MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

KHARKIV NATIONAL
UNIVERSITY OF RADIO ELECTRONICS

RADIOTEKHNIKA

All-Ukrainian
interdepartmental scientific and technical collection

ISSN 0485-8972
eISSN 2786-5525

Founded in 1965

I S S U E 2 2 2

Kharkiv
Kharkiv National
University of Radio Electronics
2025

UDC 621.3

The collection is included in the List of scientific professional publications of Ukraine, category «Б», technical and physical-mathematical sciences (approved by orders of the Ministry of Education and Science from 17.03.2020 № 409; from 02.07.2020 № 886; from 24.09.2020 № 1188) by specialties: 105 – Applied Physics and Nanomaterials; 125 – Cybersecurity and information protection; 163 – Biomedical Engineering; 171 – Electronics; 172 – Electronic communications and Radio Engineering; 173 – Avionics; 174 –Automation and Computer-Integrated Technologies and Robotics; 175 – Metrology and information-measuring technique; 176 – Micro- and Nanosystem Technology.

Website: rt.nure.ua

Registration certificate KV № 12098-969 PR dated 14. 12. 2006.

The authors are responsible for the content of the article.

Editorial Team

Members of the editorial board of foreign scientific institutions and educational institutions

Responsible for the issue: *S.O. Sheiko, PhD, Assoc. prof., I.D. Gorbenko, Dr. Sc. (Tech.), prof.*
Technical Secretary: *O.S. Polyakova.*

Recommended by the Scientific and Technical Council of Kharkiv National University of Radio Electronics, protocol № 7 dated 18.09.2025.

Address of the editorial board: Kharkiv National University of Radio Electronics (NURE), ave. Nauky, 14, Kharkiv, 61166, tel. (0572) 7021-397.

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

РАДІОТЕХНІКА

Всеукраїнський
міжвідомчий науково-технічний збірник

ISSN 0485-8972
eISSN 2786-5525

Засновано в 1965 р.

В И П У С К 2 2 2

Харків
Харківський національний
університет радіоелектроніки
2025

УДК 621.3

Збірник включено до Переліку наукових фахових видань України, категорія "Б", технічні та фізико-математичні науки (затверджено наказами МОНУ від 17.03.2020 № 409; від 02.07.2020 № 886; від 24.09.2020 № 1188) за спеціальностями: 105 – Прикладна фізика та наноматеріали; 125 – Кібербезпека та захист інформації; 163 – Біомедична інженерія; 171 – Електроніка; 172 – Електронні комунікації та радіотехніка; 173 – Авіоніка; 174 – Автоматизація, комп'ютерно-інтегровані технології та робототехніка; 175 – Метрологія та інформаційно-вимірювальні технології; 176 – Мікро- та наносистемна техніка.

Сайт: rt.nure.ua

Реєстраційне свідоцтво КВ № 12098-969 ПР від 14. 12. 2006.

За зміст статті відповідальні автори.

Відповідальні за випуск: *С.О. Шейко, канд. техн. наук, доц., І.Д. Горбенко, д-р техн. наук, проф.*
Технічний секретар: *О.С. Полякова.*

Рекомендовано Науково-технічною радою Харківського національного університету радіоелектроніки, протокол № 7 від 18.09.2025.

Адреса редакційної колегії: Харківський національний університет радіоелектроніки (ХНУРЕ), просп. Науки, 14, Харків, 61166, тел. (0572) 7021-397.

CONTENT

SYSTEMS AND METHODS OF INFORMATION PROTECTION

RADIO ELECTRONIC SYSTEMS

ELECTRONIC COMMUNICATIONS

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS

BIOMEDICAL ENGINEERING

ЗМІСТ

СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

РАДІОЕЛЕКТРОННІ СИСТЕМИ

ЕЛЕКТРОННІ КОМУНІКАЦІЇ

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

БІОМЕДИЧНА ІНЖЕНЕРІЯ

УДК 004.056.5 **DOI:10.30837/rt.2025.3.222.01**

І.Д. ГОРБЕНКО, д-р техн. наук, О.Г. КАЧКО, канд. техн. наук, Я.А. ДЕРЕВ'ЯНКО

ОПТИМІЗАЦІЯ ОПЕРАЦІЙ ОБЧИСЛЕННЯ ТА ПЕРЕВІРКИ ЦИФРОВОГО ПІДПИСУ ДЛЯ СТАНДАРТУ FIPS 205. 2 ЧАСТИНА

Вступ

В [1, 2] показано методи та результати оптимізації базових операцій на основі алгоритмів SHA та SHAKE. Показано, що загальні алгоритми генерації ключів, вироблення та перевірки електронного підпису складаються з послідовних кроків, кожний з яких застосовує результат попереднього кроку, що виключає можливість застосування паралельних обчислень для цих алгоритмів.

Автори алгоритму Sphincs [3], який покладено в основу стандарту FIPS 205, запропонували оптимізовану версію алгоритму без застосування AVX команд (папка Optimized_Implementationta з застосуванням AVX (папка Additional_Implementations). Можливість застосування паралельного виконання за рахунок потоків не розглядається.

В даній роботі розглядаються засоби та результати оптимізації, в тому числі за рахунок паралельних потоків при реалізації окремих кроків алгоритмів. Оптимізація за рахунок застосування операцій AVX не розглядається.

Для виміру результатів оптимізації в якості базового застосовують реалізацію, надану в [3], папка Additional_Implementations. Для порівняння реалізації авторів Стандарту та авторів статті виконано на комп'ютері: процесор 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz, OS Windows 11, Version 24H2, Microsoft Visual Studio Community 2019, Version 16.11.21.

Для завдання результатів застосовують прискорення, тобто відношення часу виконання до оптимізації та часу виконання після оптимізації. Час виконання вимірюється в тактах процесора. Прискорення несуттєво залежить від криптостійкості, тому при викладанні результатів для окремих схем будуть наводитись результати тільки для максимального рівня криптостійкості, кінцеві результати для функцій генерування ключів, вироблення та перевірки електронного підпису будуть наведені для усіх рівнів криптостійкості.

Для зручності застосування імена функцій співпадають з іменами зі Стандарту [1].

Нагадаємо, що в якості відкритого ключа PK застосовують компоненти:

PK_SEED – seed для відкритого ключа;

PK_ROOT – корінь дерева.

В якості секретного ключа SK застосовують компоненти:

SK_SEED – seed для секретного ключа;

SK_PRF – компонент для генерації дайджеста повідомлення;

компоненти відкритого ключа.

В якості базових операцій застосовують функції PRF, Tl, H, F, оптимізація яких розглянуто в попередній статті.

Далі розглянуто засоби та результати оптимізації для окремих схем.

1. Оптимізація функцій для роботи з одноразовим підписом (Winternitz One-Time Signature, WOTS)

Функції даного модуля застосовують при генерації ключів, електронного підпису та для його перевірки. В якості параметрів цей модуль застосовую параметри:

n – довжина в байтах повідомлення для формування одноразового підпису, компонентів

секретного та відкритого ключів. В Стандарті *n* приймає значення 16, 24 та 32 в залежності від рівня криптостійкості (2, 3, 5 відповідно);

lgw – параметр Winternitz, довжина ланцюжка в бітах, визначає граничні значення цілих чисел, якими кодується повідомлення для підпису (lgw = 4, w = 16, граничні значення 0..15). Ці цілі числа (позначимо їх s[i]) і є секретними ключами для одноразового підпису. Кількість таких чисел в стандарті дорівнює len = 2n + 3. Вибір параметру lgw = 4 значно спрощує формування секретного ключа, кожному байту вхідного повідомлення відповідають 2 цілих числа: старші 4 біта – перше число, молодші – друге, саме такий спосіб обчислення застосовують далі

1.1. Генерація електронного підпису. Функція wots_sign Стандарту

Функцію застосовують при генерації електронного підпису. Підпис складається з len послідовностей, кожна послідовність завдовжки n байтів. Одному цілому числу відповідає один компонент електронного підпису, тобто підпис складається з len компонентів. Псевдокод для алгоритму обчислення компонентів підпису представлено на рис. 1.

```
Вхід.
M – повідомлення для підпису завдовжки n байтів;
SK_SEED – компонент SEED секретного ключа;
PK_SEED – компонент SEED відкритого ключа;
ADR – поточна інформація про дерево.
Вихід.
WOTS_SIG – масив з len байтових послідовностей завдовжки n байтів кожна.
1 Генерація масиву s цілих чисел завдовжки len чисел.
2 for i := 0 to len do                          Цикл генерації секретних ключів для підпису
3       TADRS :=ADR;                            Адресна структура для секретного ключа
        TADRS.type:= WOTS_PRF
        TADRS. KeyPairAddress = ADR. KeyPairAddress
        TADRS.ChainAddress := i
4       sk[i]:= PRF (PK_SEED || skADRS || SK_SEED)    Обчислення ключа
5 end for
6 for i := 0 to len do                          Цикл генерації компонентів підпису
7         TADRS :=ADR;                          Адресна структура для підпису
          TADRS.ChainAddress := i
8         tmp:= sk [i];
9         for j:= i to i + s[i] do              Ланцюжок гешів завдовжки s[i]
10              TADRS.HashAddress := j;
11              tmp  = F (PK_SEED || TADRS|| tmp)
12        end for
13        WOTS_SIG [i] = tmp
14 end for
```

Рис. 1. Функція wots_sign. Псевдокод

Алгоритм оптимізовано за рахунок паралельного виконання циклів, обмежених рядками 2 – 5 та 6 – 14, а також за рахунок більш ефективного виконання функцій PRF та F, що визначено в роботі [2].

1.2. Генерація відкритого ключа по електронному підпису. Функція wots_pkFromSig Стандарту

Функцію застосовують для перевірки електронного підпису, який складається з len рядків завдовжки n байтів. Результатом роботи функції є рядок байтів завдовжки n байтів. При генерації підпису обчислювався ланцюжок гешів s[i] разів (цикл, обмежений рядками

9 – 12). Для відновлення відкритого ключа обчислюється ланцюжок гешів w–s[i]–1 разів, що забезпечує загальну кількість разів, яка залишається постійною і дорівнює w–1.

Псевдокод функції представлено на рис. 2.

```
Вхід.
        WOTS_SIG –підпис, масив з len послідовностей завдовжки n байтів кожна.
        M – повідомлення для підпису завдовжки n байтів;
        PK_SEED – компонент SEED відкритого ключа;
        ADR – поточна інформація про дерево.
Вихід.
        pk – відкритий ключ, рядок байтів завдовжки n байтів.
        1 Генерація масиву s цілих чисел завдовжки len чисел.
        2 for i := 0 to len do                    Цикл відновлення елементів відкритого ключа
                  TADRS :=ADR;                          Адресна структура для відновлення
                  TADRS.ChainAddress := i
        3         tmp [i]:= WOTS_SIG [i];
        4         for j:= s[i] to w – s[i] – 1 do        Ланцюжок гешів завдовжки s[i]
        5                 TADRS.HashAddress := j;
        6                 tmp[i] := F (PK_SEED || TADRS|| tmp [i])
        7          end for
        8 end for
        9 TADRS :=ADR;                            Адресна структура для відновлення
        10 TADRS.type := WOTS_PK
        11 TADRS.KeyPairAddress = ADR. KeyPairAddress
        12     := Tl(PK.seed, wotspkADRS, tmp)          Загальний ключ
```

Рис. 2. Функція wots_pkFromSig. Псевдокод

Для оптимізації цикл, обмежений рядками 2 – 8, виконується паралельно, застосовується оптимізована функція Tl [2].

1.3. Обчислення відкритого ключа по секретному ключу. Функція wots_pkGen Стандарту

В попередніх двох функціях для кожного з len значень s[i] спочатку виконується ланцюжок обчислення гешів s[i] разів (першій алгоритм), а потім цей ланцюжок продовжується w–s[i]–1 разів. Тобто загальна довжина ланцюжка дорівнює w–1 незалежно від значення s[i]. Наступний алгоритм також обчислює відкритий ключ за допомогою ланцюжка завдовжки w–1. Таким чином, значення відкритого ключа може бути обчислено як за допомогою секретного ключа, що виконується саме цим алгоритмом і відновлено по підпису, що виконується при перевірці підпису. Псевдокод алгоритму наведено на рис. 3.

```
Вхід.
        SK_SEED – компонент SEED секретного ключа;
        PK_SEED – компонент SEED відкритого ключа;
        ADR – поточна інформація про дерево.
Вихід.
        pk – відкритий ключ, рядок байтів завдовжки n байтів.
1       for i:=0 to len do                        Цикл для секретних ключів
2                 TADRS :=ADR;                        Адресна структура для обчислення
                  TADRS.Type = WOTS_PRF
                  TADRS. KeyPairAddress = ADR. KeyPairAddress
                  TADRS. ChainAddress:= i
3                 sk [i] = PRF (PK_SEED ||  TADRS || SK_SEED)
4        end for
5        for i:=0 to len do                       Цикл для обчислення компонентів ключа
6                 TADRS :=ADR;                        Адресна структура для обчислення
                  TADRS. ChainAddress:= i
7                 for j:= 0 to w – 1 do                Ланцюжок гешів завдовжки w–1
8                         TADRS.HashAddress := j;
```

```
9                    tmp[i] := F (PK_SEED || TADRS|| sk[i])
10               end for
11           end for
12           TADRS :=ADR;                          Загальний ключ. Адресна структура
             TADRS.type = WOTS_PK
             TADRS.KeyPairAddress = ADRS.KeyPairAddress
13           pk:= Tl(PK_SEED, TADRS, sk, len)      Загальний ключ
```

Рис. 3. Функція wots_ pkGen. Псевдокод

Як і в попередніх алгоритмах цикли, обмежені рядками 1 – 4 та 5 – 11, виконуються паралельно. Застосовується оптимізована функція Tl [2]

1.4. Результати оптимізації функцій для роботи з одноразовим підписом

Результати оптимізації наведено в табл. 1.

Ім'я режиму включає алгоритм для обчислення гешу (SHAKE або SHA) та ознаку режиму оптимізації (пам'ять -s, час – t). Ім'я функції після оптимізації доповнюється символом _ в кінці, наприклад, wots_sign та wots_sign_. Для кожної функції визначається прискорення, наприклад S (wots_sign), яке визначається відношенням значень в попередніх двох рядках.

Таблиця 1

Одноразовий підпис

	SHAKE256s	SHA256s	SHAKE256f	SHA256f
wots_sign	898991	448186	730929	454118
wots_sign_	376800	190132	314923	243752
S (wots_sign)	2.39	2.36	2.32	1.86
wots_pkFromSig	928681	490984	799897	447234
wots_pkFromSig_	393050	197531	331191	223814
S(wots_pkFromSig)	2.36	2.49	2.42	2.00
wots_pkGen	1857485	1027216	1775210	909103
wots_pkGen_	588189	383740	506253	294956
S (wots_pkGen)	3.16	2.68	3.5	3.08

Прискорення практично для усіх функцій 2 та більше 2.

2. Оптимізація функцій для роботи з розширеним деревом Мерклі (eXtended Merkle Signature Scheme (XMSS))

Це розширення дозволяє застосовувати одноразовий підпис для підпису декількох повідомлень.

Дерево Мерклі – двійкове дерево, в якому на нижньому рівні в якості вузлів знаходяться листи, пара яких об'єднуються в один вузол для наступного рівня.

Параметри.

h' – висота дерева. Визначає геометричні параметри дереву, а саме кількість його листів leafs = $2^{h'}$, кожний лист дерева розглядається як відкритий ключ, для перевірки якого можна застосовувати один ключ, який є коренем дерева. Кожний ключ можна застосовувати для підпису одного повідомлення, тобто схему можна застосовувати для $2^{h'}$ повідомлень.

2.1. Генерування дерева Мерклі. Функція xmss_node

Відповідна функція рекурсивна, для нижнього рівня обчислює геш листа, наступні рівні виконують обчислення гешу для пари вузлів. Кінцевий результат – корінь дерева. Псевдокод функції наведено на рис. 4.

```
Вхід.
        SK_SEED – компонент SEED секретного ключа;
        PK_SEED – компонент SEED відкритого ключа;
        ADR – поточна інформація про дерево.
         – номер вузла;
         – номер рівня.
Вихід.
        PK_ROOT – корінь дерева.
        1 if z = 0 then                  Це лист. Обчислюємо відповідний pk
        2          ADR.type =  WOTS_HASH; ADR. KeyPairAddress = i
         3         wots_pkGen_(PK_ROOT, SK_SEED, ADR)
         4 else                          Обробка пари листів
         5          lnode :=xmss_node_(SK_SEED, PK_SEED, ADR, 2 * i, z – 1); Лівий
         6          rnode:=xmss_node_(SK_SEED, PK_SEED, ADR, 2 * i + 1, z – 1); Правий
         7          ADRS.Type :=TREE; ADRS.TreeHeight :=z; ADRS.TreeIndex :=i;
         8         Tmp [0]:= lnode; Tmp [1]:= rnode; Root:= H(PK_SEED, ADR, Tmp);
         9 end if
```

Рис. 4. Функція xmss_node. Псевдокод

Оптимізація функції виконується за рахунок застосування оптимізованих варіантів функції wots_pkGen, H.

2.2. Генерування підпису. Функція xmss_sign

Підпис – це WOTS підпис, який відповідає одному з шляхів по дереву Мерклі, починаючи з рівня 0 до кореня дерева. По номеру вузла idx визначається номер додаткового вузла, який відповідає парі. Для кожного рівня, крім самого верхнього містить додатково значення вузла пари, який дозволить перейти на наступний рівень. Додаткові вузли формують шлях аутентифікації. Довжина підпису дорівнює довжині відповідного WOTS підпису та довжині відповідного шляху аутентифікації PATH для рівнів 0, 1, h’ – 1, тобто (len * n) + (h’ * n). Псевдокод функції наведено на рис. 5.

```
Вхід.
        M – повідомлення для підпису, рядок байтів завдовжки n байтів
        SK_SEED – компонент SEED секретного ключа;
        PK_SEED – компонент SEED відкритого ключа;
ADR – поточна інформація про дерево.
         dx – номер вузла;
Вихід.
        XMSS_SIG – підпис, XMSS_SIG = WOTS_SIG || PATH
1 Визначення адреси для підпису і для шляху аутентифікації
        P_XMSS_SIG = XMSS_SIG; P_PATH = P_XMSS_SIG + len * n
2 Для усіх рівнів крім останнього визначення номера додаткового вузла і обчислення його
значення
        for j := 0 to h’ do
                k:= (ind / 2^j) mod 1              Номер суміжного вузла
                xmss_node_( P_PATH[j], SK_seed, k, j, PK_seed, PK_seed_n, adr);
        end for
3 Формування адресної структури
        ADR.Type = WOTS_HASH;    ADR. KeyPairAddress = idx
4 Обчислення WOTS_SIG
        wots_sign_( WOTS_SIG, M, SK_SEED, PK_SEED, ADR);
```

Рис. 5. Функція xmss_sign. Псевдокод

Оптимізація функції xmss_sign виконується за рахунок:

- попереднього визначення адрес початку підпису і шляху аутентифікації, що попереджує копіювання даних значних розмірів;
- застосування оптимізованих версій функцій xmss_node_, wots_sign_

2.3. Обчислення відкритого ключа по підпису. Функція xmss_pkFromSig

Фактично ця функція обчислює корінь дерева Мерклі за підписом.

Для обчислення відкритого ключа для нульового рівня застосовують функцію wots_pkFromSig, яка обчислює відкритий ключ, який відповідає WOTS_SIG. Цей ключ застосовують як лист для дерева Мерклі. Для обчислення кореня цього дерева застосовують шлях аутентифікації PATH з підпису (рис. 6).

Оптимізацію функції виконують за рахунок застосування виключення операції копіювання окремих полів WOTS та застосування оптимізованих функцій для обчислення листа дерева Мерклі (функція wots_pkFromSig_) та функції H.

```
Вхід.
        WOTS_SIG – підпис, рядок байтів завдовжки (len * n) +  (h' * n)
        M – повідомлення для якого підпис, рядок байтів завдовжки n байтів
        PK_SEED – компонент SEED відкритого ключа;
        ADR – поточна інформація про дерево.
         dx – номер вузла, для якого підпис;
Вихід.
        root – корень відповідного дерева
        1 Визначення адреси для підпису і для шляху аутентифікації
                P_XMSS_SIG = XMSS_SIG; P_PATH = P_XMSS_SIG + len * n
        2 Формування інформаційної структури для WOTS ключа
        ADR.Type:= WOTS_HASH; ADR. KeyPairAddress := idx
        3 Обчислення відкритого ключа WOTS  по підпису
        node:= wots_pkFromSig_( P_XMSS_SIG, M, PK_SEED, ADR);
        4 Формування інформаційної структури для  роботи з деревом
        ADR.Type = TREE; ADR. TreeIndex:= idx
        5 for k:= 0  to h' do     Для усіх рівнів дерева Мерклі
        6         ADR.TreeHeight:= k + 1        Коректування інформаційної структури
        7         if  idx / 2^k парне                Пошук пари та обчислення вузла
        8                 ADR. TreeIndex := ADR. TreeIndex/2
        9                 node:=H(PK_seed, ADR, node, P_PATH [k]);
        10      else
        11                ADR. TreeIndex := (ADR. TreeIndex – 1)/2
        12                node:=H(PK_seed, ADR, P_PATH [k], node);
        13      endif
        14 end for
        15 root:=node
```

Рис. 6. Функція xmss_pkFromSig. Псевдокод

2.4. Результати оптимізації функцій для розширеного дерева Мерклі

Функції для розширеного дерева Мерклі в явному вигляді в реалізації Sphincs не застосовують, тому в табл. 2 наведено тільки дані для реалізації авторів. Результати для розширеного дерева Мерклі наведено в табл. 2.

Таблиця 2

Розширене дерево Мерклі

	SHAKE256s	SHA256s	SHAKE256f	SHA256f
xmss_node_	53125445	68369655	3183055	1951479
xmss_sign_	100795718	68369655	5889182	3812993
xmss_pkFromSig_	258114	157727	239068	157114

3. Оптимізація функцій для роботи з гіпердеревом (Hypertree)

Гіпердерево є двійковим деревом, кожний вузел якого є розширеним деревом Мерклі. Висота цього дерева визначається параметром d (кількість рівнів дерев Мерклі). Так як висота дерева Мерклі дорівнює h', загальна висота гіпердерева дорівнює h = d * h'. На останньому рівні з номером d-1 знаходиться одне дерево Мерклі, корінь якого співпадає з PK_ROOT, тому для останнього рівня корінь дерева не обчислюють. На попередньому рівні знаходяться 2 дерева, на рівні з номером 0 знаходиться 2^d або $2^{h-h'}$ дерев Меркли

3.1. Генерація підпису для гіпердерева. Функція ht_sign

Псевдокод для генерації підпису для гіпердерева (рис. 7).

```
Вхід.
        M – повідомлення для підпису;
        SK_seed – seed для секретного ключа;
        PK_seed – seed для відкритого ключа;
        tree_idx – індекс дерева;
        leaf_idx –індекс листа.
Вихід.
    HT_SIG – підпис для гіпердереві
                                                        Рівень 0 гіпердерева
    1 ADRS:=0; ADRS.TreeAddress:= TreeIdx               Інформаційна структура
    2 HT_SIG:= xmss_sign (M, SK_sign, PK_seed, ADR, leaf_idx);  Підпис для вузла дерева
    3 root:= xmss_pkFromSig (HT_SIG, M, PK_seed, ADRS, leaf_idx)   Корінь дерева
    4 for j = 1 to d do                                 Цикл для решти рівнів
    5       leaf_idx := tree_idx mod 2^h'                   Індекс листа
    6       tree_idx:= tree_idx / 2^h'                      Індкс дерева
    7       ADRS.LayerAddress:= j                           Інформаційна
    8       ADRS.TreeAddress:= tree_idx                       структура
    9       SIG:=xmss_sign (root, SK_sign, PK_seed, ADRS, leaf_idx);  Підпис
    10      if j ≠ d – 1 then                       Корень для неостаннього рівня
    11              root:= xmss_pkFromSig (SIG, root, PK_seed, ADRS, leaf_idx)
    12      end if
    13      HT_SIG:= HT_SIG || SIG      Підпис для HT дерева
    14 end for
```

Рис. 7. Функція ht_sign. Псевдокод

Підпис для гіпердерева складається з підписів дерев Мерклі для кожного рівня і має довжину d * XMSS_SIG = d * ((len * n) + (h' * n)).

В якості повідомлення M для рівня 0 застосовують відкритий ключ, сформований для лісу дерев, в якості індексу дерева (tree_idx) та індексу листа (leaf_idx), застосовують значення, які визначаються дайджестом повідомлення, що підписується. Для решти рівнів ідентифікаторами листів є наступні h' бітів індексу дерева tree_idx, корінь дерева попереднього рівня застосовують для обчислення підпису для наступного рівня.

Оптимізація функції виконується за рахунок застосування оптимізованих версій функцій xmss_sign, xmss_pkFromSig

3.2. Перевірка підпису для гіпердерева. Функція ht_verify

Псевдокод для перевірки підпису для гіпердерева (рис. 8).

Для перевірки підпису спочатку виконується ініціалізація адресної структури та початкових значень адреси підпису її довжини (кроки 1 – 3), а далі обчислюється відкритий ключ по цифровому підпису (функція xmss_pkFromSig, крок 4). Виконується d кроків формування кореня та порівняння кореня для останнього рівня з компонентом PK_root відкритого ключа.

Для оптимізації функції ht_verify застосовують оптимізовані версії функції xmss_pkFromSig.

```
Вхід.
        M – повідомлення для підпису;
        HT_SIG – підпис для гіпердерева;
        PK_seed – seed для відкритого ключа;
        PK_root – root для відкритого ключа;
        tree_idx – індекс дерева;
        leaf_idx –індекс листа.
Вихід.
        Success (OK в разі успіху та ERROR в разі помилки)
1 Формування інформаційної структури        Рівень 0 гіпердерева
        ADRS:=0; ADRS.TreeAddress:= TreeIdx
2 xmss_sign_len:= (len * n) +  (h' * n)                    Розмір підпису
3 SIG:= SubST (HT_SIG, 0, xmss_sign_len)                   Підпис
4 pk:= xmss_pkFromSig (SIG, M, PK_seed, ADRS, leaf_idx)    pk
5 for j = 1 to d do                                 Решти рівнів
6       SIG:= SubST (HT_SIG, j * xmss_sign_len, xmss_sign_len)  Підпис
7       leaf_idx := tree_idx mod 2^h'               Індекс листа
8       tree_idx:= tree_idx / 2^h'
9       ADRS.LayerAddress:= j                       Інформаційна
10      ADRS.TreeAddress:= tree_idx                 структура
11      pk:= xmss_pkFromSig (SIG, pk, PK_seed, ADRS, leaf_idx)   pk
12 end for
13 if pk = PK_ROOT then success = OK else success = ERROR endif
```

Рис. 8. Функція ht_verify

3.3. Результати оптимізації функцій для гіпердерева

Результати оптимізації для функцій ht_sign, ht_verify наведено в табл. 3.

Таблиця 3

Результати оптимізації функцій ht_sign, ht_verify

	SHAKE256s	SHA256s	SHAKE256f	SHA256f
ht_sign	3625906354	2108292458	490611920	285012219
ht_sign_	900021510	551139349	114492525	70366362
S (ht_sign)	4.03	3.83	4.29	4.05
ht_verify	7777288	4134225	15733905	8917948
ht_verify_	2097061	1344683	3894410	2619106
S (ht_ verify)	3.71	3.07	4.04	3.40

Прискорення для усіх функцій більше ніж в три рази.

4. Ліс дерев (Forest of Random Subsets (FORS))

FORS – багаторазовий підпис для дайджеста повідомлення, який формується безпосередньо з повідомлення для підпису і, можливо, випадкового рядка, застосовує параметри:

a – висота дерева Мерклі, саме з цих дерев складається ліс;

$t = 2^a$ – кількість листів в дереві Мерклі;

k – кількість дерев Мерклі в лісі. Кожному дереву відповідає множина особистих ключів, кількість яких визначається кількістю листів (t) дерева Мерклі. Загальна довжина особистих ключів дорівнює t * n байтів. Ці рядки псевдовипадково генеруються з SK_SEED.

4.1. Генерація секретного ключа. Функція fors_skGen

Псевдокод функції fors_skGen наведено на рис. 9.

```
Вхід.
        SK_seed – seed для секретного ключа;
        PK_seed – seed для відкритого ключа;
        ADRS – інформаційна структура;
        idx – індекс ключа.
Вихід.
        sk – рядок байтів завдовжки n
1 Формування інформаційної структури
        SK_ADRS := ADRS
        SK_ADRS.Type = FORS_PRF
        SK_ADRS.KeyParaAddress := ADRS. KeyParaAddress
        SK_ADRS.TreeIndex:= idx
2 Обчислення sk
        Sk:= PRF (PK_seed, SK_seed, SK_ADRS)
```

Рис. 9. Функція fors_skGen. Псевдокод

Оптимізація функції виконується за рахунок застосування оптимізованого варіанту функції PRF.

Для обчислення секретного ключа для усіх вузлів можна застосовувати паралельне виконання, але в даному випадку навантаження на кожну гілку буде недостатнім для отримання виграшу від паралелізму.

4.2. Генерація дерева Мерклі для лісу. Функція fors_node

Псевдокод функції fors_node показан на рис. 10.

Як і попередні функції генерації дерева Мерклі функція рекурсивна. Якщо рівень дорівнює 0, то функція обчислює лист дереву – відповідний секретний ключ (fors_skGen).

Якщо рівень більше 0, то виконується рекурсивний виклик функції для попереднього рівня z – 1 і номера вузла 2 * i, що відповідає лівому вузлу (lnode) та номера вузла 2 * i + 1, що відповідає правому вузлу (rnode). Після завершення рекурсивного виклику будуть отримані значення для лівого та правого вузлів, для яких обчислюється об’єднане значення (функція H).

4.3. Генерація підпису. Функція fors_sign

Псевдокод функції fors_ sign наведено на рис. 11.

По заданому рядку байтів *md* функція формує масив цілих чисел *s* завдовжки *k* чисел, під кожне число виділяються наступні *a* бітів.

```
Вхід.
        SK_seed – seed для секретного ключа;
        PK_seed – seed для відкритого ключа;
        ADRS – інформаційна структура;
          – номер вузла;
          – номер рівня.
Вихід.
        root – корінь дерева, рядок байтів завдовжки n
1 if z = 0 then
2       sk:= fors_skGen (SK_seed, PK_Seed, ADRS, i)
3       ADRS.TreeHeight = 0; ADRS. TreeIndex = i
4       node:= F(PK_seed, ADRS, sk);
5 else
6       lnode := fors_node (SK_seed, PK_seed, ADRS, 2 * i, z – 1);
7       rnode := fors_node (SK_seed, PK_seed, ADRS, 2 * i + 1, z – 1);
8       ADRS.TreeHeight = z; ADRS.TreeIndex = i
9       node:= H (PK_seed, ADRS, lnode || rnode);
10 end if
11 return node
```

Рис. 10. Функція fors_node

```
Вхід.
        SK_seed – seed для секретного ключа;
        PK_seed – seed для відкритого ключа;
        ADRS – інформаційна структура;
        md – рядок байтів завдовжки ── байтів
Вихід.
        FORS_SIG – підпис, рядок байтів завдовжки k * (n + a * n)
1 Формування масиву s
2 for i = 0 to k do                                  Цикл для усіх чисел масиву s
3       L_ADRS := ADRS
4       pSK:= FORS_SIG + i * (1 + a) * n             Адреса для sk
5       pAuth:= pSK + n                              Адреса для Auth
6       pSK  := fors_skGen (SK_seed, PK_seed, i * t + s[j])
7       for j = 0 to a do                            for AUTH
8               r:= ─                                Номер суміжного вузла
9               pAuth []j] := fors_node ( SK_seed, PK_seed, ADRS,  r)
10      end for
12  end for
```

Рис. 11. Функція fors_sign. Псевдокод

Для формування підпису виконується цикл k раз, в якому для кожного числа з масиву *s* формується секретний ключ (функція fors_skGen).

Адреса для секретного ключа (pSK) визначається згідно з номером ітерації, за рахунок цього забезпечено можливість виконання циклу паралельно. Після генерації секретного ключа виконується цикл для визначення гешів відповідних пар для кожного з *a* рівнів.

Структура підпису для i-го значення з масиву s:

pSK

pAuth [0], pAuth [1], … pAuth [a – 1]

Оптимізація функції за рахунок паралельного виконання зовнішнього циклу і застосування оптимізованого варіанту для функцій fors_skGen, fors_node.

4.4. Генерація відкритого ключа з підпису. Функція fors_pkFromSig

Псевдокод функції наведено на рис. 12.

```
Вхід.
        FORS_SIG – підпис завдовжки k * (1 + a)*n
        PK_seed – seed для відкритого ключа;
        ADRS – інформаційна структура;
        md – рядок байтів завдовжки ── байтів
Вихід.
        FORS_PK – відкритий ключ згідно з підписом (рядок байтів завдовжки n)
1 Формування масиву s цілих чисел з рядка md.
2 for i = 0 to k do                  Цикл для усіх чисел масиву s
3       L_ADRS := ADRS; ind:= s [i]
4       pSK:= FORS + i * (1 + a) * n;   pAuth:= pSK + n Адреси sk,  Auth; sk:= pSK[i]
5       adr_l.TreeHeight:= 0; adr_l.TreeIndex:= i * (1 << A) + ind
6       pnode0:= F(PK_seed, adr_l, pSK)
7       for j = 0 to a do                    for AUTH
8               adr_l. TreeHeight:= j + 1; ti:= adr_l. TreeIndex
9               if ind / 2^j парне then
10                      adr_l. TreeIndex:= ti / 2
11                      p[0] := pnode0; p[1] := pAuth[j]
```

```
12              else
13                      adr_l. TreeIndex := (ti – 1) / 2;
14                      p[1] := pAuth[j];  p[0] := pnode0
15              end if
16              pnode0:= H (PK_seed, adr_l, p);
17      end for
18      root [i]:= pnode0
19 end for
20 tadr:= ADRS; tadr.type:= FORS_ROOTS; tadr.KeyPair:= adr.KeyPair.
21 FORS_PK:=Tl (PK_seed, tadr, root, k)
```

Рис. 12. Функція fors_pkFromSig. Псевдокод

Після формування масиву цілих *s* з *k* чисел (крок 1) виконується цикл для кожного з цих чисел (крок 2). Для кожної ітерації циклу встановлюється в якості початкової адресна структура, яка задається в списку параметрів (крок 3). Визначаються адреси для поточного секретного ключа та шляху аутентифікації, а також поточний секретний ключ (крок 4).

Згідно з секретним ключем обчислюють вузол для рівня 0 (pnode0) (кроки 5, 6), а далі згідно зі шляхом аутентифікації (кроки 7 – 18) – корінь для відповідного дерева Меркли (root [i]).

По значенням root [i] для *k* значень обчислюється загальний відкритий ключ (кроки 20, 21).

Саме відкритий ключ, сформований функцією fors_pkFromSig, застосовують в якості повідомлення для підпису в гіпердереві (дивись функції для гіпердерева).

Для оптимізації застосовують паралельне виконання зовнішнього циклу та оптимізовані внутрішні функції.

4.5. Результати оптимізації функцій для лісу

Функції fors_skGen та fors_node займають порівняльно менше часу, ніж решта функцій цієї групи, тому далі наведено результати тільки для останніх двох функцій.

Результати оптимізації для функцій fors_sign та fors_pkFromSig наведено в табл. 4

Таблиця 4

Результати оптимізації функцій fors_sign та fors_pkFromSig

	SHAKE256s	SHA256s	SHAKE256f	SHA256f
fors_sign	1721324445	1247174926	96257811	72954353
fors_sign__	380254469	263020840	18306279	13198027
S (fors_sign)	4.53	4.74	5.26	5.53
fors_pkFromSig	489274	534699	758640	603552
fors_pkFromSig__	134006	109762	132002	110104
S (fors_node)	3.65	4.87	5.75	5.48

Прискорення для функцій змінюється в діапазоні 3.65 – 5.75.

5. Комплексні функції

Розглядаються функції генерування ключів, вироблення та перевірки електронного підпису, які застосовують функції для усіх схем, що розглянуті вище.

5.1. Генерація ключів. Функція slh_keygen_internal

Псевдокод наведено на рис. 13.

Функція приймає в якості вхідних даних випадкові значення SK_seed, SK_prf та PK_seed і формує значення PK_root як корінь дерева Меркле висотою *d* (кроки 1 – 3).

Компоненти відкритого та секретного ключа записуються в рядки PK, SK відповідно (кроки 4 – 5).

Вхід.
SK_seed – seed для секретного ключа;
SK_prf – prf для секретного ключа;
PK_seed – seed для відкритого ключа.
Вихід.
SK – секретний ключ завдовжки 4 * n байтів;
PK – відкритий ключ завдовжки 2 * n байтів.
1. ADRS:= 0 Формування інформаційної структури
2. ADRS.LayerAddress:= d – 1
3. PK_root:= xmss_node (SK_seed, PK_seed, ADRS, 0, h')
4. PK:= PK_seed || PK_root
5. SK:= SK_seed || SK_prf || PK

Рис. 13. Функція slh_keygen_internal. Псевдокод

5.2. Вироблення електронного підпису. Функція slh_sign_internal

Електронний підпис складається:

- з псевдовипадкового масиву байтів R завдовжки n байтів, який залежить від повідомлення для підпису, секретного ключа, і, можливо, випадкового рядка байтів, наявність якого робить підпис неконстантним;
- підпису для лісу;
- підпису для гіпердерева.

Псевдокод для функції slh_sign_internal (рис. 14).

Вхід.
MSG – повідомлення для підпису;
MSG_len –довжина повідомлення;
SK –секретний ключ;
addrnd – визначає, чи застосовується константний підпис, = PK_SEED для константного підпису та випадковому рядку завдовжки n байтів, якщо підпис неконстантний.
Вихід.
SIG – підпис
SIG_len[1] –довжина підпису
1 R:= PRF (SK_prf || addrnd) Компонент підпису R
2 digest := HMsg(R, PK_seed, PK_root, MSG, MSG_len) Дайджест повідомлення
3 md, idxtree, idxleaf,:= DigestParse(digest)
4 ADR:=0 Ініціалізація ADR
5 ADR. TreeAddress:= idxtree
6 ADR.Type:= FORS_TREE
7 ADR. KeyPairAddress:= idxleaf
Підпис для обраного дерева лісу та його відкритий ключ
8 SIG_FORS := fors_sign (R, md, SK_seed, PK_seed, ADR)
9 PK_fors := fors_pkFromSig (SIG_FORS, md, PK_seed, ADR);
Підпис для гірердерева, для якого повідомлення дорівнює PK_fors
10 SIG_HT := ht_sign_(PK_fors, SK_seed, PK_seed, idxtree, idxleaf);
11 SIG:= R || SIG_FORS || SIG_HT

Рис. 14. Функція slh_sign_internal. Псевдокод

Після генерації R та дайджеста (digest) (кроки 1 – 3) інформація про дерево та його лист записується в інформаційну структуру ADR (кроки 4 – 7). Для заданого дерева лісу виробляють підпис і обчислюють відповідний відкритий ключ PK_fors (кроки 8, 9).

PK_fors далі застосовують в якості повідомлення для вироблення підпису гіпердерева.

Для оптимізації застосовують оптимізовані функції для виконання кожного кроку алгоритму.

[1] Параметри MSG_len, SIG_len визначаються, якщо програмний засіб для реалізації не передбачає автоматичне визначення довжини.

5.3. Перевірка електронного підпису. Функція slh_verify_internal

Псевдокод для функції slh_verify_internal задано на рис. 15.

Після перевірки довжини підпису (крок 2) виконується виділення окремих компонентів підпису, а саме R, підпису для лісу (SIG_FORS) та підпису для гіпердереву (SIG_HT).

```
Вхід.
        MSG – повідомлення для підпису;
        MSG_len – довжина повідомлення;
        PK – відкритий ключ;
        SIG – підпис;
        SIG_len–довжина підпису.
Вихід.
        Success – ознака успішності перевірки підпису. OK – успіх, ERROR – помилка
   1    Success:= OK
   2    If SIG_len = SIGSIZE then
   3            R:=Subst (SIG, 0, n)
   4            SIG_FORS:= Subst (SIG, n, k * (n + a * n))
   5            SIG_HT:= Subst (SIG, n + k * (n + a * n), d * (h' + len) * n)
                Дайджест повідомлення та його компоненти
   6            digest :=HMsg( R, PK_seed_, PK_root, M_, M_len, buf);        Дайджест
   7            md, idxtree, idxleaf,:= DigestParse(digest)
   8            ADR:=0                                          Ініціалізація ADR
   9            ADR. TreeAddress:= idxtree
   10           ADR.Type:= FORS_TREE
   11           ADR. KeyPairAddress:= idxleaf
                Формування повідомлення для гіпердерева
   12           PK_FORS := fors_pkFromSig (SIG_FORS, md, PK_seed, ADR) ;
                Перевірка підпису. Підпис повинен співпадати з PK_root
   13           Success:= ht_verify(PK_FORS, SIG_HT, PK_seed, idxtree, idxleaf,
                PK_root);
   14    end if
   15    return Success
```

Рис. 15. Функція slh_verify_internal. Псевдокод

Виконується формування дайджесту та його компонентів як для функції вироблення підпису (SIG_HT), кроки 3 – 5.

Формується дайджест повідомлення та його компоненти (кроки 6 – 8).

Поля інформаційної структури задаються як для вироблення підпису (кроки 8 – 11).

Генерується відкритий ключ за підписом для лісу (крок 12). Цей ключ застосовують як вхідне повідомлення для перевірки підпису для гіпердерева (крок 13) .

Для оптимізації застосовують оптимізовані функції для виконання кожного кроку алгоритму.

5.4. Результати оптимізації для комплексних функцій

Результати оптимізації представлені для усіх режимів застосування функцій, а саме криптостійкості 1, 3, 5 та внутрішніх функцій SHAKE, SHA, і містяться в табл. 5 – 7.

Таблиця 5

Результати оптимізації функцій для криптостійкості 1

	SHAKE128s	SHA128s	SHAKE128f	SHA128f
slh_keygen_internal	452180489	264897957	7431865	4739014
slh_keygen_internal__	109852714	69452559	1630782	1082316
S(Keygen)	4.12	3.81	4.56	4.38
slh_sign_internal	3495259228	1995710992	179101309	99506389
slh_sign_internal_	860563546	559322878	39058311	25313915
S(Sign)	4.06	3.57	4.59	3.93
slh_verify_internal	3745263	1909084	8968145	5458235
slh_verify_internal __	1141852	890416	3027383	1892022
S(Verify)	3.28	2.14	2.96	2.88

Таблиця 6

Результати оптимізації функцій для криптостійкості 3

	SHAKE192s	SHA192s	SHAKE192f	SHA192f
slh_keygen_internal	658694406	381176539	12869939	6998364
slh_keygen_internal __	162067955	98458804	2365196	1529277
S(Keygen)	4.06	3.87	5.44	4.58
slh_sign_internal	6086417491	3596133587	277799523	164115610
slh_sign_internal_	1556043730	1007007039	61471367	39621163
S(Sign)	3.91	3.57	4.52	4.14
slh_verify_internal	4695843	3053975	13664574	8326024
slh_verify_internal __	1626014	1160354	3943224	2619513
S(Verify)	2.89	2.63	3.47	3.18

Таблиця 7

Результати оптимізації функцій для криптостійкості 5

	SHAKE256s	SHA256s	SHAKE256f	SHA256f
slh_keygen_internal	444175185	256038156	26198709	19781809
slh_keygen_internal __	99349338	64456624	6241461	3985216
S(Keygen)	4.47	3.97	4.20	4.96
Sign	5582069024	3242691723	555418484	329285834
Sign_	1276065854	829180516	123821818	81778392
S(Sign)	4.37	3.91	4.49	4.02
slh_verify_internal	6803091	4435498	13412197	8411551
slh_verify_internal __	2177669	1547537	4083640	2750761
S(Verify)	3.12	2.87	3.28	3.06

Висновки

1 Для усіх функцій і усіх режимів отримано прискорення не менше ніж в два рази, для більшості функцій та режимів прискорення більше ніж в три рази.

2 Переважна більшість сучасних процесорів багатоядерна. Застосування паралельних обчислень за рахунок застосування багатоядерних процесорів суттєво збільшує продуктивність функцій для схем wots та fors, а також функцій, які їх застосовують.

3 Ефективність застосування AVX операцій для реалізації функцій алгоритму буде розглянуто в наступній частині.

Список літератури:

1. Stateless Hash-Based Digital Signature Standard, FIPS 205, 2024 [Online]. Available: https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.205.pdf

2. Gorbenko I., Kachko O., & Derevianko Y. (2025). Optimization of digital signature calculation and verification operations for the FIPS 205 standard // Radiotekhnika. 2025. No 221. P. 7–13. https://doi.org/10.30837/rt.2025.2.221.01

3. NIST PQC. Round 3 Submissions. Алгоритм SPHINCS, Optimized_Implementation. Available: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/round-3-submissions

Надійшла до редколегії 15.05.2025

Відомості про авторів:

Горбенко Іван Дмитрович – д-р техн. наук, професор, Харківський національний університет імені В.Н. Каразіна, професор кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-науковий інститут комп'ютерних наук та штучного інтелекту; АТ «Інститут інформаційних технологій», Голова наглядової ради; Україна; e-mail: i.d.gorbenko@karazin.ua; ORCID: https://orcid.org/0000-0003-4616-3449

Качко Олена Григорівна – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри програмної інженерії, факультет комп'ютерних наук; АТ «Інститут інформаційних технологій», член наглядової ради; Україна; e-mail: iit@iit.kharkov.ua, ORCID: https://orcid.org/0000-0001-9249-0497

Дерев'янко Ярослав Андрійович – Харківський національний університет імені В. Н. Каразіна, аспірант кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-науковий інститут комп'ютерних наук та штучного інтелекту, АТ «Інститут Інформаційних технологій», науковий співробітник-консультант; Україна; e-mail: yarik0009258@gmail.com; ORCID: https://orcid.org/0000-0002-3290-3373

УДК 004.05 **DOI:10.30837/rt.2025.3.222.02**

В.В. БОРОДАВКА, В.І. ЄСІН, д-р техн. наук

ВПРОВАДЖЕННЯ АРХІТЕКТУРИ НУЛЬОВОЇ ДОВІРИ НА ОСНОВІ ЗАПРОПОНОВАНОЇ МОДЕЛІ ДЛЯ ЗАБЕЗПЕЧЕННЯ КІБЕРБЕЗПЕКИ ПІДПРИЄМСТВА

Вступ

В умовах стрімкої цифрової трансформації та зростання загроз у кіберпросторі актуальність питань організації кібербезпеки підприємств набуває особливої значущості. Зростаючий обсяг пристроїв, що підключаються до мережі, зокрема через технологію Інтернету речей (Internet of Things), швидкий розвиток нових технологій штучного інтелекту (ШІ), генеративного ШІ, квантових обчислень та постійне зростання складності кібератак створюють необхідність реалізації інноваційних підходів до забезпечення безпеки даних. Традиційні методи захисту на основі периметру виявляються неефективними перед сучасними кіберзагрозами. Відповіддю на ці виклики стала концепція нульової довіри, як один із найефективніших підходів для протидії сучасним кіберзагрозам. Вона ґрунтується на принципі «ніколи не довіряй, завжди перевіряй» і передбачає, що жоден користувач, пристрій або процес не повинен автоматично вважатися безпечним, навіть якщо вони знаходяться всередині корпоративної мережі.

Концепція нульової довіри була запропонована вперше S. P. Marsh у 1994 р. [1]. Проте значний крок у її розвитку відбувся лише в 2003 р., коли консорціум Jericho Forum представив концепцію [2], яка передбачала використання принципів нульової довіри як стратегії безпеки для деперіметризації (de-perimeterization), що фактично означало усунення межі між організацією та зовнішнім світом. У 2009 р. Google впровадив модель BeyondCorp [3], яка поклала початок застосуванню концепції нульової довіри у корпоративному середовищі. Дана модель передбачала автентифікацію на основі перевірки користувача та пристрою, що дозволяло відмовитися від привілейованих корпоративних мереж. У 2010 р. J. Kindervag (аналітик компанії Forrester) використав концепцію нульової довіри для контролю доступу [4]. Він наголосив, що традиційні методи безпеки є недостатніми для протидії сучасним кіберзагрозам. Саме тоді було сформульовано ключовий принцип концепції нульової довіри: «ніколи не довіряй, завжди перевіряй» [4]. Подальший розвиток концепції призвів до появи у 2017 р. розширеної концепції нульової довіри (Zero Trust Extended – ZTX) від компанії Forrester [5], що охоплювала ширше коло потоків даних, зокрема тих, що проходять через локальні мережі, хмарні сервіси, зовнішні застосунки, сайти та різні види кінцевих пристроїв. Компанія Gartner [6], ґрунтуючись на принципі безперервної адаптивної оцінки ризиків та довіри (Continuous Adaptive Risk and Trust Assessment – CARTA), також висунула ідею про ZTX. Надалі, додатковий імпульс розвитку концепції надали публікації NIST [7] та NCCoE (National Cybersecurity Center of Excellence) [8], які визначали основні принципи та компоненти нульової довіри, і акцентували увагу на архітектурі нульової довіри (Zero Trust Architecture – ZTA) для захисту корпоративних даних [9]. Подальший розвиток концепції нульової довіри відзначається інтеграцією системи виявлення та реагування на кіберзагрози в ZTA, включно із системами захисту кінцевих точок (Endpoint Detection and Response – EDR), розширеним виявленням та реагуванням (eXtended Detection and Response – XDR), а також мережевим виявленням та реагуванням (Network Detection and Response – NDR). Ці технології відіграють ключову роль у забезпеченні корпоративної безпеки, оскільки спрямовані на виявлення, аналіз і нейтралізацію атак, що здатні обходити традиційні засоби захисту, зокрема платформи захисту кінцевих пристроїв (Endpoint Protection Platform – EPP). Іншим важливим аспектом впровадження ZTA в умовах поширення гібридних робочих середовищ є забезпечення безпеки віддаленого доступу (Remote Access Security), що реалізується за допомогою багатофакторної автентифікації (Multifactor Authentication – MFA), політик

контролю доступу та наскрізного шифрування мережевого трафіку. Застосування ZTA є ефективним заходом протидії атакам на ланцюги постачання (Supply Chain Attacks), які дедалі частіше використовуються як вектор компрометації корпоративних мереж, що зумовлює необхідність обмеження доступу сторонніх постачальників до критично важливих систем та впровадження детального моніторингу їхньої активності з метою зниження ризику несанкціонованого втручання. Окремої уваги заслуговують атаки на відкриті програмні інтерфейси (Application Programming Interface – API), які можна використовувати для несанкціонованого доступу до корпоративних даних. В цьому контексті застосування шлюзів для автентифікації та мікросегментація дозволяє значно зменшити ризики, пов'язані з вразливостями API. Проте, окрім корпоративних мереж, ZTA має критичне значення для захисту промислових систем управління (Industrial Control Systems – ICS) та операційних технологій (Operational Technology). Дані системи є частиною критичної інфраструктури, які є постійними цілями для кібератак, в свою чергу ZTA може допомогти мінімізувати ризики кібератак на ці системи. Наприклад, під час атаки NotPetya у 2017 р. [10], яка завдала значних фінансових та операційних збитків підприємствам, впровадження принципів нульової довіри дозволило б обмежити поширення шкідливого програмного забезпечення (ПЗ) шляхом бічного переміщення (lateral movement) та мінімізувати його вплив на інфраструктуру. Водночас, ефективність реалізації ZTA значною мірою залежить від здатності системи своєчасно ідентифікувати загрози та реагувати на них. Зокрема, важливим аспектом є те, що успішна атака можлива навіть після численних невдалих спроб, тоді як захист повинен працювати безперебійно, запобігаючи будь-якій компрометації. У зв'язку з цим зростає необхідність впровадження механізмів, що забезпечують безперервний моніторинг трафіку, аналіз поведінки користувачів та пристроїв, а також адаптивне реагування на потенційні загрози. Оскільки ймовірність кібератаки неминуча, а абсолютний захист є лише теоретичним [11], ключовим завданням стає мінімізація ризиків шляхом раннього виявлення та миттєвого реагування.

На сьогодні концепція нульової довіри набуває все більшого поширення, її застосовують у різних сферах, від хмарних сервісів до 5G-мереж для медичних пристроїв [12]. За даними аналітичної компанії MarketsandMarkets [13] прогнозується, що ринок ZTA зросте від 36,5 мільярдів доларів США у 2024 р. до 78,7 мільярдів доларів у 2029 р., з середньорічним темпом зростання 16,6 % протягом даного періоду. Це свідчить про високу ефективність моделі нульової довіри у зменшенні ризиків від кібератак, підвищенні загальної безпеки інформаційних систем та про значне поширення ZTA як у вертикальному (різні галузі), так і у горизонтальному (різні регіони) вимірах. Таким чином, подальший розвиток і впровадження ZTA є важливим кроком у підвищенні загальної стійкості кібербезпеки підприємств, що дозволить створювати динамічні системи захисту, здатні не лише протистояти сучасним атакам, але й адаптуватися до нових викликів.

