O. Pavlov, L. Tovstenko. Studying the properties of different materials using Terahertz 3D Imager Radar. *Proceedings of the International Conference on Antenna Theory and Techniques (ICATT’2017*). 24–27 May. 2017. Kyiv. Ukraine. 2017. pp. 406–410.

23. 2023 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo) - #61577 is now published in IEEE Xplore.

24. Ильченко М. Е., Кравчук С. А. Телекоммуникационные системы на основе высотных аэроплатформ. К.: НПП «Издательство «Наукова думка» НАН Украины». 580 с.

25. Рекомендация ITU-R P.838-3. Модель погонного ослабления в дожде, используемая в методах прогнозирования. (Вопрос МСЭ-R 201/3).

26. White paper on RF enabling 6G – Opportunities and challenges from technology to spectrum. *6G Research Visions*. April 2021. No. 13.

27. Jornet J. M., Akyildiz I. F. Channel Modeling and Capacity Analysis for Electromagnetic Wireless Nanonetworks in the Terahertz Band. *IEEE Transactions on Wireless Communications.* 2011. Vol. 10. Pp. 3211–3221.

28. Priebe S., Kurner T. Stochastic Modeling of THz Indoor Radio Channels. *IEEE Transac-tions on Wireless Communications.* 2013. Vol. 12. Pp. 4445–4455.

29. Han C., Bicen A. O., Akyildiz I. F. Multi-Ray Channel Modeling and Wideband Characteriza-tion for Wireless Communications in the Terahertz Band. *IEEE Transactions on Wireless Communications.* 2015. Vol. 14. Pp. 2402–2412.

30. 3GPP TS 37.340 V15.2.0: NR: Multi-connectivity; Overall description, Rel. 15 – 2018. URL: https://www.3gpp.org/ /ftp/Specs/archive/37_series/37.340/ (accessed: 31.07.2019).

31. Сайко В., Наритник Т. Модель побудови бездротової терагерцової мережі з підвищеною надійністю зв’язку. *International Science Journal of Engineering & Agriculture.* 2023. № 2(2), С. 166–181. URL: https://doi.org/10.46299/j.isjea.20230202.16 (дата звернення: 10.03.2024).

32. Сайко В. Г., Наритник Т. М., Баховський П. Ф. Модель підвищення показників якості обслуговування гетерогенної мережної інфраструктури терагерцового діапазону. *Вчені записки Таврійського національного університету імені В. І. Вернадського. серія: технічні науки*. 2023. Том 34 (73). № 1. С. 51–55.

33. Hu S., Rusek F., Edfors O. Beyond Massive MIMO: The Potential of Positioning With Large Intelli gent Surfaces. *IEEE Trans. Signal Process*. 2018. 66 (Apr), 1761–1774.

34. He J., Wymeersch H., Sanguanpuak T., Silvén O., Juntti M. Adaptive beamforming design for mmwave RIS-aided joint localization and communication. *IEEE Wireless Communications and Networking Conference Workshops (WCNCW). 2020.*

DOI: 10.58254/viti.5.2024.14.154

УДК 004.7

д-р техн. наук, професор Стрельбіцький М. А. ORCID: 0000-0001-8030-3228 (НАДПСУ)
канд. техн. наук Хоптинський Р. П. ORCID: 0000-0001-9351-7938 (НАДПСУ)
Ваврічен О. А. ORCID: 0000-0001-7777-9188 (НАДПСУ)
Городиський Р. О. ORCID: 0000-0002-0918-864X (НАДПСУ)

ПІДХІД ДО СТВОРЕННЯ МОДЕЛІ ПРОГРАМНО-ОРІЄНТОВАНОЇ КОМУНІКАЦІЙНОЇ МЕРЕЖІ ДЕРЖАВНОЇ ПРИКОРДОННОЇ СЛУЖБИ УКРАЇНИ

Підвищення рівня захисту інформаційних ресурсів Державної прикордонної служби України (далі – Держприкордонслужби) як складової сил оборони України є важливим завданням, особливо в умовах воєнного стану. Ефективна та надійна корпоративна мережа прикордонного відомства стає ключовою складовою для забезпечення безпеки, координації та оперативності реагування на внутрішні та зовнішні загрози. Одним із аспектів розвитку інформаційного простору став розвиток транспортної мережі Держприкордонслужби. Аналіз існуючих технологій, визначення їхніх переваг та недоліків дозволив виокремити технологію Software-Defined Wide Area Networking як таку, що найбільш повно відповідає критеріям відповідності, що визначені експертами Управління зв'язку та інформаційних систем. Зазначена технологія дозволяє не лише оптимізувати мережевий трафік та забезпечити високу швидкість передачі даних, але й гарантує гнучкість в управлінні мережею та підвищує рівень безпеки. Для Держприкордонслужби, яка відповідає за безпеку кордонів країни, використання передових технологій Software-Defined Wide Area Networking може мати вирішальне значення для оптимізації роботи та забезпечення найвищого рівня захисту інформації.

У статті приділено увагу розробці концептуальної моделі для побудови корпоративної мережі Держприкордонслужби на основі технології Software-Defined Wide Area Networking. Основною метою дослідження є розгляд можливості впровадження цієї технології в контексті її застосування силами оборони України та аналіз її переваг і недоліків. Завданням дослідження було вивчення технічних аспектів впровадження Software-Defined Wide Area Networking, а саме: конфігурації мережевих пристроїв, управління трафіком, моніторингу безпеки та захисту даних.

Слід зазначити, що Cisco Software-Defined Wide Area Networking застосувала оптимальний підхід до архітектури мережі, який відокремлює площину керування від площини даних усіх периферійних маршрутизаторів і реалізує всі функції керування в централізованому програмному контролері під назвою vSmart. Практичні рекомендації дослідження дозволили забезпечити безперервне мережеве обслуговування шляхом зміни маршруту в обхід збоїв і потенційних причин простою.

Ключові слова: *модель, комунікаційна мережа, Держприкордонслужба, технологія програмно-орієнтованих мереж, управління мережею, Software-Defined Wide Area Networking.*

M. Strelbitskyi, R. Khoptynskyi, O. Vavrichen, R. Horodyskyi Approach to creation model of software-oriented communication network of the State Border Guard Service of Ukraine.

Enhancing the level of information resource protection of the State Border Guard Service of Ukraine, as a component of Ukraine's defense forces, is an important task, especially in times of war. An effective and reliable corporate network of the border agency becomes a key component for ensuring security, coordination, and responsiveness to both internal and external threats. One aspect of developing the information space is the development of the transport network of the State Border Guard Service. Analysis of existing technologies, identifying their advantages and disadvantages, allowed for the identification of Software-Defined Wide Area Networking technology as the one that best meets the conformity criteria defined by the experts of the Communication and Information Systems Management. This technology not only allows for optimizing network traffic and ensuring high-speed data transmission but also guarantees flexibility in network management and enhances security. For the State Border Guard Service, responsible for the security of the country's borders, the use of advanced Software-Defined Wide Area Networking technologies can be crucial for optimizing operations and ensuring the highest level of information security.

The article focuses on developing a conceptual model for building the corporate network of the State Border Guard Service based on Software-Defined Wide Area Networking technology. The main goal of the research is to consider the possibility of implementing this technology in the context of its application by Ukraine's defense forces and to analyze its advantages and disadvantages. The research task was to study the technical aspects of implementing Software-Defined Wide Area Networking, namely: configuration of network devices, traffic management, security monitoring, and data protection.

Cisco's Software-Defined Wide Area Networking has applied an optimal approach to network architecture, which separates the control plane from the data plane of all peripheral routers and implements all control functions in a

centralized software controller called vSmart. Practical recommendations from the research allowed for ensuring continuous network service by rerouting around failures and potential causes of downtime.

Keywords: *model, communication network, State Border Guard Service, Software-Defined Networking technology, network control, Software-Defined Wide Area Networking.*

Постановка проблеми у загальному вигляді. Функціонування інформаційно-комунікаційних систем в умовах воєнного стану передбачає значного підвищення рівня керованості мережевою складовою та безпеки інформаційних ресурсів Держприкордонслужби як складової сил оборони України. В умовах постійних викликів, які виникають при виконанні прикордонним відомством функціональних завдань, ефективна та надійна корпоративна мережа стає ключовою складовою для забезпечення безпеки, координації та оперативності реагування на внутрішні та зовнішні загрози.

У цьому контексті технологія програмно-орієнтованої комунікаційної мережі SDN (Software – Defined Networking) дозволяє не лише оптимізувати мережевий трафік та забезпечувати високу швидкість передачі даних, але й забезпечує гнучкість управління мережею та підвищує рівень безпеки. Для Держприкордонслужби, яка відповідає за безпеку кордонів держави, використання передових технологій SDN може мати вирішальне значення для оптимізації роботи та забезпечення високого рівня захисту інформації.

Аналіз останніх публікацій. Протягом останніх років ситуація в інформаційних та комунікаційних системах суттєво змінювалась завдяки імплементації хмарних технологій [1–3]. Розвиток комунікаційних засобів сприяв збільшенню швидкостей обміну даними, що спонукало застосуванню технології Ethernet як стандарту для всіх каналів зв'язку. Це призвело до того, що різниця між локальною мережею (LAN) і глобальною мережею (WAN) стала практично мінімальною. Разом із тим, незмінними залишились основні функції комутації та маршрутизації, які зазвичай реалізовані на апаратному автономному пристрої. З точки зору даних ці пристрої є самодостатніми стосовно завдань комутації або маршрутизації.

З метою збереження переваги на ринку комунікаційних засобів різні компанії, такі як Cisco, Juniper і HP, вдосконалили свої пристрої функціями, що призвели до появи мережевих засобів, які ґрунтуються на застарілих протоколах, наприклад, Border Gateway Protocol (BGP) [4]. Використання цих протоколів повинно було забезпечити інкапсуляцію заголовків на різних рівнях. Разом із тим, такий підхід зменшив максимальний розмір передачі пакетів (MTU). Зазначимо, що таке нашарування абстракцій не сприяє управлінню мережею, де шаблони трафіку визначаються в кожному шарі незалежно.

У цьому контексті варто зазначити потребу використання протоколу BGP для обміну інформацією про маршрутизацію та доступність вузлів. Це пов'язано із тим, щоб міграція всіх складових корпоративної мережі SDN відбувалася поступово і таким темпом, який не становить загрози безпеці та керованості інформаційно-комунікаційних систем, які функціонують на основі Intranet-мережі [5].

Інтегрована інформаційно-комунікаційна система Держприкордонслужби забезпечує автоматизацію практично всіх завдань інформаційного характеру, які виконують органи та підрозділи прикордонного відомства. Особливістю організаційної структури Держприкордонслужби є значна розосередженість прикордонних підрозділів, що, в свою чергу, вимагає використання різнорідної якості обслуговування передачі даних з метою забезпечення функціонування засобів автоматизації. Постійне вдосконалення складових мережі призводить до її варіантності стосовно типів комунікаційного обладнання та складності керування такою мережею.

Тому для Держприкордонслужби стає актуальним застосування автоматизованого програмного підходу при підключенні до інформаційно-комунікаційної систем її підрозділів. Застосування програмно-керованої глобальної мережі (SD-WAN)SD-WAN розширить програмно-визначені мережі (SDN) в рішення, яке підрозділи зможуть використовувати для

швидкого створення інтелектуальної гібридної Wan мережі, що включить MPLS, LTE, широкосмуговий Інтернет і бездротові підключення підрозділів.

Метою статті, є розробка адаптованої концептуальної моделі архітектури SD-WAN комунікаційної мережі Держприкордонслужби на основі технологій інтелектуального управління.

Виклад основного матеріалу дослідження.

Запобігання ризикам і загрозам у прикордонній сфері потребує ефективної та надійної комунікаційної інфраструктури. Швидкість передачі даних, безпека і доступність є критично важливими аспектами у контексті функціонування прикордонного відомства. Аналіз стратегій впровадження технології SD-WAN для оптимізації передачі даних та забезпечення стабільної та безпечної роботи комунікаційної мережі дозволить визначити конкретні вимоги і потреби Держприкордонслужби, розглянути варіанти конфігурацій та здійснити вибір оптимальних рішень для створення високоефективної, масштабованої та безпечної мережі, яка забезпечує високу надійність комунікацій у державних структурах.

При вирішенні питання про перенесення традиційної архітектури WAN на програмно-визначену WAN застосовують поетапний процес, зокрема розгортання контролерів, міграція основних центрів обробки даних і вузлів і, нарешті, віддалених сайтів, таких як підрозділи та філії. Основні засади виконання такої послідовності полягають в тому, щоб сайти-концентратори направляли трафік між SD-WAN і не-SD-WAN-вузлами протягом періоду міграції. Однією з головних переваг програмно-визначеної глобальної мережі є можливість розгортання контролерів у загальнодоступній хмарі. Це може значно зменшити витрати і підвищити загальну доступність і резервування площини керування/контролю. Cisco пропонує на вибір наступні варіанти: хмара, розміщена на Cisco; загальнодоступна хмара; розгортання контролерів у центрах обробки Держприкордонслужби. Саме останній варіант пропонується для вибору з причини «чутливості» прикордонного інформаційного ресурсу. Такий підхід передбачає відповідальність підрозділів зв’язку за резервне копіювання та аварійне відновлення.

За умови налаштування та запуску контролерів, вони повинні встановити безпечні з’єднання між собою. На сьогодні є два варіанти на вибір, коли мова йде про базовий безпечний протокол: TLS, який використовує транспорт TCP, і DTLS, який використовує транспорт UDP. За замовчуванням усі контролери використовують параметр DTLS. При розгортанні SD-WAN у середовищі нульової довіри інформація передається для всіх постійних з’єднань між контролерами. Слід зазначити, що кожне ядро на vManage та vSmart створює постійне з’єднання DTLS з vBond, що призводить до чотирьох з’єднань між vManage та vBond і двох з’єднань між vSmart і vBond.

Безпечне підключення граничних пристроїв WAN є дуже важливою частиною рішення SD-WAN. Рішення Cisco SD-WAN використовує модель білого списку для автентифікації та довіри до пристроїв vEdge [6]. Це означає, що перед тим, як прикордонному маршрутизатору WAN буде дозволено приєднатися до рівня керування, він повинен бути заздалегідь відомий усім контролерам SD-WAN. Кожен пристрій унікально ідентифікується ідентифікатором шасі та серійним номером сертифіката. Коли контролери SD-WAN розгорнуті та отримали дійсні сертифікати, периферійні маршрутизатори WAN можуть почати процес адаптації. На цьому етапі найважливіше переконатися, що пристрої vEdge мають доступ до всіх контролерів через усі доступні транспортні засоби.

Розглянемо типовий сценарій, коли віддалений сайт має один приватний канал MPLS і одне широкосмугове підключення до Інтернету. Є три поширені реалізації, які вирішують таке завдання:

1. MPLS має доступ до публічної хмари завдяки маршрутизації через центр обробки даних або регіональний хаб, який має обидва транспортні засоби.

2. Загальнодоступні маршрутизовані IP-адреси контролерів перерозподіляються в хмару MPLS, а граничний маршрутизатор постачальника повідомляє про них vEdge.

3. Встановлюється з'єднання з площиною управління лише через Інтернет-з'єднання. Edge зможе приєднатися до структури SD-WAN, але не матиме резервування площини керування, тому цей підхід взагалі не рекомендується.

Фундаментальна мета кожної мережі – забезпечити безперервне мережеве обслуговування шляхом зміни маршруту в обхід збоїв і потенційних причин простою. Cisco SD-WAN забезпечує високу доступність завдяки поєднанню чотирьох принципів:

- резервування контролерів;
- vEdges-резервування;
- надійний дизайн мережі;
- аварійного відновлення.

Щодо резервування контролерів, то усі контролери SD-WAN працюють як віртуальні машини або контейнери. Незалежно від методу розгортання основним принципом високої доступності є наявність кількох контролерів кожного типу, бажано розгорнутих у розрізнених географічних місцях. Це гарантує, що централізована площина управління залишається стійкою, якщо один із контролерів виходить з ладу.

Оскільки кожен контролер SD-WAN виконує різні функції, то кожен їхній тип використовує різні технології масштабування та підключення, як показано на рисунку 1 [7].

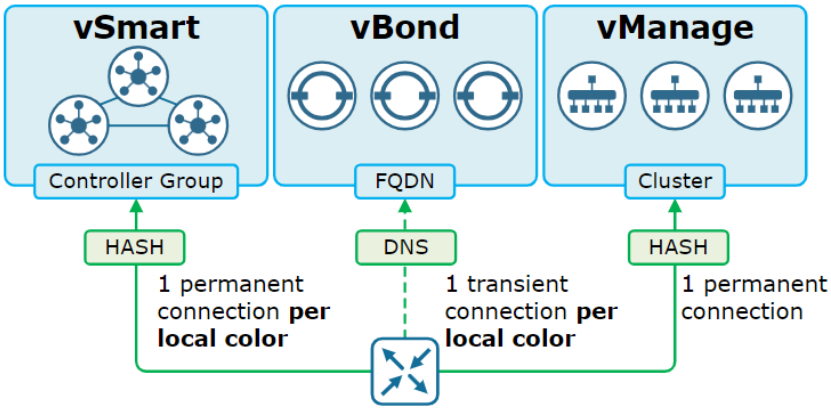


Рис. 1. Принципи високої доступності Cisco SD-WAN

Оркестратор Cisco vBond виконує дві основні функції в домені SD-WAN:

- перевіряє та автентифікує всі пристрої, які намагаються приєднатися до домену SD-WAN;
- керує встановленням керуючих з'єднань між контролерами та маршрутизаторами vEdge.

Оркестратор Cisco vBond працює як віртуальна машина локально або в хмарі. Однак це єдиний контролер, який також може працювати на звичайному маршрутизаторі vEdge, налаштованому для роботи як оркестратор vBond.

Високодоступну мережу Cisco SD-WAN, яка має кілька контролерів vBond, що працюють в режимі активний/активний, бажано розгортати в різних локальних географічних розташуваннях або хмарних регіонах. Кожен із пристроїв SD-WAN посилається на оркестратор vBond за одним іменем FQDN у своїй системній конфігурації.

Контролери Cisco vSmart керують централізованою площиною керування мережею. Вони встановлюють постійні з'єднання DTLS/TLS з усіма пристроями SD-WAN у домені. Через ці канали керування вони регулярно обмінюються даними про мережевий домен, щоб забезпечити синхронізацію їхніх централізованих таблиць маршрутизації.

Високодоступна мережа Cisco SD-WAN має кілька контролерів vSmart, які бажано розгортати в різних географічних локаціях або хмарних сервісах. Кожен маршрутизатор vEdge за замовчуванням встановлює керуючі з'єднання з двома контролерами vSmart, як показано на рисунку 2 [7]. Коли один із контролерів виходить з ладу, інший безперешкодно бере на себе функції рівня керування мережею. Площина керування мережею працює безперебійно, доки в домені SD-WAN працює один контролер.

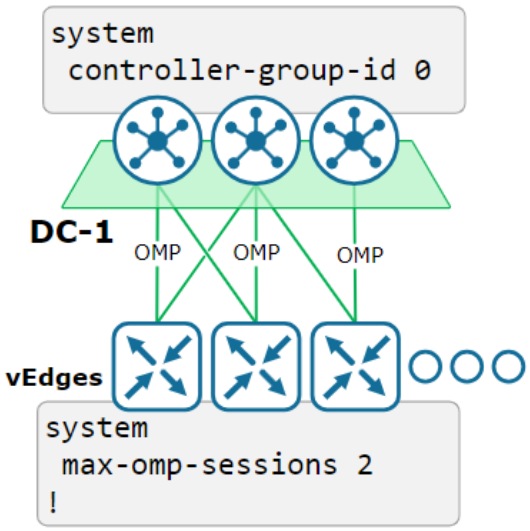


Рис. 2. Забезпечення високої доступності vSmart

Слід зазначити, що контролери vSmart встановлюють і підтримують повну мережу керуючих з'єднань між собою, над якою вони формують повну мережу сеансів OMP. Усі контролери синхронізують свою інформаційну базу маршрутизації, обмінюючись маршрутами, маршрутами TLOC, політиками та ключами шифрування. Крім того, кожен контролер встановлює постійне керуюче з'єднання з кожним оркестратором vBond. Потім ці канали керування використовуються оркестратором vBond для відстеження того, які vSmarts працюють у домені. Коли один із контролерів виходить з ладу, оркестратор перестане надавати IP-адресу недоступного vSmart маршрутизаторам vEdge, які приєднуються до домену SD-WAN.

Оскільки домен SD-WAN охоплює кілька географічних регіонів, зазвичай для відмовостійкості додається більше контролерів vSmart. Як правило, прикордонне відомство розгортає контролери в більш ніж двох регіонах, тому важливо переконатися, що vEdges підключаються до vSmarts в тому самому або суміжному географічному регіоні. Наприклад, існує три центри обробки даних: один на сході, один посередині та один на заході. У кожному центрі обробки даних існує окрема група контролерів vSmart, як показано на рисунку 3 [7].

Ієрархічна SD-WAN – це варіант дизайну використання Cisco SD-WAN, який надає можливість розділяти мережу WAN на незалежні регіони. Подібно до логіки області OSPF, ієрархічна мережа SD-WAN завжди повинна мати основну область, яка називається областю «0», що з'єднує всі інші області доступу. Одна з істотних переваг ієрархічної архітектури SD-WAN полягає в тому, що вона чітко відокремлює внутрішньо-регіональний трафік від міжрегіонального трафіку. Міжрегіональний трафік повністю обробляється набором виділених маршрутизаторів, які називаються прикордонними маршрутизаторами (BR), що створюють основну область «0». Подібно до OSPF, трафік, що надходить з одного регіону і призначений для іншої області, завжди проходить через основну область «0». Це вирішує сценарій використання роз'єднаних постачальників послуг. Основний регіон забезпечує зв'язок між двома регіонами, які використовують незв'язаних провайдерів.

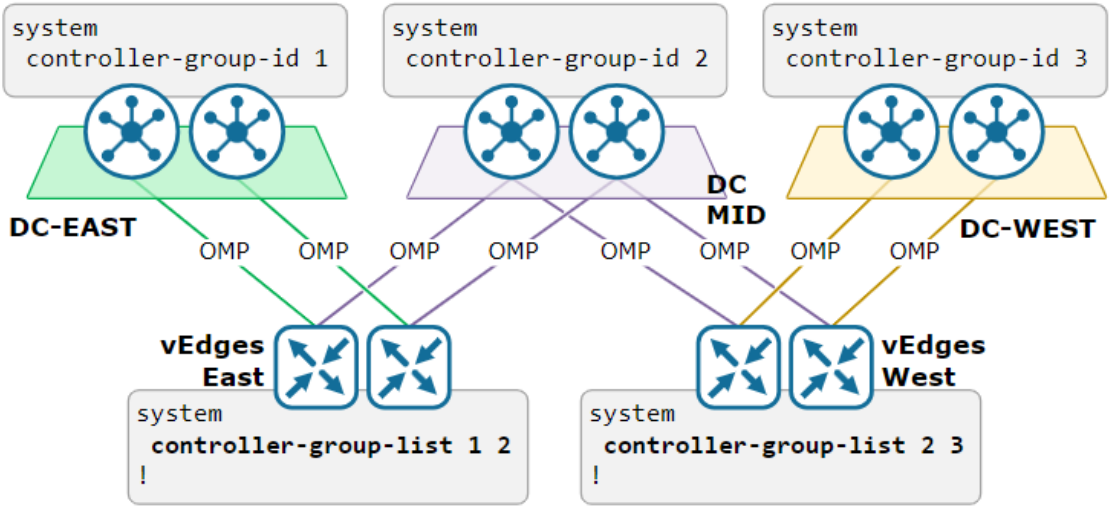


Рис. 3. Спорідненість контролера vSmart

Розділення міжрегіонального трафіку в незалежний регіон дозволяє Держприкордонслужбі використовувати різних постачальників послуг для магістральних каналів зв'язку і застосовувати різні політики до магістральних мереж передачі даних.

Ще однією великою перевагою ієрархічної архітектури є гнучкість використання різних постачальників послуг у кожному окремому регіоні. Це дозволяє вибрати найбільш економічно ефективного постачальника в кожному географічному регіоні без шкоди для доступності та керованості мережі. Крім того, можливо використовувати різні інфраструктури маршрутизації та різні політики трафіку для кожного регіону, що притаманно прикордонному відомству.

Ієрархічна архітектура SD-WAN дозволяє організації призначати різні виділені контролери vSmart для кожного незалежного регіону. Це значно спрощує розробку політики організації. Технічно кажучи, якщо регіон містить лише невелику кількість маршрутизаторів vEdge, пару контролерів можна призначити кільком регіонам. Однак для простоти та кращої керованості настійно рекомендується, щоб у кожному регіоні обслуговувалися спеціальні контролери vSmart.

Регіони доступу – це логічне групування маршрутизаторів vEdge, які обслуговують певну географічну область. Наразі максимальна кількість підтримуваних регіонів для розгортання SD-WAN становить вісім.

Ієрархічна архітектура SD-WAN накладає деякі дуже серйозні обмеження на домен накладання:

1. Мережеві пристрої, налаштовані на використання ієрархічної SD-WAN, не можуть підключатися до звичайних пристроїв SD-WAN. Для використання ієрархічної SD-WAN слід налаштувати всі пристрої або жоден із них. Таким чином, увімкнення ієрархічної SD-WAN для існуючого мережевого середовища є достатньо складним процесом.
2. Наскрізна маршрутизація з урахуванням програми не працює. AAR слід використовувати лише в межах регіону.
3. Мультитенантність не підтримується в ієрархічній SD-WAN.
4. Агрегація маршрутів протоколу керування накладенням (OMP) на прикордонних маршрутизаторах.
5. Міжрегіональна багатоадресна IP-адреса не підтримується в ієрархічній SD-WAN.

Програмно-орієнтовані рішення Cisco WAN пропонують багато функцій безпеки, орієнтованих на глобальну мережу та корпоративні мережі. Проте слід зазначити, що безпека SD-WAN є лише частиною ширшої стратегії безпеки більшості організацій.

Secure Access Service Edge (SASE) – це новий підхід до глобальної інфраструктури організації, який поєднує програмно-визначену глобальну мережу з розширеними функціями безпеки хмарної доставки. Інфраструктура SASE має на меті поєднати функції мережі, безпеки та ідентифікації в єдине уніфіковане рішення, що надається як послуга. SASE є відповіддю на поточні виклики переходу технологій до децентралізованих мереж і безпеки. Важливо розуміти, що SASE – це не окремий продукт, рішення чи функція. Це структура або філософія, яка поєднує пакет різних технологій. Таким чином, він не конкурує безпосередньо з будь-якими рішеннями і не замінює будь-який конкретний продукт.

Нині існує величезна складність, пов'язана з мережею та безпекою Cisco, а саме:

вибір – різні мережеві рішення і рішення безпеки (Meraki, Viptela, Umbrella, ThousandEyes, Duo тощо);

придбання – різні моделі ліцензування;

розгортання – кожне окреме рішення має складний процес навчання та вимагає високого рівня досвіду під час розгортання;

експлуатація – кожне окреме рішення потребує окремого набору інженерів для його експлуатації;

масштабування/оновлювання – з багатьма з'єднаними рішеннями стає дуже важко вносити зміни в архітектуру.

За допомогою SASE Cisco намагатиметься об'єднати все в одне рішення та знизити загальну складність на всіх можливих рівнях.

Програмно-визначена глобальна мережа не призначена для вирішення всіх проблем, з якими стикається мережа та інфраструктура безпеки під час переходу до децентралізованої хмарної моделі мережі. Ось деякі з помітних недоліків:

1. SD-WAN потребує надійної основи. Програмно-визначені рішення WAN створюють оверлейну структуру поверх основної інфраструктури глобальної мережі. Організаціям все ще потрібна надійна мережева магістраль. Управління та захист основної інфраструктури у великомасштабних багаторегіональних варіантах побудови все ще може бути складним і дорогим. Сама собою SD-WAN не вирішує основних проблем.

2. Віддалений працівник/віддалений доступ. Програмно-визначена накладна структура забезпечує безпечне та надійне підключення «сайт до будь-якого сайту» та «сайт до будь-якої хмари». Однак це рішення не призначене для надання віддаленого SSL VPN для мобільних працівників або для захисту конфіденційних корпоративних даних від віддаленого доступу. Сама собою SD-WAN не вирішує проблеми віддаленого працівника/віддаленого доступу.

3. Відсутність повного портфоліо безпеки. Головним завданням більшості програмно-визначених рішень WAN є допомога в автоматизації та масштабуванні інфраструктури глобальної мережі. Так, портфоліо Cisco SD-WAN містить багато можливостей безпеки. Однак SD-WAN сама собою не може вирішити всі проблеми безпеки, з якими стикаються організації під час впровадження децентралізованої хмарної інфраструктури та віддаленого доступу. Вона може інтегруватися з хмарними провайдерами безпеки, такими як Umbrella та Duo.

Поточна структура Cisco Secure Access Service Edge має три основні компоненти: Cisco Software-defined WAN, Cisco Umbrella, Cisco Duo.

На рисунку 4 показана діаграма високого рівня повної інтеграції трьох основних рішень [8].

Cisco Umbrella Secure Internet Gateway (SIG) об'єднує численні функції безпеки в одному хмарному рішенні, яке традиційно вимагало локальних пристроїв безпеки (FW, IPS, проксі). Umbrella поєднує в собі хмарний брандмауер, захищений вебшлюз (SWG), інспекцію рівня DNS, брокер безпеки доступу до хмари (CASB), запобігання втраті даних (DLP) і віддалену ізоляцію браузера (RBI) в одну хмарну службу, яка легко інтегрується з Cisco SD-WAN.

Мережевий шлюз Cisco Duo (DNG) дозволяє кінцевим користувачам отримувати доступ до локальних ресурсів, не турбуючись про керування обліковими даними VPN, а також забезпечує безпеку за допомогою Duo Multi-factor Authentication (MFA). Duo забезпечує детальний контроль доступу для програми та групи користувачів. Це гарантує, що лише довірені користувачі та кінцеві точки можуть отримати доступ до внутрішніх ресурсів організації.

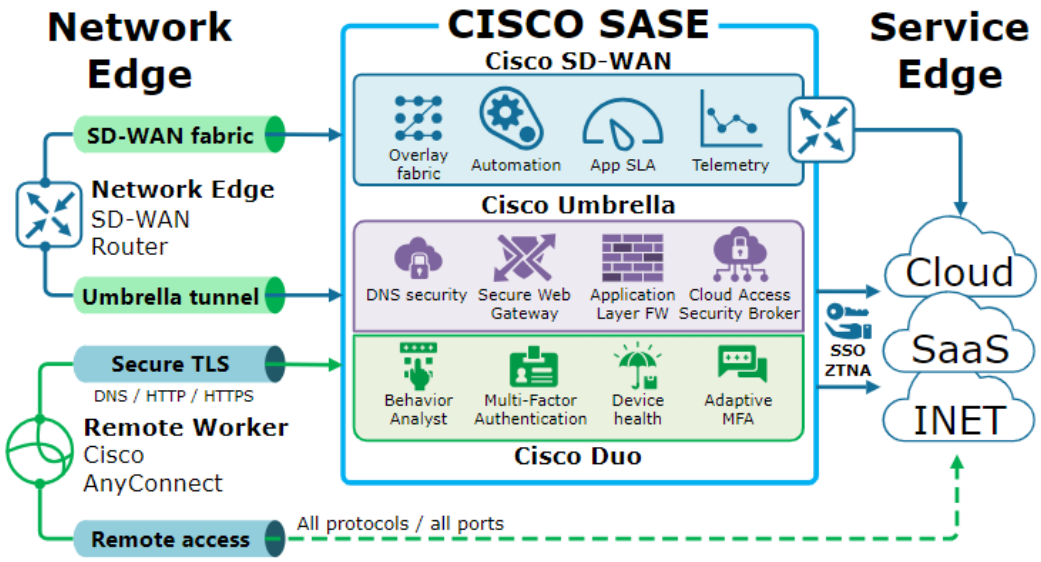


Рис. 4. Cisco SD-WAN and SASE

Разом і тим, рішення щодо високої доступності та безпеки часто суперечать одне одному. Cisco застосувала зовсім інший підхід до безпеки WAN, який базується на трьох основних принципах:

1. Fabric Security – рішення гарантує, що всі пристрої, які беруть участь у мережі, є справжніми та надійними. Весь зв'язок між мережевими пристроями автоматично шифрується.

2. Інтегрована безпека – рішення інтегрує всі функції безпеки, такі як брандмауер, IPS і AMP, у мікропрограму маршрутизаторів, усуваючи потребу в окремих виділених апаратних пристроях, які виконують функції безпеки.

3. Хмарна безпека – забезпечує повну інтеграцію з кількома постачальниками хмарної безпеки, що робить перехід до гібридної моделі безпеки дуже легким.

Першим кроком у стратегії мережевої безпеки є забезпечення надійності всіх мережевих пристроїв. Із традиційними механізмами безпеки, такими як «цей маршрутизатор розгорнуто нашим довіреним інженером» і «ми використовуємо автентифікацію BGP/OSPF/PIM/HSRP», постає питання, чи є впевненість, що використовується надійна інфраструктура? У цьому контексті варто зауважити, що забезпечення безпеки повинно починатись задовго до того, як спрацюють механізми безпеки програмного забезпечення, такі як автентифікація протоколу. Останніми роками спостерігається стрімке зростання кількості атак і експлойтів, наприклад, багато апаратних компонентів, таких як оперативна пам'ять, твердотільні накопичувачі та процесори, замінюються на скомпрометовані з попередньо встановленими троянськими програмами. Сьогодні безпека починається з обладнання під час виробництва. Усі фізичні пристрої Cisco мають модуль довіри (TAm), встановлений під час виробництва. Модуль TAM надає кілька надійних технологій:

1. Безпечне завантаження. Мікрозавантажувач, встановлений у модулі TAM, відстежує процес запуску та захищає від шкідливого коду під час завантаження.

2. Безпечне зберігання. Модуль TAm забезпечує безпечне зберігання криптографічних ключів, паролів, облікових даних клієнтів та іншої важливої інформації безпеки для пристрою. Однією з його переваг є можливість зберігати приватні ключі шифрування та паролі. Також можливе виділення безпечного сховища за межами модуля TrustAnchor.

3. Захищений унікальний ідентифікатор пристрою (SUDI). SUDI – це сертифікат SSL (X.509v3), який містить серійний номер пристрою та ідентифікатор продукту. Його встановлюють під час виробництва та перевіряють загальнодоступним центром кореневих сертифікатів. Сертифікат SUDI SSL разом із пов’язаними ключами зберігається в апаратному чипі Trust Anchor Module (TAm). Крім того, приватний ключ ніколи не може бути розкритий, оскільки пара ключів криптографічно прив’язана до конкретного чипа TAm.

4. Захист під час виконання (RTD). RTD захищає від впровадження шкідливого коду в мікропрограму під час виконання.

Ідентифікація для всіх пристроїв SD-WAN забезпечується сертифікатами SSL (x509v3). Усі пристрої постачаються з попередньо встановленими кореневими сертифікатами DigiCert, Symantec і Cisco. Крім того, усі фізичні маршрутизатори постачаються з попередньо завантаженим сертифікатом пристрою, встановленим під час виробництва та захищеним апаратним чипом (SUDI). Усі програмні пристрої, такі як хмарні маршрутизатори та контролери, не мають попередньо завантажених сертифікатів пристроїв і повинні пройти процес запиту на підписання сертифіката (CSR). Важливим моментом тут є те, що кожен пристрій SD-WAN повинен мати дійсний сертифікат SSL, який криптографічно підтверджує його ідентифікацію. Коли контролери встановлюють керуючі з’єднання один з одним, вони обмінюються SSL-сертифікатами своїх пристроїв під час процесу встановлення зв’язку DTLS/TLS. Потім кожен контролер перевіряє такі параметри, а саме:

перевіряє довіру для отриманого сертифіката пристрою за допомогою його попередньо встановлених кореневих сертифікатів;

порівнює серійний номер сертифіката зі списком авторизованих контролерів, який розповсюджується з vManage;

порівнює назву організації в полі OU отриманого сертифіката з локально налаштованою назвою організації.

Слід зазначити, що існує різниця в процесі перевірки для віртуальних і фізичних маршрутизаторів. Сертифікат SSL фізичного периферійного маршрутизатора встановлюється під час виробництва. Таким чином, неможливо закодувати назву організації в сертифікаті, оскільки на момент виготовлення невідомо, яка саме організація використовуватиме цей маршрутизатор.

З іншого боку, віртуальні маршрутизатори генерують запит на підписання сертифіката (CSR) після того, як ім’я організації налаштовано на маршрутизаторі, тому поле OU буде присутнє в сертифікаті SSL після того, як CSR буде підписано центром сертифікації. Якщо всі етапи перевірки пройдено успішно, маршрутизатор встановлює постійне з’єднання DTLS/TLS із контролером SD-WAN.

Коли два пристрої SD-WAN встановлюють захищене керуюче з’єднання DTLS/TLS, цілісність рівня керування забезпечується двома елементами безпеки: дайджестами повідомлень AES-GCM та парою відкритих і закритих ключів.

AES-GCM забезпечує як шифрування, так і можливість перевіряти цілісність повідомлень контрольної площини, надісланих через тунелі DTLS/TLS. Він генерує дайджести повідомлень (просто дайджести) для кожного пакета, надісланого через захищені тунелі DTLS/TLS. Пристрій-одержувач потім генерує дайджест для пакета, і якщо вони збігаються, пакет перевіряється. Це підтверджує, що вміст пакета не було змінено під час передавання.

Другий компонент – це використання відкритих і закритих ключів. Коли встановлено з’єднання з площиною управління, локальний пристрій надсилає виклик віддаленому.

Віддалений пристрій шифрує виклик, підписуючи його своїм закритим ключем, і надсилає виклик на локальний пристрій, який потім використовує відкритий ключ віддаленого пристрою, щоб перевірити, що отриманий виклик відповідає надісланому виклику.

Ключі використовуються для того, щоб переконатися, що пакети були надіслані надійним пристроєм. Автентичність кожного пакета перевіряється шляхом шифрування та дешифрування за допомогою симетричних ключів, якими обмінювалися під час встановлення контрольного з'єднання.

Безпечна площина керування гарантує, що власна безпека мережевих протоколів, таких як SNMP і NETCONF, не викликає занепокоєння для організації, оскільки весь зв'язок на рівні керування проходить через з'єднання DTLS/TLS, отже, автентифікується та шифрується.

Площина даних (також називається площиною пересилання) забезпечує інфраструктуру для надсилання трафіку користувача через мережу накладання SD-WAN. Трафік площини даних зазвичай проходить у безпечних тунелях Internet Security (IPsec).

Висновки. Проведені дослідження та аналіз вимог до комунікаційної мережі Держприкордонслужби дозволили визначити потребу в параметрах високої доступності, які є ключовими для успішної реалізації цієї архітектури. Представлена ієрархічна модель архітектури цифрової мережі SD-WAN дозволяє оптимально розподіляти трафік і забезпечувати високу продуктивність мережі. Ця модель стане основою для подальшого розгортання та вдосконалення мережі Державної прикордонної служби України, забезпечуючи надійний та швидкий обмін даними.

Розроблена концептуальна модель дозволить підвищити ефективність та надійність комунікаційної інфраструктури Державної прикордонної служби України. Запровадження адаптованої концептуальної моделі архітектури цифрової мережі SD-WAN є кроком у цьому напрямку. Дані дослідження допоможуть у подальших роботах із розгортання та оптимізації мережі, забезпечуючи безперервний доступ до важливої інформації та покращуючи комунікаційні можливості Державної прикордонної служби України.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Borovick L., Mehra R. Architecting the Network for the Cloud. White paper. *IDC. Analize the Future.* 2011. URL: https://www.cisco.com/c/dam/global/en_ca/solutions/midsize/docs/idc_architecting_the_network_for_the_cloud.pdf (date of access: 15.02.2024).

2. Azodolmolky S., Wieder P., Yahyapour R. Cloud computing networking: Challenges and opportunities for innovations. *IEEE Communications Magazine.* 2013. Vol. 51, no. 7. P. 54–62.

3. Banikazemi M. et al. Meridian: an SDN platform for cloud network services. *IEEE Communications Magazine.* 2013. Vol. 51, no. 2. P. 120–127.

4. Diarmuid Seosamh Ó. Briain. Department of Aerospace, Mechanical and Electronic Engineering. 2015.

5. BGP Routing Interoperability. URL: https://www.silver-peak.com/products/unity-edge-connect/bgp-routing (date of access: 15.02.2024).

6. How Cisco SD-WAN works? *NetworkAcademy.* URL: https://www.networkacademy.io/ccie-enterprise/sdwan/how-cisco-sd-wan-works (date of access: 15.02.2024).

7. Cisco SD-WAN High Availability. *NetworkAcademy.* URL: https://www.networkacademy.io/ccie-enterprise/sdwan/high-availability (date of access: 15.02.2024).

8. What is SASE? *NetworkAcademy.* URL: https://www.networkacademy.io/ccie-enterprise/sdwan/what-is-sase (date of access: 15.02.2024).

9. Monaco, Matthew, Oliver Michel and Eric Keller. Applying operating system principles to SDN controller design. // Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks. ACM, 2013.

10. Kannan, Kalapriya and Subhasis Banerjee. Compact TCAM: Flow entry compaction in TCAM for power aware SDN. // International Conference on Distributed Computing and Networking. Springer Berlin Heidelberg, 2013.

11. Banikazemi, Mohammad et al. Meridian: an SDN platform for cloud network services. // IEEE Communications Magazine 51.2. 2013. P. 120–127.

12. Hu, Hongxin et al. Towards a reliable SDN firewall. Presented as part of the Open Networking Summit 2014 (ONS 2014). 2014.

13. Masoud, Mohammad Z., Yousf Jaradat and Ismael Jannoud. On preventing ARP poisoning attack utilizing Software Defined Network (SDN) paradigm. Applied Electrical Engineering and Computing Technologies (AEECT), 2015 IEEE Jordan Conference on. IEEE, 2015.

DOI: 10.58254/viti.5.2024.15.165

УДК 004.056(53+57)

д-р техн. наук, професор Субач І. Ю. ORCID: 0000-0002-9344-713X
(ІСЗЗІ НТУУ «КПІ ім. Ігоря Сікорського»)
Власенко О. В. ORCID: 0000-0001-6671-870X (ВІТІ ім. Героїв Крут)

НЕЧІТКІ МОДЕЛІ ВИЯВЛЕННЯ КІБЕРІНЦИДЕНТІВ У БАЗАХ ДАНИХ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ СИСТЕМ ВІЙСЬКОВОГО ПРИЗНАЧЕННЯ

Захист баз даних інформаційно-комунікаційних систем військового призначення є надзвичайно важливим завданням у сучасній сфері кібербезпеки. Зростаючі загрози від кібератак, необхідність ефективного виявлення, протидії та запобігання їм вимагають застосування нових, більш ефективних моделей та методів. Основні недоліки існуючих моделей і методів включають недостатню чутливість до нових загроз, велику кількість помилок виявлення, низьку відповідь на нові загрози, можливість обходу захисних заходів та низьку масштабованість, що є ключовими викликами для подальшого вдосконалення та розвитку кібербезпеки. У статті проведено аналіз існуючих нечітких моделей виявлення кіберінцидентів, виокремлено їхні недоліки та наголошено на необхідності їхнього подальшого удосконалення та розвитку. Запропоновано удосконалену нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення та удосконалену нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення зі зваженими правилами, на основі розширення ознак кіберінцидентів шляхом отримання їх з різних рівнів кіберзахисту баз даних. До основних рівнів кіберзахисту баз даних потрібно віднести: рівень операційної системи, рівень мережі та рівень системи керування базами даних. Для усунення недоліків, пов'язаних з помилковим спрацьовуванням правил виявлення кіберінцидентів та складністю їх налаштування в умовах ландшафту кібератак, що динамічно змінюється, а також розмірністю бази знань системи управління інформацією та подіями безпеки, запропоновано нечітку модель виявлення кіберінцидентів у базах даних інформаційно-комунікаційних систем військового призначення із вагами антецедентів правил. Показано доцільність застосування розробленої моделі.

Ключові слова: *база даних, інформаційно-комунікаційна система, кіберзахист, кіберінцидент, SIEM-система, теорія нечітких множин, нечіткі правила.*

I. Subach, O. Vlasenko Fuzzy models for cyber incident detection in military information and communication systems databases

Protecting databases of military information and communication systems is an extremely important task in the modern cybersecurity sphere. Growing threats from cyberattacks, the need to effectively detect, counteract and prevent them require the use of new, more effective models and methods. The main disadvantages of existing models and methods include insufficient sensitivity to new threats, a large number of detection errors, low response to new threats, the possibility of bypassing protective measures, and low scalability, which are key challenges for further improvement and development of cybersecurity. The article analyzes the existing fuzzy models for detecting cyber incidents, identifies their shortcomings and emphasizes the need for their further improvement and development. An improved fuzzy model for detecting cyber incidents in databases of military information and communication systems and an improved fuzzy model for detecting cyber incidents in databases of military information and communication systems with weighted rules based on the expansion of cyber incident signs by obtaining them from different levels of cyber security of the data are proposed. The main levels of database cybersecurity include: the operating system level, the network level, and the database management system level. To eliminate the shortcomings associated with the false triggering of cyber incident detection rules and the complexity of their configuration in a dynamically changing cyberattack landscape, as well as the dimensionality of the knowledge base of the information and security event management system, a fuzzy model for detecting cyber incidents in databases of military information and communication systems with weights of rule antecedents is proposed. The expediency of applying the developed model is shown.

Keywords: *database, information and communication system, cyber protection, cyber incident, SIEM system, fuzzy set theory, fuzzy rules.*

Постановка завдання. З кожним днем цифрові технології підтверджують свою критичну роль у військовій справі, а отже питання кібербезпеки стає все більш актуальнішим. Через стрімкий розвиток інформаційних технологій та діджиталізацію процесів військової діяльності основою для координації військових операцій, прийняття стратегічних рішень та забезпечення національної безпеки стають інформаційно-комунікаційні системи військового

призначення (далі – ІКСВП) [1]. Зі зростанням значення ІКСВП у сучасній військовій сфері збільшується і загроза кібератак на них, що може знизити ефективність та надійність проведення військових операцій у цілому. Внаслідок цього особливої ваги набуває важливість розвитку методів та засобів кіберзахисту ІКСВП.

Сучасні кібератаки, використовуючи передові технології й методи, стають все більш складними та вишуканими, що надзвичайно ускладнює їх виявлення та запобігання їм. Зловмисники постійно вдосконалюють свої навички, використовуючи нові методи інтелектуального зламу, соціальної інженерії та експлуатацію нових вразливостей програмного забезпечення.

Тому, для ефективного захисту від сучасних кібератак необхідно постійно розвивати та вдосконалювати методи та засоби кіберзахисту, враховуючи нові виклики, що виникають в інформаційній інфраструктурі органів військового управління.

Аналіз наукових публікацій. Швидкий та постійний розвиток інформаційних технологій супроводжується ростом кількості кібератак, які ставлять під загрозу конфіденційність, цілісність та доступність інформації. У цьому контексті системи управління інформацією та подіями безпеки (далі – SIEM-системи) є одними з ключових інструментів в побудові екосистеми кіберзахисту ІКСВП [2]. SIEM-система поєднує в собі функціонал аналізу, виявлення та реагування на події безпеки, забезпечуючи комплексний підхід до захисту інформації та інфраструктури від кібератак. Ефективність SIEM-систем роблять їх важливим інструментом для будь-якого підрозділу, який прагне забезпечити безпеку ІКСВП.

У роботі [3] досліджено та аргументовано важливість використання інтелектуальних SIEM-систем для створення ефективної системи кіберзахисту інформаційно-комунікаційних систем (далі – ІКС). Архітектуру інтелектуальної SIEM-системи для виявлення кіберінцидентів у базах даних (далі – БД) наведено у публікації [4]. Ключовим компонентом SIEM-системи, який відповідає за обробку та аналіз подій безпеки, є підсистема аналізу даних. Вона включає в себе функції виявлення аномалій, ідентифікації потенційно небезпечних подій, а також кореляції даних із різних джерел. У цій підсистемі застосовуються унікальні високоспеціалізовані алгоритми та складні методи аналізу, які переважно недоступні та вважаються комерційною таємницею. Основними функціями, які покладаються на таку підсистему, є наступні:

Виявлення кібератак/кіберінцидентів. Адміністратори можуть налаштовувати правила, що визначають певні ситуації або шаблони поведінки користувачів, які можуть вказувати на потенційні загрози функціонуванню системи керування базами даних (далі – СКБД). Наприклад, використання несправжніх облікових записів, спроби неуспішного входу для роботи з БД системи, зміни конфігурації СКБД тощо, що допомагає виявляти кібератаки та кіберінциденти, пов'язані з роботою БД, а також надає інформацію про їхні наслідки.

Машинне навчання та аналіз відхилень. Деякі сучасні SIEM-системи використовують методи машинного навчання для виявлення аномальних патернів, які можуть вказувати на зловмисну діяльність.

Кореляція подій. Здійснюється аналіз і кореляція події з різних джерел для виявлення складних атак на БД або зловмисних дій, які можуть бути приховані в окремих записах.

Профілювання користувачів і програмних застосунків. Здійснення аналізу поведінки користувачів і програмних застосунків у мережі для виявлення їхніх незвичних або підозрілих дій по відношенню до БД та СКБД.

Шляхом аналізу інформації про події безпеки та використання методів інженерії знань і штучного інтелекту SIEM-система може прогнозувати майбутні кіберінциденти та допомагати у прийнятті оперативних, обґрунтованих рішень щодо запобігання їм [5]. Це досягається завдяки розробці та впровадженню нових або удосконалених моделей і методів виявлення та

ідентифікації кіберінцидентів, що відбуваються у СКБД та пов'язаних з несанкціонованим доступом до БД ІКСВП.

Існуючи у теперішній час найефективніші моделі та методи виявлення кіберінцидентів стають застарілими у швидкозмінному кіберпросторі. Стрімке зростання кількості та складності кіберзагроз вимагає постійної адаптації та покращення захисних стратегій [6].

Незважаючи на велику кількість наукових досліджень, присвячених використанню методів штучного інтелекту, машинного навчання та інших новітніх інформаційних технологій у SIEM-системах для виявлення кіберінцидентів [7–11], питання організації знань у базах знань інтелектуальних SIEM-систем та їх використання для ідентифікації кіберінцидентів, пов'язаних з БД ІКСВП, зокрема в умовах неповноти та невизначеності інформації, залишаються не повністю вирішеними.

Метою статті є побудова моделі виявлення кіберінцидентів, пов'язаних з базами даних інформаційно-комунікаційних систем військового призначення на основі застосування моделей і методів теорії нечітких множин та лінгвістичних термів.

Виклад основного матеріалу дослідження. Аналіз моделей і методів нечіткої ідентифікації кіберінцидентів [12; 13], показує, що рішення задачі щодо їхнього розпізнавання полягає у знаходженні відображення (1):

$$F^* = (f_1^*, f_2^*, \dots, f_n^*) \to c_j \in C = (c_1, c_2, \dots, c_m), \quad (1)$$

де F^* – множина ознак кіберінциденту, пов'язаного з функціонуванням БД ІКСВП;

C – множина можливих кіберінцидентів.

Проте в [14] показано, що ефективність рішення задачі виявлення кіберінцидентів у БД ІКСВП підвищується шляхом застосування багаторівневого кіберзахисту БД від кібератак (рис. 1).

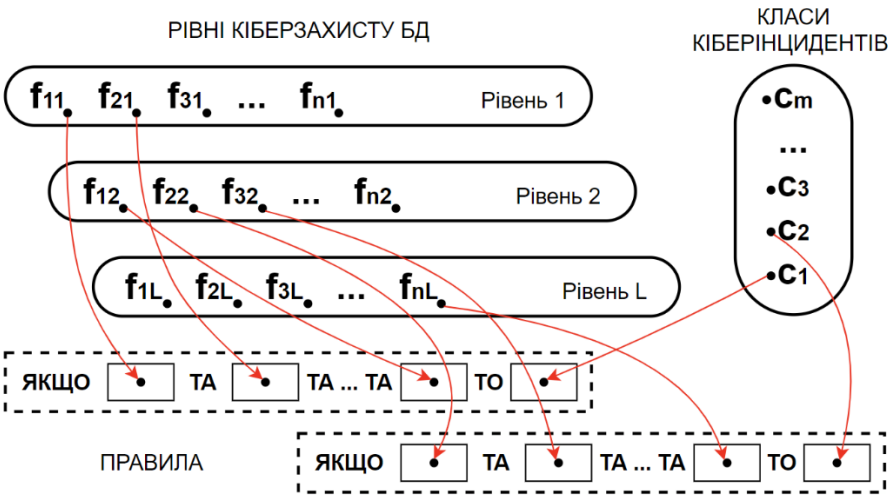


Рис. 1. Розширення ознак кіберінцидентів на основі багаторівневого кіберзахисту БД

Тоді задача розпізнавання кіберінцидентів у БД ІКСВП прийме вигляд (2):

$$F_l^* = (f_{1l}^*, f_{2l}^*, \dots, f_{nl}^*) \to c_j \in C = (c_1, c_2, \dots, c_m), \quad (2)$$

де F_l^* – множина ознак кіберінциденту, пов'язаного з функціонуванням БД ІКСВП, отриманих з різних рівнів кіберзахисту БД [14].

Область зміни ознак кіберінцидентів $f_{il} \in \left[\underline{f_{il}}, \overline{f_{il}}\right],\ i = 1 \dots n,\ l = 1 \dots L$, отриманих на різних рівнях кіберзахисту БД $l = 1 \dots L$ і вихідного значення ідентифікації $c_i \in \left[\overline{c}_j, \underline{c_j}\right]$, $j = 1 \dots m$.

Відповідно $\underline{f_{il}}, (\overline{f_{il}})$ – нижнє (верхнє) значення ознак кіберінциденту f_{il}, $\underline{c_j}$ $(\overline{c_j})$ – нижнє (верхнє) значення результату ідентифікації c_j.

Для вирішення задачі, вхідні і вихідні змінні розглядаються як лінгвістичні змінні, що задані на універсальних множинах [15; 16]:

$$f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right],\ c_j = \left[\underline{c_j},\ \overline{c_j}\right]. \tag{3}$$

Для оцінки (3) доцільно використовувати якісні терми, які входять до множин термів:

$A_{il} = \{a_{il}^1, a_{il}^2, \dots, a_{il}^k\}$ – терм-множина змінної f_{il}, де a_{il}^k – k-й лінгвістичний терм змінної f_{il}, $k = 1 \dots k_i,\ i = 1 \dots n$ l-го рівня кіберзахисту БД $l = 1 \dots L$;

$\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$ – терм-множина змінної c_j, де δ_j, $j = 1 \dots m$ – лінгвістичний терм змінної c_j, m – число можливих класів кіберінцидентів, пов'язаних із БД.

Таким чином, лінгвістичні терми $a_{il}^k \in A_{il}, k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L$ та $\delta_j \in \Delta$, $j = 1 \dots m$ можна розглядати як нечіткі множини, які задані на універсальних множинах f_{il}, c_j.

У свою чергу, нечіткі множини a_{il}^k та δ_j можна визначити так [17–19]:

$$\boldsymbol{a_{il}^k} = \int_{\underline{f_{il}}}^{\overline{f_{il}}} \boldsymbol{\mu^{\alpha_{il}^k}(f_{il})/f_{il}}\ , \tag{4}$$

$$\delta_j = \int_{\underline{c_j}}^{\overline{c_j}} \mu^{\delta_j}(c_j)/c_j\ , \tag{5}$$

де $\mu^{\alpha_{il}^k}(f_{il})$ – функція належності значення змінної $f_{il} \in \left[\underline{f_{il}}, \overline{f_{il}}\right],\ i = 1 \dots n,\ l = 1 \dots L$ терму $\alpha_{il}^k \in A_{il},\ k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L$;

$\mu^{\delta_j}(c_j)$ – функція належності значення змінної $c_j = \left[\underline{c_j}, \overline{c_j}\right]$ терму – класу кіберінциденту, пов'язаному з БД $\delta_j \in \Delta, = 1 \dots m$.

Зауважимо, що у виразах (4) та (5) знак інтегралу означає об'єднання пар $\mu(\omega)/\omega$ [20; 21]. Експертні дані можуть бути представлені у вигляді багатовимірної матриці знань про кіберінциденти (табл. 1). Вона має наступні властивості [15; 16]:

кожний рядок матриці є комбінацією вхідних значень ознак кіберінцидентів, що пов'язані з функціонуванням БД ІКСВП $f_{il},\ i = 1 \dots n, l = 1 \dots L$ та які належать різним рівням безпеки БД, яка віднесена експертом до одного з його класів δ_j, причому перші 11 рядків відповідають класу δ_1, а останні $1m$ рядків – класу δ_m;

перші n стовпчиків матриці відповідають вхідним значенням ознак кіберінцидентів з різних рівнів кіберзахисту БД $f_{il},\ i = 1 \dots n,\ l = 1 \dots L$, а $(n+1)$-ий стовпчик відповідає вихідному значенню – класу кіберінциденту c, пов'язаному з БД;

на перетині i-го стовпчика та jk_j-го рядку знаходиться елемент $\alpha_{il}^{jk_j}$, який відповідає лінгвістичній оцінці ознаки кіберінциденту f_{il} у рядку матриці jk_j, яка належить терм-множині відповідної ознаки f_{il}: $\alpha_{il}^{k} \in A_i,\ k = 1 \dots k_i,\ i = 1 \dots n,\ l = 1 \dots L$.

Таблиця 1

Багатовимірна таблиця ознак кіберінцидентів і класів, що їм відповідають

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Клас кіберінциденту
	j_1	j_2	…	j_i	…	j_n	c
11	α_{1l}^{11}	α_{2l}^{11}	…	α_{il}^{11}	…	α_{nl}^{11}	δ_1
12	α_{1l}^{12}	α_{2l}^{12}	…	α_{il}^{12}	…	α_{nl}^{12}	
…	…	…	…	…	…	…	
$1k_1$	$\alpha_{1l}^{1k_1}$	$\alpha_{1l}^{1k_2}$	…	$\alpha_{il}^{1k_1}$	…	$\alpha_{nl}^{1k_1}$	
…	…	…	…	…	…	…	…
j1	α_{1l}^{j1}	α_{2l}^{j1}	…	α_{il}^{j1}	…	α_{nl}^{j1}	δ_j
j2	α_{1l}^{j2}	α_{2l}^{j2}	…	α_{il}^{j2}	…	α_{nl}^{j2}	
…	…	…	…	…	…	…	
jk_j	$\alpha_{1l}^{jk_j}$	$\alpha_{2l}^{jk_j}$	…	$\alpha_{il}^{jk_j}$	…	$\alpha_{nl}^{jk_j}$	
…	…	…	…	…	…	…	…
m_1	α_{1l}^{m1}	α_{2l}^{m1}	…	α_{il}^{m1}	…	α_{nl}^{m1}	δ_m
m_2	α_{1l}^{m2}	α_{2l}^{m2}	…	α_{il}^{m2}	…	α_{nl}^{m2}	
…	…	…	…		…		
mk_m	$\alpha_{1l}^{mk_m}$	$\alpha_{2l}^{mk_m}$	…	$\alpha_{il}^{mk_m}$	…	$\alpha_{nl}^{mk_m}$	

Описана вище матриця знань про кіберінциденти, пов'язані з БД ІКСВП, може бути представлена у вигляді системи нечітких правил виду «ЯКЩО – ТО» [15; 16], які зв'язують значення вхідних ознак кіберінцидентів $f_{il},\ i = 1 \dots n,\ \ l = 1 \dots L$ з одним з їхніх можливих класів $\delta_j \in \Delta,\ j = 1 \dots m$:

ЯКЩО ($f_{il} = \alpha_{1l}^{11}$) ТА ($f_{2l} = \alpha_{2l}^{11}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{11}$) АБО
($f_{il} = \alpha_{1l}^{12}$) ТА ($f_{2l} = \alpha_{2l}^{12}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{12}$) АБО
($f_{il} = \alpha_{1l}^{1k_1}$) ТА ($f_{2l} = \alpha_{2l}^{1k_1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{1k_1}$), ТО (с=$\delta_1$),…
…, ЯКЩО ($f_{il} = \alpha_{1l}^{j1}$) ТА ($f_{2l} = \alpha_{2l}^{j1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j1}$) АБО (6)
($f_{il} = \alpha_{1l}^{j2}$) ТА ($f_{2l} = \alpha_{2l}^{j2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j2}$) АБО
($f_{il} = \alpha_{1l}^{k_j}$) ТА ($f_{2l} = \alpha_{2l}^{jk_j}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{jk_j}$), ТО (с=$\delta_j$),…
…, ЯКЩО ($f_{il} = \alpha_{1l}^{m1}$) ТА ($f_{2l} = \alpha_{2l}^{m1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{m1}$) АБО
($f_{il} = \alpha_{1l}^{m2}$) ТА ($f_{2l} = \alpha_{2l}^{m2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{m2}$) АБО
($f_{il} = \alpha_{1l}^{mk_m}$) ТА ($f_{2l} = \alpha_{2l}^{mk_m}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{mk_m}$), ТО (с=$\delta_m$),

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту $f_{il},\ i = 1 \dots n,\ l = 1 \dots L$ у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta,\ j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту, пов'язаного з БД, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$.

Таким чином, вираз (6), заданий у вигляді сукупності нечітких правил виду «ЯКЩО – ТО», які ґрунтуються на матриці знань про кіберінциденти (табл.1), являє собою

модель нечіткої ідентифікації кіберінцидентів, пов'язаних з БД ІКСВП інтелектуальною SIEM-системою з урахуванням різних рівнів кіберзахисту БД.

Якщо лінгвістичні оцінки α_{il}^{jk} змінних $f_{1l},\ f_{2l}, \dots ,\ f_{nl}$ та $\delta_j,\ j = 1 \dots m$ з (6) розглянути як нечіткі множини, що визначені на універсальних множинах $f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right], c = \left[\underline{c_j}, \overline{c_j}\right]$, $i = 1 \dots n,\ j = 1 \dots m,\ \ l = 1 \dots L,$ то $\mu^{\alpha_{il}^{jk}}$ – функція належності ознаки кіберінциденту $f_{il} = \left[\underline{f_{il}}, \overline{f_{il}}\right]$ нечіткому терму $\alpha_{il}^{jk},\ i = 1 \dots n,\ j = 1 \dots m, k = 1 \dots k_i,\ \ \ l = 1 \dots L,$ а $\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl})$ – функція належності вектора ознак кіберінцидентів, пов'язаних з БД $F = \{f_{1l}, f_{2l}, \dots, f_{nl}\},\ i = 1 \dots n,$ *l=1…L*, значенню вихідної оцінки $c = \delta_j,\ j = 1 \dots m$ [15–17].

Зв'язок між ними визначається через НМЗ про кіберінциденти (табл. 1) та шляхом заміни лінгвістичних термів на їхні функції належності, а також заміни логічних операцій ТА чи АБО на операції ∧ та ∨ може бути представленим у наступному вигляді [15; 16]:

$$\begin{aligned} \mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \left[\mu^{\alpha_{1l}^{j1}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{j1}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{j1}}(f_{nl})\right] \vee \\ \vee \left[\mu^{\alpha_{1l}^{j2}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{j2}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{j2}}(f_{nl})\right] \vee \\ \vee \left[\mu^{\alpha_{1l}^{jk_j}}(f_{1l}) \wedge \mu^{\alpha_{2l}^{jk_j}}(f_{2l}) \wedge \dots \wedge \mu^{\alpha_{nl}^{jk_j}}(f_{nl})\right],\ j = 1 \dots m,\ l=1 \dots L. \end{aligned} \tag{7}$$

Шляхом згортання вираз (6) може бути представлено наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \ \bigvee_{k=1}^{k_j}\left[\bigwedge_{i=1}^{n}\mu^{\alpha_{il}^{jk}}(f_{il})\right],\ j = 1 \dots m, l = 1 \dots L. \tag{8}$$

Така удосконалена модель може бути основою для розробки правило-орієнтованого методу виявлення кіберінцидентів, пов'язаних з БД SIEM-системою із врахуванням ознак кіберінцидентів, отриманих із різних рівнів кіберзахисту БД ІКСВП.

Головним недоліком запропонованої моделі (8) є те, що впевненість експертів у кожному правилі «ЯКЩО – ТО», які входять до нечіткої бази знань (6), може бути різною. Цей недолік може бути усунений шляхом введення ваги правила, яка й буде характеризувати значимість того чи іншого правила під час ідентифікації кіберінцидентів. Тут, ґрунтуючись на роботах [16–20], під вагою правила будемо розуміти число в інтервалі [0, 1], яке характеризує суб'єктивну міру впевненості експерта в тому чи іншому правилі.

Ураховуючи це, багатовимірна таблиця ознак кіберінцидентів, що відбуваються у БД ІКСВП і класів, що їм відповідають (табл. 1), з урахування впевненості експерта у тому чи іншому правилі, яка задається за допомогою ваги правила, прийме наступний вигляд (табл. 2):

Таблиця 2

Багатовимірна таблиця ознак кіберінцидентів у БД ІКСВП і класів, що їм відповідають з урахуванням ваги правил

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Вага правила	Клас кіберінциденту
	f_{1l}	f_{2l}	…	f_{il}	…	f_{nl}	ω	c
11	α_{1l}^{11}	α_{2l}^{11}	…	α_{il}^{11}	…	α_{nl}^{11}	ω_{11}	δ_1
12	α_{1l}^{12}	α_{2l}^{12}	…	α_{il}^{12}	…	α_{nl}^{12}	ω_{12}	
…	…	…	…	…	…	…		
$1k_1$	$\alpha_{1l}^{1k_1}$	$\alpha_{1l}^{1k_2}$	…	$\alpha_{il}^{1k_1}$	…	$\alpha_{nl}^{1k_1}$	ω_{1k_1}	
…	…	…	…	…	…	…		…

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД						Вага правила	Клас кіберінциденту
	f_{1l}	f_{2l}	…	f_{il}	…	f_{nl}	ω	c
$j1$	α_{1l}^{j1}	α_{2l}^{j1}	…	α_{il}^{j1}	…	α_{nl}^{j1}	ω_{j1}	δ_j
$j2$	α_{1l}^{j2}	α_{2l}^{j2}	…	α_{il}^{j2}	…	α_{nl}^{j2}	ω_{j2}	
…	…	…	…	…	…	…	…	
$j\mathrm{k}_j$	$\alpha_{1l}^{jk_j}$	$\alpha_{2l}^{jk_j}$	…	$\alpha_{il}^{jk_j}$	…	$\alpha_{nl}^{jk_j}$	ω_{jk_j}	
…	…	…	…	…	…	…	…	…
m_1	α_{1l}^{m1}	α_{2l}^{m1}	…	α_{il}^{m1}	…	α_{nl}^{m1}	ω_{m1}	δ_m
m_2	α_{1l}^{m2}	α_{2l}^{m2}	…	α_{il}^{m2}	…	α_{nl}^{m2}	ω_{m2}	
…	…	…	…	…	…	…	…	
mk_m	$\alpha_{1l}^{mk_m}$	$\alpha_{2l}^{mk_m}$	…	$\alpha_{il}^{mk_m}$	…	$\alpha_{nl}^{mk_m}$	ω_{mk_m}	

Тоді, ґрунтуючись на підході з [15; 16] із врахуванням ваг правил, нечітка база знань, яка представлена сукупністю зважених нечітких правил «ЯКЩО – ТО», що зв'язують лінгвістичні оцінки ознак кіберінцидентів у БД ІКСВП з результатами їхньої ідентифікації, прийме наступний вигляд:

ЯКЩО ($f_{il} = \alpha_{1l}^{11}$) ТА ($f_{2l} = \alpha_{2l}^{11}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{11}$) з вагою w_{11} АБО
($f_{il} = \alpha_{1l}^{12}$) ТА ($f_{2l} = \alpha_{2l}^{12}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{12}$) з вагою w_{12} АБО
($f_{il} = \alpha_{1l}^{1k_1}$) ТА ($f_{2l} = \alpha_{2l}^{1k_1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{1k_1}$) з вагою w_{1k_1},
ТО (с=δ_1),…
…, ЯКЩО ($f_{il} = \alpha_{1l}^{j1}$) ТА ($f_{2l} = \alpha_{2l}^{j1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j1}$) з вагою w_{j1} АБО
($f_{il} = \alpha_{1l}^{j2}$) ТА ($f_{2l} = \alpha_{2l}^{j2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j2}$) з вагою АБО
($f_{il} = \alpha_{1l}^{jk_j}$) ТА ($f_{2l} = \alpha_{2l}^{jk_j}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{jk_j}$) з вагою w_{jk_j},
ТО (с=δ_j),… (9)
…, ЯКЩО($f_{il} = \alpha_{1l}^{m1}$) ТА ($f_{2l} = \alpha_{2l}^{m1}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{m1}$) з вагою w_{m1} АБО
($f_{il} = \alpha_{1l}^{m2}$) ТА ($f_{2l} = \alpha_{2l}^{m2}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{m2}$) з вагою w_{m2} АБО
($f_{il} = \alpha_{1l}^{mk_m}$) ТА ($f_{2l} = \alpha_{2l}^{mk_m}$) ТА … ТА ($f_{nl} = \alpha_{nl}^{mk_m}$) з вагою w_{mk_m},
ТО (с=δ_m),

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту у БД ІКСВП $f_{il},\ i = 1 \dots n$; l=$1 \dots L$ у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta,\ j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту у БД ІКСВП, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$, w_{jk} – вага правила.

З урахуванням ваг правил, нечітка база знань (3.13) може бути представлена модифікованою системою нечітких рівнянь (3.11) наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \vee_{k=1}^{k_j} \left\{ \omega_{jk} \left[\wedge_{i=1}^{n} \mu^{\alpha_{il}^{jk}}(f_{il}) \right] \right\}, j = 1 \dots m. \quad (10)$$

Шляхом заміни операцій $\wedge$ та $\vee$ на операції min та max [14–19], які їм відповідають, отримаємо модифіковану модель нечіткої ідентифікації кіберінцидентів у БД ІКСВП SIEM-системами (15) зі зваженими правилами:

$$\mu^{\delta_j}(f_{il}, f_{2l}, \dots, f_{nl}) = \max_{k = \overline{1, k_j}} \left\{ \omega_{jk} \min_{i = \overline{1, n}} \left\{ \mu^{\alpha_{il}^{jk}}(f_{il}) \right\}, j = 1 \dots m \right\}. \quad (11)$$

Така удосконалена модель дозволяє усунути недоліки, що були присутні в моделі (8), шляхом урахування впевненості експерта у тому чи іншому правилі.

Проте слід зауважити, що вплив окремих ознак кіберінцидентів (антецедентів) правила на результат ідентифікації, з точки зору впевненості експерта, ця модель не враховує. Крім того, сучасні бази знань SIEM-систем можуть містити десятки тисяч правил, тому на практиці експерту важко визначити вагу того чи іншого правила. Виходячи з цього, цілком доцільно удосконалити модель (11) шляхом надання ознакам кіберінцидентів вагових коефіцієнтів в межах правил, що визначаються експертами (табл. 3).

Таблиця 3

Багатовимірна таблиця ознак кіберінцидентів у БД ІКСВП і класів, що їм відповідають, з урахуванням ваги ознак кіберінцидентів у межах окремих правил

Номер вхідної комбінації значень	Ознаки кіберінцидентів, отримані з різних рівнів кіберзахисту БД										Клас кібер-інциденту
	f_{1l}	Ω_{1l}	f_{2l}	Ω_{2l}	…	f_{il}	Ω_{il}	…	f_{nl}	Ω_{nl}	c
11	α_{1l}^{11}	$\Omega_{1l}{}^{11}$	α_{2l}^{11}	$\Omega_{2l}{}^{11}$	…	α_{il}^{11}	$\Omega_{il}{}^{11}$	…	α_{nl}^{11}	$\Omega_{nl}{}^{11}$	δ_1
12	α_{1l}^{12}	$\Omega_{1l}{}^{12}$	α_{2l}^{12}	$\Omega_{2l}{}^{12}$	…	α_{il}^{12}	$\Omega_{il}{}^{12}$	…	α_{nl}^{12}	$\Omega_{nl}{}^{12}$	
…	…		…		…	…		…	…		
1k₁	$\alpha_{1l}^{1k_1}$	$\Omega_{1l}{}^{1k}{}_1$	$\alpha_{1l}^{1k_2}$	$\Omega_{2l}{}^{1k}{}_2$	…	$\alpha_{il}^{1k_1}$	$\Omega_{il}{}^{1k}{}_1$	…	$\alpha_{nl}^{1k_1}$	$\Omega_{nl}{}^{1k}{}_1$	
…	…		…		…	…		…	…		…
j1	α_{1l}^{j1}	$\Omega_{1l}{}^{j1}$	α_{2l}^{j1}	$\Omega_{2l}{}^{j1}$	…	α_{il}^{j1}	$\Omega_{il}{}^{j1}$	…	α_{nl}^{j1}	$\Omega_{nl}{}^{j1}$	δ_j
j2	α_{1l}^{j2}	$\Omega_{1l}{}^{j2}$	α_{2l}^{j2}	$\Omega_{2l}{}^{j2}$	…	α_{il}^{j2}	$\Omega_{il}{}^{j2}$	…	α_{nl}^{j2}	$\Omega_{nl}{}^{j2}$	
…	…		…		…	…		…	…		
jk$_j$	$\alpha_{1l}^{jk_j}$	$\Omega_{1l}{}^{jk}{}_j$	$\alpha_{2l}^{jk_j}$	$\Omega_{2l}{}^{jk}{}_j$	…	$\alpha_{il}^{jk_j}$	$\Omega_{il}{}^{jk}{}_j$	…	$\alpha_{nl}^{jk_j}$	$\Omega_{nl}{}^{jk}{}_j$	
…	…		…		…	…		…	…		…
m$_1$	α_{1l}^{m1}	$\Omega_{1l}{}^{m1}$	α_{2l}^{m1}	$\Omega_{2l}{}^{m1}$	…	α_{il}^{m1}	$\Omega_{il}{}^{m1}$	…	α_{nl}^{m1}	$\Omega_{nl}{}^{m1}$	δ_m
m$_2$	α_{1l}^{m2}	$\Omega_{1l}{}^{m2}$	α_{2l}^{m2}	$\Omega_{2l}{}^{m2}$	…	α_{il}^{m2}	$\Omega_{il}{}^{m2}$	…	α_{nl}^{m2}	$\Omega_{nl}{}^{m2}$	
…	…		…		…		…	…			
mk$_m$	$\alpha_{1l}^{mk_m}$	$\Omega_{1l}{}^{mk}{}_m$	$\alpha_{2l}^{mk_m}$	$\Omega_{2l}{}^{mk}{}_m$	…	$\alpha_{il}^{mk_m}$	$\Omega_{il}{}^{mk}{}_m$	…	$\alpha_{nl}^{mk_m}$	$\Omega_{nl}{}^{mk}{}_m$	

Тоді, з врахуванням ваг ознак кіберінцидентів у БД ІКСВП, НБЗ, яка представлена сукупністю нечітких правил «ЯКЩО – ТО», що зв'язують лінгвістичні оцінки ознак кіберінцидентів з результатами їхньої ідентифікації, прийме наступний вигляд:

ЯКЩО ($f_{il} = \alpha_{1l}^{11}$з вагою $\Omega_{1l}{}^{11}$) ТА ($f_{2l} = \alpha_{2l}^{11}$ з вагою $\Omega_{2l}{}^{11}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{11}$з вагою $\Omega_{nl}{}^{11}$) АБО

($f_{il} = \alpha_{1l}^{12}$з вагою $\Omega_{1l}{}^{12}$) ТА ($f_{2l} = \alpha_{2l}^{12}$з вагою $\Omega_{2l}{}^{12}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{12}$з вагою $\Omega_{nl}{}^{12}$) АБО

($f_{il} = \alpha_{1l}^{1k_1}$з вагою $\Omega_{1l}{}^{1k}{}_1$) ТА ($f_{2l} = \alpha_{2l}^{1k_1}$з вагою $\Omega_{2l}{}^{1k}{}_1$) ТА…ТА ($f_{nl} = \alpha_{nl}^{1k_1}$з вагою $\Omega_{nl}{}^{1k}{}_1$), ТО (c=δ_1),…

…, ЯКЩО ($f_{il} = \alpha_{1l}^{j1}$з вагою $\Omega_{1l}{}^{j1}$) ТА ($f_{2l} = \alpha_{2l}^{j1}$з вагою $\Omega_{2l}{}^{j1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j1}$з вагою $\Omega_{nl}{}^{j1}$) АБО

($f_{il} = \alpha_{1l}^{j2}$з вагою $\Omega_{1l}{}^{j2}$) ТА ($f_{2l} = \alpha_{2l}^{j2}$з вагою $\Omega_{2l}{}^{j2}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{j2}$з вагою $\Omega_{nl}{}^{j2}$) АБО

($f_{il} = \alpha_{1l}^{jk_j}$з вагою $\Omega_{1l}{}^{jk}{}_j$) ТА ($f_{2l} = \alpha_{2l}^{jk_j}$з вагою $\Omega_{2l}{}^{jk}{}_j$) ТА…ТА ($f_{nl} = \alpha_{nl}^{jk_j}$з вагою $\Omega_{nl}{}^{jk}{}_j$), ТО (c=δ_j),…

…, ЯКЩО ($f_{il} = \alpha_{1l}^{m1}$ з вагою $\Omega_{1l}{}^{m1}$) ТА ($f_{2l} = \alpha_{2l}^{m1}$з вагою $\Omega_{2l}{}^{m1}$) ТА…ТА ($f_{nl} = \alpha_{nl}^{m1}$з вагою $\Omega_{nl}{}^{m1}$) АБО

($f_{il} = \alpha_{1l}^{m2}$з вагою Ω_{il}^{m2})ТА($f_{2l} = \alpha_{2l}^{m2}$з вагою Ω_{2l}^{m2})ТА…ТА($f_{nl} = \alpha_{nl}^{m2}$з вагою Ω_{nl}^{m2})

АБО

($f_{il} = \alpha_{1l}^{mk_m}$з вагою $\Omega_{il}^{mk_m}$)ТА($f_{2l} = \alpha_{2l}^{mk_m}$з вагою $\Omega_{2l}^{mk_m}$)ТА…ТА($f_{nl} = \alpha_{nl}^{mk_m}$з вагою $\Omega_{nl}^{mk_m}$),

ТО (c=δ_m), (12)

де α_{il}^{jk} – лінгвістична оцінка ознаки кіберінциденту у БД ІКСВП $f_{il},\ i = 1 \dots n$; l=$1 \dots L$ у рядку k j-ої диз'юнкції, що визначається на терм-множині $A_i = \{\alpha_{il}^1, \alpha_{il}^2, \dots, \alpha_{il}^{k_i}\}$;

$\delta_j \in \Delta,\ j = 1 \dots m$ – лінгвістична оцінка класу кіберінциденту у БД ІКСВП, що визначається на терм-множині $\Delta = \{\delta_1, \delta_2, \dots, \delta_m\}$;

Ω_{nl}^{mk} – вага ознак кіберінцидентів у межах правил.