Виходячи з зазначеного, концепція нульової довіри є ключовим підходом до побудови сучасних систем кібербезпеки, що дозволяє значно підвищити рівень захищеності інформаційних ресурсів підприємства. Проте, попри очевидні переваги концепції нульової довіри, процес її впровадження в корпоративні інформаційні системи супроводжується значними труднощами як з технічної, так і з організаційної точки зору. Як видно з досліджень [14, 15], існує проблема, пов'язана з відсутністю підходу до впровадження ZTA, який враховував би як технічні, так і організаційні аспекти. Дана робота націлена на вирішення цієї проблеми шляхом узагальнення наявних досліджень та досвіду різних міжнародних компаній, а також розробки на основі цього моделі впровадження ZTA для організації кібербезпеки підприємства. Це дозволить підприємствам ефективно адаптувати підхід нульової довіри відповідно до своїх потреб та особливостей власної інфраструктури. У статті розглядаються рекомендації щодо вибору моделей розгортання ZTA для різного роду діяльності підприємств, а також модель впровадження ZTA, що допоможе зрозуміти фундаментальні зміни у підході до організації кібербезпеки, а також ефективно впровадити концепцію нульової довіри з урахуванням технічних та організаційних можливостей і вимог конкретного IT-підприємства.

1. Рекомендації щодо вибору моделей розгортання архітектури нульової довіри для різних підприємств

Захист корпоративних даних і ресурсів стає дедалі складнішим завданням. Багатьом користувачам потрібен доступ з будь-якого місця, в будь-який час, з будь-якого пристрою, для забезпечення функціональної діяльності організації. В свою чергу, дані створюються, зберігаються, передаються та обробляються в різних корпоративних мережах і розподілені між локальними та хмарними середовищами, щоб задовольнити потреби бізнесу, які постійно змінюються. Вже неможливо просто захистити дані та ресурси за периметром корпоративного середовища або припустити, що всім користувачам, пристроям, застосункам і сервісам в ньому можна довіряти.

ZTA є ефективним підходом до забезпечення кібербезпеки, який дозволяє організаціям захищати свої дані та ресурси незалежно від їхнього місцезнаходження. Концепція нульової довіри ґрунтується на ризик-орієнтованому підході до управління доступом, що передбачає постійний моніторинг, оцінку та перевірку умов і запитів на доступ із подальшим наданням або обмеженням доступу залежно від рівня ризику. Кожен запит на доступ у межах ZTA явно перевіряється з урахуванням контексту, що включає як статичну інформацію, так і динамічні фактори. Якщо запит відповідає визначеним політикам, створюється безпечний сеанс доступу до даних. ZTA також забезпечує безперервний моніторинг та оцінку ризиків у реальному часі, що дозволяє динамічно застосовувати корпоративні політики безпеки. Варто зазначити, оскільки впровадження ZTA є складним процесом, організації потребують структурованого підходу до її реалізації. Насамперед необхідно провести комплексну оцінку поточних ресурсів, визначити сильні та слабкі сторони існуючої інфраструктури, а також окреслити ключові етапи переходу до ZTA. Поетапне впровадження дозволяє поступово адаптувати корпоративне середовище до вимог ZTA, враховуючи наявні ризики, витрати, доступні ресурси та стратегічні пріоритети. Важливим аспектом цього процесу є забезпечення балансу між підвищенням рівня безпеки та збереженням ефективності бізнес-процесів. При цьому слід зазначити, що не існує універсального підходу до впровадження ZTA, який був би однаково ефективним для всіх підприємств. ZTA – це набір керівних принципів та концепцій, а не набір технічних специфікацій, яких можна дотримуватися. Тому впровадження ZTA передбачає не фіксований набір вимог, а безперервний процес удосконалення процесів, механізмів контролю доступу та політик безпеки відповідно до принципів ZTA та специфіки кожного підприємства.

Основними перевагами від впровадження концепції нульової довіри на підприємстві, як зазначається у роботах [9, 15, 16], є:

1. Поліпшення видимості мережі, виявлення зловживань або порушень та керування вразливостями (покращення видимості того, які користувачі, коли, як і звідки отримують доступ до тих чи інших ресурсів та які дії виконують).

2. Перешкода поширенню шкідливого ПЗ, а саме більш ефективне виявлення, реагування та відновлення після інцидентів, що дозволяє мінімізувати наслідки витоків даних, а також мінімізація ризику бічного переміщення.

3. Оптимізація та скорочення витрат, пов’язаних із забезпеченням безпеки, а також скорочення обсягу та вартості робіт, необхідних для дотримання та забезпечення нормативних вимог і стандартів безпеки.

4. Усунення конфліктів між підрозділами при розслідуванні інцидентів. Наприклад, у випадку деяких інцидентів мережеві спеціалісти можуть звинувачувати службу безпеки у створенні перешкод у роботі, тоді як служба безпеки може вказувати на недоліки в мережевій інфраструктурі. Концепція нульової довіри сприяє налагодженню взаємодії між робочими групами, дозволяючи усунути подібні протиріччя шляхом чіткої регламентації доступу та відповідальності.

5. Підвищення рівня обізнаності, розуміння та контролю даних.

6. Комплексний захист інформаційних ресурсів і критичних активів, що включає запобігання витоку важливих даних до зловмисників шляхом обмеження їх можливостей, зниження рівня внутрішніх загроз, застосування надійного шифрування для захисту важливих (sensitive) корпоративних даних як під час передачі, так і в стані спокою. Додатково передбачено динамічний контроль доступу з урахуванням ризиків шляхом постійного моніторингу активності, безперервної переоцінки всіх операцій та сеансів доступу, збору інформації, отриманої в результаті періодичної повторної автентифікації та авторизації користувачів, постійної оцінки стану пристроїв, аналізу поведінки, виявлення аномалій у режимі реального часу та інших аналітичних даних щодо безпеки. Наприклад, система може автоматично обмежити доступ до конфіденційних файлів, якщо користувач входить із незвичного місця або використовує потенційно скомпрометований пристрій.

7. Забезпечення цифрової трансформації бізнесу, що включає підтримку віддаленої роботи, надання працездатних пристроїв від постачальників, гнучке керування доступом для співробітників і підвищення якості обслуговування кінцевих користувачів.

Проте, під час впровадження ZTA можуть виникати деякі специфічні загрози [15], які мають унікальні особливості (наприклад, спотворення процесу прийняття рішень; відмова в обслуговуванні або порушення роботи мережі; крадіжка облікових даних (інсайдерська загроза); видимість у мережі; зберігання системної та мережевої інформації; залежність від пропрієтарних/власних форматів даних або рішень; використання сутностей, які не є фізичними особами, при адмініструванні ZTA), для яких також мають місце певні рішення у межах реалізації певної ZTA. При цьому слід розуміти, що при розгортанні рішень, які задовольняють вимогам концепції нульової довіри, як зазначається у роботах [17 – 23], організації можуть зіткнутися з численними технічними та організаційними викликами (табл. 1), які повинні враховуватися під час впровадження ZTA.

Таблиця 1

Виклики при розгортанні рішень на основі концепції нульової довіри

Організаційні	**Технічні**
✓ Помилкове розуміння, що ZTA підходить лише для великих організацій і вимагає значних інвестицій, замість розуміння того, що ZTA – це набір керівних принципів, придатних для організацій будь-якого розміру. ✓ Занепокоєння, що ZTA може негативно вплинути на функціонування IT-середовища або на роботу кінцевих користувачів. ✓ Недостатність ресурсів для розробки необхідних політик та експериментальної реалізації, необхідних для формування плану переходу до ZTA. ✓ Ефективне використання наявних ресурсів та баланс пріоритетів під час переходу до ZTA. ✓ Відсутність розуміння того, які додаткові навички та навчальні програми можуть знадобитися адміністраторам, співробітникам служби безпеки, постачальникам послуг, кінцевим користувачам. ✓ Недостатня розвиненість стандартизованих політик для впровадження, керування та забезпечення дотримання вимог ZTA, що призводить до фрагментованості середовища або несумісності окремих компонентів в рамках архітектури. ✓ Відсутність універсального підходу до реалізації та впровадження ZTA потребує розробки індивідуальних стратегій та політик впровадження, що враховують специфіку організаційних процесів, рівень прийнятного ризику, поточну технологічну базу та фінансові можливості.	✓ Відсутність належної системи інвентаризації та управління активами, необхідної для повного розуміння бізнес-інфраструктури, активів та процесів, які потребують захисту, а також відсутність чіткого розуміння критичності даних ресурсів. ✓ Відсутність або обмежене розуміння процесів, що відбуваються між суб'єктами, активами, застосунками та сервісами організації, а також відсутність даних, необхідних для ідентифікації цих процесів та їхніх конкретних потоків. ✓ Складність інтеграції різнорідних доступних технологій, оцінки їхньої відповідності вимогам ZTA, а також виявлення технологічних недоліків, що необхідні для формування комплексної ZTA. ✓ Проблеми сумісності ZTA із застарілими/успадкованими системами, що потребують додаткових ресурсів для адаптації або заміни. ✓ Відсутність належного цифрового опису, керування та контролю ролей користувачів в організації, необхідних для впровадження детальної політики доступу до певних застосунків та сервісів

Одним із ключових аспектів впровадження ZTA є необхідність поступового підходу, що дозволяє мінімізувати ризики та забезпечити адаптацію інфраструктури підприємства до нових умов. У цьому контексті першочерговим завданням є визначення стратегії переходу, яка відповідно до рекомендацій NIST [7], повинна містити поетапний підхід із впровадженням експериментального проєкту. На початковому етапі визначаються об'єкти та системи, які першочергово підлягають міграції, після чого процес поступово масштабується на інші компоненти інфраструктури. Крім того, акцент робиться на ідентифікації активів, бізнес-процесів і керуванні ризиками. В свою чергу, дослідження, проведені компанією Cisco [24], акцентують увагу на механізмах автентифікації, моніторингу користувачів та пристроїв, а також окреслюють три ключові аспекти впровадження ZTA: стратегічний, управлінський і операційний. Водночас у дослідженні недостатньо деталізовано методи та технічні аспекти міграції, що ускладнює практичну реалізацію ZTA. Прикладом технічного підходу до міграції є реалізація Google BeyondCorp [3, 25], де основна увага приділяється впровадженню керування користувачами, пристроями та мережевою інфраструктурою в рамках концепції нульової довіри. Міграція реалізується поетапно, з початковим тестуванням на обмеженій кількості систем, після чого проводиться аналіз ефективності та поступова інтеграція в масштабах усієї інфраструктури.

У свою чергу, підхід Microsoft [26] орієнтується на управлінські аспекти міграції, зокрема на визначення меж та етапів реалізації моделі нульової довіри. Згідно з цією стратегією, Microsoft структурно розподіляє впровадження ZTA на ключові компоненти, такі як ідентифікація користувачів, керування пристроями, доступ до ресурсів та сервісів, визначаючи критерії оцінки ефективності впровадження ZTA.

NCCoE [8] активно співпрацює з численними постачальниками технологій, наприклад, такими як AWS, Broadcom, Cisco, Google Cloud, IBM, Ivanti, Microsoft, Okta, Palo Alto Networks, Tenable, що дозволяє інтегрувати комерційні і відкриті рішення (open-source) та забезпечити комплексну безпеку згідно з принципами концепції нульової довіри. У рамках цього підходу багато досліджень [27 – 29] із впровадження ZTA визначають основні етапи, які необхідно пройти для забезпечення надійного захисту підприємства. Серед них першочерговим етапом є ідентифікація об'єктів захисту, що охоплюють дані, активи, застосунки та сервіси. Наступним кроком є визначення потоків даних, що взаємодіють із цими об'єктами, для забезпечення належного контролю доступу та запобігання несанкціонованому доступу до критичних ресурсів. Далі важливим є архітектурне проектування мережі, яке повинно враховувати принципи мікросегментації. Завершальним етапом є розробка та впровадження політики нульової довіри, що передбачає безперервну автентифікацію та авторизацію всіх запитів на доступ до ресурсів, незалежно від їхнього походження або місцезнаходження користувача чи сервісу, що дозволяє забезпечити безпечний та ефективний захист інфраструктури підприємства на всіх етапах її життєвого циклу.

Проте, попри широке визнання переваг ZTA, значна частина досліджень обмежується лише загальним описом основних етапів її впровадження, без детального обґрунтування підходів до міграції. Крім того, багато робіт зосереджуються на вузьких технічних або організаційних аспектах, ігноруючи динамічні та комплексні методи, що забезпечують поступовий і ефективний перехід до ZTA. На думку фахівців NIST, ZTA може різнитися залежно від потреб компанії. Для цього вони розглядають можливість застосування кількох різних підходів (шляхів), за допомогою яких підприємство може запровадити ZTA для робочих процесів, зокрема: вдосконалене/покращене управління ідентифікацією (enhanced identity governance), логічну мікросегментацію (logical micro-segmentation) та сегментацію на основі мережі (network-based segmentation) [7]. У роботі [15] докладно описані дані підходи та визначена різниця в них між компонентами, що використовуються, і основним джерелом правил політики для організації. При цьому варто зазначити, що одні підходи найбільше підходять для одних випадків, тоді як інші доцільно використовувати в інших ситуаціях. Тому підприємство, яке прагне розробити та впровадити ZTA, може виявити, що обраний ним варіант використання

та існуючі політики виділяють один підхід серед інших існуючих. Однак це не означає, що інші підходи не працюватимуть. Зважаючи на все, це лише вказує на те (і не більше), що інші підходи можуть бути більш важкими для реалізації і можуть вимагати кардинальніших змін в організації розвитку бізнесу.

З іншого боку, розглядати потенційні рішення з нульовою довірою доцільно з погляду їх моделей розгортання, які можуть бути корисною основою, за допомогою якої підприємства зможуть оцінити потенційних постачальників відповідних рішень, аналізуючи їх плюси та мінуси. Вважається, що багато з корпоративних моделей нульової довіри, що надаються постачальниками, будуть відповідати одній або декільком моделям розгортання. Наразі відомі такі поширені варіанти розгортання ZTA на підприємствах [7, 15, 30]:

1. Модель агента пристрою/шлюзу (Device Agent/Gateway Model; іноді цю модель називають моделлю розгортання на основі ресурсів – Resource-Based Deployment Model).

2. Модель розгортання з мікросегментацією (Micro-Segmentation Deployment Model).

3. Модель розгортання на основі анклаву (Resource Enclave Deployment Model).

4. Модель розгортання на основі порталу ресурсів (Resource Portal-Based Deployment).

5. Модель розгортання з використанням хмарної маршрутизації (Cloud-Routed Deployment Model).

6. Модель розгортання на основі використання «пісочниці» (sandboxing) застосунків.

Дані моделі можуть оцінити та надати новий рівень деталізації того, як насправді можуть бути розгорнуті системи з ZTA. Хоча при цьому не слід забувати, що реальна архітектура розгортання, звичайно ж, буде залежати від можливостей обраної технології [15] та може відбуватися з використанням декількох моделей залежно від особливостей підприємства та його технологічних можливостей. Варто зазначити, що будь-яке IT-середовище може бути спроектоване з урахуванням принципів нульової довіри, а більшість організацій вже мають певні елементи ZTA в своїй корпоративній інфраструктурі або знаходяться на шляху до впровадження політик і передових практик. З огляду на це, важливим аспектом є врахування різних сценаріїв при розгортанні ZTA в корпоративне середовище. Малоймовірно, що будь-яке підприємство може перейти до ZTA за один цикл оновлення технологій. Оскільки повний перехід до концепції нульової довіри є складним і тривалим процесом, у більшості організацій протягом певного часу можуть співіснувати як традиційні механізми захисту, так і елементи ZTA, що впроваджуються поступово для окремих бізнес-процесів. Успішна реалізація цього підходу вимагає забезпечення гнучкості основних компонентів, таких як керування ідентифікацією, контроль пристроїв і реєстрація подій для ефективного функціонування як у межах класичної архітектури безпеки на основі периметра, так і в гібридній архітектурі. Крім того, обрані рішення мають бути сумісними з наявною корпоративною інфраструктурою, щоб уникнути значних ускладнень під час впровадження. Варто також зазначити, що міграція до ZTA вимагатиме (як мінімум) часткової реорганізації робочих процесів в організації. Це, водночас, створює умови для впровадження сучасних методів та практик безпечної розробки систем відповідно до міжнародних стандартів [31], якщо IT-підприємства ще не зробили цього для робочих процесів.

Деякі сценарії розгортання та варіанти використання дозволяють швидко впровадити ZTA на підприємстві. Зокрема, концепція нульової довіри є ефективним підходом для організацій із розподіленою інфраструктурою або значною кількістю віддалених користувачів. Водночас будь-яке підприємство може застосовувати принципи нульової довіри для підвищення рівня безпеки, незалежно від структури чи масштабу. На практиці перехід до нової моделі безпеки зазвичай здійснюється поступово, тому в корпоративному середовищі певний час можуть співіснувати як традиційні механізми захисту на основі периметру, так і елементи ZTA.

Далі розглянемо деякі рекомендації щодо практичного впровадження ZTA, з урахуванням можливих сценаріїв, викликів і оптимальних підходів до її інтеграції в корпоративне середовище для різного роду діяльності IT-підприємств.

1.1. Підприємство з віддаленими структурними підрозділами

Найпоширенішим сценарієм є підприємство, що має центральний офіс і один або декілька географічно розподілених підрозділів (філіалів), які не з'єднані фізичною мережею, що належить підприємству (рис. 1).

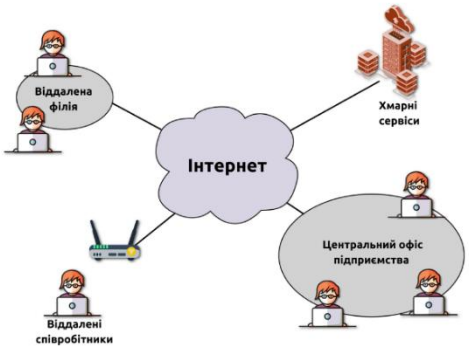


Рис. 1. Підприємство з віддаленими працівниками

При цьому працівники у віддалених підрозділах або офісах можуть не мати повноцінної локальної мережі підприємства, проте їм все одно необхідний доступ до корпоративних ресурсів для виконання своїх завдань. У таких випадках підприємство може використовувати багатопротокольну комутацію за мітками (Multiprotocol Label Switching – MPLS) для зв'язку з центральним офісом, однак пропускна здатність цього каналу може бути недостатньою для обробки всього трафіку. Крім того, підприємство може вважати недоцільним спрямовувати весь трафік, призначений для хмарних сервісів, застосунків або сервісів через центральний офіс. Аналогічна ситуація виникає, коли співробітники працюють віддалено або перебувають поза межами основної інфраструктури підприємства, використовуючи як корпоративні, так і особисті пристрої (BYOD). У таких випадках підприємство може надавати доступ до певних ресурсів, таких як електронна пошта чи корпоративний календар, водночас обмежуючи або забороняючи доступ до більш конфіденційних ресурсів (наприклад, база даних співробітників).

У цьому сценарії механізм політики (Policy Engine – PE) та адміністратор політики (Policy Administrator – PA) [7, 15] реалізується у вигляді хмарного сервісу (який зазвичай забезпечує високий рівень доступності та не вимагає від віддалених користувачів підключення до корпоративної інфраструктури для доступу до хмарних ресурсів), а кінцеві пристрої мають встановлений агент (відповідно до моделі розгортання на основі агента пристрою/шлюзу, де зазвичай у системі суб'єкта є розгорнутий користувальницький агент, що діє як точка застосування політики (Policy Enforcement Point – PEP) агента користувача) або доступ до ресурсного порталу (відповідно до моделі розгортання на основі порталу ресурсів, де PEP є єдиним компонентом, який виконує роль шлюзу для відповідних запитів). Розміщення PE/PA у локальній мережі підприємства у цьому випадку є менш ефективним, оскільки змусить віддалені філії та співробітників перенаправляти весь трафік через центральний офіс для доступу до хмарних сервісів та застосунків.

З огляду на відсутність єдиного захищеного периметра, одним із ключових викликів у такому сценарії є забезпечення безпеки корпоративних баз даних і сховищ, що можуть бути розміщені як у локальному сегменті мережі підприємства, так і у хмарному середовищі. Концепція нульової довіри вимагає, щоб кожен запит на доступ до даних, незалежно від джерела, супроводжувався ретельною і повторною перевіркою ідентичності користувача та контексту доступу. Зокрема, для ресурсів, що зберігають конфіденційну/чутливу інформацію (наприклад, персональні дані співробітників, фінансові документи чи інтелектуальну власність), впроваджується модель багаторівневого контролю доступу з динамічною адаптацією політик. Це може включати як ізоляцію критичних сховищ у сегментованих віртуальних локальних мережах (Virtual Local Area Network – VLAN), так і застосування моделей контролю

доступу на основі ролей (Role-based Access Control – RBAC) та/або атрибутів (Attribute-based Access Control – ABAC).

Крім того, у контексті доступу до конфіденційних даних дедалі актуальнішими стають механізми, що базуються на концепції «нульового знання/розголошення» (Zero-Knowledge – ZK) [32 – 36] та які можуть бути впроваджені в процеси автентифікації або перевірки дозволів, повноважень, прав. Дані механізми дозволяють довести право на доступ до ресурсу без потреби передавати або розкривати самі дані, що значно підвищує рівень конфіденційності та захищає від атак типу «людина посередині» (man-in-the-middle) або витоку даних у процесі автентифікації. Наприклад, використовуючи концепцію ZK, система може перевірити, чи має користувач необхідні повноваження для доступу до певних чутливих даних бази даних (наприклад, до даних про співробітників підприємства), не розкриваючи при цьому жодних деталей про самі повноваження або вміст ресурсу. У хмарних середовищах така технологія може бути інтегрована з постачальниками ідентифікації (Identity Provider – IdP), які здійснюють перевірку повноважень користувача без необхідності доступу до самих даних, що додатково відповідає вимогам конфіденційності та відповідності нормативним стандартам (GDPR, HIPAA тощо).

1.2. Підприємство, що функціонує у багатохмарному середовищі

Поширеним сценарієм застосування ZTA є функціонування підприємства у багатохмарному середовищі, коли використовується інфраструктура кількох хмарних провайдерів (рис. 2).

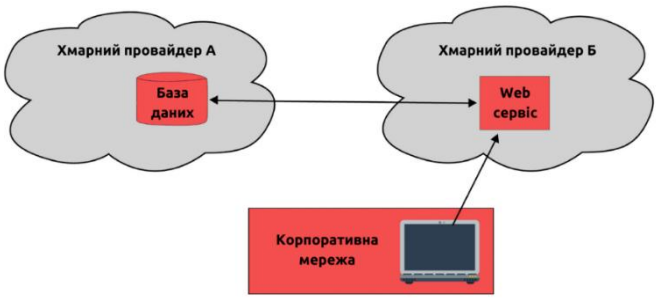


Рис. 2. Підприємство з багатохмарним середовищем

У цьому випадку підприємство має локальну мережу, проте для розміщення застосунків, сервісів і баз даних задіяні два або більше постачальників хмарних послуг. Нерідко застосунок або сервіс функціонує в одній хмарній інфраструктурі, тоді як база даних розташована в іншій. Для забезпечення ефективної роботи системи та спрощення адміністрування необхідно, щоб застосунок, розгорнутий в інфраструктурі одного хмарного провайдера *А*, мав можливість безпосередньо взаємодіяти з базою даних іншого хмарного провайдера *Б*, уникаючи маршрутизації трафіку через локальну мережу підприємства.

Цей сценарій є реалізацією серверної взаємодії (server-to-server) відповідно до специфікації програмно-визначуваного периметра (Software Defined Perimeter – SDP) від Cloud Security Alliance [37]. Проте широке застосування хмарних технологій виявило обмеження традиційної моделі безпеки, яка базується на захисті мережевого периметра та є неефективною в межах хмарного середовища. В свою чергу, ZTA передбачає, що внутрішня та зовнішня інфраструктури мають однаковий рівень довіри, тобто фізичне розташування ресурсів не повинно впливати на рівень безпеки доступу (не повинно бути ніякої різниці між мережевою інфраструктурою, що належить і експлуатується підприємством, і інфраструктурою, що належить і експлуатується будь-яким іншим постачальником послуг). Тому, для забезпечення контролю доступу в багатохмарному середовищі PEP розміщуються на рівні застосунків, сервісів і баз даних. В свою чергу PE/PA можуть бути інтегровані як у межах одного із задіяних хмарних провайдерів, так і в незалежному середовищі третьої сторони. Користувачі отримують доступ до ресурсів через портал або встановлений агент, звертаючись до відпові-

дних PEP. Такий підхід дає змогу підприємству контролювати доступ до своїх ресурсів незалежно від місця їх розташування.

Варто зазначити, що у процесі проєктування та розгортання ZTA в багатохмарному середовищі важливим аспектом є забезпечення безпеки баз та сховищ даних, які використовуються на різних етапах обробки, передачі та зберігання інформації. У контексті моделі, наведеної на рис. 2, особливу увагу слід приділити захисту баз та сховищ даних, що знаходяться під контролем хмарного провайдера *А*, але обслуговують застосунки, розгорнуті в інфраструктурі хмарного провайдера *Б*. Такий формат взаємодії між хмарними провайдерами створює додаткові вектори загроз, які не можуть бути належним чином враховані у межах традиційної моделі безпеки. У зв'язку з цим обов'язковою умовою є впровадження механізмів доступу до баз і сховищ даних, які базуються на принципах мінімальних привілеїв, автентифікації та авторизації на основі контексту, а також на криптографічних механізмах, що підтримують концепцію ZK. Такі механізми дозволяють підтвердити права доступу без безпосереднього розкриття ідентифікаційних даних, що суттєво знижує ризики компрометації у випадку перехоплення або несанкціонованого втручання в канали зв'язку (це дозволяє істотно підвищити стійкість системи до атак типу «людина посередині», атак повторного відтворення (replay attacks), а також до витоку ідентифікаційних даних у результаті компрометації одного з провайдерів). Сьогодні існує різні підходи доказів із нульовим розголошенням (Zero-knowledge proofs – ZKP), зокрема, інтерактивні та неінтерактивні [38]. Для хмарного середовища більш релевантними є неінтерактивні ZKP, які не вимагають багатоетапної взаємодії між сторонами, що знижує затримки при обробці запитів і дозволяє реалізовувати масштабовані рішення для контролю та керування доступу до розподілених ресурсів. Одним із відомих прикладів таких протоколів є *zk-SNARK* [39], які використовуються в низці проєктів, орієнтованих на конфіденційність. Єдине, що в даному випадку є важливим, це те, що їхнє впровадження в архітектуру доступу до хмарних сховищ забезпечує можливість підтвердження автентичності без необхідності передачі паролів, токенів або інших сенситивних атрибутів. Крім того, всі бази та сховища даних, незалежно від їх фізичного або хмарного розміщення, повинні розглядатися як середовища з невизначеним рівнем довіри. Це передбачає, що всі дані мають зберігатися та передаватися виключно у зашифрованому вигляді, з використанням сучасних алгоритмів симетричного та асиметричного шифрування, протоколів TLS 1.3 або/та QUIC для захисту даних на транспортному рівні, а також з обов'язковим застосуванням наскрізного шифрування для захисту даних в стані спокою, під час обробки та під час передавання. Додатково, необхідно впровадити регулярний аудит політик доступу, що забезпечить своєчасне виявлення аномалій та спроб несанкціонованого доступу. Інтеграція таких засобів із PEP, які контролюють кожен запит до ресурсу, дозволяє забезпечити дотримання принципів концепції нульової довіри незалежно від середовища розгортання (навіть у середовищі, де кілька хмарних провайдерів відповідають за різні компоненти інфраструктури підприємства). У сукупності ці підходи дозволяють створити стійку до компрометації архітектуру, здатну гарантувати безпечну взаємодію між компонентами підприємства навіть у розподіленому хмарному ландшафті з неоднорідними вимогами до безпеки з боку різних постачальників послуг.

Таким чином, у даному сценарії критичним є забезпечення повноцінного контролю над кожним етапом доступу до ресурсів підприємства, від початкового підключення до мережі – до виконання конкретного запиту до внутрішніх ресурсів. Реалізація принципів нульової довіри у такому середовищі вимагає застосування комплексної моделі захисту, яка поєднує мікросегментацію інфраструктури, динамічну автентифікацію на основі контекстних факторів, надійні криптографічні механізми, включаючи ZK, а також постійний моніторинг активності користувачів. Впровадження таких механізмів у межах політики найменших привілеїв, а також RBAC та/або ABAC, забезпечує не лише захист конфіденційної інформації, але й унеможливлює бічне переміщення всередині мережі та несанкціоноване підвищення привілеїв. Як результат, навіть за умови різного ступеня довіри до користувачів і систем та у

межах розподіленої структури доступу, підприємство зберігає контроль над своїми внутрішніми ресурсами, знижуючи ризики як зовнішніх, так і внутрішніх загроз.

1.3. Підприємство з контрактними послугами та/або доступом для позаштатних користувачів

Одним із поширених сценаріїв є підприємство, яке надає обмежений доступ до своїх ресурсів відвідувачам та/або стороннім постачальникам послуг, що виконують контрактні зобов'язання (рис. 3).

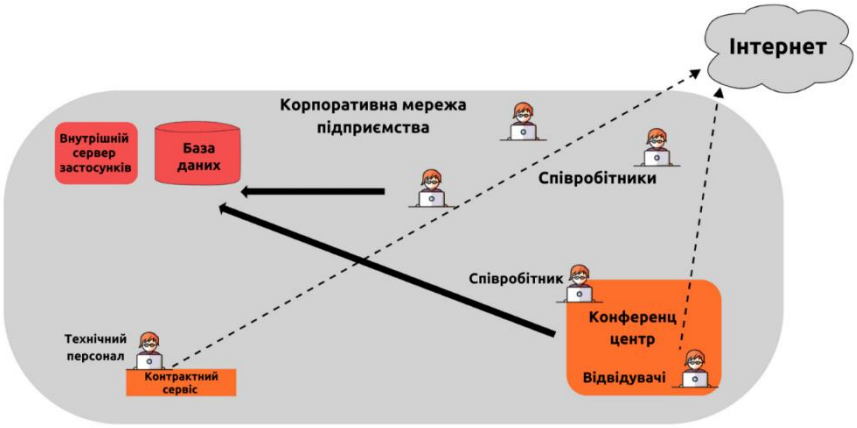


Рис. 3. Підприємство з доступом для позаштатних працівників

Таке підприємство має власні внутрішні застосунки, сервіси, бази даних та інші активи. Деякі з них можуть обслуговуватися за контрактом залученими постачальниками, які можуть періодично відвідувати підприємство для забезпечення технічного обслуговування, наприклад, системи керування розумного опалення або освітлення, що знаходяться у власності та під управлінням зовнішніх компаній. Для виконання своїх завдань такі постачальники та відвідувачі потребують підключення до мережі. У межах ZTA підприємство може організувати даний процес таким чином, щоб забезпечити доступ технічному персоналу, який відвідує підприємство, або його пристроям до мережі Інтернет, водночас унеможливлюючи несанкціонований доступ до своїх внутрішніх ресурсів.

Крім того, підприємство може мати конференц-центр, де відвідувачі взаємодіють зі співробітниками. У рамках ZTA, реалізованої за допомогою концепції SDP, пристрої та суб'єкти автентифікації розмежовуються, що дозволяє співробітникам отримувати доступ до необхідних корпоративних ресурсів, тоді як відвідувачі можуть користуватися лише доступом до мережі Інтернет. Більше того, вони не мають змоги виявити внутрішні сервери або сервіси підприємства за допомогою мережевого сканування, що запобігає активній розвідці мережі та бічному переміщенню. У цьому сценарії PE/PA можуть бути реалізовані як хмарний сервіс або розміщені локально (за умови, що хмарні сервіси майже не використовуються на підприємстві або використовуються тільки для окремих сервісів). Доступ до корпоративних ресурсів може здійснюватися через встановлений агент на кінцевих пристроях користувачів (відповідно до моделі розгортання на основі агента пристрою/шлюзу, де зазвичай у системі суб'єкта є розгорнутий користувальницький агент, що діє як PEP агента користувача) або через портал (відповідно до моделі розгортання на основі порталу ресурсів, де PEP є єдиним компонентом, який виконує роль шлюзу для відповідних запитів). При цьому механізм PA гарантує, що всі некорпоративні активи (тобто пристрої, які не мають встановленого агента або не можуть під'єднатися до порталу) не матимуть доступу до локальних ресурсів, а лише до мережі Інтернет.

Окрім зазначеного вище, в умовах реалізації ZTA в даному сценарії, ключове значення має безпека корпоративних баз і сховищ даних, що інтегровані у внутрішню мережу підприємства та взаємодіють із внутрішніми сервісами і авторизованими суб'єктами доступу. Відповідно до принципів нульової довіри доступ до таких ресурсів повинен здійснюватися

виключно після автентифікації суб'єкта та динамічної авторизації кожного окремого запиту із залученням механізмів оцінки контексту доступу, що враховує не лише ідентифікаційні атрибути суб'єкта, а й технічні характеристики пристрою, зокрема тип пристрою, часові параметри запиту, місцезнаходження, попередню поведінку при доступі, рівень ризику, поведінкові атрибути та інші параметри, які визначаються динамічною політикою контролю доступу для забезпечення глибокого аналізу подій у випадку прийняття рішення про надання доступу або при виникненні інцидентів безпеки.

В свою чергу, внутрішні сховища даних повинні бути ізольованими від загальної корпоративної мережі на рівні маршрутизації та міжмережевих політик, із застосуванням принципу мікросегментації. Усі запити до таких ресурсів повинні проходити виключно через PEP, які, у взаємодії з точкою прийняття рішення про політику (Policy Decision Point – PDP) та відповідно до принципу найменших привілеїв, забезпечують надання доступу лише до тієї інформації, яка є необхідною для виконання конкретного службового завдання, без можливості несанкціонованого доступу до інших частин системи, переміщення між сегментами мережі або підвищення привілеїв (privilege escalation). Водночас, навіть за умов впровадження мікросегментації мережевої інфраструктури та ізоляції внутрішніх сховищ даних, постає потреба у застосуванні додаткових механізмів безпеки, орієнтованих безпосередньо на забезпечення конфіденційності та цілісності інформації, яка зберігається. У цьому контексті ключову роль відіграють сучасні криптографічні механізми, зокрема методи шифрування та протоколи підтвердження прав доступу на основі концепції ZK, які дозволяють підтверджувати достовірність даних або ідентичність суб'єкта без передачі самих даних стороні, що здійснює перевірку та прийняття рішення. Це значно знижує ризики витоку конфіденційної інформації, зокрема у випадках, коли певні компоненти інфраструктури виявляються вразливими або скомпрометованими. Застосування зазначених механізмів має бути інтегрованим у загальну модель контролю доступу, що реалізується в межах ZTA, із дотриманням принципу найменших привілеїв, а також із використанням моделей контролю доступу RBAC та/або ABAC. Це забезпечує гнучке та динамічне керування доступом із урахуванням поточного контексту запиту та рівня довіри до суб'єкта, що мінімізує ризики бічного переміщення або підвищення привілеїв. Крім того, усі запити на отримання доступу до баз і сховищ даних повинні фіксуватися у спеціалізованих журналах подій із забезпеченням цілісності відповідних записів, їх подальшою перевіркою та автоматизованим аналізом у рамках постійного моніторингу. Даний підхід дозволяє своєчасно виявляти спроби порушення політик доступу, ознаки аномальної поведінки або інші події, що можуть свідчити про потенційну загрозу.

Таким чином, у контексті сценарію з підвищеним рівнем ризику, зокрема за участі зовнішніх підрядників, ефективне керування доступом до внутрішніх сховищ даних вимагає впровадження багаторівневої моделі безпеки. Така модель має ґрунтуватися на мікросегментації інфраструктури, динамічній авторизації, застосуванні криптографічних механізмів захисту, включно з ZK та безперервному моніторингу. Комплексне поєднання цих підходів забезпечує реалізацію принципів нульової довіри та сприяє формуванню адаптивного й стійкого механізму захисту корпоративних інформаційних ресурсів.

1.4. Співпраця між підприємствами

Співпраця між підприємствами є ще одним поширеним сценарієм (наприклад, реалізація спільного проєкту може передбачати залучення співробітників двох підприємств (рис. 4)). Ці підприємства можуть бути окремими державними установами або навіть державною установою та приватним підприємством. У цьому випадку, відповідно до рис. 4, *Підприємство А* керує базою даних, необхідною для виконання проєкту, і повинне надати доступ до певної інформації окремим співробітникам *Підприємства Б*. Найпростішим підходом може бути створення спеціальних облікових записів для співробітників *Підприємства Б* з доступом лише до необхідних даних, проте такий підхід ускладнює керування доступом у міру зростання кількості користувачів. Використання централізованої системи керування ідентифіка-

цією дозволяє спростити цей процес, за умови, що PEP обох підприємств підтримують автентифікацію суб'єктів у межах спільної системи керування ідентифікацією.

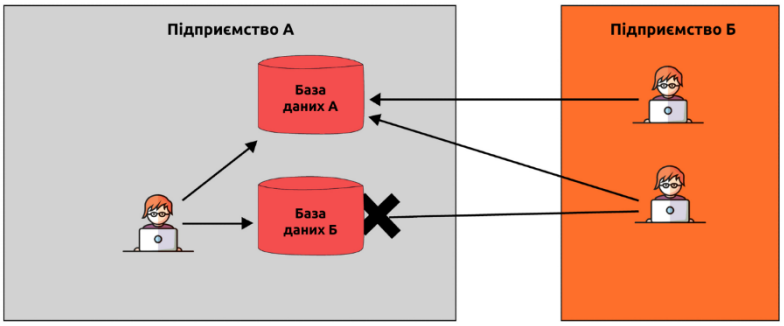


Рис. 4. Сценарій взаємодії між підприємствами

Даний сценарій частково подібний до сценарію підприємства з віддаленими структурними підрозділами, оскільки співробітники обох підприємств можуть працювати поза межами корпоративної мережі, а необхідний для них ресурс може бути розміщений як у локальному середовищі підприємства, так і в хмарному середовищі. Це усуває необхідність налаштовувати складні правила для міжмережевих екранів або створювати розширені списки контролю доступу (Access Control List – ACL), що визначають певні IP-адреси *Підприємства Б,* які можуть отримати доступ до ресурсів *Підприємства А* на основі політик доступу *Підприємства А*. Так само, як і у сценарії підприємства з віддаленими структурними підрозділами, використання PE/PA, які розгорнуті у хмарному середовищі, може забезпечити доступність ресурсів для співробітників обох підприємств без необхідності створення віртуальної приватної мережі (Virtual Private Network – VPN) або подібних рішень. Співробітники *Підприємства Б* можуть отримати доступ до необхідних ресурсів шляхом встановлення клієнтського ПЗ (наприклад, агента на кінцевому пристрої) або через відповідний безпечний/захищений веб-шлюз (Secure Web Gateway – SWG), відповідно до моделі розгортання на основі порталу ресурсів, де PEP є єдиним компонентом, який виконує роль шлюзу для відповідних запитів.

Однак у випадку співпраці між підприємствами захист баз і сховищ даних вимагає впровадження додаткових механізмів, які враховують використання корпоративних ресурсів зовнішніми користувачами. Навіть за умови обмеження доступу до частини інформації, залишається ризик порушення цілісності або розголошення даних унаслідок помилок у налаштуванні прав доступу або зламу засобів захисту. Для мінімізації таких ризиків, бази даних мають проєктуватися відповідно до принципу найменших привілеїв, що передбачає розмежування даних за типом інформації, ролями та/або атрибутами користувачів. На практиці це, наприклад, може досягатися шляхом використання віртуальних таблиць (уявлень – view) у системах керування традиційними базами даних, або впровадження політик обмеження доступу до окремих полів у нереляційних сховищах. Зовнішнім користувачам надається доступ лише до певних уявлень, причому сам доступ здійснюється через відповідні шлюзи PEP, які перевіряють не тільки облікові дані користувача, але й додаткові атрибути, серед яких можуть бути, наприклад, роль, тип пристрою, часові параметри запиту, місцезнаходження, попередня поведінка при доступі, тип та стан сеансу, поведінкові атрибути тощо. Це дозволяє застосовувати динамічні правила доступу, які враховують поточні умови взаємодії.

Проте захист даних вимагає також збереження конфіденційності не лише їхнього змісту, але й фактів звернення до них. У випадках обробки фінансової, службової або технічної інформації навіть доступ до метаданих може становити загрозу витоку. Для усунення подібних ризиків у рамках ZTA повинні застосовуватися механізми, що забезпечують безпеку даних на основі концепції ZK, тобто механізмів, які забезпечують підтвердження прав доступу без передачі облікових даних або їх атрибутів. Зокрема, використання протоколів на основі *zk-SNARK* або *zk-STARK* [40] у межах хмарних шлюзів дозволяє перевіряти рівень доступу користувача без розкриття його ідентифікатора чи змісту даних, до яких здійснюється запит.

Інтеграція механізмів ZK в ZTA забезпечує керований доступ до інформації незалежно від місцезнаходження користувача або типу пристрою. Всі перевірки здійснюються на рівні PEP, що поєднують політики безпеки організації із атрибутами (характеристиками) користувача та середовища доступу. Завдяки цьому ресурси залишаються ізольованими, а доступ надається лише за умов, які є достатніми й обґрунтованими з позиції безпеки. Однак постійний розвиток IT-технологій потребує ще глибшої інтеграції таких підходів безпосередньо у внутрішню архітектуру інформаційних систем, зокрема, інтеграції механізмів ZK безпосередньо у системи керування базами даних. Щодо цього, цікавий підхід до використання концепції ZK для баз даних запропонований у роботі [41]. Такі рішення відкривають нові можливості для побудови середовищ із підвищеним рівнем довіри навіть в інфраструктурі із спільним використанням ресурсів або під час обробки даних у спільних інформаційних системах.

1.5. Підприємство, що надає публічні або сервісні послуги

Підприємства, що надають публічні або сервісні послуги, можуть працювати з користувачами, які не є частиною організації, але взаємодіють із її цифровими ресурсами. Доступ до таких ресурсів може вимагати або не вимагати реєстрації користувачів, тобто створення ними облікових записів або отримання наданих підприємством облікових даних. Такі сервіси можуть бути розраховані на широке коло користувачів (наприклад, загальнодоступні інформаційні портали) або обмежену групу (наприклад, клієнтів компанії чи певних зовнішніх суб'єктів). У всіх цих випадках є ймовірність, що доступ до ресурсів здійснюється з пристроїв, які не належать підприємству, а отже, можливості застосування корпоративних політик кібербезпеки є обмеженими.

Якщо доступ до ресурсу здійснюється без реєстрації (наприклад, загальнодоступний веб-сайт або веб-сервіс), застосування принципів ZTA є обмеженим, оскільки підприємство не може контролювати стан пристроїв, із яких надходять запити, або встановлювати власні політики доступу. В разі, якщо користувачі проходять процедуру реєстрації, підприємство може запроваджувати власні політики безпеки, серед яких вимоги до складності паролів, їхнього життєвого циклу та інших параметрів, а також використання MFA. Водночас можливості впровадження більш детальних політик залишаються обмеженими для цієї категорії користувачів, оскільки підприємство не контролює їх зовнішні пристрої. В такому разі, аналіз вхідних запитів може відігравати важливу роль у моніторингу стану сервісу та виявленні потенційних кібератак, що імітують дії легітимних користувачів. Наприклад, якщо зареєстровані користувачі зазвичай отримують доступ до порталу за допомогою певного набору типових веб-браузерів, а система фіксує аномальне зростання запитів із невідомих типів браузерів або відомих застарілих версій, що може свідчити про автоматизовану атаку. У такому випадку підприємство може вжити заходи для обмеження запитів від виявлених клієнтів або впровадити додаткові механізми перевірки для подібних запитів. Крім того, підприємство повинне враховувати чинні законодавчі норми та вимоги щодо збору, обробки та зберігання даних користувачів, які взаємодіють із сервісом.

Окрему увагу в цьому сценарії необхідно приділити безпеці баз і сховищ даних, що використовуються для зберігання облікових записів, журналів доступу, історії транзакцій та інших персональних і конфіденційних даних користувачів. Використання концепції нульової довіри передбачає обмеження доступу до таких ресурсів навіть серед внутрішніх суб'єктів підприємства, тобто кожен запит до бази даних повинен здійснюватися лише суб'єктами, які пройшли автентифікацію та відповідати визначеним політикам безпеки. Наприклад, замість надання безперервного доступу для сервісу до бази даних користувачів, можна впровадити тимчасові облікові токени автентифікації з обмеженим терміном дії та визначеними правами доступу, що зменшує ризики у разі компрометації окремої компоненти системи. Крім того, підвищення рівня захисту може бути досягнуто завдяки впровадженню механізмів ZK, які дозволяють системі з нульовою довірою переконатися, що суб'єкт є справжнім (автентичним), а запит дійсним. PDP/PEP приймає належне рішення, щоб дозволити суб'єкту отрима-

ти доступ до ресурсу. Це знижує ризик витоку критичної інформації у разі перехоплення мережевого трафіку або успішної атаки на один із проміжних вузлів інфраструктури. В свою чергу керування доступом залежить від стану безпеки пристрою (механізму, засобу) та інших ситуативних факторів (наприклад, часу та місцезнаходження, попередньої поведінки при доступі тощо), які можуть вплинути на рівень довіри до того, як доступ до ресурсу буде наданий відповідно до певних політик. Загалом, підприємствам необхідно розробити та підтримувати динамічну політику доступу до ресурсів, що базується на оцінці ризиків, і налаштувати систему, яка гарантуватиме правильне та послідовне застосування цих політик для окремих запитів на доступ. Підприємству не слід покладатися на передбачувану надійність, коли суб'єкт відповідає базовому рівню автентифікації (наприклад, при вході до системи), а усі наступні запити до ресурсів вважаються однаково дійсними. Проте важливо розуміти, що при впровадженні таких підходів необхідно забезпечити баланс між рівнем безпеки та зручністю для користувачів, що особливо критично для ресурсів, орієнтованих на широку аудиторію.

В цілому реалізація ZTA є досить складним процесом, причому це швидше за все шлях (рух), ніж повна заміна інфраструктури або технологічних процесів. Тому підприємство повинне прагнути до поступового впровадження принципів нульової довіри, змін у процесах та технологічних рішень, які захищають її найцінніші активи даних. Більшість підприємств, швидше за все, продовжуватимуть працювати у комбінованому режимі з використанням нульової довіри та периметра протягом невизначеного періоду часу, продовжуючи вкладати кошти у постійну модернізацію IT [7]. При цьому наявність плану модернізації IT, що включає перехід до ZTA, може допомогти підприємству сформувати дорожні карти для здійснення невеликих (поступових) переходів на нові робочі процеси. Зрештою, те, як підприємство переходитиме на концепцію нульової довіри, буде залежати від його поточного стану кібербезпеки та операційної діяльності. Причому підприємство має досягти базового рівня підготовленості (базовий рівень включає визначення та класифікацію активів, суб'єктів, бізнес-процесів, потоків даних і відображення залежностей для підприємства), перш ніж стане можливим розгорнути масштабну систему, орієнтовану на нульову довіру. Підприємству необхідна ця інформація, щоб визначити список бізнес-процесів-кандидатів та суб'єктів/активів, які будуть залучені до цього процесу. При цьому, найбільшою проблемою, що перешкоджає впровадженню успішних рішень в області нульової довіри, на думку фахівців ACT-IAC (American Council for Technology and Industry Advisory Council – Американська рада з технологій та Консультативна рада з питань промисловості), може бути загальний недостатній рівень кібербезпеки [42]. Ними зазначається, наприклад, що більшість державних установ не мають фундаментальних основ (таких як, політик, процесів та інструментів), необхідних для розгортання систем, що відповідають концепції нульової довіри.

Таким чином, можна зробити висновок, що процес впровадження ZTA повинен бути поступовим та орієнтованим на розвиток існуючих технологій та IT-середовищ в організації. Для цього необхідно розробляти і тестувати різні практичні рішення ZTA, керуючись загальноприйнятими принципами та рекомендаціями в галузі кібербезпеки. Важливим аспектом є детальний аналіз архітектури кожного рішення, використаних технологій, специфічних конфігурацій та інтеграцій, а також їх відповідності сучасним стандартам кібербезпеки та регуляторним вимогам. Тому, кожна організація повинна розробляти власну стратегію впровадження ZTA, оскільки єдиного універсального підходу не існує. ZTA є не технічним стандартом, а концептуальною моделлю, що передбачає постійне вдосконалення політик і процесів керування доступом відповідно до її принципів. При цьому доречно відразу ж звернути увагу на те, що керівництво IT-підприємства має прагнути до поступового впровадження принципів нульової довіри, зміни процесів та технологічних рішень, що захищають його найцінніші активи даних. Такий поступовий підхід дозволить знизити ризик відмов і помилок у системі, допоможе зрозуміти подальші процеси розгортання елементів системи, а також полегшить перехід персоналу до нової архітектури.

2. Пропонована модель впровадження архітектури нульової довіри

Залежно від того як налаштовано корпоративну мережу для різних бізнес-процесів на одному підприємстві можуть використовуватися кілька моделей розгортання ZTA. У цьому контексті для розуміння доцільно звернутися до розробленої відповідно до загальноприйнятих галузевих та академічних принципів нульової довіри Міністерством оборони США так званої моделі стовпів та можливостей [43, 44] (див. рис. 5).

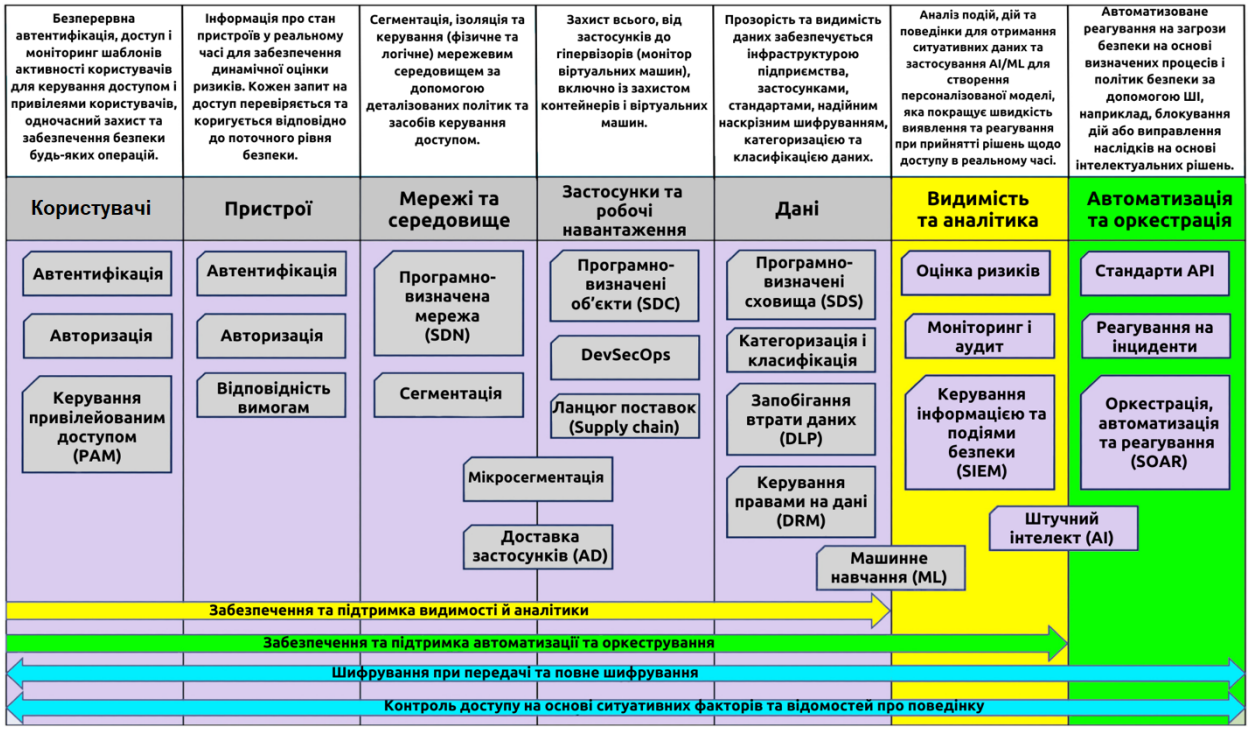


Рис. 5. Стовпи та можливості моделі впровадження архітектури нульової довіри

Стовп (*pillar*) – це ключова область для впровадження систем контролю/захисту нульової довіри У моделі, що розглядається, їх сім: користувачі, пристрої, мережі та середовища, застосунки та робочі навантаження, дані, видимість та аналітика, а також автоматизація та оркестрація. *Можливості* (*capabilities*) – це здатність досягати бажаного ефекту відповідно до певних стандартів та умов за допомогою комбінації способів і засобів (дій та ресурсів) для виконання певного набору дій. Стовпи асоціюються з такими можливостями, як автентифікація ідентифікаційних даних (особистості) та програмно-визначуване підприємство (Software Defined Enterprise). Загалом, запропонована вище модель передбачає використання різних механізмів (способів, засобів) автентифікації та авторизації користувачів і пристроїв, керування привілейованим доступом (Privileged Access Management – PAM), мікросегментацію мережі, програмно-визначену інфраструктуру (SDN, SDC, SDS), контроль доступу до даних (DRM, DLP), підхід DevSecOps (Development, Security, and Operations), а також рішення для моніторингу, оцінки ризиків і реагування на інциденти (SIEM, SOAR, AI/ML) тощо. Усі ці елементи працюють у взаємозв'язку, забезпечуючи видимість і аналітику, автоматизацію та оркестрацію процесів безпеки, шифрування при передачі та повне шифрування, а також ефективний контроль доступу на основі ситуативних факторів та відомостей про поведінку. Для досягнення належної інтеграції між стовпами та можливостями потрібно всеосяжне управління. Основні стовпи та можливості забезпечують максимальну видимість та захист даних, що є ключовим моментом за будь-якої реалізації нульової довіри. У цьому сенсі, стовпи моделі нульової довіри можна представити у вигляді взаємопов'язаних частин пазла навколо стовпа даних, оточеного стовпами захисту [44]. Усі стовпи захисту працюють у комплексі для ефективного захисту стовпа даних. Взаємодія цих компонентів забезпечує функціонування корпоративного середовища відповідно до принципів ZTA.