З урахуванням ваг ознак кіберінцидентів у межах правил, нечітка база знань (11) може бути представлена модифікованою системою нечітких рівнянь (12) наступним чином:

$$\mu^{\delta_j}(f_{1l}, f_{2l}, \dots, f_{nl}) = \bigvee_{k=1}^{k_j} \left\{ \bigwedge_{i=1}^{n} \left[\mu^{\alpha_{il}^{jk}}(f_{il})\ \Omega_{il}^{jk} \right] \right\}, j = 1 \dots m. \quad (13)$$

Шляхом заміни операцій ∧ та ∨на операції min та max, які їм відповідають, отримаємо модифіковану модель нечіткої ідентифікації кіберінцидентів SIEM-системами (13) зі зваженими ознаками правил:

$$\mu^{\delta_j}(f_{il}, f_{2l}, \dots, f_{nl}) = \underset{k = \overline{1, k_j}}{max} \left\{ \underset{i = \overline{1, n}}{min} \left\{ \mu^{\alpha_{il}^{jk}}(f_{il})\ \Omega_{il}^{jk} \right\},\ j = \overline{1, m} \right\}. \quad (14)$$

Така модель дозволяє певною мірою усунути недоліки, що були присутні в моделі (11), шляхом урахування впевненості експерта у тій чи іншій ознаці кіберінциденту у правилі.

Як приклад, розглянемо кіберінцидент, який має бути розглянутий офіцером безпеки, щодо кількості невдалих спроб підключення до БД з однієї або декількох IP-адрес.

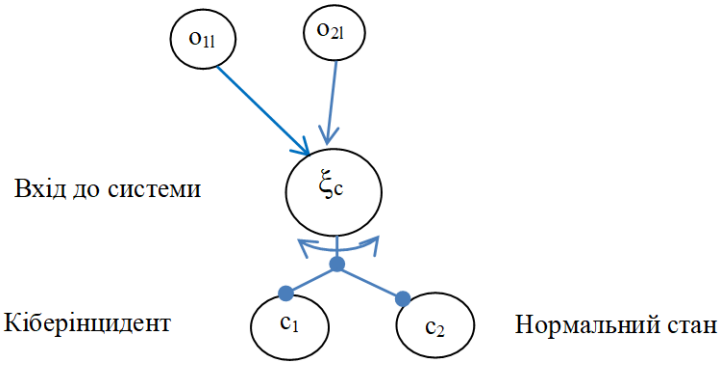


Рис. 2. Дерево логічного виводу для кіберінциденту

На рисунку 2 наведено дерево логічного виводу для цього кіберінциденту.

Для ідентифікації зазначеного кіберінциденту доцільно дослідити його ознаки у вигляді вхідних лінгвістичних змінних:

numb_attempts (o_{1l}) – кількість невдалих спроб вводу пароля для входу до БД, який відноситься безпосередньо до рівня кіберзахисту БД [15];

count_ip (o_{2l}) – кількість IP-адрес, з яких вводився пароль, що відповідає рівню кіберзахисту мережі.

Для опису подій, які відбуваються в системі з БД, використаємо наступні їхні типи та позначимо через c_1 та c_2.

Таблиця 4

Тип події, що відбувається в системі

Подія	Зміст події
C_1	Кіберінцидент
C_2	Нормальний стан системи

Структура дерева логічного виводу відповідає відношенню (14):

$$c = \xi_c(o_{1l}, o_{2l}). \quad (14)$$

Для оцінки значень лінгвістичних змінних o_{1l} та o_{2l} застосовується єдина шкала якісних термів: Н – низька, В – висока. Кожний з цих термів задається відповідною функцією належності.

Таблиця 5

Нечіткий опис станів

Ознаки кіберінциденту		Тип кіберінциденту
numb_attempts (o_{1l})	*count_ip* (o_{2l})	
Н	Н	C_2
Н	В	
В	Н	C_1
В	В	

Дослідження ознак проведемо на основі функцій належності з наступними значеннями:
numb_attempts (o_{1l}) – {«Н-низька [1,2,3]», «В-висока [3,4,5]»} на універсумі [1,5];
count_ip (o_{2l}) – {«Н-низька [1,2]», »В-висока [2,3,4,5]»} на універсумі [1,5].

Значення функції належності для деякого фіксованого вектора вхідних даних, що відповідають лінгвістичній змінній – *numb_attempts* (o_{1l}), наведено на рисунку 3.

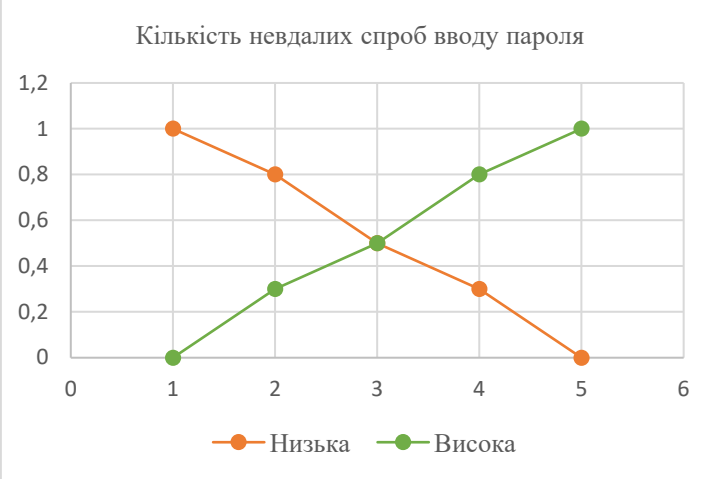


Рис. 3. Значення функції належності лінгвістичної змінної «Кількість невдалих спроб вводу пароля»

Значення функції належності для деякого фіксованого вектора вхідних даних, що відповідають лінгвістичній змінній – *count_ip* (o_{2l}), наведено на рисунку 4.

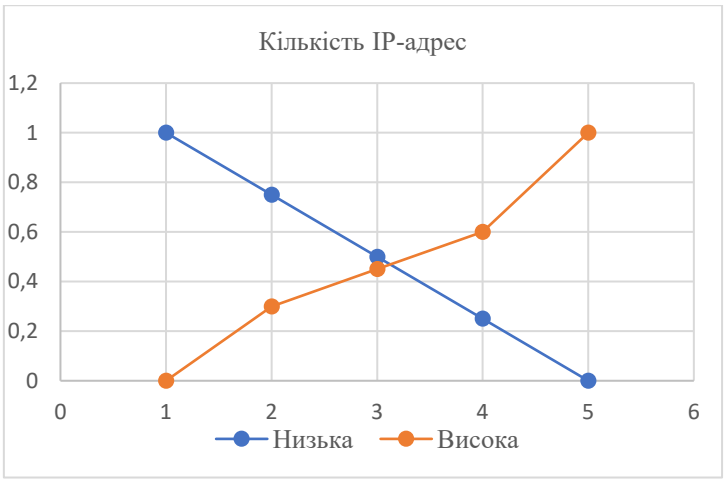


Рис. 4. Значення функції належності лінгвістичної змінної «Кількість IP-адрес, з яких вводився пароль»

З метою порівняння раніше розглянутих моделей (8), (11) та (14), розглянемо деякі варіанти спрацювання нечітких правил. Функції належності, що відповідають цим ознакам, наведено в таблиці 6 та в таблиці 7.

Таблиця 6

Результати обчислень функцій належності вхідного вектора

Лінгвістична змінна	o_{il}	$\mu^H(o_{il})$	$\mu^B(o_{il})$
numb_attempts (o_{1l})	1	1	0
	2	0,8	0,3
	3	0,5	0,5
	4	0,3	0,8
	5	0	1

Таблиця 7

Результати обчислень функцій належності вхідного вектора

Лінгвістична змінна	o_{il}	$\mu^H(o_{il})$	$\mu^B(o_{il})$
count_ip (o_{2l})	1	1	0
	2	0,75	0,3
	3	0,5	0,45
	4	0,25	0,6
	5	0	1

Для визначення максимальної відповідності досліджуваних ознак до шаблону подій, які відбуваються в системі, обчислюємо функції належності на основі (8), а логічні операції кон'юнкції та диз'юнкції заміняємо на нечіткі кон'юнкцію та диз'юнкцію на основі максимінного підходу:

$$\mu(a) \wedge \mu(b) = \min[\mu(a), \mu(b)],$$
$$\mu(a) \vee \mu(b) = \max[\mu(a), \mu(b)].$$

Результати розрахунків для моделі (8) будуть наступними:
Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 \wedge 0{,}75) \vee (0{,}8 \wedge 0{,}3) = 0{,}75,$$

$$\mu^{c_1} = (0{,}3 \wedge 0{,}75) \vee (0{,}3 \wedge 0{,}3) = 0{,}3.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}75) \vee (0{,}5 \wedge 0{,}3) = 0{,}5,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}75) \vee (0{,}5 \wedge 0{,}3) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}75) \vee (0{,}3 \wedge 0{,}3) = 0{,}3,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}75) \vee (0{,}8 \wedge 0{,}3) = 0{,}75.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}75) \vee (0 \wedge 0{,}3) = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}75) \vee (1 \wedge 0{,}3) = 0{,}75.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}5) \vee (0{,}5 \wedge 0{,}45) = 0{,}5,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}5) \vee (0{,}5 \wedge 0{,}45) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}5) \vee (0{,}3 \wedge 0{,}45) = 0{,}3,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}5) \vee (0{,}8 \wedge 0{,}45) = 0{,}5.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}5) \vee (0 \wedge 0{,}45) = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}5) \vee (1 \wedge 0{,}45) = 0{,}5.$$

Результати розрахунків, опираючись на модель (12), допустимо випадок, коли значення ваги правил співпадає і має – 0,8:

Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 \wedge 0{,}75) * 0,8 \vee (0{,}8 \wedge 0{,}3) * 0,8 = 0{,}75,$$
$$\mu^{c_1} = (0{,}3 \wedge 0{,}75) * 0,8 \vee (0{,}3 \wedge 0{,}3) * 0,8 = 0{,}3.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}75) * 0,8 \vee (0{,}5 \wedge 0{,}3) * 0,8 = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}75) * 0,8 \vee (0{,}5 \wedge 0{,}3) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}75) * 0,8 \vee (0{,}3 \wedge 0{,}3) * 0,8 = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}75) * 0,8 \vee (0{,}8 \wedge 0{,}3) * 0,8 = 0{,}6.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}75) * 0,8 \vee (0 \wedge 0{,}3) * 0,8 = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}75) * 0,8 \vee (1 \wedge 0{,}3) * 0,8 = 0{,}6.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 \wedge 0{,}5) * 0,8 \vee (0{,}5 \wedge 0{,}45) * 0,8 = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 \wedge 0{,}5) * 0,8 \vee (0{,}5 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 \wedge 0{,}5) * 0,8 \vee (0{,}3 \wedge 0{,}45) * 0,8 = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 \wedge 0{,}5) * 0,8 \vee (0{,}8 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 \wedge 0{,}5) * 0,8 \vee (0 \wedge 0{,}45) * 0,8 = 0,$$
$$\mu^{c_1} = (1 \wedge 0{,}5) * 0,8 \vee (1 \wedge 0{,}45) * 0,8 = 0{,}4.$$

Припустимо, що були введені наступні значення ваги антецедентів:

кількість невдалих спроб вводу пароля = низька – 0,8;

кількість невдалих спроб вводу пароля = висока – 1;

кількість IP-адрес, з яких вводився пароль = низька – 0,8;

кількість IP-адрес, з яких вводився пароль = висока – 0,9.

Отже, результати розрахунків на основі моделі (14):

Кількість спроб підключення до БД – 2 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}8 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}8 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}6,$$
$$\mu^{c_1} = (0{,}3 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}3 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}27.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}5 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}5 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}5 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}5.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0{,}3 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0{,}3 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 * 1 \wedge 0{,}75 * 0{,}8) \vee (0{,}8 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}6.$$

Кількість спроб підключення до БД – 5 і відбуваються вони з двох IP-адрес:

$$\mu^{c_2} = (0 * 0{,}8 \wedge 0{,}75 * 0{,}8) \vee (0 * 0{,}8 \wedge 0{,}3 * 0{,}9) = 0,$$
$$\mu^{c_1} = (1 * 1 \wedge 0{,}75 * 0{,}8) \vee (1 * 1 \wedge 0{,}3 * 0{,}9) = 0{,}6.$$

Кількість спроб підключення до БД – 3 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}5 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0{,}5 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0{,}4,$$
$$\mu^{c_1} = (0{,}5 * 1 \wedge 0{,}5 * 0{,}8) \vee (0{,}5 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}405.$$

Кількість спроб підключення до БД – 4 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0{,}3 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0{,}3 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0{,}24,$$
$$\mu^{c_1} = (0{,}8 * 1 \wedge 0{,}5 * 0{,}8) \vee (0{,}8 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}405.$$

Кількість спроб підключення до БД - 5 і відбуваються вони з трьох IP-адрес:

$$\mu^{c_2} = (0 * 0{,}8 \wedge 0{,}5 * 0{,}8) \vee (0 * 0{,}8 \wedge 0{,}45 * 0{,}9) = 0,$$
$$\mu^{c_1} = (1 * 1 \wedge 0{,}5) \vee (1 * 1 \wedge 0{,}45 * 0{,}9) = 0{,}5.$$

Внаслідок проведених розрахунків (табл. 8) та аналізу виявлено, що у випадку застосування моделей (8) та (11) у процесі прийняття рішень виникають ситуації, коли кілька правил активуються одночасно (в поточних розрахунках це відображається, коли кількість невдалих спроб входу відповідає трьом (o_{1l})). Це призводить до невизначеності у прийнятті рішень офіцерами безпеки щодо ідентифікації кіберінциденту, пов'язаного з БД.

Для усунення цієї проблеми запропоновано використовувати ваги антецедентів (логічних умов у правилі), що дозволяє враховувати важливість кожного антецедента при прийнятті рішення. Застосування ваг антецедентів дозволить уникнути ситуацій, коли правила з однаковою вагою конфліктують між собою, та надати особам, що приймають рішення (ОПР), рекомендації для прийняття обґрунтованих рішень у критичних ситуаціях.

Таблиця 8

Порівняння результатів розрахунків

Значення o_{2l}	Значення o_{1l}	Модель (8)		Модель (11)		Модель (14)	
		μ^{c_2}	μ^{c_1}	μ^{c_2}	μ^{c_1}	μ^{c_2}	μ^{c_1}
2	2	0,75 +	0,3	0,5 +	0,3	0,6 +	0,27
	3	0,5 ?	0,5 ?	0,4 ?	0,4 ?	0,4	0,5 +
	4	0,3	0,75 +	0,24	0,6 +	0,24	0,6 +
	5	0	0,75 +	0	0,6 +	0	0,6 +
3	3	0,5 ?	0,5 ?	0,4 ?	0,4 ?	0,4	0,405 +
	4	0,3	0,5 +	0,24	0,4 +	0,24	0,405 +
	5	0	0,5 +	0	0,6+	0	0,5 +

Використання ваг антецедентів у моделі (14) ідентифікації кіберінцидентів SIEM-системою (рис. 5) призвело до покращення порівняно з моделлю (8) та моделлю з вагами правил (12).

Так, з рисунку 5 видно, що під час трьох невдалих спроб введення пароля до БД, у випадку використання моделі (14), усувається невизначеність в результатах ідентифікації кіберінциденту/нормального стану системи при заданих значеннях функцій належності нечітких термів (див. рис. 3). Зокрема, вочевидь буде прийнято рішення щодо віднесення цієї події до нормального стану функціонування інформаційної системи, до складу якої входить БД, що у багатьох випадках відповідає дійсності. Проте це не можна сказати у випадку застосування моделей (8) та (11).

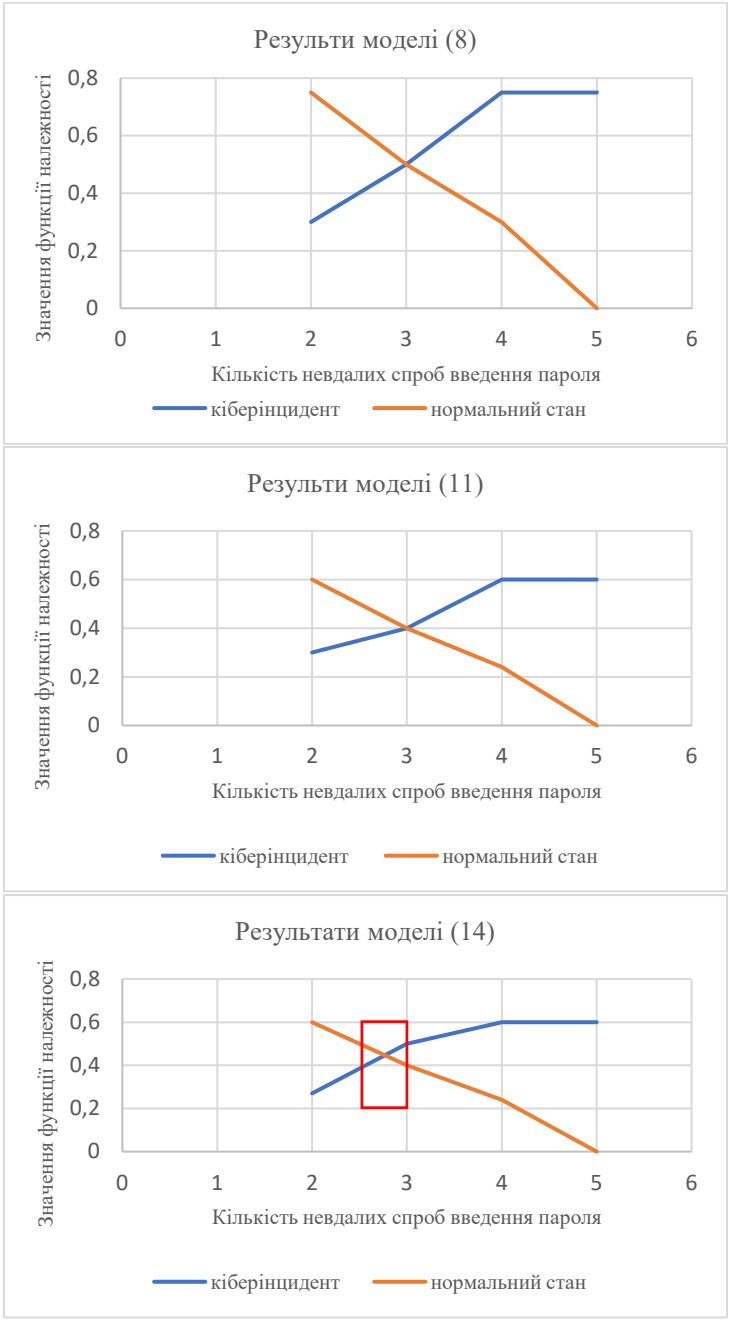


Рис. 5. Результати ідентифікації кіберінцидентів/нормального стану роботи системи у випадках застосування моделей (8), (11) та (14)

Це свідчить про доцільність у деяких випадках врахування ваг антецедентів в усуненні проблеми неоднозначності, що виникає при спрацюванні протилежних правил. Отже, результати підтверджують, що модель з використанням ваг антецедентів, при певних умовах,

може бути більш адекватною порівняно з іншими розглянутими моделями виявлення (ідентифікації) кіберінцидентів у БД ІКСВП. Це може мати важливі наслідки для розробки ефективних систем кіберзахисту та забезпечення кіберзахисту БД в ІКСВП.

Висновки. Удосконалення та впровадження нечітких моделей виявлення кіберінцидентів у БД суттєво покращує ефективність застосування систем управління інформацією та подіями безпеки в контурі захисту ІКСВП в умовах неповноти та невизначеності інформації про них. Аналіз існуючих моделей показав їхні недоліки, пов'язані зі складністю вирішення задачі ідентифікації кіберінцидентів у деяких випадках, зокрема у випадку невдало налаштованих функцій належності нечітких термів, що потребує додатково уточнення їх по експериментальним даним. Проте в умовах ландшафту кібератак, що постійно змінюється, це не завжди є можливим. Крім цього успішне застосування розглянутих моделей є можливим на основі удосконалення їх шляхом розширення простору ознак кіберінцидентів, отриманих з різних рівнів кіберзахисту БД.

Ще одним із недоліків, пов'язаним із застосування нечіткої моделі ідентифікації кіберінцидентів зі зваженими правилами, є великий обсяг бази знань, що призводить до складності рішення цієї задачі, а також значних часових та обчислювальних витрат.

Виходом з цієї ситуації, у деяких випадках, може стати застосування нечіткої моделі ідентифікації кіберінцидентів зі зваженими антецедентами (ознаками кіберінцидентів) правил, яка запропонована в роботі.

Перспективним напрямком подальших досліджень є розробка методу визначення ваги антецедентів нечітких правил.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Герасимов Б. М., Субач І. Ю., Хусаінов П. В., Міщенко В. О. Аналіз задач моніторингу інформаційних мереж та методів підвищення ефективності їх функціонування. *Сучасні інформаційні технології у сфері безпеки та оборони.* 2008. № 3 (3). С. 24–27.

2. Гнатюк С. О. Система корелювання подій та управління інцидентами кібербезпеки на об'єктах критичної інфраструктури. *Кібербезпека: освіта, наука, техніка*. 2023. Т. 3. № 19. С. 176–196. URL: https://doi.org/10.28925/2663-4023.2023.19.176196.

3. Субач І., Кубрак В., Микитюк А. Архітектура та функціональна модель перспективної проактивної інтелектуальної системи SIEM-системи для кіберзахисту об'єктів критичної інфраструктури. *Information Technology and Security.* 2019. № 7 (2). P. 208–215. URL: https://doi.org/10.20535/2411-1031.2019.7.2.190570.

4. Субач І., Власенко О. Архітектура інтелектуальної SIEM-системи для виявлення кіберінцидентів у базах даних інформаційно-телекомунікаційних системах військового призначення. *Збірник наукових праць ВІТІ.* 2023. № 4. С. 82–92. URL: https://doi.org/10.58254/viti.4.2023.07.82.

5. Seyed M. Z. Analysis of Security Information and Event Management (SIEM) – Evasion and Detection Methods. Tallinn University of Technology, Faculty of Information Technology, Tallinn, Estonia, Master Thesis, 2016.

6. Granadillo, Gustavo Gonzalez. Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. *Sensors (Basel, Switzerland)* 21. 2021: n. pag. DOI: 10.3390/s21144759.

7. Suarez-Tangil Guillermo, Palomar Esther, Ribagorda Arturo, Sanz Ivan. *Providing SIEM systems with self-adaptation*. 2015.

8. Anastasov Igor, Davcev Danco. SIEM implementation for global and distributed environments. *Computer Applications and Information Systems* (WCCAIS). 2014 World Congress, 2014.

9. Rafał Leszczyna, Michał R. Wróbel. Evaluation of Open Source SIEM for Situation Awareness Platform in the Smart Grid Environment. *Factory Communication Systems* (WFCS), IEEE World Conference on, 2015.

10. Hanemann A., Marcu P. Algorithm Design and Application of Service Oriented Event Correlation, In Proceedings of Conference BDIM 2008, 3rd IEEE/IFIP International Workshop on Business-Driven IT Management. 2011. P. 61–70.

11. Elshoush H., Osman I. M. Alert correlation in collaborative intelligent intrusion detection systems. A survey. Applied Soft Computing, 2011. P. 4349–4365.

12. Субач І., Фесьоха В. Модель виявлення аномалій в інформаційно-телекомунікаційних мережах органів військового управління на основі нечітких множин та нечіткого логічного виводу. *Збірник наукових праць ВІТІ*. 2017. № 3. С. 158–164. URL: http://nbuv.gov.ua/UJRN/Znpviti_2017_3_21.

13. Subach I., Fesokha V. Model of detecting cybernetic attacks on information-telecommunication systems based on description of anomalies in their work by weighed fuzzy rules. *Collection «Information Technology and Security»*, 2017. № 5 (2). P. 145–152. URL: https://doi.org/10.20535/2411-1031.2017.5.2.136984.

14. Субач І., Власенко О. Інформаційні технології захисту баз даних від кібератак в інформаційних системах військового призначення. *Collection «Information Technology and Security».* 2022. № 10 (2). С. 177–193. URL: https://doi.org/10.20535/2411-1031.2022.10.2.270412.

15. Rotshtein A. P. Medical diagnostics using fuzzy logic. Vinnitsa: Continent-PRIM, 1996. 132 p.

16. Rothstein A. Intelligent identification technologies: fuzzy sets, genetic algorithms, neural networks. Vinnytsia: UNIVERSUM, 1999.

17. RothsteinO., Chernovolyk G., Laryushkin E. Method of constructing membership functions of fuzzy sets. Bulletin of VPI, 1996. Vol. 3. P. 72–75.

18. Mityushkin Y., Mokin B., Rothstein O. Soft Computing: identification of patterns of fuzzy knowledge bases: a monograph. Vinnytsia: UNIVERSUM-Vinnytsia, 2002.

19. Rotshtein A. Design and Tuning of Fuzzy Rule-Based Systems for Medical Diagnosis. In N.-H. Teodorescu (ed): Fuzzy and Neuro - Fuzzy Systems in Medicine. CRC Press. 1998. P. 243–289.

20. Zaichenko Y. P. Operations Research: Fuzzy Optimisation. Vyshcha Shkola, 1991.

21. Borisov A. N., Krumberg O. A., Fedorov I. P. Decision-Making on the Basis of Fuzzy Models: Examples of Use. Zinatne. 1990.

DOI: 10.58254/viti.5.2024.16.181

УДК 004.056.57

д-р філософії Фесьоха В. В. ORCID: 0000-0001-6612-1970 (ВІТІ ім. Героїв Крут)
Кисиленко Д. Ю. ORCID: 0000-0001-5491-6231 (ВІТІ ім. Героїв Крут)
д-р філософії Фесьоха Н. О. ORCID: 0000-0002-9797-5589 (ВІТІ ім. Героїв Крут)

ОБҐРУНТУВАННЯ ВИБОРУ ПІДХОДУ ДО ВИЗНАЧЕННЯ ІНВАРІАНТНОЇ КОМПОНЕНТИ У ПОВЕДІНЦІ ПОЛІМОРФНОГО (МЕТАМОРФНОГО) ШКІДЛИВОГО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ НА ОСНОВІ ЗНИЖЕННЯ РОЗМІРНОСТІ ПРОСТОРУ ОЗНАК

Еволюція сценаріїв застосування шкідливого програмного забезпечення зумовлює потребу в розробці дієвих стратегій нейтралізації їхнього деструктивного впливу. Одним із найбільш загрозливих типів шкідливого програмного забезпечення є поліморфні (метаморфні) віруси, оскільки значною мірою спроможні уникати виявлення системами виявлення вторгнень, управління інформаційною безпекою (подіями безпеки), антивірусними програмами та системами проактивного виявлення нетипових загроз і цільових атак на кінцевих точках завдяки властивості змінювати власну сигнатуру. До того ж, протягом останнього часу зафіксовано стрімке збільшенням кіберінцидентів, пов'язаних із застосуванням поліморфного (метаморфного) шкідливого програмного забезпечення. Основна причина цього зросту – доступність технологій штучного інтелекту, які дозволяють зловмисникам досить швидко та ефективно модифікувати код вже класифікованих шкідливих програм, не потребуючи значних спеціалізованих технічних компетенцій.

Проведено порівняльний аналіз існуючих підходів до виявлення поліморфного, олігоморфного і метаморфного шкідливого програмного забезпечення. Виявлено, що ні одна група методів не використовує на свою користь ключову особливість поліморфного (метаморфного) шкідливого програмного забезпечення – інваріантну поведінку за певною підмножиною ознак, яка характеризує один і той самий вектор деструктивного впливу шкідливого програмного забезпечення.

З метою нівелювання властивості модифікації власного коду поліморфним (метаморфним) шкідливим програмним забезпеченням запропоновано підхід до визначення його інваріантної компоненти під час поведінкового аналізу на основі поєднання переваг поведінкового аналізу та техніки машинного навчання – зменшення розмірності досліджуваного простору ознак. Такий підхід потенційно дозволить визначати інваріантну поведінку шкідливого програмного забезпечення у вигляді підмножини досліджуваних ознак для кожного відомого його типу, що у свою чергу формує підґрунтя для реалізації нового підходу до ефективного виявлення модифікованого (удосконаленого) шкідливого програмного забезпечення.

Ключові слова: *кібербезпека*, *шкідливе програмне забезпечення, поліморфні (метаморфні) віруси, поведінковий аналіз, простір ознак, штучний інтелект, машинне навчання, зменшення розмірності даних, генетичні алгоритми.*

V. Fesokha, D. Kysylenko, N. Fesokha. Justification of the choice of the approach to the determination of the invariant component in the behavior of polymorphic (metamorphic) malware on the basis of reducing the dimensionality of the sign space

The evolution of malware use scenarios necessitates the development of effective strategies to neutralise their destructive impact. One of the most threatening types of malware is polymorphic (metamorphic) viruses, as they are largely able to evade detection by intrusion detection systems, information security management (security events), antivirus software and systems for proactive detection of atypical threats and targeted attacks on endpoints due to their ability to change their own signature. In addition, there has been a rapid increase in recent cyber incidents involving the use of polymorphic (metamorphic) malware. The main reason for this growth is the availability of artificial intelligence technologies that allow attackers to modify the code of already classified malware quickly and efficiently, without requiring significant specialised technical competence.

A comparative analysis of existing approaches to detecting polymorphic, oligomorphic and metamorphic malware is carried out. It is found that no group of methods uses to its advantage the key feature of polymorphic (metamorphic) malware – invariant behaviour by a certain subset of features that characterise the same vector of destructive impact of malware.

With a view to neutralising the property of modification of its own code by polymorphic (metamorphic) malware, the article proposes an approach to determining its invariant component during behavioural analysis based on a combination of the advantages of behavioural analysis and machine learning techniques – reducing the dimensionality of the studied feature space. Such an approach will potentially allow determining the invariant behaviour of malware as

a subset of the studied features for each known type of malware, which in turn forms the basis for implementing a new approach to the effective detection of modified (advanced) malware.

Keywords: *cybersecurity, malware, polymorphic (metamorphic) viruses, dynamic analysis, feature space, artificial intelligence, machine learning, data dimensionality reduction, genetic algorithms.*

Актуальність та постановка завдання в загальному вигляді. Серед численних завдань забезпечення безпеки в кібернетичному просторі, розробка ефективних методів виявлення і нейтралізації шкідливого програмного забезпечення (ШПЗ) відіграє критично важливу роль. Згідно зі статистичними даними, наведеними у [1–3], зафіксовано збільшення кількості кіберінцидентів, пов'язаних із застосуванням ШПЗ на 123 % у 2023 році порівняно з попередніми роками. Такий стрімкий зріст застосування ШПЗ пояснюється відносно простим і досить ефективним підходом до здійснення прихованого інформаційно-руйнівного впливу, обумовленого його поліморфізмом (мутацією).

Ключову роль у процесі створення поліморфного, олігоморфного і метаморфного ШПЗ відіграють все більш доступні та широко впроваджені системи і технології штучного інтелекту (ШІ), що дозволяє зловмисникам досить швидко та ефективно модифікувати код та/або структуру вже відомих шкідливих програм, не потребуючи при цьому значних спеціалізованих технічних компетенцій [4].

Дослідження канадської телекомунікаційної компанії BlackBerry показують, що 48 % зловмисників використовують ChatGPT для створення нових зразків ШПЗ, тоді як 46 % – модифікують вже існуюче ШПЗ [5]. Так, на початку 2023 року зафіксовано факти використання зловмисниками системи генеративного ШІ, зокрема мовної моделі ChatGPT для створення адаптивного ШПЗ та інструментів шифрування.

Як зазначає компанія Check Point – розробник програмного і апаратного забезпечення для кіберзахисту, – російські кіберзлочинці активно тестують різноманітні способи обходу обмежень OpenAI – лабораторії досліджень технологій ШІ, щоб використовувати ChatGPT у незаконних цілях [6]. Одним із підрозділів компанії Check Point зафіксовано численні випадки маніпулювання прикладним програмним інтерфейсом ChatGPT з метою генерації шкідливих кодів і створення фішингових електронних листів [7].

Наступним зафіксованим випадком використання ChatGPT у зловмисних цілях є модифікація ШПЗ, відомого як Infostealer [8], суть застосування якого зводиться до пошуку поширених типів файлів в операційній системі об'єкта кібератаки з подальшим їх архівуванням та відправкою на віддалений FTP-сервер.

У звітах компанії CyberArk [9], діяльність якої тісно пов'язана з інформаційною безпекою, представлено результати експерименту щодо генерації поліморфного (метаморфного) ШПЗ засобами ChatGPT на основі коду відомих шкідливих програм, що дозволило отримати кілька варіантів одного інформаційно-руйнівного впливу без демонстрації підозрілої активності.

На початку 2023 року ІТ-спеціалісти компанії HYAS, яка спеціалізується на протидії кіберзагрозам, представили докази на основі концепції PoC (Proof of Concept – демонстрація практичної реалізації) щодо використання технологій-ШІ у сфері кібербезпеки. Перший PoC-експлойт, BlackMamba [10], використовує технології ШІ для синтезу функцій поліморфного кейлоггера – програмного забезпечення (ПЗ) реєстрації дій користувача, який здатний динамічно модифікувати безпечний код під час виконання, без залучення командно-контрольованої інфраструктури.

Розроблений компанією HYAS PoC-експлойт – EyeSpy [11] являє собою новий тип поліморфного (метаморфного) ШПЗ, який використовує технології ШІ для прийняття обґрунтованих рішень та розширення власних можливостей. Ця шкідлива програма постійно модифікує власний код з метою уникнення виявлення системами захисту, а також використовує стратегії, які постійно розвиваються для ускладнення процесу виявлення.

Очевидно, головна мета неперервного вдосконалення шкідливих програм – заподіяння інформаційно-руйнівного впливу об'єкту атаки за умови запобігання ідентифікації системами виявлення вторгнень (кібератак) – IDS/IPS (Intrusion Detection and Prevention System), системами управління інформаційною безпекою (подіями безпеки) – SIEM (Security Information and Event Management), антивірусними програмами та системами проактивного виявлення нетипових загроз і цільових кібератак на кінцевих точках – EDR (Endpoint Detection and Response). Так, хоч зазначені системи постійно вдосконалюються, проте ефективність застосування поліморфного і метаморфного ШПЗ залишається високою протягом багатьох років [12]. Успіх їхнього застосування обумовлено, насамперед, тим, що існуючі методики і стратегії удосконалення систем кіберзахисту ґрунтуються на дослідженні інцидентів безпеки постфактум (після їх виникнення). По-друге, використання інструментів статичного аналізу ШПЗ часто не є ефективним, внаслідок використання прийомів обфускації та шифрування коду ШПЗ. По-третє, існуючі підходи до динамічного аналізу ШПЗ часто зосереджені на досліджені обмеженого кола дій (спроб доступу до певних файлів або системних ресурсів) [13].

Враховуючи темпи розвитку сучасних інформаційних технологій, слід зазначити, що використання систем і технологій ШІ значно прискорює еволюцію поліморфного (метаморфного) ШПЗ, роблячи процес його створення автоматизованим, значно швидшим та доступнішим, що у свою чергу дозволяє ефективно здійснювати приховану деструктивну діяльність протягом тривалого часу.

Таким чином, ефективне виявлення і нейтралізація поліморфного (олігоморфного) і метаморфного ШПЗ вимагають постійного аналізу та адаптації до нових сценаріїв їх застосування, що у свою чергу зумовлює необхідність розробки інноваційних підходів за цією тематикою.

Аналіз попередніх досліджень. Науковим дослідженням щодо виявлення ШПЗ, яке постійно самомодифікується (еволюціонує) та/або використовує різноманітні методи обходу існуючих систем захисту, присвячено значну кількість робіт [14–25], аналіз найперспективніших із яких викладено нижче.

У дослідженнях [14–15] розглядаються перспективи використання ***статичного аналізу*** для вивчення програмного коду, усіх можливих його станів та дефектів без фактичного виконання. Основна перевага зазначеного підходу полягає в здатності досліджувати всі можливі шляхи виконання коду та значень змінних, що в подальшому сприяє виявленню аномальної поведінки. Враховуючи те, що статичний аналіз базується на раніше набутому досвіді, він є ефективним інструментом для виявлення відомого ШПЗ, однак має суттєві обмеження у виявленні нових, поліморфних (метаморфних) зразків ШПЗ.

У роботах [14–15] використовується ***динамічний аналіз*** для вивчення ШПЗ безпосереднього під час його фактичного виконання в ізольованому середовищі (пісочниці). Цей підхід формує поведінкові шаблони ПЗ у реальному часі, які в подальшому можуть бути використані для виявлення нових зразків ШПЗ, але з певними обмеженнями у часі. До того ж, такий підхід неспроможний надавати латентні (приховані/неочевидні) шаблони поведінки ШПЗ, що у свою чергу не дозволяє глибоко зрозуміти моніторингові дані та обґрунтовувати відповідні прийняті рішення з належним рівнем впевненості.

У дослідженнях [16–17] описується використання ***евристичного аналізу*** для виявлення ШПЗ. Такий підхід ґрунтується на застосуванні набору правил (евристик) для ідентифікації підозрілої поведінки програм, що може вказувати на потенційну присутність ШПЗ у системі. Це дозволяє виявляти як відоме, так і певною мірою поліморфне, олігоморфне і метаморфне ШПЗ. Проте цей підхід має досить велику кількість хибних спрацювань.

У роботах [15; 18] використовується метод ***аналізу поведінки*** для виявлення ШПЗ, що передбачає постійний моніторинг активності програм та системи з подальшим порівнянням з

еталонними зразками поведінки. Будь-яке відхилення від типової поведінки свідчить про можливе функціонування ШПЗ. Хоча аналіз поведінки ефективний у виявленні відомих зразків ШПЗ, проте він менш ефективний у виявленні поліморфного (метаморфного) ШПЗ. Однак характеризується відсутністю обмежень, пов'язаних із необхідністю попереднього знання про конкретні віруси або їхні сигнатури.

У роботах [19; 20] описується застосування методів ***машинного навчання*** для виявлення ШПЗ. Ці методи здатні розпізнавати закономірності (шаблони), що властиві ШПЗ у великих наборах даних з великою ефективністю. Поряд з цим, їхня ефективність значно залежить від тренувального набору даних (навчальної вибірки), для якого часто не піднімається питання актуальності і дисбалансу класів.

У дослідженнях [21–24] запропоновано використання ***декларативного підходу*** для виявлення ШПЗ. Так, передбачається порівняння поведінки відомого ШПЗ та нового ПЗ на предмет виявлення збіжностей і повторюваних шаблонів у трасах системних викликів, які властиві шкідливим програмам. Реалізація такого підходу забезпечується засобами Declare – мовою моделювання процесів як множини обмежень у системі і послідовності дій. Основними недоліками цього підходу є складність інженерії коректних трас системних викликів, а також необхідність реєстрації трас системних викликів для кожної операційної платформи окремо.

У роботі [25] зазначаються дві стратегії виявлення поліморфного (метаморфного) ШПЗ: виявлення аномальної діяльності на основі наявних шаблонів функціонування системи, яка підлягає захисту, та виявлення за ознаками відомого ШПЗ. Однак з огляду на те, що більшу кількість поліморфного (метаморфного) ШПЗ складають модифіковані зразки вже існуючого шкідливого коду, доцільним є використання останньої стратегії.

Підсумовуючи вищезазначене, можна зробити висновок, що кожен з аналізованих підходів до виявлення поліморфного, олігоморфного і метаморфного ШПЗ характеризується тією чи іншою мірою певною множиною обмежень. До того ж, ні одна група методів не використовує на власну користь ключову особливість поліморфного (метаморфного) ШПЗ – інваріантну поведінку за певною підмножиною ознак, яка характеризує один і той самий вектор впливу ШПЗ.

У зв'язку з цим виникає актуальне завдання пошуку ефективного рішення щодо виявлення поліморфного і метаморфного ШПЗ, яке б забезпечило нівелювання властивості модифікації коду ШПЗ, а також дозволило аналізувати такі властивості ШПЗ, які характеризують його вектор інформаційно-руйнівного впливу (є інваріантними для конкретного типу ШПЗ). Відтак вирішення цього завдання доцільно здійснювати шляхом визначення інваріантної компоненти у поведінці ШПЗ на основі поєднання переваг поведінкового аналізу та техніки машинного навчання – зменшення розмірності простору ознак (Dimensionality reduction [26] – перетворення досліджуваних даних, що полягає у зменшенні числа ознак (проєкції на нову розмірність) шляхом отримання множини найбільш значущих із них).

Метою статті є обґрунтування вибору підходу до визначення інваріантної компоненти у поведінці поліморфного (метаморфного) ШПЗ на основі зниження розмірності простору ознак.

Обґрунтування вибору підходу до визначення інваріантної компоненти у поведінці ШПЗ. Вирішення завдання нівелювання результатів метапрограмування коду поліморфним (метаморфним) ШПЗ шляхом застосування методів зниження розмірності досліджуваного простору ознак обумовлено потенційним паритетом існуючих систем кіберзахисту та ефективністю уникати виявлення новими зразками ШПЗ, які побудовано на основі раніше відомого (класифікованого) ШПЗ. Важливість цього підходу полягає у можливості адаптивної реакції на постійно змінювані тактики ШПЗ, що вимагає постійного вдосконалення захисних систем без втручання експертів. Методи зниження розмірності дозволяють виділити

статистично значущі, але неочевидні патерни поведінки, що є ключовими для розпізнавання зловмисного ПЗ навіть у випадку його значних модифікацій. До того ж, застосування таких методів дозволяє покращити розуміння структури даних, що представляються навчальною вибіркою про поведінку ШПЗ. Ключовою складовою такого підходу є використання розширеного аналізу простору ознак з метою ідентифікації таких, які залишаються стабільними незалежно від зовнішніх модифікацій коду. Оскільки переважну більшість екземплярів поліморфного (метаморфного) ШПЗ складають модифіковані зразки вже існуючого шкідливого коду [13], то за умови збереження вектора здійснення інформаційно-руйнівного впливу залишається незмінною певна підмножина ознак для кожного типу (класу) ШПЗ. Звідси шукана підмножина ознак описує інваріантну поведінку для кожного типу (класу) ШПЗ, тоді як решта ознак описують його поліморфну (метаморфну) компоненту.

На рисунку 1 зображено узагальнену схему визначення інваріантної компоненти у поведінці ШПЗ на основі зниження розмірності простору досліджуваних ознак, де x_i – ознака.

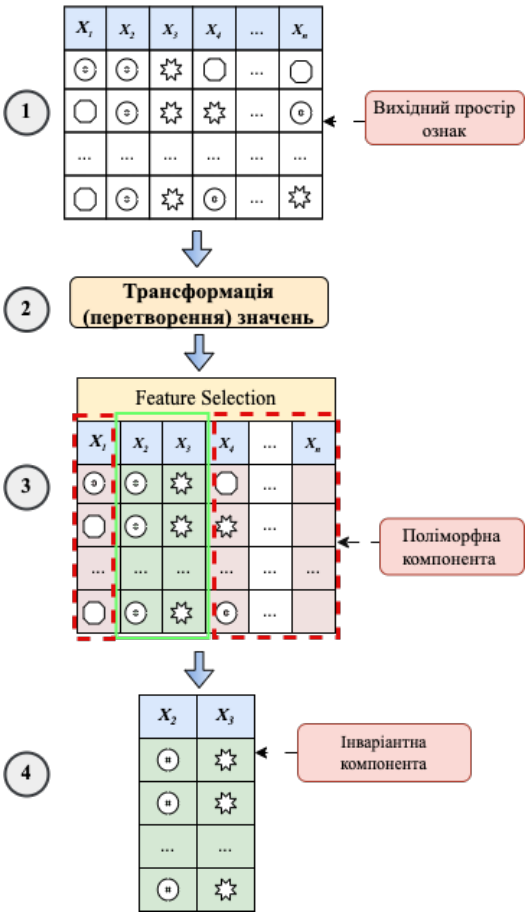


Рис. 1. Узагальнена схема визначення інваріантної компоненти у поведінці ШПЗ на основі зниження розмірності простору досліджуваних ознак

На *першому* кроці ініціалізується вихідний простір ознак $x_1 - x_n$, який описує поведінку ШПЗ, представлену навчальною вибіркою.

На *другому* кроці здійснюється трансформація (перетворення) значень ознак шляхом кодування, наприклад, нечіткими лінгвістичними термами з метою подальшого ефективного формування як множин нечітких правил для кожного відомого класу ШПЗ, так і підмножин нечітких інваріантних компонент для поліморфного (метаморфного) ШПЗ, асоційованого з цими класами ШПЗ.

Таблиця 1

Порівняльна характеристика методів зниження розмірності простору ознак

№	Метод	Основна ідея	Стійкість до шуму	Часова складність	Типи вихідних даних	Результат зменшення розмірності	Можливість інтерпретації отриманого результату правилами	Лінійність
1.	Метод фільтрації	Виділення найбільш інформативних ознак з великої кількості ознак	Чутливий	Середня	- числові; - категоріальні; - текстові; - зображення	Оцінка та вибір найбільш інформативних ознак з урахуванням статистичних метрик, без створення нових ознак	+	Лінійний
2.	Метод опорних ознак	Вибір найкращої підмножини ознак для використання в моделі машинного навчання, виходячи з якості продуктивності моделі на конкретних комбінаціях ознак	Чутливий	Висока	- числові; - категоріальні	Оцінка та вибір найкращої підмножини ознак, без створення нових ознак	+	Гібридний
3.	Топологічний аналіз даних (TDA)	Виявлення складних зав'язків (структур) між даними	Стійкий	Висока	- числові; - зображення; - категоріальні; - геометричні; - текстові; - мережеві; - сигнали; - часові ряди; - множини даних	Топологічні структури (нові дані)	+/-	Нелінійний
4.	Факторний аналіз (FA)	Виявлення латентних зав'язків між змінними	Чутливий	Середня	- числові; - категоріальні	Проєкція існуючих змінних на нижчу розмірність. В результаті отримуються нові змінні, які являють собою лінійні комбінації вихідних ознак	+	Лінійний
5.	Дискримінантний аналіз	Знаходження дискримінантних функцій	Чутливий	Низька	- числові	Перетворення даних на новий простір ознак, використовуючи лінійні комбінації вихідних ознак	+/-	Лінійний
6.	Т-розподілене вкладення стохастичної близькості (t-SNE)	Знаходження нелінійних залежностей між вихідними даними	Чутливий	Висока	- числові; - категоріальні	Подання даних для візуалізації на зменшеному просторі, створення нових ознак	+	Нелінійний
7.	Метод головний компонент (PCA)	Знаходження головних компонент	Чутливий	Низька	- числові	Нові ознаки	+/-	Лінійний
8.	Генетичний алгоритм	Моделювання процесів природного відбору та еволюції для пошуку оптимальних рішень	Стійкий	Висока	- числові	Відбір оптимальної комбінації ознак, яка зберігає найбільшу кількість інформації при зниженні розмірност	+	Гібридний
9.	Розклад невід'ємних матриць (NMF)	Розклад матриці на невід'ємні складові	Чутливий	Низька	- текст; - числові; - зображення; - сигнали; - спектральні	Інтерпретація та відображення латентних структур даних, які є новими ознаками	+/-	Нелінійний
10.	Сингулярний розклад матриці (SVD)	Розклад матриці на такий спосіб, щоб виділити основні характеристики даних	Чутливий	Висока	- числові; - текстові; - сигнали; - спектральні	Нові ознаки, які є лінійними комбінаціями вихідних ознак	+	Лінійний
11.	Канонічно-кореляційний аналіз (CCA)	Виявлення кореляційних зав'язків	Чутливий	Висока	- числові	Нові змінні	+/-	Лінійний

На *третьому, четвертому* кроках застосовуються методи зниження розмірності досліджуваного простору ознак з метою визначення інваріантної компоненти у поведінці поліморфного (метаморфного) ШПЗ для кожного відомого класу ШПЗ. Досягнення зазначеної цілі ґрунтується на тому, що кожен окремий екземпляр конкретного класу ШПЗ за своєю суттю є поліморфним (олігоморфним/метаморфним) іншим екземплярам цього класу ШПЗ.

З метою реалізації зазначеного необхідно здійснити порівняльний аналіз існуючих методів зниження розмірності простору ознак для подальшого використання найефективнішого з них.

В математичній статистиці та деяких розділах штучного інтелекту, зокрема у машинному навчанні, методи зниження розмірності простору ознак дозволяють відібрати (Feature selection [26]) або згенерувати (Feature engineering [27]) найбільш інформативні ознаки з великих масивів досліджуваних даних шляхом виключення неінформативних, дубльованих і шумових ознак, зберігаючи при цьому важливу інформацію (приховану структуру даних у процесі проєкції даних на меншу розмірність ознак). У таблиці 1 наведено порівняльну характеристику найбільш поширених методів зниження розмірності простору ознак за характеристиками: стійкості до шуму у даних, часової складності, типів даних, можливості інтерпретації отриманого результату і лінійності.

Метод фільтрації [28] в машинному навчанні здійснює відбір підмножини найбільш інформативних ознак з вихідної множини. Цей метод оцінює і ранжує ознаки за мірою їхньої інформативності для подальшого навчання моделі, використовуючи статистичні метрики або евристики, такі як дисперсія, інформаційний приріст та взаємна інформація. Після відбору підмножини ознак здійснюється оцінка ефективності моделі на предмет наявності випадкових та/або шумових ознак.

Переваги: швидкість та простота реалізації, зменшення ризику перенавчання моделі.

Недоліки: не враховуються взаємозв’язки між ознаками, неефективний у роботі з великими наборами даних, чутливий до шуму даних.

Метод опорних ознак (Support Vector Machines, SVM) [29] – ефективний алгоритм машинного навчання для класифікації та регресії ознак на основі навчальної вибірки даних. Метод спрямований на пошук оптимального розміщення гіперплощини для ефективної сепарабельності вихідної множини ознак на класи, максимізуючи відстань між опорними векторами кожного класу. Для досягнення високої точності вводиться параметр, який контролює допустиму кількість невірно обраних ознак, що дозволяє ефективніше регулювати розподільну гіперплощину.

Переваги: здатність працювати у лінійному та нелінійному просторах, обробка великих наборів даних, висока точність та швидкість класифікації.

Недоліки: значна часова складність, значна залежність від шуму у даних.

Топологічний аналіз даних (Topological Data Analysis, TDA) [30] є інноваційним підходом до аналізу багатовимірних даних, що ґрунтується на принципах топології, вивчаючи властивості об’єктів (петель, порожнеч, групувань) під час неперервних змін. Суть використання TDA в задачах зниження розмірності простору ознак полягає у виявленні складних неочевидних структур (форм) та взаємозв’язків, які можуть бути приховані за шумом та нелінійністю даних. Особливості виявлених топологічних структур представляються меншою кількістю ознак.

Переваги: виявлення складних структур, стійкість до шуму, робота з різними типами даних.

Недоліки: значна часова складність, складна інтерпретація результатів.

Факторний аналіз [31] – це статистичний метод, який використовується для обробки великих наборів змінних (ознак). Цей метод дозволяє описати загальну варіацію даних вихідного набору ознак на основі підходу Feature engineering. Так, завдання зниження

простору ознак досягається шляхом опису вихідного простору ознак меншою кількістю знайдених прихованих взаємозв'язків – факторів (лінійних комбінацій ознак) із кореляційної матриці. В подальшому отримані фактори представляються новими ознаками.

Переваги: виявлення прихованих взаємозв'язків між ознаками, збереження прихованої структури у даних.

Недоліки: складна інтерпретація результатів, складність вибору оптимальної кількості факторів для аналізу, у випадку побудови класифікатора з'являється необхідність постійного перетворення вхідних даних.

Дискримінантний аналіз [32] – це статистичний метод, що призначений для класифікації об'єктів на групи на основі їхніх ознак. Суть застосування такого підходу в задачах зниження розмірності простору ознак зводиться до пошуку таких лінійних комбінацій вихідних ознак (дискримінантних функцій), які найефективніше класифікують об'єкти. Пошук дискримінантних функцій ґрунтується на визначенні напряму, вздовж яких максимізується відношення міжкласової дисперсії до внутрішньокласової дисперсії, що дозволяє ефективніше розділяти класи.

Переваги: збереження прихованої структури у даних.

Недоліки: чутливий до аномалій, у випадку побудови класифікатора з'являється необхідність постійного перетворення вхідних даних.

T-розподілене вкладення стохастичної близькості (T-Distributed Stochastic Neighbor Embedding, T-SNE) [33] – метод нелінійного зменшення розмірності даних та їх візуалізації у меншому просторі. В багатовимірному просторі на основі нормального закону розподілу обчислюються подібності між точками в просторі координат, які представляють досліджувані дані з метою подальшої мінімізації різниці між попарними подібностями. Після оптимізації відстаней координати точок представляють зменшену розмірність даних, зберігаючи при цьому початкову структуру і відношення.

Переваги: ефективна візуалізація, збереження локальних структур, чутливість до збіжностей.

Недоліки: низька масштабованість, значна часова складність.

Метод головних компонент (Principal Component Analysis, PCA) [32] – один із найбільш поширеніших лінійних методів зниження розмірності даних, що дозволяє зберегти важливу інформацію (структуру) з вихідного набору даних. Основна ідея полягає у знаходженні ортогональних компонент, які максимально відображають варіацію у вихідних даних, шляхом їх проєкції, щоб всі вісі захоплювали максимальну дисперсію даних.

Переваги: збереження основної структури при проєкції на зменшену розмірність, проєкція для максимальної дисперсії.

Недоліки: чутливість до викидів та шуму, у випадку побудови класифікатора з'являється необхідність постійного перетворення вхідних даних.

Генетичні алгоритми (Genetic Algorithms) [34] є еволюційним підходом до вирішення задачі зниження розмірності простору ознак, який базується на імітації природного процесу еволюції шляхом відбору підмножини найбільш пристосованих ознак з усієї множини простору ознак. Суть застосування цього підходу полягає у виборі таких рішень – хромосом (підмножин ознак), побудованих з генів, – ознак, які найбільше відповідають функції пристосованості

Основні кроки включають створення початкової популяції хромосом, де кожна хромосома являє собою комбінацію ознак, оцінку пристосованості кожної хромосоми та відбір найбільш пристосованих для наступного покоління, генерацію нового покоління шляхом мутації або кросоверу (схрещування) та оцінку нового покоління для подальшої заміни в популяції. Такий процес повторюється до досягнення критерію зупинки, такого як кількість ітерацій або досягнення заданого рівня пристосованості.

Переваги: пошук глобального оптимуму, гнучкість, стійкість до шуму.

Недоліки: значна часова складність.

Розклад невід'ємних матриць (Non-negative Matrix Factorization, NMF) [35] – алгоритми багатовимірного аналізу даних, які використовуються для зменшення розмірності простору ознак шляхом розкладання вихідної матриці даних на дві невід'ємні матриці меншої розмірності. Вихідна матриця містить об'єкти (рядки) та їхні ознаки (стовпці). Під час розкладання матриці отримуємо: матрицю базисних векторів, яка визначає базові ознаки досліджуваних об'єктів, та матрицю зі зваженими ознаками.

Переваги: легка інтерпретація результатів, збереження важливої інформації.

Недоліки: вибір початкових значень матриці, локальні мінімуми – менш точні результати, чутливий до шуму.

Сингулярний розклад матриці (Singular Value Decomposition, SVD) [32] – метод розкладання матриці для аналізу та обробки даних, зокрема для зниження розмірності даних. Перший етап SVD полягає в побудові вихідної матриці даних, де рядки відповідають об'єктам, а стовпці – їхнім ознакам. Для коректності розкладу застосовується нормалізація даних, оскільки SVD є дуже чутливим до масштабування. Наступним етапом є розкладання вихідної матриці на добуток трьох ранг-матриць:, кожна з яких представляє окрему частину інформації вихідної матриці. Головні компоненти, які необхідно зберегти, обираються на основі важливості інформації та використовуються у формуванні нової матриці зі зменшеною розмірністю.

Переваги: стійкий до шуму, робота з різними типами даних.

Недоліки: складний для розуміння та реалізації, значна часова складність, складна інтерпретація результатів, часткова втрата інформації.

Канонічно-кореляційний аналіз (Canonical-Correlation Analysis, CCA) [36] – статистичний метод дослідження взаємозв'язків між наборами даних. Основною метою CCA є знаходження лінійних комбінацій (нових змінних) кожного набору даних, які максимально корелюють між собою. Це дозволяє виокремити найважливіші кореляційні структури для подальшої інтерпретації результатів. Під час зниження розмірності за допомогою CCA обирається кількість канонічних змінних, які найефективніше пояснюють взаємозв'язки між наборами даних, що дозволяє забезпечити оптимальне співвідношення між розмірністю даних та їхнім інформаційним вмістом.

Переваги: максимальна кореляція, виокремлення важливої інформації.

Недоліки: складність інтерпретації результатів, залежність від лінійності, вразливість до викидів.

На основі проведеного аналізу зазначених методів зниження розмірності простору досліджуваних ознак можна зробити наступні висновки:

1. Для вирішення поставленого завдання найбільш доцільними є застосування методів, суть застосування яких зводиться до використання підходу ***Feature selection***: *метод фільтрації, метод опорних ознак*, *генетичні алгоритми*, оскільки на відміну від підходу ***Feature engineering***: *топологічний аналіз даних*, *факторний аналіз*, *дискримінантний аналіз*, *Т-розподілене вкладення стохастичної близькості, метод головних компонент, розкладання невід'ємних матриць, сингулярне розкладання матриці, канонічно-кореляційний аналіз*, не передбачається перетворення вихідного простору ознак на нові ознаки.

2. З-поміж методів ***Feature selection*** найбільшої уваги заслуговують методи, функціональне ядро яких передбачає використання генетичних алгоритмів, оскільки дозволяють потенційно досягти глобального оптимуму в задачах пошуку оптимальної підмножини найбільш значущих ознак.

3. Використання генетичних алгоритмів для вирішення завдання визначення інваріантної і поліморфної компонент у поведінці ШПЗ на основі зниження розмірності

простору досліджуваних ознак не є чутливим до шуму (аномалій) в наборах даних, які обираються у якості навчальної вибірки, що значно підвищує показник точності моделі.

Таким чином, для підвищення ефективності виявлення поліморфного, олігоморфного і метаморфного ШПЗ запропоновано визначати інваріантну компоненту в поведінці відомих його типів на основі зниження розмірності простору ознак засобами генетичних алгоритмів. Особливістю застосування цього підходу є збереження основної структури (топологічної форми) даних, прихованої у навчальній вибірці значень про поведінку ШПЗ у системі в умовах наявності шуму у даних, а також використання ключової характеристики поліморфного (метаморфного) ШПЗ – інваріантної поведінки, властивої конкретному типу ШПЗ, описану певною підмножиною ознак. Такий підхід дозволяє сформувати фундамент для пошуку неочевидних збіжностей у поведінці різних екземплярів деструктивного ПЗ (вірусів).

Висновки. У статті вирішується завдання обґрунтування вибору підходу до визначення інваріантної компоненти у поведінці поліморфного (метаморфного) ШПЗ на основі зниження розмірності простору ознак.

Актуальність зазначеного завдання обумовлено стрімким збільшенням кіберінцидентів протягом останнього часу, пов’язаних із застосуванням поліморфного (метаморфного) ШПЗ. Основна причина цього зросту – доступність технологій ШІ, які дозволяють зловмисникам досить швидко та ефективно модифікувати код вже класифікованих шкідливих програм, не потребуючи значних спеціалізованих компетенцій.

Проведені дослідження існуючих підходів до виявлення поліморфного, олігоморфного і метаморфного ШПЗ показали наявність певних обмежень, що значно знижують їхню ефективність. До того ж, ні одна група методів не використовує на свою користь ключову особливість поліморфного (метаморфного) ШПЗ – інваріантну поведінку за певною підмножиною ознак, яка характеризує один і той самий вектор деструктивного впливу ШПЗ. З урахуванням зазначеного, запропоновано підхід до вирішення завдання виявлення поліморфного (метаморфного) ШПЗ, який ґрунтується на ідеї визначення інваріантної компоненти у поведінці ШПЗ на основі поєднання переваг поведінкового аналізу та зменшення розмірності досліджуваного простору ознак ШПЗ.

Проведено аналіз існуючих методів зниження розмірності простору ознак, результати аналізу якого виділяють генетичні алгоритми як групу найефективніших еволюційних методів пошуку оптимальної підмножини значущих ознак.

Такий підхід потенційно дозволить значно ефективніше виявляти функціонування поліморфного (метаморфного) ШПЗ на основі інформації про типову інваріантну поведінку відомих класів ШПЗ у вигляді підмножини досліджуваних ознак, що у свою чергу формує підґрунтя для реалізації нового підходу до ефективного виявлення модифікованого (удосконаленого) ШПЗ, в тому числі і засобами технологій ШІ.

Таким чином, отримані результати є підґрунтям для подальших наукових досліджень, які полягають у розробці математичної моделі визначення інваріантної компоненти у поведінці ШПЗ під час динамічного аналізу на основі зменшення розмірності простору ознак засобами генетичних алгоритмів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Системи виявлення вразливостей і реагування на кіберінциденти та кібератаки. Оперативний центр реагування на кіберінциденти державного центру кіберзахисту Державної служби спеціального зв’язку та захисту інформації України. 2023. 14 с.

2. Російські кібероперації. Аналітика за перше півріччя 2023 року. Державна служба спеціального зв’язку та захисту інформації України. 2023. 23 с.

3. Кібербезпека в інформаційному суспільстві: Інформаційно-аналітичний дайджест / відп. ред. О. Довгань; упоряд. О. Довгань, Л. Литвинова, С. Дорогих; Державна наукова установа «Інститут інформації, безпеки і права НАПрН України»; Національна бібліотека України ім. В. І. Вернадського. К., 2024. № 1 (січень). 327 с.

4. Islam M. The next frontier: AI and the evolution of polymorphic malware. *LinkedIn: Log In or Sign Up*. URL: https://www.linkedin.com/pulse/next-frontier-ai-evolution-polymorphic-malware-moinul-islam-ov4nc?trk=public_post_main-feed-card_feed-article-content.

5. ChatGPT AI technology of the century or potential weapon in the hands of cybercriminals? URL: https://blackberry.bakotech.com/chatgpt-en.

6. Generative AI is the pride of cybercrime services. *Check Point*. URL: https://blog.checkpoint.com/research/generative-ai-is-the-pride-of-cybercrime-services/.

7. Ben-Moshe S., Gekker G., Cohen G. OpwnAI: AI that can save the day or HACK it away - check point research. *Check Point Research*. URL: https://research.checkpoint.com/2022/opwnai-ai-that-can-save-the-day-or-hack-it-away/.

8. Kaspersky: more than 36 million AI & gaming credentials compromised by infostealers in 3 years. *www.kaspersky.com*. URL: https://www.kaspersky.com/about/press-releases/2024_kaspersky-more-than-36-million-ai-gaming-credentials-compromised-by-infostealers-in-3-years.

9. Shimony E., Tsarfati O. Chatting our way into creating a polymorphic malware. *Identity Security and Access Management Leader. CyberArk*. URL: https://www.cyberark.com/resources/threat-research/chatting- our-way-into-creating-a-polymorphic-malware.

10. Sims J. BlackMamba: using AI to generate polymorphic malware. *HYAS The Authority on Cyber Threat Adversary Infrastructure*. URL: https://www.hyas.com/blog/blackmamba-using-ai-to-generate-polymorphic- malware.

11. Sims J. EyeSpy Proof-of-Concept. *HYAS The Authority on Cyber Threat Adversary Infrastructure*. URL: https://www.hyas.com/blog/eyespy-proof-of-concept.

12. Sharma S. ChatGPT creates mutating malware that evades detection by EDR. *CSO Online*. URL: https://www.csoonline.com/article/575487/chatgpt-creates-mutating-malware-that-evades-detection-by-edr.html.

13. Фесьоха В. В., Кисиленко Д. Ю., Нестеров О. М. Аналіз спроможності існуючих систем антивірусного захисту та покладених у їхню основу методів до виявлення нового шкідливого програмного забезпечення у військових інформаційних системах. *Системи і технології зв'язку, інформатизації та кібербезпеки*. 2023. Т. 3. С. 143–151.

14. The differences between static and dynamic malware analysis. *Bitdefender Blog*. URL: https://www.bitdefender.com/blog/businessinsights/the-differences-between-static-malware-analysis-and-dynamic-malware-analysis/.

15. A state of the art survey on polymorphic malware analysis and detection techniques / E. Masabo et al. *RUFORUM Institutional Repository | RUFORUM Institutional Repository*. URL: https://repository.ruforum.org/sites/default/files/IJSC_Vol_8_Iss_4_Paper_9_1762_1774.pdf.

16. What is heuristics evasion? Outsmarting heuristic antivirus systems. *ReasonLabs Cyberpedia*. URL: https://cyberpedia.reasonlabs.com/EN/heuristics%20evasion.html.

17. Shetty S. What is heuristic analysis and why is it important for cybersecurity? *TechGenix*. URL: https://techgenix.com/heuristic-analysis-cybersecurity/.

18. Deng X., Mirkovic J. Malware behavior through network trace analysis. *Lecture Notes in Networks and Systems. Selected Papers from the 12th International Networking Conference*. 2020. Vol. 180. P. 3–18. URL: https://doi.org/10.1007/978-3-030-64758-2.

19. Selamat N. S., Al F. H. M. Polymorphic malware detection based on supervised machine learning. *Journal of positive school psychology*. 2022. Vol. 6, no. 3. P. 8538–8547. URL: https://journalppw.com/index.php/jpsp/issue/view/30.

20. Akhtar M. S., Feng T. Malware analysis and detection using machine learning algorithms. *Symmetry*. 2022. Vol. 14, no. 11. P. 2304. URL: https://doi.org/10.3390/sym14112304.

21. Using discriminative rule mining to discover declarative process models with non-atomic activities/ M. Bernardi et al. *International web rule symposium*. 2014. URL: https://www.semanticscholar.org/paper/Using-Discriminative-Rule-Mining-to-Discover-Models-Bernardi-Cimitile/ef426bfa04caac0c91e9e3fc476d938f27321db8.

22. Bernardi M. L., Cimitile M., Mercaldo F. Process mining meets malware evolution: a study of the behavior of malicious code. *International symposium on computing and networking – across practical development and theoretical research*. 2016. URL: https://www.semanticscholar.org/paper/Process-Mining-Meets-Malware-Evolution:-A-Study-of-Bernardi-Cimitile/7838664913ba2ab34d78f6120188293bd77a7fb3.

23. Ardimento P., Bernardi M. L., Cimitile M. Malware phylogeny analysis using data-aware declarative process mining. *IEEE conference on evolving and adaptive intelligent systems (EAIS)*. 2020. URL: https://www.semanticscholar.org/paper/Malware-Phylogeny-Analysis-using-Data-Aware-Process-Ardimento-Bernardi/859dd8a091b4af71426a189225ee09a3a2e78a69.

24. Data-Aware declarative process mining for malware detection / P. Ardimento et al. *EEE international joint conference on neural network*. 2020. URL: http://vigir.missouri.edu/~gdesouza/Research/Conference_CDs/IEEE_WCCI_2020/IJCNN/Papers/N-21418.pdf.

25. Субач І., Фесьоха В., Фесьоха Н. Фесьоха Н. О. Аналіз існуючих рішень запобігання вторгненням в інформаційно-телекомунікаційні мережі. *Information Technology and Security*. 2017. Т. 5, № 1. С. 29–41. URL: http://nbuv.gov.ua/UJRN/inftech_2017_5_1_6.

26. Sanjyal A. Dimensionality reduction VS feature selection. *Medium*. URL: https://medium.com/@asanjyal81/dimensionality-reduction-vs-feature-selection-e68f91aa8724.

27. Kumar B. What is feature engineering in dimensionality reduction - 360digitmg. *360digitmg.com*. URL: https://360digitmg.com/blog/feature-engineering-in-dimensionality-reduction.

28. Calledda C. Focus on filter methods for feature selection. *LinkedIn: Log In or Sign Up*. URL: https://www.linkedin.com/pulse/focus-filter-methods-feature-selection-carlo-calledda-7kene.

29. Banerjee S. From high dimensions to clarity: unraveling complex data with support vector machines and principal. *Medium*. URL: https://shekhar-banerjee96.medium.com/from-high-dimensions-to-clarity-unraveling-complex-data-with-support-vector-machines-and-principal-78d3871af248.

30. Munch E. A user's guide to topological data analysis. *Journal of learning analytics*. 2017. Vol. 4, no. 2. P. 47–61. URL: https://doi.org/10.18608/jla.2017.42.6.

31. Factor analysis: how to reduce the complexity and dimensionality of your data - fastercapital. *FasterCapital*. URL:https://fastercapital.com/content/Factor-Analysis--How-to-Reduce-the-Complexity-and-Dimensionality-of-Your-Data.html.

32. Baruah I. D. Dimensionality reduction techniques – PCA, LCA and SVD. *Medium*. URL: https://medium.com/nerd-for-tech/dimensionality-reduction-techniques-pca-lca-and-svd-f2a56b097f7c#:~:text=SVD%20allows%20for%20dimensionality%20reduction,significant%20singular%20values%20and%20vectors.&text=SVD%20is%20used%20in%20data,storage%20requirements%20of%20a%20matrix.&text=By%20using%20only%20the%20most,of%20noise%20in%20the%20data.

33. Pajak A. T-SNE: t-distributed stochastic neighbor embedding. *Medium*. URL: https://medium.com/@pajakamy/dimensionality-reduction-t-sne-7865808b4e6a.

34. Метод виявлення кіберзагроз на основі еволюційних алгоритмів / С. М. Лисенко, Д. І. Стопчак, В. В. Самотес. *Вісник Хмельницького національного університету. Технічні науки.* 2017. № 6. С. 81–88. URL: http://nbuv.gov.ua/UJRN/Vchnu_tekh_2017_6_15.

35. Olaya J., Otman C. Non-negative matrix factorization for dimensionality reduction. *ITM web of conferences*. 2022. Vol. 48. P. 03006. URL: https://doi.org/10.1051/itmconf/20224803006.

36. Karwowska Z. Canonical Correlation analysis−simple explanation and python example. *Medium*. URL: https://medium.com/@pozdrawiamzuzanna/canonical-correlation-analysis-simple-explanation-and-python-example-a5b8e97648d2.

DOI: 10.58254/viti.5.2024.17.193

УДК 621.391

д-р філософії Фомін М. М. ORCID: 0000-0002-6864-4238 (ВІТІ ім. Героїв Крут)

МОДЕЛЬ УДОСКОНАЛЕНОГО АЛГОРИТМУ NOMA НА ОСНОВІ РЕТРАНСЛЯЦІЇ І МОДУЛЯЦІЇ

Для успішного впровадження мереж мобільного зв'язку у спектральному діапазоні 30–300 ГГц необхідно вирішити кілька проблем, пов'язаних із кінцевими абонентськими пристроями. І однією з головних проблем є блокування невеликих об'єктів передачі прямої видимості. Шляхами вирішення цієї задачі є ущільнення стільникової інфраструктури терагерцового діапазону за допомогою рухомих вузлів мобільного зв'язку, а також застосування одночасного підключення абонентських терміналів до двох і більше стільників одночасно [1; 2].

У цій роботі пропонується для рішення такої задачі удосконалений алгоритм NOMA на основі ретрансляції і модуляції під назвою M-CO-NOMA. На відміну від відомих він відрізняється тим, що, по-перше, перед випромінюванням у радіоефір у першому часовому слоті проводиться кластеризація термінальних пристроїв у зоні обслуговування базової станції мобільної мережі із застосуванням алгоритму оцінки фрактальної розмірності і, по-друге, при формуванні загального транспортного сигналу для термінальних пристроїв абонентів базова станція модулює сигнали віддалених користувачів на квадратурну, а сигнали ближніх користувачів на реальній складовій сузір'я QPSK.

Запропонований удосконалений алгоритм M-CO-NOMA забезпечує усунення основних завад NOMA, ввівши ортогональність на стадії підготовки сигналу замість включення її в спектр. Показано, що пропонований удосконалений алгоритм ефективний за такими показниками, як зниження SER, обчислювальної складності, завад порівняно зі звичайним NOMA.

Крім того, інтегроване рішення на базі M-NOMA та CO-NOMA дозволило отримати нові інноваційні можливості по забезпеченню передачі даних при блокуванні прямої передачі в терагерцовому діапазоні частот завдяки допоміжній ретрансляції сигналів між абонентами. Розрахунок інформаційної ефективності розробленого інтегрованого рішення показав, що для систем з NOMA та M-CO-NOMA цей показник дорівнює відповідно 4,50e^{-2} та 6,00 e^{-2}. Отже, розраховане значення інформаційної ефективності системи M-CO-NOMA в 1,5 рази більше, ніж у NOMA.

Ключові слова: *терагерцовий діапазон, блокування передачі, модуляція, ретрансляція, інтегровані рішення.*

M. Fomin. A model of the improved NOMA algorithm based on representation and modulation.

For the successful implementation of mobile communication networks in the spectral range of 30-300 GHz, it is necessary to solve several problems related to terminal subscriber devices. And one of the main problems is to be blocked by small line-of-sight transmission objects. Ways to solve this problem are densification of cellular infrastructure in the terahertz range with the help of mobile mobile communication nodes, as well as the use of simultaneous connection of subscriber terminals to two or more cells at the same time [1; 2].

In this work, an improved NOMA algorithm based on relaying and modulation called M-CO-NOMA is proposed to solve this problem. Unlike the known ones, it differs in that, firstly, before broadcasting into the radio air in the first time slot, terminal devices are clustered in the service area of the base station of the mobile network using the algorithm for estimating the fractal dimension and, secondly, during the formation of the general transport signal for terminal devices of subscribers, the base station modulates the signals of remote users on the quadrature, and the signals of nearby users on the real component of the QPSK constellation.

The proposed improved M-CO-NOMA algorithm eliminates the main drawbacks of NOMA by introducing orthogonality at the stage of signal preparation instead of including it in the spectrum. It is shown that the proposed improved algorithm is effective in terms of such indicators as the reduction of SER, computational complexity, and interference compared to the usual NOMA.

In addition, the integrated solution based on M-NOMA and CO-NOMA made it possible to obtain new innovative opportunities for ensuring data transmission when blocking direct transmission in the terahertz frequency range due to auxiliary signal relaying between subscribers. Calculation of the information efficiency of the developed integrated solution showed that for systems with NOMA and M-CO-NOMA, this indicator is equal to 4.50 e-2 and 6.00 e-2, respectively. Thus, the calculated information efficiency value of the M-CO-NOMA system is 1.5 times greater than that of NOMA.

Keywords: *terahertz range, transmission blocking, modulation, retransmission, integrated solutions.*

Постановка завдання у загальному вигляді

Для успішного впровадження мереж мобільного зв’язку у спектральному діапазоні 30–300 ГГц необхідно вирішити кілька проблем, пов’язаних із кінцевими абонентськими пристроями [3]. Поточні термінали мобільного зв’язку підтримують різні радіоінтерфейси на різних частотних діапазонах, здатні функціонувати від 800 МГц до 2,7 ГГц. Проте для використання вищих частотних діапазонів потрібне значне вдосконалення апаратної частини терміналів, а також оптимізація антенних систем кінцевих пристроїв для ефективної роботи в будь-якому діапазоні від 800 МГц до 300 ГГц.

Дослідження в області терагерцового діапазону показали, що передача на високих частотах може блокуватися невеликими об’єктами через коротші довжини хвиль [4]. Це призводить до падіння потужності на приймальній стороні і робить з’єднання менш надійними. Однак можливе збільшення надійності з’єднання завдяки ущільненню стільникової інфраструктури терагерцового діапазону за допомогою рухомих вузлів мобільного зв’язку, а також застосуванню одночасного підключення абонентських терміналів до двох і більше стільників одночасно.

Важливо зауважити, що блокування з’єднання в терагерцовому діапазоні є випадковою подією, яка потребує негайної реакції системи зв’язку, щоб уникнути перерви в активній сесії. Тому потрібно керувати одночасним підключенням абонентських терміналів до двох і більше з’єднань у терагерцовому діапазоні для оперативного перенаправлення сесії на іншу точку доступу у разі блокування поточного з’єднання. Такий підхід є новим функціоналом мережі доступу, що відрізняється від існуючих підходів щодо розподілу абонентського трафіку на інші технології радіодоступу.

З цієї точки зору представляє великий інтерес використання нової технології NOMA для забезпечення рішення такої задачі [5; 6]. У існуючих технологіях, таких як OFDMA, сигнали користувачів поділяються з використанням частоти або часу, або того й іншого. Технології NOMA є особливо привабливими внаслідок високої спектральної ефективності, яка може бути досягнута шляхом усунення ортогональності, яка використовується існуючими технологіями для поділу доступного спектра з метою запобігання завадам. Однак повне усунення ортогональності також викликає завади, які можуть бути видалені з використанням алгоритмів SIC, що виконуються на приймачі NOMA. Кількість процедур SIC збільшує обчислювальну складність звичайної системи NOMA.

Неортогональний множинний доступ (NOMA) є однією з інновацій, здатних у перспективі забезпечити бездротовий зв’язок 5-го і наступного поколінь (5G і B5G). Для забезпечення системи зв’язку в існуючих системах використовується поділ доступного природного спектра, що на сьогодні призвело до вичерпання більшої частини спектра. NOMA не потрібно розділяти спектр для надання зв’язку користувачам та виконання інших системних вимог, він може ефективно використовувати вихід спектра для виконання всіх сучасних вимог зв’язку поряд з Інтернетом речей (IoT), BigData та тактильним Інтернетом. Ця ефективна система використовує безліч різних механізмів подолання розриву у розділенні спектра, що дозволяє надавати послуги зв’язку вищої якості. Використання NOMA забезпечує кращу секретність та конфіденційність, високу швидкість передачі даних, пропускну здатність, енергоефективність, спектральну ефективність, рівну доступність зв’язку для користувачів (системну справедливість), знижує ймовірність простою, затримку, кількість помилок на біт (BER) тощо, може бути інтегрований з будь-яким з існуючих алгоритмів системи виходу. Поряд з ефективною стороною NOMA стикається з певними обмеженнями, які вимагають уваги, в першу чергу з великим числом завад. Це призводить до певних недоліків, включаючи складність системи, проблеми із забезпеченням безпеки та конфіденційності. Перед передачею NOMA поділяє сигнали користувачів за частотою і часом, використовуючи FDMA (Frequency Division Multiple Access – множинний доступ із поділом

каналів за частотою) та TDMA (Time Division Multiple Access – багаторазовий доступ із тимчасовим поділом). Щоб диференціювати користувачів, NOMA виділяє різну кількість потужності для різних користувачів, що підвищує рівень завад та проблеми безпеки між сигналами (рис. 1).