Аналіз представленої на рис. 5 моделі дозволяє зробити висновок, що для її впровадження необхідно використовувати кілька моделей розгортання ZTA (див. п. 1), зокрема, моделі розгортання з мікросегментацією, з використанням хмарної маршрутизації, на основі ресурсів. Тобто цю модель певною мірою через охоплення різних моделей розгортання можна вважати комплексною, яка потребує узгодженого застосування технологій та механізмів безпеки для забезпечення належного рівня захисту корпоративного середовища. У зв'язку з цим постає питання стандартизації впровадження принципів нульової довіри. Попри відсутність єдиної загальноприйнятої методології, все ж таки деякий підхід впровадження було запропоновано Міністерством оборони США [45]. У цьому підході/моделі застосовуються/передбачаються два списки стандартів, а саме так звані профільні стандарти (Standards Profile – StdV-1) і стандарти прогнозування (Standards Forecast – StdV-2), що забезпечує системний підхід до впровадження ZTA в інформаційних системах. Профільні стандарти визначають набір технологій, нормативно-правових актів, політик, а також тактик, технік і процедур, що регулюють функціонування компонентів нульової довіри. Стандарти прогнозування охоплюють технологічні, операційні та бізнес-стандарти, що визначають перспективні напрями розвитку можливостей нульової довіри, а також майбутні вимоги до їхньої реалізації. Це забезпечує структурований підхід до впровадження ZTA, поєднуючи як поточні регуляторні вимоги, так і стратегічне планування розвитку архітектури. У табл. 2 представлено набори можливостей, запропоновані Міністерством оборони США [44], що відповідають різним рівням забезпечення безпеки та захисту *даних, застосунків, активів* та *послуг* (*data, applications, assets and services* – DAAS) для кожного стовпа нульової довіри, пов'язані з поетапністю впровадження ZTA на підприємстві.

Цільовий рівень (*Target Level*) – це мінімальний набір можливостей нульової довіри та видів діяльності (activities), необхідних для забезпечення безпеки та захисту даних, застосунків, активів та послуг. Кожна можливість нульової довіри відповідає одному із семи стовпів нульової довіри. Незважаючи на те, що повний набір можливостей охоплює діапазон від цільового (*Target*) до просунутого/розширеного рівня (*Advanced*), деякі можливості досягаються лише на цільовому рівні, а деякі лише на розширеному. Більшість можливостей мають асоційовані активності (види діяльності) як на цільовому рівні, так і на просунутому/розширеному рівні нульової довіри. Подібний підхід дозволяє поступово впроваджувати необхідні заходи безпеки, починаючи з критично важливих механізмів та завершуючи їхньою повною інтеграцією у всі аспекти керування кіберзахистом підприємства. *Цільовий* (*Target*) або *базовий рівень* охоплює критично необхідні механізми захисту та передбачає впровадження інвентаризації користувачів і пристроїв, уніфіковане керування кінцевими пристроями, визначення потоків даних та застосування принципу найменших привілеїв тощо. Дані механізми є основою для подальшого розвитку архітектури безпеки. В свою чергу *середній* (*Target & Advanced*) *рівень* розширює функціональність шляхом впровадження додаткових механізмів ідентифікації, MFA, контролю відповідності пристроїв вимогам безпеки, програмно-визначених мереж, мікросегментації тощо. Також реалізуються механізми авторизації пристроїв у режимі реального часу, контроль їх відповідності політикам безпеки та розширений моніторинг активності користувачів, що дозволяє оперативно реагувати на потенційні загрози. *Розширений* (*Advanced*) *рівень* включає динамічні методи контролю доступу, такі як безперервна автентифікація та аналіз поведінкових факторів. Здійснюється автоматизація процесів кібербезпеки, впровадження інтелектуальних механізмів реагування та запобігання втраті даних. Додатково впроваджуються ШІ та автоматизація політик безпеки, що підвищує рівень адаптивності системи безпеки. Водночас, варто зазначити, що окрім вище вказаного визначаються інструменти реалізації, що охоплюють як технічні, так і організаційні аспекти при впровадженні ZTA, зокрема стратегічне планування, матеріальне забезпечення, навчання персоналу та розробку політик безпеки. Тому, враховуючи поступовий характер впровадження, більшість функціональних можливостей реалізується поетапно,

охоплюючи як базовий, так і середній рівні, що забезпечує ефективне впровадження та адаптацію організації до принципів нульової довіри.

Таблиця 2

Набори можливостей моделі впровадження архітектури нульової довіри

	Цільовий рівень, що відповідає певному набору можливостей нульової довіри, необхідних для забезпечення безпеки та захисту DAAS		
	Target (мінімальний набір)	***Target & Advanced*** (мінімальний і розширений набір)	***Advanced*** (розширений набір)
Користувачі	✓ Інвентаризація користувачів. ✓ Принцип найменших привілеїв.	✓ Керування привілейованим доступом (PAM) та контрольований доступ користувачів. ✓ Багатофакторна автентифікація. ✓ Механізм Identity Federation. ✓ Контроль доступу на основі ситуативних факторів, відомостей про поведінку, біометрії. ✓ Безперервна автентифікація. ✓ Впровадження платформи ICAM.	
Пристрої	✓ Частково та повністю автоматизоване керування активами, вразливостями та виправленнями. ✓ Уніфіковане керування кінцевими та мобільними пристроями.	✓ Авторизація пристроїв з перевіркою в режимі реального часу. ✓ Контроль пристроїв та відповідність вимогам. ✓ Інвентаризація пристроїв та віддалений доступ. ✓ Захист кінцевих точок і розширене виявлення та реагування (EDR і XDR).	
Мережі та середовище	✓ Визначення потоків даних. ✓ Макросегментація.	✓ Програмно-визначені мережі. ✓ Мікросегментація.	
Застосунки та робочі навантаження	✓ Інвентаризація застосунків. ✓ Керування ризиками програмного забезпечення.	✓ Безпечна розробка та впровадження програмного забезпечення. ✓ Авторизація та інтеграція ресурсів.	✓ Безперервний моніторинг та постійні авторизації.
Дані	✓ Керування ризиками баз, сховищ даних. ✓ Керування корпоративними даними.	✓ Категоризація і класифікація даних. ✓ Моніторинг та аналіз даних, керування доступом до даних. ✓ Шифрування даних та керування правами. ✓ Запобігання втрати даних (DLP).	
Видимість та аналітика	✓ Ведення журналу всього трафіку та впровадження розвідки загроз. ✓ Загальна аналітика безпеки та ризиків.	✓ Керування інформацією та подіями безпеки (SIEM). ✓ Аналіз поведінки користувачів і суб'єктів (UEBA).	✓ Автоматизовані динамічні політики.
Автоматизація та оркестрація	✓ Машинне навчання. ✓ Стандартизація API.	✓ Точка прийняття рішень (PDP) та оркестрування політик. ✓ Автоматизація процесів. ✓ Оркестрація, автоматизація та реагування (SOAR). ✓ Операційний центр безпеки (SOC) і реагування на інциденти (IR).	✓ Штучний інтелект.

Спираючись на досвід впровадження систем нульової довіри відомими у світі різними організаціями та компаніями, у роботі [15] були сформульовані деякі рекомендації щодо успішного впровадження ZTA на типовому IT-підприємстві у вигляді послідовності деяких кроків/етапів. Орієнтуючись на роботу [15], а також на перевірені практики та дослідження викладені вище, можна розробити більш ефективну модель для впровадження ZTA, яка в тому числі використовуватиме і можливості вже існуючої інфраструктури інформаційної

системи підприємства. Пропонований далі підхід дозволить ретельно та продумано застосовувати засоби контролю ZTA, які найкраще захищають бізнес, водночас не створюючи перешкод для його основних операційних процесів та гнучкої роботи в умовах сучасного цифрового середовища.

Як відомо, перш ніж розпочати впровадження ZTA на підприємстві, необхідно провести детальний аналіз активів, суб'єктів, потоків даних та робочих процесів. Це дозволить отримати повне уявлення про поточний стан операційної діяльності та визначити необхідні зміни в інфраструктурі. Крім того, щоб IT-підприємство, побудоване на основі принципів нульової довіри, могло успішно функціонувати, PE повинен мати знання про суб'єктів підприємства, серед яких можуть бути як фізичні користувачі, так і нефізичні сутності, зокрема облікові записи сервісів, що взаємодіють із ресурсами [15]. Водночас, концепція нульової довіри може бути ефективно інтегрована з підходом розробки ПЗ та IT-операцій (Development and Operations – DevOps), який сприяє автоматизації процесів життєвого циклу ПЗ, включно з його розробкою, тестуванням, інтеграцією та розгортанням. DevOps забезпечує швидку адаптацію інформаційних систем до змін, підвищуючи їхню ефективність та стійкість [19, 46, 47]. Додатково, інтеграція механізмів безпеки на всіх етапах життєвого циклу можлива завдяки розширеному підходу DevSecOps, який орієнтований на проактивне виявлення загроз та мінімізацію вразливостей ще до розгортання рішень [48, 49]. Інтеграція DevSecOps у процес впровадження ZTA дозволяє створити циклічний процес управління безпекою, що включає безперервний моніторинг, ефективне реагування на кіберзагрози, аналіз ризиків, адаптивне оновлення політик доступу. Такий підхід підвищує ефективність міграції до ZTA, знижує ймовірність збоїв під час переходу та забезпечує узгодженість роботи інформаційних систем із сучасними вимогами кібербезпеки. Крім того, цей підхід дозволяє підприємству підвищити якість і стабільність роботи сервісів, а також скоротити час розгортання продуктів завдяки швидкому отриманню зворотного зв'язку від користувачів. З точки зору впровадження ZTA, підхід DevSecOps дозволяє оптимізувати процес міграції, оскільки обидва підходи базуються на поступових ітераціях, високій залученості зацікавлених сторін та постійному моніторингу результатів. Таким чином, використання DevSecOps у поєднанні з ZTA дозволяє не лише покращити керованість процесом міграції, а й мінімізувати будь-які ризики, пов'язані з його впровадженням.

Впровадження ZTA потребує структурованого підходу, що враховує як технічні, так і організаційні аспекти. Враховуючи виклики та вимоги щодо впровадження ZTA, пропонується модель впровадження, яка включатиме кілька ключових процесів. Назвемо їх: 1) стратегія нульової довіри, 2) оцінка середовища підприємства, 3) підготовка до впровадження ZTA, 4) перехід до ZTA, 5) моніторинг, обслуговування та оптимізація ZTA.

Оскільки підхід DevSecOps орієнтований на автоматизацію, ітеративний підхід і постійний контроль безпеки, запропонована модель передбачає інтеграцію даних принципів на всіх етапах впровадження ZTA. Це дозволяє не лише підвищити стійкість та ефективність інформаційних систем, а й забезпечити відповідність сучасним стандартам кібербезпеки. На рис. 6 наведено пропоновану модель поетапного впровадження ZTA на IT-підприємстві, яка включає в себе, у тому числі, концепції підходу DevSecOps та охоплює компоненти розширеної моделі нульової довіри ZTX [5], а саме оточуючі елементи (робочі навантаження (workloads), мережі (networks), пристрої (devices) і люди (people), які є провідниками даних і, отже, також потребують захисту), а також видимість та аналітика/аналіз (visibility and analytics), автоматизація (automation) та оркестрування/оркестрація (orchestration). Окрім цього при впровадженні ZTA необхідно враховувати дотримання організаціями вимог галузевих практик і стандартів.

Також слід відмітити, що поетапне впровадження ZTA в рамках запропонованої моделі, окрім основних процесів, передбачає виконання низки підпроцесів, які деталізують кожен процес реалізації ZTA. У табл. 3 наведено структуру цих процесів, їхній зміст у вигляді підп-

роцесів і послідовність виконання етапів, необхідних для впровадження та подальшого ефективного функціонування ZTA.

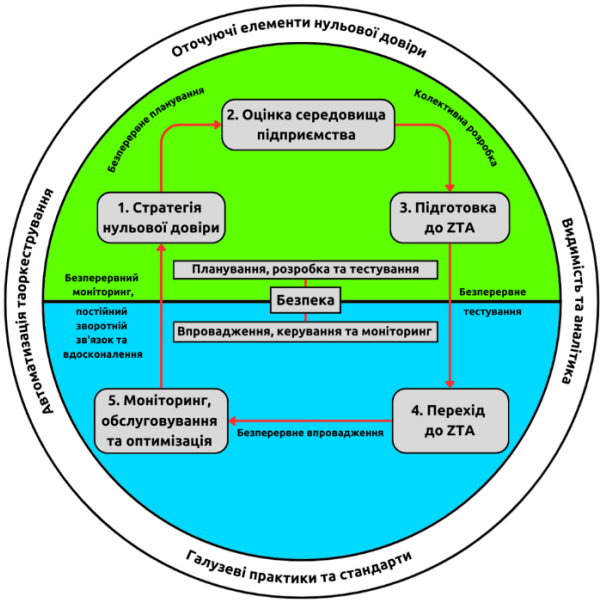


Рис. 6. Пропонована модель впровадження архітектури нульової довіри

Таблиця 3

Процеси та підпроцеси поетапного впровадження архітектури нульової довіри

Процес	Підпроцеси
Стратегія нульової довіри підприємства	Етап 1. Створення стратегії нульової довіри. Етап 2. Створення команди для впровадження ZTA.
Оцінка середовища підприємства	Етап 3. Виявлення та інвентаризація активів підприємства. Етап 4. Розробка політик доступу. Етап 5. Визначення існуючих можливостей та технологій забезпечення безпеки. Етап 6. Створення плану міграції.
Підготовка до впровадження ZTA	Етап 7. Усунення недоліків у політиках та процесах забезпечення нульової довіри. Етап 8. Підготовка пристроїв, користувачів та мережі.
Перехід до ZTA	Етап 9. Впровадження компонентів ZTA, поступове використання існуючих безпекових рішень для досягнення кінцевої мети. Етап 10. Перевірка реалізації для підтвердження підсумкових досягнень під час розгортання ZTA та виправлення виявлених помилок після впровадження.
Моніторинг, обслуговування та оптимізація ZTA	Етап 11. Забезпечення безперервного моніторингу та технічного обслуговування екосистеми нульової довіри відповідно до виявлених помилок, актуальних загроз та вимог безпеки. Етап 12. Впровадження автоматизації процесів, постійне вдосконалення та розвиток відповідно до змін характеру загроз, завдань, технологій та нормативних документів.

2.1. Стратегія нульової довіри

Першим кроком для впровадження ZTA є розробка відповідної стратегії, яка базується на принципах нульової довіри. Чітке формулювання цієї стратегії є критично важливим для успішного впровадження ZTA на підприємстві. Після її розробки необхідно досягти підтримки з боку ключових зацікавлених сторін, зокрема керівництва та кінцевих користувачів, які мають усвідомити важливість переходу до ZTA та активно сприяти цьому процесу. Крім того, для ефективного переходу до ZTA необхідно створити спеціалізовані команди, що включатимуть фахівців з впровадження рішень нульової довіри та осіб, які приймають рішення в сфері інформаційних технологій всередині підприємства. Результатом цього етапу є розробка комплексного плану переходу до ZTA, який визначатиме основні цілі та напрями впровадження ZTA для всіх учасників процесу [50].

Етап 1. Створення стратегії нульової довіри. На початковому етапі впровадження ZTA підприємство повинно сформулювати стратегічне бачення та визначити ключові цілі реалізації цього підходу. Важливим аспектом є забезпечення підтримки з боку керівництва, постачальників, партнерів та кінцевих користувачів. Основним завданням є формування єдиного розуміння необхідності переходу до ZTA, що зумовлено потребою підвищення адаптивності бізнес-процесів, оптимізації керування IT-інфраструктурою та відповідності нормативним вимогам. Окрім внутрішніх заходів, підприємство має забезпечити відповідність стандартам безпеки та нормативно-правовим вимогам, що регулюють впровадження ZTA. Це передбачає взаємодію з державними регуляторами та органами сертифікації, які встановлюють вимоги до інформаційної безпеки, а також врахування рекомендацій галузевих стандартів та міжнародних протоколів. Забезпечення ефективної комунікації з цими суб'єктами сприятиме усуненню можливих розбіжностей у трактуванні концепції ZTA та мінімізації ризиків затримки при її впровадженні [30]. Стратегічний план переходу до ZTA повинен містити комплексний аналіз поточного середовища, оцінку доступних технологічних рішень та визначення масштабу впровадження [51]. На основі цього аналізу, доцільно здійснити тестування на обмеженій частині IT-інфраструктури, що дозволить оцінити ефективність обраних рішень, виявити потенційні вразливості та сформувати оптимальний підхід до подальшого масштабування ZTA. Це також дозволить організації адаптувати ZTA відповідно до актуальних потреб та вимог підприємства [30].

Етап 2. Створення команди для впровадження ZTA. Для успішного впровадження ZTA важливим елементом є створення відповідної команди, яка забезпечить безперешкодний перехід до реалізації концепції нульової довіри. Команда повинна бути міжфункціональною, що передбачає залучення спеціалістів з різних підрозділів, таких як керування бізнес-процесами та IT. Команда повинна об'єднувати як керівників, які приймають стратегічні рішення, бізнес-аналітиків, а також технічних експертів з безпеки, зокрема які мають досвід у сфері захисту даних (включно із безпекою баз і сховищ даних), застосунків, мережі та інфраструктури, ідентифікації та керування користувачами і пристроями тощо [7, 29]. До основних завдань команди належить визначення стратегії впровадження ZTA, контроль за виконанням планів і забезпечення підтримки на всіх етапах реалізації стратегії нульової довіри. Важливим є ухвалення рішень щодо впровадження нових технологій та ініціатив, перевірка поточних рішень, визначення архітектури підприємства, а також просування технічних рішень у реальне IT-середовище. Окрім того, важливою складовою є ефективна комунікація в процесі впровадження ZTA з усіма сторонами, зокрема з користувачами, оскільки недостатнє інформування може призвести до непорозумінь і неспроможності ефективно реалізовувати концепцію нульової довіри на підприємстві. Для цього організація повинна розробити відповідний план комунікацій або провести кампанію, що забезпечить підвищення обізнаності та інформування користувачів на ранніх етапах впровадження, що сприятиме отриманню їхньої підтримки [52]. Для комунікації можна застосовувати різноманітні методи, зокрема особисті зустрічі, конференції на рівні компанії, електронну пошту, інформаційні сервіси, відео та блоги.

2.2. Оцінка середовища підприємства

Оцінка середовища є комплексним процесом при впровадженні ZTA та передбачає аналіз поточного стану підприємства, зокрема виявлення та інвентаризацію активів, розробку політик доступу, оцінку існуючих заходів безпеки, а також розробку стратегії міграції до нової архітектури. Основною метою оцінки є визначення ключових активів та ресурсів, формування деталізованого уявлення про поточний рівень захищеності підприємства, а також ідентифікація потенційних ризиків і вразливостей. Важливим компонентом цього процесу є розробка політик доступу для підтримки завдань та корпоративних сценаріїв використання, а також тестування на проникнення (penetration testing), яке дозволяє виявити слабкі місця в інфраструктурі та оцінити ефективність існуючих механізмів захисту. Результатом оцінки

середовища є створення базового плану міграції до ZTA, який визначає поточний стан підприємства та слугує основою для подальшого впровадження концепції нульової довіри.

Етап 3. Виявлення та інвентаризація активів підприємства. Першочерговим завданням команди впровадження ZTA є ідентифікація всіх активів, що використовуються в IT-середовищі підприємства. Це передбачає визначення наявних ресурсів, зокрема обладнання, ПЗ, застосунків, даних і сервісів. Для досягнення цієї мети може знадобитися впровадження спеціалізованих засобів моніторингу трафіку, які дозволяють виявити активні елементи інфраструктури та ресурси, до яких здійснюється доступ. Важливо отримати повне уявлення про всі ресурси, як локальні, так і хмарні, та провести їх систематизований облік. Це включає визначення кількості, місця розташування, рівня захисту, значущості та впливу кожного ресурсу на функціонування підприємства, адже саме вони стануть основними об'єктами, які необхідно захищати в межах ZTA. Додатково слід приділити увагу пристроям, що не належать підприємству, але можуть бути підключені до його мережевої інфраструктури або мати доступ до корпоративних ресурсів. Після складання повного переліку активів необхідно оцінити їхній поточний стан: чи відповідають вони сучасним вимогам, чи потребують оновлення, а також чи виникають у користувачів труднощі з доступом до них. Наявність неврахованих ресурсів несе ризик їхньої недостатньої захищеності в рамках концепції нульової довіри, що може зробити їх вразливими до витоку даних, несанкціонованої модифікації, видалення, відмови в обслуговуванні чи інших атак [15]. Таким чином, після проведення виявлення та інвентаризації активів наступним кроком є розробка політик доступу, які визначатимуть умови використання цих ресурсів у межах ZTA. Це забезпечить контрольоване та безпечне керування доступом відповідно до принципів найменших привілеїв та поділу обов'язків.

Етап 4. Розробка політик доступу для підтримки завдань та корпоративних сценаріїв використання. Після ідентифікації всіх ресурсів, що потребують захисту, та визначення їхнього розташування, необхідно сформулювати політики доступу в межах ZTA. Вони мають регламентувати, хто та за яких обставин отримує доступ до кожного ресурсу. Політики доступу повинні відповідати принципам найменших привілеїв і поділу обов'язків, а їх розробка має враховувати категорії користувачів, які потребують доступу, їхні посади, умови контрактів, типи пристроїв та моделі власності (наприклад, особиста, корпоративна) тощо. Усі ці фактори впливають на формування політик [15]. Дозволи на доступ можуть бути обмежені залежно від місцезнаходження особи, яка запитує доступ, часу доби або інших параметрів, які можуть додатково обмежувати доступ без втручання у роботу підприємства. Формування політик має ґрунтуватися на критичності захищених ресурсів, зокрема з урахуванням особливостей реалізації багаторівневих політик у середовищах, що використовуються для створення, розгортання та керування сучасними хмарними застосунками [53]. Однак у процесі розробки політик підприємства можуть зіткнутися з певними труднощами. Зокрема, ZTA зазвичай складається з численних компонентів, які можуть виконувати функції як PE, так і PA. В результаті політика доступу не може бути централізована в одному місці, оскільки правила можуть розподілятися між різними компонентами [15]. Наприклад, деякі з них у системах EDR, інші в системах керування ідентифікацією, обліковими даними та доступом, треті – у компоненті мережевої безпеки або захисту даних. Відсутність єдиного централізованого сховища для всіх правил може ускладнити повне розуміння політик доступу, і підтримку структурованого та узгодженого керування. Тому, щоб ефективно керувати політиками доступу, підприємство повинно не лише чітко відстежувати правила, а й розуміти, в яких саме компонентах вони визначені. З огляду на це, наступним важливим кроком є визначення наявних елементів інфраструктури, що вже виконують функції безпеки. Це дозволить підприємству інтегрувати нові політики в існуючі технологічні рішення та засоби кібербезпеки. Тому після розробки політик доступу необхідно здійснити оцінку наявних можливостей та технологій забезпечення безпеки, щоб визначити, які з них можна адаптувати чи інтегрувати в нову модель безпеки.

Етап 5. Визначення існуючих можливостей та технологій забезпечення безпеки. Якщо підприємство планує впроваджувати ZTA в абсолютно новому середовищі, не маючи попередньо створеної IT-інфраструктури чи засобів безпеки, які потребували б адаптації, цей етап можна пропустити. Однак більшість організацій не розпочинають процес впровадження ZTA з нуля, а використовують вже наявні технологічні рішення та засоби безпеки, для них важливим є врахування існуючих елементів інфраструктури, які вже виконують певні функції захисту. Як правило, підприємства мають базові засоби безпеки, такі як мережеві брандмауери та системи виявлення вторгнень, що забезпечують захист периметра. Крім того, вони часто використовують рішення для керування ідентифікацією та контролю доступу, які дозволяють автентифікувати користувачів і надавати їм відповідні права доступу на основі ідентифікаційних даних і визначених ролей. На робочих станціях, ноутбуках та/або мобільних пристроях можуть бути встановлені системи EDR, що виконують функції брандмауера, антивірусного захисту тощо. Також багато організацій використовують інструменти безпеки для ведення журналів подій, контролю конфігурацій, управління вразливостями та інших процесів, пов'язаних із забезпеченням кібербезпеки. Крім цього, в структурі підприємства, як правило, функціонує центр керування безпекою, який відповідає за моніторинг, аналіз загроз та координацію заходів реагування [15].

На даному етапі командою проводиться визначення рівня захищеності інфраструктури підприємства та виявлення вразливих місць у системі безпеки. Першим кроком на цьому етапі є оцінка існуючих заходів безпеки, яка передбачає аналіз наявних механізмів захисту для перевірки відповідності принципам нульової довіри. Підприємство має ідентифікувати та описати наявні компоненти й функціональні можливості систем безпеки, щоб визначити, які механізми захисту вже впроваджені. Далі необхідно оцінити, чи можуть ці елементи й надалі забезпечувати належний рівень безпеки в межах впровадження ZTA або ж потребують модифікації (заміни). З метою оптимізації витрат підприємство буде прагнути продовжувати використовувати або оновлювати вже наявні технології, не жертвуючи при цьому безпекою. Проте, продовження використання існуючих технологій вимагатиме від підприємства розуміння того, як існуюча система захисту буде інтегрована та взаємодіятиме з потенційними компонентами та рішеннями нульової довіри. Будь-які нові елементи, що будуть придбані спеціально для впровадження в рамках ZTA, в ідеалі повинні інтегруватися з компонентами засобів безпеки, які організація вже має та планує продовжувати використовувати [15]. Після цього проводиться аналіз розривів (gap analysis), який дозволяє оцінити поточний стан безпеки підприємства та визначити напрямки для вдосконалення [54].

Наступним кроком є проведення оцінки ризиків безпеки, що дозволяє ідентифікувати зовнішні та внутрішні загрози, а також фактичні та потенційні ризики. Цей процес включає визначення ключових активів, які потребують захисту, та розробку стратегії керування ризиками. Результати оцінки ризиків можуть бути використані для проєктування заходів для захисту, зокрема впровадження принципу найменших привілеїв, ZK та обмеження доступу до критичних ресурсів. Крім того, оцінка ризиків допомагає визначити пріоритети при впровадженні ZTA, зосереджуючись на процесах з найнижчим рівнем ризику [18, 23]. Ключовим елементом цієї оцінки є тестування на проникнення, яке дозволяє імітувати дії зловмисника з метою виявлення вразливих місць у механізмах захисту та інформаційних системах підприємства, оцінити їх рівень ефективності та визначити потенційні шляхи обходу політик нульової довіри. Його результати не лише дають змогу оцінити ефективність чинних заходів безпеки, а й виявити критичні слабкі місця, що можуть бути використані для компрометації корпоративної інфраструктури [55]. Тому інтеграція тестування на проникнення у модель розгортання ZTA є ключовим інструментом оцінки ефективності засобів захисту шляхом моделювання реальних сценаріїв кібератак.

Одним із пріоритетних завдань тестування на проникнення є виявлення вразливостей у системах контролю доступу. У ході контрольованих атак перевіряється, чи здатен зловмисник обійти механізми автентифікації або скористатися надмірними або неналежно налашто-

ваними правами доступу. Окрема увага приділяється оцінці ефективності впровадження мікросегментації, а саме проводитися моделювання бічного переміщення в мережі, що дозволяє оцінити здатність архітектури до ефективного стримування та локалізації інцидентів. Тестуванню на проникнення підлягають також системи керування ідентифікацією та контролю доступу, де перевіряються потенційні шляхи підвищення привілеїв, недоліки у механізмах автентифікації та помилки в обробці сеансів доступу. У контексті гібридних і хмарних середовищ, де ZTA реалізується між локальними та хмарними компонентами, важливо оцінити рівень безпеки при передачі даних між середовищами. Крім того, з огляду на постійну перевірку справжності/ідентифікації суб'єкта, а також виконання автентифікації та авторизації суб'єктів в рамках ZTA, тестуванню на проникнення підлягають кінцеві пристрої та застосунки, включно з виявленням вразливостей, пов'язаних із застарілим ПЗ або неправильними налаштуваннями.

З іншого боку, слід мати на увазі, що впровадження тестування на проникнення в ZTA супроводжується низкою викликів. Зокрема, динамічні політики безпеки, які адаптуються залежно від поведінки користувача та контексту, ускладнюють створення постійної методології тестування. Мікросегментація й посилений контроль доступу обмежують можливості традиційного тестування на проникнення, яке покладається на бічне переміщення, а засоби безперервного моніторингу з підтримкою ШІ можуть розцінювати дії при тестуванні як реальну загрозу, активуючи автоматизовані механізми реагування, що перешкоджає проведенню тестування на проникнення. Ще одним викликом є інтеграція тестування на проникнення в процеси DevSecOps протягом циклу розгортання ZTA, а команда тестування на проникнення повинна тісно співпрацювати з командою впровадження ZTA. Тому, з метою усунення зазначених викликів, варто дотримуватися низки запропонованих практик. Передусім, необхідно чітко визначати цілі тестування, які повинні узгоджуватися зі специфікою архітектури, а саме перевіркою механізмів автентифікації, ефективності сегментації мережі, політик контролю доступу, налаштувань безпеки гібридної або хмарної інфраструктури. При цьому варто застосовувати механізми ШІ та засоби автоматизації для ефективного аналізу великих обсягів даних та моделювання сценаріїв атак. В свою чергу, використання підходів «червоної команди» (red team) і «блакитної команди» (blue team) у форматі взаємодії (purple team) дає змогу не лише моделювати реальні кібератаки, але й узгоджено перевіряти ефективність заходів безпеки. Також, особливу увагу слід приділити перевірці механізмів керування доступом, де ключовими аспектами при тестуванні на проникнення є потенційні методи обходу MFA, ризики перехоплення сеансів користувачів, перевірка політики паролів, а також вразливість до атак соціальної інженерії. У цьому контексті важливо враховувати, що взаємодія в межах ZTA дедалі частіше реалізується через API-з'єднання та хмарні провайдери. Відтак, актуальним стає тестування механізмів автентифікації для API, виявлення неправильних налаштувань хмарних сервісів, а також оцінка відповідності вимогам галузевих стандартів безпеки (наприклад, NIST 800-207, CIS Benchmarks). Однак ZTA передбачає наявність загроз не лише ззовні, тому важливим елементом є моделювання дій інсайдера. Зокрема це можуть бути спроби ексфільтрації (несанкціонована передача) даних, підвищення привілеїв чи зловживання легітимними обліковими записами. Тому, з огляду на постійний розвиток концепції нульової довіри та ускладнення цифрового середовища підприємств, стратегії тестування на проникнення повинні регулярно оновлюватися. Це дає змогу своєчасно виявляти нові вектори атак і підтримувати належний рівень кібербезпеки підприємства.

Таким чином, проведення тестування на проникнення має ключову роль при впровадженні ZTA та дозволяє не лише оцінити ефективність наявних засобів захисту та дотримання принципів нульової довіри, але й виявити слабкі місця, що можуть бути використані для компрометації IT-інфраструктури, а також встановити пріоритети для вдосконалення реалізації ZTA. Такий підхід забезпечує обґрунтовану й цілеспрямовану оптимізацію політик безпеки, орієнтовану на запобігання несанкціонованому доступу та зміцнення загальної кіберстійкості підприємства. В свою чергу дані, отримані в результаті аналізу на даному етапі

встановлюють, на якій стадії впровадження ZTA перебуває підприємство: чи воно функціонує за традиційною моделлю безпеки, чи вже інтегрує окремі елементи нульової довіри, чи наближається до повного її впровадження, що передбачає застосування передових технологій, зокрема автоматизованих систем аналізу загроз на основі ШІ [26, 56]. Отримані результати інтегруються у загальний план керування ризиками та слугують основою для формування стратегії поступового переходу до концепції нульової довіри. Оцінка поточного стану підприємства є невіддільним етапом підготовки до впровадження ZTA, оскільки вона дозволяє виявити прогалини в наявних механізмах захисту, визначити рівень ризиків та встановити пріоритетні напрями їх мінімізації. На основі отриманих результатів та оцінки формується детальний план міграції до ZTA, який передбачає модернізацію інфраструктури, перегляд та вдосконалення політик безпеки, а також реалізацію комплексних заходів для досягнення необхідного рівня захищеності корпоративної інформаційної системи.

Етап 6. Створення плану міграції. Процес впровадження ZTA передбачає розробку плану міграції, який підприємство використовуватиме для реалізації ZTA. Перед початком впровадження ZTA необхідно здійснити детальний аналіз доступних рішень і технологій, що відповідають вимогам підприємства та концепції нульової довіри [53]. Вибір постачальників має ґрунтуватися на всебічному вивченні технічних вимог ZTA, що забезпечить відповідність придбаних рішень встановленим критеріям. Одним із ключових аспектів вибору технологій є забезпечення сумісності ПЗ, що дозволить уникнути ефекту «прив'язки до постачальника». Для запобігання цій проблемі необхідно враховувати можливості адаптації послуг, зокрема наступні параметри: якість обслуговування (Quality of Service) та угода про рівень послуг (Service-Level Agreement). Це сприятиме гнучкості у подальшій експлуатації та модернізації ZTA. Розробка плану міграції передбачає структурований розподіл завдань на етапи та визначення їх пріоритетності відповідно до потреб підприємства [57]. Ефективне планування дає змогу раціонально розподілити ресурси, що сприяє зменшенню витрат та оптимізації часових рамок при впровадження ZTA. Крім того, необхідно створити тестові плани та план оцінювання ефективності впровадження ZTA, що забезпечить можливість об'єктивного аналізу успішності міграції. Для досягнення швидких результатів підприємство може використовувати підхід, заснований на конкретних сценаріях використання (use-case), що дозволить вирішити окремі проблеми за допомогою невеликих команд та обмеженого бюджету.

2.3. Підготовка до впровадження архітектури нульової довіри

Підготовка до впровадження ZTA вимагає комплексного та послідовного підходу, що охоплює пристрої, користувачів і мережеву інфраструктуру. Перш за все необхідно усунути недоліки у політиках та процесах забезпечення нульової довіри. Це здійснюється через застосування підходу, заснованого на оцінці ризиків і цінності даних, що дозволяє оптимізувати існуючі політики доступу та забезпечити їх відповідність принципам ZTA. Виявлені недоліки в процесах і політиках потребують коригування для забезпечення максимальної ефективності та безпеки. Після вдосконалення політик і процесів наступним кроком є підготовка пристроїв, користувачів та мережі до інтеграції з ZTA. На цьому кроці передбачається налаштування відповідних компонентів інфраструктури, що дозволить безперешкодно реалізувати контроль доступу згідно з новими політиками і підвищити загальний рівень безпеки організації. Таким чином, процес підготовки до впровадження ZTA можна поділити на такі етапи.

Етап 7. Усунення недоліків у політиках та процесах забезпечення нульової довіри шляхом застосування підходу, заснованого на оцінці ризиків та цінності даних. Після того як буде визначено перелік ресурсів, що потребують захисту, а також оцінено середовище підприємства та наявні засоби безпеки, наступним кроком є планування технології захисту доступу. Вона передбачає ухвалення рішень щодо визначення рівня захисту кожного ресурсу та сегментації інфраструктури. Ефективне впровадження технології захисту доступу ґрунтується на підході, орієнтованому на ризики: критично важливі ресурси повинні бути ізольовані у

власних зонах довіри, де контроль здійснюється через PEP, тоді як менш важливі ресурси можуть розміщуватися у одній зоні довіри. Для успішного захисту IT-підприємства, необхідно оцінити рівень ризику для кожного ресурсу та надати їм рівні. Оцінка включає аналіз ймовірності виникнення загрози та потенційного збитку, який вона може спричинити для підприємства. Після документування ризиків стає зрозуміло, які ресурси є найбільш критично важливими для роботи підприємства та які вразливості вони мають. Подібний підхід відповідає принципам дії у межах системи керування ризиками (Risk Management Framework), оскільки впровадження ZTA є насамперед процесом зниження ризиків для бізнес-функцій підприємства [15].

На етапі розробки технології захисту доступу підприємству слід визначити, які PEP відповідатимуть за захист кожного ресурсу, а також які допоміжні технології будуть залучені до ухвалення рішень про доступ. Початковий стан мережі може не передбачати сегментації, фактично до впровадження ZTA, коли підприємство все ще покладається на захист на основі периметра, таку технологію можна представити як централізований захист усіх ресурсів підприємства за допомогою однієї PEP, наприклад, брандмауер на периметрі. Проте поступове впровадження ZTA передбачає сегментування інфраструктури на дрібніші частини, що дозволяє обмежити вплив потенційних атак або порушень та спростити моніторинг мережевого трафіку. Захист доступу має бути багаторівневим і охоплювати контроль на рівні застосунків, вузлів і мережевої інфраструктури. Водночас ефективність реалізованих політик і заходів безпеки значною мірою залежить від готовності підприємства до їхнього впровадження. Наступним кроком у цьому процесі є підготовка пристроїв, користувачів та мережі до роботи в умовах нульової довіри.

Етап 8. Підготовка пристроїв, користувачів та мережі. Для забезпечення безпечного функціонування пристроїв у межах корпоративної мережі необхідно впровадити систему керування обліком пристроїв (Device Inventory Management), яка дозволяє вести централізований облік усіх пристроїв, підключених до корпоративного середовища [24]. Це дає змогу контролювати перелік активних пристроїв і забезпечувати доступ лише тим, які відповідають встановленим вимогам безпеки [3, 26]. Крім того, усі пристрої повинні відповідати встановленим стандартам безпеки, мати необхідне ПЗ для захисту від кібератак та перебувати під постійним моніторингом для своєчасного виявлення потенційних загроз [58]. З метою запобігання несанкціонованому доступу до інформаційних активів підприємства необхідно впровадити механізми оцінки відповідності пристроїв (compliance check). Це дозволить визначати пристрої з високим рівнем ризику та обмежувати їхній доступ до критично важливих ресурсів. Важливим аспектом є облік як корпоративних, так і особистих пристроїв (BYOD), що забезпечує гнучкість у роботі співробітників, водночас знижуючи ризики за рахунок використання механізмів керування мобільними пристроями (Mobile Device Management) та мобільністю підприємства (Enterprise Mobility Management).

Ідентифікація та автентифікація користувачів є фундаментальним аспектом безпеки в рамках ZTA. Для цього організація повинна створити централізовану базу даних облікових записів, що містить актуальну інформацію про користувачів, їхні ролі та рівень доступу [24]. Дотримання принципу найменших привілеїв забезпечує мінімізацію можливих загроз, пов’язаних із компрометацією облікових записів. Користувачі можуть бути класифіковані за категоріями, а саме: співробітники підприємства, клієнти, партнери та адміністратори, які мають розширені права для налаштування доступу [3]. Для захисту облікових записів слід застосовувати MFA, що значно знижує ризик несанкціонованого доступу. Додаткові заходи безпеки включають використання механізмів керування привілейованими обліковими записами (Privileged Identity Management) та PAM, які унеможливлюють використання підвищених привілеїв у разі компрометації облікового запису [50].

Корпоративна мережа повинна бути структурована таким чином, щоб забезпечити захищений обмін даними та мінімізувати потенційні загрози. Одним із важливих підходів є мікросегментація мережі, що ґрунтується на створенні ізольованих сегментів із контрольованим

обміном даними між ними. Це дозволяє обмежити розповсюдження можливих атак і забезпечити доступ до критично важливих ресурсів лише тим користувачам і пристроям, які мають на це відповідні дозволи [59]. Окрім мікросегментації, важливим аспектом є забезпечення контрольованого доступу шляхом застосування шлюзів безпеки, які фільтрують увесь вхідний і вихідний трафік. Для підвищення рівня безпеки слід впровадити механізми розширеного моніторингу мережевої активності, що включає виявлення аномальної поведінки користувачів та пристроїв, а також аналіз трафіку в реальному часі. Повна видимість активів у мережі та дотримання встановлених стандартів кібербезпеки є критично важливими умовами ефективного впровадження ZTA. В свою чергу, всі запити на доступ до ресурсів повинні проходити через механізми ідентифікації та авторизації, що гарантує відповідність кожної сесії вимогам безпеки. Це сприяє створенню динамічного механізму керування доступом, який адаптується до змін у середовищі загроз та поточних бізнес-процесів підприємства [3].

2.4. Перехід до архітектури нульової довіри

Впровадження ZTA вимагає комплексного підходу, що вимагає ретельної та поетапної підготовки. Цей перехід охоплює всі аспекти цифрового середовища підприємства, включаючи користувачів, пристрої та мережеву інфраструктуру. Основним завданням на цьому етапі є забезпечення безперервності бізнес-процесів підприємства та мінімізація впливу змін на роботу користувачів. Під час переходу до ZTA необхідно поступово адаптувати застосунки та сервіси, забезпечуючи їхню сумісність із сучасними механізмами контролю доступу та протоколами безпеки. Важливим є зменшення залежності від традиційних рішень на основі периметру, таких як VPN, шляхом надання доступу лише для тих користувачів, хто потребує цього для виконання своїх завдань. Усі застосунки та сервіси мають працювати через визначені шлюзи, а їхнє використання контролюватися відповідно до встановлених політик доступу. Також необхідно реалізувати комплексний моніторинг мережевого трафіку, аналіз активності користувачів і пристроїв, щоб забезпечити відповідність діям заданим правилам безпеки. Подальші кроки передбачають оцінку ефективності реалізованих заходів та адаптацію політик безпеки на основі отриманих результатів. Виявлені проблеми потребують коригування та повторного тестування, щоб гарантувати стабільність та функціональність системи. Після підтвердження стабільної роботи впроваджених рішень, впровадження ZTA розширюється на всю інфраструктуру підприємства, забезпечуючи поступову інтеграцію сучасних механізмів контролю доступу та виведення з експлуатації застарілих рішень. На цьому етапі починається впровадження ключових компонентів ZTA. Далі здійснюється перевірка реалізації, оцінка підсумкових досягнень та виправлення можливих виявлених недоліків для остаточного завершення процесу переходу до ZTA.

Етап 9. Впровадження компонентів ZTA, поступове використання існуючих безпекових рішень для досягнення кінцевої мети. Щойно підприємство досягне чіткого розуміння існуючого середовища з точки зору ресурсів, які потребують захисту, та вже розгорнутих засобів безпеки, сформулює політики доступу, які підходять для підтримки його завдань та бізнес-показників, розробить технологію захисту доступу із зазначенням рівня деталізації, з яким захищатиметься доступ до різних ресурсів і допоміжних технологій, які будуть використовуватись у PDP, підприємство може безпосередньо розпочати поступове впровадження ZTA [15].

Зважаючи на те, що нині існують різні моделі розгортання нульової довіри [15], важливо визначити, яка саме архітектура найбільш підходить для конкретного IT-підприємства. Наприклад, під час вибору міжмережевих екранів із PDP слід враховувати функціональні ролі брандмауерів як елементів інфраструктури – шлюзів (PEP), які відіграють ключову роль у реалізації детальних мережевих політиках та політиках рівня ідентифікації. Відповідно до рекомендацій NIST [53], можна виділити кілька типів шлюзів, що забезпечують функціональність у межах ZTA: вхідний шлюз (Ingress gateway) регулює вихідні запити застосунків за межі кластеру/групи, контролюючи імена, сертифікати, порти, протоколи та кінцеві точки, які обслуговуються за межами кластера; вихідний шлюз/шлюз виходу (Egress gateway) керує

взаємодією застосунків у межах кластеру/групи із зовнішнім середовищем, може використовуватися для традиційної фільтрації та реєстрації вихідних повідомлень, як проксі Squid, але також може реалізовувати політики на основі ідентифікації, дозволяючи або обмежуючи обмін обліковими даними; граничний/крайовий шлюз (Edge gateway) працює на межі між мережею та вхідним шлюзом, оптимізуючи трафік, забезпечуючи балансування навантаження між групами або вузлами, також використовується для переривання зовнішнього трафіку та підвищення стійкості до відмов на рівні інфраструктури; окрім цього, є Sidecar gateway, який функціонує поруч із кожним екземпляром застосунка, перехоплюючи весь вхідний і вихідний трафік та обробляючи внутрішні комунікації між сервісами в інфраструктурі.

Ключовими аспектами, які підприємство має враховувати при розгортанні ZTA, є ідентифікація, автентифікація та авторизація суб'єктів, а також перевірка відповідності кінцевих пристроїв встановленим політикам. Враховуючи, що прийняття та виконання рішень щодо доступу є основними складовими ZTA, підприємству доцільно використовувати існуючі або нові рішення керування ідентифікацією, обліковими даними та доступом (Identity, Credential, Access Management – ICAM) як базовий компонент реалізації ZTA. Важливим кроком є впровадження MFA для всіх користувачів та інтеграція EDR (або аналогічне рішення, що може оцінювати стан пристрою) із системою ICAM. Початкова ZTA, що базується на цих компонентах, зможе забезпечити контрольовану ідентифікацію та авторизацію суб'єктів, а також стан і відповідність вимогам кінцевих пристроїв, що запитують, як основу для прийняття рішень про доступ, що згодом можна розширити додатковими компонентами та функціями безпеки для задоволення більшої кількості вимог ZTA. Подальший розвиток ZTA залежить вже, безпосередньо, від пріоритетів підприємства. Якщо основним завданням є захист даних, першочергово слід впроваджувати компоненти безпеки даних. Якщо ж акцент робиться на виявленні аномалій з урахуванням поведінки, варто встановити рішення для моніторингу і систему аналізу на основі ШІ. ZTA має розвиватися поступово, включаючи додаткові допоміжні компоненти, функції та можливості, щоб крок за кроком забезпечити її повну функціональність. При цьому варто зазначити, що розгортання ZTA може супроводжуватися складнощами (протидія змінам), зокрема пов'язаними з людським фактором. Опір змінам може виникати через технічні упередження, культуру, емоційну орієнтацію на існуючі інструменти, архітектуру безпеки або процеси компанії [30]. Ця проблема може здаватися незначною у порівнянні з фінансовими та апаратними складовими, однак це може викликати серйозні проблеми із впровадженням. Щоб подолати ці труднощі, необхідно інвестувати у навчання персоналу, підвищуючи їх кваліфікацію та рівень обізнаності щодо принципів нульової довіри та методів їх реалізації. Важливим є розуміння того, що ефективне впровадження ZTA безпосередньо залежить від рівня підготовки користувачів та їхнього усвідомлення переваг нової архітектури, знання основних принципів нульової довіри, які застосовуються до ситуації, а також методів, необхідних для забезпечення дотримання цих принципів.

Завершальним етапом впровадження є перевірка реалізованої архітектури, що дозволить підтвердити досягнення запланованих цілей та здійснити корекцію виявлених недоліків після впровадження ZTA.

Етап 10. Перевірка реалізації для підтвердження підсумкових досягнень під час розгортання ZTA та виправлення виявлених помилок після впровадження. Для оцінки успішності впровадження ZTA та ефективності його роботи мають застосовуватися плани тестування та відповідні метрики. Вибрані для тестування сценарії використання повинні відповідати тим, які найбільш точно відображають повсякденний доступ користувачів підприємства до його ресурсів. В ідеалі підприємство може створити набір тестів, які можна використовувати для перевірки можливостей у рамках ZTA не лише перед розгортанням кожної нової функціональності в процесі поступового впровадження ZTA, а й на основі регулярного тестування після завершення розгортання ZTA. Наприклад, можуть проводитися тестування на проникнення, перевірка стійкості до викрадення даних, протистояння фішинговим атакам, фальсифікації даних тощо [15]. Водночас підприємство повинно передбачити механізми або канали для

усунення можливих помилок, що можуть виникнути під час експлуатації системи. Наприклад, можуть бути впроваджені засоби самостійного вирішення типових проблем користувачами шляхом використання покрокових інструкцій або керівництв, а також організовані інформаційні сервіси, які надаватимуть відповіді на поширені питання (Frequently Asked Questions – FAQ). Для своєчасного інформування про оновлення, пов'язані з впровадженням ZTA, та можливі тимчасові зміни в доступі до ресурсів, може бути налаштована система електронних сповіщень для користувачам. У разі виникнення складніших (не типових) проблем команда впровадження ZTA повинна бути готова оперативно реагувати, аналізувати причини збоїв та забезпечувати швидке усунення несправностей, мінімізуючи їх вплив на користувачів. Для цього члени команди мають пройти відповідне навчання, щоб ефективно допомагати користувачам у разі перебоїв і сприяти швидкому відновленню роботи. Після того як команда впровадження підтвердить стабільне функціонування всіх компонентів ZTA та усуне виявлені недоліки, процес впровадження переходить до наступного етапу, а саме безперервного моніторингу та обслуговування ZTA, що дозволяє забезпечити її довгострокову ефективність та адаптивність до нових викликів.

2.5. Моніторинг, обслуговування та оптимізація архітектури нульової довіри

Ефективне функціонування підприємства в рамках ZTA передбачає безперервний моніторинг і технічне обслуговування, що забезпечує контроль за станом системи, оцінку її продуктивності та відповідності вимогам безпеки. Моніторинг ZTA спрямований на досягнення повної прозорості мережевої активності, своєчасного виявлення потенційних загроз, забезпечення стабільності роботи всіх компонентів та визначення необхідності їхнього технічного обслуговування. У цьому контексті особливого значення набуває впровадження автоматизації процесів, а також постійне вдосконалення та розвиток архітектури відповідно до змін характеру загроз, завдань, технологій та нормативних документів. Здатність ZTA до динамічного реагування на зміни у середовищі функціонування, а також на зовнішні та внутрішні загрози є ключовим чинником не тільки підтримки її ефективності, але й її здатності до адаптації та збереження належного рівня безпеки в умовах динамічного кіберсередовища.

Етап 11. Забезпечення безперервного моніторингу та технічного обслуговування екосистеми нульової довіри відповідно до виявлених помилок, актуальних загроз та вимог безпеки. Після впровадження ZTA необхідно здійснювати безперервний моніторинг та технічне обслуговування екосистеми нульової довіри відповідно до виявлених вразливостей, актуальних загроз і встановлених вимог безпеки. Це передбачає контроль доступності мережевих сервісів, моніторинг мережевого трафіку в режимі реального часу, а також регулярну перевірку відповідності системи чинним стандартам безпеки [29, 60]. Окрім цього, необхідним є систематичне відстеження стану всієї IT-інфраструктури підприємства, зокрема її критичних компонентів, що забезпечують функціонування ZTA. В свою чергу, моніторинг відповідності нормативним вимогам включає проведення регулярного сканування вразливостей, виявлення потенційних витоків даних і аналіз стану реалізованих політик безпеки. Використання сучасних аналітичних інструментів (із використанням ШІ) сприяє поглибленому аналізу подій у межах ZTA, що дає змогу своєчасно виявляти аномалії та підвищувати рівень кіберзахисту. Для ефективної інтеграції таких засобів необхідно оцінювати їхню відповідність логічній архітектурі системи та забезпечувати їхнє оптимальне розміщення в межах інфраструктури підприємства [27, 61]. У разі оновлення засобів аналітики перевагу слід надавати рішенням, які сумісні з принципами нульової довіри та можуть бути інтегровані у вже функціонуючу екосистему без порушення її стабільності та функціональності.

Оцінка ефективності безпеки після впровадження ZTA є також важливим аспектом в рамках моніторингу та забезпечення безперервного функціонування ZTA. Для цього проводиться аналіз впливу впроваджених механізмів на робочі процеси та визначення їх ефективності у запобіганні кіберзагрозам [62]. Одним із ключових показників ефективності є частота виникнення інцидентів безпеки та їхній вплив на функціонування системи та роботу корис-

тувачів. Після впровадження ZTA кількість таких інцидентів має зменшитися, а також має знизитися їхній вплив на бізнес-процеси. Наприклад, можна відстежувати зменшення часу простою критичних сервісів через кіберінциденти або скорочення кількості випадків компрометації облікових записів. Крім того, важливим аспектом є оцінка навантаження на IT-відділ, що включає аналіз частоти звернень користувачів щодо проблем із доступом, часу, що витрачається на ручні перевірки та обробку сповіщень безпеки, а також кількості запитів, пов'язаних із відновленням доступу до систем. Ефективна робота ZTA сприяє автоматизації рутинних операцій безпеки та мінімізації потреби у ручному втручанні. Оцінка рівня впровадження MFA також є важливим показником ефективності: слід аналізувати відсоток працівників, які регулярно використовують MFA, а також випадки труднощів з автентифікацією через невідповідність політикам безпеки. Додаткові метрики включають кількість запитів на автентифікацію, частку співробітників, що використовують єдиний вхід (Single sign-on – SSO), динаміку використання корпоративних застосунків, а також кількість пристроїв, застосунків і сервісів, що відповідають політикам безпеки. Оптимізація ресурсів безпеки передбачає зменшення кількості інструментів, що виконують однакові функції, скорочення потреби у їх складній ручній інтеграції, а також зниження часу, витраченого на обробку хибнопозитивних (false positive) сповіщень. Використання аналітичних інструментів на основі ШІ в таких випадках дозволяє зменшити рівень помилкових сповіщень, а також своєчасно виявляти аномальні шаблони поведінки користувачів, пристроїв, мережевого трафіку та реагувати на реальні загрози. Це, у свою чергу, сприяє стабільному та ефективному функціонуванню ZTA, забезпечуючи високий рівень кіберзахисту без надмірного впливу на бізнес-процеси.