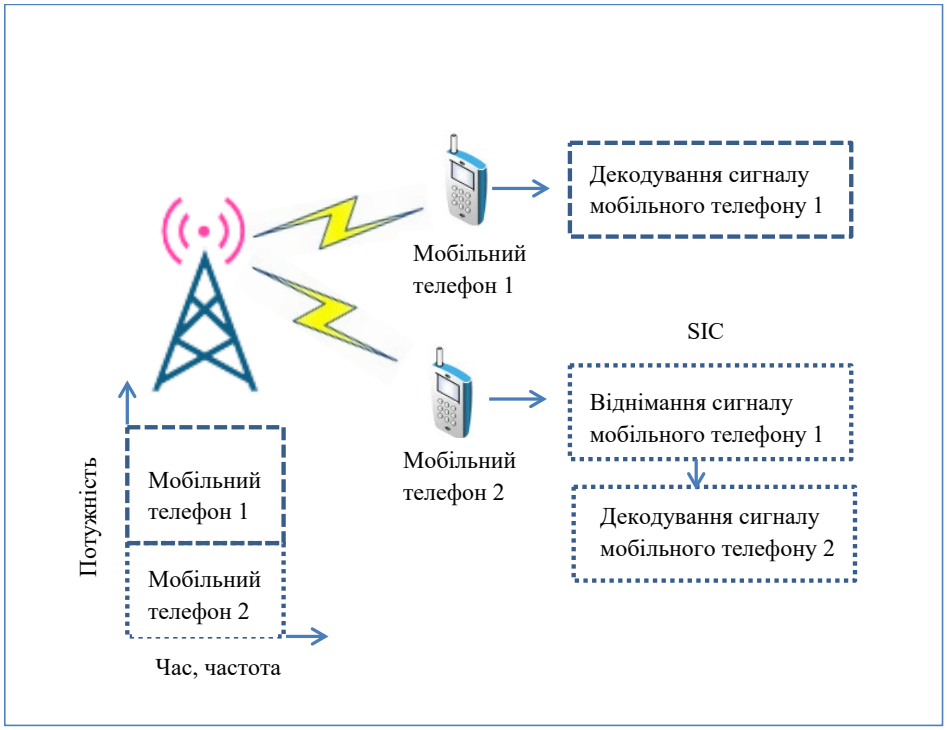


Рис. 1. Процес функціонування з технологією NOMA

Кожен користувач отримує комбінований сигнал повідомлення з завадами від усіх інших користувачів. Отже, кожен користувач повинен виконати SIC (Successive Interference Cancellation – послідовне придушення завад) для сигналів усіх інших користувачів. Більша кількість завад з боку користувачів збільшує обчислювальну складність SIC зі збільшенням кількості сигналів, що заважають. Дослідження NOMA для системи бездротового зв’язку нового покоління продовжуються.

Аналіз публікацій за темою досліджень

Для задоволення високих технологічних вимог NOMA пропонується як претендент на майбутню технологію 6G [7]. Він може забезпечити ряд переваг, включаючи надзвичайно високу швидкість, кращу спектральну ефективність, меншу затримку, чудову швидкість ергодичної суми та множинного доступу і т. ін. потужності. Виділення як субпередавачі користувачів із гіршими умовами каналу істотно впливає на спектральну ефективність системи в OMA. Використання NOMA в 5G може подолати цей недолік завдяки більш ефективному використанню спектра всіма користувачами разом із використанням різних або звичайних стратегій розподілу енергії [8; 9]. В методі множинного ортогонального доступу (OMA) у певних випадках обслуговувалися лише користувачі зі стандартними умовами каналу, тоді як у NOMA всі користувачі є рівнозначними, що забезпечує рівноцінне обслуговування поряд із високою зв’язністю та низькою затримкою [10]. Методи підвищення спектральної ефективності MIMO-NOMA та зв’язку між когнітивним радіо та NOMA обговорюються в [10]. Переваги NOMA порівняно з OMA (включаючи OFDMA) з використанням результатів моделювання продемонстровано в [11]. Компроміс між пропускною здатністю та енергоефективністю для енергетичної області OMA та кілька проблем для NOMA відображено у [12].

Інші нові ідеї в областях, пов’язаних з NOMA, охоплюють такі питання, як спільне NOMA, що використовує кодування розподілених просторово-часових блоків [13], вплив хаотично розподілених користувачів на NOMA [20], нові методи спільного використання спектра для 5G з використанням NOMA [14; 19], дослідження продуктивності збоїв [15] та NOMA на основі просторової модуляції [16–18]. Віртуальне сполучення користувачів OMA з прикордонних зон зв’язку з центральними користувачами NOMA під назвою VP-NOMA [21] надало можливість зв’язати одного центрального користувача з кількома користувачами прикордонних зон OMA. Для VP-NOMA користувачів периферійних осередків має бути більше, ніж центральних користувачів. Результат моделювання довів, що VP-NOMA працює краще ніж NOMA і OMA.

Постановка завдання

Метою цієї статті є розробка алгоритму підвищення ефективності NOMA з урахуванням завад між сигналами користувача, обчислювальної складності SIC та надійності функціонування децентралізованих мереж мобільного зв’язку 5G при динамічному блокуванні з’єднання прямої видимості рухомими перешкодами під час передачі у терагерцовому діапазоні частот.

Виклад основного матеріалу

Загальний принцип функціонування запропонованого алгоритму NOMA

Удосконалений алгоритм NOMA низхідного каналу зв’язку на основі ретрансляції та модуляції (M-CO-NOMA) для двох груп абонентів має покрокову наступну послідовність:

Крок 1. На базовій станції мобільної мережі відбувається формування загального транспортного (групового) сигналу NOMA для термінальних пристроїв абонентів.

Крок 2. Після чого сигнал випромінюється в радіоефір у першому часовому слоті, який приймається термінальним пристроєм першого абоненту.

Крок 3. Здійснюється демодуляція та компенсація сигналу для термінального пристрою другого абонента з подальшою демодуляцією власного сигналу за допомогою алгоритму SIC демодуляції, при цьому регенерований сигнал для термінального пристрою другого абонента зберігається у пам’яті до моменту ретрансляції.

Крок 4. Далі протягом другого тимчасового слота відбувається передача сигналу від першого до другого термінального пристрою абонента.

Крок 5. Після прийому сигналу, що ретранслює, на стороні другого абонента відбувається гнучка демодуляція двох прийнятих сигналів.

Удосконалений алгоритм на основі ретрансляції та модуляції відрізняється від відомих [8–18] тим, що для підвищення ефективності функціонування алгоритму NOMA у ньому застосовуються наступні допоміжні технологічні операції:

– перед випромінюванням у радіоефір у першому часовому слоті проводиться кластеризація термінальних пристроїв у зоні обслуговування базової станції мобільної мережі із застосуванням алгоритму оцінки фрактальної розмірності;

– при формуванні загального транспортного сигналу для термінальних пристроїв абонентів базова станція модулює найближчих користувачів за допомогою дійсної компоненти комплексних символів сузір’я QPSK (чотирипозиційної фазової модуляції), а далеких – за допомогою уявної компоненти.

1. Алгоритм кластеризації мобільних терміналів для реалізації удосконаленого алгоритму M-CO-NOMA

Принцип запропонованого алгоритму, побудованого на основі методу аналізу фрактальної розмірності мережі [22], полягає у нарощуванні кластера та контролю приросту кількості вузлів (мобільних терміналів) Q_r. Зменшення приросту кількості вузлів до малої величини Q_0 свідчить про виділення першого кластера. Усі зв’язкові вузли позначаються як

вузли першого кластера та виключаються з подальшого розгляду. Далі процес циклічно повторюється з випадкової позиції до моменту локалізації наступного кластера. Процес локалізації кластерів повторюється, доки всі вузли в заданій області не будуть приписані до своїх кластерів.

Алгоритм методу кластеризації наведено рисунку 2.

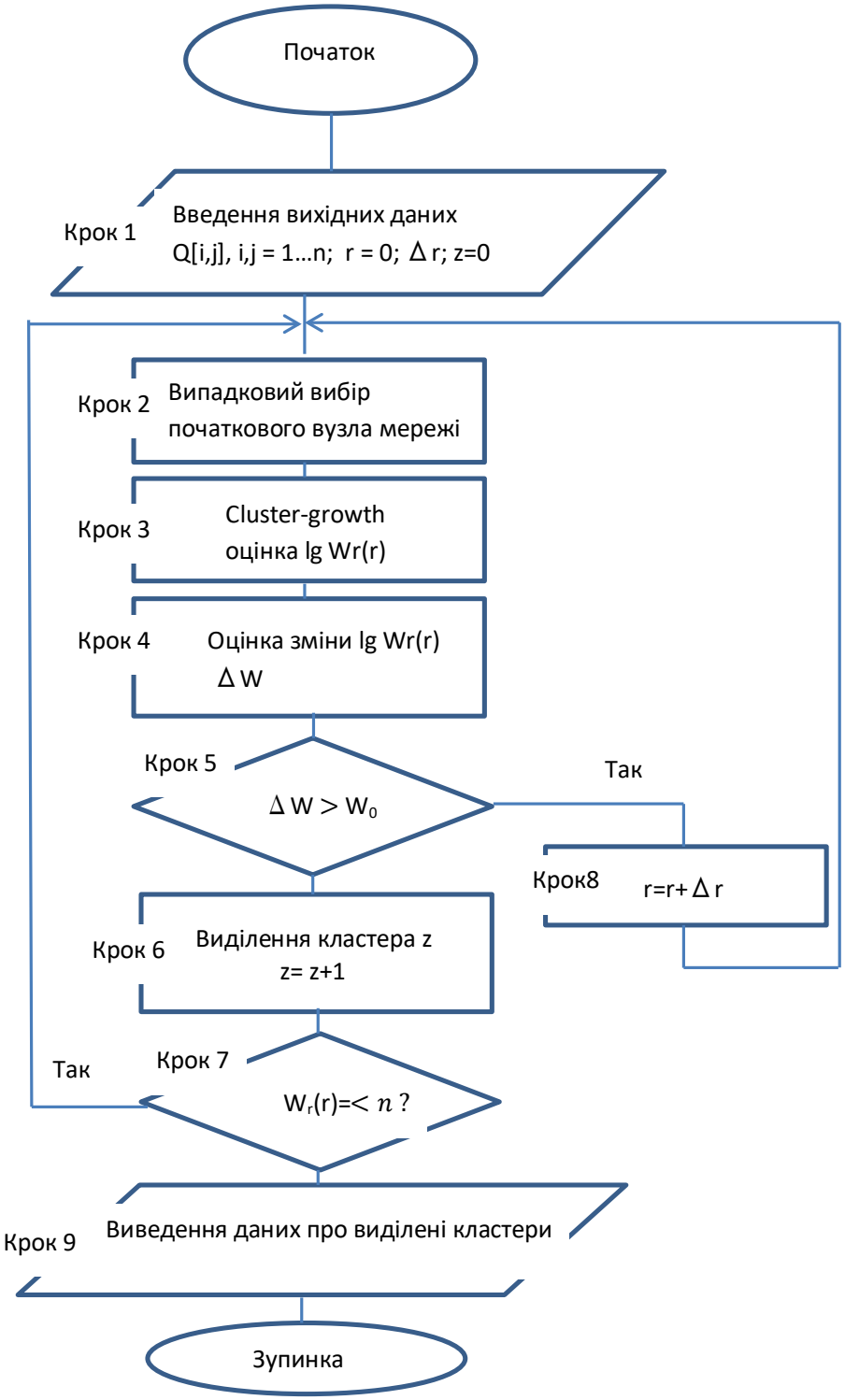


Рис. 2. Алгоритм методу кластеризації мережі

Мережа описана матрицею Q розміру n × n:

$$Q = \begin{bmatrix} d_{11} & Z & d_{1n} \\ W & d_{ij} & W \\ d_{n1} & Z & d_{nn} \end{bmatrix}.$$

Елементи матриці d_{ij} є метрикою, що характеризує досяжність вузла j з вузла i. Як було зазначено вище, такою метрикою може бути відстань або інші метрики, отримані на основі рівня потужності сигналу або відношення сигнал/шум, згасання.

У запропонованому алгоритмі d_{ij} це рівень потужності сигналу, що може бути виміряний для реальної мережі, а у моделі обчислений з урахуванням моделі згасання і потужності передавача:

$$d_{ij} = p_j - \propto_{ij},$$

де p_j – рівень потужності передавача j-го вузла (дБм), $\propto_{ij}$ – загасання сигналу між вузлами i та j (дБ).

Збір даних про основні параметри функціонування мережної інфраструктури здійснюється з використанням системи моніторингу. Ключовою відмінністю цього рішення від традиційних сучасних рішень, які використовуються операторами мобільного зв'язку, є те, що в основі процесу моніторингу є кінцеві термінали абонентів, які здійснюють збір даних про показники ефективності функціонування мережі. Таке рішення неможливо реалізувати існуючими засобами моніторингу, які передбачені у стандарті 5G, оскільки вони передбачають централізоване функціонування в межах домену одного оператора, що суттєво ускладнить процес узгодження результатів моніторингу між операторами для формування консенсусної бази даних.

Збирання даних про основні параметри роботи мережної інфраструктури відбувається за допомогою системи моніторингу. Головною відмінністю цього підходу від традиційних методів, що використовуються операторами мобільного зв'язку, є те, що в процесі моніторингу використовуються кінцеві термінали абонентів для збору даних про продуктивність мережі. Такий підхід не можна реалізувати за допомогою існуючих інструментів моніторингу, передбачених у стандарті 5G, оскільки вони передбачають централізоване управління в межах одного оператора, що ускладнює процес узгодження результатів моніторингу між операторами для створення єдиного джерела даних.

Вибір метрики віддзеркалюється на правилі формування кластера, що використовується методом Cluster-growth. Кластер формується, починаючи з випадкового стартового вузла, приєднанням нових вузлів, якщо значення метрики не менше r. Побудова кластера може бути реалізована пошуком углиб [23].

Розмір Δr є збільшенням, а r при побудові залежності розміру мережі від величини метрики підбирається емпірично.

Вихідними даними (крок 1) для даного алгоритму є дані про кількість вузлів мережі n матриці Q, що містить значення метрики, величина кроку зміни Δr, лічильник числа кластерів z.

Далі (крок 2) випадковим чином вибирається вузол, який використовується як стартовий для наступного кроку алгоритму.

На наступному етапі виконується побудова кластера з поточною величиною метрики r. Результатом є логарифм кількості вузлів у побудованому кластері, яке може набувати значень від 0 до lg n.

На кроці 4 проводиться оцінка зміни кількості вузлів у кластері порівняно з попереднім циклом. Зміна є випадковою, а її величина залежить від Δr і розподілу метрики в матриці $\boldsymbol{R}$, тому прийняття рішення на наступному кроці 5 про те, що приріст розміру мережі малий, робиться за декількома послідовними результатами.

На кроці 5 приймається рішення про виділення кластера або подальше продовження його локалізації. Кластер виділяється, якщо приріст розміру мережі не перевищує величини W_0, інакше значення метрики збільшується на величину r (крок 8) і процес повертається до кроку 2.

Якщо критерій виконується, то на кроці 6 проводиться виділення вузлів локалізованого кластера та виключення їх із подальшого розгляду. Кількість локалізованих кластерів z зростає на одиницю.

На наступному кроці проводиться перевірка кількості вузлів мережі, не віднесених до одного кластеру, якщо вона не нуль, то процес виділення кластерів повторюється, тобто провадиться перехід до кроку 2.

В іншому випадку процес кластеризації завершено, на кроці 9 проводиться висновок результатів кластеризації і зупиняється.

Порівняння запропонованого алгоритму з алгоритмами FOREL, k-середніх [24] за вимогами до вихідних даних наведено в таблиці 1.

Таблиця 1

Алгоритм			
	Кількість кластерів	Розмір кластера	Щільність вузлів у кластері
FOREL	-	+	-
k-середніх	+	-	-
Пропонований алгоритм	-	-	-

З наведеної таблиці видно, що запропонований алгоритм вимагає найменшої кількості вихідних даних, що дозволяє його застосовувати у випадках із невідомою кількістю кластерів, їхніх розмірів і форми, а також за різної щільності вузлів у кластерах. Найбільш близьким із наведених алгоритмів можна вважати FOREL, проте для застосування в умовах різної щільності вузлів у кластерах потрібна його модифікація.

2. Особливості реалізації методів CO-NOMA та M-NOMA

2.1. Метод CO-NOMA

Як технологія ретрансляція сигналів на базі групи методів NOMA у зарубіжних джерелах називається Cooperative-NOMA (CO-NOMA). Метод CO-NOMA полягає у додатковій ретрансляції сигналу від одного абонента до іншого за умови, що канал зв'язку між абонентами є сприятливим. Процедура ретрансляції при цьому може здійснюватися у додатковий частотно-часовий ресурс, наприклад, у тимчасовий слот. Застосування ретрансляції з урахуванням NOMA найефективніше (на відміну OFDMA) оскільки груповий сигнал NOMA містить у собі інформацію про сигнали всіх абонентів, тому не потрібно виділення додаткового ресурсу передачі для того, що ретранслюється. Протягом першого тимчасового слота сигнал, сформований за технологією NOMA, передається від базової станції до абонентів MT1 та MT2. Протягом другого тимчасового слота сигнал ретранслюється від MT1 до MT2. На стороні абонента MT2 відбувається гнучка демодуляція двох прийнятих сигналів – сигналу від базової станції, прийнятого в перший часовий слот, і ретрансльованого сигналу від іншого абонента, прийнятого в другий часовий слот. Таким чином, відношення сигнал/шум на вході демодулятора повідомлення MT2 підвищується і досягається покращення якості зв'язку.

Технологія CO-NOMA запропонована зовсім недавно і являє інтерес для моделювання та експериментального дослідження. Особливо дуже актуальним питання є підвищення завадостійкості. Найбільш перспективними, з нашої точки зору, виглядають комбіновані методи, що використовують одночасно кілька ресурсних просторів для поділу користувачів і забезпечують багатовимірний доступ до радіоефіру. Такі методи множинного доступу повинні

бути динамічними та адаптивними, що дозволяють гнучко вибирати ті чи інші ресурсні простори залежно від кількості користувачів, що обслуговуються в конкретний момент, рівня завад, умов розповсюдження сигналу та інших факторів. Тому дослідження комплексного рішення із застосуванням методів множинного доступу CO-NOMA та методів з бітовим поділом NOMA на основі модуляції (modulation based NOMA, M-NOMA) є актуальною науково-технічною задачею.

2.2. Метод M-NOMA

У методі M-NOMA використовується модуляція для створення ортогональності між користувачами. Цей метод як спрощує систему, так і поліпшує її загальну продуктивність порівняно з існуючими методами NOMA. В M-NOMA базова станція поділяє ближніх та далеких користувачів за допомогою модуляції. Найближчі користувачі модулюються за допомогою дійсної компоненти комплексних символів сузір'я QPSK (чотирипозиційної фазової модуляції), а далекі – за допомогою уявної компоненти. Це забезпечує ортогональність, або кластеризацію між сигналами ближніх і віддалених користувачів. Кластеризація запобігає завадам між віддаленими та ближніми користувачами. В алгоритмі M-NOMA користувачі, розташовані ближче до базової станції, отримують сигнали повідомлення без завад, викликаних сигналами віддалених користувачів.

Метод M-NOMA дозволяє зменшити частоту символьних помилок, міжстільникові та міжкластерні завади, затримку та обчислювальну складність. Він має кращу продуктивність порівняно зі звичайним NOMA з погляду швидкості передачі даних і завад.

Для того щоб усунути розраховані на багато користувачів завади на приймачі, використовують метод послідовної компенсації завад (Successive Interference Cancellation, SIC). Крім того, SIC зменшує відношення «сигнал – завада плюс шум» (SINR) через природний ефект близької відстані або при зміні розподілу потужності між користувачами на передавачі.

При цьому ближнім та віддаленим користувачам необхідно виконувати операції SIC лише в межах одного кластера ближніх або віддалених користувачів. Для порівняння, ближні користувачі алгоритму M-NOMA виконують у N/2 t рази менше операцій SIC, ніж алгоритму NOMA, але віддалені користувачі виконують таку саму кількість операцій SIC.

3. Системна модель удосконаленого алгоритму M-CO-NOMA

Розглянемо сценарій системи мобільного зв'язку, у якому є один стільник з однією базовою станцією (БС) в центрі зони обслуговування і чотирма користувачами. Для пропонованої схеми можемо мати N користувачів у зоні обслуговування БС. Однак для простоти ми розглядаємо 4 користувачів. Користувачі відокремлені від БС на відповідні відстані. Загальна кількість користувачів поділена на дві групи кластерів. Половина (N/2) користувачів вважається ближніми, а половина, що залишилася, – віддаленими.

3.1. Принцип функціонування удосконаленого алгоритму M-CO-NOMA

Розглянемо спільну комунікацію у двох кластерах – ближньому та віддаленому. Кластер 1 – це кластер із кращим відношення потужності корисного сигналу до потужності шуму SNR (Signal-to-Noise Ratio) щодо кластера 2. W_1 – мобільний телефон найближчого користувача у кластері 1, а W_3 – мобільний телефон найближчого користувача у кластері 2 відносно БС. БС модулює сигнали користувачів кластера 1 на реальній компоненті, а сигнали користувачів кластера 2 на уявній компоненті сузір'я QPSK. Кожен користувач мобільного телефону знає, на якій компоненті знаходиться його повідомлення.

У кластері 1 мобільний термінал користувача W_1 приймає сигнал усіх інших користувачів у часовому інтервалі 1 за допомогою алгоритму CO-NOMA-M, декодує весь сигнал та виконує операції SIC для отримання власного сигналу. Після декодування сигналу користувача W_2 він пересилає сигнал мобільному телефону користувачеві W_2 у другому інтервалі часу з використанням режиму ретрансляції. При цьому при блокуванні передачі

прямої видимості в терагерцовому діапазоні мобільний телефон користувача W_2 отримує два сигнали, один від БС, а інший від користувача W_1. Система MRC (комбінуванням максимального співвідношення) декодує свій сигнал. Аналогічно, в кластері 2 мобільний телефон користувача W_3 отримує сигнал користувача W_4 разом зі своїм власним сигналом. Він також декодує та пересилає сигнал користувачу W_4.

Отже, запропоноване інноваційне рішення підвищує надійність функціонування децентралізованих мереж мобільного зв'язку 5G/6G при динамічному блокуванні з'єднання прямої видимості рухомими перешкодами під час передачі у терагерцовому діапазоні частот на основі спільного використання частотно-часових та енергетичних ресурсів із використанням ресурсів мобільних терміналів відповідних кластерів.

3.2. Оцінка ефективності функціонування удосконаленого алгоритму M-CO-NOMA

В режимі налаштування БС модулює сигнали віддалених користувачів на квадратурні компоненти, а сигнали ближніх користувачів на реальну складову сузір'я QPSK. Це робить сигнали повідомлень ортогональними, що запобігає завадам між сигналами віддалених та ближніх користувачів. У M-CO-NOMA найближчі користувачі не приймають сигнали повідомлень із завадами сигналів віддалених користувачів. Отже, їм потрібно застосовувати операції SIC лише для віднімання сигналів повідомлень ближніх користувачів. Тому ближнім та віддаленим користувачам необхідно виконувати операції SIC ($N/2 - 1$) разів. Порівняно з NOMA, CO-NOMA-M ближні користувачі виконують в N/2 рази менше операцій SIC, але віддалені користувачі виконують таку саму кількість операцій SIC. У випадку чотирьох груп користувачів у стільнику, найближчі користувачі виконують операції SIC три рази на NOMA і тільки один раз на CO-NOMA-M. Слід зазначити, що з чотирьох груп користувачів CO-NOMA-M знижує обчислювальну складність до однієї третини (1/3) значень порівняно з NOMA. У випадку віддалених користувачів кількість виконаних операцій SIC однакова, але у NOMA віддалені користувачі отримують сигнали з завадами не тільки від інших віддалених користувачів, але і від ближніх користувачів. Завади віддаленого користувача можуть бути вираховані за допомогою операцій SIC, але SIC не може бути виконаний для ближнього сигналу користувача через дуже низьку потужність цього сигналу, прийнятого віддаленим користувачем. У M-CO-NOMA-M сигнали ближніх користувачів в ідеалі не створюють завади для декодування віддалених користувачів, оскільки сигнали віддалених та ближніх користувачів ортогональні один одному. Отже, у CO-NOMA-M завади приймача для віддаленого користувача зменшуються в N/2 рази відносно рівня базового алгоритму NOMA. У випадку N користувачів завади виникають у кожного користувача приймача (N – 1) користувачів NOMA і (N/2 – 1) користувачів CO-NOMA-M. Для чотирьох користувачів завади сигналу користувача викликані трьома користувачами алгоритму NOMA і тільки одним користувачем CO-NOMA-M. Отже, алгоритм CO-NOMA-M знижує обчислювальну складність до трьох разів для ближніх користувачів. Це також зменшує загальні завади системи у 6 разів.

На рисунку 3 представлено результат моделювання алгоритму CO-NOMA-M. На ньому зображено криві операцій SER для NOMA, M-NOMA та CO-NOMA-M.

У налаштуванні для моделювання використовували МАТЛАБ 2021a та наступні вихідні дані: канал загасання Релея, втрати сигналу 2, сумарна потужність передачі становить 1 Вт, коефіцієнти розподілу потужності вибираються відповідно до відстані до кожного мобільного телефону користувача, такі як $a_1 = 0{,}15$, $a_2 = 0{,}20$, $a_3 = 0{,}25$, $a_4 = 40$.

Дані моделювання показують кращий результат для C-NOMA-M W4 порівняно з M-NOMA W3 і M-NOMA W4, а CO-NOMA-M W2 перевершує NOMA W3. Зрештою, CO-NOMA-M показує кращі результати порівняно з відомими методи NOMA.

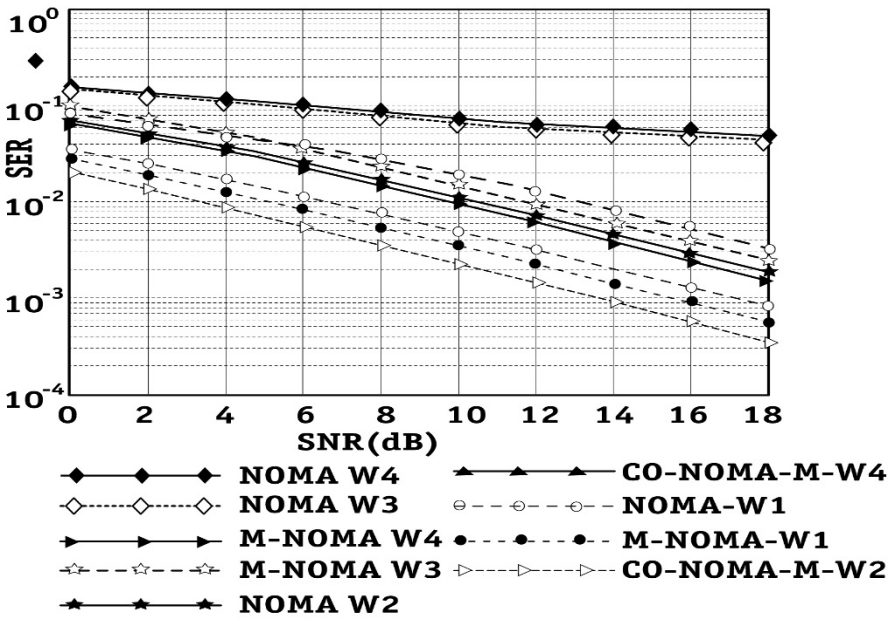


Рис. 3. Порівняння ефективності SER з використанням алгоритмів NOMA, M-NOMA та запропонованого CO-NOMA-M

Висновки

1. У роботі запропоновано для забезпечення надійності функціонування децентралізованих мереж мобільного зв’язку 5G при динамічному блокуванні з’єднання прямої видимості рухомими перешкодами під час передачі у терагерцовому діапазоні частот удосконалений алгоритм NOMA на основі ретрансляції і модуляції.

2. На відміну від відомих він відрізняється тим, що, по-перше, перед випромінюванням у радіоефір у першому часовому слоті проводиться кластеризація термінальних пристроїв у зоні обслуговування базової станції мобільної мережі із застосуванням алгоритму оцінки фрактальної розмірності і, по-друге, при формуванні загального транспортного сигналу для термінальних пристроїв абонентів базова станція модулює сигнали віддалених користувачів на квадратурну, а сигнали ближніх користувачів на реальній складовій сузір’я QPSK.

3. Показано, що пропонований удосконалений алгоритм ефективний за такими показниками, як зниження SER, обчислювальної складності, завад порівняно зі звичайним NOMA.

4. Інтегроване рішення на базі M-NOMA та CO-NOMA дозволило отримати нові інноваційні можливості щодо забезпечення передачі даних при блокуванні прямої передачі в терагерцовому діапазоні частот завдяки допоміжній ретрансляції сигналів між абонентами.

Напрямком подальшої роботи є розробка алгоритмів підвищення ефективності NOMA з урахуванням завад між сигналами користувача, проблем безпеки, енергоефективності, досяжні швидкості передачі даних та затримки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. 3GPP TS 37.340 V15.2.0: NR: Multi-connectivity; Overall description, Rel. 15 - 2018. URL: https://www.3gpp.org/ /ftp/Specs/archive/37_series/37.340/ (date accessed: 31.07.2019).

2. Moltchanov D. Improving Session Continuity with Bandwidth Reservation in mmWave Communications / D. Moltchanov, A. Samuylov, V. Petrov et al. *IEEE Wireless Communications Letters.* 2018. No. 7. P. 1–4.

3. Сайко В. Г. Мережі мобільного зв’язку нового покоління 4G/5G/6G: монографія / В. Г. Сайко, Р. С. Одарченко, А. О. Абакумова, Т. М. Наритник, В. С. Наконечний, В. М. Домрачев, С. В. Толюпа, В. Ю. Заблоцький, П. Ф. Баховський. К.: ТОВ «Про формат», 2021. 200 с.

4. Сайко В., Наритник Т. Модель побудови бездротової терагерцової мережі з підвищеною надійністю зв'язку. *International Science Journal of Engineering & Agriculture. 2023.* № 2 (2). С. 166–181. URL: https://doi.org/10.46299/j.isjea.20230202.16 (дата звернення: 10.03.2024).

5. Бакулин М. Г. Технология NOMA с кодовым разделением в 3GPP: 5G или 6G. *T-Comm: Телекоммуникации и транспорт.* 2022. Том 16. С. 4–14.

6. Arachchillage, U. S. S. S., Jayakody, D. N. K., Biswash, S. K., Dinis, R. Samaratunge. Recent Advances and Future Challenges in Non-Orthogonal Multiple Access (NOMA). *In IEEE 87th Vehicular Technology Conference (WS)*. Porto, Portugal, 2018.

7. Rappaport T. S. 6G and beyond: Terahertz communications and sensing. 2019 Brooklyn 5G Summit Keynote, Apr. 2019. [Online]. Available: URL: https://ieeetv.ieee.org/conference-highlights/keynote-tedrappaport-terahertz-communication-b5gs-2019? (date accessed: 10.01.2024).

8. Zhang N., Wang J., Kang G., Liu Y. Uplink Non-orthogonal Multiple Access in 5G Systems. *IEEE Communications Letters*. Mar. 2016. Vol. 20. No. 3. P. 458–461.

9. El-Sayed M. M., Ibrahim A. S., Khairy M. M. Power allocation strategies for Non-Orthogonal Multiple Access. *In International Conference on Selected Topics in Mobile and Wireless Networking (MoWNeT)*, Cairo, 2016.

10. Ding Z., Liu Y., Choi J., Sun Q., Elkashlan M., Chih-Lin I., Poor H. V. Application of Non-Orthogonal Multiple Access in LTE and 5G Networks. *IEEE Communications Magazine*. Feb. 2017. Vol. 55. No. 2. P. 185–191.

11. Benjebbour A., Saito K., Li A., Kishiyama Y., Nakamura T. Non- orthogonal multiple access (NOMA): Concept, performance evaluation and experimental trials. *In International Conference on Wireless Networks and Mobile Communications (WINCOM),* Marrakech, 2015.

12. Islam S. R., Avazov N., Dobre O. A., Kwak K. S. Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges. *IEEE Communications Surveys and Tutorials*. Second quarter 2017. Vol. 19. No. 2. P. 721–742.

13. Jamal M. N., Hassan S. A., Jayakody D. N. K. A New Approach to Cooperative NOMA Using Distributed Space Time Block Coding. *In IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Montreal, 2017.

14. Qureshi S., Hassan S. A., Jayakody D. N. K. Divide-and-Allocate: An Uplink Successive Bandwidth Division NOMA System. *Transactions on emerging telecommunications technologies*. 19 Jun. 2012. Vol. 2. No. 10.

15. Xu P., Yuan Y., Ding Z., Dai X., Schober R. On the Outage Performance of Non-Orthogonal Multiple Access With 1-bit Feedback. *IEEE Transactions on Wireless Communications*. Oct. 2016. Vol. 15. No. 10. P. 6716–6730.

16. Wang Z., Cao J. NOMA-Based Spatial Modulation. *IEEE Access*. 2017. Vol. 5. P. 3790–3800.

17. Zhong C., Hu X., Chen X., Ng D. W. K., Zhang Z. Spatial Modulation Assisted Multi-Antenna Non-Orthogonal Multiple Access. *IEEE Wireless Communications*. April 2018. Vol. 25. No. 2. P. 61–67.

18. Wang X., Wang J., He L., Song J. Spectral Efficiency Analysis for Downlink NOMA Aided Spatial Modulation With Finite Alphabet Inputs. *IEEE Transactions on Vehicular Technology*. Nov. 2017. Vol. 66. No. 11. P. 10562–10566.

19. Ding Z., Lei X., Karagiannidis G. K., Schober R., Yuan J., Bhargava V. K. A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends. *IEEE Journal on Selected Areas in Communications*. Oct. 2017. Vol. 35. No. 10. P. 2181–2195.

20. Ding Z., Yang Z., Fan P., Poor H. V. On the Performance of Non-Orthogonal Multiple Access in 5G Systems with Randomly Deployed Users. *IEEE Signal Processing Letters*. Dec. 2014. Vol. 21. No. 12. P. 1501–1505.

21. Shahab M. B., Shin S. Y. On the Performance of a Virtual User-pairing Scheme to Efficiently Utilize the Spectrum of Unpaired Users in NOMA. *Physical Communication*. Dec. 2017. Vol. 25. P. 492–501.

22. Морозов А. Д. Введение в теорию фракталов. М.: Институт компьютерных исследований, 2002. 160 с.

23. Берж К. Теория графов и ее применение. М.: Издательство иностранной литературы, 1962. 320 с.

24. Снитюк В. Є. Прогнозування. Моделі. Методи. Алгоритми. К.: Маклаут, 2008. 364 с.

DOI: 10.58254/viti.5.2024.18.204

УДК 681.35

Терещенко Т. П. ORCID: 0000-0002-9659-7897 (ВІТІ ім. Героїв Крут)
канд. техн. наук Остапчук В. М. ORCID: 0000-0001-5686-0198 (ВІТІ ім. Героїв Крут)
канд. техн. наук Хусаінов П. В. ORCID: 0000-0002-0675-0369 (ВІТІ ім. Героїв Крут)
Черниш Ю. О. ORCID: 0000-0002-6626-5656 (ВІТІ ім. Героїв Крут)

ОЦІНКА ТА ЗАПОБІГАННЯ ПРОЯВУ ВРАЗЛИВОСТІ SERVER-SIDE WEB APPLICATION

Вразливість (vulnerability) системи (технологічної, комунікаційної) об'єкта кіберзахисту до здійснення негативного технічного ефекту (negative technical impact) завжди базується на експлуатації її дефектів. Дефект (weakness) – властивість програмних, апаратних, програмно-апаратних або системних компонентів, які за певних умов можуть призвести до вразливості. Атака (attack) – спроба експлуатації дефекту (дефектів) системи (об'єкта кіберзахисту) для здійснення негативного технічного ефекту (несанкціоноване читання/запис даних, неправильна робота, обхід захисних механізмів, аномальне споживання ресурсів; виконання нав'язаних команд тощо) за умови прояву вразливості з використанням експлойта (exploit).

Поява дефектів притаманна всім етапам життєвого циклу системи (об'єкта кіберзахисту). Так, розрізняють дефекти етапу проєктування, реалізації, налаштування при введенні в експлуатацію, супроводженні, завершення роботи. Існування у складі компонентів системи хоча б одного дефекту, який дозволяє здійснити негативний технічний ефект, робить її вразливою. Своєчасний пошук та усунення дефектів зменшує ймовірність компрометації системи від використання, принаймні, відомих експлойтів.

Доцільність, раціональність та обґрунтованість рішень з пошуку дефектів базується на застосуванні апробованих джерел експертного досвіду світового рівня. Так, Adversarial Tactics, Techniques & Common Knowledge визначають системи класу Server-side Web Application привабливим для хакерів об'єктом атак з найбільшою середньою кількістю потенційних дефектів у складі принаймні трьох компонентів: Web Server, Web Application Server, DBMS Server. Для з'ясування розподілу дефектів Server-side Web Application за характерними класами необхідно скористатися Open Worldwide Application Security Project, для детального ознайомлення з ними – Common Weakness Enumeration. Застосування National Vulnerability Database та Common Vulnerabilities and Exposures доповнює оцінку знайденого дефекту.

Ключові слова: *кіберзахист, компрометація системи, оцінка вразливостей*

T. Tereshchenko, V. Ostapchuk, P. Khusainov, Y. Chernysh Assessment of vulnerability of Server-side Web Application and prevention of their manifestation.

The vulnerability of the system (technological, communication) of the cyber protection object to negative technical impact is always based on the exploitation of its defects. Defect is a property of software, hardware, software-hardware or system components, which under certain conditions can lead to vulnerability. Attack – an attempt to exploit a defect(s) of the system (object of cyber protection) for the implementation of a negative technical effect (unauthorized reading/writing of data, incorrect operation, circumvention of protective mechanisms, abnormal consumption of resources; execution of imposed commands, etc.) under the manifestation of vulnerability using an exploit (exploit).

The appearance of defects is inherent in all stages of the system (cyber protection object) life cycle. Thus, defects are distinguished at the design stage, implementation, adjustment during commissioning, maintenance, and completion of work. The existence of at least one defect in the components of the system, which allows a negative technical effect, makes it vulnerable. Timely detection and elimination of defects reduces the probability of system compromise using at least known exploits.

The expediency, rationality and reasonableness of solutions for finding defects is based on the application of proven sources of world-class expert experience. Thus, Adversarial Tactics, Techniques & Common Knowledge defines Server-side Web Application class systems as an attractive attack target for hackers with the highest average number of potential defects in at least three components: Web Server, Web Application Server, DBMS Server. To find out the distribution of Server-side Web Application defects by characteristic classes, it is necessary to use the Open Worldwide Application Security Project, for a detailed study of them - Common Weakness Enumeration. The application of the National Vulnerability Database and Common Vulnerabilities and Exposures complements the evaluation of the found defect.

Keywords: *cyber protection, system compromise, vulnerability assessment.*

Постановка завдання. Суб'єкти забезпечення кібербезпеки у межах своєї компетенції повинні забезпечувати періодичний аудит інформаційної безпеки систем (об'єктів

кіберзахисту) із запровадженням кращих світових практик та міжнародних стандартів. Базуючись на принципах пропорційності, адекватності і пріоритетності запобіжних заходів кіберзахисту реальним та потенційним ризикам аудит інформаційної безпеки, перш за все, спрямований на запобігання порушень у системі (об'єкта кіберзахисту): режиму сталого, надійного та нормального функціонування служб (функцій, компонентів); конфіденційності, цілісності та доступності інформаційних ресурсів [1–3].

Обов'язковою складовою аудиту інформаційної безпеки є пошук та виявлення потенційної вразливості (vulnerability) системи (технологічної, комунікаційної) об'єкта кіберзахисту. Вразливість системи – властивість системи, через використання якої створюється загроза для її безпеки, порушується сталий, надійний та штатний режим функціонування системи, здійснюється несанкціоноване втручання в її роботу, створюється загроза для безпеки (захищеності) електронних інформаційних ресурсів, конфіденційності, цілісності, доступності таких ресурсів. Дослідник потенційної вразливості (далі – дослідник) – фізична або юридична особа, яка здійснює пошук потенційної вразливості системи.

Організація пошуку потенційної вразливості системи здійснюється її власником. Пошук потенційної вразливості системи здійснюється на підставі публічної пропозиції. У публічній пропозиції, зокрема, визначається:

інформація про систему, пошук потенційної вразливості якої здійснюється;

дії дослідника щодо системи, які йому заборонено проводити;

порядок надання дослідником звіту, вимоги до його підготовки, форми.

Після завершення пошуку потенційної вразливості системи дослідник повідомляє про результати власнику системи або координатору згідно з умовами публічної пропозиції та подає йому звіт. За результатами перевірки звіту власник системи оцінює можливі наслідки використання вразливості системи для її безпеки, порушення сталого, надійного та штатного режиму її функціонування, здійснення несанкціонованого втручання в її роботу, створення загрози для безпеки (захищеності) електронних інформаційних ресурсів, конфіденційності, цілісності, доступності таких ресурсів (далі – наслідки вразливості системи) та приймає рішення щодо внесення або невнесення змін до системи.

Звіт про вразливість повинен принаймні містити опис дефекту (дефектів) та негативного технічного ефекту (negative technical impact) для системи. Вразливість системи до здійснення негативного технічного ефекту завжди базується на експлуатації її дефектів. Дефект (weakness) – властивість програмних, апаратних, програмно-апаратних або системних компонентів, які при певних умовах можуть призвести до вразливості. Атака (attack) – спроба експлуатації дефекту (дефектів) системи (об'єкта кіберзахисту) для здійснення негативного технічного ефекту за умови прояву вразливості на основі використання експлойта (exploit). Викладення дефекту (дефектів), негативного технічного ефекту спрямоване на формування технічного та контекстного розуміння вразливості.

Поява дефектів притаманна всім етапам життєвого циклу системи (об'єкта кіберзахисту). Так, розрізняють дефекти етапу проєктування, реалізації та налаштування (при введенні в експлуатацію, супроводженні, завершенні функціонування). Існування у складі компонентів системи хоча б одного дефекту, який дозволяє здійснити негативний технічний ефект, робить її вразливою. Своєчасний пошук та усунення дефектів зменшує ймовірність компрометації системи з використання, принаймні, відомих.

Для з'ясування тактик, технік, процедур (роботи експлойтів) компрометації систем (об'єктів кіберзахисту) слід звернутися до класифікаторів Adversarial Tactics, Techniques & Common Knowledge (ATT&CK) та Common Attack Pattern Enumerations and Classifications (CAPEC), які сформовані на основі розслідування відомих кіберінцидентів/кібертатак. Компрометація системи – порушення сталого, надійного та штатного режиму функціонування систем та/або порушення конфіденційності, цілісності, доступності інформації (інших

властивостей інформації), що обробляється в цих системах. Тактики, техніки, процедури – ієрархічна модель опису поведінки та дій актора (користувача, адміністратора, зловмисника) щодо будь-якої діяльності в системах/мережах, а також спроб діяльності з метою впливу на ці системи [4–7].

Методичним базисом та обов'язковим інструментарієм діяльності дослідника для аналізу ознак вразливості, оцінок, наслідків (прояв негативного технічного ефекту) тощо є Common Vulnerabilities and Exposures (CVE), Common Weakness Enumeration (CWE) [8–9].

Аналіз публікацій про вразливості та дефекти систем. Станом на 06.05.2024 облікована 233 151 вразливість продуктів та технологічних рішень у сфері інформаційних технологій. Розрізняють 11 категорій вразливості, які уособлюють технологічну основу здійснення негативного технічного ефекту. Вразливість є наслідком відсутності або неправильності роботи (недостатності, помилковості, некоректності) процедур контролю (запобігання) непередбаченого використання невід'ємних функцій компонентів системи:

Overflow – перевищення ємності місця їх розташування даних;

Memory Corruption – пошкодження пам'яті;

Sql Injection – поява в запитах мовою Structured Query Language зовнішніх даних;

Cross-Site Scripting (XSS) – обробка компрометованого гіперпосилання;

Directory Traversal – зміна робочої локації (каталогу) у файловій системі;

File Inclusion – нав'язування обробки файлу;

Cross-Site Request Forgery (CSRF) – підміна суб'єкта доступу за запитом;

XML External Entity (XXE) – обробка зовнішніх об'єктів eXtensible Markup Language;

Server-side Request Forgery (SSRF) – підміна суб'єкта доступу до довіреного сервісу;

Open Redirect – модифікація шляху доступу;

Input Validation – порушення фільтрації вхідних даних.

Середній процент загальної ваги категорії вразливості за 10 років: Overflow – 20 %; Memory Corruption – 20 %; Sql Injection – 9 %; Cross-Site Scripting – 25 %; Directory Traversal – 5 %; File Inclusion – 1,5 %; Cross-Site Request Forgery – 6 %; XML External Entity – 1,5 %; Server-side Request Forgery – 1,5 %; Open Redirect – 1,5 %; Input Validation – 9 %.

Окрім опису технологічної основи здійснення негативного технічного ефекту вразливість має доповнення у вигляді 5 форм доведеної практичної експлуатації:

Code Execution – виконання команд, програм із зовнішніх джерел;

Bypass – подолання захисних механізмів;

Privilege Escalation – розширення повноважень;

Denial of Service – порушення (припинення) функціонування;

Information Leak – порушення конфіденційності інформації.

За останні 10 років обліковано 69 833 нових вразливості, з них Code Execution – 19 052; Bypass – 7 114; Privilege Escalation – 9 920; Denial of Service – 22 644; Information Leak – 11 103.

Розрізняють 938 типових дефектів програмних, апаратних та програмно-апаратних компонентів систем (об'єкта кіберзахисту) за 374 категоріями, які були притаманні протягом 17 виділених фаз життєвого циклу системи (проєктування – 3, реалізація – 6, введенні в експлуатацію – 2, супроводження – 5, звершення роботи – 1) [7–9].

Формулювання мети статті. Оприлюднити результати проведеного дослідження категорій вразливості Server-side Web Application на основі оцінок за визначеними показниками та обґрунтовану дорожню карту заходів для запобігання їх прояву.

Виклад основного матеріалу. Поняття Web Application (далі також – Web-застосування) є сучасною практичною реалізацією ідеї, принципів та технологій доступу користувача комп'ютера до інформації, яка оформлена у вигляді документа мовою Hyper Text Markup Language (далі – HTML-документ) з використанням Hyper Text Transfer Protocol (далі – HTTP-протокол). Взаємодія двох програм за HTTP-протоколом забезпечує доступ до

HTML-документа шляхом передачі вмісту відповідного файлу. Абстрагуючись від несуттєвих особливостей, алгоритм роботи однієї з взаємодіючих програм передбачає здійснення запиту на одержання файлу HTML-документа (далі – HTTP-клієнт), інша програма – HTTP-сервер – здійснює його відправку (за запитом) зі своєї файлової системи (рис. 1). Головною особливістю мови HTML-документа є можливість використання, так званих, гіперпосилань на інші HTML-документи за місцем розташування та іменами відповідних їх файлів. Якщо файли, на які вказують гіперпосилання, знаходяться у тій самій файловій системі, то говорять про його локальне розташування відносно файлу початкового HTML-документа з іменем index.html, якщо у файловій системі іншого комп'ютера – зовнішнє. Унікальне та однозначне іменування файлів (у файловій системі) HTML-документів здійснюється у формі Uniform Resource Locator (далі – URL-адреса) завдяки включення доменного імені або мережевої адреси розташування комп'ютера.

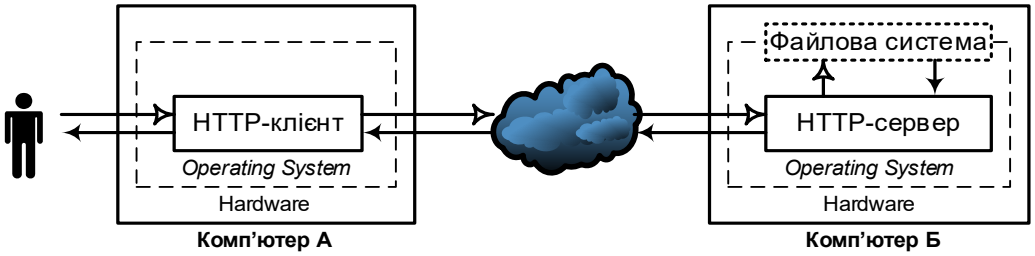


Рис. 1. Основні компоненти взаємодії комп'ютерів за Hyper Text Transfer Protocol

Web-застосування використовують набір серверних сценаріїв (Java, C#, Ruby, PHP тощо) і клієнтських сценаріїв (HTML, JavaScript тощо) для виконання програми. Робота вебдодатка залежить від його архітектури, яка включає апаратне та програмне забезпечення, яке виконує такі завдання, як читання запиту, а також пошук, збір і відображення необхідних даних. Архітектура включає різні пристрої, Web-браузери та зовнішні служби. Розрізняють архітектурні рівні клієнта або презентації, бізнес-логіки, бази даних. Базуючись на мовах програмування, які підтримуються браузером, наприклад, JavaScript, HTML і CSS, у поєднанні з іншими мовами програмування, такими як SQL, для доступу до даних із баз даних (рис. 2).

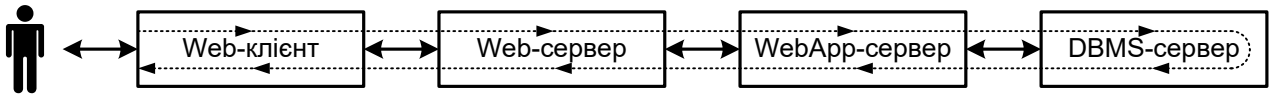


Рис. 2. Комплекс компонентів реалізації Web-застосування

Web-сервіс – це програма або програмне забезпечення, яке розгортається через Інтернет. Він використовує стандартний протокол обміну повідомленнями (наприклад, SOAP) для забезпечення зв'язку між додатками, розробленими на різних платформах. Наприклад, сервіси на основі Java можуть взаємодіяти з PHP-додатками. Ці вебдодатки інтегруються з SOAP, UDDI, WSDL і REST через мережу (рис. 3).

Архітектура Web-сервісу описує взаємодію між постачальником послуг, запитувачем послуг і реєстром послуг. Ця взаємодія складається з трьох операцій, а саме: публікація, пошук і зв'язування. Всі ці ролі та операції працюють разом над артефактами вебсервісу, відомими як програмні модулі (сервіси) та їхні описи. Постачальники послуг пропонують вебсервіси. Вони розгортають і публікують описи вебсервісів у реєстрі сервісів. Запитувачі знаходять ці описи в реєстрі послуг і використовують їх, щоб зв'язатися з постачальником вебсервісу і викликати реалізацію вебсервісу. В архітектурі вебсервісу розрізняють ServiceProvider (постачальник послуг); ServiceRequester (запитувач послуг); ServiceRegistry (реєстр послуг). Logic tier містить код, який зчитує дані з браузера та повертає результати. Рівень бізнес-логіки

включає функціональну логіку Web-застосування, яка реалізована за допомогою таких технологій, як .NET, Java та «проміжне програмне забезпечення». Він визначає потік даних, відповідно до якого розробник будує додаток за допомогою мов програмування. Він зберігає дані програми та інтегрує застарілі програми з найновішими функціями програми. Серверу потрібен певний протокол для доступу до запитуваних користувачем даних із його бази даних. Цей рівень містить програмне забезпечення та визначає кроки для пошуку та отримання даних. Data tier складається з сервера бази даних, який надає виробничі дані організації в структурованій формі (наприклад, MS SQL Server, сервер MySQL).

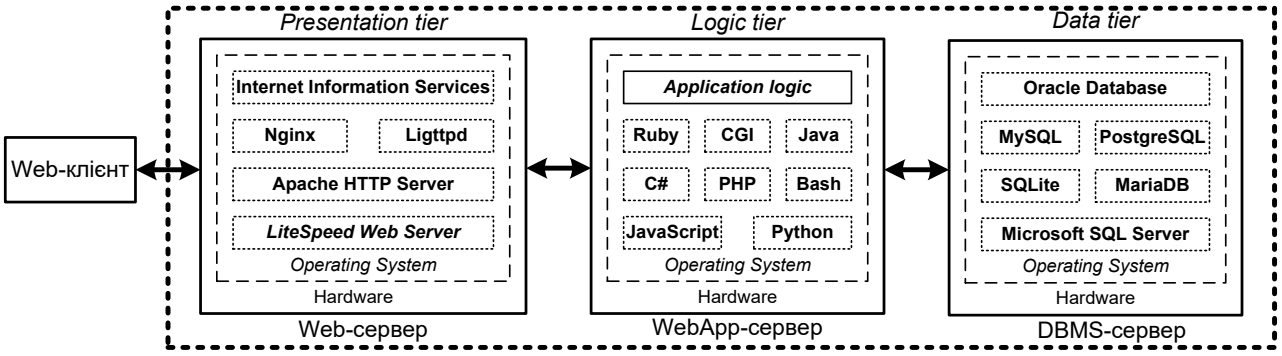


Рис. 3. Компоненти Server-Side Web Application

Існують невід'ємні ризики безпеки, пов'язані з вебсерверами, локальними мережами (LAN), у яких розміщені вебсайти, і кінцевими користувачами, які отримують доступ до цих вебсайтів за допомогою браузерів. Уразливі місця неправильно налаштованих вебсерверів можуть бути пов'язані з конфігурацією, програмами, файлами, сценаріями або вебсторінками. Зловмисник шукає такі вразливі вебсервери для здійснення атак. Неправильна конфігурація вебсервера надає зловмисникові шлях до цільової мережі організації. Неправильна конфігурація вебсервера відноситься до слабких місць конфігурації вебінфраструктури, які можна використати для запуску різних атак на вебсервери, таких як обхід каталогу, вторгнення на сервер і крадіжка даних. Адміністратори, які неправильно налаштовують вебсервери, можуть залишити серйозні лазівки у вебсервері, таким чином надаючи зловмиснику можливість використати неправильно налаштований вебсервер, щоб поставити під загрозу його безпеку та отримати конфіденційну інформацію.

Неправильна конфігурація відноситься до слабких місць конфігурації, які можна використати для запуску різних атак, таких як обхід каталогу, вторгнення на сервер і крадіжка даних. Типові недоліки конфігурації:

детальні повідомлення про налагодження/помилки;
анонімні або стандартні користувачі/паролі;
файли конфігурації та сценарії за замовчуванням;
функції віддаленого адміністрування;
увімкнено непотрібні служби;
неправильно налаштовані сертифікати;
неправильно налаштовані параметри шифрування;
сертифікати за замовчуванням;
використання самопідписаних сертифікатів;
неналежні дозволи доступу для файлів і каталогів;
увімкнено непотрібні служби,
увімкнено керування вмістом;
безпека суперечить вимогам компанії до простоти використання;
відсутність належної політики безпеки, процедур і обслуговування;

неналежна автентифікація за допомогою зовнішніх систем;

облікові записи за замовчуванням із паролями за замовчуванням або без них;

непотрібні файли за замовчуванням, резервні копії або зразки;

помилки в серверному програмному забезпеченні;

неправильно налаштована реєстрація подій;

адміністративні функції увімкнені або доступні на вебсерверах.

Оцінки категорій вразливості Server-side Web Application згідно з The Open Worldwide Application Security Project(OWASP) [11]:

CWEs Mapped – кількість типових дефектів, що мали місце для прояву вразливості;

Max Incidence Rate, Avg Incidence Rate – максимальний та середній процент існуючих Web Application-продуктів, які мають дефекти з переліку CWE для цієї категорії вразливості;

Max Coverage, Avg Coverage – максимальний та середній процент Web Application-продуктів, які мають хоча б один дефект з переліку CWE для цієї категорії вразливості відносно всіх Web Application-продуктів всіх організацій;

Total Occurrences – загальна кількість ідентифікованих Web Application-продуктів, які мають хоча б один дефект з переліку CWE для цієї категорії вразливості;

Total CVEs – загальна кількість CVE, для яких є характерним згадування хоча б одного з переліку CWE для цієї категорії вразливості

Broken Access Control. Недостатність (недосконаліть, некоректність, неправильна робота, відсутність) процедур контролю доступу може призвести до:

порушення принципу надання найменших привілеїв за замовчуванням;

обходу процедур перевірки повноважень;

несанкціонованого розширення повноважень, використання маркерів доступу;

несанкціонованого використання функцій інтерфейсу прикладного програмування (Application Programming Interface, API), посилань на об'єкти.

Оцінки Broken Access Control: CWEs Mapped = 34; Max Incidence Rate = 55,97 %; Avg Incidence Rate = 3,81 %; Max Coverage = 94,55 %; Avg Coverage = 47,72 %; Total Occurrences = 318487; Total CVEs = 19013.

Cryptographic Failures. Недостатність (недосконаліть, некоректність, неправильна робота, відсутність) процедур криптографічного перетворення може призвести до:

передбаченого, за замовчуванням, перемикання на використання незахищених комунікаційних протоколів, відмови від утворення захищеного каналу;

передбаченого, за замовчуванням, перемикання на використання менш стійких криптографічних алгоритмів, ключів, векторів ініціалізації або гам шифру тощо;

відсутності контролю дійсності ключів, цифрових підписів, програм (програмних компонентів) у репозиторіях, у ланцюгах підтвердження дійсності ідентифікаційних сертифікатів тощо;

небажаної (небезпечної, надлишкової) інформативності реакції компонентів при обробці некоректних (неправильних) комбінацій зашифрованих значень.

Оцінки Cryptographic Failures: CWEs Mapped = 29; Max Incidence Rate = 46,44 %; Avg Incidence Rate = 4,49 %; Max Coverage = 79,33 %; Avg Coverage = 34,85 %; Total Occurrences,= 233788; Total CVEs = 3075.

Injection. Недостатність (недосконаліть, некоректність, неправильна робота, відсутність) процедур контролю коректності вхідних даних може призвести до:

обробки небезпечних (некоректних) вхідних даних;

некоректного співставлення та інтерпретації об'єктів;

небажаної (небезпечної, надлишкової) інформативності реакції компонентів при обробці некоректних (неправильних) запитів до системи керування базою даних.

Оцінки Injection: CWEs Mapped = 33; Max Incidence Rate = 19,09 %; Avg Incidence Rate = 3,37 %; Max Coverage = 94,04 %; Avg Coverage = 47,90 %; Total Occurrences = 274228; Total CVEs = 32078.

Insecure Design. Небезпечний дизайн уособлює сукупність порушень моделі циклу безпечної розробки OWASP Software Assurance Maturity Model (SAMM), яка базується на здійсненні систематичного аналізу та усуненні загроз на всіх етапах розробки компонентів на основі використання вже апробованих безпечних з позиції безпеки елементів.

Оцінки Insecure Design: CWEs Mapped = 40; Max Incidence Rate = 29,19 %; Avg Incidence Rate = 3,00 %; Max Coverage = 77,25 %; Avg Coverage = 42,51 %; Total Occurrences = 262407; Total CVEs = 2691.

Security Misconfiguration. Недостатність (недосконаліть, некоректність, помилкове полягання на параметри за замовчуванням) технологічних процедур таких аспектів організації безпеки, як:

використання хмарних сервісів;

запобігання роботі компонентів та доступності облікових записів, які не використовуються, у тому числі тих, що інстальовані (створені) за замовчуванням;

запобігання (ускладнення) одержанню суб'єктом атаки відомостей про архітектуру server-side на основі аналізу відповідей за запити, у тому числі некоректні, недоречні тощо;

контроль та запобігання утворенню незахищених каналів комунікації, застосуванню менш стійких режимів захищених комунікаційних протоколів;

контроль використання актуальних версій та оновлень безпеки програмних компонентів;

систематичний аналіз ризиків.

Оцінки Security Misconfiguration: CWEs Mapped = 20; Max Incidence Rate = 19,84 %; Avg Incidence Rate = 4,51 %; Max Coverage = 89,58 %; Avg Coverage = 44,84 %; Total Occurrences = 208387; Total CVEs = 789.

Vulnerable and Outdated Components. Недостатність (недосконаліть, помилковість) організації технологічних процедур систематичного пошуку вразливостей, контролю використання застарілих (неактуальних) версій, систематичного оновлення реалізації програмних компонентів Web-серверів, WebApp-серверів, DBMS-серверів (операційних систем, системних бібліотек функцій, інтерпретаторів мов програмування тощо).

Оцінки *Vulnerable and Outdated Components*: CWEs Mapped = 3; Max Incidence Rate = 27,96 %; Avg Incidence Rate = 8,77 %; Max Coverage = 51,78 %; Avg Coverage = 22,47 %; Total Occurrences = 30457; Total CVEs = 0.

Identification and Authentication Failures. Недостатність (недосконаліть, некоректність, помилкова впевненість в параметрах за замовчуванням) технологічних процедур ідентифікації та автентифікації може призвести до набуття суб'єктом атаки повноважень в системі на основі, а саме:

підбору автентифікаційного параметрів (одноразових паролів, маркерів, ідентифікаторів тощо) в процесі віддаленої комунікації між обчислювальними процесами Web-клієнта та компонентами Web-сервера, WebApp-сервера, DBMS-сервера, а також в процесі внутрішньої комунікації між ними;

підбору нестійких до угадування паролів користувачів (внаслідок відсутності відповідного контролю при створенні), у тому числі двофакторної автентифікації (через нестійкість до угадування та/або прогнозування можливих значень);

компрометації значення паролю через критичне порушення логіки (при проєктуванні та створенні) процесу відновлення пароля, у тому числі можлива відмова від його використання при певних умовах;

витоку незашифрованих значень паролів або зашифрованих з використанням нестійких алгоритмів хешування паролів;

повторного використання одноразових паролів (маркерів, ідентифікаторів тощо) через критичне порушення логіки (при проєктуванні та створенні) процесу їх видання.

Оцінки Identification and Authentication Failures: CWEs Mapped = 22; Max Incidence Rate = 14,84 %; Avg Incidence Rate = 2,55 %; Max Coverage = 79,51 %; Avg Coverage = 45,72 %; Total Occurrences = 132195; Total CVEs = 3897.

Software and Data Integrity Failures. Недостатність (недосконаліть, некоректність, помилкова впевненість в параметрах за замовчуванням, відсутність) технологічних процедур організації контролю автентичності, цілісності, безпечності даних (включаючи у складі серіалізованих об'єктів, конвеєрів обробки даних) програмного забезпечення, які надходять з репозиторіїв (постачаються на зовнішніх носіях) для інсталяції (оновлення) компонентів Web-сервера, WebApp-сервера, DBMS-сервера.

Оцінки Software and Data Integrity Failures: CWEs Mapped = 10; Max Incidence Rate = 16,67 %; Avg Incidence Rate = 2,65 %; Max Coverage = 75,04 %; Avg Coverage = 45,35 %; Total Occurrences = 47972; Total CVEs = 1152.

Security Logging and Monitoring Failures. Недостатність (недосконаліть, некоректність, помилкова впевненість в параметрах за замовчуванням) технологічних процедур реєстрації, обліку та систематичного аналізу повідомлень про події функціонування Web-сервера, WebApp-сервера, DBMS-сервера, а також введення архівного (резервного) зберігання журнальних файлів. Найбільш суттєві аспекти прояву негативного впливу цієї категорії дефектів:

низька інформативність (прагматична цінність) повідомлень про події через неузгодженість їх форматів (синтаксис, семантика повідомлень) від різнотипних джерел;

недоступність резервних або архівних копій журнальних файлів для аналізу у разі їх цілеспрямованої модифікації (пошкодження, фальсифікація, знищення) суб'єктом атаки;

неможливість застування ретроспективного аналізу архівних журнальних файлів при розслідуванні кіберінцидентів;

потрапляння критичної технологічної інформації, значень параметрів ідентифікації, автентифікації тощо до вмісту повідомлень про події, що є передумовою для її неконтрольованого розповсюдження та витоку.

Оцінки Security Logging and Monitoring Failures: CWEs Mapped = 4; Max Incidence Rate = 19,23 %; Avg Incidence Rate = 6,51 %; Max Coverage = 53,67 %; Avg Coverage = 39,97 %; Total Occurrences = 53615; Total CVEs = 242.

Server-Side Request Forgery. Недостатність (недосконаліть, некоректність, неправильна робота, відсутність) процедур контролю небезпечного (некоректного) вмісту запитів Web-сервера до інших компонентів Server-side Web Application, які формуються на основі запитів Web-клієнта до Web-сервера, а також систематичного аналізу та усунення відповідних дефектів такого класу.

Оцінки Server-Side Request Forgery: CWEs Mapped = 1; Max Incidence Rate = 2,72 %; Avg Incidence Rate = 2,72 %; Max Coverage = 67,72 %; Avg Coverage = 67,72 %; Total Occurrences = 9503; Total CVEs = 387.

Injection. На третій позиції з 94 % визначених продуктів та 274 тис. випадків. Серед найпомітніших загальних вразливостей CWE-79, CWE-89, CWE-73.

Ознаки вразливості до атаки:

дані, що надаються користувачем, не перевіряються, не фільтруються та не очищуються програмою;

динамічні запити або непараметризовані виклики без контекстно-залежного екранування використовуються безпосередньо в інтерпретаторі;

вхідні дані використовуються у параметрах пошуку об'єктно-реляційного відображення для вилучення додаткових, чутливих записів;

вхідні дані використовуються безпосередньо або конкатенуються.

До цього класу атак відноситься як традиційний SQL, так і NoSQL, нав'язування команд операційній системі, об'єктно-реляційне відображення. Концепція ідентична для всіх інтерпретаторів. Аналіз вихідного коду є найкращим методом виявлення вразливості. Наполегливо рекомендується автоматизоване тестування всіх параметрів, заголовків, URL, файлів cookie, вхідних даних.

Щоб запобігти ін'єкціям, потрібно зберігати дані окремо від команд і запитів. Кращим варіантом є використання безпечного API, який повністю уникає використання інтерпретатора, надає параметризований інтерфейс або мігрує до інструментів відображення об'єктно-реляційних зв'язків. Навіть параметризовані збережені процедури можуть спричинити SQL-ін'єкцію. Для будь-яких залишкових динамічних запитів екрануйте спеціальні символи, використовуючи спеціальний синтаксис екранування для цього інтерпретатора. Структури SQL, такі як назви таблиць, стовпців тощо, не можна екранувати, тому назви структур, введені користувачем, є небезпечними. Це поширена проблема у програмах для написання звітів.

Доступність початкового тексту є дуже важливим фактором захисту та запобігання дефектам ін'єкції. Лише менша частина всіх програм компанії/підприємстві розробляється власними силами, а більшість додатків – із зовнішніх джерел. Програми з відкритим кодом дають принаймні можливість виправити проблеми, але програми із закритим кодом потребують іншого підходу до недоліків впровадження.

Недоліки впровадження виникають, коли програма надсилає ненадійні дані інтерпретатору. Недоліки ін'єкцій дуже поширені, особливо в застарілому коді. Вади легко виявити під час вивчення коду, але важче під час тестування. Сканери та фазери можуть допомогти зловмисникам їх знайти.

Залежно від доступності потрібно виконати різні дії, щоб їх виправити. Це завжди найкращий спосіб вирішити проблему в самому вихідному коді або навіть змінити дизайн деяких частин програми. Але якщо вихідний код недоступний або просто неекономно виправляти застаріле програмне забезпечення, лише віртуальне виправлення має сенс.

Найвідомішою формою впровадження є SQL-ін'єкція, коли зловмисник може змінювати існуючі запити до бази даних. Щоб отримати додаткові відомості, перегляньте шпаргалку щодо запобігання впровадження SQL. Атака SQL-ін'єкції полягає у вставці або «ін'єкції» часткового або повного SQL-запиту через дані, що вводяться або передаються від клієнта (браузера) до вебдодатка.

Успішна атака SQL-ін'єкцій може зчитувати конфіденційні дані з бази даних, змінювати дані бази даних (вставити/оновити/видалити), виконати операції адміністрування бази даних, відновити вміст певного файлу, записувати файли у файлову систему, а в деяких випадках видавати команди операційній системі.

Атаки SQL Injection можна розділити на такі три класи:

1. Дані витягуються за допомогою того самого каналу, який використовується для впровадження коду SQL. Це найпростіший вид атаки, у якому отримані дані представлено безпосередньо на вебсторінці програми.

2. Дані отримуються за допомогою іншого каналу (наприклад, електронний лист із результатами запиту створюється та надсилається тестувальнику).

3. Фактичної передачі даних немає, але тестер може реконструювати інформацію, надсилаючи певні запити та спостерігаючи за результатом поведінки сервера бази даних

Усі мови сценаріїв, що використовуються у вебдодатках, мають форму виклику eval, який отримує код під час виконання та виконує його. Якщо код створено з використанням неперевірених і неекранованих введених користувачем кодів, може статися ін'єкція коду, що дозволить зловмиснику підірвати логіку програми та зрештою отримати локальний доступ.

Якщо мова сценаріїв має недолік у коді обробки даних, вектори атаки «введення нульового байта» можуть розгортатись для отримання доступу до інших областей пам’яті, що призводить до успішної атаки. Впровадження команд операційної системи – це техніка, яка використовується через інтерфейс для виконання команд. Маючи можливість виконувати команди ОС, користувач може завантажувати шкідливі програми або навіть отримувати паролі. Впровадженню команд ОС можна запобігти, якщо під час проєктування та розробки програм наголошується на безпеці.

Дорожня карта заходів запобігання прояву вразливості Server-side Web Application:
виправляти та оновлювати програмне забезпечення;
застосовувати всі оновлення, незалежно від їхнього типу, за потребою;
тестувати пакети оновлень перед розгортанням;
здійснювати резервне копіювання;
розробити план повернення до початкового стану після невдалого оновлення;
планувати періодичні оновлення пакетів;
вимкнути невикористані зіставлення розширень сценаріїв;
усунути використання конфігурацій за замовчуванням;
планувати аварійного відновлення для вирішення проблем керування виправленнями;
блокуйте непотрібні порти;
блокувати непотрібні протоколи;
використовувати безпечну автентифікацію;
використовувати тунелювання та шифрування каналів комунікації;
використовувати безпечні протоколи для зв’язку з вебсервером;
ізолювати каталоги сервісів;
видалити непотрібні модулі та розширення програм;
вимкнути облікові записи користувачів за замовчуванням;
видалити користувачів бази даних і збережені процедури;
дотримуватися принципу найменших привілеїв для бази даних;
використовувати безпечні дозволи;
використовувати політики надійних паролів;
здійснювати перевірки стійкості паролів;
реєструвати помилки автентифкації;
надавати процесам найменші привілеї;
створювати мінімальну кількість привілейованих облікових записів;
обмежити доступ адміністратора;
здійснювати реєстрацію подій у зашифрованому вигляді;
вимкнути усі неінтерактивні облікові записи;
вилучити конфіденційну конфігураційну інформацію з байт-коду;
вилучити зіставлення віртуальних каталогів з різними серверами або через мережу;
здійснювати аналіз журналів обліку подій;.
вимкнути обслуговування списків каталогів;
вилучити застосування непотрібних файлів і нецільових розширень;
вимкнути обслуговування певних типів файлів;
забезпечити ізольоване зберігання файлів;
використання ізольованого програмного середовища;
уникайте посилань на всі типи невебфайлів в URL-адресі.

Висновки. Впровадженню SQL найкраще запобігти за допомогою параметризованих запитів. Кращим варіантом є використання безпечного API, який повністю уникає використання інтерпретатора, надає параметризований інтерфейс або мігрує до інструментів

відображення об'єктно-реляційних зв'язків. Зверніть увагу, що багато фреймворків і бібліотек на стороні клієнта пропонують параметризацію запитів на стороні клієнта. Ці бібліотеки часто просто створюють запити з конкатенацією рядків перед надсиланням необроблених запитів на сервер.

Подальші дослідження передбачають розробку розширеної дорожньої карти заходів, необхідної для запобігання дії додаткових вразливостей Server-side Web Application, які неодмінно будуть виявлені в майбутньому.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Про основні засади забезпечення кібербезпеки України: Закон України від 05.10.2017 № 2163-VIII. *Офіційний вісник України.* 2017. № 91. С. 91.

2. Про захист інформації в інформаційно-комунікаційних системах: Закон України від 05.07.1994 № 80/94-ВР. *Відомості Верховної Ради України.* 1994. № 31. Ст. 286.

3. Про затвердження Загальних вимог до кіберзахисту об'єктів критичної інфраструктури: постанова Каб. Міністрів України від 19.06.2019 № 518. *Офіційний вісник України.* 2019. 05 лип. (№ 50). С. 53.

4. Про затвердження Порядку пошуку та виявлення потенційної вразливості інформаційних (автоматизованих), електронних комунікаційних, інформаційно-комунікаційних систем, електронних комунікаційних мереж: постанова Каб. Міністрів України від 16.05.2023 № 497. *Офіційний вісник України.* 2023. 02 черв. (№ 52). С. 72.

5. Про затвердження Методичних рекомендацій щодо реагування суб'єктами забезпечення кібербезпеки на різні види подій у кіберпросторі: наказ Адміністрації Державної служби спеціального зв'язку та захисту інформації України від 03.072023 № 570. *GOV.ua.* URL: https://cip.gov.ua/ua/news/nakaz-administraciyi-derzhspeczv-yazku-vid-03-07-2023-570-pro-zatverdzhennya-metodichnikh-rekomendacii-shodo-reaguvannya-sub-yektami-zabezpechennya-kiberbezpeki-na-rizni-vidi-podii-u-kiberprostori.

6. Adversarial Tactics, Techniques & Common Knowledge. URL: https://attack.mitre.org.

7. Common Attack Pattern Enumerations and Classifications. URL: https://capec.mitre.org.

8. Common Vulnerabilities and Exposures. URL: https://cve.mitre.org.

9. Common Weakness Enumeration. URL: https://cwe.mitre.org.

10. Open Worldwide Application Security Project. URL: https://owasp.org.

DOI: 10.58254/viti.5.2024.19.215

УДК 004.056.5

д-р техн. наук Чевардін В. Є. ORCID: 0000-0002-1070-4568 (ВІТІ ім. Героїв Крут)
Лаврик І. В. ORCID: 0000-0002-3433-9083 (ВІТІ ім. Героїв Крут)

КРИПТОСИСТЕМИ НА ОСНОВІ ІЗОМОРФНИХ ПЕРЕТВОРЕНЬ ЕЛІПТИЧНИХ КРИВИХ

У статті розглянуто напрями розробки та вдосконалення постквантових криптографічних систем, заснованих на ізоморфних перетвореннях еліптичних кривих, потенційно стійких до квантового криптоаналізу. Проведено аналіз недоліків та переваг існуючих асиметричних криптосистем, в тому числі таких, які побудовані на основі ізоморфних перетворень. Досліджено підходи до побудови криптографічних алгоритмів на основі ізогеній еліптичних кривих, що можуть стати основою для створення стійких до квантових атак криптосистем.

У процесі проведених досліджень було розроблено програмні функції для реалізації операцій над ізогеніями еліптичних кривих різного порядку, які забезпечать зазначені в стандарті [12] рівні безпеки: 256, 384, 512. Розроблена програмна реалізація операцій скалярного множення точки кривої та операцій над ізогеніями еліптичної кривої, на основі якої отримано експериментальні значення часу на обчислення скалярного добутку з використанням розпаралелювання. Проведено експерименти порівняння з класичного множення точки кривої з представленням скаляру k у вигляді послідовності 4-бітових слів, що дозволило прискорити операцію скалярного множення в 30 разів, для 8-бітових слів прискорення склало 18,8 разів.

Напрямком подальших досліджень є розробка методів генерації та верифікації цифрового підпису, на основі перетворень над точками ізогенії еліптичної кривої з використанням розпаралелювання операцій скалярного множення точки кривої.

Ключові слова: *асиметричні криптосистеми, еліптична крива, ізоморфні перетворення еліптичної кривої, ізогенія еліптичної кривої.*

V. Chevardin, I. Lavryk Cryptosystems based on isomorphic transformations of elliptic curve points.

The article presents research in the field of development and improvement of cryptographic systems based on elliptic curves isomorphic transformations potentially resistant to quantum cryptanalysis. Analysis results of existing asymmetric cryptosystems disadvantages and advantages, including those based on isomorphic transformations, are presented. The approaches to the construction of cryptographic algorithms based on isogenies of elliptic curves, which can become the basis for creating cryptosystems resistant to quantum attacks, are investigated.

In the course of the research, program functions were developed to implement operations on elliptic curves isogenies of different orders, which will ensure the security levels specified in the standard: 256, 384, 512. A software implementation of the operations of a curve point scalar multiplication and operations on elliptic curve isogenies has been developed, on the basis of which experimental values of the time to perform the scalar product using parallelization have been obtained. Experiments have been conducted to compare the classical multiplication of a curve point with the representation of the scalar k as a sequence of 4-bit words, which made it possible to speed up the scalar multiplication operation by 30 times, for 8-bit words the speedup was 18.8 times.

The direction of further research is the development of methods for the generation and verification of a digital signature, based on transformations over the isogeny points of the elliptic curve using the parallelization of operations of scalar multiplication of the curve point.

Keywords: *asymmetric cryptosystems, elliptic curve, isomorphic transformations of an elliptic curve, isogeny of an elliptic curve.*

Постановка проблеми та актуальність дослідження

Одним із найважливіших досягнень цього століття стала поява квантових комп'ютерів, що викликало реальну потребу оцінки стійкості існуючих асиметричних криптопримітивів, які використовуються в системах електронно-цифрового підпису, системах та комплексах автентифікації та розмежування доступу, генерації загальних секретних ключів для шифрування та автентифікації даних. У більшості існуючих та широко розповсюджених алгоритмів використовують стандартизовані перетворення в групі точок еліптичної кривої, які визначені стандартами ДСТУ 4145-2002, IEEE P.1363, AIS 2.0. Основні напрямки досліджень розділилися. Частина робіт була спрямована на підвищення стійкості класичних асиметричних криптосистем, що базуються на складності рішення задач DLP, ECDLP та інших

еквівалентних ним [3]. Інша спрямована на розробку нових криптопримітивів, стійких до квантового криптоаналізу [3; 4]. Враховуючи, що розробка принципово нових криптопримітивів викликає потребу в додаткових витратах на впровадження в інформаційні системи, реалізацію нових програмно-апаратних засобів та систем, актуальною науково-технічною задачею є вдосконалення існуючих криптоперетворень з підвищеними показниками стійкості до квантового криптоаналізу.