Етап 12. Впровадження автоматизації процесів, постійне вдосконалення та розвиток відповідно до змін характеру загроз, завдань, технологій та нормативних документів. Після того як ZTA буде розгорнута та буде вважатися завершеною, екосистема нульової довіри повинна продовжувати адаптуватися до умов, що змінюються. Проведена оцінка ефективності ZTA використовується для визначення подальших заходів щодо її модернізації. Якщо технологічні компоненти, що використовуються в ZTA, застарівають, їх слід замінити. І навпаки, якщо з'являються нові інноваційні технології, підприємство повинне розглянути можливість їх інтеграції в існуючу ZTA, щоб скористатися перевагами нових засобів захисту, методів та механізмів, які можуть підвищити загальний рівень безпеки підприємства [15]. Також слід враховувати, що здатність ZTA до ефективного функціонування значною мірою залежить від рівня автоматизації механізмів керування загрозами, що, у свою чергу, передбачає використання технічних рішень, спрямованих на зменшення використання ручних операцій і впровадження механізмів автоматизованого реагування. Автоматизація та оркестрування процесів безпеки також забезпечує підвищення точності виявлення атак і скорочення часу реагування на інциденти, що зумовлює необхідність інтеграції відповідних рішень у систему керування безпекою, а також відіграє ключову роль у забезпеченні безперервного вдосконалення ZTA. Завдяки автоматизованому моніторингу інфраструктури, аналізу загроз у реальному часі та адаптивному оновленню політик безпеки підприємства можуть оперативно реагувати на зміни у сфері кіберзахисту. У цьому контексті інтеграція сучасних систем безпеки для оркестрації, автоматизації та реагування (Security Orchestration, Automation, and Response – SOAR) у поєднанні з системою керування інформацією та подіями безпеки (Security Information and Event Management – SIEM) є ефективним рішенням для підвищення рівня безпеки. SOAR забезпечує централізоване керування інцидентами, автоматизацію при аналізі загроз і координацію дій між різними системами безпеки, що дозволяє мінімізувати затримки в реагуванні та зменшити навантаження на команди безпеки. SOAR-системи у поєднанні зі ШІ та машинним навчанням, демонструють значний потенціал у виявленні, пом'якшенні та запобіганні кіберзагрозам. Ось чому постачальники даних інструментів намагаються інтегрувати алгоритми ШІ та машинного навчання для підвищення ефективності команд безпеки [63]. Іншими словами, використання сучасних підходів до автоматизації та оркестрування проце-

сів безпеки дозволяє підприємствам не лише підтримувати стабільність роботи ZTA, а й досягати високого рівня адаптивності до змін у сфері кіберзахисту [60, 61].

Надалі, якщо цілі підприємства в сфері безпеки змінюються, або внаслідок зміни завдань, або внаслідок змін у нормативних документах, може знадобитися зміна політик та самої ZTA, щоб найкраще відповідати новим цілям. Тобто в рамках цього безперервного процесу валідації та вдосконалення підприємства повинні забезпечити контроль актуальності політик безпеки, технології та топології сегментації мережі, постійний автоматизований моніторинг стану мережі та іншої IT-інфраструктури, щоб переконатися, що вони залишаються ефективними. Саме автоматизація процесів безпеки дозволяє забезпечити безперервну перевірку та вдосконалення архітектури ZTA, знижуючи вплив людського фактору та забезпечуючи ефективність роботи екосистеми нульової довіри.

Наприкінці, хочеться зазначити, що розгортання та впровадження ZTA є складним, тривалим і багатоетапним процесом, який потребує ретельного аналізу та кропіткої роботи для досягнення рішень високої якості у забезпеченні безпеки IT-підприємства. Незважаючи на те, що основні принципи ZTA сьогодні вже вважаються визначеними, питання їхнього практичного впровадження залишається відкритим. Насамперед складнощі виникають відносно того, як узгодити та інтегрувати різні технологічні рішення з вимогами ZTA [64]. Тому важливим напрямом подальших досліджень є розробка технічних методів ефективного та узгодженого розгортання компонентів безпеки у межах вимог ZTA. У майбутньому необхідно більш детально дослідити підходи до впровадження таких компонентів, як мікросегментація, механізми автентифікації та системи керування доступом. Це дозволить розробляти більш ефективні способи інтеграції технологій та підвищити загальний рівень захищеності інфраструктури підприємства. Крім того, слід звернути увагу ще на один актуальний виклик – адаптацію ZTA до хмарних середовищ, оскільки сучасні підприємства все частіше використовують комбіновані архітектури, що поєднують локальні та хмарні сервіси. Такий підхід створює додаткові ризики, пов’язані з крадіжкою ідентифікаційних даних та витоками інформації. Тому також перспективним напрямом досліджень слід вважати – удосконалення методів інтеграції ZTA з хмарною інфраструктурою, зокрема, впровадження механізмів контролю доступу до ресурсів, орієнтованих на ідентифікаційні (identity-centric) та динамічні дані у хмарному середовищі.

Висновки

1. Відповідно до сучасних викликів, парадигма безпеки «нульової довіри» є оптимальним та ефективним підходом для захисту інформаційних систем цифрових підприємств від новітніх загроз. Дана концепція забезпечує безпечний доступ до корпоративних ресурсів з будь-якого місця та у будь-який час.

2. Існуючі рішення побудови систем на основі ZTA та досвід їх використання свідчать про істотні переваги нової концепції перед традиційними архітектурами, орієнтованими на захист по периметру. Проте, впровадження ZTA є не просто повною заміною технологічних процесів або інфраструктури IT-підприємства, а реалізацією плану кібербезпеки, що використовує концепцію нульової довіри та охоплює планування робочих процесів, політики доступу та зв’язки логічних компонентів.

3. Будь-яке підприємство може застосовувати принципи нульової довіри для підвищення рівня безпеки, незалежно від структури чи масштабу. В роботі розглянуті деякі рекомендації щодо практичного впровадження ZTA, з урахуванням можливих сценаріїв, викликів і оптимальних підходів до інтеграції принципів нульової довіри в корпоративне середовище IT-підприємства.

4. Враховуючи організаційні та технічні виклики, а також вимоги щодо розгортання ZTA, в даній роботі запропонована модель впровадження ZTA. Для її ефективної реалізації запропоновано п’ять основних процесів: 1) стратегія нульової довіри, 2) оцінка середовища підприємства, 3) підготовка до впровадження ZTA, 4) перехід до ZTA, 5) моніторинг, обслу-

говування та оптимізація ZTA. Важливими складовими пропонованої моделі поетапного впровадження ZTA на IT-підприємстві є концепція підходу DevSecOps та компоненти розширеної моделі нульової довіри ZTX. Окрім цього при впровадженні ZTA, запропонована модель враховує дотримання вимог галузевих практик і стандартів.

5. Представлені в роботі результати покликані допомогти спеціалістам з безпеки використати надані рекомендації щодо практичного впровадження ZTA на своїх IT-підприємствах відповідно до запропонованої моделі. Встановлено, що подальші дослідження доцільно зосередити на розробці технічних методів розгортання компонентів безпеки у межах вимог ZTA, інтеграції ZTA з хмарною інфраструктурою та більш детальному аналізі підходів впровадження таких компонентів, як мікросегментація, механізми автентифікації та системи керування доступом.

6. Запропонована модель впровадження ZTA може слугувати орієнтиром для ефективного переходу до ZTA в сучасних цифрових підприємства, а її використання відкриває значні перспективи для розвитку більш надійних та масштабованих систем захисту інформаційних ресурсів IT-підприємства, що є надзвичайно актуальним у сучасному цифровому кіберпросторі.

Список літератури:

1. Marsh S.P. Formalising Trust as a Computational Concept. Department of Computing Science and Mathematics University of Stirling. 1994. 184 p.
2. Welborn R., Kasten V. The Jericho principle: how companies use strategic collaboration to find new sources of value. John Wiley & Sons. 2003. 288 p.
3. Ward R., Beyer B. Beyondcorp // A new approach to enterprise security. 2014. 39(6). P. 6–11.
4. Gilman E., Barth D. Zero Trust Networks: Building Secure Systems in Untrusted Networks. 2017. 315 p.
5. Cunningham C., Balaouras S., Barringham B., Dostie P. The Zero Trust eXtended (ZTX) Ecosystem. Extending Zero Trust Security Across Your Digital Business. Forrester Research, Inc. Cambridge, MA. 2018. URL: https://www.cisco.com/c/dam/m/en_sg/solutions/security/pdfs/forrester-ztx.pdf.
6. Fisher B. Forrester's Zero Trust or Gartner's Lean Trust? 2019. URL: https://blogs.cisco.com/security/forresters-zero-trust-or-gartners-lean-trust.
7. Rose S., Borchert O., Mitchell S., Connelly S. Zero Trust Architecture // NIST Special Publication 800-207. 2020. https://doi.org/10.6028/NIST.SP.800–207.
8. National Cybersecurity Center of Excellence (NCCoE). Implementing a Zero Trust Architecture. URL: https://www.nccoe.nist.gov/projects/implementing-zero-trust-architecture.
9. Єсін В. І., Вілігура В. В., Узлов Д. Ю. Огляд існуючих моделей та основних принципів нульової довір // Радіотехніка. 2024. Вип. 217. С. 39–54. https://doi.org/10.30837/rt.2024.2.217.03.
10. Greenberg A. The Untold Story of NotPetya, the Most Devastating Cyberattack in History // Wired. (2018). URL: https://www.wired.com/story/notpetya-cyberattack-ukraine-russia-code-crashed-the-world.
11. Yampolskiy R., Spellchecker M. Artificial Intelligence Safety and Cybersecurity: a Timeline of AI Failures // arXiv preprint arXiv:1610.07997. 2016. 12 p. https://doi.org/10.48550/arXiv.1610.07997.
12. Chen B., Qiao S., Zhao J., Liu D., Shi X., Lyu M., Chen H., Lu H., Zhai Y. A Security Awareness and Protection System for 5G Smart Healthcare Based on Zero-Trust Architecture // IEEE Internet of Things Journal. 8(13). 2020. P. 10248–10263. https://doi.org/10.1109/JIOT.2020.3041042.
13. MarketsandMarkets. Zero Trust Security Market. (2024). URL: https://marketsandmarkets.com/PressReleases/zero-trust-security.
14. Bertino E. Zero Trust Architecture: Does It Help? // IEEE Security & Privacy. 2021. 19(05). P. 95–96. https://doi.org/10.1109/MSEC.2021.3091195.
15. Єсін В. І., Вілігура В. В., Узлов Д. Ю. Архітектура нульової довіри: проблеми та рекомендації щодо успішного впровадження // Радіотехніка. 2024. Вип. 218. С. 7–34. https://doi.org/10.30837/rt.2024.3.218.01.
16. Cunningham C., Holmes D., Pollard J. The eight business and security benefits of zero trust // Forrester Research, Inc. 2019. URL: https://www.forrester.com/report/the-eight-business-and-security-benefits-of-zerotrust/RES134863.
17. Fernandez E. B., Brazhuk A. A critical analysis of Zero Trust Architecture (ZTA) // Computer Standards & Interfaces. 2024. Vol. 89. 103832. https://doi.org/10.1016/j.csi.2024.103832.
18. Teerakanok S., Uehara T., Inomata A. Migrating to Zero Trust Architecture: Reviews and Challenges // Security and Communication Networks. 2021. Vol. 6. P. 1–10. https://dx.doi.org/10.1155/2021/9947347.
19. Phiayura P., Teerakanok S. A Comprehensive Framework for Migrating to Zero Trust Architecture // IEEE Access. 2023. Vol. 11. P. 19487–19511. https://10.1109/ACCESS.2023.3248622.

20. Ali B., Hijjawi S., Campbell L. H., Gregory M. A., Li S. A Maturity Framework for Zero-Trust Security in Multiaccess Edge Computing // Security and Communication Networks. 2022. P. 1–14. https://doi.org/10.1155/2022/3178760.

21. Haber M. J. Privileged Attack Vectors: Building Effective Cyber-Defense Strategies to Protect Organizations, 2nd ed. New York, NY, USA: Apress. 2020. 384 p.

22. Foltz K. E., Simpson W. R. Zero Trust Technology Integration Issues. Institute for Defense Analyses. 2021. P. 34. URL: https://www.jstor.org/stable/resrep34846.

23. Bertino E., Brancik K. Services for Zero Trust Architectures – A Research Roadmap // IEEE International Conference on Web Services (ICWS), Chicago, IL, USA, 2021. P. 14–20. https://doi.org/10.1109/ICWS53863.2021.00016.

24. CISCO. From MFA to Zero Trust: A Five-Phase Journey to Securing the Federal Workforce. 2021. URL: https://www.cisco.com/c/dam/global/en_uk/products/collateral/security/zero-trust/mfa-zero-trust-five-phase-journey-securing-workforce.pdf.

25. Osborn B., McWilliams J., Beyer B., Saltonstall M. BeyondCorp: Design to deployment at Google. 2016. 41(1). P. 28–34.

26. Hirning D. Implementing a Zero Trust Security Model at Microsoft. 2025. URL: https://www.microsoft.com/insidetrack/blog/implementing-a-zero-trust-security-model-at-microsoft/.

27. AWS Prescriptive Guidance: Embracing Zero Trust: A strategy for secure and agile business transformation. 2025. URL: https://docs.aws.amazon.com/pdfs/prescriptive-guidance/latest/strategy-zero-trust-architecture/strategy-zero-trust-architecture.pdf.

28. ON2IT. A Hands-on Approach to Zero Trust Implementation. 2020. URL: https://www.cymbel.com/wp-content/uploads/2020/10/A-hands-on-approach-to-Zero-Trust-implementation.pdf.

29. Best Practices Implementing Zero Trust with Palo Alto Networks. Palo Alto Networks, Inc. 2024. URL: https://docs.paloaltonetworks.com/content/dam/techdocs/en_US/pdf/best-practices/zero-trust-best-practices/zero-trust-best-practices.pdf.

30. Garbis J., Chapman J. W. Zero Trust Security: An Enterprise Guide. Berkeley, CA : Apress, 2021. 300 p.

31. Ross R., Winstead M., McEvilley M. Engineering Trustworthy Secure Systems // NIST Special Publication. NIST SP 800-160v1r1. 2022. 195 p. https://doi.org/10.6028/NIST.SP.800-160v1r1.

32. Goldwasser S., Micali S., Rackoff C. The knowledge complexity of interactive proof-systems // Proceedings of the seventeenth annual ACM symposium on Theory of computing (STOC '85). Association for Computing Machinery, New York, NY, USA, 1985. P. 291–304. https://doi.org/10.1145/22145.22178.

33. Fiege U., Fiat A., Shamir A. Zero knowledge proofs of identity // Proceedings of the nineteenth annual ACM symposium on Theory of computing. 1987. P. 210–217. https://doi.org/10.1145/28395.28419.

34. Blum M., Feldman P., Micali S. Non-interactive zero-knowledge and its applications // Proceedings of the twentieth annual ACM symposium on Theory of computing (STOC '88). Association for Computing Machinery, New York, NY, USA, 1988. P. 103–112. https://doi.org/10.1145/62212.62222.

35. Goldreich O. Foundations of Cryptography. Basic Tools (Vol. 1). Cambridge University Press, 2001. 372 p.

36. Ben-Sasson E., Chiesa A., Tromer E., Virza, M. Succinct Non-Interactive Zero Knowledge for a von Neumann Architecture // 23rd USENIX Security Symposium (USENIX Security 14). August 20-22, 2014. San Diego, CA. 2014. P. 781–796. URL: https://eprint.iacr.org/2013/879.pdf.

37. Cloud Security Alliance (CSA). Software-Defined Perimeter (SDP) Specification v2.0. (2022). URL: https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2.

38. Chailloux A., Ciocan D. F., Kerenidis I., Vadhan S. (2008). Interactive and Noninteractive Zero Knowledge are Equivalent in the Help Model // Canetti R. (eds) Theory of Cryptography. TCC 2008. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 2008. Vol. 4948. P. 501–534. https://doi.org/10.1007/978-3-540-78524-8_28.

39. Junkai L., Daqi H., Pengfei W., Yunbo Y., Qingni S., Zhonghai W. SoK: Understanding zk-SNARKs: The Gap Between Research and Practice. arXiv. arXiv:2502.02387. 2025. 24 p. https://doi.org/10.48550/arXiv.2502.02387.

40. Ben-Sasson E., Bentov I., Horesh Y., Riabzev M. Scalable, transparent, and post-quantum secure computational integrity // Cryptology ePrint Archive. 2018. 046.

41. Liskov M. Updatable Zero-Knowledge Databases // Roy B. (eds) Advances in Cryptology – ASIACRYPT 2005. ASIACRYPT 2005. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 2005. Vol. 3788. P. 174–198. https://doi.org/10.1007/11593447_10.

42. American Council for Technology and Industry Advisory Council (ACT-IAC). Zero Trust Cybersecurity Current Trends. 2019. 29 p. URL: https://www.actiac.org/documents/zero-trust-cybersecurity-current-trends.

43. Department of Defense (DOD). Zero Trust Reference Architecture. Version 1.0. 2021. URL: https://www.texasre.org/Documents/Resource%20Hub/Cybersecurity/Zero%20Trust%20Reference%20Architecture.pdf.

44. Department of Defense (DoD). Zero Trust Strategy. The US Department of Defense. Version 2.0. 2022. URL: https://dodcio.defense.gov/Portals/0/Documents/Library/DoD-ZTStrategy.pdf.

45. Department of Defense (DoD). Zero Trust Reference Architecture. Version 2.0. Defense Information Systems Agency (DISA) and National Security Agency (NSA) Zero Trust Engineering Team. 2022. 104 p. URL: https://dodcio.defense.gov/Portals/0/Documents/Library/(U)ZT_RA_v2.0(U)_Sep22.pdf.

46. Badshah S., Khan A. A., Khan B. Towards process improvement in DevOps: A systematic literature review // Proceedings of the 24th International Conference on Evaluation and Assessment in Software Engineering. (EASE '20). Association for Computing Machinery, New York, NY, USA. 2020. P. 427–433. http://dx.doi.org/10.1145/3383219.3383280.

47. Davis J., Daniels R. Effective DevOps: Building a Culture of Collaboration, Affinity, and Tooling at Scale 1st Ed. O'Reilly Media, Inc. 2016. 408 p.

48. Сусукайло В. А. Використання підходу DevSecOps для аналізу сучасних загроз інформаційної безпеки // Кібербезпека: освіта, наука, техніка. 2021. 2(14). С. 26–35. https://doi.org/10.28925/2663-4023.2021.14.2635.

49. Prates L., Pereira R. DevSecOps practices and tools // International Journal of Information Security. 2024. Vol. 24. 11. https://doi.org/10.1007/s10207-024-00914-z.

50. Turner S., Holmes D., Cunningham C., Budge J., McKay P., Cser A., Shey H., Maxim M. A. A Practical Guide To A Zero Trust Implementation. Forrester Research, Inc. 2021. 14 p. URL: https://www.forrester.com/report/a-practical-guide-to-a-zero-trust-implementation/RES157736.

51. Uttecht K. D. Zero trust (ZT) concepts for federal government architectures // MIT Lincoln Laboratory. 2020. 58 p. URL: https://apps.dtic.mil/sti/pdfs/AD1108910.pdf.

52. Peck J., Beyer B., Beske C., Saltonstall M. Migrating to BeyondCorp // Maintaining productivity while improving security. 2017. 42(2). P. 1–7.
URL: https://www.usenix.org/system/files/login/articles/login_summer17_10_peck.pdf.

53. Chandramouli R., Butcher Z. NIST Special Publication 800-207A. A Zero Trust Architecture Model for Access Control in Cloud-Native Applications in Multi-Location Environments // National Institute of Standards and Technology. 2023. 31 p. https://doi.org/10.6028/NIST.SP.800-207A.

54. Collier Z. A., Sarkis J. The zero trust supply chain: Managing supply chain risk in the absence of trust // International Journal of Production Research. 2021. 59(11). P. 3430–3445.
https://doi.org/10.1080/00207543.2021.1884311.

55. Sin L.W., Samsudin A.E., Zengeni I.P., Zolkipli M.F. Zero Trust Security Models in Penetration Testing // International Journal of Advances in Engineering and Management (IJAEM). 2024. 6(7). P. 442–450. URL: https://ijaem.net/issue_dcp/Zero%20Trust%20Security%20Models%20in%20Penetration%20Testing.pdf.

56. Akamai Security. A Blueprint for Building a Zero Trust Architecture. 2024. URL: https://www.akamai.com/site/en/documents/white-paper/a-blueprint-for-building-a-zero-trust-architecture-white-paper.pdf.

57. Adahman Z., Malik A.W., Anwar Z. An analysis of zero-trust architecture and its cost-effectiveness for organizational security // Computers & Security. 2022. 122(1). 102911. https://doi.org/10.1016/j.cose.2022.102911.

58. Mavroudis V. Zero-Trust Network Access (ZTNA). arXiv. 2024. 10 p. https://doi.org/10.48550/arXiv.2410.20611.

59. Al-Ofeishat H., Alshorman R. Build a Secure Network Using Segmentation and Micro-segmentation Techniques // International Journal of Computing and Digital Systems. 2024. 16(1). P. 1499–1508. http://dx.doi.org/10.12785/ijcds/1601111.

60. Zero Trust Architecture: A Paradigm Shift in Cybersecurity and Privacy. 2021. URL: https://www.pwc.com/sg/en/publications/assets/page/zero-trust-architecture.pdf.

61. Ahmed I., Nahar T., Urmi S. S., Taher K. A. Protection of Sensitive Data in Zero Trust Model. In Proceedings of the International Conference on Computing Advancements (ICCA '20) // Association for Computing Machinery, New York, NY, USA. 2020. 63(1). P. 1–5. https://doi.org/10.1145/3377049.3377114.

62. Microsoft. Zero Trust deployment plan with Microsoft 365. 2025. URL: https://learn.microsoft.com/en-us/microsoft-365/security/microsoft-365-zero-trust.

63. Kinyua J., Awuah L. AI/ML in security orchestration, automation and response: Future research directions // Intelligent Automation & Soft Computing. 2021. 28(2). P. 527–545. https://doi.org/10.32604/iasc.2021.016240.

64. He Y., Huang D., Chen L., Ni Y., Ma X. A survey on zero trust architecture: Challenges and future trends. Wireless Communications and Mobile Computing. 2022. 13 p. https://doi.org/10.1155/2022/6476274.

Надійшла до редколегії 07.06.2025

Відомості про авторів:

Єсін Віталій Іванович – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Україна; e-mail: v.i.yesin@karazin.ua; ORCID: https://orcid.org/0000-0003-1977-7269

Бородавка Владислав Вячеславович – Харківський національний університет імені В. Н. Каразіна, аспірант кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Україна; e-mail: vladyslav.borodavka@karazin.ua; ORCID: https://orcid.org/0009-0002-3885-1364

УДК 621.396:002.53 **DOI:10.30837/rt.2025.3.222.03**

В.М. БЕЗРУК, д-р техн. наук, Ю.М. ГОЛОБОРОДЬКО, канд. техн. наук,
В.І. ЗАБОЛОТНИЙ, канд. техн. наук, М.С. СКИБЕНКО

РАДІОКОНТРОЛЬ ВИПРОМІНЮВАНЬ РАДІОЕЛЕКТРОННИХ ЗАСОБІВ. ПРОБЛЕМИ ТА ШЛЯХИ ВИРІШЕННЯ

Вступ

Радіоконтроль – це система технічних засобів та заходів, спрямованих на виявлення, перехоплення, аналіз та оцінку радіовипромінювань з метою контролю за використанням радіочастотного спектру; виявлення загроз, пов'язаних з використанням радіоелектронних засобів. Радіоконтроль (РК) може включати дії щодо порушення роботи радіоелектронних засобів шляхом радіоелектронної протидії.

Надійність радіозв'язку, особливо в декаметровому (короткохвильовому) діапазоні (Short Waves), критично залежить від ефективності використання частотного спектра. Особливості поширення декаметрових хвиль, які надходять до приймача як поверхневою (ближня зона), так і просторовою (дальня зона, відбитою від іоносфери) хвилями, створюють значні виклики [1, 2]. Це призводить до високого завантаження діапазону та ускладнює виявлення цільових радіовипромінювань (РВ) [3].

Завдання оперативного радіоконтролю завантаження декаметрового діапазону вимагає не лише виявлення РВ, а і точного визначення зони їхнього розміщення та розпізнавання заданих сигналів серед великої кількості невідомих, що не представляє оперативного інтересу. Методи радіоконтролю [4, 5], які часто базуються на ручному пошуку та візуально-апаратному аналізі, є надто повільними. Час, що витрачається оператором на аналіз одного випромінювання (від секунд до хвилин), робить виявлення короткочасних РВ майже неможливим і змушує різко зменшувати ширину діапазону, що обслуговується системою РК. Крім того, ці методи вимагають аналізу всіх виявлених сигналів, що істотно знижує швидкодію системи.

У відповідь на ці виклики актуальною стає розробка автоматизованих систем радіоконтролю, здатних ефективно виділяти (селектувати) РВ за зоною розміщення джерела та видом випромінювання. Метою статті є аналіз існуючих проблем і задач радіоконтролю декаметрового діапазону та обґрунтування застосування підходів, заснованих на використанні апріорної інформації про РВ, для автоматизації процесів їх виявлення та розпізнавання.

Наразі це завдання зазвичай вирішується так: за допомогою панорамного радіоприймального пристрою (РПУ) здійснюється пошук радіовипромінювань у заданому діапазоні частот. Потім усі виявлені випромінювання піддаються слуховим або візуально-апаратним способам аналізу для виявлення заданих РВ. Для визначення зони розміщення джерела радіовипромінювання використовують або антенні системи з великою апертурою для вимірювання вертикальних кутів приходу електромагнітної хвилі (кутів місця), або непрямі методи, що ґрунтуються на тривалому спостереженні за РВ (оцінка федінгу, стійкості пеленгу, тощо).

У зв'язку з тим, що за існуючих методів виявлення оператору пред'являються на аналіз усі виявлені РВ (як задані, так і множину невідомих), швидкодія таких систем РК не задовольняє сучасним вимогам. Зокрема, час, який витрачається оператором на аналіз одного випромінювання, становить від одиниць секунд до одиниць хвилин. При такому способі РК ймовірність виявлення короткочасних РВ прагне до нуля і для реалізації можливості виявляти короткочасні випромінювання доводиться різко знижувати ширину діапазону частот, що обслуговується одним оператором.

Аналіз завантаження діапазону та апріорна інформація

Завантаження декаметрового радіодіапазону залежить від багатьох факторів і дуже впливає на надійність радіозв'язку [1]. Для підвищення надійності радіозв'язку необхідно здійснювати оперативний радіоконтроль завантаження декаметрового діапазону радіочастот. З цією метою слід із множини РВ, які існують в ефірі, виділити задані.

У декаметровому діапазоні хвиль існує множина РВ, що відрізняються:

- частотою;
- розміщенням джерела в просторі (у ближній зоні – зоні поверхневого поширення електромагнітної хвилі, і в дальній – зоні просторового поширення електромагнітної хвилі);
- видом радіовипромінювання (вид модуляції та маніпуляції, параметри модулюючого сигналу, тощо);
- рівнем сигналу;
- часом появи;
- тривалістю.

У ряді випадків радіоконтролю діапазону частот оперативний інтерес представляють РВ, для яких апріорі відомі:

- вид радіовипромінювання;
- зона розміщення джерела радіовипромінювання (ближня);
- ділянки діапазону або номінали частот, на яких задані РВ не можуть з'явитися.

Аналіз поставленої задачі РК показує, що компромісне рішення може бути отримано при послідовному огляді заданого діапазону з проріджуванням всієї сукупності аналізованих каналів з урахуванням наявної апріорної інформації про задані РВ. Зокрема, ряд частотних каналів не потрібно аналізувати, оскільки в них задані РВ не можуть з'явитися. Крім того, можна виключити з аналізу частотні канали з РВ дальньої зони. Це дозволяє зменшити час послідовного огляду до величини, що визначається заданими обмеженнями на час вирішення завдання виявлення та розпізнавання, що визначаються часом існування РВ.

Необхідність автоматизації та оптимізації радіоконтролю

Потрібно при сформульованих умовах, введеному секторі показників якості та обмежених на показник якості оптимізувати систему за сукупністю якості. Особливості вирішуваного завдання такі: передбачається, що відбувається послідовний огляд декаметрового діапазону частот за допомогою автоматизованого панорамного комплексу (РПК) [6 – 8]. При аналізі частотного каналу за сигналами з виходу панорамного приймача має прийматися рішення про наявність чи відсутність одного із заданих РВ. Причому сигнали носять випадковий характер через випадковий в загальному випадку характер повідомлень, що передаються, а також дію перешкод у частотному каналі. Тому завдання прийняття рішення у частотному каналі зводиться до багатоальтернативного виявлення заданих випадкових сигналів за наявності невідомих сигналів. Для проріджування числа аналізованих каналів має бути використана апріорна інформація про характерні особливості заданих РВ.

За умовами завдання багатоальтернативне виявлення заданих РВ повинно проводитись в умовах підвищеної апріорної невизначеності. Зокрема, статистичні характеристики сигналів із виходу РПК апріорі невідомі. Однак можуть бути отримані навчальні вибірки сигналів для заданих РВ. Для інших (невідомих, таких, що заважають) РВ навчальні вибірки неможливо отримати, або вони не є представницькими, тобто вони є недостатніми для синтезу алгоритмів обробки.

При оптимізації структури алгоритмів виявлення має бути враховано сукупність показників якості виявлення та реалізованих витрат з урахуванням використання бортових засобів обчислювальної техніки. Реалізованість засобами обчислювальної техніки забезпечить гнучкість структури отриманих виявників. Для скорочення (проріджування) числа каналів, у яких має проводитися багатоальтернативне виявлення заданих РВ, необхідно

використовувати різні характерні ознаки РВ, що ґрунтуються на обліку апріорних даних про задані РВ (ближня – дальня зона, номінал частоти тощо).

У ряді випадків вирішення поставленої задачі виявлення заданих РВ у частотному каналі виникають ситуації, коли задані РВ свідомо відсутні, і необхідно вирішувати завдання лише виявлення нових (нових, що з'являються на тлі існуючих) РВ з невідомими характеристиками. У деяких працях показано, що з отриманого рішення задачі багатоальтернативного виявлення заданих сигналів за наявності невідомих сигналів як окремий випадок випливає рішення задачі виявлення нових (невідомих) сигналів за умови, що задані сигнали відсутні. На даний час для вирішення задачі виявлення нових РВ використовується енергетичний принцип виявлення. При цьому забезпечується дуже низька можливість правильного виявлення нових РВ (близько 0,3 – 0,4), що не задовольняє реальним вимогам.

Зі змістовної точки зору, завдання розпізнавання заданих РВ може бути сформульоване так: у деяких частотних каналах у невідомий час з'являються задані РВ. Зазвичай у цих каналах діє велике число РВ, зокрема й дуже близьких до заданих, але які не мають оперативного інтересу. Потрібно, аналізуючи конкретний частотний канал, виявити та розпізнати задані РВ. При цьому необхідна автоматизація процесу виявлення та розпізнавання заданих сигналів при підвищенні порівняно з відомими як швидкодії системи, так і якості розпізнавання.

Подібна задача може розглядатися як завдання розпізнавання сигналів в умовах підвищеної апріорної невизначеності, коли розпізнаються М заданих сигналів, а інші сигнали, про які немає достатніх даних для їх відмінності або вони не представляють оперативного інтересу, належать до об'єднаного М +1 – й класу. Це завдання є окремим випадком завдання багатоальтернативного виявлення заданих сигналів [9].

Як показано, розв'язання задачі побудови комплексу радіоконтролю з урахуванням суперечливих обмежень на час вирішення задачі та на апаратні витрати можлива при послідовному огляді декаметрового діапазону частот з проріджуванням потоку аналізованих РВ. При цьому для скорочення часу огляду всього діапазону на аналіз з метою розпізнавання заданих видів РВ надходять не всі виявлені РВ, а тільки ті, що пройшли проріджування за деякими апріорно відомим ознаками РВ, зокрема, таким, як несуча частота, зона розміщення джерела тощо.

Як відомо, внаслідок маневреності та економічності радіоліній у декаметровому діапазоні частот, парк діючих у цьому діапазоні радіостанцій зростає з кожним роком, а частотний радіоспектр залишається незмінно обмеженим. Наприклад, як показано в [1], відсоток однокілогерцових смуг із середнім рівнем перешкод (під перешкодами в [1] маються на увазі всі радіосигнали, що існують у точці прийому) ≥30 дБ для ділянки від 4 до 6 МГц в літньо-осінній період для нічного часу становить 90 %, для ділянки від 12 до 13 МГц – 40 %, а для ділянки від 14 до 15 МГц – 10 %. Якщо в якості аналізованих смуг прийняти не однокілогерцеві, а двокілогерцевих, відсоток двокілогерцових смуг із середнім рівнем перешкод ≥40 дБ буде близько 50 % для всього декаметрового діапазону, тобто для ділянки сумарною шириною 4 МГц (2000 смуг) близько 1000 смуг будуть з перешкод. Апріорно можуть бути відомі до 100 двокілогерцевих смуг, що не становлять оперативного інтересу для вирішення поставленого завдання.

Досліджено в реальних умовах способи проріджування потоку РВ, що надходять на розпізнавання при послідовному огляді декаметрового діапазону частот. Проріджування засноване на ознаках частоти та зони розміщення джерела РВ. Запропоновані способи проріджування дозволяють у кілька десятків разів зменшити потік РВ, що надходять на розпізнавання, і тим самим скоротити час огляду широкого діапазону частот. Методи проріджування, засновані на ознаках зони (ближня зона – зона поверхневого розповсюдження електромагнітної хвилі, дальня – зона просторового розповсюдження електромагнітної хвилі) розміщення джерела РВ, розглянуті у [6 – 8]. Як правило, у точці прийому радіовипромінювань декаметрового діапазону частот, співвідношення кількості

випромінювань дальньої зони до випромінювань ближньої зони становить кілька десятків, і навіть сотень. Якщо оператора цікавлять випромінювання лише ближньої зони, час огляду при автоматичній селекції (дискримінації) випромінювань дальньої зони зменшується майже в стільки ж разів [6].

Класичні задачі виявлення і розпізнавання сигналів

Виявлення та розпізнавання сигналів це важливі задачі обробки сигналів у системах автоматизованого радіоконтролю. Як випливає з проведеного аналізу, ці задачі обробки сигналів вирішуються у складній сигнально-завадовій обстановці. Для спрощення їх вирішення попередньо проводиться селекція РВ за просторовим положенням, частотною ознакою, ознакою новизни і т.д. Це дозволяє проводити аналіз і обробку сигналів в окремих каналах спостереження.

У задачі виявлення сигналів на фоні використовується вирішальне правило

$$\frac{W(\mathbf{x}/s)}{W(\mathbf{x}/0)} \geq h. \qquad (1)$$

Тут $W(\mathbf{x}/s), W(\mathbf{x}/0)$ – функції правдоподібності за умови, що сигнал відповідно присутній або відсутній у каналі спостереження; $\mathbf{x} - L$-мірний вектор представлення спостережень.

У задачі розпізнавання сигналів спостерігається суміш одного з сигналів $s^i(t),\ i=\overline{1,M}$ із завадою $n(t)$: $x(t)=s^i(t)+n(t),\ i=\overline{1,M}$. Необхідно прийняти рішення про те, який із сигналів присутній у спостереженні. Для цього використовується вирішальне правило

$$\frac{W(\mathbf{x}/s^i)}{W(\mathbf{x}/s^l)} \geq h_{il},\ \ i,l=\overline{1,M}. \qquad (2)$$

У цих правилах порогові значення залежать від обраного критерію оптимальності (байєсовського, ідеального спостерігача, максимальної правдоподібності).

В реальних умовах при обробці сигналів має місце апріорна невизначеність відносно щільностей ймовірності сигналів і завад. При цьому у вирішальних правилах використовуються відповідні оцінки щільності ймовірності розподілів, які отримуються з використанням навчальних вибірок заданих сигналів і завад.

Задачі виявлення та розпізнавання сигналів
в умовах підвищеної апріорної невизначеності

Розглянуто класичні задачі виявлення та розпізнавання сигналів неадекватні умовам проведення автоматизованого РК. Для заданих РВ відомі вид та параметри модуляції, вид кодування, характеристики повідомлення, або можуть бути проведені додаткові дослідження з використанням навчальних вибірок відповідних реальних сигналів. Для багатьох інших РВ, що не становлять інтересу для РК, об'єднуються в $M+1$-й клас. Ймовірнісний опис відповідних їм сигналів відсутній і не можуть бути отримані їх навчальні вибірки. Є лише відомості, що РВ із $M+1$-го класу відрізняються від заданих РВ.

Специфіка таких задач обробки сигналів така, що виникає необхідність виявляти і розпізнавати задані сигнали на фоні завад при наявності невідомих сигналів, для яких невідомі ймовірнісні характеристики і відсутня можливість отримати їх оцінки. При цьому потрібно виявити і розпізнати задані сигнали, а також віднести в $M+1$-й клас сигнали, які не цікавлять РК або прийняти рішення, що це нові сигнали і підлягають подальшому аналізу.

Щодо обробки сигналів у каналі спостереження можна ухвалити рішення на користь однієї з гіпотез:

$H^i: x(t)=s^i(t)+n(t),\ i=\overline{1,M}$ – спостерігається один із заданих сигналів;

$H^{M+1}: x(t)=s^i(t)+n(t),\ i=\overline{1,M+1}$ – спостерігається сигнал із класу невідомих.

Розглянута задача фактично є задачею розпізнавання M заданих сигналів при наявності $M+1$-го класу невідомих сигналів. У цій задачі ймовірність помилкових рішень складається з трьох складових: $P_{ош(M)}$ – ймовірності переплутування заданих сигналів між собою, $P_{ош(M+1/M)}$ – ймовірності помилкового прийняття рішень на користь невідомих сигналів при дії одного із M ладанних сигналів, $P_{ош(M/M+1)}$ – ймовірності помилкового прийняття рішень на користь одного із M ладанних сигналів при дії невідомих сигналів.

Нерандомізоване вирішальне правило розпізнаваня виконує розбиття вибірочного простору спостереження сигналів на $(M+1)$-ну області, що не перекриваються. З урахуванням цього перша складова ймовірності помилок за рахунок переплутування M заданних сигналів між собою, друга складова – за рахунок відненсення заданих сигналів до $(M+1)$-го класу невідомих сигналів, третя складова – за рахунок віднесення сигналів із $(M+1)$-го класу до оного із M заданих сигналів.

Відповідно до наявної апріорної інформації можна знайти оцінки лише перших двух складових ймовірності похибок. Оцінити величину третьої складової неможливо. Для урахування третьої складової вводится показник обєму критичної області відхилення гіпотези H_0 про дію $(M+1)$-го сигналу. Ця область має сенс власної області M заданих сигналів. Із змістовної точки зору така задача розпізнавання сигналів заключається у прийнятті рішення про дію одного із M заданих сигналів при наявності $(M+1)$-класу невідомих сигналів.

При вирішенні такої оптимізаційної задачі синтезу методом множників Лагранжа можна отримати наступне вирішальне правило розпізнаванння M заданих сигналів при наявності $(M+1)$-класу невідомих сигналів [13]:

– якщо хоча б для одного значення i $(i=\overline{1,M})$ виконується нерівність

$$P_i W\ \varepsilon/H^i, \alpha^i\ \geq \lambda^i, \tag{3а}$$

то приймається рішення на користь заданих сигналів;

– якщо при всіх значеннях $i=\overline{1,M}$ виконується нерівність

$$P_I W\ \varepsilon/H^i, \alpha^i\ < \lambda^i, \tag{3б}$$

то приймається рішення на користь невідомого сигналу із $(M+1)$-го класу.

На другому етапі при виконанні умови (3а) розпізнаються задані сигнали згідно з наступною системою нерівностей

$$P_i W\ \varepsilon/H^i, \alpha^i\ \geq P_l W\ \varepsilon/H^l,\ \alpha^l\ ,\ \ l=\overline{1,M},\ \ l\neq i. \tag{3в}$$

Вирішальне правило (3) базується на побудові власних областей окремо для кожного із M заданих сигналів. При отриманні цього вирішального правила не використовується інформація про щільності ймовірностей розподілу сигналів із $(M+1)$-го класу і не вимагається наявність їх навчальних вибірок.

Постановка і вирішення розглянутої задачі розпізнавання сигналів в умовах підвищеної апріорної невизначеності – це формалізація вимоги про необхідність розпізнати M заданих сигналів і віднести в $(M+1)$-й клас всі інші сигнали, інформація про які недостатня для їх розпізнавання.

Якщо вводяться втрати лише за рахунок переплутування $(M+1)$-го сигналу з M заданим (не має значення з яким конкретно), приходять до розрізнення двох гіпотез: H^i – про дію одного із M заданих сигналів; H^0 – про дію $(M+1)$-го невідомого сигналу. При цьому отримується наступне вирішальне правило виявлення (селекції) M заданих сигналів в умовах підвищеної апріорної невизначеності

$$H^0:\ \sum_{i=1}^{M} P_i W(\varepsilon / H^i) \le \lambda,$$
$$H^i:\ \sum_{i=1}^{M} P_i W(\varepsilon / H^i) > \lambda. \qquad (4)$$

Неважко бачити, що це вирішальне правило можливо використати і для вирішения протилежної задачі – виявлення нових (невідомих) сигналів.

Конкретизація структури наведених вирішувальних правил (3), (4) виконується урахуванням ймовірних моделей, які вибрані для описування сигналів. Це визначає аналітичний вид цільностей розподілу ймовірностей сигналів і завад у вирішальних правилах.

Оптимізація алгоритмів виявлення та розпізнавання сигналів за сукупністю показників якості

При вирішенні задач виявлення та розпізнавання в процесі автоматизованого РК необхідно враховувати сукупність показників якості

$$\overset{\vee}{k} = (k_1, k_2, k_3, k_4, k_5, k_6)\ . \qquad (5)$$

Ці показники визначають ймовірності помилкових рішень, зокрема, ймовірності переплутування заданих сигналів між собою, ймовірності помилкового прийняття рішень на користь невідомих сигналів при дії одного із M ладанних сигналів, ймовірності помилкового прийняття рішень на користь одного із M ладанних сигналів при дії невідомих сигналів. а також показників якості, що визначають реалізаційні затрати, зокрема, тривалість прийняття ріщень та обчислювальні витрати. Ці показники тісно зв’язані між собою та носять антагоністичний характер, тобто при покращення одних показників інші показники погіршуються.

Для вибору оптимальних алгоритмів виявлення та розпізнавання сигналів слід використовувати основні положення багатокритеріальної оптимізації [14]. При цьому спочатку формується множина допустимих алгоритмів, серед яких у критерійному просторі показників якості (5) за безумовним критерієм переваги видаляються безумовно гірші варіанти і знаходиться підмножина Парето-оптимальних алгоритмів виявлення і розпізнавання сигналів. Для реалізації на практиці може бути вибраний любий із Парето-оптимальних алгоритмів виявлення і розпізнавання сигналів оскільки вони є незрівнянними між собою. Для вибору алгоритму виявлення і розпізнавання сигналів для реалізації може бути використана додаткова експертна інформація, яка враховує відносну важливість показників якості або інша інформація, що є важливою для проведення автоматизованого РК.

Висновки

Проведений аналіз демонструє, що ефективний радіоконтроль декаметрового діапазону є життєво важливим для забезпечення надійності радіозв'язку [10 – 12], проте існуючі методики РК стикаються зі значними обмеженнями швидкодії та ефективності. Основна проблема полягає у необхідності аналізу великої кількості радіовипромінювань, більшість з яких не представляють оперативного інтересу, а також у складності виявлення і розпізнавання короткочасних сигналів.

Запропонований підхід, що базується на використанні апріорної інформації про РВ та проріджування потоку РВ для аналізу, дозволяє значно підвищити швидкість і якість радіоконтролю. Зокрема, використання таких ознак, як несуча частота та зона розміщення джерела (ближня чи дальня), дає змогу істотно скоротити кількість каналів, що підлягають детальному аналізу. Дослідження показали, що автоматична селекція РВ дальньої зони, які становлять переважну більшість, може зменшити час огляду діапазону в десятки або навіть сотні разів.

Розробка алгоритмів виявлення та розпізнавання заданих випадкових сигналів в умовах підвищеної апріорної невизначеності, з урахуванням можливості отримання навчальних

вибірок для цільових РВ, є ключовим напрямком для подальшої автоматизації. Це дає можливість виявляти і розпізнавання заданих сигналів при наявності невідомих сигналів і дозволить створити гнучку та високоефективну систему радіоконтролю. Впровадження бортових засобів обчислювальної техніки забезпечить необхідну реалізованість та адаптивність таких систем до динамічних умов декаметрового діапазону.

Список літератури;

1. Комарович В. Ф., Сосунов В. М. Случайные радиопомехи и надежность КВ связи. Москва : Связь, 1977. 135 с.
2. Короткохвильовий радіозв'язок: Настанова. Київ : Центр учбової літ-ри, 2024. 83 с.
3. Заславець В.П., Долина М.П., Чечуй О.В. Особливості розрахунку завадозахищеності ліній радіозв'язку в умовах радіоподавлення (радіоелектронного конфлікту // Системи озброєння і військова техніка. 2020. № 1(61). С. 7–12. https://doi.org/10.30748/soivt.2020.61.01
4. Довідник військового зв'язківця. Засоби радіоелектронної боротьби та розвідки, які використовуються російською федерацією. Київ : ЦУЛ, 2024. 64 с.
5. Майборода І. М., Власов К. В., Глущенко М. О. Застосування і перспективи розвитку мобільних засобів радіоелектронної розвідки тактичної ланки сил сектору безпеки та оборони України // Честь і закон. 2024. № 2 (89). С. 83–90.
https://dspace.nlu.edu.ua/jspui/bitstream/123456789/20200/1/Vlasov_83%E2%80%9390.pdf.
6. Обухов Н. П., Кикоть А. В., Голобородько Ю. Н., Горбачинский И. С. Авторское свидетельство № 177720 от 09.04.82 «Устройство автоматического панорамного обнаружения и пеленгования с дискриминацией радиоизлучений источников дальней или ближней зоны».
7. Голобородько Ю. Н., Горбачинский И. С., Авторское свидетельство № 177874 от 01.09.82 «Панорамное приемное устройство с дискриминацией зоны размещения источника».
8. Гурьев В. И., Горбачинский И. С., Голобородько Ю. Н., Авторское свидетельство № 210497 от 26.10.84 «Панорамное радиоприемное устройство для определения зоны нахождения источника радиоизлучения».
9. Omelchenko V. A. Balabanov V. V. Bezruk V. M. Goloborod'ko Ju. N., Detection of Changes in the Random Signal Properties by Spectral Methods Under the Conditions of A Priori Uncertainty // Telecommunications and Radio Engineering. 1999. № 53 (9-10). P. 1–10.
10. Голобородько Ю. Н., Кузниченко В. С. Перспективи розвитку засобів радіоконтролю // Зб. тез наук.-практ. конф. „Проблеми забезпечення внутрішньої безпеки держави". Харків, 2005.
11. Молодцов В.А., Писарєв А.В., Радченко І.О., Лисенко О.В. Підвищення ролі радіоелектронної боротьби збройних сил США в асиметричних діях повномасштабної сетецентричної війни / Національна академія Національної гвардії України // Молодий вчений. 2022. № 1 (101). С. 128–134. doi: https://doi.org/10.32839/2304-5809/2022-1-101-27,https://molodyivchenyi.ua/index.php/journal/issue/view/38, (дата звернення: 20.04.2025).
12. Голобородько Ю. М., Захаров В. М., Русинов М. Г., Снігуров А. В., Ніколаєв М. О., Кузніченко В. С., Повтарев В. І., Комплексна система радіомоніторингу засобів радіозв'язку // Честь і закон. 2005. №3. С.18–22.
13. Безрук В.М., Певцов Г.В. Теоретические основы проектирования систем распознавания сигналов для автоматизированного радиоконтроля. Харьков : Коллегиум, 2007. 430 с.
14. Березовский, Б.А., Барышников Ю.М., Борзенко В.И., Кепнер Л.М. Многокритериальная оптимизация. Математические аспекты. Москва : Наука, 1986. 128 с.

Надійшла до редколегії 11.06.25

Відомості про авторів:

Безрук Валерій Михайлович – доктор технічних наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри інформаційно-мережної інженерії, Україна; email: valerii.bezruk@nure.ua; ORCID: https://orcid.org/0000-0002-5482-9960

Голобородько Юрій Миколайович – кандидат технічних наук, Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій, Україна; email: yurii.holoborodko@nure.ua; ORCID:https://orcid.org/0009-0007-4040-9467

Заболотний Володимир Ілліч – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри безпеки інформаційних технологій, Україна; email: volodymyr.zabolotnyi@nure.ua; ORCID: https://orcid.org/0000-0003-3258-8489

Скибенко Микола Сергійович – старший викладач кафедри безпеки інформаційних технологій, факультет комп'ютерної інженерії та управління; Харківський національний університет радіоелектроніки; Україна; e-mail: mykola.skybenko@nure.ua, ORCID: https://orcid.org/0009-0002-4838-9329

УДК 004.056.5 **DOI:10.30837/rt.2025.3.222.04**

І.В. ЛИСИЦЬКА, д-р техн. наук, К.Є. ЛИСИЦЬКИЙ, І.М. ГАЛЬЦЕВА,
Є.П. КОЛОВАНОВА, канд. техн. наук

ОСОБЛИВОСТІ ПОБУДОВИ НЕЛІНІЙНИХ ПЕРЕТВОРЕНЬ БЛОКОВИХ СИМЕТРИЧНИХ ШИФРІВ

Вступ

Нелінійність є життєво важливою для криптографічної стійкості, оскільки вона ускладнює лінійні та алгебраїчні атаки. Без нелінійних перетворень шифри були б вразливими до відносно простих криптоаналітичних методів, які могли б ефективно зламати шифр.

Історія виникнення нелінійних перетворень у блокових симетричних шифрах бере свій початок у фундаментальних принципах Шеннона. У роботі "Communication Theory of Secrecy Systems" Шеннон заклав основи сучасної криптографії. Він розглянув важливі загальні концепції розсіювання і перемішування. Для досягнення цих властивостей Шеннон пропонував використання "комбінованих шифрів", що містять простіші операції, такі як шифри заміни (substitution) і шифри перестановки (permutation).

У 1970 р. Хорст Файстель став ключовою фігурою у розробці перших сучасних блокових шифрів. Його робота над шифром LUCIFER в IBM (1971) заклала основу для архітектури, відомої як мережа Файстеля. У цих мережах шифрування відбувається в кілька раундів, де кожен раунд включає нелінійну функцію, яка залежить від ключа.

У ранніх розробках блокових шифрів Файстель вже використовував механізм S-блоків. S-блоки є основним компонентом, що реалізує нелінійні заміни у симетричних шифрах. Вони приймають певну кількість вхідних бітів і перетворюють їх у певну кількість вихідних бітів нелінійним чином.

У 1977 р. Національне бюро стандартів США (зараз NIST) запропонувало блоковий симетричний шифр. DES [1], який базується на мережі Файстеля, активно використовує S-блоки як єдиний нелінійний елемент. Конструкція S-блоків DES спочатку трималася в секреті. Пізніше стало відомо, що ці S-блоки були ретельно спроєктовані, щоб забезпечити стійкість до нового тоді методу криптоаналізу – диференціального криптоаналізу.

Цей потужний метод криптоаналізу був запропонований Біхамом і Шаміром у 1990 р., хоча його використання було відоме розробникам DES раніше [2]. Лінійний криптоаналіз був відкритий Мацуї в 1993 р. [3].

Еволюція блокових симетричних шифрів після DES рухалася в напрямку збільшення довжини ключа, розміру блоку та вдосконалення внутрішньої структури для стійкості до нових криптоаналітичних атак. З'являлись нові ідеї побудови нелінійних перетворень з покращеними криптографічними показниками. Національний інститут стандартів і технологій США (NIST) ініціював відкритий конкурс на новий стандарт блокового шифру, який мав би замінити DES. Цей конкурс тривав з 1997 по 2000 р. і сприяв розробці безлічі нових та інноваційних алгоритмів. Деякі з помітних кандидатів: Serpent, Mars, Twofish, RC6, Blowfish, IDEA, Skipjack, CAST-128 / CAST-256 [4].

Крім структури типу мережа фейстеля були запропоновані структура "мережа Лая-Массі" (Lay-Massey network) [5] та SPN-мережа.

SPN (Substitution-Permutation Network), або підстановочно-перестановочна мережа, є однією з фундаментальних архітектур, що використовуються для побудови симетричних блокових шифрів. Найвідомішим прикладом шифру, що базується на SPN, є AES (Advanced Encryption Standard) – фіналіст конкурсу NIST [6].