Аналіз останніх публікацій та наукових результатів

Початком розвитку систем асиметричного шифрування, цифрового підпису, інкапсуляції криптографічних ключів тощо стала поява однонаправлених функцій, для яких задача обчислення прообразу потребує надзвичайно великого обсягу обчислювальних витрат. Першим прикладом таких задач стали системи Діффі – Геллмана, RSA та подібні їм [1]. Наприклад, криптосистема RSA базується на складності задачі факторизації. Знаходження ефективних алгоритмів факторизації цілого числа N призвели до потреби постійно збільшувати його бітову довжину.

З часом з'явилася альтернатива цим алгоритмам. Замість RSA-подібних перетворень стали застосовувати перетворення, які базуються на еліптичних кривих, що розглянуті в роботах [5–9]. Для побудови криптосистем на еліптичних кривих використовують нормальну форму (форму Веєрштрасса [5]) чи канонічну форму кривої з параметром $q \neq 2,3$, $(q \neq 2^m)$ у спрощених видах (криві Монтгомері, криві Коблиця, криві Едвардса та інші). Перехід від RSA-подібних систем до систем на еліптичних кривих дозволив отримати можливість зниження довжини ключа криптосистеми з 1024 біт до 160 біт без зниження криптостійкості системи. Найбільш цікавими результатами вдосконалення або розробки нових перетворень на еліптичних кривих є використання ізоморфних трансформацій нормальної форми кривої до інших скорочених форм [10–12]. Зазначений підхід надає можливість спростити число примітивних операцій для виконання скалярного множення точок і підвищити швидкодію криптоперетворень. З іншого боку, перехід накладає додаткові обмеження на параметри кривої та умови виконання операції скалярного множення. Наприклад, для кривих Едвардса [6] обираються криві, порядок яких не є простим числом і є кратним 4, що вважається уразливістю для криптографічно стійких кривих. Перехід до еліптичних кривих у формі Едвардса дозволяє в середньому підвищити швидкодію операцій скалярного множення в два рази, але не забезпечує стійкість до квантового криптоаналізу.

Наведемо деякі положення з теорії еліптичних кривих.

Криві Веєрштрасса

Крива (1) називається кривою Веєрштрасса або нормальною формою еліптичної кривої над скінченним полем F_q.

$$y^2 + a_1xy + a_2y = a_3x^3 + a_4x^2 + a_5x + a_6,\ a_i \subset F_p. \quad (1)$$

Гладкою (неособливою) еліптичною кривою порядку n над полем F_q називається множина точок, які задовольняють рівнянню (1), де многочлен ступеня[1] m з коефіцієнтами $A \subset F_p$, $B \subset F_p$, де $q = p$ – просте число. Для криптографічних цілей для еліптичної кривої виконуються вимоги: дискримінант $\Delta(E) \neq 0$, j-інваріант $j \neq \{0,12^3\}$.

Представлення кривої (1) над полем характеристики $q = p$ з фіксованими параметрами $a_1 = 0$, $a_2 = 0$, $a_3 = 1$, $a_4 = 0$ дає канонічну форму кривої (скорочену форму Веєрштрасса):

$$y^2 = x^3 + Ax + B. \quad (2)$$

[1] Ступінь багаточлена є максимальним ступенем одночленів, з яких він складається.

Крива (1) над полем характеристики $q = 2^m$ трансформується до форми (3):

$$y^2 + xy = x^3 + Ax^2 + B \quad (3)$$

Криві (2, 3) є більш розповсюдженими для застосування в криптографічних додатках та системах електронного цифрового підпису, а саме в системі ЕЦП на основі ДСТУ 4145-2002, в протоколі генерації спільного секрету Діффі – Геллмана IEEE 1363–2000 та інших стандартизованих криптопротоколах стандарту 1363а-2004. Від основної форми Веєрштрасса (1) можемо перейти до інших форм еліптичних кривих завдяки ізоморфним трансформаціям. Найбільш популярними з відомих трансформацій кривої (1) стали криві Монтгомері та Едвардса [6; 7], які дозволили зменшити обчислювальну складність скалярного множення точок кривої. Так, криві Едвардса дозволили прискорити скалярне множення точки кривої в два рази.

Скалярним множенням точки кривої є k-кратне додавання точки кривої, $\boldsymbol{k * P = P + P + \ldots + P}$. Скалярне множення є складно оберненою операцією, для якої зворотним перетворенням є дискретне логарифмування в групі точок еліптичної кривої. Враховуючи математичну конструкцію скалярного множення точки кривої, існує два основних напрямки вдосконалення криптоперетворень у групі точок еліптичних кривих, а саме підходи, спрямовані на зменшення обчислювальної складності, та підходи, спрямовані на підвищення криптографічної стійкості. Розглянемо переваги перетворень у групі точок еліптичної кривої.

1. Використовуючи ізоморфізм $E_p \equiv G^*{}_{p'}$, кожна точка $Q \in E_p$ може бути представлена як цілочисельний елемент поля $c \in F_{p'}$. Це дозволяє переходити від операцій над точками кривих до операцій над цілими числами в простому полі. Зазначена властивість викликала вимогу до перебільшення бітової довжини поля p в 30 разів над бітовою довжиною p'. Це дозволило перейти від перетворень у скінченному полі до перетворень у групі точок еліптичної кривої.

2. Використання в асиметричних алгоритмах перетворень над точками еліптичної кривої дозволяє змінювати параметри криптосистеми шляхом лише зміни коефіцієнтів A та B з фіксованим значенням p. Так, еліптичні криві з різними коефіцієнтами, але з однаковим значенням (4) вважаються ізоморфними і дозволяють змінювати криві зі сталим значенням характеристики поля E_p.

$$j(E) = \frac{1728(4A^3)}{4A^3 + 27B^2} \neq 0 \cdot \quad (4)$$

3. Перетворення в групі точок еліптичної кривої дозволяє отримати можливість розпаралелювання найбільш обчислювально складної операції скалярного множення точки кривої на потоки, що надає можливість збільшити швидкість множення точок кривої на скаляри [8; 9]. Так, якщо зафіксуємо еліптичну криву E з коефіцієнтами A та B над скінченним полем $F(p)$, то для довільної точки кривої $P \in E_{A,B,p}$ можна виконати скалярне множення точки кривої, а саме $Q = k * P$. Нехай існує умовний l_k-розрядний регістр, що реалізує скалярне множення точки P на число k, довжина якого дорівнює l_k. Якщо представити значення k у вигляді суми цілих чисел довжиною l_{ki}, а саме $k = k_1 + k_2 + \ldots + k_n = \sum_{i=0}^{i=n-1} k_i$, де $n = l_k / l_{ki}$, тоді замість множення $k * P$ з l_k -розрядним регістром виникає можливість виконати n множень з використанням l_{ki}-розрядних регістрів. Результат обчислення точки Q в такому випадку матиме вигляд: $Q = k * P = \sum_{i=0}^{i=n-1} k_i * P$. У такому випадку реалізація алгоритму скалярного множення у вигляді паралельних потоків, які будуть виконувати операції $k_i * P$, дозволить потенційно прискорити процес множення в n разів.

Розглянемо недоліки або обмеження перетворень у групі точок еліптичної кривої.

1. Для генерації параметрів еліптичної кривої необхідно обрати такі значення A, B та p, які забезпечать вимоги до порядку еліптичної кривої. Визначення порядку кривої потребує обчислювальних витрат, які викликані складністю алгоритму Schoof-Atkin-Elkies (SEA) [8; 9].

2. Можливість розпаралелювання процесу скалярного множення точок кривої викликає підвищення апаратно-програмної складності реалізації скалярного множення. З іншого боку, можливість розпаралелювання операції скалярного множення з появою квантових комп'ютерів створило загрозу швидкого зламу існуючих асиметричних криптосистем з використанням алгоритму Шора, а для симетричних з використанням алгоритму Гровера [4].

3. Здійснення MOV-атаки [12] завдяки зведенню операцій в групі точок еліптичної кривої до операцій в скінченному полі цілих чисел. Так, для забезпечення рівня стійкості 128 біт для алгоритмів на основі еліптичних кривих ключ сьогодні повинен мати довжину 240 біт, а для алгоритмів RSA – 2800 біт.

4. Здійснення атаки на основі аналізу потужності, яка описана в [13]. Як правило, цей тип атаки називають Simple Power Analysis (SPA). Він використовується також для більшості асиметричних криптосистем, захист від яких включає, як правило, організаційно-технічні аспекти.

Враховуючи переваги та недоліки перетворень у групі точок еліптичної кривої, одним з перспективних напрямків вважається використання ізоморфних трансформацій, які дозволяють отримати зменшення кількості примітивних операцій для скалярного добутку точки кривої та генерації еліптичної кривої з однієї сторони, а з іншої, збільшити криптографічну стійкість криптоперетворень у групі точок еліптичної кривої завдяки збільшенню кількості варіантів шифрованих текстів. Використання множини ізоморфних кривих [7–8] дозволило відкрити новий шлях збільшення стійкості криптоперетворень, у тому числі і до квантового криптоаналізу.

Так, під час генерації еліптичної кривої генеруються параметри ізоморфної трансформації кривої, що дозволяє покращити показники стійкості алгоритмів генерації ПВП до відтворення та передбачення. Використання ізоморфних трансформацій еліптичної кривої надало змогу збільшити число внутрішніх станів генератора без суттєвих втрат часу та продовжити використання генераторів ПВП цього класу до переходу на постквантову криптографію.

Найбільш ефективним застосуванням ізоморфних трансформацій еліптичної кривої сьогодні є побудова криптоалгоритмів на основі ізогенії еліптичної кривої, яка є ядром ізоморфної трансформації кривої. Це дозволяє збільшити множину як сеансових, так і спільних ключів в алгоритмі Діффі – Геллмана при фіксованому розмірі характеристики скінченного поля. Сьогодні схема обміну ключів Діффі – Геллмана вже запропонована з використанням ізогенії еліптичної кривої, а саме схема Supersingular Isogeny Key Exchange (SIKE). Порівняно з іншим постквантовим кандидатом NTRUEncrypt з бітовою довжиною ключа 600 байтів, SIKE дозволяє використовувати 330-байтові ключі для однакового рівня безпеки 128 біт. Це робить SIKE більш привабливим для реалізації в системах електронного документообігу, мережах Bitcoin, Tor та інших. Існуючі програмні реалізації SIKE запропоновано з параметрами: SIKEp182, SIKEp217, SIKEp503, SIKEp610 і SIKEp751. Однак, в роботі [20] представлено алгоритм відтворення ключа цього алгоритму за реальний час, а саме для SIKEp503 – 20 хв, SIKEp610 – 55 хв та SIKEp751 – 3 год 15 хв. Це створило умови для виключення цього алгоритму зі списків постквантових кандидатів з метою опрацювання і вдосконалення. У зв'язку з цим, виникла потреба у пошуку шляхів для вдосконалення криптографічних схем та перетворень на основі ізоморфних трансформацій еліптичної кривої, що викликало необхідність цього дослідження.

Метою роботи є визначення шляхів удосконалення сучасних криптосистем, представлених в якості постквантових кандидатів, побудованих з використанням операцій над ізогеніями еліптичних кривих.

Викладення основного матеріалу

Для досягнення поставленої мети необхідно оцінити можливості щодо зменшення обчислювальних витрат на виконання криптографічних операцій та здатності зазначених операцій до збільшення криптографічної стійкості. Для цього проведемо аналіз перетворень в групі та підгрупі точок кривої, наведемо відомий підхід прискорення операцій завдяки розпаралелюванню скалярного множення точки кривої, побудуємо операції над ізогеніями еліптичної кривої, оцінимо потужність простору ізогеній кривої з метою використання параметру ізоморфної трансформації для збільшення криптостійкості.

Група та підгрупи точок еліптичної кривої

Позначимо скінченне поле F характеристики q як F_q [5], порядок якого $q = ord(F)$. Тоді якщо $q = p$, де p – просте число, скінченне поле назвемо простим полем F_p.

Позначимо еліптичну криву E з коефіцієнтами A та B визначену виразом (2) над полем F_p як $E_{p,A,B}$. Крива $E_{p,A,B}$ є скінченною множиною точок $P_i \in E_{p,A,B}$, для яких можемо побудувати операцію додавання (подвоєння) точок кривої, деталі отримання виразів додавання та добутку точок кривої, обрання параметрів еліптичної кривої наведено в роботах [5; 6; 9].

Приклад. Зафіксуємо криву $E_{11,3,4}: y^2 = x^3 + 3x + 4$. Так, як порядок кривої $ord(E_{11,3,4}) = 14$, то група точок складається з підгруп простого порядку 2 та 7. Позначимо підгрупу порядку 7 як subE1, а підгрупу порядку 2 як subE2. Наведемо точки, що входять до цих підгруп з урахуванням нейтрального елементу точки O.

Група subE1: {(0,2);(0,9);(4,5);(8,10);(8,1);(4,6);O}, з порядком підгрупи $ord(subE) = 7$,

Група subE2: {(10,0);O}, з порядком підгрупи $ord(subE) = 2$.

Порядок точок кривої $E_{11,3,4}$ наведено в таблиці 1.

Таблиця 1

Порядок точок кривої $E_{11,3,4}$

P_i	(5,10)	(4,6)	(7,4)	(8,10)	(9,1)	(0,9)	(10,0)	(0,2)	(9,10)	(8,1)	(7,7)	(4,5)	(5,1)	O
$ord(P_i)$	14	7	14	7	14	7	2	7	14	7	14	7	14	1

Точки порядку 14 не використовують, так як точки (4,6), (4,5), (8,10), (8,1), (0,9), (0,2) створюють циклічну підгрупу порядку 7, що зменшує кількість операцій, які повинні виконати криптоаналітику. В зазначеному прикладі точка (10,0) має порядок 2 і в якості генератора групи не використовується також. Для криптографічних додатків завжди обирають циклічні підгрупи точок кривих простого порядку.

Особливістю операцій над точками еліптичної кривої є можливість розпаралелювання операції скалярного множення точки кривої. Розглянемо сутність процесу розпаралелювання операції скалярного множення точки кривої.

Розпаралелювання скалярного множення точки кривої

Нехай, еліптична крива $E_{p,A,B}$ має скінченну множину точок $P_i \in E_{p,A,B}$. Точка $P \in E_{p,A,B}$ є базовою точкою (має великий простий порядок). Тоді операцію скалярного множення $k * P$ можна представити як $k * P = (k_1 + k_2) * P = k_1 * P + k_2 * P$ на основі асоціативності операції додавання точок кривої. Для виконання операцій $k_1 * P$ та $k_2 * P$ використовують два різних обчислювача (регістра) або ядра процесора, що суттєво прискорює

виконання операції $k * P$. Є різні підходи щодо використання розпаралелювання скалярного множення точки кривої з використанням комбінації Double-and-add та Halve-and-add алгоритмів та Montgomery-halving алгоритму, результати аналізу ефективності яких наведено в роботах [11], а саме скалярне множення точки кривої визначається виразом:

$$I = (\omega + l)I_{dbl} + (\frac{l}{\omega+1} + 2^{\omega-1})I_{add},$$

де I_{dbl} – складність операції подвоєння точки 2P кривої; I_{add} – складність операції додавання точок P+Q кривої.

Сутність прискорення скалярного множення з використанням цих алгоритмів полягає в різноманітних формах розкладання цілого числа k, що скорочує кількість операцій додавання та подвоєння точки кривої, які здійснюються під час виконання основної операції множення $k * P$. Так,

$$k = k' \times 2^{-\delta} mod N = k_1 + k_2 = (k'_l 2^{l-\delta} + \ldots +_{\square} k'_\delta) + (k'_l 2^{-1} + \ldots +_{\square} k'_0 2^{-\delta}) \text{ mod N},$$

звідки

$$k * P = (k'_l 2^{l-\delta} + \ldots +_{\square} k'_\delta) * P + (k'_l 2^{-1} + \ldots +_{\square} k'_0 2^{-\delta}) * P$$

скалярний добуток можна реалізувати δ паралельними процесами (потоками).

Розглянемо модель реалізації скалярного добутку з використанням розпаралелювання скалярного множення точки еліптичної кривої.

Нехай для скалярного множення обрано скаляр k, який поданий у вигляді бітової строки довжиною n. Представимо значення скаляру k послідовністю ω-бітових слів α_i, довжина якої дорівнює δ, $k = \alpha_1||\alpha_2||\alpha_3||\ldots||\alpha_\delta$. В такому випадку число $\delta = \frac{n}{\omega}$.

Позначимо процес множення точки на скаляр, що виконує обчислювач (регістр) або ядро процесора, pr_i, де $1 \le i \le N_{pr}$, N_{pr} – кількість процесів, якими реалізований скалярний добуток (складність обчислювальної системи). Позначимо час виконання скалярного множення точки P кривої як t_k. З метою зменшення обчислювальної складності або часу t_k необхідно розрахувати значення параметру N_{pr} залежно від значення ω та δ. Проведемо обчислення залежності t_k від ω та δ на ПЕОМ з параметрами Processor 11th Gen Intel(R) Core(TM) i9-11900H @ 2.50GHz, 2496 Mhz, 8 Core(s), 16 Logical Processor(s) Installed Physical Memory (RAM) 40.0 GB.

Приклад наведено у таблицях 2, 3.

Таблиця 2

Параметри еліптичної кривої та базової точки для проведення експерименту

p:	10061
A:	6451
B:	1036
Px:	10056
Py:	9389
scalar 256 bit:	870342851277024034039875644609635332673029392016573708158509864187358908571 22
Qx:	3498
Qy:	6874

Таблиця 3

Параметри еліптичної кривої та базової точки для проведення експерименту

ω (біт)	t_k (сек)
4	0.0000419
8	0.0000676
16	0.0001070
32	0.0001969
64	0.0003782
128	0.0006585
256	0.0012763

За результатами експерименту встановлено, що представлення скаляру *k* послідовністю 4-бітових слів надає можливість прискорити операцію скалярного множення *k*P* в 30 разів, завдяки встановленню значення $N_{pr} = 30$.

Операції над ізогеніями еліптичної кривої

Нехай E_1 та E_2 – гладкі еліптичні криві (2) над полем *F*, які визначаються рівнянням (2) з відповідними коефіцієнтами. Кожна крива визначена значеннями: #*E*, Δ, *j*(*E*). Існування ізоморфізмів для еліптичних кривих надає можливість використовувати весь простір еквівалентних кривих для створення більш стійких модифікацій існуючих криптоалгоритмів [14–18]. Ізогенія еліптичної кривої є одним із різновидів ізоморфних трансформацій кривої в еквівалентну.

Визначення 1. Ізогенія еліптичної кривої є неконстантним раціональним відображенням кривої E_1 над скінченним полем *F* в криву E_2, яке також називається груповим гомоморфізмом та подається у вигляді:

$$\phi(x;y) \to \left(\frac{f_1(x;y)}{f_2(x;y)}, \frac{g_1(x;y)}{g_2(x;y)}\right) = \left(\frac{p_{(x)}}{q_{(x)}}, yr(x)\right),$$

де f_1, f_2, g_1, g_2 – поліноми.

Одним із найважливіших для криптографічної стійкості криптоперетворень, які будуються на основі ізогеній, є їхній ступінь, який визначає розмір множини ізоморфних трансформацій.

Визначення 2. Степінь ізогенії – є степенем раціонального відображення, що знаходиться як максимум зі степенів поліномів $p(x)$ та $q(x)$:

$$deg(\phi(x;y)) = max(deg(p(x), deg(q(x)))),$$

де *p(x)*, *q(x)* – поліноми.

Для сепарабельних ізогеній $deg(\phi(x;y)) = \#ker\phi(x;y)$. Якщо криві $E1 = E2,$ то $\phi(x;y)$ – ендоморфізм.

Теорема Tate [19]. Нехай E_1 та E_2 – криві над скінченним полем *F*. Тоді криві E_1 та E_2 є ізогенними кривими тоді і тільки тоді, коли порядки їх груп дорівнюють $\#E_1 = \#E_2$.

Приклад. Для обраної раніше кривої $E_{11,3,4}$ порядку 14 оберемо циклічну підгрупу простого порядку 7. Для обраної кривої над полем F_{11} отримаємо всі ізоморфні криві

порядку 14, як мають у своєму складі підгрупу порядку 7. Кількість таких кривих для визначених умов дорівнює кількості пар коефіцієнтів A та B, для яких порядок циклічної підгрупи однаковий. Для обраних параметрів в таблиці 4 наведено трансформації всіх кривих ізоморфних базовій кривій $E_{11,3,4}$, де n – порядок кривої над полем характеристики 11.

Таблиця 4

n	*6*	*7*	*8*	*9*	*10*	*11*	*13*	*14*	*15*	*16*	*17*	*18*
A, B	*1, 8*	*2, 7*	*1, 9*	*1, 4*	*1, 10*	*1, 5*	*1, 6*	*1, 1*	*1, 7*	*1, 2*	*2, 4*	*1, 3*
	3, 10	*6, 6*	*2, 10*	*2, 2*	*2, 5*	*2, 8*	*2, 3*	*2, 6*	*2, 9*	*2, 1*	*6, 5*	*3, 1*
	4, 2	*7, 2*	*3, 3*	*3, 5*	*3, 7*	*3, 9*	*3, 2*	*3, 4*	*3, 6*	*3, 8*	*7, 9*	*4, 9*
	5, 6	*8, 10*	*4, 5*	*4, 1*	*4, 8*	*4, 4*	*4, 7*	*4, 3*	*4, 10*	*4, 6*	*8, 1*	*5, 5*
	9, 7	*10, 8*	*5, 4*	*5, 3*	*5, 2*	*5, 1*	*5, 10*	*5, 9*	*5, 8*	*5, 7*	*10,3*	*9, 4*
			6, 7	*6, 8*	*6, 9*	*6, 10*	*6, 1*	*6, 2*	*6, 3*	*6, 4*		
			7, 6	*7, 10*	*7, 3*	*7, 7*	*7, 4*	*7, 8*	*7, 1*	*7, 5*		
			8, 8	*8, 6*	*8, 4*	*8, 2*	*8, 9*	*8, 7*	*8, 5*	*8, 3*		
			9, 1	*9 ,9*	*9, 6*	*9, 3*	*9, 8*	*9, 5*	*9, 2*	*9, 10*		
			10, 2	*10, 7*	*10, 1*	*10, 6*	*10, 5*	*10, 10*	*10, 4*	*10, 9*		

Зафіксуємо значення порядку кривої $n = 6$ та отримаємо всі ізоморфні трансформації точок для кожної ізоморфної кривої (табл. 5).

Таблиця 5

A_i, B_i	$\{P_i\}$	$\{P_i\}$	$\{P_i\}$	$\{P_i\}$	$\{P_i\}$	$\{P_i\}$
1, 1	*(0, 1)*	*(3, 3)*	*(6, 6)*	*(6, 5)*	*(3, 8)*	*(0, 10)*
2, 6	*(1, 3)*	*(10, 6)*	*(5, 3)*	*(5, 8)*	*(10, 5)*	*(1, 8)*
3, 4	*(0, 2)*	*(4, 6)*	*(8, 1)*	*(8, 10)*	*(4, 5)*	*(0, 9)*
4, 3	*(0, 5)*	*(5, 4)*	*(10, 8)*	*(10, 3)*	*(5, 7)*	*(0, 6)*
5, 9	*(0, 3)*	*(1, 9)*	*(2, 7)*	*(2, 4)*	*(1, 2)*	*(0, 8)*
6, 2	*(3, 5)*	*(5, 6)*	*(6, 10)*	*(6, 1)*	*(5, 5)*	*(3, 6)*
7, 8	*(3, 1)*	*(8, 2)*	*(4, 1)*	*(4, 10)*	*(8, 9)*	*(3, 10)*
8, 7	*(1, 4)*	*(9, 7)*	*(2, 8)*	*(2, 3)*	*(9, 4)*	*(1, 7)*
9, 5	*(0, 4)*	*(9, 1)*	*(7, 2)*	*(7, 9)*	*(9, 10)*	*(0, 7)*
10, 10	*(4, 2)*	*(7, 4)*	*(9, 2)*	*(9, 9)*	*(7, 7)*	*(4, 9)*

Обчислимо ізогенію для ізоморфних кривих E_1: $y^2 = x^3 + x + 1$ та E_2: $y^2 = x^3 + 4x + 13$, побудованих над скінченним полем F_{19}, та перевіримо правильність ізоморфної трансформації точок кривої E_1.

Приклад. Нехай скінченне поле буде F_{19}, а криві над цим полем E_1: $y^2 = x^3 + x + 1$ та E_2: $y^2 = x^3 + 4x + 13$. Порядок кривих E_1 та E_2 дорівнює $\#E_1 = \#E_2 = 21$, інваріанти кривих дорівнюють один одному, $\mathrm{j}(E_1) = \mathrm{j}(E_2)$. Для заданої кривої була обчислена гомоморфна трансформація, а саме вирази $f1(x;y)$, $f2(x;y)$, $g1(x;y)$, $g2(x;y)$:

1. $f1(x; y) = x^3 - 4x^2 - 8x - 8$;
2. $f2(x; y) = x^3 - 4x + x$;
3. $g1(x; y) = x^3y - 6x^2y + 5xy - 6y$;
4. $g2(x; y) = x^3 - 6x^2 - 7x - 8$.

Ступінь знайденої ізогенії дорівнює 3.

Обчислення ізогенії еліптичної кривої

Побудуємо ізогенію кривої E_1: $y^2 = x^3 + x + 1 \; mod \; 19$ за допомогою алгоритму Велю з ядром ізогенії C: $\{O, (2, 7), (2, 12)\}$, де ядро ізогенії – це циклічна підгрупа простого порядку.

Алгоритм Велю для ядра C: $\{O, (2,7), (2,12)\}$ кривої E_1

1. Відкинути точку на нескінченності.

2. Знайти C_2 – множини точок парного порядку. R – решта точок. Точок парного порядку в підгрупі C немає.

3. Розбити R на дві частини – R_+ та R_-. Для R_+ обрано точку (2,7). Точка (2,12) обернена до неї, так як $7 = -12 \; mod \; 19$.

4. Отримати множину $S = \{(2,7)\}$. Для кожної точки $Q = (x_Q, y_Q)$ із S знайти v та w.

Цикл виконується лише один раз, оскільки множина S містить лише одну точку $Q = (2,7)$, з координатами $x_Q = 2, y_Q = 7$.

$g_Q^x = 3 * 2^2 + 1 = 13, g_Q^y = -2 * 7 = 5, v_Q = 2 * 13 = 7, u_Q = 5^2 = 6, v = 7,$

$w = 6 + 2 * 7 = 1.$

5. Обчислити коефіцієнти A' та B' для ізогенної кривої E'.

$A' = 1 - 5 * 7 = 4, B' = 1 - 7 * 1 = 13.$

6. Обчислити формулу для раціонального відображення $(x, y) \to (\alpha, \beta)$ з використанням ядра C: $\{O, (2,7), (2,12)\}$ кривої (1):

$$\alpha = x + \sum_{Q \in S} \left(\frac{v_Q}{(x - x_Q)} + \frac{u_Q}{(x - x_Q)^2} \right), \alpha = x + \frac{7}{x-2} + \frac{6}{(x-2)^2} = \frac{x^3 - 4x^2 + 11x - 8}{x^2 - 4x + 4};$$

$$\beta = y - \sum_{Q \in S} \left(u_Q \frac{2y}{(x - x_Q)^3} + v_Q \frac{y - y_Q}{(x - x_Q)^2} - \frac{g_Q^x g_Q^y}{(x - x_Q)^2} \right), \beta = \frac{x^3 y - 6x^2 y + 5xy - 6y}{x^3 - 6x^2 + 12x - 8}.$$

7. Обчислити $(x, y) \to (\alpha, \beta)$ з використанням C: $\{O, (2,7), (2,12)\}$ кривої E_1: $y\text{^}2 = x\text{^}3 + x + 1 \; mod \; 19$ на E_2: $y^2 = x^3 + 4x + 13 \; mod \; 19$.

Обрано точки кривої E_1 та виконано перевірку правильності ізоморфної трансформації точок кривої. Взято випадкові точки кривої $P_1 = (9;6)$ та $P_2 = (14;2)$. $P_1 + P_2 = P_3 = (5;6)$. Ізоморфною точкою для P_1 на кривій E_2 буде точка $Q_1 = (14;1)$, відповідно $Q_2 = (17;4)$. Обчислимо $Q_1 + Q_2 = Q_3 = (8;5)$. Трансформація точки $P_3 = (5;6)$ на криву E_2 дає точку Q_3, що підтверджує правильність ізоморфної трансформації.

$$\alpha = \frac{x^3 - 4x^2 + 11x - 8}{x^2 - 4x + 4} = 72 * \; 9^{-1} = 72 * 8 \; mod \; 19 = 8 \, , x = 5.$$

$$\beta = \frac{x^3 y - 6x^2 y + 5xy - 6y}{x^3 - 6x^2 + 12x - 8} = -36 * \; 27^{-1} = -36 * 12 \; mod \; 19 = 5, x = 5, \text{y} = 6.$$

Відображення точок кривої E_1: $y^2 = x^3 + x + 1 \; mod \; 19$ на криву E_2: $y^2 = x^3 + 4x + 13 \; mod \; 19$ подано в таблиці 6. З використанням отриманого раніше перетворення φ: $(x, y) \to (\alpha, \beta)$ на основі значень $P_i \in E_1$ були отримані значення $Q_i \in E_2$.

Таблиця 6

Відображення точок кривої E_1 на криву E_2

№	1	2	3	4	5	6	7	8	9	10	11	12
P_i	*(0,1)*	*(7,16)*	*(14,17)*	*(0,18)*	*(7,3)*	*(14,2)*	*(5,6)*	*(10,2)*	*(16,16)*	*(5,13)*	*(10,17)*	*(16,3)*
Q_i	*(17, 15)*			*(17, 4)*			*(8, 5)*			*(8, 14)*		
№	13	14	15	16	17	18	19	20	21			
P_i	*(9,6)*	*(13,11)*	*(15,3)*	*(9,13)*	*(13,8)*	*(15,16)*	*(2,7)*	*(2,12)*	*O*			
Q_i	*(14, 1)*			*(14, 18)*			*O*					

Таким чином, ізогенія кривої знайдена, побудована операція перетворення точок з використанням ізогенії для E_1 та E_2, що надає можливість її застосування в криптопримітивах.

Оцінка кількості ізогеній для фіксованої форми еліптичної кривої

Для оцінки кількості ізогеній скористаємось кривою (2) над F_{11} і F_{19} та перебираючи всі коефіцієнти кривої A та B, отримаємо кількість ізогеній для кожного випадку, але використаємо лише ізогенії простого порядку. В таблиці 7 наведено фрагмент отриманих даних.

Таблиця 7

F_{11}										
коефіцієнт A	1	1	1	1	1	1	1	1	1	1
коефіцієнт B	1	2	3	4	5	6	7	8	9	10
порядок кривої #E	14	16	18	9	11	13	15	6	8	10
порядок ізогеній	(**2**,**7**)	(**2**,4,8)	(**2**,**3**,6,9,18)	(3,9)	(11)	(13)	(**3**,**5**,15)	(**2**,**3**,6)	(**2**,4)	(**2**,**5**,10)
кількість ізогеній простого порядку N_ϕ	2	1	2	1	1	1	2	2	1	2

Були проведені обчислення для інших значень полів F_{11}, F_{113}, F_{257}, результати яких наведені на рисунках 1–3.

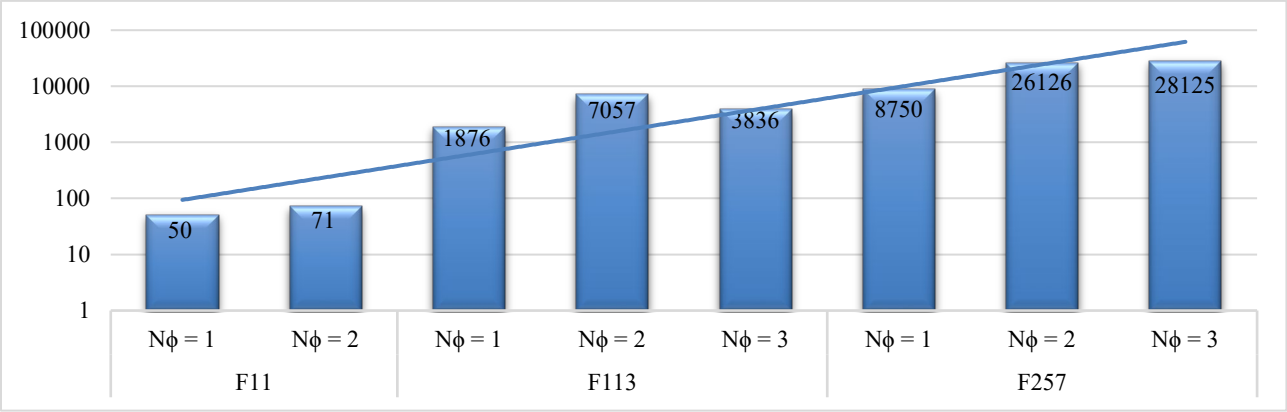


Рис. 1. Оцінка кількості кривих, які мають ядро ізогенії простого порядку (кількість циклічних підгруп простого порядку), $N_\phi = 1,2,3$

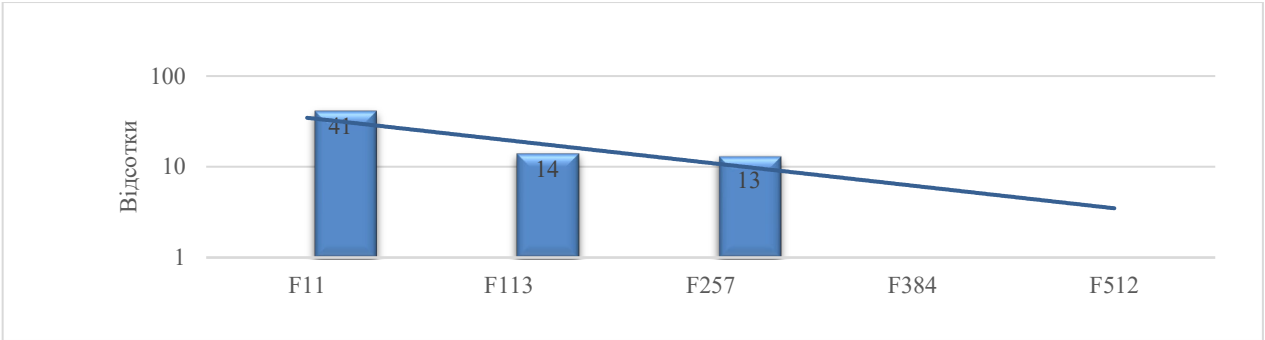


Рис. 2. Залежність кількості кривих, які мають ядро ізогенії простого порядку від #E кривої для $N_\phi = 1$

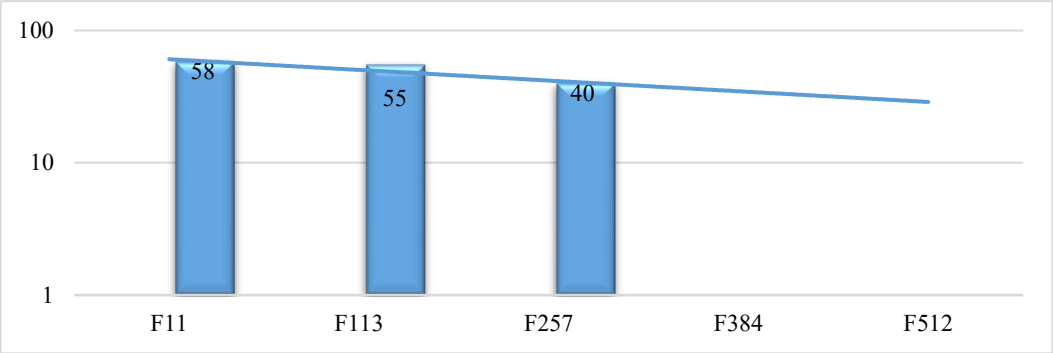


Рис. 3. Залежність кількості кривих, які мають ядро ізогенії простого порядку від #E кривої для N_ϕ = 2

Висновки

Отже, в процесі проведених досліджень було розроблено програмні функції для реалізації операцій над ізогеніями еліптичних кривих різного порядку, які забезпечать зазначені в стандарті рівні безпеки: 256, 384, 512 з урахуванням зростання потужності квантових комп'ютерів. Отримано значення кількості ізогенних еліптичних кривих над простим полем F11, які склали 41 % (50) і 58 % (71) для N_ϕ = 1 і N_ϕ = 2 відповідно, над полем F113 ці значення складають 14 % (1876) і 55 % (7057) для N_ϕ = 1, N_ϕ = 2 відповідно, над полем F257 ці значення складають 13 % (8750) і 40 % (26126) для N_ϕ = 1 і N_ϕ = 2 відповідно від значення характеристики скінченного поля. Визначено, що удосконалення сучасних криптосистем, представлених в якості постквантових кандидатів і побудованих з використанням операцій з ізогеніями еліптичних кривих, можливо завдяки використанню розпаралелювання операцій в групі точок еліптичної кривої та з ізогенією кривої, використанню різноманітних ізогеній еліптичної кривої, потужність яких достатня для використання як додаткових секретних параметрів. Побудовано ізогенію еліптичної кривої третього порядку та отримано ізоморфну трансформацію еліптичної кривої. Отримано апроксимовані значення кількості ізогенних кривих для полів F512, N_ϕ = 2 – 28 % (73400), N_ϕ = 1 – 4 % (10485), F384, N_ϕ = 2 – 33 % (48660), N_ϕ = 1 – 8 % (11796). Підтверджено правильність побудованої ізогенії. Побудовано та перевірено операції для обчислення ізогенії заданого порядку еліптичної кривої, які дозволяють використовувати в майбутньому стандартні рівні безпеки.

Отримані експериментальні значення скалярного добутку з використанням розпаралелювання скалярного множення точки еліптичної кривої дозволили оцінити залежності часу виконання скалярного множення точки кривої t_k від параметрів ω та δ. Розроблена програмна реалізація операцій скалярного множення точки кривої, операцій над ізогеніями еліптичної кривої дозволяє обчислювати значення часу скалярного множення з визначеною потужністю обчислювальної системи, а саме з фіксованим значенням N_{pr} кількості регістрів (ядер процесору). За результатами експериментів було встановлено, що порівняно з класичним підходом до множення точки кривої, представлення скаляру k послідовністю 4-бітових слів прискорило операцію скалярного множення $k*P$ в 30 разів, завдяки встановленню значення складності обчислювальної системи $N_{pr} = 30$. Для 8-бітових слів прискорення склало 18,8 разів.

Перспективою подальших досліджень є розробка методів генерації та верифікації цифрового підпису, генераторів псевдовипадкових послідовностей на основі перетворень над точками ізогенії еліптичної кривої з використанням розпаралелювання операцій скалярного множення точки кривої.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Rivest R., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems // Communications of the ACM. New York City: Association for Computing Machinery. 1978. Vol. 21, Iss. 2. P. 120–126. ISSN 0001-0782; 1557-7317. DOI: 10.1145/359340.359342.

2. Bernstein D., Lange T., Niederhagen R. Dual EC: A Standardized Back Door // Cryptology ePrint Archive, Report 2015. P. 767. URL: https://projectbullrun.org/dual-ec/documents/dual-ec-20150731.pdf.