S-блоки (Substitution-boxes) є серцем нелінійності в багатьох симетричних блокових шифрах (особливо в SPN-структурах). Їхня якість має вирішальне значення для стійкості шифру до криптоаналітичних атак.

Основні властивості S-блоків, які забезпечують криптографічно стійке нелінійне перетворення

1. Висока нелінійність (High Non-linearity): Це найважливіша властивість. S-блок має бути максимально нелінійним. Це протистоїть лінійному криптоаналізу.

2. Низька диференціальна однорідність (Low Differential Uniformity): Для будь-якої ненульової вхідної різниці Δx, вихідні різниці $\Delta y = S(x) \oplus S(x \oplus \Delta x)$ повинні розподілятися якомога рівномірніше. Це протистоїть диференціальному криптоаналізу. Ідеально – мати диференціальну однорідність 2 (це означає, що для будь-якої вхідної різниці та будь-якої вихідної різниці, існує не більше двох пар вхідних значень, що дають таку різницю).

3. Бієктивність (Bijectivity): S-блок має бути бієктивною функцією, тобто кожне вхідне значення повинне відображатися на унікальне вихідне значення, і навпаки. Це забезпечує оборотність процесу дешифрування.

4. Збалансованість (Balancedness): Кожне можливе вихідне значення S-блоку має з'являтися однакову кількість разів. Це запобігає витоку інформації через статистичний зсув.

5. Лавинний ефект (Avalanche Effect). Невелика зміна на вході S-блока (наприклад, зміна одного біта) повинна призводити до значної зміни на виході (багато вихідних бітів повинні змінити своє значення). Це властивість важлива для поширення змін по всьому шифротексту, що робить його стійким до статистичних атак.

6. Відсутність фіксованих точок та антифіксованих точок (No Fixed Points and Antifixed Points).

7. Високий алгебраїчний ступінь (High Algebraic Degree): Функція S-блока повинна мати високий алгебраїчний ступінь та бути складною для опису за допомогою простих алгебраїчних виразів. Це протистоїть алгебраїчним атакам [7].

S-блоки можуть бути сконструйовані різними способами, і кожен підхід має свої переваги та недоліки.

Різні структури (методи побудови) S-блоків

1. Детерміновані S-блоки (Fixed S-boxes) [6, 8, 9]

Це S-блоки, які є жорстко заданими таблицями або математичними функціями, і їхні значення не змінюються під час шифрування. Типовими прикладами таких S-блоків є S-блоки алгоритмів DES, Serpent, Camellia, ARIA, AES та деякі інші.

Переваги цього підходу:

- простота реалізації та швидкість. Це робить їх придатними для апаратних реалізацій та високопродуктивних систем;

- ретельне проєктування та перевірка: S-блоки, які використовуються в широко відомих та перевірених алгоритмах розроблені таким чином, щоб мати бажані криптографічні властивості, такі як висока нелінійність, низька диференціальна однорідність та стійкість до відомих криптоаналітичних атак (лінійний та диференціальний криптоаналіз);

- стабільність та передбачуваність властивостей: Оскільки S-блоки є фіксованими, їхні криптографічні властивості є постійними та передбачуваними. Це дозволяє розробникам шифрів точно оцінювати безпеку алгоритму;

- сумісність та стандартизація: Використання фіксованих S-блоків полегшує стандартизацію алгоритмів шифрування, забезпечуючи сумісність між різними реалізаціями та системами.

Недоліки:

- вразливість до атак на основі структури: якщо S-блоки мають приховану математичну структуру, це може бути використано в алгебраїчних атаках або інших видах криптоаналізу для пошуку вразливостей;

- потенційна вразливість до атак по стороннім каналам (Side-channel attacks): оскільки S-блоки є таблицями пошуку, їх використання може бути вразливим до атак по стороннім каналам, таким як атаки по часовим затримкам (cache-timing attacks). Це пов'язано з тим, що

час доступу до пам'яті може залежати від того, до якої частини S-блоку відбувається звернення, що може дати зловмиснику інформацію про дані або ключ;

- відсутність адаптивності: фіксовані S-блоки не змінюються залежно від ключа або раунду шифрування. Це може зробити криптоаналіз простішим у порівнянні з системами, які використовують динамічні (ключозалежні) S-блоки, оскільки атакуючому не потрібно обчислювати або дізнаватися структуру S-блоку для кожного окремого випадку;

- ризик "бекдорів" (Backdoors): хоча це малоймовірно для широко відомих стандартів, теоретично існує ризик того, що фіксовані S-блоки могли бути розроблені з прихованими "бекдорами", які дозволяють полегшити криптоаналіз за певних умов;

- розробка криптографічно сильних фіксованих S-блоків є дуже складним завданням.

Пошук S-блоків, які задовольняють всім необхідним криптографічним критеріям (висока нелінійність, низька диференціальна однорідність, хороший лавинний ефект тощо), вимагає значних досліджень та обчислювальних ресурсів.

2. Динамічні S-блоки (Dynamic/Key-Dependent S-boxes) [10, 11, 12]

Це S-блоки, які генеруються або змінюються під час виконання алгоритму, зазвичай, на основі секретного ключа. Типовими представниками цього підходу є Blowfish, Twofish, Khufu/Khafre.

Переваги динамічних S-блоків:

- підвищена стійкість до криптоаналізу "білої скриньки" (White-box Cryptanalysis). Оскільки структура S-блоків змінюється для кожного ключа, криптоаналітику набагато складніше провести атаки, що базуються на статичній структурі, оскільки він не знає, які саме S-блоки використовуються для конкретного ключа. Це ускладнює пошук "бекдорів" або прихованих вразливостей, які могли би бути вбудовані у фіксовані S-блоки [13];

- ускладнення універсальних атак. Атаки, які покладаються на фіксовані властивості S-блоків (наприклад, лінійні або диференціальні апроксимації), стають значно складнішими, оскільки ці властивості змінюються з кожним ключем [14];

- зменшення ризику "бекдора". Якщо S-блоки генеруються криптографічно надійною функцією з використанням ключа, то розробник алгоритму не може навмисно вбудувати приховану вразливість, оскільки вона залежатиме від ключа, який йому невідомий. Це підвищує довіру до прозорості розробки.

Недоліки динамічних S-блоків:

- високі обчислювальні витрати на генерацію. Процес генерації ключозалежних S-блоків може бути обчислювано дорогим і займати значний час, особливо якщо S-блоки повинні мати високі криптографічні властивості. Це може уповільнити фазу ініціалізації шифру [10];

- складність забезпечення криптографічних властивостей. Гарантувати, що згенеровані S-блоки завжди матимуть бажані криптографічні властивості (високу нелінійність, низьку диференціальну однорідність тощо) для будь-якого ключа, є складним завданням. Якщо механізм генерації може створювати "слабкі" S-блоки для певних ключів, це може створити вразливість [15];

- складність криптографічного аналізу. Хоча динамічні S-блоки ускладнюють атаки для криптоаналітика, вони також ускладнюють криптографічний аналіз шифру. Довести безпеку шифру зі змінними S-блоками набагато важче, оскільки необхідно аналізувати алгоритм генерації S-блоків та їхні властивості в цілому, а не лише одну фіксовану таблицю [16];

- збільшення складності реалізації. Алгоритми з динамічними S-блоками є більш складними для реалізації порівняно з тими, що використовують фіксовані таблиці, оскільки вони вимагають додаткового коду для процедури генерації S-блоків.

Таким чином, використання динамічних S-блоків є компромісом. Хоча вони додають шар складності для криптоаналітика, вони також створюють проблеми для розробника шифру у забезпеченні гарантованої безпеки для всіх можливих випадків.

3. Хаотично згенеровані S-блоки

Деякі дослідження пропонують використовувати хаотичні системи для генерації S-блоків. Хаотично згенеровані S-блоки (Substitution-boxes) використовують властивості хао-

тичних систем (такі як чутливість до початкових умов, непередбачуваність, ергодичність та топологічне перемішування) для створення S-блоків, що демонструють високі криптографічні властивості. Це робить їх привабливим напрямком у створенні сучасних шифрів. Існують розробки у таких напрямках:

S – блоки, засновані на логістичному відображенні (Logistic Map);

S – блоки, засновані на системах Чуа (Chua's Circuit);

S – блоки, засновані на Хенон-відображенні (Hénon Map) або інших багатовимірних хаотичних відображеннях;

S – блоки, що використовують гіперхаотичні системи.

Переваги хаотично згенерованих S-блоків:

- висока нелінійність та стійкість до криптоаналізу. Хаотичні системи за своєю природою є нелінійними, що дозволяє генерувати S-блоки з високими показниками нелінійності. Це є ключовою властивістю для забезпечення стійкості до лінійного та диференціального криптоаналізу [17, 18];

- підвищена чутливість до ключа та розповсюдження помилок (Avalanche Effect): навіть мінімальні зміни у вхідних даних (наприклад, у секретному ключі) призводять до значно різних вихідних даних. Це забезпечує сильний лавинний ефект (avalanche effect), де зміна одного біта у відкритому тексті або ключі призводить до зміни близько половини бітів у шифротексті, що ускладнює криптоаналіз [17, 19];

- додаткове заплутування (Confusion): Хаотичні відображення, хоча й детерміновані, демонструють надзвичайно складну динаміку. Ця складність посилює властивість заплутування (confusion), яка є однією з двох фундаментальних властивостей безпечного шифрування (за Шенноном) [17, 20];

- ергодичність та статистична однорідність: Властивість ергодичності хаотичних відображень дозволяє їм з часом охоплювати весь простір станів, що сприяє статистичній однорідності зашифрованих вихідних даних. Це підвищує стійкість до криптоаналізу, який базується на статистичних властивостях шифротексту [17];

- потенціал для ефективності (у певних реалізаціях). Хаотичні системи можуть бути використані для створення S-блоків, які є ефективними з точки зору використання ресурсів та пропускної здатності [17, 18].

Недоліки:

- складність гарантування криптографічних властивостей: важко довести, що хаотичні відображення завжди генерують S-блоки з бажаними криптографічними властивостями (нелінійність, диференціальна однорідність тощо) для всіх можливих початкових умов;

- реалізаційні складнощі: вимагають точної реалізації хаотичних систем, що може бути чутливим до помилок округлення в комп'ютерних системах;

- менш досліджені: хоча є багато досліджень, вони поки не так широко інтегровані в основні стандарти, як фіксовані S-блоки.

Більшість цих "хаотичних" шифрів залишаються в царині досліджень і академічних публікацій. Вони демонструють цікаві концепції та потенціал, але їм, як правило, бракує суворого криптоаналізу та стандартизації, які необхідні для прийняття в якості надійних криптографічних примітивів для широкого використання. Ці шифри часто є "одноразовими" пропозиціями в наукових статтях і рідко отримують загальновизнані назви.

Загалом, використання хаотичних систем у генерації S-блоків є перспективним напрямом для розробки криптографічно стійких алгоритмів, особливо для застосувань, де потрібна висока нелінійність та стійкість до відомих атак.

4. Використання випадкових S-блоків

Окремий напрямок, який було розроблено у чисельних роботах, серед яких [21 – 23]. Використання "випадкових" S-блоків у блочних шифрах – це концепція, де таблиця заміни не є жорстко зафіксованою або детерміновано математично побудованою, натомість генерується або вибирається таким чином, що її конкретна структура виглядає випадковою або є непередбачуваною. Запропоновано в якості S-блоків шифрів використовувати випадкові підстановки з виходу генератора випадкових підстановок, що проходять перевірку на відповідність додатним критеріям відбору.

Важливий висновок міститься в тому, що для випадково взятої булевої функції більшість її криптографічних параметрів близькі до оптимальних.

В результаті набуває актуальності задача побудування шифрів, в яких без зниження стійкості можуть бути використані випадкові S-блоки, яка знайшла своє перше вирішення в роботі [24] та ряді інших.

Переваги використання випадкових S-блоків:

- захист від "бекдорів" та прихованих вразливостей. Якщо S-блоки генеруються випадковим або псевдовипадковим чином, це усуває можливість для розробника шифру навмисно вбудувати приховану слабкість або "бекдор", оскільки навіть сам дизайнер не знатиме точної структури S-блоків для конкретного ключа. Це підвищує довіру до шифру;
- стійкість до криптоаналізу. Для випадково взятої булевої функції більшість її криптографічних параметрів близькі до оптимальних;

Недоліки використання випадкових S-блоків:

- існує ризик, що для певних випадкових початкових значень (або ключів) можуть бути згенеровані "слабкі" S-блоки, що скомпрометують шифр.
- обчислювальні витрати на генерацію. Процес генерації S-блоків на основі ключа або випадкових даних може займати значний час.

Порівняння за основними властивостями різних типів S-блоків надано у табл. 1.

Таблиця 1

Властивостями різних типів S-блоків

Властивість / Тип S-блоку	Детерміновані S-блоки	Динамічні S-блоки	Хаотично згенеровані S-блоки	Випадкові S-блоки
Генерація	Фіксовані, заздалегідь визначені таблиці підстановки розроблені з використанням складних математичних конструкцій або пошукових алгоритмів для досягнення бажаних криптографічних властивостей	Генеруються або модифікуються під час виконання алгоритму шифрування, часто залежать від ключа або стану	Генеруються за допомогою хаотичних систем (наприклад, хаотичних відображень, клітинних автоматів), використовують детерміновану, але непередбачувану поведінку хаосу	Генеруються випадково (наприклад, як випадкові підстановки)
Змінюваність під час шифрування	Не змінюються	Змінюються	Зазвичай змінюються	Можуть бути статичними або динамічними
Нелінійність	Висока: зазвичай розроблені з максимально можливою нелінійністю для протидії лінійному криптоаналізу	Може бути високою: залежить від методу динамічної генерації	Залежить від хаотичної системи	Висока але не гарантовано оптимальна
Диференціальна однорідність	Низька	Може бути низькою: залежить від методу динамічної генерації	Залежить від хаотичної системи: повинна бути низькою через властивості хаосу, але може вимагати перевірки	В середньому, випадкові S-блоки мають прийнятну диференціальну однорідність, але не гарантовано оптимальну
Бієктивність (зворотність)	Так	Може бути так/ні	Може бути так/ні	Так
Складність реалізації	Низька	Висока	Висока	Низька
Складність проєктування	Висока	Середня	Середня	Низька
Залежність від ключа	Ні	Так	Так:	Може бути так/ні

У контексті використання симетричних блокових шифрів (БСШ) в умовах постквантової криптографії основна увага зосереджена на протидії алгоритму Гровера. Цей алгоритм дозволяє квантовому комп'ютеру здійснити пошук по неструктурованій базі даних (що включає перебір ключів) з квадратичним прискоренням. Тобто, якщо для класичного комп'ютера потрібно операцій для повного перебору n-бітного ключа, то для квантового комп'ютера – приблизно операцій. Це означає, що для забезпечення "постквантової стійкості" симетричних шифрів найпростішим і найпоширенішим рішенням є збільшення (подвоєння) довжини ключа. Щодо S-блоків, у постквантовій криптографії основні вимоги до них залишаються тими ж, що й у класичній.

Висновки

Вибір структури S-блоку завжди є компромісом між безпекою, продуктивністю та простотою реалізації. Вибір підходу до побудови S-блоку залежить від багатьох факторів, включаючи бажаний рівень безпеки, обчислювальні ресурси, доступні для проєктування, та специфічні вимоги до продуктивності. Сучасні шифри, як правило, віддають перевагу алгебраїчним S-блокам через їхні доведені криптографічні властивості та елегантність.

Однак гібридні підходи, що поєднують алгебраїчні та евристичні методи, також можуть бути використані для досягнення оптимальних результатів.

Щодо умов квантової криптографії. Підходи до розробки самих S-блоків кардинально не змінюються. Умови квантової криптографії впливають на вибір серед існуючих, добре відомих S-блоків. Якщо деякі з них виявляться значно складнішими для квантової реалізації та менш вразливими до певних типів квантових атак.

Постквантові БСШ, що пропонуються для стандартизації або широкого використання, як правило, покладаються на ретельно проаналізовані, фіксовані S-блоки. Причини ті ж самі, що й у класичному випадку: складність аналізу безпеки, обчислювальні витрати на генерацію та проблеми з відтворюваністю.

Таким чином, "різні типи S-блоків" (фіксовані, ключово-залежні, хаотично згенеровані, випадкові) залишаються предметом досліджень, але в контексті практичних постквантових БСШ переважає підхід використання перевірених S-блоків з достатньою довжиною ключа для захисту від атаки Гровера.

Список літератури:

1. FIPS PUB 46-3 (Federal Information Processing Standards Publication 46-3), Data Encryption Standard (DES). [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/publications/detail/fips/46/3/archive/1999-10-25
2. Biham E. and Shamir A. Differential Cryptanalysis of DES-Like Cryptosystems ; Menezes A.J. and Vanstone S.A., Eds. // Advances in Cryptology-CRYPTO'90, Springer. Berlin, 1990. P. 2–21. [Електронний ресурс]. Режим доступу: https://www.scirp.org/reference/referencespapers?referenceid=2235215
3. Matsui M. Linear Cryptanalysis Method for DES Cipher ; T. Helleseth (Ed.) // Advances in Cryptology-EUROCRYPT '93 (LNCS. 1994. Vol. 765. P. 386–397). Springer, Berlin, Heidelberg.
4. National Institute of Standards and Technology. (n.d.). AES Development. Computer Security Resource Center. Retrieved from https://csrc.nist.gov/projects/cryptographic-standards-and-guidelines/archived-crypto-projects/aes-development
5. Lai X., Massey J. L., & Murphy S. A proposal for a new block encryption standard ; T. Helleseth (Ed.). // Advances in Cryptology-EUROCRYPT '90 (LNCS. 1991. Vol. 473. P. 389–404). Springer, Berlin, Heidelberg.
6. Daemen J., Rijmen V. The Design of Rijndael-The Advanced Encryption Standard. Springer-Verlag; Berlin Heidelberg, New York, 2002. 10.1007/978-3-662-04722-4.
7. Chew L., C. N., & Ismail E. S. S-box Construction Based on Linear Fractional Transformation and Permutation Function // Sensors. 2020. No 20(5). P. 826.
8. Canteaut A., Duval S., and Leurent G. Construction of LightweightS-Boxes Using Feistel and MISTY Structures // Lecture Notes in Com-puter Science. Springer International Publishing. 2016. P. 373–393. doi:10.1007/978-3-319-31301-6_22.
9. Jing-Mei L., Bao-Dian W., Xiang-Guo C., Xin-Mei W. Cryptanalysis of Rijndael S-box and improvement // Appl Math Comput. 2005. No 170. P. 958–975. 10.1016/j.amc.2004.12.043.
10. Blowfish Algorithm with Examples. (n.d.). GeeksforGeeks. Retrieved June 28, 2025, from https://www.geeksforgeeks.org/blowfish-algorithm-with-examples/.

11. The Twofish Encryption Algorithm: A 128-Bit Block Cipher. (1999). Bruce Schneier, John Kelsey, Doug Whiting, David Wagner, Chris Hall, Niels Ferguson.
12. Farwa S., Sohail A., Muhammad N. A novel application of elliptic curves in the dynamical components of block ciphers // Wirel. Pers. Commun. 2020. Vol. 115. P. 1309–1316.
13. Mustafa R. A., Jalab H. A., & Al-Qurashi M. S. Dynamic S-Box Construction Using Mordell Elliptic Curves over Galois Field and Its Applications in Image Encryption // Mathematics. 2023. Vol. 12(4). P. 587. https://doi.org/10.3390/math12040587
14. Keliher L., Meijer D., & Tavares E. Key-dependent S-boxes and differential cryptanalysis // International Workshop on Fast Software Encryption. 2004. P. 37–51. Springer, Berlin, Heidelberg.
15. Husain I., & Mahmood M. Construction of High Quality Key-dependent S-boxes // International Journal of Computer Science and Engineering (IJCSE). 2017. Vol. 6(3). P. 47–53.
16. Mustafa R. A., & Ahmad N. R. SECURITY ANALYSIS BETWEEN STATIC AND DYNAMIC S-BOXES IN BLOCK CIPHERS // Journal of Information Systems and Technology Management. 2018. Vol. 11(1). P. 19–32.
17. Huda M. A. A., & Al-Qurashi M. S. Chaos-Based S-Boxes as a Source of Confusion in Cryptographic Primitives // Electronics. 2025. Vol.14(11). P. 2198. https://doi.org/10.3390/electronics14112198
18. Zhang M., Han F., Sun K., Zhou C., & Xu F. A Novel S-Box Design Algorithm Based on a New Compound Chaotic System // Symmetry. 2020. Vol. 12(9). P. 1469.
19. Singh A. P., Kumar R., & Kumar R. Creation of S-Box based One-Dimensional Chaotic Logistic Map: Colour Image Encryption Approach // Open Access Journal of Information Technology. 2022. Vol. 1(1). P.1–11.
20. Belkhode V., & Raut R. A Chaos Based Method for Efficient Cryptographic S-box Design // International Journal of Computer Applications. 2016. Vol.133(1). P. 18–21.
21. Dolgov V.I., Lisitska I.V., Lisitskyi K.Ye. The new concept of block symmetric ciphers design // Telecommunications and Radio Engineering. 2017. Vol. 76, issue 2. P. 157–184.
22. Лисицький К., & Лисицька І. Mathematical model of random substitution // Radiotekhnika. 2020. No 3(202). P. 116–124. DOI: https://doi.org/10.30837/rt.2020.3.202.12
23. Lisickiy, K., Dolgov, V., Lisickaya, I., & Kuznetsova, K. Block Symmetric Cipher with Random S-boxes // International Journal of Computing. 2019. Vol. 18, iss. 1. P. 89–100. DOI: 10.47839/ijc.18.1.1278.
24. Dolgov V.I., Lisitska I.V., Lisitskyi K.Ye. The new concept of block symmetric ciphers design // Telecommunications and Radio Engineering. 2017. Vol. 76, issue 2. P. 157–184.

Надійшла до редколегії 17.06.2025

Відомості про авторів:

Лисицька Ірина Вікторівна – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри кібербезпеки інформаційних систем, мереж і технологій, навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Харківський національний університет радіоелектроніки, професор кафедри безпеки інформаційних технологій, Україна; e-mail: ivlisitska@karazin.ua; ORCID: https://orcid.org/0000-0001-6758-9516

Лисицький Костянтин Євгенійович – PhD, Харківський національний університет імені В. Н. Каразіна, доцент кафедри математичного моделювання і аналізу даних навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Національний аерокосмічний університет "Харківський авіаційний інститут", старший викладач кафедри комп'ютерних систем, мереж і кібербезпеки; Україна; e-mail: constantin.lisickiy@gmail.com; ORCID: https://orcid.org/0000-0002-7772-3376;

Гальцева Ірина Михайлівна – Харківський національний університет імені В. Н. Каразіна, старший викладач кафедри кібербезпеки інформаційних систем, мереж і технологій навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Україна; e-mail: irina.galceva@karazin.ua

Колованова Євгенія Павлівна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, доцент кафедри кібербезпеки інформаційних систем, мереж і технологій навчально-наукового інституту комп'ютерних наук та штучного інтелекту; Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій; Україна; e-mail: e.kolovanova@gmail.com; ORCID: https://orcid.org/0000-0002-0326-2394

УДК 004.056.55 **DOI:10.30837/rt.2025.3.222.05**

Р.І. МОРДВІНОВ, канд. техн. наук

ПРОТОКОЛИ З НУЛЬОВИМ РОЗГОЛОШЕННЯМ: ТЕОРЕТИЧНІ ОСНОВИ ТА ЗАСТОСУВАННЯ В СУЧАСНІЙ КРИПТОГРАФІЇ

Вступ

У сучасному цифровому світі, де обмін інформацією відбувається постійно, забезпечення конфіденційності та цілісності даних є надзвичайно важливим. Криптографія відіграє ключову роль у захисті інформації, надаючи інструменти для шифрування, аутентифікації та забезпечення безпечного зв'язку. Протоколи з нульовим розголошенням (ZKP) є однією з найбільш інноваційних та перспективних концепцій у криптографії, що дозволяє одній стороні (доказувачу) переконати іншу сторону (верифікатора) у правдивості певної заяви, не розкриваючи при цьому жодної додаткової інформації, крім самого факту її істинності [1].

Поняття доказів з нульовим розголошенням сформувалося в середині 1980-х років завдяки піонерським роботам Ш. Гольдвассер, С. Мікалі та Ч. Ракова, які ввели концепцію «складності знання» в інтерактивних доказах [2]. На той час було показано, що за наявності односторонніх функцій будь-яка мова класу NP має інтерактивний доказ з нульовим знанням, тобто можна переконувати у твердженнях без розкриття секретної інформації. Класичними прикладами кінця 1980-х стали протоколи автентифікації, побудовані за схемою трьохходового обміну (так звані σ-протоколи). Зокрема, схема Фіата–Шаміра і протокол Шнорра дозволили переконливо доводити знання секретного ключа без його розголошення – ці роботи заклали основу поняття Zero-Knowledge у прикладних сценаріях. Важливим етапом розвитку стало усунення інтерактивності: 1988 р. М. Блум, П. Фельдман і С. Мікалі вперше запропонували концепцію неінтерактивних доказів з нульовим розголошенням, коли доведення здійснюється за допомогою спільного випадкового рядка без обміну запитами. Цей результат продемонстрував, що взаємодія не є обов’язковою для збереження нульового розголошення, і відкрив шлях до подальших удосконалень протоколів.

У 1990-х роках теорія ZKP стрімко розвивалася: було доведено повноту ZKP для всього класу NP, розроблялися багатоверифікаторні (multi-prover) системи і застосовувалися нові техніки, такі як докази з перевіркою за допомогою поліномів (ідеї PCP). Ці дослідження підготували ґрунт для появи стислих доказів знання. Починаючи з 2010-х фокус змістився на підвищення ефективності: з’явилися перші практичні реалізації NIZK-доказів для довільних обчислень. Зокрема, було запропоновано конструкції типу zk-SNARK, які дали можливість генерувати дуже короткі докази й перевіряти їх у часі, не залежному від складності твердження. Наприклад, до 2016 р. Й. Грот розробив перевірюваний за мілісекунди SNARK з постійного розміру доказом [5], що стало проривом для використання ZKP на практиці. У 2018 р. з’явилася альтернатива на основі лише геш-функцій – zk-STARK, що не потребує довіреної установки та є стійким до квантових атак [4]. Ці досягнення збіглися з вибухом інтересу до ZKP у контексті блокчейну: якщо перші концепції 1980-х були суто теоретичними, то з середини 2010-х докази з нульовим розголошенням впроваджуються у реальні системи. Яскравий приклад – криптовалюта Zcash, у якій технологія zk-SNARK забезпечила повну приватність транзакцій. Відправник, отримувач і сума платежу приховані, але достовірність транзакції можна перевірити публічно. На платформі Ethereum також з’явилися рішення, що використовують ZKP (проєкти Aztec, StarkWare та ін.) для підвищення конфіденційності і масштабованості смарт-контрактів. Таким чином, станом на зараз ZKP перетворилися з академічного концепту на ключову технологію у сфері криптовалют і розподілених систем.

Базове визначення протоколу з нульовим розголошенням включає три ключові властивості:

- повнота (Completeness). Якщо твердження є істинним, чесний доказувач зможе переконати чесного версифікатора;
- надійність (Soundness). Якщо твердження є хибним, жоден (навіть шахрайський) доказувач не зможе переконати чесного верифікатора (за винятком незначної ймовірності);
- нульове розголошення (Zero-Knowledge). У процесі доведення верифікатор не дізнається нічого, крім того, що твердження є істинним.

Дослідження в галузі протоколів з нульовим розголошенням є досить широким і включає значну кількість наукових праць. Крім фундаментальної роботи Гольдвассер, Мікалі та Ракоффа [2], важливими є праці, що досліджують різні типи ZKP та їхні застосування. Огляд різних конструкцій ZKP можна знайти в роботі [3]. Питання практичної реалізації та ефективності ZKP, особливо в контексті блокчейн-технологій, активно досліджуються в роботі [4]. Останні досягнення в області неінтерактивних доказів з нульовим розголошенням (NIZK), таких як zk-SNARKs та zk-STARKS, детально розглядаються в [5]. Також існують роботи, присвячені вивченню потенційних загроз та обмежень ZKP, включаючи питання масштабованості та стійкості до квантових атак [6, 7].

Протоколи з нульовим розголошенням можуть бути класифіковані на кілька типів залежно від їхніх характеристик.

Інтерактивні та неінтерактивні докази

У інтерактивних протоколах верифікатор та доказувач обмінюються серією повідомлень для підтвердження істинності твердження. Класичним прикладом є протокол Фіата–Шаміра, що використовується для аутентифікації [8]. У такому протоколі верифікатор надсилає випадкові запитання доказувачу, і той має надати правильні відповіді, не розкриваючи секретної інформації.

Неінтерактивні протоколи (Non-Interactive Zero-Knowledge, NIZK) дозволяють доказувачу створити доказ, який може бути перевірений будь-ким без необхідності взаємодії з доказувачем. Перевага NIZK полягає в їхній зручності для практичного застосування, оскільки вони не вимагають постійної присутності обох сторін.

zk-SNARK та zk-STARK

Серед неінтерактивних протоколів особливе місце займають zk-SNARK (Zero-Knowledge Succinct Non-Interactive ARgument of Knowledge) та zk-STARK (Zero-Knowledge Scalable Transparent ARgument of Knowledge).

zk-SNARK є одним з найбільш використовуваних типів NIZK. "Succinct" означає, що розмір доказу є малим і час його перевірки також невеликий, незалежно від складності твердження. "ARgument of Knowledge" вказує на те, що доказувач не просто стверджує, що знає певну інформацію, а й надає криптографічний доказ цього знання. Для створення zk-SNARK часто потрібен початковий етап довіреної установки (Trusted Setup), в якому генеруються спільні параметри для доказувача та верифікатора. Недоліком цього етапу є потенційний ризик, якщо інформація, використана під час установки, потрапить до зловмисників.

zk-STARK є альтернативою zk-SNARK, яка не потребує довіреної установки, що робить їх більш прозорими та безпечними з точки зору криптографічних припущень. "Scalable" вказує на те, що час генерації та перевірки доказів зростає полілогарифмічно зі зростанням складності твердження, що робить їх більш придатними для великомасштабних застосувань. Однак розмір доказів zk-STARK часто більший, ніж у zk-SNARK.

Основні математичні засади

В основі протоколів з нульовим розголошенням лежать різні криптографічні примітиви та математичні концепції.

- Доведення знання дискретного логарифма. Це один з базових прикладів ZKP, де доказувач може переконати верифікатора, що йому відоме значення x таке, що

y = g^x mod p (де g є генератором групи, а p є простим числом), не розкриваючи при цьому самого значення x.

- Гомоморфне шифрування. Деякі протоколи ZKP використовують властивості гомоморфного шифрування, яке дозволяє виконувати певні операції над зашифрованими даними, не розшифровуючи їх. Результат такої операції при розшифруванні відповідає результату тієї ж операції, виконаної над відкритими даними.
- Поліноміальні зобов'язання (Polynomial Commitments). У zk-SNARK та zk-STARK часто використовуються поліноміальні зобов'язання, які дозволяють доказувачу зафіксувати поліном і потім надавати докази щодо значень цього полінома в певних точках, не розкриваючи сам поліном.
- Гешування. Криптографічні хеш-функції використовуються для створення стислих і незворотних представлень даних, що є важливим для забезпечення надійності протоколів.
- Еліптичні криві. У багатьох сучасних криптографічних системах, включаючи ZKP, використовуються еліптичні криві завдяки їхнім високим рівням безпеки при відносно невеликих розмірах ключів.

Приклади практичного застосування

Протоколи з нульовим розголошенням знаходять широке застосування в різних сферах, де потрібна конфіденційність та верифікація даних.

- Криптовалюти. Одним з найбільш відомих прикладів є криптовалюта Zcash, яка використовує технологію zk-SNARK для забезпечення повністю приватних транзакцій. Користувачі можуть здійснювати перекази, приховуючи інформацію про відправника, одержувача та суму транзакції, при цьому зберігаючи можливість перевірити легітимність транзакції. Проєкти на базі Ethereum, такі як Aztec та StarkWare, також активно використовують zk-SNARK та zk-STARK для підвищення приватності та масштабованості транзакцій [4].
- Автентифікація. ZKP можуть бути використані для створення безпечних та приватних систем автентифікації. Користувач може довести, що він володіє певним секретом (наприклад, паролем або біометричними даними), не розкриваючи сам секрет верифікатору. Це може бути використано для покращення безпеки онлайн-акаунтів та інших систем доступу.
- Цифрові ID. Концепція цифрових ідентифікаторів, що зберігають персональну інформацію користувачів, може значно виграти від використання ZKP. Користувач зможе надавати лише ту інформацію, яка необхідна для конкретної ситуації, не розкриваючи зайвих персональних даних. Наприклад, при оренді автомобіля можна підтвердити наявність водійського посвідчення та достатній вік, не показуючи повну копію документа.
- Голосування. Протоколи з нульовим розголошенням можуть бути використані для створення анонімних та верифікованих систем електронного голосування. Виборець може довести, що він має право голосувати, не розкриваючи свій вибір до завершення голосування.
- Доведення володіння секретами. ZKP можуть бути використані для доведення володіння певним секретом, наприклад приватним ключем у криптовалютній системі, без його розкриття.

Сучасні напрями досліджень

Постквантові докази з нульовим розголошенням. Одним з актуальних напрямів є розробка квантостійких ZKP-протоколів. Більшість класичних схем (наприклад, на основі дискретного логарифма чи факторизації) потенційно вразливі перед квантовими алгоритмами. Сучасні дослідження зосереджені на альтернативах, стійких до квантових атак, зокрема на протоколах, що покладаються лише на геш-функції або на складні задачі теорії ґраток. Прикладом є поява прозорих протоколів zk-STARK, які використовують лише криптографічні хеші і вважаються витривалими до квантових обчислень. Інший підхід – ґраткові ZKP: використовуються математичні задачі на ґратках, схожі до тих, що лягли в основу постквантових

підписів. Нещодавно було запропоновано нові методи побудови доказів знання на ґратках, що значно зменшують розмір і покращують продуктивність таких доказів [9]. Хоча поки що ґраткові ZKP поступаються класичним за ефективністю, прогрес у цій галузі триває – вже створено прототипи квантобезпечних схем, придатних для реалізації в протоколах приватності. Постквантовий напрям є критично важливим, адже забезпечить довгострокову безпеку ZKP в умовах появи квантових комп'ютерів.

Zero-Knowledge і машинне навчання (ZKML). Інша передова сфера – поєднання доказів з нульовим розголошенням з методами машинного навчання. ZKML – це новітня технологія, що дозволяє перевіряти результати виконання ML-моделей або певні властивості даних без розкриття самих моделей чи даних. Фактично, ZKP надають інструмент для доведення правильності роботи алгоритму штучного інтелекту, зберігаючи в таємниці і сам алгоритм, і використані для навчання дані. Це відкриває можливості для приватного та верифікованого AI: наприклад, вузол мережі може довести, що виконав Inference нейронної мережі без і закладених «бекдорів», не показуючи модель цілком. Останні дослідження демонструють значний прогрес у цій галузі – розробляються спеціальні фреймворки і бібліотеки, що оптимізують генерацію SNARK-доказів для моделей глибокого. Вже показано, що за допомогою таблиць пошуку й інших прийомів можна прискорити докази для нелінійних функцій (наприклад, ReLU, softmax та ін.) на порядки без втрати точності. Хоча ZKML все ще перебуває на ранніх етапах, він розглядається як ключовий компонент майбутніх *довірених AI*-систем, де користувачі зможуть перевіряти чесність моделей штучного інтелекту, не отримуючи доступу до їх внутрішніх параметрів.

Конфіденційність у Web3 та децентралізованих додатках. Zero-knowledge протоколи відіграють дедалі більшу роль у забезпеченні приватності та безпеки в екосистемі Web3. По-перше, вони лежать в основі самої ідеї *децентралізованої ідентичності* (Self-Sovereign Identity): користувач може мати цифровий ID і доводити окремі атрибути (вік, громадянство, наявність ліцензій тощо) без розкриття повного профілю. Це дозволяє будувати селективне розкриття даних – наприклад, підтвердити право доступу чи виконання регуляторних вимог, не передаючи особистої інформації контрагенту. По-друге, ZKP активно досліджуються для приватності смарт-контрактів і децентралізованих фінансових протоколів. Існують концепції «приватних DeFi», де транзакції та стан контрактів шифруються, а їхня коректність доводиться нульовим розголошенням. Таким чином, можна створювати біржі, кредитні платформи чи аукціони, що зберігають таємницю комерційних даних (балансів, заявок тощо), але залишаються перевіряльними. По-третє, ZKP пропонуються як рішення для регуляторних задач у блокчейні. Зокрема, концепція *zkKYC* передбачає, що користувач проходить перевірку KYC/AML у довіреного провайдера і отримує криптографічне доказ цього, яке може пред'являти в смарт-контрактах. При цьому самі персональні дані (ім'я, документи) не розкриваються блокчейну, але контракт упевнений, що користувач дотримується вимог законодавства. Подібні рішення покликані знайти баланс між анонімністю Web3 та необхідністю відповідати нормативним правилам. Загалом, напрям приватності у Web3 охоплює широкий спектр досліджень – від анонімних облікових записів і соціальних мереж до захищеного обміну даними між організаціями на блокчейні – і Zero-Knowledge докази є ключовою технологією для втілення бачення «приватного децентралізованого Інтернету».

Переваги й виклики

Використання протоколів з нульовим розголошенням має значні переваги, але також пов'язане з певними викликами.

Переваги:

- Підвищення приватності. Основною перевагою ZKP є можливість довести істинність твердження, не розкриваючи жодної зайвої інформації. Це є критично важливим для багатьох застосувань, де конфіденційність даних є пріоритетом.

• Зниження потреби в довірі. У деяких випадках ZKP можуть зменшити або усунути потребу в довірених посередниках. Верифікатор може бути впевнений у правдивості твердження лише на основі наданого доказу.

• Покращення безпеки. Завдяки можливості підтверджувати володіння секретами без їхнього розкриття, ZKP можуть підвищити безпеку різних систем автентифікації та контролю доступу.

Виклики:

• Масштабованість. Генерація та перевірка доказів з нульовим розголошенням може бути обчислювально інтенсивною, особливо для складних тверджень. Покращення масштабованості є однією з ключових проблем у цій галузі.

• Складність реалізації. Розробка та впровадження протоколів з нульовим розголошенням є складним завданням, що вимагає глибоких знань в області криптографії та математики.

• Довірена установка (для zk-SNARK). Необхідність довіреної установки в zk-SNARK створює потенційну вразливість, якщо параметри установки будуть скомпрометовані. Розробка альтернатив, таких як zk-STARK, спрямована на вирішення цієї проблеми.

• Постквантові аспекти. Як і багато сучасних криптографічних систем, деякі ZKP можуть бути вразливі до атак з використанням квантових комп'ютерів. Розробка квантовостійких протоколів з нульовим розголошенням є важливим напрямком досліджень.

Висновки

Протоколи з нульовим розголошенням є потужним інструментом у сучасній криптографії, що відкриває нові можливості для забезпечення приватності, безпеки та достовірності інформації в цифровому світі. Від забезпечення анонімності транзакцій у криптовалютах до створення безпечних систем цифрової ідентифікації, ZKP демонструють свій значний потенціал у різних галузях. Незважаючи на існуючі виклики, такі як питання масштабованості та складності реалізації, активні дослідження та розробки в цій області продовжують просувати технологію вперед. З появою більш ефективних та безпечних протоколів, таких як zk-STARK, можна очікувати ще ширшого застосування протоколів з нульовим розголошенням у майбутніх криптографічних системах.

Список літератури:

1. Goldwasser S., Micali S., Rackoff C. The knowledge complexity of interactive proof systems // SIAM Journal on Computing. 1989. Vol. 18, No. 1. P. 186–208.
2. Goldwasser S., Micali S., Wigderson A. Proofs that yield nothing but their validity and a methodology of cryptographic protocol design // 27th Annual Symposium on Foundations of Computer Science. IEEE, 1986. P. 174–187.
3. Camenisch J., Stadler M. Proof systems for general statements in discrete logarithms. ETH Zurich, 2003. [Електронний ресурс]. Режим доступу: https://crypto.ethz.ch/publications/files/CamSta97b.pdf
4. Ben-Sasson E., Bentov I., Horesh Y., Riabzev M. Scalable, transparent, and post-quantum secure computational integrity // IACR Cryptology ePrint Archive. 2018. No. 046. [Електронний ресурс]. Режим доступу: https://eprint.iacr.org/2018/046.pdf
5. Groth J. Short pairing-based non-interactive zero-knowledge arguments // Advances in Cryptology – ASIACRYPT 2010. Lecture Notes in Computer Science, Vol. 6477. Springer, 2010. P. 321–340. [Електронний ресурс]. Режим доступу: https://www.iacr.org/archive/asiacrypt2010/6477323/6477323.pdf
6. Yu R., Liu J. K. A survey of zero-knowledge proof systems // Journal of Computer Science and Technology. 2021. Vol. 36, No. 4. P. 705–727.
7. Mosca M. Cybersecurity in an Era with Quantum Computers: Will We Be Ready? // IEEE Security & Privacy. 2018. Vol. 16. P. 38–41.
8. Fiat A., Shamir A. How to prove yourself: Practical solutions to identification and signature problems // Advances in Cryptology – CRYPTO'86. Lecture Notes in Computer Science. Vol. 263. Springer, 1987. P. 186–194. [Електронний ресурс]. Режим доступу: https://mit6875.github.io/PAPERS/Fiat-Shamir.pdf
9. Zero knowledge proofs IBM research project [Електронний ресурс]. Режим доступу: https://research.ibm.com/projects/zero-knowledge-proofs

Надійшла до редколегії 20.05.2025

Відомості про автора:
Мордвінов Руслан Ігорович – кандидат технічних наук, Україна; email: rmordvinov@gmail.com; ORCID: https://orcid.org/0000-0003-1229-2840

УДК 004.056 **DOI:10.30837/rt.2025.3.222.06**

Д.М. МОРГУЛЬ, О.П. НАРЄЖНІЙ, канд. техн. наук, Т.О. ГРІНЕНКО, канд. техн. наук

РОЗРОБКА ТИПОВОЇ ІНФРАСТРУКТУРИ ДЛЯ ВЕБ-СЕРВІСУ КВАНТОВОГО ГЕНЕРАТОРА ВИПАДКОВИХ ЧИСЕЛ

Вступ

Квантові генератори випадкових чисел (QRNG, quantum random number generator) є критично важливими для забезпечення високої ентропії у криптографічних системах, моделюванні та наукових дослідженнях. Локальні рішення гарантують максимальну безпеку, але не завжди доступні через вартість і технічні обмеження. Публічні веб-сервіси QRNG, що надають доступ до квантової випадковості через інтерфейс прикладного програмування (API, application programming interface), дозволяють швидко інтегрувати її у прототипи, навчальні проєкти та дослідницькі системи. Водночас відсутність уніфікованого підходу до інфраструктури таких сервісів ускладнює їх порівняння, інтеграцію та аудит.

Сучасні веб-сервіси ґрунтуються на високоякісній випадковості для криптографічних протоколів, керування сесіями, токенів доступу, seed-значень і моделювання. Криптографічно стійкі детерміновані генератори випадкових бітів (DRBG, deterministic random bit generator) критично залежать від секретності початкових станів і коректності реалізації, тоді як істинно випадкові генератори (TRNG, true random number generator) можуть зазнавати систематичних похибок або деградації джерела ентропії. QRNG, які відносяться до TRNG, вирізняються тим, що спираються на фундаментальну невизначеність вимірювань у квантовій механіці, забезпечуючи принципово непередбачувану ентропію за умови коректної побудови вимірювальної схеми та постобробки [1, 2]. Питання інтеграції QRNG як мережевого сервісу (веб-сервісу QRNG) є актуальною задачею масштабування довіри до випадковості для широкого кола прикладних напрямків – від критичної інфраструктури держави до кібербезпеки хмарних платформ.

Перехід до формату веб-сервісу породжує нові ризики і вимоги [3]. По-перше, джерело ентропії має бути фізично надійним та валідованим відповідно до процедур оцінювання min-ентропії відповідно до рекомендацій NIST SP 800-90B щодо створення моделі джерела ентропії, контролю якості бітів і виявлення відмов [4]. По-друге, результати екстракції повинні мати гарантії наближення до рівномірного розподілу. На практиці для QRNG часто використовують екстрактори на основі Toeplitz-гешування та пов’язане з ними доведення через LHL (Leftover Hash Lemma) [5]. Наприклад, для реалізації високих сервісних навантажень використовується апаратна реалізація типу FPGA-based Toeplitz Strong Extractor для QRNG [6]. По-третє, навіть за правильної екстракції потрібно здійснювати health-тести (continuous tests) вихідних потоків джерела ентропії за допомогою рекомендацій NIST SP 800-90B [4].

Мета статті – розробити та обґрунтувати типову інфраструктуру веб-сервісу QRNG, яка включатиме функціональні компоненти, вимоги до безпеки, інтерфейси доступу (API), методи контролю якості випадковості та рекомендації щодо масштабування. Запропонована інфраструктура має слугувати основою для створення сумісних, безпечних і продуктивних веб-сервісів джерел квантової ентропії.

1. Вимоги до типової інфраструктури публічного веб-сервісу QRNG

Із погляду безпеки потоку даних веб-сервісу QRNG потрібно використовувати TLS 1.3 з мінімізацією поверхні атаки [7]. Для цифрової ідентифікації користувача на веб-сервісі QRNG пропонується керуватися вимогами NIST SP 800-63B-4 [8] щодо життєвого циклу автентифікаторів і застосовувати біометрію у складі багатофакторної автентифікації (MFA, multi-factor authentication). Для авторизації та делегування доступу пропонується використовувати протокол OAuth 2.0 з короткоживучими токенами та форматом JWT (JSON Web Token) для перенесення підписаних повноважень [9, 10]. На рівні периметра

інфраструктура має забезпечувати відокремлення зон (демілітаризовані зони (DMZ, demilitarized zone), підмережі для бекендів і сховищ), передбачувану поведінку NAT (Network Address Translation) і Internet-gateway та контрольований вихідний трафік через керований вихідний проксі згідно з відповідними RFC-вимогами [11 – 14].

Теоретичні моделі довіри до QRNG також еволюціонують. Поряд із класичною (повністю довіреною) моделлю популярності набувають незалежний від джерела (SI, source-independent) та незалежний від пристрою (DI, device-independent) підходи, де випадковість може бути сертифікована експериментально через порушення нерівностей Белла навіть за обмеженої довіри до елементів апаратури [15 – 18]. У роботі [19] демонструється суттєвий прогрес у швидкостях DI-генерації з одночасною сертифікацією квантової випадковості, що поступово знімає бар'єри до використання таких схем у сервісних сценаріях. У дослідженні розглянуто SI/DI як цільові напрями розвитку QRNG і показано, як закласти для них «каркас» у типовій інфраструктурі без шкоди для продуктивності поточної (класичної) реалізації [19].

Узагальнюючи, дослідницька задача цієї роботи полягає у побудові типової інфраструктури веб-сервісу QRNG, яка задовольняє наступним вимогам:

1. Гарантує обґрунтовану якість випадковості через моделювання джерела, коректну екстракцію та статистичну валідацію [4 – 6, 20].

2. Забезпечує сучасний захист комунікації, ідентифікацію та делегування доступу (TLS 1.3; NIST SP 800-63B-4; OAuth 2.0/JWT) [7 – 10].

3. Мінімізує периметрові ризики і контролює вихідні потоки (NAT, Internet-gateway, outbound-proxy) [11 – 14].

4. Має чітку дорожню карту постквантової міграції відповідно до FIPS 203, FIPS 204, FIPS 205 і супровідних рекомендацій NIST [17, 21 – 24].

5. Допускає поступову інтеграцію SI/DI-сертифікації без радикальної перебудови всієї системи [5, 15 – 19].

2. Функціональна схема типової інфраструктури веб-сервісу QRNG та опис основних елементів

На рис. 1 наведена запропонована функціональна схема типової інфраструктури веб-сервісу QRNG.

Опис основних елементів функціональної схеми типової інфраструктури веб-сервісу QRNG та їх функції:

1. Internet-gateway і 15. NAT-gateway.

Internet-gateway забезпечує керований вихід/вхід трафіку між локальною мережею і мережею інтернет. NAT-gateway транслює приватні адреси, маскуючи внутрішній простір і дозволяючи масштабовану вихідну комунікацію. Поведінкові вимоги до NAT для UDP/TCP формалізовані в RFC 4787 та RFC 5382, а базова термінологія в RFC 2663; оновлення вимог наведено в RFC 7857 [11 – 13,25]. Для веб-сервісу QRNG це критично, оскільки згенерована на бекенді телеметрія, оновлення прошивки екстракторів мають виходити предиктивно, без витоку внутрішньої адресації.

2. Локальна мережа (зонування) (Local Network).

Сегментація на підмережі (DMZ для веб-сервісу; окремі VLAN/VPC-підмережі для QRNG-бекенду та сховищ) мінімізує бічний рух у разі компрометації. Завершення TLS розміщується або на edge-проксі, або безпосередньо у веб-сервісі з примусовою підтримкою TLS 1.3.

3. Модуль автентифікації (біометрія як фактор) (Authentication module (Biometric)).

Користувацька автентифікація (оператори, інтегратори, адміністратори веб-сервісу QRNG) реалізується відповідно до NIST SP 800-63B-4 (Authentication Assurance Level 2 та Level 3) [8]. Біометричні фактори можуть використовуватися у складі багатофакторних схем, але потребують управління життєвим циклом автентифікаторів, захисту шаблонів і валідації

придатності каналів (вимоги NIST SP 800-63B-4 щодо Authentication Assurance Level, Presentation Attack Detection (PAD) – загроз тощо).

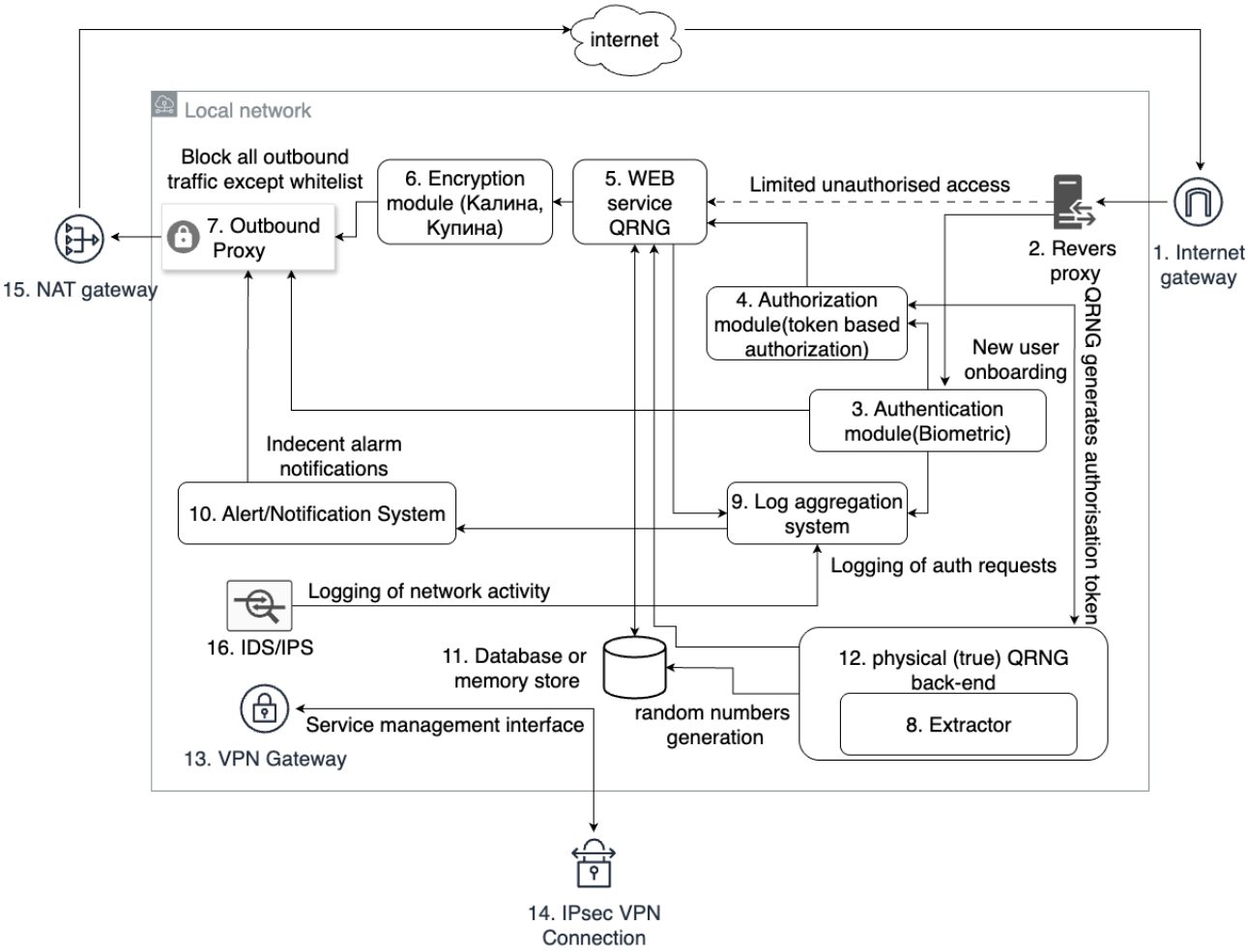


Рис. 1. Функціональна схема типової інфраструктури веб-сервісу QRNG

4. Модуль авторизації (токен-базована авторизація) (Authorization module (token based authorization)).

Внутрішній Authorization Server реалізує протокол OAuth 2.0 (RFC 6749) з видачею короткоживучих токенів доступу. Ресурси приймають токени у форматі JWT (RFC 7519) з відповідною політикою електронних підписів (RSASSA-PSS та ECDSA або постквантові варіанти «гібридних» токенів у перехідний період) [9, 10]. Для машинних інтеграцій рекомендується використання сервісних аккаунтів із обмеженим рівнем доступу, для користувачів – авторизаційний код відповідно протоколу OAuth 2.0 (PKCE, Proof Key for Code Exchange).

5. Веб-сервіс QRNG (API/портал) (WEB service QRNG).

Функціональність: запит на отримання бітів, параметри постобробки (екстракції), довідкова телеметрія (обчислення min-ентропії, температурні сенсори, оптичні сенсори), опціональний «віртуальний пул» з буферизацією. Для міжсервісної взаємодії обов'язкове використання TLS 1.3 з взаємною автентифікацією [7]. Підрахунок запитів контролюється через обмеження швидкості і квоти; аудиторські події фіксуються в журналі.

6. Модуль шифрування (дані під час передачі/дані на носіях) (Encryption module (Калина, Купина)).

Під час передачі даних використовується TLS 1.3. Дані на носіях шифруються за допомогою апаратних модулів HSM (Hardware Security Module) або KMS (Key Management Service). В умовах Української юрисдикції необхідно використовувати національні стандарти (наприклад, блоковий симетричний шифр «Калина» та геш-функція «Купина»).

7. Вихідний проксі (Outbound Proxy).

Всі вихідні HTTP(S)-потоки веб-сервісу QRNG (оновлення прошивок, залежності тощо) прямують через керований проксі з TLS-інспекцією, а також з обмеженням цільових доменів/підмереж. Таке обмеження сприяє інформаційному контролю потоків і відповідає загальним практикам захисту мережі, доповнюючи NAT-рівень.

8. Буфер постобробки/екстракції випадковості (Extractor).

Raw-біти QRNG піддаються оцінюванню min-ентропії відповідно до NIST SP 800-90B та екстракції для отримання рівномірного розподілу [4]. Інформаційно-теоретично доведена екстракція реалізується Toeplitz-гешуванням LHL, яке широко використовується в QRNG-реалізаціях і може бути апаратно прискорене (FPGA/SoC) [5, 26]. Оперативна перевірка вихідних даних після екстракції проводиться за допомогою статистичних тестів NIST SP 800-22.

9. Система агрегації логів (Log aggregation system).

Це програмне рішення, яке збирає, фільтрує, трансформує та об'єднує журнали (логі) з різних джерел (серверів, додатків, мережевих пристроїв) в єдину систему для спрощення аналізу та моніторингу. Журнали автентифікації, авторизації, запити API, телеметрія QRNG, результати тестів випадковості та адміністративні дії централізовано збираються і зберігаються згідно з встановленим періодом зберігання та Hot/Worm-політиками. Це забезпечує відтворюваність і ланцюжок довіри для аудитів (відповідність вимогам до обліку доступу, реагування на інциденти та розслідуваня).

10. Система оповіщення та нотифікацій (Alert/Notification System).

Це механізм для миттєвого інформування адміністраторів веб-сервісу про події (тригери), новини, оновлення чи критичні стани компонентів системи. Тригери на основі правил/кореляції (зниження min-ентропії нижче порогів SP 800-90B; аномальний rate API; невдачі TLS-перевірок; відхилення параметрів діода/лазера) сповіщають адміністраторів. Сповіщення надсилаються через надійні канали (залежно від політики ризику – через зашифровані канали або внутрішній мессенджер).

11. Сховище метаданих/«пулу» бітів (Database or memory store).

Пул зберігає попередньо екстраговані біти з криптографічним контролем цілісності та мітками часу/джерела. Політика «не повторювати» забезпечується індексуванням блоків/посиланням на «вичерпані» сегменти. Для зовнішніх клієнтів краще надавати потоки «на вимогу» без довготривалого зберігання.

12. Фізичний (істинний) QRNG-бекенд (physical (true) QRNG back-end).

Оптичні, фазові, вакуумні флуктуації чи інші квантові джерела ентропії генерують raw-біти, при цьому моніториться стан середовища джерела ентропії. Для доведеної безпекової моделі розглядаються підходи SI і DI, де випадковість може бути сертифікована навіть за обмеженої довіри до компонентів.

13. VPN-шлюз (VPN Gateway) та 14. IPsec VPN-з’єднання (IPsec VPN Connection).

Операційний доступ інженерів, приватні інтеграції з довіреними партнерами та кластерний менеджмент можуть додатково ізолюватися через IPsec-тунелі між визначеними сегментами мереж.

16. Система виявлення і запобігання вторгненням (IDS/IPS).

IDS/IPS є окремим шаром контролю, що моніторить мережевий трафік і події елементів інфраструктури з метою виявлення зловмисної активності та негайного блокування. IDS/IPS охоплює широкий простір протоколів і поведінкових патернів таких як сканування, «бічний рух» (Lateral Movement), аномалії у TLS-рукостисканнях тощо. Понятійна база наведена у NIST SP 800-94 (IDPS) [27] і спосіб контролю SI-4 (System Monitoring) у NIST SP 800-53 Rev.5 [28], які рекомендують як мережеві, так і хостові сенсори з централізованою кореляцією подій.

3. Обґрунтування ефективності типової інфраструктури веб-сервісу QRNG

Квантова випадковість і довірена постобробка. Багаторічні огляди і сучасні результати підтверджують, що QRNG на основі квантових процесів (детекція фотонів, фазові флуктуації тощо) забезпечують фундаментально непередбачувану випадковість, якщо вимірювальна система належно модельована і контрольована [1, 2]. Для зменшення впливу класичного шуму та систематик використовують теоретично доведені екстрактори: Toeplitz-гешування разом із LHL гарантує, що вихідні дані мають розподіл близький до рівномірного за наявності достатньої min-ентропії [5, 26]. Використання підходів SI і DI гарантує, що випадковість може бути сертифікована експериментально, але це знизить вимоги довіри до апаратури [15 – 18]. Останні експерименти демонструють мегабітні DI-швидкості з одночасною сертифікацією квантової випадковості, що робить їх більш практичними для сервісних сценаріїв [19].

Оцінка джерела ентропії повинна виконуватися згідно з NIST SP 800-90B (моделювання джерела, збір raw-даних, оцінювання min-ентропії, виявлення збоїв) із подальшим статистичним тестуванням потоків (SP 800-22 Rev.1a) як додатковим санітарним бар'єром [4, 20]. Послідовність тестування NIST SP 800-90B → екстракція → NIST SP 800-22 є де-факто стандартним безпековим підходом щодо оцінки джерела квантової ентропії [4, 5, 20].

Захищений транспортний потік та ідентифікація. TLS 1.3 лімітує використання застарілих шифрів, забезпечуючи захищений канал між клієнтом і веб-сервісом QRNG [7]. Це особливо важливо під час передачі токенів доступу та потоків випадкових бітів. Політика ідентифікації відповідає NIST SP 800-63B-4, а саме: біометрія застосовується як частина MFA (Authentication Assurance Level 2 та Level 3), автентифікатори керуються протягом усього життєвого циклу, згодом відкликаються і ревалідовуються. Це підтримується технологічно через Fast Identity Online-сумісні автентифікатори або апаратні ключі, однак, саме NIST SP 800-63B-4 визначає вимоги щодо рівнів запевнення (Assurance Level) і політик відновлення.

Авторизація та делегування доступу. Авторизація та делегування доступу забезпечується протоколом OAuth 2.0. Це забезпечує стандартизоване делегування з різними потоками (авторизаційні коди разом з PKCE для інтерактивних клієнтів; клієнтські повноваження для server-to-server), а JWT надає компактний формат перенесення прав із електронним підписом та шифруванням [9, 10]. Для веб-сервісу QRNG важлива короткоживучість токенів, вузька сфера застосування, чітке обмеження швидкості на ключ або клієнта та аудит усіх дій.

Постквантова готовність. У постквантовий період на веб-сервісі QRNG необхідно використовувати стандарти типу FIPS 203, FIPS 204, FIPS 205 і супровідні матеріали NIST [17, 21 – 24]. Наприклад, ML-KEM для обміну ключами, а ML-DSA/SLH-DSA для електронних підписів.

Периметр і керування вихідним трафіком. NAT/Internet-gateway разом із outbound-proxy обмежують поверхню атак і створюють спостережний контроль за зовнішніми залежностями. Стандарти RFC 4787, RFC 5382, RFC 2663, RFC 7857 регламентують поведінку NAT, що забезпечує передбачуваність з'єднань.

Спостережність і відповідальність. Централізоване журналювання з кваліфікованою електронною позначкою часу (NTP autokey), що дозволяє відстежити інциденти: падіння min-ентропії, аномалії запитів, збої TLS, спроби перевищення квот. Це критично для відповідності вимогам кібербезпеки і для незалежної перевірки веб-сервісів QRNG.

Загрози і їх пом'якшення. Вихід з ладу джерела квантової ентропії: постійний моніторинг min-ентропії згідно з NIST SP 800-90B. Атака Man-in-the-Middle: використання TLS 1.3 разом з mTLS для міжсервісних зв'язків; суворий контроль SSL сертифікатів. Компрометація токенів: короткі життєві цикли, обмежені сфери застосування, ротація ключів електронного підпису JWT, аудит випуску токенів. Зловживання API: квотування, обмеження швидкості, розгалуження доступів відповідно бізнес моделі, відсікання аномалій. Апаратні збої: резервування джерела квантової ентропії; контроль допусків сенсорів; аварійний канал DRBG з

чітким маркуванням походження бітів (і забороною «змішування» без явної політики). Периметрові ризики: відмова від відкритих вихідних маршрутів, зазначені можливі маршрути через outbound-proxy, конфігурація NAT згідно з RFC 4787, RFC 6888, RFC 7857.

Висновки

Представлена інфраструктура веб-сервісу QRNG є типовою і має адаптуватися до галузевих та регуляторних вимог, зокрема, національної нормативно-правової бази в галузі кібербезпеки. Частина джерел про DI-генерацію залишається експериментальною, їх безшовна інтеграція у високонавантажені веб-сервіси QRNG потребує додаткових інженерних досліджень продуктивності й відмовостійкості. Оцінка економічного ефекту впровадження типової інфраструктури веб-сервісу QRNG у постквантовий період залежить від профілю клієнтів і життєвого циклу даних.

Напрями майбутніх робіт: по-перше, перспективним є створення відкритих апаратних (FPGA/SoC) еталонних наборів телеметрії QRNG для поточної оцінки джерела квантової ентропії згідно з рекомендаціями NIST SP-800-90B; по-друге, формалізація цілі рівня обслуговування та угоди про рівень надання послуг для веб-сервісів джерел квантової ентропії. Наприклад, гарантований min-ентропійний бюджет на запит та прозорість журналів аудиту. По-третє, розробка гібридних протоколів TLS з постквантовим рукостисканням, оптимізованих для навантажених REST API. Наприклад, використання апаратних прискорювачів екстракції (FPGA/SoC) з доведеними властивостями та відкритими описами проєктів.

Таким чином, запропонована інфраструктура веб-сервісу QRNG забезпечує збалансований компроміс між науковою обґрунтованістю, інженерною практичністю та регуляторною відповідністю. Вона дає організаціям чітку дорожню карту: від відомих прототипів QRNG з формальною екстракцією і сучасним стеком безпеки – до поступового впровадження SI/DI-сертифікації та повної постквантової готовності. Такий підхід дозволяє підвищити рівень довіри до веб-постачання випадкових бітів з джерел квантової ентропії і зробити цей ресурс надійним фундаментом для широкого спектра криптографічних та прикладних задач у майбутньому.

Список літератури:

1. Ma X., Yuan X., Cao Z., Qi B., and Zhang Z. Quantum random number generation // Quantum Information. 2016. Vol. 2. Available: https://www.nature.com/articles/npjqi201621 (Nature)
2. Mannalath V., Mishra S., Pathak A. A Comprehensive Review of Quantum Random Number Generators. arXiv:2203.00261, 2022. Available: https://arxiv.org/abs/2203.00261 (arXiv)
3. Morhul D., Nariezhnii O., & Hrinenko T. Threat and adversary models for QRNG web services // Radiotekhnika. 2025. No 221. P. 31–38. https://doi.org/10.30837/rt.2025.2.221.04
4. Turan M. S. et al. Recommendation for the Entropy Sources Used for Random Bit Generation, NIST SP 800-90B (Final), 2018. DOI: 10.6028/NIST.SP.800-90B. PDF: https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-90b.pdf (NIST Publications)
5. Ma X. et al. Postprocessing for quantum random-number generators: entropy evaluation and randomness extraction // Phys. Rev. A. 2013. Vol. 87, 062327. DOI: 10.1103/PhysRevA.87.062327 (Physical Review)
6. Chouhan S. et al. FPGA-based Toeplitz Strong Extractor for Quantum Random Number Generators. arXiv:2505.02868, 2025. Available: https://arxiv.org/abs/2505.02868 (arXiv)
7. Rescorla E. The Transport Layer Security (TLS) Protocol Version 1.3. RFC 8446, Aug. 2018. Available: https://datatracker.ietf.org/doc/html/rfc8446 (IETF Datatracker)
8. Temoshok D. et al. Digital Identity Guidelines: Authentication and Authenticator Management NIST SP 800-63B-4, 2025. PDF: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-63b-4.pdf (NIST Publications)
9. Hardt D. The OAuth 2.0 Authorization Framework // RFC 6749, Oct. 2012. Available: https://datatracker.ietf.org/doc/html/rfc6749 (IETF Datatracker)
10. Jones M., Bradley J., and Sakimura N. JSON Web Token (JWT) // RFC 7519, May 2015. Available: https://datatracker.ietf.org/doc/html/rfc7519 (IETF Datatracker)
11. Audet F. and Jennings C. Network Address Translation (NAT) Behavioral Requirements for Unicast UDP // RFC 4787, Jan. 2007. Available: https://datatracker.ietf.org/doc/html/rfc4787 (IETF Datatracker)
12. Guha S. et al. NAT Behavioral Requirements for TCP // RFC 5382, Oct. 2008. Available: https://datatracker.ietf.org/doc/html/rfc5382 (IETF Datatracker)

13. Srisuresh P. and Holdrege M. IP Network Address Translator (NAT) Terminology and Considerations // RFC 2663, Aug. 1999. Available: https://datatracker.ietf.org/doc/html/rfc2663 (IETF Datatracker)
14. Penno R. et al. Updates to NAT Behavioral Requirements // RFC 7857, Apr. 2016. Available: https://www.rfc-editor.org/rfc/rfc7857.html (RFC Editor)
15. Cao Z. et al. Source-Independent Quantum Random Number Generation // Phys. Rev. X. Vol. 6, 011020, 2016. DOI: 10.1103/PhysRevX.6.011020 (Physical Review)
16. Wang C. et al. Provably-secure quantum randomness expansion with uncharacterized measurement devices // Nature Communications, 2023. DOI: 10.1038/s41467-022-35556-z (Nature)
17. Bamps C., Massar S., Pironio S. Device-independent randomness generation with low-power states // Quantum, 2018; Pironio et al., Nature, 2010. Available: https://quantum-journal.org/papers/q-2018-08-22-86/ (quantum-journal.org) DOI:10.22331/q-2018-08-22-86
18. Joch D. et al. Certified random-number generation from quantum steering // Phys. Rev. A. Vol. 106, L050401, 2022. DOI: 10.1103/PhysRevA.106.L050401 (Physical Review)
19. Kumar A. Nai, Kumar V. Device-independent, megabit-rate quantum random number generation with live quantumness certification. arXiv:2412.18285, 2024. Available: https://arxiv.org/abs/2412.18285 (arXiv)
20. Rukhin A. et al. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications, NIST SP 800-22 Rev. 1a, 2010. PDF: https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf (NIST Publications)
21. NIST, FIPS 203: Module-Lattice-Based Key-Encapsulation Mechanism (ML-KEM), 2024. Web/PDF: https://csrc.nist.gov/pubs/fips/203/final; https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.203.pdf (NIST Computer Security Resource Center, NIST Publications) DOI: 10.6028/NIST.FIPS.203
22. NIST, FIPS 204: Module-Lattice-Based Digital Signature Standard (ML-DSA), 2024. Web/PDF: https://csrc.nist.gov/pubs/fips/204/final; https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.204.pdf (NIST Computer Security Resource Center, NIST Publications) DOI:10.6028/NIST.FIPS.204
23. NIST, FIPS 205: Stateless Hash-Based Digital Signature Standard (SLH-DSA), 2024. Web/PDF: https://csrc.nist.gov/pubs/fips/205/final; https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.205.pdf (NIST Computer Security Resource Center, NIST Publications) DOI:10.6028/NIST.FIPS.205
24. NIST News, “NIST releases first 3 finalized post-quantum encryption standards,” Aug. 13, 2024. Available: https://www.nist.gov/news-events/news/2024/08/nist-releases-first-3-finalized-post-quantum-encryption-standards (NIST)
25. IETF, “Best Current Practice 127: Network Address Translation (NAT) Behavioral Requirements,” comprising RFC 4787 (UDP), RFC 6888 (CGN) and RFC 7857 (Updates), 2007–2016. Available: IETF Datatracker (BCP 127 info page https://datatracker.ietf.org/doc/bcp127/).
26. Zhang X.-G., Nie Y.-Q., Zhou H., Liang H., Ma X., Zhang J., and Pan J.-W. Note: Fully integrated 3.2 Gbps quantum random number generator with real-time extraction // Review of Scientific Instruments. 2016. Vol. 87, no. 7. P. 076102. DOI:10.1063/1.4958663. (Preprint: arXiv:1606.09344).
27. Scarfone K. and Mell P. Guide to Intrusion Detection and Prevention Systems (IDPS), NIST SP 800-94, 2007. DOI:10.6028/NIST.SP.800-94
28. Joint Task Force. Security and Privacy Controls for Information Systems and Organizations, NIST SP 800-53 Rev.5, 2020. DOI:10.6028/NIST.SP.800-53r5

Надійшла до редколегії 17.05.2025

Відомості про авторів:
Моргуль Дмитро Миколайович – Харківський національний університет імені В. Н. Каразіна, аспірант кафедри кібербезпеки інформаційних систем, мереж і технологій, Україна; e-mail: dmitriymdn85@gmail.com; ORCID: https://orcid.org/0009-0007-5272-1634
Нарєжній Олексій Павлович – канд. техн. наук, Харківський національний університет імені В. Н. Каразіна, доцент кафедри кібербезпеки інформаційних систем, мереж і технологій, Україна; e-mail: o.nariezhnii@karazin.ua; ORCID: https://orcid.org/0000-0003-4321-0510
Гріненко Тетяна Олексіївна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій, Україна; e-mail: tetiana.grinenko@nure.ua; ORCID: https://orcid.org/0000-0002-8251-8991

УДК 004.056.5:005.8 **DOI:10.30837/rt.2025.3.222.07**

Т.І. КОРОБЕЙНІКОВА, канд. техн. наук, А.Б. ЯМНИЧ

ПРОЦЕСНА МОДЕЛЬ ДИНАМІЧНОГО АНАЛІЗУ ТА ПРОГНОЗУВАННЯ РИЗИКІВ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ ДЛЯ ПЕРСОНАЛУ

Вступ

Цифровізація бізнес-процесів, гібридні моделі роботи та зростання залежності від розподілених сервісів суттєво підвищили вагу людського фактора та персоналу зокрема в кіберзагрозах [1 – 3]. Традиційні періодичні процедури оцінювання ризиків та статичні політики доступу не встигають за динамікою середовища, де контекст доступу, рівень навантаження, поведінкові патерни користувачів і загальна загрозова картина змінюються у реальному (майже) часі [4 – 6]. Саме тому організаціям потрібні моделі, що поєднують безперервний моніторинг, динамічну оцінку ризику і негайне застосування контрзаходів, узгоджених з принципами Zero Trust [7, 8] та вимогами аудитованості й прозорості. Нормативні орієнтири (зокрема NIST SP 800-207) [9, 10] прямо підкреслюють необхідність відмови від «імпліцитної довіри» і переходу до рішень, що щоразу верифікують суб'єкта, пристрій і запитуваний ресурс з урахуванням контексту, ризику та політик організації [11, 12].

Запропонована у поданому тексті процесна модель акумулює ці потреби й тренди в єдину архітектуру: 1) формує багатовимірну матрицю класифікації ресурсів; 2) збирає технічні та поведінкові дані; 3) нормалізує їх у вектор ознак Q; 4) будує цифровий двійник користувача для прогнозування ризиків; 5) генерує адаптивні контрзаходи та навчальний контент; 6) доставляє персоналізовані політики доступу; 7) збирає результати впровадження; 8) самонавчається, коригуючи ваги, моделі та правила RBAC-блокчейну. Така інтеграція безпосередньо задовольняє актуальні запити ринку: оперативність, персоніфікацію, прозорість аудиту та стале підвищення кіберстійкості.

Метою роботи є підвищення точності та оперативності оцінювання ризиків інформаційної безпеки для персоналу, покращення керованості доступу через адаптивні політики на основі поведінкової аналітики та цифрових двійників, а також вдосконалення прозорості й відтворюваності аудиту за рахунок інтеграції RBAC-блокчейну в замкнений цикл самонавчання системи.

1. Аналіз останніх джерел в галузі ІБ в контексті діяльності компанії

Наукові огляди динамічного оцінювання ризику (Dynamic Risk Assessment, DRA) показують: статичні підходи не забезпечують потрібного рівня оперативності й точності у сучасних мережевих і хмарних середовищах; натомість моделі DRA інтегрують потоки подій (IDS/SIEM), контекстні атрибути та машинне навчання для безперервного перерахунку ризикових показників і суттєво підсилюють прийняття рішень щодо доступу й реагування [13].

Паралельно еволюціонують інструменти поведінкової аналітики (UEBA), що будують профілі користувачів і виявляють аномалії, зокрема для інсайдерських загроз – критичної категорії ризиків у корпоративному середовищі. Дослідження демонструють дієвість UEBA для безперервної автентифікації та виявлення девіацій у діях працівників і сервісів [14]. У промислових і кіберфізичних системах, де наслідки інсайдерської активності особливо дороговартісні, огляди підтверджують потребу в системних підходах до детекції, що поєднують технічні та поведінкові ознаки [15].

Ще один актуальний трендом є використання «цифрових двійників» (Digital Twins) для моделювання поведінки суб'єктів/систем і прогнозування небажаних сценаріїв. Сучасні огляди фіксують зрілість підходу та його релевантність саме до задач безпеки: симуляція сценаріїв, оцінювання вразливостей, прогнозування ризикових подій і тестування політик до їхнього застосування у кінцевих продуктах [16].

Нарешті, для вимог незмінності журналів доступу, відтворюваності аудиту й довіри між підсистемами зростає роль блокчейн-парадигм у контролі доступу (RBAC/ABAC із реєстра-

цією транзакцій доступу в розподіленому реєстрі). Систематичні огляди підтверджують потенціал блокчейну підвищувати довіру, прослідковуваність і автоматизацію політик, включно зі смарт-контрактами та ризик-адаптивними правилами [17 – 20].

Динамічне оцінювання ризику (DRA). Систематичний огляд Cheimonidis & Rantos (2023) узагальнює 50 моделей DRA і показує, що в сучасних умовах статичні процедури не забезпечують потрібної гнучкості [3]; провідні підходи використовують потоки подій (IDS), вразливості (NVD) і ML/BN для перерахунку ризику в (майже) реальному часі. Автори також підкреслюють потребу поєднувати DRA з розвідданими (CTI) та працювати у Zero Trust-контексті. Це добре корелює з місцем модуля f_4 (прогнозування) та f_8 (самонавчання) у запропонованій моделі.

Zero Trust і ризик-адаптивний контроль доступу. NIST SP 800-207 формулює перехід від периметрових припущень довіри до постійної верифікації суб'єктів, активів і запитів на доступ із урахуванням контексту, що підтримує ідею ризик-адаптивних політик (f_5 –f_6). Додатково SP 800-207A моделює архітектуру доступу у ZTA та демонструє, як будувати рішення з тонким урахуванням ризику у каналі прийняття рішень [1].

UEBA та інсайдерські загрози. На рівні поведінкових ознак сучасні праці підтверджують ефективність UEBA для безперервної автентифікації і виявлення аномалій. Martín et al. (2022) [4] у Knowledge-Based Systems показали, що комбінування клавіатурної та мишкової динаміки дає стійкі сигнали для безперервної автентифікації і це відповідає нашій функції f_2 –f_3 (збір/нормалізація) та побудові вектора Q. Для контексту загроз у CPS, де інсайдери особливо небезпечні, систематичний огляд (Computers & Electrical Engineering, 2024) показує прогалини у виявленні і підкреслює потребу мультиджерельних поведінкових і технічних сигналів і саме те, що робить запропонована ПМ.

Цифрові двійники у безпеці. Огляд IEEE Communications Surveys & Tutorials (Alcaraz & Lopez, 2022) систематизує загрози та контрзаходи для DT і підкреслює їхню корисність для моделювання та «what-if» аналізу безпеки. У нашій ПМ цифровий двійник користувача (f_4) симулює поведінку в штатних та аномальних сценаріях і генерує матрицю ймовірностей атак R, тобто підхід, що резонує з висновками огляду про прогностичну цінність DT [6].

Блокчейн-підсилення контролю доступу. Огляди й систематичні рев'ю [7 – 10] демонструють зрілість інтеграції блокчейну з ABAC/RBAC для підвищення довіри, незмінності логів і автоматизації політик (смарт-контракти, аудит, децентралізація). Для IoT/хмар показано класифікації парадигм доступу, сценарії застосування та виклики (масштабованість, приватність, продуктивність), що робить відстежуваність транзакцій доступу в нашому RBAC-блокчейні (f_4 , f_8) технічно обґрунтованою й методично захищеною.

Таким чином, поточна наукова картина підтверджує необхідність DRA у ZTA-рамці, дієвість UEBA для профілювання персоналу, прогностичну користь цифрових двійників, доречність блокчейну для незмінного аудиту та розподіленого управління політиками доступу. У сукупності це створює надійне підґрунтя для запропонованої процесної моделі з замкненим контуром самонавчання.

2. Процесна модель динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу

У сучасних системах інформаційної безпеки (ІБ) основним елементом є персонал, водночас ефективність технічного захисту безпосередньо залежить від дій користувачів. Особливо це прослідковується під час контролю доступу до інформаційних ресурсів, де рішення про надання прав мають ґрунтуватися на об'єктивних критеріях: посаді, функціональних обов'язках, рівні відповідальності та контексті використання даних. Такий підхід забезпечує формалізовану та перевірену основу для визначення допустимого рівня доступу, перетворюючи довіру з абстрактного поняття на вимірюваний параметр в управлінні ризиками.

Таким чином, виникає потреба у розробці технологічних етапів динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу, де рівень довіри визначається

через систему конкретних показників. Реалізація такої процесної моделі дає змогу автоматично формувати та модифікувати політики доступу на основі динамічного аналізу багатовимірного вектора ознак Q, що поєднує технічні та поведінкові параметри користувача з контекстом взаємодії з ресурсами. Оновлення прав відбувається в режимі реального часу із застосуванням алгоритмів прогнозування на основі цифрового двійника (f_4), тоді як RBAC-блокчейн гарантує криптографічно захищену реєстрацію кожної транзакції доступу в розподіленому реєстрі. Така архітектура дає змогу оперативно змінювати права відповідно до виявлених аномалій, і водночас гарантує верифікованість усіх змін доступу, мінімізуючи ризики несанкціонованих модифікацій.

Запропонована процесна модель (ПМ) динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу: від збору даних до їхньої аналітичної обробки та прийняття рішень. Подана ПМ складається із процесів, і кожен з них відповідає за конкретну дію або процедуру. Така побудова дає системі необхідну гнучкість і здатність адаптуватися до реальних загроз, що виникають у повсякденній роботі з людьми, а не лише в межах теоретичних сценаріїв (рис. 1).

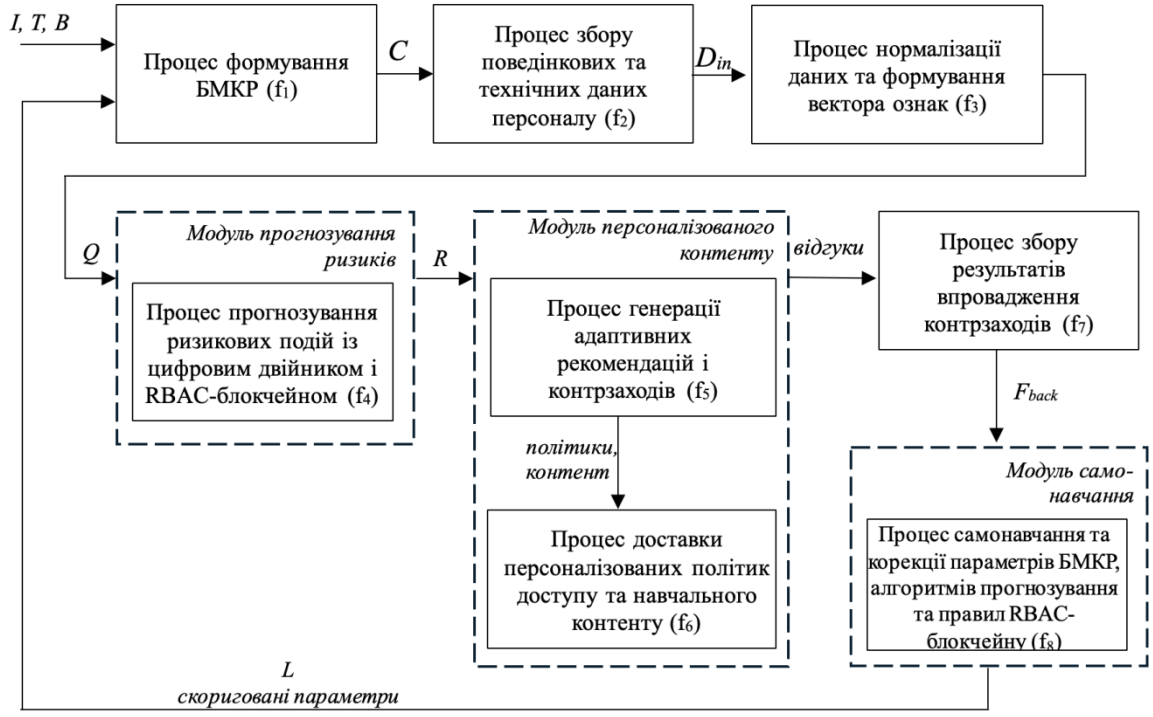


Рис. 1. Процесна модель динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу

У запропонованій ПМ з персоніфікованим підходом дані поетапно перетворюються на адаптивні управлінські рішення, спрямовані на підвищення рівня захисту інформаційних ресурсів.

Процес формування багатовимірної матриці класифікації ресурсів.

На етапі формування багатовимірної матриці класифікації ресурсів (функція f_1) інформаційні ресурси структуруються за визначеними ознаками, серед яких можуть бути конфіденційність, цінність, доступність тощо. Вхідними даними для процесу формування БМКЛ є множини: I – інформаційні ресурси, T – технічні параметри і B – поведінкові метрики персоналу. Кожна ознака має вагу, яка показує її значущість у загальній системі. Методика відповідає підходам [22], заснованим на багатовимірному аналізі та експертному оцінюванні.

Процес збору поведінкових та технічних даних персоналу.

Процес збору поведінкових та технічних даних персоналу (функція f_2) передбачає збір та обробку даних C про поведінкову та технічну активність персоналу від БМКР. Джерелами

таких даних є журнали подій (логи), телеметричні показники, результати анкетування або тестування користувачів.

Процес нормалізації даних та формування вектора ознак.

В результаті роботи процесу нормалізації даних та формування вектора ознак (за функцією f_3) зібрані дані (нормалізовані та уніфіковані) формують вектор ознак Q. Цей вектор є формалізованим представленням поведінкових характеристик окремого суб'єкта доступу та використовується як вхідна змінна для подальших обчислень. Побудова вектора Q ґрунтується на поєднанні класичних статистичних методів і сучасних алгоритмів машинного навчання. Такий підхід відповідає рекомендаціям ризик-менеджменту в інформаційній безпеці.

Процес прогнозування ризикових подій із цифровим двійником і RBAC-блокчейном.

Процес прогнозування ризикових подій із цифровим двійником і RBAC-блокчейном (функція f_4) передбачає, що після формування вектора Q дані надходять до модуля прогнозування, де створюється так званий цифровий двійник користувача (ЦДК) – персоніфікована математична модель, що симулює його поведінку в умовах, наближених до реальних. ЦДК дає змогу оцінювати поточну активність і прогнозувати потенційні небажані дії. Модель ЦДК інтегрується з системами рольового контролю доступу (RBAC) та блокчейном, який фіксує дії користувача в незмінному реєстрі. Зазначені підходи обґрунтовано в дослідженнях із контролю доступу та застосування технологій блокчейн [20, 21].

На основі результатів симуляції формується матриця ймовірностей атак (R). Вона оцінює ймовірність різних типів порушень інформаційної безпеки конкретним користувачем щодо певних ресурсів.

Процес генерації адаптивних рекомендацій і контрзаходів та процес доставки персоналізованих політик доступу та навчального контенту є складовими Модуля персоналізованого контенту.

Модуль персоналізованого контенту передбачає, що на основі матриці R система генерує персоналізовані управлінські рішення, зокрема: адаптивні політики доступу, рекомендації щодо дій користувача (процес генерації адаптивних рекомендацій і контрзаходів за функцією f_5); а також навчальні матеріали з підвищення інформаційної грамотності (процес доставки персоналізованих політик доступу та навчального контенту за функцією f_6).

Процес збору результатів впровадження контрзаходів.

Заходи реалізуються (за функцією f_7) через інтерфейс взаємодії з користувачем, після чого активується модуль зворотного зв'язку. Його функція – моніторинг ефективності впроваджених рішень та збір коригувальних даних (F_{back}).

Процес самонавчання та корекції параметрів БМКР, алгоритмів прогнозування та правил RBAC-блокчейну.

Система використовує ці дані для самонавчання (за функцією f_8): оновлює вагові коефіцієнти, удосконалює прогнозні моделі та уточнює політики доступу. Отже, система працює як замкнутий адаптивний контур управління безпекою з персоналізованим оцінюванням ризиків.

Ідея, що лежить в основі запропонованої системи оцінювання ризиків інформаційної безпеки, полягає в комплексній інтеграції традиційних і сучасних методологічних підходів. Йдеться про поєднання класичних статистичних методів, алгоритмів машинного навчання, моделей цифрової поведінки користувачів і технологій блокчейн. Такий гібридний підхід поєднує реактивну та проактивну стратегії: система не лише фіксує загрози, а й заздалегідь прогнозує потенційні ризики.

Зокрема, упровадження блокчейн-технологій у цю систему виконує критично важливу функцію забезпечення цілісності та незмінності логів доступу до інформаційних ресурсів. Оскільки записи в блокчейні не підлягають редагуванню без фіксації цифрового сліду, це створює прозору, верифіковану історію взаємодії користувача із системами. Така незмінність підвищує рівень довіри до механізмів аудиту й сприяє обґрунтованому ухваленню рішень у контексті ризик-менеджменту. Фактично йдеться про формування надійного джерела даних

у цифровому середовищі, що є особливо важливим в умовах багаторівневої корпоративної інфраструктури.

3. Складові процесної моделі динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу

На першому етапі відбувається процес формування матриці класифікації ресурсів (функція f_1) і слугує методологічною основою всієї системи. Побудова матриці передбачає аналіз наявних в організації інформаційних активів для їхнього структурованого опису за низкою релевантних ознак, серед яких технічна цінність ресурсу, його критичність для бізнес-процесів, рівень конфіденційності та контекст використання. Вагові коефіцієнти призначаються на основі комбінованого підходу, що поєднує елементи експертного оцінювання (із залученням фахівців з інформаційної безпеки, аналітиків і керівників підрозділів) та статистичні методи, такі як кореляційний аналіз або метод головних компонент (PCA, Principal component analysis), що мінімізує суб'єктивність. Принцип адаптації матриці полягає в періодичному перерахунку вагових коефіцієнтів зі зміною структури підприємства, складу інформаційних ресурсів або бізнес-процесів. Це реалізується через повторне застосування обраної методики вагового аналізу до оновленого набору ознак, що гарантує актуальність матриці щодо поточної організаційної конфігурації.

Другим етапом є процес збору й уніфікації поведінкових та технічних даних персоналу, позначений функцією f_2 . Його метою є формування повного та релевантного масиву даних, що відображає реальну активність користувачів у цифровому середовищі організації. Збір інформації відбувається з різних джерел: журнали подій серверів і прикладного ПЗ, телеметричні дані з мережевого обладнання, результати анкетування та опитувань, показники продуктивності праці. Отримані дані проходять обробку для очищення, нормалізації та синхронізації з урахуванням видалення дублікатів, виправлення форматних помилок та вирівнювання часових міток. Для уніфікації застосовуються формати CEF, JSON, LEEF, а передача між модулями відбувається через захищені канали (TLS).

На етапі роботи процесу нормалізації даних та формування вектора ознак f_3 виконується формування вектора ознак Q – числового представлення профілю користувача, що забезпечує його подальший аналіз математичними моделями. Формалізовано, $Q = \{x_1, x_2, \ldots, x_n\}$, де кожен x_i є нормалізованим і, за потреби, закодованим показником, отриманим із поведінкових чи технічних характеристик. На цьому етапі числові змінні масштабуються (Min–Max або z-нормалізація), а категоріальні – кодуються (one-hot encoding). Для кожної ознаки розраховуються статистичні характеристики (середнє, медіана, стандартне відхилення, коефіцієнт варіації, частота подій), після чого за допомогою Random Forest Feature Importance відбираються найбільш інформативні параметри. Отже, f_2 виконує збір даних, а f_3 перетворює їх у стандартизовану форму Q для моделювання.

Процес прогнозування ризикових подій із цифровим двійником і RBAC-блокчейном використовує функцію f_4 . Вектор Q надходить на вхід модуля прогнозування ризиків, де будується цифровий двійник працівника, який моделює його поведінку в штатних і потенційно аномальних сценаріях. У результаті формується матриця ймовірностей ризиків R, де кожен елемент відображає ймовірність реалізації певної загрози щодо конкретного інформаційного активу.

Функція f_5 процесу генерації адаптивних рекомендацій і контрзаходів має чітку мету – трансформувати матрицю R у набір конкретних контрзаходів, спрямованих на зниження ризиків, з урахуванням їхніх категорій та порогових значень. Розрізняють три категорії контрзаходів: превентивні (запобігають інциденту), детективні (виявляють його) та коригувальні (усувають наслідки). Активація відбувається, якщо ймовірність у R перевищує визначений поріг P^t, який налаштовується залежно від класу активу та допустимого рівня ризику (наприклад, $P^t = 0{,}7$ для критичних систем, 0,5 – для середньої важливості та 0,3 – для менш важливих). Якщо ризик перевищує поріг, запускаються адаптивні дії: від додаткової багато-

факторної автентифікації та тимчасового обмеження доступу до призначення обов'язкової ручної перевірки. Паралельно формується персоналізована програма підвищення обізнаності співробітника з питань кібербезпеки, контент якої динамічно адаптується до його актуального профілю ризику.

Таким чином, система усуває поточні загрози й знижує ризики в довгостроковій перспективі завдяки розвитку цифрової компетентності та кіберграмотності персоналу. Це створює додатковий рівень стійкості інформаційної системи підприємства, заснований на проактивному управлінні впливом діяльності людини.

Після того як система сформувала рекомендації з управління ризиками та навчальні матеріали для підвищення обізнаності користувачів, наступним етапом є процес доставки персоналізованих політик доступу та навчального контенту (що забезпечується функцією f_6) для ефективного доставлення цього контенту безпосередньо до кінцевих суб'єктів. Цей процес виконує Модуль персоналізованого контенту.

Функціонування цього модуля ґрунтується на мультиканальному підході до комунікації. Кожен користувач бачить персоналізовану інформаційну панель, адаптовану не лише до його ролі в організації, але й до виявлених ризиків. Інтерфейс структуровано так, щоб акцентувати на пріоритетних діях, зокрема на завданнях, які потребують негайного виконання, термінах їхнього завершення, способах підтвердження тощо. Такий підхід мінімізує часовий лаг між отриманням інформації та діями користувача, що зменшує «вікно ризику» – період, упродовж якого вразливість може бути використана.

В результаті роботи функції f_7 активується процес збору результатів впровадження контрзаходів, що збирає та аналізує реакцію користувача на отриманий контент і рекомендації. Мета цього етапу – не лише зафіксувати факт взаємодії, а й оцінити її якість, повноту та наслідки для зміни поведінки або рівня ризику.

Крім об'єктивних цифрових показників процес збору результатів впровадження контрзаходів використовує інструменти якісного аналізу: міні опитування, анкетування щодо зручності та зрозумілості контенту, а також відкриті поля для фідбеку. Інформація структурується та агрегується в матрицю зворотного зв'язку F_{back}, що містить кількісні (наприклад, середній час реагування) та якісні дані (наприклад, задоволеність користувача освітнім процесом). Це дає змогу виявляти не лише технічні уразливості, а й елементи культури безпеки в межах організації, зокрема мотивацію працівників, ступінь залученості та рівень цифрової відповідальності.

Останнім, але концептуально найважливішим компонентом системи є модуль самонавчання (функція f_8), завданням якого є динамічна адаптація всієї системи на основі даних, накопичених у матриці F_{back}. Це означає, що система не лише фіксує реакцію користувачів на застосовані контрзаходи, але й використовує цю інформацію для корекції власних моделей і правил.

Процес самонавчання та корекції параметрів БМКР, алгоритмів прогнозування та правил RBAC-блокчейну починається зі збору прикладів успішних і помилкових спрацювань прогнозної моделі (зокрема, випадків хибнопозитивних та хибнонегативних класифікацій). Для оцінювання якості моделі використовується функція втрат L, яка є зваженою комбінацією Binary Cross-Entropy та Focal Loss. Такий підхід дає змогу приділити увагу рідкісним, але критично важливим класам загроз, мінімізуючи при цьому кількість помилкових спрацьовувань.

Після обчислення L модуль f_8 ініціює процедуру автоматичного перенавчання, застосовуючи алгоритми глибокого навчання. Якщо аналіз матриці помилок свідчить про значний відсоток хибнопозитивних спрацювань для певної категорії подій, система коригує ваги ознак у багатовимірній матриці класифікації ресурсів (подає їх на вхід функції f_1) та оновлює відповідні RBAC-правила, знижуючи чутливість для цієї категорії. Якщо ж домінують хибнонегативні випадки, навпаки, вага підвищується, а політики доступу посилюються.

Механізм корекції RBAC-правил реалізується через модифікацію контекстних умов (наприклад, обмеження часу чи геолокації доступу) та рівнів авторизації (додавання або вилучення вимоги багатофакторної автентифікації). Усі зміни фіксуються в блокчейн-журналі доступу, що гарантує прозорість і неможливість несанкціонованого редагування політик.

Параметри моделі ризиків оновлюються двома способами: повним перенавчанням або частковим донавчанням (fine-tuning) із застосуванням підходу transfer learning, що дає змогу зберегти вже накопичені знання та швидко адаптувати модель до нових умов без значних обчислювальних витрат. Таким чином, функція f_8 виконує роль замикального механізму зворотного зв'язку, що підвищує точність та адаптивність системи в динамічному середовищі загроз.

Отже, повний цикл, від генерації контенту до отримання зворотного зв'язку й самокорекції, формує адаптивне середовище керування інформаційними ризиками, здатне оперативно реагувати на нові загрози, забезпечуючи гнучкість і стійкість системи. Інтеграція блокчейн-механізмів і класичних підходів до контролю доступу гарантує прозорість, верифікованість і захищеність усіх транзакцій та змін у системі. Така циклічна структура з елементами самонавчання й користувацької взаємодії забезпечує не просто стабільність, а еволюційний розвиток системи, що є критично важливим у динамічному середовищі інформаційної безпеки.

Висновки

Запропонована процесна модель адресує ключові виклики сучасної ІБ: динаміку загроз, вагу людського фактора, потребу в персоналізації й аудитованості. Завдяки поетапній трансформації даних у вектор ознак Q, побудові цифрового двійника користувача та матриці ризиків R, а також модулю персоналізованого контенту й механізму самонавчання система забезпечує підвищення точності оцінювання ризику, покращення оперативності реагування через ризик-адаптивні політики доступу та вдосконалення прозорості змін і журналів за рахунок RBAC-блокчейну. Порівняння з сучасними підходами (DRA/UEBA/ZTA/блокчейн-АС) підтверджує методологічну узгодженість і практичну релевантність моделі для корпоративних середовищ, зокрема там, де критичною є відтворюваність аудиту та стале зниження інсайдерських ризиків.

Список літератури:

1. Papanikolaou A., Varvarousi E., & Gavala E. Postal sector digitalisation: Security and vulnerabilities // International Journal of Applied Systemic Studies. 2024. 11(1). P. 42–51. https://doi.org/10.1504/IJASS.2024.139211(inderscience.com).

2. Sèdes F., & Degrace J. (2024). Social engineering and security: From human vulnerabilities to malicious threats // 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). 2024. P. 301–305. IEEE. https://doi.org/10.1109/WiMob61911.2024.10770451.

3. Marushchak L., Pavlykivska O., Khrapunova Y., Kostiuk V., & Berezovska L. The economy of digitalization and digital transformation: Necessity and payback11th International Conference on Advanced Computer Information Technologies (ACIT). 2021. P. 305–308. IEEE. https://doi.org/10.1109/ACIT52158.2021.9548529.

4. Alnafea F. S. M., Sengar A. S., Hamid N. K., Selladurai K. M., Hassan S. I., & Saravanakumar R. Improving multi-factor authentication security with deep learning-based user behaviour analysis // 3rd International Conference on Integrated Circuits and Communication Systems (ICICACS). 2025. P. 1–5. IEEE. https://doi.org/10.1109/ICICACS65178.2025.10968961.

5. Kaur M., & Garg P. Exploring behavioral patterns for security in cloud computing: A case study // 3rd International Conference on Advancement in Computation & Computer Technologies (InCACCT). 2025. P. 720–725. IEEE. https://doi.org/10.1109/InCACCT65424.2025.11011337.

6. Terumalasetti S., & Reeja S. R. Enhancing social media user's trust: A comprehensive framework for detecting malicious profiles using multi-dimensional analytics // IEEE Access. 2025. Vol. 13. P. 7071–7093. https://doi.org/10.1109/ACCESS.2024.3521951.

7. Korobeinikova T. I. The zero trust model: Theory, practice, and prospect. In Heritage of European Science 2025: Innovative technology, computer science, cybernetics and automation, security systems, transport development, architecture and construction, physics and mathematics (Monographic series "European Science," Book 37, Part 1, pp. 126–147). European Science. ISBN 978-3-98924-080-3.

8. Коробейнікова Т., Журавель І., Бодак А., & Бороденко Д. Концепція нульової довіри: сучасні методи забезпечення кібербезпеки в корпоративних мережах // Вісник Львів. держ. ун-ту безпеки життєдіяльності. 2025. Т. 30. С. 67–77. Retrieved із https://journal.ldubgd.edu.ua/index.php/Visnuk/article/view/2769.

9. Lee S., Huh J.-H., & Woo H. Security System Design and Verification for Zero Trust Architecture. Electronics. 2025. Vol. 14(4). P.643. https://doi.org/10.3390/electronics14040643.

10. Molina M., Betarte G., & Luna C. Consent validation for personal data access control using ABAC // Proceedings of the 13th Latin-American Symposium on Dependable and Secure Computing (LADC '24). 2024. P. 30–31. Association for Computing Machinery. https://doi.org/10.1145/3697090.3699803.

11. Rose S., Borchert O., Mitchell S., & Connelly S. Zero Trust Architecture (NIST Special Publication 800-207) // National Institute of Standards and Technology. 2020. https://doi.org/10.6028/NIST.SP.800-207.

12. Chandramouli R., Badger L., & O'Rourke D. SP 800-207A: A Zero Trust Architecture Model for Access Control in Cloud-Native Applications. NIST. 2023. https://csrc.nist.gov/pubs/sp/800/207/a/final (csrc.nist.gov).

13. Cheimonidis P., & Rantos K. Dynamic risk assessment in cybersecurity: A systematic literature review // Future Internet. 2023. Vol. 15(10). P. 324. https://doi.org/10.3390/fi15100324 (MDPI).

14. Martín A. G., Martín-de-Diego I., Fernández-Isabel A., et al. Combining user behavioural information at the feature level to enhance continuous authentication systems // Knowledge-Based Systems. 2022. Vol. 244. P. 108544. https://doi.org/10.1016/j.knosys.2022.108544 (ScienceDirect).

15. Al-Mhiqani M. N., Alsboui T. A. A., Al-Shehari T., Abdulkareem K. H., et al. Insider threat detection in cyber-physical systems: A systematic literature review // Computers & Electrical Engineering. 2024. Vol. 119, P. 109489. https://doi.org/10.1016/j.compeleceng.2024.109489 (ResearchGate).

16. Alcaraz C., & Lopez J. Digital twin: A comprehensive survey of security threats // IEEE Communications Surveys & Tutorials. 2022. Vol. 24(3). P. 1475–1503. https://doi.org/10.1109/COMST.2022.3171465 (NICS Lab).

17. Ullah S. S., Oleshchuk V., & Pussewalage H. S. G. A survey on blockchain-envisioned attribute-based access control for Internet of Things: Overview, comparative analysis, and open research challenges // Computer Networks. 2023. Vol. 235. P. 109994. https://doi.org/10.1016/j.comnet.2023.109994 (ACM Digital Library).

18. Punia A., Hoda M., Kaushik K., & Tomar D. A systematic review on blockchain-based access control systems in cloud environment // Journal of Cloud Computing. 2024. Vol. 13. P. 62. https://doi.org/10.1186/s13677-024-00697-7(PMC).

19. Namane S., Derhab A., Guerroumi M., & Challal Y. Blockchain-based access control techniques for IoT: A systematic review // Electronics. 2022. Vol. 11(14). P. 2225. https://doi.org/10.3390/electronics11142225 (MDPI).

20. Ямнич А. Б. Модель контролю доступу персоналу до інформаційних ресурсів підприємств на основі RBAC та технології BLOCKCHAIN / А.Б. Ямнич, Т.І. Коробейнікова // Вісник Хмельниц. нац. ун-ту. 2024. Т. 343, №6(1). С. 380–386. ISSN 2307–5732 https://doi.org/10.31891/2307-5732-2024-343-6-56.

21. Коробейнікова Т. І. Багатовимірна матриця класифікації інформації для оцінки ризиків інформаційної безпеки / Т. І. Коробейнікова, А. Б. Ямнич // Інформаційні технології та комп'ютерна інженерія. 2024. №2. С. 91–106. ISSN 1999–9941.

22. Korobeinikova T., Tachenko I., Romanyuk O., Romanyuk S., Stakhov O. and Reyda O. Assessing Network Security Risks: a Technological Chain Perspective // 14th International Conference on Advanced Computer Information Technologies (ACIT), Ceske Budejovice, Czech Republic. 2024. P. 565–570.
doi: 10.1109/ACIT62333.2024.10712586.

Надійшла до редколегії 03.06.2025

Відомості про авторів:

Коробейнікова Тетяна Іванівна – канд. техн. наук, Національний університет «Львівська політехніка», доцент кафедри безпеки інформаційних технологій; Україна; e-mail: tetianakorobeinikova@gmail.com; ORCID: https://orcid.org/0000-0003-2487-8742

Ямнич Андрій Богданович – Національний університет «Львівська політехніка», аспірант кафедри безпеки інформаційних технологій; Україна; e-mail: andrii.b.yamnych@lpnu.ua; ORCID: https://orcid.org/0009-0005-7226-1896

УДК 004.738.5:519.816:004.056.5 **DOI:10.30837/rt.2025.3.222.08**

L. MELNIKOVA, Ph.D. (Tech. Sci.), A. MARCHUK (Tech. Sci.),
S. SHTANGEI, Ph.D. (Tech. Sci.)

UKRAINIAN INTERNET SERVICE PROVIDERS RANKING: MULTI-CRITERIA MODEL INCORPORATING CYBERSECURITY

Introduction

Internet services have become an indispensable facet of modern life, imposing stringent requirements on both quality and security. Selecting an Internet Service Provider (ISP) is a complex task because it involves many heterogeneous often mutually conflicting criteria. High demands for speed, stability and safety, combined with cost constraints, necessitate an approach that can reconcile diverse requirements within a single decision-making framework [1].

Multi-criteria optimization is well suited to this task, as it seeks solutions that are jointly optimal across several criteria. The present study aims to demonstrate a practical method for ISP selection based on multi-criteria analysis, supported by heuristic procedures for refining candidate optima [2].

Because clients invariably seek the “best” service, yet must choose a single provider from among many options and parameters, the selection process is formidable. Beyond technical specifications, one must consider provider ratings, user reviews, pricing, and contractual terms. Factoring these technical and commercial indicators into the analysis enables a more informed choice of provider capable of delivering reliable, high-speed connectivity [3–5].