3. Shor P. W. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. Foundations of Computer Science: Conference Publications. 1997. P. 1484–1509.

4. Alkim E., Ducas L., Pöppelmann T., Schwabe P. Post-quantum key exchange – a new hope // IACR Cryptology ePrint Archive, Report 2015/1092, 2015.

5. Husemöller D., Theisen S., Forster O., Lawrence R. Elliptic Curves, Second Edition // Springer. 2002. P. 487.

6. Edwards H. A normal form for elliptic curves // Bulletin of the American Mathematical Society. Vol. 44, № 3. 2007. P. 393–422.

7. Bernstein D. J., Lange T. Inverted Edwards coordinates // Applied algebra, algebraic algorithms and error-correcting codes: 17th international symposium, AAECC-17, Bangalore, India, December 16-20, 2007, proceedings. LNCS 4851, Springer. 2007. P. 20–27.

8. Schoof R. Elliptic curves over finite fields and the computation of square roots modulo p. Bordeaux: Math. Comput. 1985. № 44. P. 483–494.

9. Schoof R. Counting points on elliptic curves over finite fields. J. Theor. Nombres. Bordeaux 7. 1995. P. 219–254.

10. Чевардин В. Е. Изоморфные трансформации эллиптической кривой над конечным полем // Международный научно-теоретический журнал «Кибернетика и системный анализ». 2013. Том 49, № 3. С. 168–171.

11. Christophe Negre, Jean-Marc Robert. New Parallel Approaches for Scalar Multiplication in Elliptic Curves over Fields of Small Characteristic // IEEE Transactions on Computers. 2015. № 64 (10). P. 2875–2890. URL: https://hal.science/hal-00908463v1/file/parallelization-ecsm8.pdf.

12. Federal Office for Information Security (BSI). BSI – Technical Guideline. Cryptographic Mechanisms: Recommendations and Key Lengths. BSI TR-02102-1. V. 2023-01.

13. Goubin L. A Refined Power-Analysis-Attack on Elliptic Curve Cryptosystems, Proceedings of Public-Key-Cryptography – PKC 2003, Lecture Notes in Computer Science 2567, Springer Verlag, 2003.

14. Stolbunov A. Constructing public-key cryptographic schemes based on class group action on a set of isogenous elliptic curves. Adv. in Math. of Comm. 2010. No. 4(2). P. 215–235.

15. Galbraith S., Stolbunov A. Improved algorithm for the isogeny problem for ordinary elliptic curves // Applicable Algebra in Engineering, Communication and Computing. 2013. No. 24(2). P. 107–131.

16. De Feo L., Jao D., Plˆut J. Towards quantum-resistant cryptosystems from supersingular elliptic curve isogenies // Journal of Mathematical Cryptology (to appear). 2014. URL: http://eprint.iacr.org/2011/506.

17. Costello C., Longa P., Naehrig M. Efficient algorithms for supersingular isogeny Diffie-Hellman // CRYPTO 2016. URL: https://eprint.iacr.org/2016/413.pdf.

18. Rostovtsev A., Stolbunov A. Public-key cryptosystem based on isogenies. URL: https://eprint.iacr.org/2006/145.pdf.

19. Tate J. Endomorphisms of abelian varieties over finite fields // Inventiones Mathematica. 1966. No. 2. P. 134–144.

20. Castryck W., Decru Th. An efficient key recovery attack on SIDH (PDF). In Carmit Hazay; Martijn Stam (eds.). Advances in Cryptology – EUROCRYPT 2023. International Association for Cryptologic Research. Lecture Notes in Computer Science. Vol. 14008. Springer. 2023. P. 423–447. DOI: 10.1007/978-3-031-30589-4_15. ISBN 978-3-031-30589-4.

DOI: 10.58254/viti.5.2024.20.227

УДК 003.26:519.688

канд. техн. наук Яковлєв С. В. ORCID: 0000-0002-5647-5043 (НТУУ «КПІ ім. Ігоря Сікорського»)
Кріпака І. А. (ННФТІ НТУУ «КПІ ім. Ігоря Сікорського»)

МЕТОДИ УСІЧЕННЯ ЦИФРОВОГО ПІДПИСУ ДЛЯ СХЕМ ТИПУ ЕЛЬ-ГАМАЛЯ ТА ДСТУ 4145-2002

Усічення цифрового підпису має важливе значення для реалізації криптографічних систем для малоресурсних пристроїв, зокрема у системах, де підписи зберігаються довгий час, але перевіряються відносно нечасто, таких як апаратні журнали аудиту чи сховища захищених документів. Скорочення підпису також може бути використане у тих випадках, коли обмежено пам'ять для представлення чи зберігання самого підпису (наприклад, у QR-кодах). У відкритих джерелах запропоновано методи усічення підпису для стандартів ECDSA та EdDSA.

У цій роботі запропоновано метод усічення цифрових підписів для класичної схеми Ель-Гамаля та деяких її узагальнень, які ґрунтуються на модулярній арифметиці. Запропоновані методи не вимагають зміни процедури підписування, а тому застосовні до довільних існуючих реалізацій схеми Ель-Гамаля; відновлення цифрового підпису відбувається завдяки збільшенню обчислювальних витрат з боку сторони, яка перевіряє підпис. Також запропоновано метод усічення цифрового підпису для національного стандарту ДСТУ 4145-2002; у цьому методі враховано особливості арифметики еліптичних кривих, на яких ґрунтується стандарт.

Запропоновані методи дозволяють ефективно реалізовувати усічення підписів типу Ель-Гамаля (особливо відчутно для підписів ДСТУ 4145-2002), відкидаючи до 32 бітів, але складність перевірки починає швидко зростати зі збільшенням частини підпису, що відкидається. Втім, для обраних меж довжини відкинутої частини запропоновані алгоритми мають відносно ефективну реалізацію і тому можуть бути використані для зменшення довжини підписів у протоколах для малоресурсних пристроїв.

Ключові слова: *цифровий підпис, схема Ель-Гамаля, еліптичні криві, ДСТУ 4145-2002.*

S. Yakovliev, I. Kripaka. Methods of digital signature truncation for El-Gamal-type signatures and for DSTU 4145-2002.

Digital signature truncation is important for implementing cryptographic systems for low-resource devices, particularly in systems where signatures are stored for a long term but are checked relatively infrequently. Among such systems we can mention hardware audit logs or secure document repositories. Signature truncation can also be used in cases where memory is limited to represent or store the signature itself (for example, in QR codes). Signature truncation methods for ECDSA and EdDSA standards were proposed in public sources.

This paper proposes a method of truncation of digital signatures for the classical El-Gamal scheme and some of its generalizations, which are based on modular arithmetic. The proposed methods do not require any alternations in the signing procedure, and are therefore applicable to arbitrary existing implementations of the El-Gamal scheme; the restoration of a digital signature occurs at the expense of an increase in computational costs on the part of the party verifying the signature. A digital signature truncation method for the national standard DSTU 4145-2002 is also proposed; this method takes into account the properties of the arithmetic of elliptic curves, on which the standard is based.

The proposed methods make it possible to effectively implement the truncation of signatures of the El-Gamal type (especially noticeable for DSTU 4145-2002 signatures), discarding up to 32 bits of signature, but the complexity of the verification increases rapidly with the increase in the discarded part of the signature. However, for selected limits of the length of the discarded part, the proposed algorithms have a relatively efficient implementation and therefore can be used to reduce the length of signatures in protocols for low-resource devices.

Keywords: *digital signature, El-Gamal scheme, elliptic curves, DSTU 4145-2002.*

Постановка завдання. Проблема скорочення підпису має велике значення в легкій криптографії для малоресурсних пристроїв, зокрема у системах, де підписи зберігаються довгий час, але перевіряються відносно нечасто (наприклад, в апаратних журналах аудиту чи сховищах захищених документів). Також скорочення підпису має значення у тих випадках, коли обмежено пам'ять для представлення самого підпису, наприклад, у QR-кодах. У таких ситуаціях було б бажано, щоб підпис скорочувався на стороні підписника без безпосередньої зміни алгоритму, – можливо, завдяки додатковим обчисленням на стороні перевіряючого. Також очевидною вимогою скороченого алгоритму підпису виступає збереження та незначне

зменшення стійкості, оскільки зловмисник може певним чином маніпулювати невідомою частиною підпису.

Нині відомі методи, направлені на скорочення розміру цифрового підпису для деяких типів схем цифрового підпису. Зокрема, був запропонований ефективний метод для скорочення підписів схем ECDSA та EdDSA, які побудовані на еліптичних кривих над простими полями. У цій роботі запропоновано аналогічні методи для класичних та узагальнених схем Ель-Гамаля, а також для національного стандарту цифрового підпису ДСТУ 4145-2002.

Аналіз публікацій за темою дослідження. Однією з основних сучасних схем цифрового підпису є *схема Ель-Гамаля* [3]. Наведемо її короткий опис, необхідний для викладення подальшого матеріалу.

1) Загальні параметри: p – велике просте число, g – елемент великого порядку q за модулем p. Особистим ключем виступає число $x \in \mathbb{Z}_q$, відкритим ключем – число $y = g^x \bmod p$. Використовується також геш-функція $H : \{0,1\}^* \to \mathbb{Z}_q^*$.

2) Постановка підпису описується таким алгоритмом:

– нехай k – випадково згенероване одноразове число, $k \in \mathbb{Z}_q^*$, $H(m)$ – геш-значення вхідного повідомлення m, $H(m) \in \mathbb{Z}_q^*$;
– обчислюються значення $r = g^k \bmod p$, $s = k^{-1}(H(m) - rx) \bmod q$;
– підписом повідомлення m є пара чисел (r, s);

3) Перевірка підпису виконується за співвідношенням:

$$y^r \cdot r^s \stackrel{?}{\equiv} g^{H(m)} \pmod p.$$

Якщо воно виконується, то підпис вважається вірним, інакше підпис невірний.

Варто зазначити, що довжина підпису у схемі Ель-Гамаля складає $\lceil log_2 q \rceil + \lceil log_2 p \rceil$ бітів. До прикладу, якщо обрати p як 1024-бітове число, а q як 160-бітове, то довжина підпису буде складати 1024 + 160 = 1184 бітів.

Алгебраїчні співвідношення, які лежать у основі схеми Ель-Гамаля, можуть бути замінені на інші, що дозволяє генерувати інші схеми зі зберіганням основних властивостей. Значну частину таких схем можна описати у формальний спосіб, який одержав назву *узагальненої схеми Ель-Гамаля* [4]. Такі схеми задаються параметрами A, B, C, які визначаються як певні функції від чисел $H(m)$, s, r. Постановка підпису відбувається аналогічно до класичної схеми, причому параметр s знаходиться зі співвідношення $Ak \equiv B + Cx \pmod q$. Перевірка підпису виконується за співвідношенням $r^A \stackrel{?}{\equiv} g^B y^C \pmod p$.

Зауважимо також, що схеми Ель-Гамаля (і класичну, й узагальнену) можна перевизначити на інших алгебраїчних структурах із мультиплікативною операцією, наприклад, на еліптичних кривих.

У роботі [1] було розглянуто три можливих методи скорочення підписів у схемах типу Ель-Гамаля.

1. Схема із використанням функції компресії.

У цьому підході пропонується робити скорочення параметрів завдяки деякій функції компресії. Найбільш вдала реалізація цього методу, на нашу думку, викладена у стандарті цифрового підпису DSA [5], де значення r, s обчислюються за формулами

$$r = (g^k \bmod p) \bmod q, \qquad s = k^{-1}(H(m) + xr) \bmod q,$$

а перевірочне співвідношення має такий вид:

$$r \stackrel{?}{\equiv} \left(\left(g^{H(m)s^{-1} \bmod q} \cdot y^{rs^{-1} \bmod q}\right) \bmod p\right) \bmod q.$$

Перевагою цього методу є його довжина підпису, яка складає $2\lceil log_2 q \rceil$ бітів. До прикладу, якщо обрати p як 1024-бітове число, а q як 160-бітове, то довжина підпису DSA буде складати 320 бітів проти 1184 бітів у класичній схемі Ель-Гамаля. Але варто зазначити, що застосування цього методу до схем типу Ель-Гамаля вимагає їхньої значної алгоритмічної перебудови, фактично – побудови нової схеми цифрового підпису.

2. Фіксування частини підпису.

У цьому підході для зменшення підпису фіксується частина параметру r: $r = r_0 \| const$, де r_0 – обчислювана частина підпису, $const$ – зафіксована частина підпису, яка є загальним параметром схеми. Щоб отримати такий підпис, генеруються випадкові значення k доти, доки r не матиме потрібної форми. Сам підпис при цьому буде мати форму (r_0, s) та, відповідно, мати скорочену довжину. Альтернативно можна фіксувати частину параметру s.

Зауважимо, що цей підхід вимагає суттєвого збільшення ресурсних витрат у сторони, яка підписує, і тому для малопотужних пристроїв цей метод слабко застосовний.

3. Відкидання частини підпису.

Основна ідея цього методу полягає в скороченні частини s підпису. Величина s представляється як $s = s_0 \| s_1$, сам підпис буде мати форму (r, s_0), а частина s_1 буде відновлюватись під час перевіряння підпису. На противагу попередньому способу, тут обчислювальне навантаження збільшується на стороні, яка перевіряє підпис.

Саме цей метод розглядався у [1] як основний, але Томас Порнін [2] незалежно дослідив такий метод та запропонував його ефективну реалізацію для схем цифрового підпису ECDSA/EdDSA.

Метою цієї роботи є розробка ефективних методів усічення цифрового підпису для схем типу Ель-Гамаля, зокрема для національного стандарту цифрового підпису ДСТУ 4145-2002.

Виклад основного матеріалу дослідження. Сформулюємо методи усічення підпису для класичної схеми Ель-Гамаля та його узагальнених варіантів на основі відкидання частини підпису.

Спочатку необхідно обмежити доцільну величину скорочення підпису: очевидно, що ми не можемо відкинути дуже велику кількість бітів s, тому що тоді втратиться залежність з бітами секретного ключа x та одноразового ключа k і підпис стане уразливим до атак відновлення ключа та підробки. Будемо вважати, що кількість t бітів підпису, які відкидаються, лежить в межах $8 \leq t \leq 32$, і що сторони заздалегідь домовляються про конкретне значення t.

Представимо частину *s* як $s = s_0 \| s_1 = s_0 + s_1 2^n$, де $n = \lceil log_2 q \rceil - t$, s_0 – частина підпису, що залишається, а s_1 – частина, що відкидається. Підписом у такому випадку буде пара (r, s_0).

Для відновлення підпису під час перевіряння метод верифікації потребує алгоритму ефективного знаходження можливих кандидатів разом із перевіркою на правильність перевірочного співвідношення. Застосуємо для такої задачі алгоритм великих та малих кроків (BSGS [6]), який повертає відповідь у форматі $s_1 = i + jI$, $i \in \{1, \dots I\}$, $j \in \{1, \dots J\}$, а параметри I та J визначаються налаштуваннями алгоритму BSGS.

З перевірочного співвідношення схеми Ель-Гамаля одержуємо:

$$y^r \cdot r^{s_0 + s_1 2^n} \stackrel{?}{\equiv} g^{H(m)} \ (mod\ p),$$
$$y^r \cdot r^{s_0 + i2^n} \stackrel{?}{\equiv} g^{H(m)} r^{jI \cdot 2^n} \ (mod\ p).$$

Останнє співвідношення дозволяє адаптувати алгоритм BSGS для нашої задачі.

Алгоритм 1: знаходження значення s_1 для відновлення усіченого підпису в класичній схемі Ель-Гамаля.

1. Встановити $J = 2^{t/2}$, $I = 2^{t/2-1}$.
2. Для j від 0 до J обчислюємо значення $U_j = g^{H(m)} \cdot r^{-jI\,2^n} \ mod\ p$.
3. Для i від 0 до I обчислюємо значення $V_i = y^r \cdot r^{s_0 + i\,2^n}$.
4. Шукаємо збіги між множинами $\{U_j\}$ та $\{V_i\}$. Кожен збіг $U_i = V_j$ формує кандидата на відновлений підпис $s = s_0 + (i + Ij)2^n$, якого вже перевіряємо за основним перевірочним співвідношенням схеми Ель-Гамаля. Якщо збігів не виявлено, підпис вважається невірним.

В аналогічний спосіб будується алгоритм усічення цифрового підпису для узагальненої схеми Ель-Гамаля. Зауважимо, що аналітична форма величин U_j, V_i визначається через параметри узагальненої схеми; приклади таких співвідношень для різних варіантів узагальнених схем наведено у таблиці 1. Варто додати, що у таблиці параметри, позначені як «довільні», не повинні залежати від s, інакше наведений метод вимагатиме уточнення та додаткової адаптації.

Алгоритм відновлення та перевіряння усіченого цифрового підпису в узагальненій схемі Ель-Гамаля повністю повторює Алгоритм 1, за винятком інших формул для обчислення значень U_j, V_i.

Таблиця 1

Формули обчислення величин U_j, V_i для деяких варіантів узагальнених схем Ель-Гамаля із різними параметрами A, B, C

Параметри	**Перевірочне співвідношення**	U_j	V_i
$A = s,\ B = m,\ C = -r$	$r^s = g^m y^{-r}$	$g^m r^{-(jI2^n)}$	$y^r r^{(s_0+i2^n)}$
$A = s,\ B,\ C$ — довільні	$r^s = g^B y^C$	$g^B y^C r^{-(jI2^n)}$	$r^{(s_0+i2^n)}$
$A = -s,\ B,\ C$ — довільні	$r^{-s} = g^B y^C$	$g^B y^C r^{(jI2^n)}$	$r^{-(s_0+i2^n)}$
$B = s,\ A,\ C$ — довільні	$r^A = g^s y^C$	$g^{(s_0+jI2^n)} y^C$	$r^A g^{-(i2^n)}$
$B = -s,\ A,\ C$ — довільні	$r^A = g^{-s} y^C$	$g^{-(s_0+jI2^n)} y^C$	$r^A g^{(i2^n)}$
$C = s,\ A,\ B$ — довільні	$r^A = g^B y^s$	$g^B y^{(s_0+jI2^n)}$	$r^A y^{i2^n}$
$C = -s,\ A,\ B$ — довільні	$r^A = g^B y^{-s}$	$g^B y^{-(s_0+jI2^n)}$	$r^A y^{i2^n}$
$A = rs,\ B = m,\ C = const = c$	$r^{rs} = g^m y^c$	$g^m y^c r^{-(rjI2^n)}$	$r^{r(s_0+i2^n)}$
$A = ms,\ B,\ C$ — довільні	$r^{ms} = g^B y^C$	$g^B y^C r^{-(mjI2^n)}$	$r^{m(s_0+i2^n)}$

Національний стандарт цифрового підпису ДСТУ 4145-2002 [7] описує схему цифрового підпису типу Ель-Гамаля на основі еліптичних кривих у формі Веєрштраса над полями F_{2^m} характеристики 2, що суттєво відрізняє його від алгоритмів ECDSA та EdDSA. Використовуються еліптичні криві у формі $y^2 + xy = x^3 + ax^2 + b$, де a, b є параметрами.

Деталі та нюанси арифметики еліптичних кривих над простими полями та полями характеристики 2 можна знайти, наприклад, у [8].

Наведемо необхідні для подальшого викладення відомості про схему ДСТУ 4145-2002.

1) цифровий підпис D має вид $D = (r \parallel s)$, де s, r – частини цифрового підпису;

2) головне співвідношення для перевірки підпису має вид $R = sP + rQ$, де Q – відкритий ключ (точка на еліптичній кривій), P – базова точка кривої, R – спеціально обчислена на основі підписаного повідомлення точка еліптичної кривої;

3) алгоритм перевіряння цифрового підпису складається з таких кроків:

а) відновити значення r та s з підпису D;

б) обчислити точку $R \coloneqq sP + rQ$, $R = (x_R, y_R)$;

в) обчислити елемент поля $y = hx_R$, де h є елементом поля, одержаним із геш-значення $H(M)$ повідомлення M, яке перевіряється;

г) перетворити елемент поля y у число r';

д) перевірити рівність $r \stackrel{?}{=} r'$.

Як і для класичної схеми Ель-Гамаля, будемо робити усічення частини s. Знову покладаємо $s = \mathrm{s}_0 + \mathrm{s}_1 2^n$, де s_1 – t-бітова частина, що відкидається. Підписом у такому випадку, буде $D = (r \parallel s_0)$.

Процедура перевіряння підпису ґрунтується на пошуку оригінального значення s_1; для цього необхідно перебудувати сам алгоритм перевіряння підпису.

Алгоритм 2: перевіряння усіченого підпису ДСТУ 4145-2002.

1. Відновити точки еліптичної кривої R з відомого значення r (алгоритм 3).
2. Застосувати модифікований алгоритм BSGS (алгоритм 4) для пошуку s_1.
3. Перевірити знайдені точки та кандидати у s_1 головним перевірочним співвідношенням. Підпис вважається правильним, якщо хоча б одне співвідношення виконалось.

Для відновлення точки R використовується такий алгоритм.

Алгоритм 3: відновлення точки еліптичної кривої R з відомого значення r

Вхід: – r: частина підпису за ДСТУ 4145-2002;
– a, b: параметри використаної еліптичної кривої;
– h: геш-значення вхідного повідомлення ($h = H(M)$).

Вихід: точки еліптичної кривої R та $-R$.

1. Обчислити $x_R = r \cdot h^{-1}$ у скінченному полі.
2. Розв'язати рівняння $y^2 + xy = x^3 + ax^2 + b$ зі значенням $x = x_R$ для одержання двох значень y_R. Дві пари (x_R, y_R) формують обидві точки R та $-R$.

Для пошуку кандидатів у скорочену частину підпису використовується такий алгоритм.

Алгоритм 4: знаходження значення s_1 для відновлення усіченого підпису у схемі ДСТУ 4145-2002.

Вхід: – r, s_0: частини усіченого підпису за ДСТУ 4145-2002;
– P, Q, R: базова точка кривої, відкритий ключ та відновлена точка R;
– t:довжина відкинутої частини підпису.

Вихід: відновлене значення s_1.

1. Встановити $I = 2^{t/2}, J = 2^{t/2-1}$.
2. Для j від 0 до $J - 1$ обчислити точки $U_j = s_0 P + jI(2^n P) + rQ$. Зберегти координати x обчислених точок.
3. Для i від 0 до $I - 1$ обчислити точки $V_i = R - i(2^n P)$. Зберегти координати x обчислених точок.
4. Шукаємо збіги між збереженими x-координатами точок U_j та V_i. Кожен збіг дає кандидата $s_1 = i + Ij$. Якщо збігів не виявлено, підпис вважається невірним.

Зауважимо, що $U_{j+1} = U_j + (I \cdot 2^n)P$ та $V_{i+1} = V_i - 2^n P$, отже, послідовності точок U та V можуть бути просто обчислені звичайним додаванням із передобчисленими точками.

Алгоритм 4 може знайти двох можливих кандидатів у s_1, якщо йому на вхід замість правильної точки R подати $(-R)$ – іншу точку, яка відповідає значенню r. Це можливо, оскільки перевіряючий не знає, яка саме точка з цих двох була використана під час генерування цифрового підпису. З іншого боку, для формування підпису використовується тільки x-координата точки R, яка збігається у R та $(-R)$. Втім, перевіряючий, якщо він має якісь сумніви, може перевірити головні співвідношення для кожної точки:

$$R = s_0 P + s_1(2^n)P + rQ, \qquad -R = s_0 P + s_1(2^n)P + rQ,$$

і якщо одне з них буде виконуватись, приймати підпис як вірний.

У таблиці 2 наведено оцінки ефективності запропонованого алгоритму відновлення усіченого підпису порівняно з повним перебором.

Таблиця 2

Порівняння кількості операцій алгоритмів відновлення усіченого цифрового підпису та повного перебору

t	Складність звичайного (повного) перебору s_1	Складність алгоритмів 1 та 4
8	$2^8 = 256$	$I + J = 24$
16	$2^{16} = 65536$	$I + J = 384$
24	$2^{24} = 16777216$	$I + J = 6144$
32	$2^{32} = 4294967296$	$I + J = 98304$

Запропоновані методи дозволяють ефективно реалізовувати усічення підписів типу Ель-Гамаля на певну кількість бітів (не дуже велику для класичної схеми Ель-Гамаля, але доволі суттєву для ДСТУ 4145-2002), але складність перевірки починає швидко зростати зі збільшенням частини підпису, що відкидається. Втім, для обраних меж довжини відкинутої частини t запропоновані алгоритми мають відносно ефективну реалізацію і тому можуть бути використані для зменшення довжини підписів у протоколах для малоресурсних пристроїв. Варто зауважити, що такий метод можна комбінувати із методом № 2 (фіксуванням частини r), але при цьому суттєво зросте навантаження на підписника, що може бути неприйнятним для певних випадків, і необхідно провадити ретельний аналіз ризиків безпеки, пов'язаних зі зниженням стійкості такого цифрового підпису.

Висновки. У роботі було запропоновано метод скорочення цифрового підпису для схем типу Ель-Гамаля. Були розглянуті класична схема Ель-Гамаля та її узагальнений багатопараметричний варіант. Також був запропонований метод усічення для цифрових підписів ДСТУ 4145-2002 з урахуванням особливостей арифметики еліптичних кривих, на якій ґрунтується цей стандарт. Запропоновані методи не змінюють процедуру постановки підпису; обчислений підпис скорочується шляхом відкидання певних бітів. Це дозволяє застосовувати запропоновані методи у реалізаціях криптографічних систем на малоресурсних пристроях, зокрема використовувати вже існуючі реалізації без змін. Для перевіряння усіченого підпису використовується відновлення на основі алгоритму великих та малих кроків, який дозволяє розв'язувати таку задачу із високою ефективністю.

Напрямками подальших досліджень є пошук границі для довжини усікання підпису, за якої підпис ще можна вважати практично стійким до відомих атак, а також пошук методів усікання інших типів цифрового підпису (наприклад, RSA).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. L. Akhmetzyanova, E. Alekseev, A. Babueva, S. Smyshlyaev. On Methods of Shortening ElGamal-type Signatures. Cryptology ePrint Archive, Paper 2021/148. URL: https://eprint.iacr.org/2021/148.

2. Pornin T. Truncated EdDSA/ECDSA Signatures. Cryptology ePrint Archive, Paper 2022/938. URL: https://eprint.iacr.org/2022/938.

3. ElGamal T. A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms // IEEE transactions on information theory. 1985. Т. 31. № 4. P. 469–472.

4. Schneier B. Applied Cryptography: Protocols, Algorthms, and Source Code in C. 20th Anniversary Edition. John Wiley & Sons, 2015. 784 p. ISBN 978-1-119-09672-6.

5. FIPS PUB 186-5:2019. Standards Federal Information Processing. Digital Signature Standard (DSS). URL: https://doi.org/10.6028/NIST.FIPS.186-5.

6. Menezes A. J., Van Oorschot P. C., Vanstone S. A. Handbook of Applied Cryptography. CRC Press, 1996. 816 p. ISBN 0-8493-8523-7.

7. ДСТУ 4145-2002. Національний стандарт України. Інформаційні технології. Криптографічний захист інформації. Цифровий підпис, що ґрунтується на еліптичних кривих. Формування та перевірка.

8. Chen L. et al. Recommendations for discrete logarithm-based cryptography: Elliptic curve domain parameters. 2023. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-186.pdf.

АВТОРИ НОМЕРА

1. ***Бабарика Анатолій Олександрович*** – доктор філософії, доцент кафедри зв’язку та інформаційних систем Національної академії Державної прикордонної служби України ім. Б. Хмельницького, м. Хмельницький, Україна.

2. ***Басараб Олександр Корнійович*** – кандидат технічних наук, доцент, викладач кафедри зв’язку та інформаційних систем Національної академії Державної прикордонної служби України ім. Б. Хмельницького, м. Хмельницький, Україна.

3. ***Бєляков Роберт Олегович*** – кандидат технічних наук, доцент, докторант науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

4. ***Бернацький Андрій Петрович*** – старший викладач Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

5. ***Білий Олександр Анатолійович*** – провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

6. ***Бондаренко Дмитро Михайлович*** – начальник центру Державного науково-дослідного інституту технологій кібербезпеки та захисту інформації, м. Київ, Україна.

7. ***Бондаренко Леонід Олександрович*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

8. ***Вавpічен Олексій Анатолійович*** – старший викладач кафедри Національної академії Державної прикордонної служби України ім. Богдана Хмельницького, м. Хмельницький, Україна.

9. ***Власенко Олександр Володимирович*** – ад’юнкт науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

10. ***Городиський Роман Олександрович*** – викладач кафедри Національної академії Державної прикордонної служби України ім. Богдана Хмельницького, м. Хмельницький, Україна.

11. ***Громлюк Мирослав Миколайович*** – офіцер Збройних Сил України, м. Київ, Україна.

12. ***Гроздов Андрій Анатолійович*** – офіцер Збройних Сил України, м. Київ, Україна.

13. ***Долженко Галина Вікторівна*** – ад’юнкт науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

14. ***Драглюк Олексій Вікторович*** – начальник управління Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

15. ***Захарчук Любов Володимирівна*** – асистент кафедри інформаційних технологій в телекомунікаціях Навчально-наукового інституту телекомунікаційних систем Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

16. ***Зінченко Ірина Анатолієвна*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

17. ***Івченко Микола Миколайович*** – начальник відділу Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

18. ***Ільїнов Михайло Дмитрович*** – кандидат технічних наук, доцент кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

19. ***Ільницький Анатолій Іванович*** – кандидат технічних наук, доцент, доцент кафедри інформаційних технологій в телекомунікаціях Навчально-наукового інституту телекомунікаційних систем Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

20. ***Кисиленко Дар’я Юріївна*** – ад’юнкт науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

21. ***Клімович Сергій Олегович*** – кандидат технічних наук, начальник кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

22. ***Коваленко Ілля Григорович*** – кандидат технічних наук, старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

23. ***Комаров Володимир Олександрович*** – кандидат технічних наук, провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

24. ***Кріпака Ілля Анатолійович*** – магістр Навчально-наукового фізико-технічного інституту Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

25. ***Кузавков Василій Вікторович*** – доктор технічних наук, професор, начальник кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

26. ***Куцаєв Павло Володимирович*** – молодший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

27. ***Лаврик Іван Васильович*** – ад’юнкт науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

28. ***Липський Олександр Анатолійович*** – кандидат технічних наук, доцент, головний науковий співробітник Державного науково-дослідного інституту технологій кібербезпеки та захисту інформації, м. Київ, Україна.

29. ***Масесов Микола Олександрович*** – кандидат технічних наук, начальник наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

30. ***Міночкін Анатолій Іванович*** – заслужений працівник освіти України, доктор технічних наук, професор, провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

31. ***Могилевич Дмитро Ісакович*** – доктор технічних наук, завідувач кафедри Інституту спеціального зв’язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

32. ***Наритник Теодор Миколайович*** – кандидат технічних наук, професор, директор Інституту електроніки та зв’язку Української академії наук національного прогресу, м. Київ, Україна.

33. ***Остапчук Віктор Миколайович*** – кандидат технічних наук, начальник Головного управління зв’язку та кібербезпеки Генерального штабу Збройних Сил України, м. Київ, Україна.

34. ***Овсянніков Вячеслав Володимирович*** – провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

35. ***Паламарчук Наталія Анатоліївна*** – начальник науково-дослідної лабораторії Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

36. ***Панченко Ігор В’ячеславович*** – кандидат технічних наук, начальник кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

37. ***Площик Анна Сергіївна*** – викладач кафедри зв’язку та інформаційних систем Національної академії Державної прикордонної служби України ім. Б. Хмельницького, м. Хмельницький, Україна.

38. ***Процюк Юрій Олександрович*** – провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

39. ***Радзівілов Григорій Данилович*** – кандидат технічних наук, професор, заступник начальника Військового інституту телекомунікацій та інформатизації ім. Героїв Крут з наукової роботи, м. Київ, Україна.

40. ***Романов Дмитро Олександрович*** – начальник Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

41. ***Руденко Володимир Іванович*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

42. ***Сайко Володимир Григорович*** – доктор технічних наук, професор кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

43. ***Стефанишин Ярослав Іванович*** – науковий співробітник Державного науково-дослідного інституту технологій кібербезпеки та захисту інформації, м. Київ, Україна.

44. ***Стрельбіцький Михайло Анатолійович*** – доктор технічних наук, професор, викладач кафедри Національної академії Державної прикордонної служби України ім. Богдана Хмельницького, м. Хмельницький, Україна.

45. ***Субач Ігор Юрійович*** – доктор технічних наук, професор, завідувач спеціальної кафедри Інституту спеціального зв’язку та захисту інформації Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

46. ***Табенський Сергій Миколайович*** – викладач кафедри Національної академії Державної прикордонної служби України ім. Б. Хмельницького, м. Хмельницький, Україна.

47. ***Терещенко Тетяна Павлівна*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

48. ***Ушаков Віктор Дмитрович*** – директор, фізична особа – підприємець «Ушаков Віктор Дмитрович», м. Київ, Україна.

49. ***Фесенко Олексій Дмитрович*** – доктор філософії, доцент кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

50. ***Фесьоха Віталій Вікторович*** – доктор філософії, доцент кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

51. ***Фесьоха Надія Олександрівна*** – доктор філософії, старший викладач кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

52. ***Фомін Микола Миколайович*** – доктор філософії, начальник кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

53. ***Хоменко Павло Володимирович*** – ад’юнкт науково-організаційного відділу Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

54. ***Хоптинський Руслан Петрович*** – кандидат технічних наук, доцент кафедри Національної академії Державної прикордонної служби України ім. Богдана Хмельницького, м. Хмельницький, Україна.

55. ***Хусаінов Павло Володимирович*** – кандидат технічних наук, старший викладач кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

56. ***Цимбал Ірина Володимирівна*** – науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

57. ***Чевардін Владислав Євгенійович*** – доктор технічних наук, начальник кафедри Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

58. ***Черниш Юлія Олександрівна*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

59. ***Шаціло Петро Васильович*** – кандидат технічних наук, доцент, провідний науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

60. ***Шугалій Ольга Олександрівна*** – старший науковий співробітник Наукового центру зв’язку та інформатизації Військового інституту телекомунікацій та інформатизації ім. Героїв Крут, м. Київ, Україна.

61. ***Яковлєв Сергій Володимирович*** – кандидат технічних наук, завідувач кафедри математичних методів захисту інформації Національного технічного університету України «Київський політехнічний інститут ім. Ігоря Сікорського», м. Київ, Україна.

ПАМ'ЯТКА АВТОРУ

Наукові статті у фахових виданнях повинні мати такі необхідні елементи:

анотація як стисла стаття;

постановка проблеми у загальному вигляді та її зв'язок із важливими науково-практичними завданнями;

аналіз останніх досліджень і публікацій, в яких започатковано розв'язання зазначеної проблеми і на які спирається автор, виділення невирішених раніше частин загальної проблеми, котрим присвячується стаття;

формулювання мети статті (**постановка наукового завдання**);

виклад основного матеріалу дослідження з повним обґрунтуванням отриманих наукових результатів;

висновки з цього дослідження і **перспективи** подальших досліджень у зазначеному напрямку;

список використаних джерел: 10–15 посилань терміном **не більше 10 років.**

Редакція не рекомендує використання джерел інформації держав-агресорів.

До друку приймаються оригінальні наукові праці, які не було відправлено до інших редакцій та не опубліковано раніше в інших виданнях.

Рукопис статті потрібно подавати разом з наступними документами:

довідка про результати перевірки на академічний плагіат;

акт експертної оцінки про можливість відкритого опублікування (1 примірник);

згода на публікацію статті та оприлюднення персональних даних;

відомості про автора (авторів) за формою: прізвище, ім'я та по батькові; науковий ступінь, вчене звання, посада; назва установи, де працює автор, її місце розташування (місто, країна); обліковий запис автора ORCID, електронна адреса та номер телефону.

Відомості про автора (авторів) подаються окремим супровідним файлом. Наприклад:

№ з/п	Українська мова	Англійська мова
1	**Шевченко Олександр Іванович** Кандидат технічних наук Доцент Старший викладач Військовий інститут телекомунікацій та інформатизації імені Героїв Крут; м. Київ, Україна https://orcid.org/0000-0000-0000-0000 e-mail _____, телефон _____	**Shevchenko Oleksandr** Candidate of Technical Sciences Associate Professor Senior Lecturer Military Institute of Telecommunications and Informatization Technologies named after Heroes of Kruty; Kyiv, Ukraine https://orcid.org/0000-0000-0000-0000

Рукопис подається в друкованому та електронному вигляді у текстовому редакторі Microsoft Word, а також може бути надісланий за електронною адресою: naukaviti@viti.edu.ua, naukaviti@gmail.com.

Якщо стаття подається в електронному вигляді, супровідні документи подаються у сканованому вигляді.

Обсяг рукопису – не менше 7 повних аркушів українською або англійською мовами.

Формат аркушу – А4 (210 мм × 297 мм).

Параметри сторінки: зліва – 20 мм, справа – 20 мм, зверху – 20 мм, знизу – 20 мм.

Основний шрифт статті – Times New Roman, розмір 12 пт, міжрядковий інтервал – 1,0; абзацний відступ – 1,0 см; вирівнювання – по ширині; із виключенням переносів.

У лівому кутку першої сторінки на першому рядку друкується шифр УДК, розмір 12 пт, у правому кутку першої сторінки на другому рядку – дані про авторів: науковий ступінь, вчене звання, ініціали і прізвище автора, ORCID, в дужках назва установи, де він працює. Наприклад: к. т. н. Шевченко О. І. ORCID: 0000-0000-0000-0000 (ВІТІ ім. Героїв Крут). Далі через рядок друкується назва статті (відцентрована, великими напівжирними літерами).

Пропуск (перед і після назви статті та кожного розділу) – 1 рядок.

Анотацію друкують курсивом, шрифт Times New Roman, розмір 10 пт. Анотацію та ключові слова приводять українською та англійською мовами. Обсяг кожної з них не менше **1800 знаків з пробілами**, включаючи ключові слова. Анотація повинна бути структурована таким чином: вступ, проблематика, мета, матеріали й методи, результати, висновки. Іншими словами, анотація повинна відображати послідовну логіку опису результатів, описувати основну мету дослідження та підсумовувати найбільш значимі результати. Скорочення слів в анотації не застосовувати.

Після анотації вказати 6–8 ключових слів окремо українською та англійською мовами.

Список використаних джерел оформлюється шрифтом Times New Roman, розмір 11 пт, та складається з урахуванням Національного стандарту України ДСТУ 8302:2015 «Інформація та документація. Бібліографічне посилання. Загальні положення та правила складання».

Посилання на використані джерела наводяться у тексті статті у квадратних дужках […] із зазначенням порядкового номеру бібліографічного посилання у списку.

Редакційна колегія залишає за собою право вносити зміни в рукопис редакційного характеру.

Телефон для довідок: (044) 256 – 22 – 37.

Адреса сайту збірника: https://journal.viti.edu.ua/index.php/cicst/issue/archive.

Електронна адреса для надання статей: naukaviti@viti.edu.ua, naukaviti@gmail.com.

Відповідальний за випуск *В. В. Куцаєв*

Підписано до друку 10.06.2024. Зам. 139.
Друк. арк. 30,00. Ум.-друк. арк. 27,9. Обл.-вид. арк. 25,94.
Формат паперу 60×84/8. Тираж 50 прим.

Адреса друкарні ВІТІ імені Героїв Крут: 01011, м. Київ, вул. Князів Острозьких, 45/1