Despite growing global attention to cybersecurity, this criterion is often omitted from integrated assessments in Ukrainian research. According to the 2023 market review, the telecommunications sector is among the most frequently targeted: more than 500 cybersecurity incidents were recorded that year. The National Security and Defense Council of Ukraine likewise emphasizes the importance of cyber-resilient ISP infrastructure, especially under conditions of hybrid warfare [6, 7].

Accordingly, a clear research gap exists: the absence of an integrated model that evaluates ISPs not only by classical quality parameters but also by their cybersecurity posture. A multi-criteria approach to ISP assessment can thus benefit both consumers by facilitating provider selection and service providers by guiding strategic development under competitive pressure.

The multi-criteria optimization problem that arises in selecting an effective solution requires the identification of a compromise alternative. The search process is heuristic in nature and takes place under uncertainty. At several stages, a decision maker (DM) is involved to formulate the task and interpret its results. Utility theory underlies this approach, postulating that the DM can assign a single, aggregated value or utility to each criterion.

Many numerical methods for such problems are based on selecting a point from the Pareto set corresponding to criterion weights. Choosing the best alternative requires taking account of conditional preferences, derived from additional expert information and processed according to utility theory [8].

The objective of this study is therefore to develop a multi-criteria model for evaluating Ukrainian ISPs that includes cybersecurity as an independent criterion, employs heuristic optimization methods, and involves the DM in refining criterion weights.

To this end, we propose a heuristic refinement procedure with DM participation. The procedure is an interconnected set of formalized methods for specifying the optimization criterion for the system under design, forming an initial set of design alternatives, identifying the Pareto-optimal subset, and narrowing this subset to a single best design alternative using additional expert information [9].

Methods

Before a single solution can be selected from among all feasible compromises, the fundamental principles (axiomatics) underlying the choice rule must be justified. This requires gathering supplementary data obtained by analyzing the system's objectives and formalizing its specific features.

A key stage of the analysis is to establish how important each partial criterion is; this information guides the most suitable compromise scheme. Depending on the amount and quality of information available, three decision-making scenarios are distinguished:

Scenario 1 – Exact weights known. Precise numerical values of the criterion weights are given.

Scenario 2 – Ordinal information only. Exact weights are unavailable, but the criteria can be ranked by importance; those ranks are converted into weights by means of appropriate formulae.

Scenario 3 – No weight information. Neither quantitative nor qualitative data exist; consequently, all criteria are assigned equal (or nearly equal) weights.

Under uncertainty, the role of the decision maker (DM) becomes pivotal. The DM assigns or revises criterion weights to reflect the current context; can adjust admissible intervals for criterion values in accordance with system goals, thereby shaping the Pareto domain; participates interactively, allowing the final choice to be aligned with real-world objectives and to be specified with greater precision.

Let *{I}* denote the *n*-dimensional Pareto set of feasible alternatives in the multi-criteria optimisation (MCO) problem. The quality of the *i*-th alternative is described by an *m*-dimensional vector of partial criteria $Ki=(Ki_1,Ki_2,\dots,Ki_m)^T$.

The task is to determine an alternative I^* from *{I}* that provides satisfactory values of all criteria *Kij (j=1,…,m)* and yields the best possible compromise between them, given the structure of PPP and any a priori information about criterion importance. Three information cases are considered: pre-specified weights, comparative (rank-order) importance, and no information on importance.

The decision algorithm evaluates each criterion's influence on overall system performance via weight coefficients (importance factors) *Pj (j=1,…,m)* that satisfy $\sum^m P_j = 1$

The utility of the *i*-th alternative is modelled by the general additive function

$$Q_i = \sum_{j=1}^{m} P_j q_{ij}, \tag{1}$$

where $q_{ij} = \xi_{ij}\left[K_{ij}(x)\right]$ is the normalised utility of the *j*-th partial criterion for *i*-th alternative.

To address this problem, the functional form of the partial-criterion utility function must first be justified.

Such a function must be universal yet readily adaptable to the specific features, goals, and criteria of the system under study. Accordingly, it should satisfy the following requirements:

- be dimensionless;
- have a unit measurement interval (0,1);
- be invariant with respect to the extremum direction of each criterion (min or max), so that the best value of a criterion maps to 1 and the worst to 0;
- permit characteristic non-linear dependencies.

One function that meets these requirements is proposed in [8]. A suitable form is

$$\xi_{ij}\left(K_{ij}\right) = \left[\frac{K_{ij}-K_{j\,worst}}{K_{j\,best}-K_{j\,worst}}\right], \tag{2}$$

where *Kij* is the value of the *j*-th partial criterion for the i-th alternative, and *Kj best*, *Kj worst* denote the best and worst permissible values of that criterion, respectively. Equation (2) quantifies the degree to which *Kij* approaches the local optimum for criterion *j*.

To establish the bounds of the compromise region *X* that is, the values *Kj best* and *Kj worst*, the following procedure [5] is employed.

Over the feasible set of alternatives PPP, optimization is performed separately for each partial criterion *Kj*. This yields the criterion-specific extremal solution

$$X_j^0 = \arg \underset{x \in X}{\text{extr}} K_{ij}(x),\ j = \overline{1,m}, i = \overline{1,n}, \quad (3)$$

together with the corresponding values of all partial criteria.

Consequently,

$$K_{j\,best} = K_j\left(X_j^0\right),\ K_{j\,worst} = \begin{cases} \max\limits_i K_j\left(x_j^0\right), K_j(x) \to min \\ \min\limits_i K_j\left(x_j^0\right), K_j(x) \to max \end{cases}. \quad (4)$$

Thus, *Kj best* and *Kj worst* define the image of the approximated compromise region in the space of partial criteria.

Weight determination often encounters serious difficulties and is typically reduced to expert judgment. To ease the expert's task, it is sometimes sufficient to request only an ordinal ranking of criteria by importance; the ranks are then transformed into weight coefficients via predefined formulae. In the absence of any quantitative or qualitative information from the DM, it is logical to adopt equal (or nearly equal) weights for all criteria.

Search Procedure for the Optimal Solution:

1. Initial exploration. The entire set of feasible alternatives {I} is examined, and separate optimization is carried out for each partial criterion *Kj (j=1,…,m).* The resulting maxima form the vector *Z1*. This vector contains the "best-possible'' values of the partial criteria across all alternatives in PPP and is presented to the decision maker (DM) for reference.

2. Global-utility optimization. Optimization is next performed with respect to the aggregated utility function *Qi* defined in (1). The alternative that attains the largest utility value $Q_1 = max\{Q_i\}$ is identified. Its partial-criterion values constitute the vector

$$Y1 = \begin{vmatrix} K_{l1} \\ \cdots \\ K_{lm} \end{vmatrix}$$

which is submitted to the DM as a candidate solution.

3. Satisfaction check. The DM is asked: "Are all criteria satisfactory?" The decision is based on *Z1*. If the answer is no, the DM selects the most unsatisfactory criterion *Kr* and specifies a threshold value C_r deemed acceptable for that criterion.

4. Constraint tightening.

A new feasible domain $I^* = \left\{i \left| \left(K_{ir} \le C_r\left(K_{ij} \to max\right)\right) \vee \left(K_{ir} \ge C_r\left(K_{ij} \to min\right)\right)\right.\right\}$ is defined, and Step 1 is repeated on *{I*}* to obtain an updated vector of maxima *Z2*.

The DM is then asked: "Is the reduction from Z1 to Z2 acceptable?"

If no, the threshold *Cr* is relaxed (made less demanding), and Steps 4–3 are repeated.

If yes, the compromise value *Cr* is fixed for all subsequent iterations.

5. Utility re-optimisation within the restricted set. On the refined set *{I*}* the aggregated utility function (1) is maximised again, yielding a new vector *Y2*, which is presented to the DM. The procedure returns to Step 3.

The algorithm terminates when, at Step 3, the DM responds affirmatively, i.e. when the current vector of criterion values satisfies all requirements.

Compilation of the Input Dataset

Most ISPs offer broadly similar baseline services and differ mainly in the number of value-added options and the quality of service (QoS). Consequently, these criteria are decisive for distinguishing one provider from another. Current Ukrainian market surveys [9] typically evaluate ISPs against the following parameters:

- Service cost;
- Connection speed;
- Optimal balance of services and price within the tariff plan;

– Connection stability;
– Availability of high-quality technical support and the practical ability to obtain effective assistance 24/7;
– Range of auxiliary services supplied by the ISP;
– Subscriber reviews and reputational feedback.

Ukraine hosts roughly 6 500 registered ISPs and network operators, most of which operate in large population centres. Provider categories and their characteristics are summarised in Table 1.

Table 1

Provider type	Characteristics	Remarks
Primary providers	Deliver global Internet connectivity at international traffic-exchange points (e.g., NAP – Network Access Point; MAE – Metropolitan Area Exchange; CIX – Commercial Internet Exchange).	Absent in Ukraine.
Tier-1 providers (primary regional)	Supply Internet access to local users and to downstream ISPs; purchase upstream transit from foreign carriers at international exchange points via domestic backhaul links.	≈ 20 in Ukraine.
Major providers	Tier-1 operators plus large Tier-2 companies that serve extensive subscriber bases.	≈ 40 in Ukraine.
Other providers (Tier-2/Tier-3; small and medium-sized)	Lease bandwidth of varying capacity from higher-tier providers; operate dozens to hundreds of access lines.	The majority of Ukrainian ISPs fall into this category.

These categorical distinctions and evaluation parameters form the core of the input data used in the subsequent multi-criteria analysis.

Such a high market concentration means that residents of a single multi-storey building may receive offers from five to seven different companies. This forces operators to launch aggressive promotional campaigns featuring bonus periods and temporary discounts to attract subscribers. However, these promotions usually expire quickly, after which customers are migrated to the full tariff. As for price dynamics, it is worth noting that Ukraine still enjoys some of the lowest Internet prices in the world, making service cost a critical selection criterion.

Different access technologies inevitably deliver different levels of quality of service (QoS). As the number of networked devices and the intensity of Internet usage increase, quality requirements become more stringent. Below, we outline the principal characteristics that determine this quality.

Technical parameters – the specifications that govern the performance and reliability of a service, device, or system – are central to ensuring a stable, high-speed connection. When choosing an ISP, one should pay close attention to parameters such as data-transfer rate, bandwidth, and connection stability, all of which have a direct impact on perceived Internet quality.

The most familiar indicators are upload and download speeds, routinely expressed worldwide in megabits per second (Mbps) or gigabits per second (Gbps).

Data-transfer and download speeds depend on several factors, including:

1. Channel bandwidth. Throughput is limited by the capacity of the physical link between the subscriber and the ISP. The greater the bandwidth, the higher the attainable data-rate.

2. Access technology. Cable, fibre-optic and wireless connections support different maximum speeds.

3. Distance to the provider's node. A long physical distance from the ISP's point of presence can degrade speed: the closer the subscriber is to the node, the lower the signal loss and the smaller the throughput reduction.

4. Network load. Regional or provider-side congestion may also curb both upload and download rates; simultaneous activity by many users can lead to noticeable slow-downs.

Data-rate performance determines how quickly information can be uploaded to or downloaded from the network; it affects page-load times, streaming quality, file-transfer duration and other

online activities. Average upload and download speeds for Ukrainian ISPs are summarised in Table 2.

Latency (measured in milliseconds) indicates the time required for a signal to travel from the user's device to the server and back. Lower latency enables faster interaction and better real-time responsiveness. It is therefore a critical indicator of connection quality and data-transfer efficiency. Latency influences the time needed to send requests and receive server responses and is affected by several factors:

1. Physical distance. Greater separation between sender and receiver increases round-trip time.

2. Access technology. Fibre links typically exhibit low latency, whereas wireless links may suffer higher latency.

3. Network configuration. Heavy traffic, insufficient bandwidth or misconfigured equipment can all elevate delay; the more congested or impaired the network, the greater the latency.

4. Equipment and infrastructure quality. The performance of routers, switches, cabling and other hardware also contributes to overall delay.

Latency is especially important for services requiring rapid feedback—such as online gaming or video calls—because optimal latency ensures higher quality and smoother user experience. Average latency figures for each Ukrainian ISP are presented in Table 3.

Table 2

Internet Service Provider	Transmission speed, Mbps	Download speed, Mbps
Datagroup	31.78	32.6
Fregat	39.65	39.48
Kyivstar	42.56	46.67
Lanet	64.82	66.17
O3 Freenet	48.8	51.17
Triolan	54.56	56.48
UkrTelecom	11.24	6.1
Vega Telecom	33.2	33.58
Volia	45.45	27.58

Table 3

Internet Service Provider	Delay, ms	Internet Service Provider	Delay, ms	Internet Service Provider	Delay, ms
Datagroup	59.67	Lanet	25.88	UkrTelecom	65.17
Fregat	39.8	O3 Freenet	29.22	Vega Telecom	39.88
Kyivstar	33.8	Triolan	30.64	Volia	32

Security is arguably the most critical factor when choosing an ISP. A provider must be able to safeguard both customer data and its own infrastructure. A positive indicator is the presence of a robust backup system that can restore lost or corrupted data under any unforeseen circumstance. Likewise, safe operating conditions for servers (adequate ambient temperature, reliable power supply, and so forth) must be maintained.

The need for an information-security function within organizations becomes more evident each year. Growing cyber-crime, an increasingly tense geopolitical climate, and other external pressures push information-security risks to the foreground. Threat-response time therefore serves as a key indicator of a provider's ability to defend connections against external attacks.

Selecting an optimal ISP is essential for high-quality Internet connectivity. Connection quality depends directly on download and upload speeds as well as latency. Equally important is reliable protection, because no user wants an attacker to gain access to personal information or to disrupt service by targeting the provider. Cost constitutes a further decisive factor: users want many features yet prefer to pay less, so price must be integrated into the decision model. Although dependency analysis yields general guidelines, choosing a single provider that meets all requirements remains a challenging task [12].

To tackle this challenge, a heuristic multi-criteria optimisation (MCO) procedure was employed. The principal criteria influencing ISP selection are:

– Upload speed – efficiency of transmitting large data volumes;
– Download speed – stability and convenience when accessing media resources;
– Latency – critical for real-time applications such as video conferencing and gaming;
– Service cost – key economic consideration for users;
– Threat-response time – indicator of the provider's ability to protect the connection from external threats.

Tables 4 and 5 list the partial criteria and initial dataset for the MCO task involving the five most popular Ukrainian ISPs. The problem is analysed under the assumption of known criterion weights; the relative importance of those criteria supplies additional information that constrains the selection process.

Table 4

Number of criteria m = 5 Number of vectors n = 5	CRITERIA CHARACTERISTICS			
	Number	Name	Type of extremum	Weight
	1	Upload speed	max	0.2
	2	Download speed	max	0.2
	3	Delay	min	0.2
	4	Cost	min	0.1
	5	Response speed to attack	min	0.3

Table 5

ISP	Upload speed, Mbit/s	Download speed, Mbps	Delay, ms	Cost, UAH	Response speed to attack, s
Kyivstar	42.56	46.67	34	250	5
UkrTelecom	11.24	6.1	65	260	20
Triolan	54.56	56.48	31	99	120
Volia	45.45	27.8	32	150	90
Datagroup	31.78	32.6	60	200	60

Solution Procedure and Results Analysis

A JavaScript application was developed to solve the multi-criteria optimisation (MCO) problem using the heuristic procedure; HTML and CSS were employed for visualisation [5].

Five of the most popular Ukrainian ISPs—Kyivstar, Ukrtelecom, Triolan, Volia, and Datagroup—were analysed. For each provider the values of all partial criteria were determined.

The initial run assigned highest priority to threat-response time. Under this weighting scheme, the corresponding optimal outcomes for every provider were obtained; they are presented in Fig. 1.

Результат вирішення задачі БКО

Провайдер	Швидкість завантаження, Мбит/с	Швидкість скачування, Мбит/с	Затримка, мс	Вартість, грн	Швидкість реагування на атаку, с	Qi, При заданих вагах адітивних критеріїв	Qi, При відстутності інформації о вагах адітивних критеріїв
Київстар	0.722	0.805	0.911	0.062	1	0.792	0.698
Укртелеком	0	0	0	0	0.869	0.26	0.172
Triolan	1	1	1	1	0	0.7	0.8
Volia	0.789	0.43	0.97	0.683	0.26	0.583	0.625
Datagroup	0.474	0.526	0.147	0.372	0.521	0.421	0.405

Fig. 1. Output of the program under the decision scenario "priority = threat-response time," assuming no prior information about criterion importance

Subsequent testing modified the criterion weights to emphasise service cost (Table 6), enabling identification of the alternative that offers the best quality-to-price balance. The results obtained with cost-minimisation priority are shown in Fig. 2.

Table 6

CRITERIA CHARACTERISTICS			
Number	Name	Type of extremum	Weight
1	Upload speed	max	0.2
2	Download speed	max	0.2
3	Delay	min	0.2
4	Cost	min	0.3
5	Response speed to attack	min	0.1

Результат вирішення задачі БКО

Провайдер	Швидкість завантаження, Мбит/с	Швидкість скачування, Мбит/с	Затримка, мс	Вартість, грн	Швидкість реагування на атаку, с	Qi, При заданих вагах адітивних критеріїв	Qi, При відстутності інформації о вагах адітивних критеріїв
Київстар	0.722	0.805	0.911	0.062	1	0.604	0.698
Укртелеком	0	0	0	0	0.869	0.086	0.172
Triolan	1	1	1	1	0	0.899	0.8
Volia	0.789	0.43	0.97	0.683	0.26	0.666	0.625
Datagroup	0.474	0.526	0.147	0.372	0.521	0.391	0.405

Fig. 2. Program output under the modified weighting scheme (priority = service cost)

A further experimental scenario increased the weights assigned to the speed-related criteria (Table 7), thereby identifying the optimal choice for users who prioritise maximum bandwidth. The corresponding test results are presented in Fig. 3.

Table 7

CRITERIA CHARACTERISTICS			
Number	Name	Type of extremum	Weight
1	Upload speed	max	0.3
2	Download speed	max	0.3
3	Delay	min	0.2
4	Cost	min	0.1
5	Response speed to attack	min	0.1

Результат вирішення задачі БКО

Провайдер	Швидкість завантаження, Мбит/с	Швидкість скачування, Мбит/с	Затримка, мс	Вартість, грн	Швидкість реагування на атаку, с	Qi, При заданих вагах адітивних критеріїв	Qi, При відстутності інформації о вагах адітивних критеріїв
Київстар	0.722	0.805	0.911	0.062	1	0.744	0.698
Укртелеком	0	0	0	0	0.869	0.086	0.172
Triolan	1	1	1	1	0	0.9	0.8
Volia	0.789	0.43	0.97	0.683	0.26	0.652	0.625
Datagroup	0.474	0.526	0.147	0.372	0.521	0.417	0.405

Fig. 3. Program output under the modified weighting scheme (priority = speed)

The results indicate that Kyivstar is the optimal choice for security-conscious users, whereas Triolan offers the best price-to-speed ratio for budget-oriented customers.

Conclusions

The study has developed and validated an integrated multi-criteria decision-making (MCDM) framework for selecting Ukrainian Internet Service Providers (ISPs) that explicitly incorporates cybersecurity alongside conventional quality-of-service (QoS) metrics. The main findings are as follows:

Holistic evaluation model. By combining upload/download throughput, latency, cost, and threat-response time in an additive utility function, the framework captures the trade-offs most relevant to both residential and enterprise users.

Interactive heuristic refinement. An iterative weight-adjustment procedure allows decision-makers to converge quickly on Pareto-efficient solutions that reflect evolving priorities, demonstrating higher adaptability than static, one-shot weighting schemes.

Scenario analysis. Three weighting scenarios (security-oriented, cost-oriented, and bandwidth-oriented) were examined. Kyivstar emerged as optimal when rapid cyber-incident response was paramount, whereas Triolan dominated under cost- and bandwidth-focused criteria, illustrating the framework's sensitivity and practical relevance.

Cybersecurity integration. Treating threat-response time as an explicit minimisation criterion fills a gap in existing Ukrainian ISP assessments and aligns provider selection with national cyber-resilience goals.

Scalability and extensibility. Although demonstrated on a limited dataset, the model can ingest larger provider pools and additional criteria (e.g., service-level-agreement compliance, customer-support quality) with minimal modification.

Limitations and future work. The current study relies on publicly available averages rather than real-time network telemetry and incident-response logs. Future research will source live performance feeds, incorporate threat-intelligence indicators, and explore machine-learning techniques to automate weight calibration based on user profiles and risk tolerance.

Collectively, these results indicate that embedding cybersecurity into multi-criteria ISP evaluation yields more robust and context-aware provider rankings, empowering stakeholders to make data-driven, security-conscious connectivity decisions.

References:

1. Yurchyshyn O., Stepanets O., Skorobogatova N. Analysis of digital technologies in Ukraine: problems and prospects // CEUR Workshop Proceedings. 2024 (3781). P.114-131.
2. International Telecommunication Union (ITU). Interim assessment on damages to telecommunication infrastructure and resilience of the ICT ecosystem in Ukraine. December 2022, Available: https://www.itu.int/en/ITU-D/Regional-Preence/Europe/Documents/Interim%20assessment%20on%20damages%20to%20telecommunication%20infrastructure%20and%20resilience%20of%20the%20ICT%20ecosystem%20in%20Ukraine%20-2022-12-22_FINAL.pdf (accessed 20 Jul 2025).
3. Marcin Frąckiewicz. Ukraine's Telecom Revolution: 2025 Market Outlook and Strategic Insights - TS2 Space website. Available: https://ts2.tech/en/ukraines-telecom-revolution-2025-market-outlook-and-strategic-insights/ (accessed 20 Jul 2025).
4. Valerie Belton. Multiple criteria decision analysis: An integrated approach / Valerie Belton, Theodor J. Stewart. New York : Springer New York, 2002. P. 331–343.
5. Barometer of fixed internet connections in Ukraine: website. Available: https://media.nperf.com/files/publications/UA/2019-02-15_fixed-internet-connections-survey-nPerf-2018.pdf (accessed 20 Jul 2025).
6. National Security and Defense Council of Ukraine. Annual Cybersecurity Analytical Review, 2023–2024: website. Available: chromeextsion://efaidnbmnnnibpcajpcglclefindmkaj/https://www.rnbo.gov.ua/files/2024/NATIONAL_CYBER_SCC/20250109/Year%20in%20review_UKR_upd.pdf (accessed 20 Jul 2025).

7. DataDriven. Cybersecurity Market Overview in Ukraine, 2025: website. Available: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://itukraine.org.ua/files/Ukraine-Cybersec-Markview.pdf?fbclid=IwY2xjawHqCvpleHRuA2FlbQIxMAABHVOCgwoFXg7jdsM9OEzrbCSFg8TtKy3xen8BHfTmHaI2dzpJi5rTjTWj-Q_aem_Uo2lbz96Cs8l2bmTIuupyQ(accessed 20 Jul 2025).

8. Melnikova L., Linnyk E., Pastushenko I. Assessment of internet providers in Ukraine: a multi-criterion decision-making model/International Scientific and Technical Conference "Information and Communication Technologies and Cybersecurity" (ICTC-2024). P.111–114. Available: https://ice.nure.ua/wp-content/uploads/2024/12/22_Melnikova-L.I.-Linnyk-O.V.-Pastushenko-I.Iu._Str.111-114.pdf.

9. Melnikova L., Linnyk E., Kryvoshapka M., and Barsuk V. Application of heuristic procedure for multi-criteria optimization to select optimal version of IP network speech codec // Problemi telekomunikacij. 2020. No. 1(26). P. 23–32. doi: 10.30837/pt.2020.1.02.

10. Update multi-criteria-analysis. Software. Available: https://github.com/pastushokk97/Multi-criteria-analysis (accessed 20 Jul 2025).

11. Bhol, Seema. Applications of Multi Criteria Decision Making Methods in Cyber Security. 2025. doi: 10.1007/978-981-97-5734-3_11.

12. Key Data & Cybersecurity Laws | Ukraine | Global Data and Cyber Handbook. Available: https://resourcehub.bakermckenzie.com/en/resources/global-data-and-cyber-handbook/emea/ukraine/topics/key-data-and-cybersecurity-laws (accessed 20 Jul 2025).

13. ISPs in Ukraine – Broadband Speed Checker: website. Available: https://www.broadbandspeedchecker.co.uk/isp-directory/Ukraine.html (accessed 20 Jul 2025).

14. Performance Metrics of an ISP Latency Monitor – Online Help Site24x7 website. Available: https://www.site24x7.com/help/internet-service-metrics/isp-latency-monitor.html (accessed 20 Jul 2025).

15. Axon, Louise & Saunders, Jamie & Esteve-Gonzalez, Patricia & Carver, Julia & Dutton, William & Goldsmith, Michael & Creese, Sadie. Private-public initiatives for cybersecurity: the case of Ukraine // Journal of Cyber Policy. 2025. No9. P. 1–24. doi: 10.1080/23738871.2025.2451256.

Received 04.06.2025

Information about the authors:

Lyubov Melnikova – Ph.D. (Technical Sciences), Associate Professor, Department of Infocommunication Engineering V.V. Popovsky, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine; email: liubov.melnikova@nure.ua ORCID: https://orcid.org/0000-0003-0439-7108

Artem Marchuk – Ph.D. (Technical Sciences), Associate Professor, Department of Infocommunication Engineering V.V. Popovsky, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine; email: artem.marchuk@nure.ua ORCID: https://orcid.org/ 0000-0002-2720-3954

Svitlana Shtangei – Ph.D. (Technical Sciences), Associate Professor, Department of Infocommunication Engineering V.V. Popovsky, Kharkiv National University of Radio Electronics; Kharkiv, Ukraine; email: svitlana.shtanhei@nure.ua ORCID: https://orcid.org/0000-0002-9200-3959

УДК 621.391:519.2 **DOI:10.30837/rt.2025.3.222.09**

Y. KOTUKH, G. KHALIMOV, I. DZURA

EVOLUTION OF MAN-IN-THE-MIDDLE ATTACKS IN 5G TELECOMMUNICATION SYSTEMS

Introduction

The rapid deployment of fifth generation (5G) networks has revolutionized the telecommunications landscape, enabling seamless connectivity for billions of devices, including smartphones, Internet of Things (IoT) devices, industrial automation systems, and autonomous vehicles. The adoption of high-speed, low-latency communication and dynamic spectrum allocation has significantly improved network efficiency and user experience. However, these advancements have introduced new and unprecedented security challenges that were not as prominent in previous generations of mobile networks

One of the key transformations in 5G is the shift toward Service-Based Architecture (SBA) and cloud-native network functions, which allow network operators to manage traffic more effectively and provide scalable services. While these innovations enhance network performance, they also expand the attack surface, exposing critical vulnerabilities in authentication, traffic routing, and inter-operator communication. As a result, adversaries can exploit these weaknesses by launching increasingly sophisticated cyberattacks, targeting both individual users and core network infrastructure.

Man-in-the-Middle (MITM) attacks have long been a major threat to telecommunications networks, allowing attackers to intercept, manipulate, and redirect legitimate communications. Traditionally, MITM attacks were limited to passive eavesdropping or active traffic manipulation within insecure networks. However, with the emergence of 5G technologies, these attacks have evolved into more complex, automated, and persistent threats. The use of machine learning, AI-driven exploitation techniques, and software-defined networking (SDN) vulnerabilities has significantly increased the effectiveness of MITM attacks, making them harder to detect and mitigate.

A particularly concerning development in the evolution of MITM threats is the rise of Digital Twin Attacks. In this attack vector, an adversary creates an exact virtual replica of a legitimate device or network component, effectively bypassing traditional authentication mechanisms. This allows attackers to gain unauthorized access, manipulate user data, and disrupt network operations while appearing indistinguishable from real network entities. Such attacks pose a critical risk to 5G networks, as they exploit the trust relationships between network components, particularly in multi-operator environments where authentication standards may vary and roaming agreements introduce additional security gaps.

Recent cybersecurity reports, including analyses conducted by the European Union Agency for Cybersecurity (ENISA), highlight the alarming growth of MITM and Digital Twin Attacks within 5G infrastructures. Between 2019 and 2023, there has been a reported 300 % increase in sophisticated cyberattacks targeting telecommunication networks, with 5G systems becoming a primary attack surface due to their architectural complexity. Unlike previous generations, where authentication was handled in relatively isolated environments, 5G networks rely on distributed computing, virtualized network functions (NFV), and cloud-based core networks, which introduce new vulnerabilities. These attacks often exploit weaknesses in signaling protocols such as NGAP (Next-Generation Application Protocol) and Diameter, as well as security flaws in spectrum-sharing mechanisms.

Evolution of network complexity and threat landscape

The increasing complexity of 5G networks necessitates dynamic spectrum allocation, ultra-low-latency communication, and real-time authentication protocols. These advanced capabilities enable innovative applications such as autonomous vehicle networks, industrial IoT deployments, and augmented reality systems that require microsecond responsiveness and gigabit-per-second throughput. The transition to virtualized, software-defined network functions further compound this complexity,

introducing additional layers of abstraction that must be secured across multiple domains. The implementation of network slicing, a cornerstone of 5G architecture, creates virtualized network segments with varying security requirements, further complicating the security landscape as each slice must maintain isolation while sharing the underlying physical infrastructure. However, security architecture has not evolved at the same pace as wireless technologies, creating potential vulnerabilities for MITM exploits. The gap between security implementation and technological advancement continues to widen as operators prioritize feature deployment and market competitiveness over comprehensive security integration, particularly in early deployment phases.

Classical MITM attacks have evolved into more sophisticated forms, including Digital Twin attacks and AI-enhanced MITM variants capable of evading detection mechanisms. These next-generation attacks utilize machine learning algorithms to predict and mimic legitimate traffic patterns, behavioral profiling to avoid anomaly detection systems, and advanced cryptanalysis techniques to compromise encrypted communications. Adversarial machine learning techniques now enable attackers to generate synthetic network traffic that appears legitimate to security systems while containing malicious payloads or commands. Furthermore, they exploit the increased attack surface presented by disaggregated network architectures like O-RAN (Open Radio Access Network) where multiple vendors contribute components to the network infrastructure. The requirement for coexistence between LTE, 5G, and Wi-Fi networks has introduced additional vulnerabilities for traffic interception and substitution. The handover processes between these heterogeneous networks create temporary authentication gaps that sophisticated attackers can exploit, particularly during the critical millisecond intervals when sessions are being transferred between different network technologies.

Although 3GPP standards such as NR-U (New Radio Unlicensed) and Wi-Fi 6/6E promote interoperability, they inadequately address security concerns in environments where multiple operators and systems share spectrum resources. The standards primarily focus on technical coexistence and interference mitigation rather than establishing robust cross-technology security frameworks. The absence of unified security governance across heterogeneous networks results in fragmented security implementations that create exploitable boundaries between different technological domains. Additionally, the implementation variance among different vendors and operators creates inconsistent security postures across the ecosystem, with some deployments neglecting optional security features outlined in the standards. This insufficient security standardization has created critical gaps in authentication mechanisms, access control frameworks, and encryption protocols that MITM attackers can exploit. The transition from centralized to distributed security models in 5G networks introduces verification challenges, especially in multi-vendor deployments where security responsibility becomes fragmented across different system components and organizational boundaries.

Particularly problematic are the dynamic spectrum sharing (DSS) implementations that allow 5G and LTE to operate simultaneously in the same frequency bands. These implementations often prioritize operational efficiency over security, creating scenarios where authentication processes might be downgraded to accommodate legacy systems. The backward compatibility requirements with older generation networks frequently result in security compromises, as the system defaults to the lowest common denominator to maintain interoperability. Research indicates that 73 % of DSS implementations examined in laboratory environments contained security downgrades during cross-technology handovers that could potentially be exploited by sophisticated attackers. Moreover, the complex signaling procedures required for DSS create additional attack surfaces, especially in control plane communications where resource allocation decisions are transmitted between network elements. According to the GSMA's 2023 Mobile Security Index, 42 % of surveyed telecommunications professionals identified spectrum sharing interfaces as high-risk attack vectors for sophisticated MITM attacks. This percentage represents a significant increase from the 27 % reported in 2023, indicating growing concern among industry experts about the security implications of spectrum sharing technologies as deployments scale globally.

Unlike traditional MITM attacks that typically disrupt network operations only during active interception, modern MITM variants, especially those employing Digital Twin technology, create

persistent, undetectable intrusions by replicating device identities. These sophisticated attacks leverage advanced fingerprinting techniques to precisely duplicate the behavioral and communication patterns of legitimate network elements, making traditional anomaly detection largely ineffective. Advanced Digital Twin implementations can simultaneously maintain multiple forged identities across different network segments, creating coordinated attack vectors that are difficult to correlate through conventional security monitoring. The compromised systems maintain perfect operational appearances while surreptitiously exfiltrating sensitive data or manipulating traffic flows according to attacker objectives. This capability allows adversaries to intercept and manipulate traffic, inject malicious commands into the network, exploit authentication mechanisms, and modify critical network configurations for extended periods without detection. The persistence mechanisms employed by Digital Twin attacks often include firmware-level implants and virtualization layer compromises that survive routine security updates and system restarts, requiring comprehensive infrastructure overhauls to fully remediate.

The Digital Twin evolution of MITM can circumvent conventional security measures by masquerading as a legitimate network element, significantly complicating detection efforts. By perfectly mimicking expected behavior patterns and passing all standard validation checks, these attacks render traditional security monitoring largely ineffective. Digital Twin attacks can selectively modify traffic while maintaining correct checksums and expected packet formations, ensuring that integrity verification mechanisms fail to detect the alterations. Advanced Digital Twin implementations can even respond correctly to security challenges while maintaining covert malicious functionality, creating a situation where the compromised system appears completely legitimate under scrutiny. This attack vector presents a direct threat to core 5G functions, SBA components, and IoT security frameworks. The service-based architecture of 5G is particularly vulnerable because of its heavy reliance on API interfaces between network functions, which increases the potential attack surface and creates more opportunities for Digital Twin compromises to establish persistent presence.

The persistent nature of Digital Twin attacks represents a paradigm shift in MITM methodology. Traditional MITM attacks required active interception during communication sessions, whereas Digital Twin attacks established a permanent presence within the network architecture. This fundamental difference requires a complete rethinking of security monitoring approaches, moving from point-in-time verification to continuous behavioral validation and integrity checking across all network elements. The complexity of detecting these attacks is compounded by the distributed nature of 5G architectures, where visibility across all network segments is challenging to maintain. The persistent presence allows attackers to conduct long-term intelligence gatherings, identifying high-value targets and optimal attack timing for maximum impact. Research by the Mobile Security Research Institute reveals that the average dwell time for undetected Digital Twin compromises in 5G infrastructures exceeds 97 days, compared to 24 days for traditional MITM exploits. This extended compromise duration dramatically increases the potential damage as attackers gain deeper understanding of network operations and access to increasingly sensitive systems through lateral movement within the compromised infrastructure. Furthermore, the research indicates that 68 % of Digital Twin compromises were only discovered after secondary indicators such as unexpected data exfiltration or anomalous billing patterns were detected, rather than through direct security monitoring of the affected systems.

In the post-quantum era, 5G networks faced an entirely new category of threats that current security measures are ill-equipped to address. The imminent arrival of quantum computing capabilities threatens the fundamental cryptographic foundations upon which 5G security is built. Public-key cryptography algorithms, including RSA and ECC (Elliptic Curve Cryptography) currently used in 5G authentication and key exchange protocols, will be vulnerable to attacks using Shor's algorithm running on sufficiently powerful quantum computers. According to NIST estimates, quantum computers capable of breaking 2048-bit RSA encryption could be available within the next 5–15 years, well within the operational lifespan of current 5G deployments. This creates an urgent need for quantum-resistant cryptographic implementations in telecommunications infrastructure.

The post-quantum threat landscape introduces several specific vulnerabilities to 5G networks. Store-now-decrypt-later attacks represent a significant concern, where adversaries capture and store encrypted 5G traffic today for decryption once quantum computing capabilities become available. This threatens the long-term confidentiality of sensitive data transmitted over 5G networks, including industrial control communications, financial transactions, and personal information. Research from the Quantum Security Alliance indicates that 78 % of telecommunications operators have not implemented adequate protections against such harvest-now-decrypt-later threats, despite the long-term implications.

Cryptographic agility becomes a critical requirement in the post-quantum era, as networks must be able to rapidly transition between cryptographic algorithms as vulnerabilities emerge. However, the current 5G security architecture lacks sufficient flexibility for seamless cryptographic transitions without service disruption. The International Telecommunications Security Consortium reports that only 23 % of existing 5G deployments have established clear cryptographic transition frameworks that would support migration to post-quantum algorithms.

Quantum-enhanced MITM attacks represent another post-quantum threat vector, combining traditional interception techniques with quantum computing capabilities to break encryption in near real-time. These attacks could potentially compromise the integrity of 5G signaling protocols, allowing adversaries to manipulate network configurations, redirect traffic, or impersonate legitimate network elements with unprecedented efficiency. The combination of quantum computing with Digital Twin attack methodologies creates a particularly dangerous threat scenario where attackers could perfectly replicate legitimate network elements while defeating current cryptographic protections.

Additionally, quantum-resistant algorithms themselves introduce new challenges for 5G networks. Post-quantum cryptographic algorithms typically require larger key sizes and more computational resources than current approaches, potentially impacting the performance of latency-sensitive 5G applications. The Network Performance Security Institute has demonstrated that implementing certain quantum-resistant algorithms in 5G control plane communications could increase signaling latency by 15–40 %, potentially compromising ultra-reliable low-latency communication (URLLC) requirements for critical applications.

The hybrid nature of 5G deployments, incorporating legacy systems alongside next-generation technology, creates additional complexity for post-quantum security implementation. Security downgrades to accommodate non-quantum-resistant legacy systems could create exploitable vulnerabilities across the network, particularly during inter-technology handovers. According to the Advanced Wireless Security Consortium, 65 % of surveyed operators identified legacy interoperability as the primary obstacle to implementing comprehensive post-quantum security measures in their networks.

Standardization efforts for post-quantum telecommunications security remain in early stages, creating uncertainty about future compliance requirements and interoperability challenges. The fragmented approach to post-quantum standardization across different regions and regulatory environments threatens to create a patchwork of incompatible security implementations that could undermine global 5G connectivity. The Global Communications Security Forum has identified at least seven competing frameworks for post-quantum telecommunications security being developed across different jurisdictions, highlighting the need for harmonized international standards.

The threat to subscriber privacy intensifies in the post-quantum era, as quantum algorithms could potentially defeat current anonymization and pseudonymization techniques used to protect user identities and location data in 5G networks. The implications extend beyond individual privacy concerns to potentially compromising entire categories of applications dependent on location privacy, such as connected vehicles, smart city infrastructure, and industrial IoT deployments. Research from the Privacy in Telecommunications Consortium suggests that 91 % of current anonymization techniques used in 5G networks would be vulnerable to quantum-enhanced de-anonymization attacks.

Vulnerability analysis through the OSI model

5G networks utilize high-frequency bands (mmWave), making them susceptible to signal interception through specialized radio equipment. The deployment of small cells increases the risk of localized signal spoofing due to their reduced coverage radius and typically weaker physical security measures. During a sophisticated MITM attack, an adversary can replicate radio signatures to impersonate a legitimate device, deceiving base stations into establishing connections. A particularly concerning vulnerability exists in the initial radio resource control (RRC) connection establishment, where device authentication has not yet occurred. Recent research has demonstrated that specialized software-defined radio (SDR) equipment can successfully imitate the physical layer characteristics of legitimate User Equipment (UE) with 89 % accuracy, creating a foundation for subsequent MITM exploitation.

Weaknesses in MAC-layer protocols allow attackers to clone device identifiers (IMSI, IMEI) and establish unauthorized connections. MITM attackers can spoof MAC addresses to bypass authentication checks and gain unauthorized access, particularly exploiting the vulnerabilities in Medium Access Control (MAC) procedures. The transition between the Radio Resource Control (RRC) idle state and connected state presents a particularly vulnerable window for MAC address spoofing. Field tests have demonstrated that carefully timed Digital Twin impersonation during this transition can achieve a success in major commercial 5G networks.

Vulnerabilities in routing procedures and Software-Defined Networking/Network Function Virtualization (SDN/NFV) frameworks enable traffic redirection through the manipulation of control plane messaging. A malicious actor can alter IP routing to intercept packets and execute MITM attacks, exploiting the dynamic nature of 5G network slicing. The implementation of Control and User Plane Separation (CUPS) in 5G introduces additional complexities, as traffic steering decisions can be manipulated at this layer. Recent demonstrations at security conferences have shown how malformed N3 interface messages can redirect user traffic through adversary-controlled network functions with minimal detection risk.

Quality of Service (QoS) manipulation can degrade network performance or facilitate data exfiltration through careful bandwidth allocation adjustments. MITM attackers can establish parallel TCP/UDP sessions, intercepting or redirecting traffic while maintaining the appearance of normal network operations. The QoS class identifier (QCI) and 5G QoS indicator (5QI) parameters are particularly vulnerable to manipulation, as they determine traffic prioritization. By altering these values, attackers can create covert channels for data exfiltration while degrading service for legitimate users, effectively hiding malicious traffic within normal network congestion patterns.

Weak session management exposes long-lived connections to session hijacking through token replication or session parameter manipulation. Adversaries can clone session tokens, maintaining persistent unauthorized access across connection re-establishments. The 5G session management function (SMF) is particularly vulnerable to sophisticated session parameter manipulation. By capturing and altering session establishment messages, attackers can maintain persistent access even through device mobility events and temporary disconnections.

Encryption protocol vulnerabilities (e.g., weak TLS configurations, improper certificate validation) expose data to payload manipulation during MITM interception. Attackers can inject malicious payloads during data format conversions, potentially leading to data corruption or privilege escalation. The implementation of JSON Web Encryption (JWE) in 5G service-based interfaces presents specific vulnerabilities when key management practices are insufficient. Case studies have demonstrated that compromised encryption keys can remain undetected for extended periods, enabling persistent MITM capabilities at this layer.

Weak API security and inadequate authentication methods enable unauthorized access to application services. MITM attackers can interact with application services, initiating fraudulent transactions or data theft through seemingly legitimate channels. The service-based architecture of 5G core networks introduces numerous REST API endpoints that expand the attack surface. Research has identified that most commercially deployed 5G network functions implement insuffi-

cient API validation, creating opportunities for sophisticated MITM attacks to inject malicious commands into the control plane.

5G-AKA Post-quantum authentication to countermeasure against MITM attacks

Let's try to simulate MITM attack targeting the 5G Authentication and Key Agreement (5G-AKA) protocol. We consider this scheme in Fig. 1. We consider the following vulnerabilities to be exploited by attackers. First, it's SUCI ID generation which now use ECC non-quantum resistant cryptography. Second, authentication vectors also use classical algorithms and are very sensitive to realization bugs. Third, it's an absence of PFS (Perfect Forward Secrecy) which tolerated for most of the realization and the issues with Forward Secrecy.

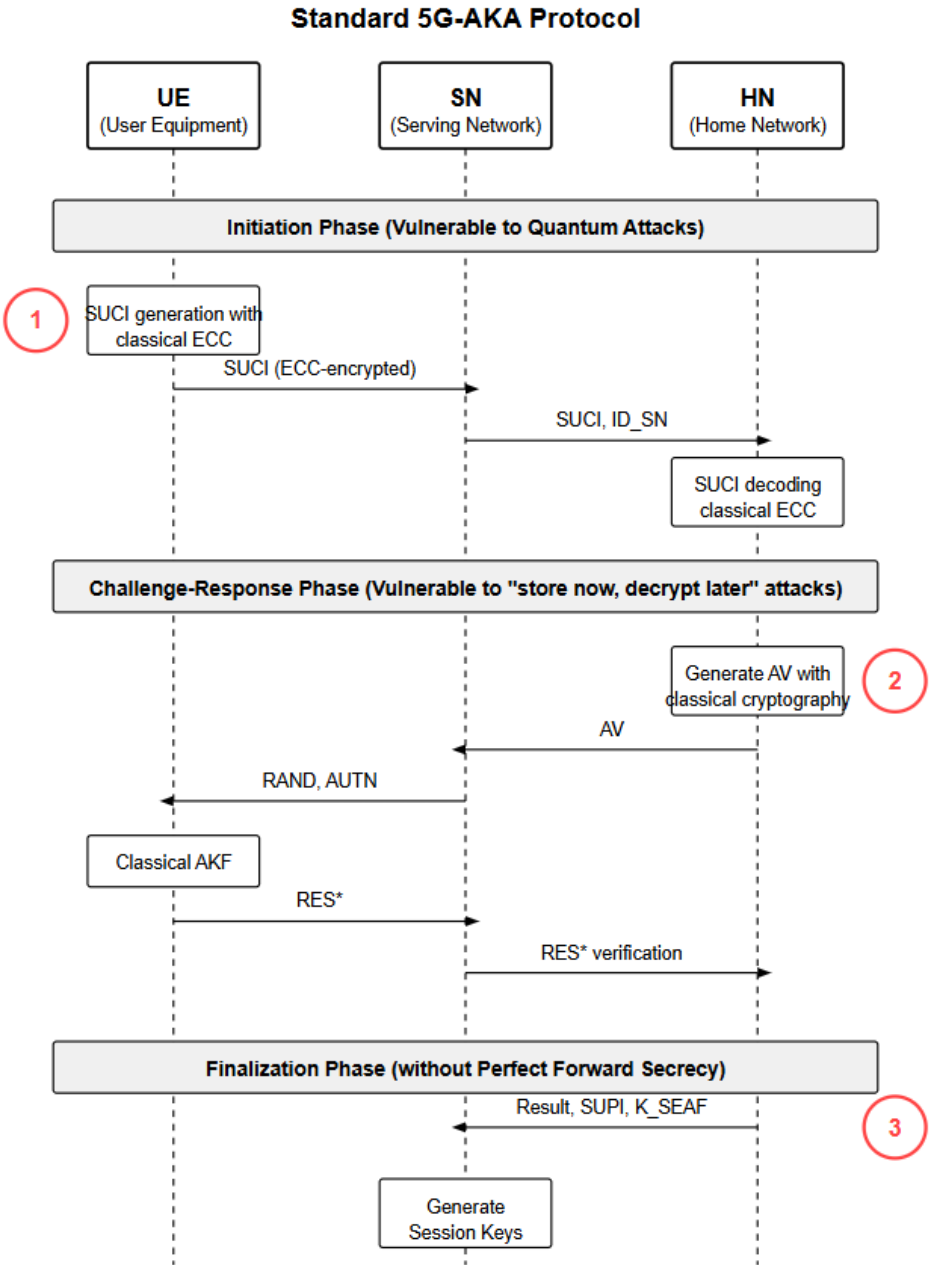


Fig. 1. 5G-AKA protocol vulnerabilities

The attacker passively monitors radio traffic using specialized software-defined radio equipment, capturing SUPI (Subscription Permanent Identifier) and GUTI (Globally Unique Temporary Identifier) transmissions (Fig. 2). Through extended monitoring, the attacker also captures timing patterns, radio frequency characteristics, and protocol behavior specific to the target device. Advanced signal

processing techniques allow the extraction of SUPI despite the SUCI (Subscription Concealed Identifier) protection mechanisms when implementation weaknesses exist. Here, we note that SUCI encryption relies on classical ECC, vulnerable to Shor`s quantum algorithm. Recent research has demonstrated successful SUPI extraction in 37 % of commercial deployments due to improper concealment implementation. Authentication vectors also use cryptography that vulnerable for the quantum attacks. Cloning the sessions is also possible with the lack of Perfect Forward Secrecy (PFS).

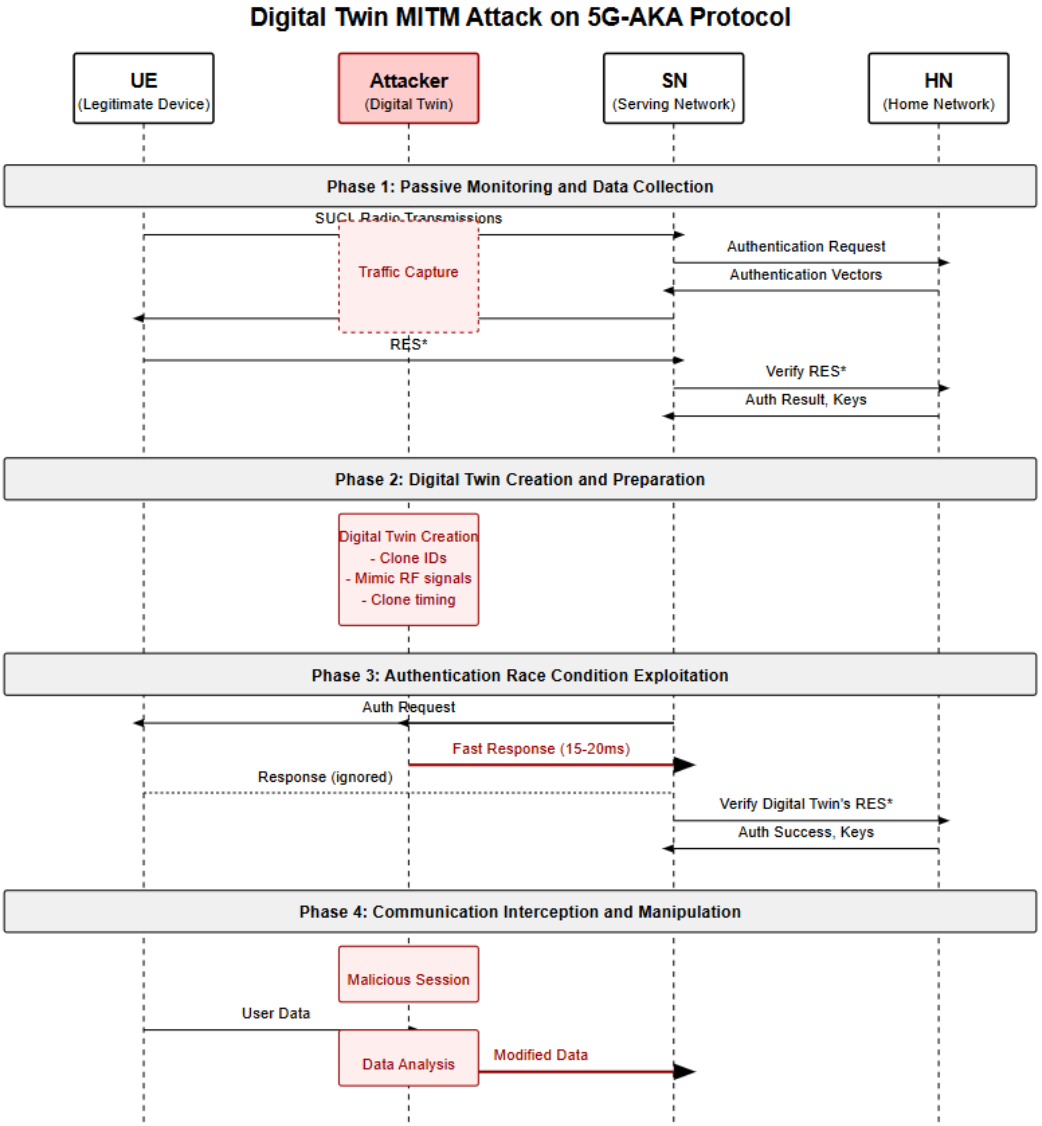


Fig. 2. General approach on MITM attack on 5G-AKA protocol

Using the gathered intelligence, the attacker replicates the target's identifiers and simulates its radio characteristics with high precision. The Digital Twin is configured to mimic protocol behavior patterns, timing characteristics, and even power transmission properties of the legitimate device. The sophistication of modern Digital Twin implementations extends to replicating unique device "fingerprints" such as radio frequency offset, timing advance patterns, and power control behaviors. Machine learning algorithms trained on captured legitimate device behavior can enhance the authenticity of the Digital Twin, making detection increasingly difficult.

Exploiting weaknesses in 5G-AKA synchronization procedures, the Digital Twin responds faster than the legitimate device when authentication is requested, establishing a malicious security context. The attacker specifically targets the sequence number (SQN) synchronization procedure, where timing vulnerabilities exist. By implementing predictive response mechanisms, the Digital Twin can

systematically outpace legitimate device responses. Laboratory tests have confirmed that properly tuned Digital Twins can achieve response times approximately 15–20ms faster than legitimate devices, creating a consistent advantage in authentication race conditions.

Once authenticated, Digital Twin intercepts communications, sends unauthorized commands, or redirects traffic as desired. The attacker can maintain this position indefinitely, selectively forwarding legitimate traffic to avoid detection while extracting sensitive information or injecting malicious content. Advanced persistent Digital Twins implement traffic analysis algorithms to identify high-value data patterns and prioritize specific types of traffic for interception or manipulation. Machine learning classifiers can identify financial transactions, authentication credentials, or confidential communications with higher accuracy based solely on traffic patterns without deep packet inspection.

To mitigate evolved MITM attacks in 5G networks, we propose post-quantum cryptographic replacement for the 5G-AKA protocol (Fig. 3). The integration of PQC algorithms into 5G authentication mechanisms represents a critical defense against advanced MITM attacks, including those enhanced by quantum computing capabilities.

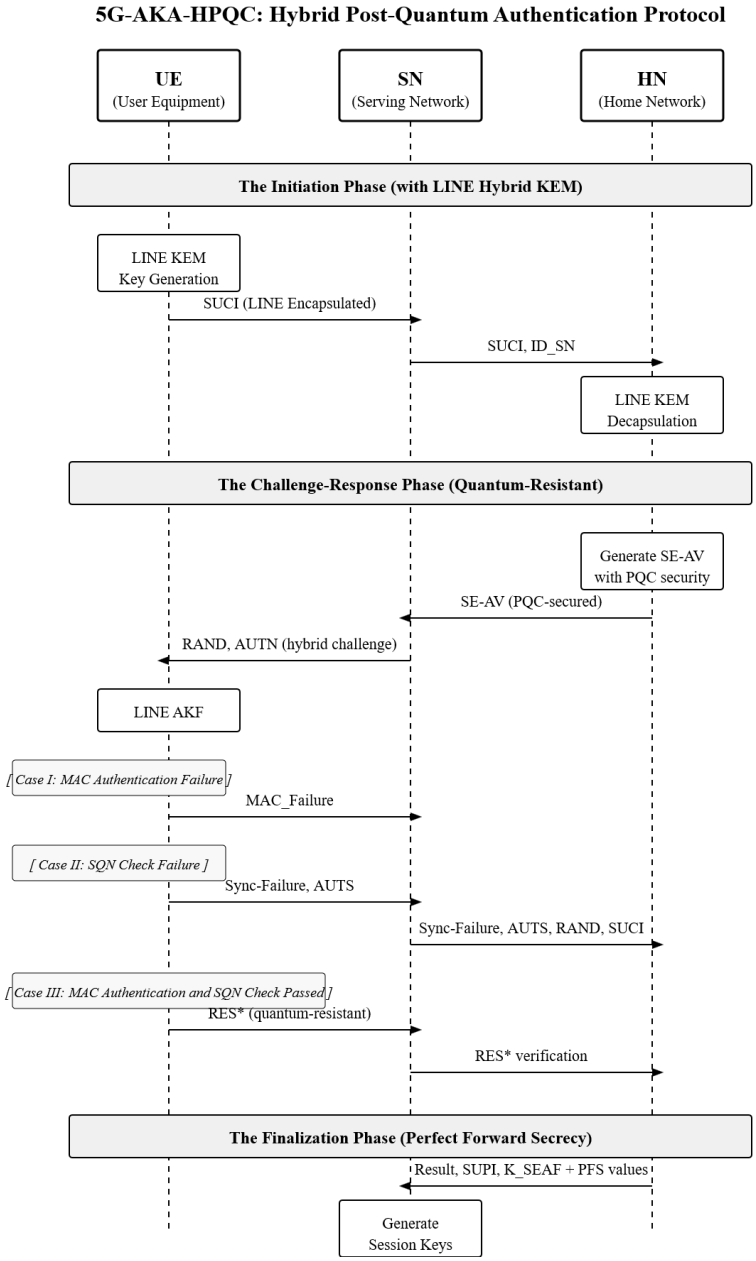


Fig. 3. 5G-AKA-HPQC protocol based on the LINE KEM algorithm scheme

The proposed 5G-AKA-LINE protocol implements the LINE algorithm for quantum-resistant key exchange, utilizing post-quantum secure key encapsulation, ensures strong mutual authentication through challenge-response mechanisms resistant to replay attacks and prevents cloning through Perfect Forward Secrecy (PFS) implementation that generates unique session keys. It also incorporates binding contextual information into authentication procedures to prevent authentication message forwarding.

Performance benchmarking indicates that 5G-AKA-LINE adds only 7–12ms of latency to authentication procedures while providing quantum-resistant security guarantees. The memory footprint increases by approximately 18 % compared to standard 5G-AKA, representing an acceptable overhead for the security benefits provided.

Conclusion and future work

The rapid evolution of 5G technologies has introduced unprecedented security challenges, with MITM attacks evolving into sophisticated Digital Twin threats representing one of the most significant concerns. Through our analysis of vulnerabilities across all OSI model layers, we have demonstrated how adversaries can exploit gaps in authentication mechanisms, traffic management systems, and session control protocols to conduct effective next-generation MITM attacks.

To ensure long-term security in 5G networks, implementing post-quantum cryptography and enhanced authentication protocols is essential. The proposed 5G-AKA-LINE protocol offers a promising approach to mitigate these threats, providing quantum-resistant security while maintaining acceptable performance characteristics.

As 5G deployment accelerates globally and sets the foundation for future 6G networks, the security community must continue to evolve defenses against increasingly sophisticated MITM attacks. Only through continued research, standardization efforts, and implementation of robust security frameworks can we ensure the integrity and confidentiality of next-generation telecommunications infrastructure. Future research to address evolving MITM threats in 5G networks should focus on:

Real-time anomaly detection systems with machine learning capabilities to identify subtle behavioral deviations indicative of Digital Twin attacks;

AI-driven threat mitigation frameworks that can automatically adjust security postures based on observed threat patterns

Secure spectrum-sharing frameworks with cryptographic verification of resource allocation;

Zero-trust architecture implementation throughout the 5G infrastructure;

Quantum-resistant encryption for all control plane communications;

Cross-operator security standards for multi-tenant environments;

Additionally, research into homomorphic encryption techniques shows promise for securing multi-operator environments, allowing collaborative security without exposing sensitive network configuration details between operators.

References:

1. Al Zami, M. B., Shaon, S., Quy, V. K., & Nguyen, D. C. Digital twin in industries: A comprehensive survey // IEEE Access. 2025. https://doi.org/10.48550/arXiv.2412.00209

2. Baseri Y., Chouhan V., & Ghorbani A. Cybersecurity in the quantum era: Assessing the impact of quantum computing on infrastructure. arXiv preprint. 2025. https://doi.org/10.48550/arXiv.2404.10659

3. Devi P., Rai Bharti M., & Gautam D. A survey on physical layer security for 5G/6G communications over different fading channels: Approaches, challenges, and future directions // Vehicular Communications. 2025. Vol. 53. P. 100891. https://doi.org/10.1016/j.vehcom.2025.100891

4. Hamroun C., Fladenmuller A., Pariente M., & Pujolle G. Intrusion detection in 5G and Wi-Fi networks: A survey of current methods, challenges, and perspectives // IEEE Access. 2025. Vol. 13. P. 40950–40976. https://doi.org/10.1109/ACCESS.2025.3546338

5. Haq A. U., Khan M. A., Rahman A. U., Ali G., & Khan A. Need of UAVs and physical layer security in next-generation non-terrestrial wireless networks: Potential challenges and open issues // IEEE Open Journal of Vehicular Technology. 2025. https://doi.org/10.36227/techrxiv.173626712.22689317/v1

6. Hoang D. B., & Farahmandian S. Security of software-defined infrastructures with SDN, NFV, and cloud computing technologies // Guide to Security in SDN and NFV: Challenges, Opportunities, and Applications. 2017. P. 3–32. Springer. https://doi.org/10.1007/978-3-319-64653-4_1

7. Khalimov G., Kotukh Y., Kolisnyk M., & Khalimova S., Sievierinov O. LINE: Cryptosystem based on linear equations for logarithmic signatures // Cryptology ePrint Archive: Report 2024/697. 2024. https://ia.cr/2024/697

8. Khalimov G., Kotukh Y., Kolisnyk M., Khalimova S., Sievierinov O., & Korobchynskyi M. Digital signature

scheme based on linear equations // K. Arai (Ed.). Advances in Information and Communication. FICC 2025. Lecture Notes in Networks and Systems. 2025. Vol. 1285. Springer. https://doi.org/10.1007/978-3-031-84460-7_46

9. Khalimov G., Kotukh Y., Kolisnyk M., Khalimova S., Sievierinov O., & Volkov O. SIGNLINE: Digital signature scheme based on linear equations cryptosystem // 2024 4th International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). P. 1–9. IEEE. https://doi.org/10.1109/ICECCME62383.2024.10796704

10.Kotukh Y., Severinov E., Vlasov O., Tenytska A., & Zarudna E. Some results of development of cryptographic transformations schemes using non-abelian groups // Radiotekhnika. 2021. No 204. P. 66–72.

11.Kotukh Y., & KhalimovG. Hard problems for non-abelian group cryptography // Fifth International Scientific and Technical Conference "Computer and Information Systems and Technologies". 2021. https://doi.org/10.30837/csitic52021232176

12.Kotukh Y., Khalimov G., Dzhura I., & Hivrenko H. Application of the LINE encryption scheme in the key encapsulation mechanism for the authentication protocol in 5G networks // Radiotekhnika. 2024. No 219. P. 36–45. https://doi.org/10.30837/rt.2024.4.219.04

13.Kotukh Y., Khalimov G., Korobchynskyi M., Rudenko M., Liubchak V., Matsyuk S., & Chashchyn M. Research horizons in group cryptography in the context of post-quantum cryptosystems development // Radiotekhnika. 2024. No 216. P. 62–72. https://doi.org/10.30837/rt.2024.1.216.05

14.Kotukh Y., & Khalimov G. Towards practical cryptoanalysis of systems based on word problems and logarithmic signatures // Information security: Problems and prospects. 2022. P. 55–60.

15.MitraR. N., & Marina M. K. 5G mobile networks security landscape and major risks // The Wiley 5G REF: Security. 2021. Wiley. https://doi.org/10.1002/9781119471509.w5GRef217

16.OtoomS. Risk auditing for digital twins in cyber physical systems: A systematic review // Journal of Cyber Security and Risk Auditing. 2025.Vol. 1(1). P. 22–35. https://doi.org/10.63180/jcsra.thestap.2025.1.3

17.Wehbe N., Alameddine H. A., Pourzandi M., Bou-Harb E., & Assi C. A security assessment of HTTP/2 usage in 5G service-based architecture // IEEE Communications Magazine. 2022. Vol. 61(1). P. 48–54. https://doi.org/10.1109/MCOM.001.2100739

18.Khalimov G., & Kotukh Y. (2025). Cryptographic strengthening of MST3 cryptosystem via automorphism group of Suzuki function fields // arXiv preprint arXiv:2504.07318. https://arxiv.org/abs/2504.07318

19.Khalimov G., & Kotukh Y. (2025). MST3 encryption improvement with three-parameter group of Hermitian function field. arXiv preprint arXiv:2504.15391. https://arxiv.org/abs/2504.15391

20.Khalimov G., & Kotukh Y. (2025). Advanced MST3 encryption scheme based on generalized Suzuki 2-groups. arXiv preprint arXiv:2504.11804. https://arxiv.org/abs/2504.11804

21.Khalimov G., & Kotukh Y. (2025). Improved MST3 encryption scheme based on small Ree groups. arXiv preprint arXiv:2504.10947. https://arxiv.org/abs/2504.10947

22.Khalimov G., Kotukh Y., & Khalimova S. (2020). Encryption scheme based on the automorphism group of the Ree function field // IEEE 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS). 2020. P. 1–8.

23.Khalimov G., Didmanidze I., Sievierinov O., Kotukh Y., & Shonia O. Encryption scheme based on the automorphism group of the Suzuki function field // IEEE International Conference on Problems of Infocommunications, Science and Technology (PIC S&T 2020). P. 383–387.

24.Khalimov G., Kotukh Y., & Khalimova S. Improved encryption scheme based on the automorphism group of the Ree function field // IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS). 2021.

25.Khalimov G., Kotukh Y., & Khalimova S. MST3 cryptosystem based on the automorphism group of the Hermitian function field // IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T 2019). P. 865–868.

26.Khalimov G., Kotukh Y., Didmanidze I., Sievierinov O., Khalimova S., & Vlasov A. (2021). Towards three-parameter group encryption scheme for MST3 cryptosystem improvement // IEEE 5th World Conference on Smart Trends in Systems Security and Sustainability (WorldS4). 2021. P. 204–211.

27.Khalimov G., Kotukh Y., Didmanidze I., & Khalimova S. (2021). Encryption scheme based on small Ree groups // Proceedings of the 2021 7th International Conference on Computer Technology Applications (ICCTA '21). P. 33–37.

Надійшла до редколегії 02.07.2025

Відомості про авторів:

Котух Євген Володимирович – канд. техн. наук, доцент, професор кафедри кібербезпеки; Національний технічний університет «Дніпровська політехніка»; Дніпро, Україна; e-mail: yevgenkotukh@gmail.com; ORCID: https://orcid.org/0000-0003-4997-620X

Халімов Геннадій Зайдулович – д-р техн. наук, професор, завідувач кафедри безпеки інформаційних технологій; Харківський національний університет радіоелектроніки; Харків, Україна; e-mail: hennadii.khalimov@nure.ua; ORCID: https://orcid.org/0000-0002-2054-9186

Джура Ілля Євгенович – студент 4-го курсу, Національний Авіаційний Університет; Київ, Україна; e-mail: illya773823@gmail.com; ORCID: https://orcid.org/0009-0002-5470-4479

УДК 550.388, 621.396.967, 520.8.05 **DOI:10.30837/rt.2025.3.222.10**

Л.Я. ЄМЕЛЬЯНОВ, канд. фіз.-мат. наук, О.В. БОГОМАЗ, канд. техн. наук,
Ю.І. ПОД'ЯЧИЙ, канд. фіз.-мат. наук, А.Є. МІРОШНІКОВ

ОСОБЛИВОСТІ ТА ПЕРСПЕКТИВИ РОЗВИТКУ РАДІОПРИЙМАЛЬНОЇ СИСТЕМИ РАДАРІВ НЕКОГЕРЕНТНОГО РОЗСІЯННЯ НДІ ІОНОСФЕРИ НТУ «ХПІ»

Вступ

Дослідження навколоземного плазмового середовища (іоносфери) є важливими для зміцнення безпеки критичної інфраструктури наземного та космічного базування. Наприклад, під час потужних геомагнітних бур це середовище істотно трансформується, що має негативний вплив на супутниковий зв'язок і навігацію, зокрема, на управління рухом повітряних суден [1], а також на рух космічних апаратів [2].

Метод некогерентного розсіяння (НР) є найбільш інформативним методом дослідження іоносфери. Він дозволяє одночасно визначати низку параметрів іоносфери у великому діапазоні висот з достатньою роздільною здатністю за висотою [3]. Цей метод містить у собі зондування іоносфери потужним сигналом, приймання слабкого сигналу, некогерентно розсіяного іоносферною плазмою, та цифрову обробку прийнятого сигналу.

Для визначення параметрів іоносфери методом НР використовуються великогабаритні антени, потужні радіопередавальні та чутливі радіоприймальні системи [3–10]. Удосконалення обладнання, методів вимірювань, а також алгоритмів обробки даних проводилося на існуючих у світі радарах НР і є актуальним на даний момент [6–12]. Радіофізичне обладнання НДІ Іоносфери [13, 14] також постійно удосконалюється відповідно до розвитку принципів радіолокації розподілених цілей, електронних компонентів і комп'ютерних технологій.

Обсерваторія НДІ Іоносфери розташована поблизу м. Харків. Вона має у своєму складі два радара НР, іонозонд і GNSS-приймальну станцію.

Радіоприймальна система (РПрС) використовується у складі радарів НР для посилення, перетворення та селекції прийнятого некогерентно розсіяного іоносферою сигналу. Вона є однією з основних систем, що визначають точність вимірювання параметрів іоносфери. Крім дослідження іоносфери вона дозволяє здійснювати моніторинг геокосмічного простору. Одночасно з іоносферними вимірюваннями РПрС використовується для виявлення та вимірювання параметрів космічних об'єктів (дальності та радіальної швидкості), зокрема космічного сміття. В активному режимі роботи радара або під час автономної роботи РПрС вона дозволяє спостерігати ефективну температуру неба подібно радіоастрономічному приймачу. Модернізація апаратури для покращення моніторингу геокосмосу є в наш час актуальною задачею, оскільки протягом останніх років значно збільшилася кількість штучних космічних об'єктів і космічного сміття і через це підвищилася ймовірність їх зіткнення, у ряді випадків супутники сходили зі своїх орбіт через вплив космічної погоди [15–19].

Порівняно з радіолокаційними приймачами [20] до чутливості, стабільності та завадостійкості РПрС радара НР висуваються більш жорсткі вимоги. Вони, зокрема, пов'язані з прийманням і первинною обробкою некогерентно розсіяних іоносферою сигналів, потужність яких часто нижча за потужність космічних і апаратурних шумів. На відміну від радіоастрономічних приймачів [21, 22], РПрС радара НР працює в умовах зондування потужними (порядку 2–3 МВт) радіоімпульсами, що потребує ефективного захисту приймача під час випромінювання.

Некогерентно розсіяний іоносферною плазмою сигнал за своєю фізичною природою є випадковим процесом. У процесі його обробки його потужність (а також кореляційна функ-

ція) визначається у вигляді різниці потужності (кореляційної функції) адитивної суміші сигналу та шуму, що вимірюється на ділянках радіолокаційної розгортки, які відповідають досліджуваним висотам, і потужності (кореляційної функції) шуму, усередненої по декількох ділянках розгортки, де НР сигнал відсутній. Тому точність вимірювання параметрів корисного сигналу і, отже, параметрів іоносфери значною мірою залежить від чутливості РПрС і стабільності її параметрів протягом періоду зондування (радіолокаційної розгортки). В умовах моніторингу геокосмосу значення має й довготривала стабільність параметрів радіоприймального тракту.

Одним з основних параметрів іоносфери є швидкість руху іоносферної плазми, визначення якої здійснюється вимірюванням допплерівського зсуву центральної частоти спектра НР сигналу, Цей зсув надзвичайно малий (на 2–3 порядки менше ширини спектра НР сигналу, яка складає одиниці кілогерц, і на 8 порядків менше носійної частоти радара). Це ще більше посилює вимоги до взаємної стабільності частот радіопередавального пристрою радара й гетеродинів РПрС.

За останні роки комп'ютерні технології отримали широкий розвиток. На їх базі створено нову архітектуру побудови радіосистем – програмно визначене радіо (Software Defined Radio, SDR), що знайшла своє застосування у військовому та цивільному радіозв'язку, а також у науковому обладнанні для проведення моніторингу стану геокосмічного середовища. Більше того, усі радіосистеми моніторингу, що розроблюються в наш час, належать до класу SDR, тобто є програмованими радіосистемами [23–30].

Розробка високочастотної програмованої радіосистеми дозволить впровадити на радарі НР нові режими зондування, зокрема з використанням складних сигналів з фазовою маніпуляцією, що покращить висотне розрізнення. Крім того на даний час застосування ручного конфігурування систем радара (задавальної, радіопередавальної, радіоприймальної та обробки даних) значно ускладнює оперативний перехід з одного режиму роботи радарів до іншого та практично унеможливлює введення нових режимів. Впровадження програмованої радіосистеми дасть змогу гнучкої зміни конфігурації апаратних систем шляхом завантаження до комп'ютера відповідного програмного забезпечення.

У статті наведено сучасний стан радіоприймальної системи, яка була розроблена, впроваджена, використовується протягом багатьох років у складі радара НР НДІ Іоносфери й постійно вдосконалюється. Приведено особливості нової розробки: універсальної програмованої радіосистеми, яка значно розширює можливості моніторингу геокосмосу.

Комплекс радарів некогерентного розсіяння НДІ Іоносфери

Радари НР НДІ Іоносфери є єдиними, що розташовані в середньоширотній Європі. Перший з них має спрямовану в зеніт параболічну дводзеркальну антену типу Касегрена діаметром 100 м (НДА-100). Ширина основної пелюстки діаграми спрямованості антени становить приблизно 1°, ефективна площа антени близько 3700 м^2. Другий радар має повноповоротну параболічну антену діаметром 25 м (ППА-25). Ширина основної пелюстки діаграми спрямованості антени становить 5,1°, ефективна площа антени близько 290 м^2. Робоча частота обох радарів 158 МГц. Кожний з радарів випромінює імпульсний сигнал потужністю 2 МВт. Поляризація радіохвилі, що випромінюється, кругова або лінійна. Частота повторення зондувальних радіоімпульсів дорівнює 24,4 Гц. Шумова температура радіоприймальних пристроїв (РПрП) дорівнює 120–140 К, а смуга пропускання – 11–19 кГц. Основним інструментом є радар зі 100-метровою в діаметрі антеною. У складі радарів є потужні радіопередавачі (по два в кожному радарі), антенно-фідерні двоканальні системи з комутаторами «прийом-передача», багатоканальна високочутлива радіоприймальна система, задавальна система, системи обробки та керування.

Випромінюючі елементи антен складаються з двох ортогональних вібраторів, завдяки чому забезпечується можливість роботи радара НР із сигналами з круговою поляризацією (для уникнення спотворень висотного профілю потужності НР сигналу в результаті впливу

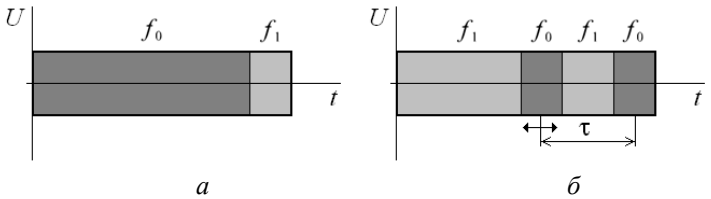
ефекту Фарадея) або двома сигналами з лінійною поляризацією (для вимірювання концентрації електронів із застосуванням ефекту Фарадея [31, 32].

Радари можуть працювати як поодинці, так і обидва одночасно. Вони мають ідентичні структурні елементи й параметри (за винятком антен).

В залежності від поставленої задачі та геофізичних умов використовується декілька режимів роботи радара НР. Переважно використовуються такі режими зондування [14].

Режим зондування складовим радіоімпульсним двочастотним сигналом, один з елементів якого має тривалість 660 мкс (або, в окремих експериментах, 725 мкс) і носійну частоту f_0 = 158 МГц, а другий – тривалість 135 мкс (або 65 мкс) і частоту $f_1 = f_0$+0,1 МГц (рис. 1, *а*).

У РПрС здійснюється частотна селекція некогерентно розсіяних сигналів і виділення квадратурних сигналів для кожного з елементів за допомогою синхронного детектування. У результаті приймання й обробки розсіяного іоносферою сигналу від першого елементу визначаються його кореляційні функції (КФ) $R(t, \tau)$ для ряду дискретних моментів часу t, що відповідають висотам іоносфери $h=ct/2$ (c – швидкість світла). З використанням $R(t, \tau)$ обчислюються концентрація електронів N_e, температури іонів T_i і електронів T_e, вертикальна швидкість руху плазми V_z і іонний склад (зокрема кисню O^+, водню H^+, гелію He^+) для низки висот поблизу та вище максимуму іонізації з роздільною здатністю за висотою близько 100 км. Сигнал розсіяння від другого елементу використовується для визначення потужності $P(h)$ НР сигналу, розсіяного плазмою в діапазоні висот 100–550 км, з роздільною здатністю за висотою 20 км (або 10 км) і розрахунку висотного профілю концентрації електронів $N_e(h)$. Цей режим роботи радара НР на теперішній час є основним.



а *б*

Рис. 1. Зондувальні сигнали в режимі випромінювання двочастотного радіоімпульсу з довгим і коротким елементами (*а*) та режимі зондування іоносфери циклічною послідовністю коротких одиночних і здвоєних радіоімпульсів (*б*)

Режим зондування іоносфери циклічною послідовністю одиночних і здвоєних радіоімпульсів використовується для вимірювання параметрів іоносфери в діапазоні висот 100–550 км. У цьому режимі передавач випромінює один (для вимірювання потужності) або два (для вимірювання КФ НР сигналу) коротких (~150 мкс) імпульси зі змінною затримкою між ними (160–1000 мкс), що змінюється від періоду до періоду випромінювання на постійну величину (~40 мкс). З метою уникнення похибок вимірювання іоносферних параметрів, спричинених неідентичністю здвоєних радіоімпульсів (в першу чергу, похибки визначення швидкості руху іоносферної плазми), використовується модернізований режим зондування здвоєними імпульсами – випромінюється складовий сигнал з частотою елементів коротких імпульсів f_0 =158 МГц і міжімпульсним ВЧ заповненням на зсунутій частоті $f_1 = f_0$+100 кГц (рис. 1, *б*).

Режим вимірювання концентрації електронів з використанням ефекту Фарадея. У цьому режимі випромінюється радіохвиля з лінійною поляризацією. Структура зондувального сигналу є такою, що наведена на рис. 1, *а*, але короткий елемент має частоту f_0, а довгий елемент – f_1. Приймання лінійно поляризованих сигналів здійснюється двома взаємно ортогональними вібраторами антени НДА-100, двома каналами фідерної системи й підключеними до них ідентичними РПрП. Обробці підлягають луна-сигнали від короткого елемента зондувального сигналу, тривалість якого складає 135 мкс. Висотний профіль концентрації електронів визначається з отриманих профілів потужностей сигналів на виходах двох приймальних трактів.

Проведення експериментів з одночасним зондуванням у вертикальному й похилому напрямках значно підвищує можливості дослідження іоносфери і, зокрема, динамічних

процесів в іоносфері методом НР [33], але споживана за цих обставин електроенергія зростає майже вдвічі. Робота двох радарів з одночасним зондуванням іоносфери двома антенами, спрямованими вертикально, дозволяє порівнювати результати, отримані незалежними радарами, але із зондованими об'ємами іоносферної плазми, що відрізняються.

Для реалізації роботи радарів з одночасним використанням зенітної та повноповоротної антен були вирішені такі задачі:

- забезпечити синхронну роботу всіх пристроїв радарної установки;
- забезпечити електромагнітну сумісність при синхронній роботі радарів;
- забезпечити вимірювання КФ сигналу НР на низькій частоті.

Першу задачу вирішено завдяки використанню спільного для обох радарів синхронізатора (пристрою керування), який формує сигнали запуску передавачів, сигнали стробування виділення зондувальних імпульсів, бланкувальні імпульси, призначені для замикання РПрП під час зондування та сигнал початку періоду посилок (початку радіолокаційної розгортки), що подається на цифрові пристрої обробки.

Електромагнітна сумісність радарів із зенітною та повноповоротною антенами досягається використанням сигналів з круговою поляризацією з протилежними напрямками обертання вектора електричного поля та екрануванням окремих вузлів апаратури.

Остання задача полягає в перенесенні в кожному радарі спектра НР сигналу на нульову частоту з точністю до доплерівського зсуву, обумовленого рухом плазми вздовж напрямку зондування, і виділення квадратурних сигналів для кореляційної обробки сигналів на низькій частоті з можливістю визначення швидкості цього руху. Це забезпечується когерентною роботою радарів.

Структурна схема комплексу радарів НР показана на рис. 2.

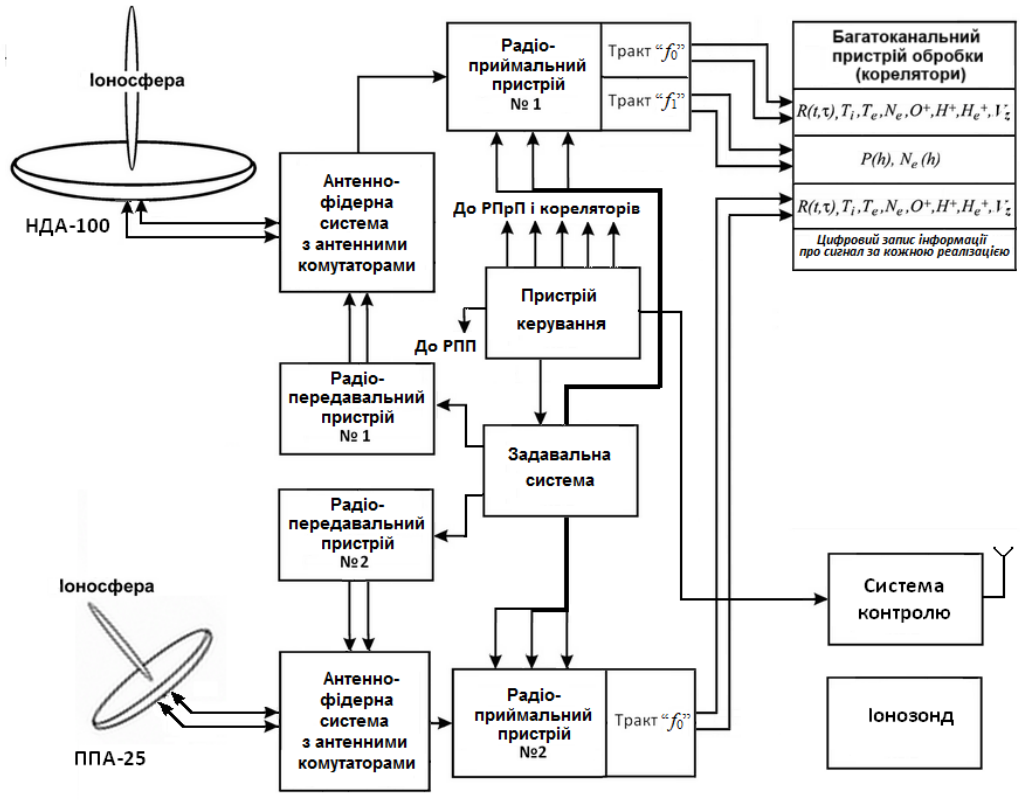


Рис. 2. Структурна схема комплексу радарів НР НДІ Іоносфери

Сформовані задавальною системою складені радіоімпульсні сигнали подаються на радіопередавальні пристрої (РПП), підсилюються й надходять фідерними трактами через антенні комутатори до антен НДА-100 і ППА-25. Антенами випромінюються сигнали відповідно у вертикальному й обраному похилому чи вертикальному напрямках. При одно-

часній роботі радарів передаються та приймаються радіохвилі кругової поляризації з протилежними напрямками обертання вектора електричного поля для усунення взаємних завад.

Розсіяні іоносферою сигнали приймаються цими ж антенами й через антенні комутатори антенно-фідерного тракту подаються на входи відповідних радіоприймальних пристроїв, де здійснюється підсилення прийнятих сигналів, їх селекція в трактах «f_0» і «f_1», фільтрація, перенесення спектра в область низьких частот і виділення квадратурних сигналів для кожного з елементів за допомогою синхронного детектування для подальшої кореляційної обробки.

Пристрій керування формує імпульсні сигнали на всі системи радарів НР.

Система контролю формує такі сигнали: білий шум, вузькосмугові радіоімпульсні шумові сигнали з відомими спектрами, що відповідають спектрам некогерентного розсіяння, гармонійний сигнал [34]. Ці сигнали подаються через контрольну антену по ефіру на антени НДА-100 та ППА-25 або безпосередньо на вхід РПрП для контролю антенно-фідерних пристроїв і приймально-обробних систем. Також здійснюється контроль радіопередавальних пристроїв та якості формування сигналів з круговою поляризацією.

Іонозонд використовується для калібрування вимірювань концентрації електронів шляхом прив'язки вимірюваної критичної частоти f_0F2 до максимуму відносного висотного профілю концентрації електронів (при використанні перших двох режимів зондування), а також для отримання додаткової інформації про іоносферу з іонограм.

Структура й особливості діючої радіоприймальної системи

Радіоприймальна система містить два РПрП, кожен з яких є супергетеродинним приймачем з потрійним перетворенням частоти. Спектр прийнятого сигналу послідовно переноситься з носійної частоти (158 МГц) в область низьких частот, де здійснюється кореляційна обробка. З метою максимального зниження втрат вхідного сигналу та зменшення впливу завад РПрС виконана компактно й розташована в безпосередній близькості від виходів антенно-фідерного тракту. Її структурну схему наведено на рис. 3.

На вході кожного приймача є система вибору поляризації сигналу (що складається з фазообертача та коаксіального суматора сигналів від двох каналів фідерного тракту), циркулятор для узгодження вхідного опору приймача з опором фідера лінії (75 Ом), а також транзисторний підсилювач високої частоти (ПВЧ1) з низьким рівнем шуму. Щоб уникнути перевантаження приймачів й АЦП пристроїв обробки, забезпечено бланкування РПрП у вхідних ланцюгах за допомогою двох швидкодіючих електронних комутаторів на PIN-діодах, що керуються сигналами формувача імпульсів бланка та забезпечують замикання приймачів на час випромінювання зондувального імпульсу.

З виходу антенно-фідерної системи сигнал надходить на вхід відповідного РПрП. Його перетворення розглянемо на прикладі РПрП №1. У вхідному пристрої приймача сигнал посилюється, піддається бланкуванню (під час випромінювання імпульсу передавача) і переноситься на першу проміжну частоту (ПЧ) ($f_{пр1}$≈15 МГц). У тракті ПЧ-1 здійснюється його посилення та поділ на два тракти: тракт виділення відгуку сигналу з носійною частотою f_0 і тракт виділення відгуку сигналу з носійною частотою f_1. Обидва тракти за своєю структурою ідентичні. У тракті ПЧ-2 спектр сигналу переноситься на другу проміжну частоту ($f_{пр2}$=972,4 МГц), і здійснюється його селекція та посилення.

Двоканальний тракт НЧ (тракт виділення сигналів для квадратурної обробки) має два однакові канали, кожен з яких містить синхронний детектор (СД), виконаний на прецизійному інтегральному помножувачі, набір комутованих фільтрів нижніх частот (ФНЧ) і підсилювач. Для здійснення синхронного детектування використовуються сигнали фазообертача третього (синхронного) гетеродина з однаковою амплітудою, частотою $f_{сг}$ та різницею фаз 90°. Виділені синхронними детекторами сигнали піддаються фільтрації у фільтрах НЧ з обраною шириною смуги пропускання й посилюються до рівня напруги, необхідної для нормальної роботи АЦП пристрою обробки.

Структура тракту «f_0» РПрП №2 така ж, як і РПрП №1.

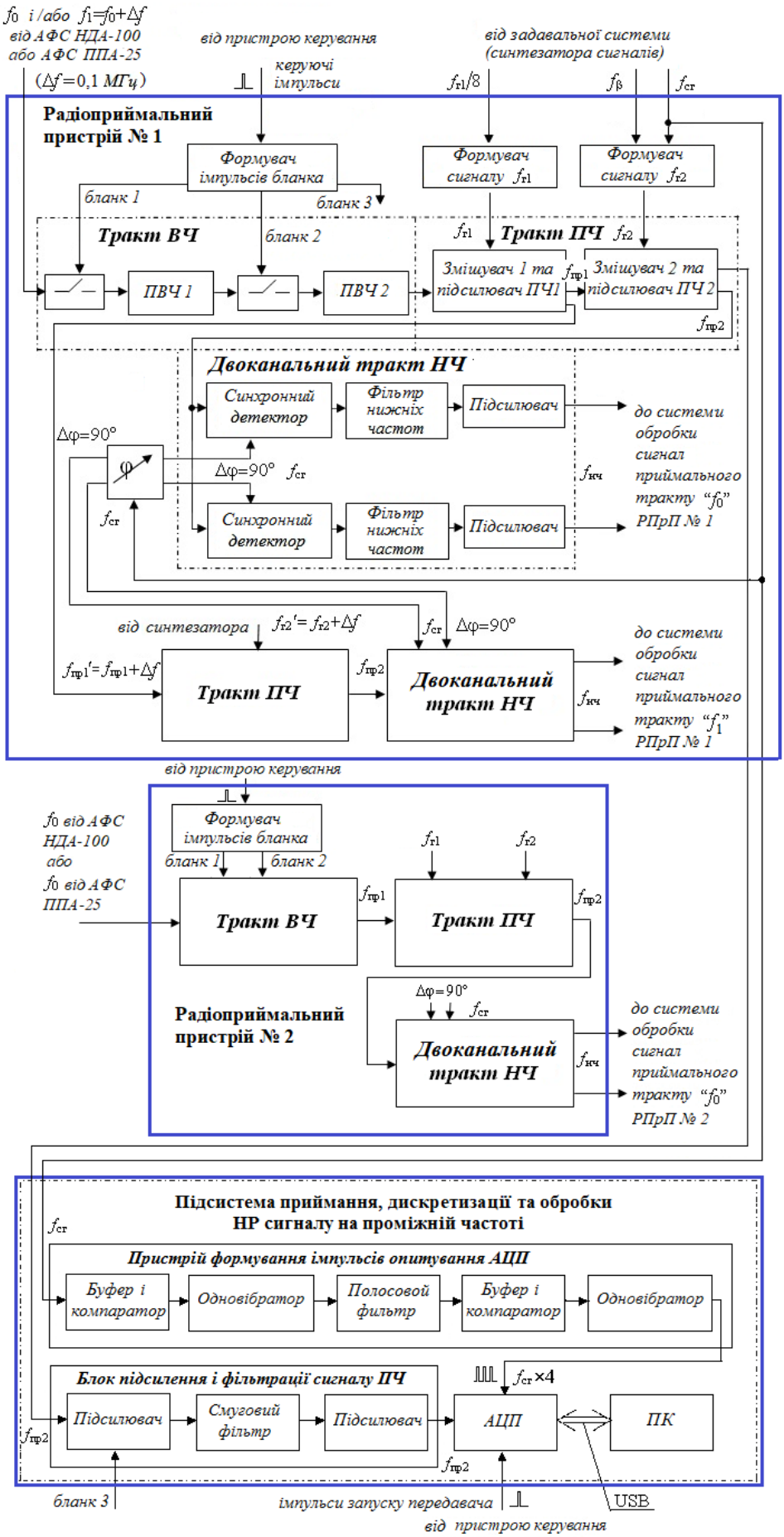


Рис. 3. Структурна схема радіоприймальної системи радара НР

Завдяки формуванню сигналів гетеродинів РПрС із сигналів задавальної системи забезпечується когерентність роботи радара НР і можливість вимірювати одночасно з іншими іоносферними параметрами швидкість руху іоносферної плазми за доплерівським зміщенням спектра НР сигналу.

Сигнал 1-го гетеродину з частотою $f_{г1}$ формується в помножувачі частоти з сигналу синтезатора з частотою ($f_{г1}$/8). Оскільки носійна частота зондувального сигналу формується синтезатором відповідно до виразу $f_0 = ((f_{г1}/8)–2f_β)\cdot 8$, перша ПЧ дорівнює $f_{пр1} = f_{г1}–f_0 = 16f_β$. Таким чином, вона відома з точністю до доплерівського зміщення спектра НР сигналу, викликаного рухом плазми. Сигнал 2-го гетеродину формується відповідно до виразу: $f_{г2} = 16f_β+f_{сг}$. Сигнал 3-го (синхронного) гетеродину формується із сигналу синтезатора з частотою $f_{сг}$ за допомогою кварцового фільтра. Тому друга проміжна частота $f_{пр2}$ збігається з частотою $f_{сг}$ синхронного гетеродину (у разі відсутності доплерівського зміщення). Унаслідок синхронного детектування, сигнал переноситься на нульову частоту з доплерівським зміщенням. Таким чином, забезпечується когерентність систем радара та можливість вимірювання швидкості руху іоносферної плазми й об’єктів у геокосмосі.

Для тракту «f_1» як другий гетеродин з частотою $f_{г2'}=f_{г2}+\Delta F$, де $\Delta F\approx 100$ кГц, використовується промисловий синтезатор, стабільність якого цілком задовільна для визначення потужності НР сигналу, виділеного в цьому тракті.

Відповідно до рівняння радіолокації [3] відношення потужності НР сигналу, розсіяного об'ємом плазми з центром на висоті h, до потужності шуму $P_ш$ дорівнює

$$q(h)=\frac{0{,}76P_i\eta c\tau_i A}{16\pi kT_c\Delta F}\cdot\frac{\sigma N_e(h)}{h^2}, \qquad (1)$$

де P_i – імпульсна потужність передавача; η – к.к.д. антенно-фідерного тракту; c – швидкість світла; τ_i – тривалість зондувального імпульсу; A – ефективна площа антени; σ – ефективний переріз розсіяння; k – постійна Больцмана; T_c – ефективна шумова температура системи включно з антенно-фідерним трактом; ΔF – ефективна ширина смуги пропускання приймача.

Ефективна шумова температура приймальної системи T_c визначається формулою [35]:

$$T_c = T_A + T_{лф}(L-1)+LT_{пр}, \qquad (2)$$

де T_A − ефективна шумова температура антени; $T_{лф}$ – фізична температура фідерної лінії ($T_{лф}\approx T_0 = 300$К); $T_{пр}$ – ефективна шумова температура приймача; $L = 1/\eta$ – втрати в тракті. Для радара з антеною НДА-100 T_c = 470–980 К в залежності від космічного шуму, який змінюється протягом доби.

Із виразів (1) і (2) видно, що при вибраній тривалості τ_i (виходячи з необхідної роздільної здатності за дальністю), при наявних параметрах P_i і η підвищити відношення сигнал/шум можна зниженням шумової температури (коефіцієнту шуму) РПрП і вибором його оптимальної ширини смуги пропускання.

Оптимізація параметрів РПрП торкнулася насамперед поліпшення їх чутливості та вибору найбільш прийнятних для конкретних умов іоносферних вимірювань параметрів фільтрів нижніх частот вихідних ланцюгів РПрП, які визначають відповідні характеристики всього приймального тракту. Для цього на входах РПрП були встановлені малошумливі підсилювачі та вузькосмугові коаксіальні фільтри високих частот, що забезпечило достатній для метрового діапазону коефіцієнт шуму (1,4–1,5) і високу захищеність від завад. У вихідних трактах квадратурних сигналів використовуються ФНЧ, найбільш прийнятні для умов експериментів: у тракті «f_0» РПрП антени НДА-100 – фільтри Кауера 7-го порядку з формою АЧХ, близькою до прямокутної, і шириною смуги пропускання за рівнем половинної потужності 9,5 кГц (для визначення параметрів іоносфери на висотах 200–1500 км); у тракті «f_1» РПрП антени НДА-100 – фільтри Чебишева 3-го порядку з пологими схилами АЧХ і шириною смуги пропускання 6,0 кГц (для визначення потужності НР сигналу на висотах нижче й поблизу

висоти максимуму іонізації); в РПрП антени ППА-25 – фільтри Кауера 7-го порядку з формою АЧХ, близькою до прямокутної, і шириною смуги пропускання за рівнем половинної потужності, що дорівнює 5,5 кГц (для визначення параметрів іоносфери на висотах поблизу максимуму іонізації і нижче, де присутні електрони й іони атомарного кисню).

Статистична похибка визначення іоносферних параметрів тим менше, чим більше відношення сигнал/шум. Якщо порівняти можливості радарів з антенами НДА-100 і ППА-25, встановленої у вертикальному напрямку зондування, за інших рівних умов відношення сигнал/шум відрізняються у 12,8 разів відповідно до виразів (1) та (2), що пов'язано з відмінністю площ їх ефективних поверхонь.

Однією з останніх модернізацій РПрС є підсистема приймання, дискретизації та обробки сигналу на проміжній частоті [36], яка використовує лише один швидкодіючий АЦП, що підключається по шині USB. Обробка сигналу НР на проміжній частоті має переваги в деяких випадках. Вона дозволяє використовувати алгоритми, адаптовані до досліджуваного діапазону висот і стану іоносфери, а також підвищити точність вимірювання параметрів НР сигналу (і, відповідно, параметрів іоносфери), виключаючи вплив низки інструментальних факторів та адаптуючи для цифрової обробки в реальному часі частоту дискретизації сигналу на ПЧ до параметрів розсіяних плазмою сигналів або сигналів, відбитих від дискретних об’єктів у геокосмосі [36, 37].

Підсистема містить у собі блок підсилення та фільтрації сигналу ПЧ (що складається з двох підсилювачів і смугового фільтра), АЦП, персональний комп'ютер (ПК) і пристрій формування опитувальних імпульсів (рис. 3). Для жорсткої прив'язки імпульсів опитування АЦП до проміжної частоти як вхідний сигнал для пристрою формування опитувальних імпульсів використовується сигнал синхронного гетеродина з частотою $f_{сг} = f_{пр2}$ =972,4 кГц. На виході пристрою формуються імпульси з частотою слідування $4f_{сг}$, тобто з періодом слідування, рівним чверті періоду сигналу синхронного гетеродина. Таким чином, сусідні відліки сигналу знаходяться в квадратурній залежності. Це дозволяє визначати ординати квадратурних складових КФ $R(\tau)$ з часовими зсувами (аргументами) τ, кратними періоду ПЧ і оптимальними для конкретних геліогеофізичних умов. Як сигнал позначення початку кожної радіолокаційної розгортки використовуються імпульси запуску передавача радара НР.

До блоку формування імпульсів зчитування АЦП пред'являються особливі вимоги до стабільності і точності множення частоти. Це пов'язано з точністю дискретизації сигналу для забезпечення квадратури його вибірок, які використовуються для подальшого розрахунку параметрів НР сигналу та параметрів іоносфери. Із декількох проаналізованих варіантів помножувачів частоти було обрано найбільш прийнятний варіант для формування імпульсного сигналу з почетвереною ПЧ. Цей варіант заснований на виділенні 4-ї гармоніки з послідовності імпульсів з оптимальним співвідношенням періоду повторення та тривалості.

Параметри іоносфери та відповідні статистичні характеристики сигналу НР істотно змінюються з висотою й залежать від космічної погоди. Від цих факторів також залежить точність оцінки кореляційної функції шумоподібного сигналу НР і, отже, точність визначення параметрів іоносфери. Було запропоновано підвищити точність такої оцінки за рахунок підвищення співвідношення сигнал/шум, використавши додаткову цифрову обробку сигналу, а саме набір цифрових фільтрів, що є оптимізованими для різних висот і стану іоносфери. Смугові фільтри з нескінченною імпульсною характеристикою були створені за допомогою бібліотеки SciPy мови програмування Python. Для аналізу були створені фільтри типу Чебишева 17-го порядку зі смугою пропускання 10, 14, 19 і 24 кГц з лінійною характеристикою в смузі пропускання та придушенням сигналу за межами смуги пропускання на рівні 60 дБ. Для недопущення зміщення вихідного сигналу після фільтрації щодо вхідного сигналу використовується метод “forward-backward” – двічі застосовується лінійний цифровий фільтр, один раз вперед і один раз назад. Комбінований фільтр має нульову фазу й порядок фільтра, який вдвічі перевищує вихідний.

Розробка універсальної програмованої радіосистеми моніторингу геокосмосу

З метою отримання нових геофізичних знань про навколоземний космічний простір, його моніторингу та спостереженню космічних об'єктів, що перебувають у ньому, було розроблено програмовану радіосистему, що функціонує під керуванням спеціалізованого програмного забезпечення. У процесі виконання роботи було вирішено такі основні задачі: аналіз особливостей формування сигналів у програмованих радіосистемах для радіофізичних досліджень атмосфери та геокосмосу; розробка апаратного та програмного забезпечення для формування й обробки сигналів; проведення тестових випробовувань розробленої радіосистеми та аналіз результатів [38].

Впровадження програмованої радіосистеми до складу обладнання НДІ Іоносфери може здійснюватися за схемою, наведеною на рис. 4. Через те, що передавач і приймач радара знаходяться в різних будівлях, передавальну та приймальну частину програмованої радіосистеми необхідно розміщувати окремо, реалізувавши на двох лінійках універсальної програмованої радіосистеми Universal Software Radio Peripheral (USRP).

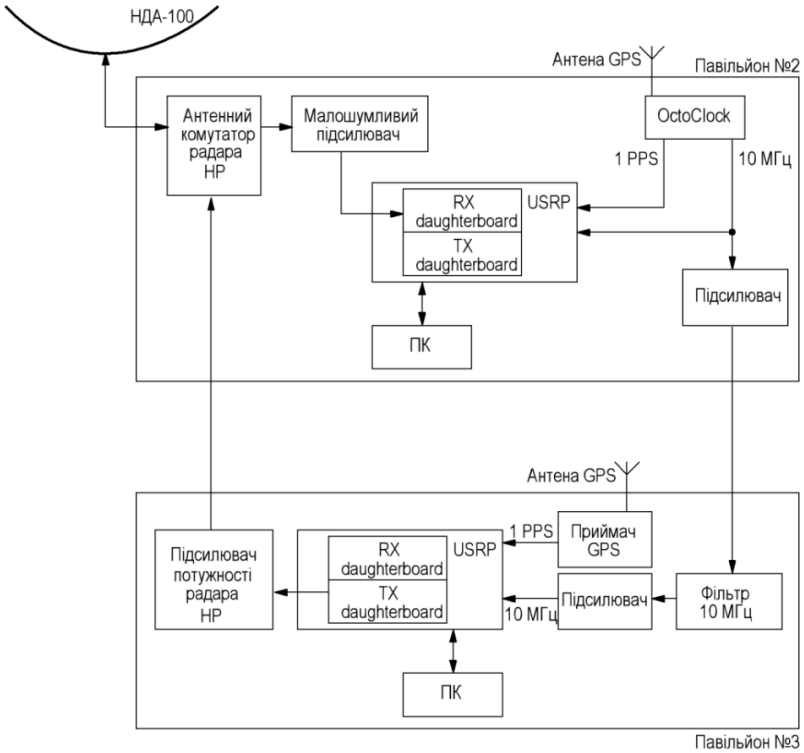


Рис. 4. Структурна схема радіосистеми на базі USRP у складі радара НР НДІ Іоносфери

До складу приймальної частини входить USRP з дочірніми платами BasicRX та BasicTX (частота, на якій передбачається робота радіосистеми, складає 158 МГц), антенний комутатор, малошумливий підсилювач, генератор сигналів з частотами 10 МГц та 1 Гц, а також персональний комп'ютер, який керує USRP, оброблює прийняті сигнали та зберігає результати обробки даних. У структурі передавальної частини присутні попередній підсилювач сигналів від передавальної дочірньої плати USRP і вихідний підсилювач потужності. Передавальна дочірня плата USRP – BasicTX. Синхронізація передавальної та приймальної підсистем відбувається за допомогою OctoClock (1 PPS та 10 МГц) і додаткового приймача GPS для формування сигналу PPS.

Радіоприймальну частину можна використовувати в пасивному режимі для дослідження космічних шумів. Замість вказаної на рис. 4 вертикально спрямованої антени НДА-100 можливе використання повноповоротної антени ППА-25.

Інтегральний GaAs pHEMT малошумливий підсилювач має забезпечувати наведені нижче вимоги:

– коефіцієнт підсилення на частоті 158 МГц – не менше 20 дБ;

– коефіцієнт шуму на частоті на 158 МГц – менше 0,7 дБ;

– коефіцієнт стоячої хвилі по входу на частоті 158 МГц без кіл узгодження на вході – не більше 1,5.

Передавальний тракт використовується тільки в активному режимі та складається з попереднього підсилювача й вихідного підсилювача потужності. Попередній підсилювач забезпечує рівень сигналу, що надходить від USRP, необхідний для роботи підсилювача потужності. Вихідний підсилювач має розвивати на навантаженні з опором 50 Ом потужність від декількох десятків до сотень ватт і може бути побудований на спеціалізованих високочастотних транзисторах, наприклад MRF240.

Для фільтрації сигналу синхронізації від гармонік, що утворюватимуться в підсилювачі, з боку USRP необхідно встановити смуговий кварцовий фільтр.

Програмне забезпечення складається з розроблених для різних режимів роботи радара програм керування модулем USRP на мові програмування C++, керування антенним комутатором тощо. Параметри кожної програми задаються у відповідному конфігураційному файлі. Запис отриманих даних розгорток дальності проводиться в бінарні файли.

Висновки

Розглянуто особливості радіоприймальної системи радара некогерентного розсіяння (НР) НДІ Іоносфери, яка є ключовим елементом для забезпечення високої точності вимірювання параметрів іоносфери та геокосмічного середовища.

Показано, що для успішного приймання та обробки слабких НР сигналів, потужність яких часто значно менша за потужність шумів, необхідна висока чутливість, стабільність і завадостійкість радіоприймальної системи.

Детально охарактеризовано структуру та принцип дії радіоприймальної системи радара НР, зокрема супергетеродинної архітектури з потрійним перетворенням частоти. Реалізація доплерівських вимірювань швидкості руху іоносферної плазми та об’єктів у геокосмосі ґрунтується на когерентній структурі радарного комплексу та узгодженій синхронізації передавальних і приймальних трактів.

Наведено такі методи модернізації діючої системи:

− оптимізація її окремих елементів, зокрема впровадження малошумливих підсилювачів, удосконалення фільтрів нижніх частот, використання високоточних методів синхронного детектування дозволило значно підвищити чутливість та завадостійкість радіоприймального тракту;

− введення системи вибору поляризації сигналу та спеціалізованих режимів зондування забезпечують гнучке налаштування експериментів відповідно до поставлених дослідницьких задач та поточних геліогеофізичних умов;

− залучення одночасного зондування двома антенами у вертикальному та похилому напрямках значно розширює діагностичні можливості комплексу радарів НР, дозволяючи досліджувати тривимірну динаміку іоносфери;

− введення до складу радара підсистеми приймання, дискретизації та обробки сигналу на проміжній частоті дозволило підвищити точність вимірювання параметрів НР сигналу (і, відповідно, параметрів іоносфери) завдяки виключенню впливу низки інструментальних факторів, характерних для аналогових трактів виділення квадратурних сигналів; адаптації частоти дискретних вибірок сигналу до параметрів розсіяних плазмою сигналів з різних висотних діапазонів; використанню цифрових смугових фільтрів і алгоритмів, що адаптовані до сигналів, відповідних конкретним геліогеофізичним умовам і висотам дослідження.

Розроблено апаратну та програмну складові багатоцільової програмованої радіосистеми на базі SDR (Software Defined Radio) технології, яка призначена безпосередньо для моніторингу геокосмосу й значно розширює можливості радіофізичного обладнання обсерваторії НДІ Іоносфери.

Представлені технічні рішення відповідають вимогам сучасних систем моніторингу геокосмосу та є корисними для дослідження іоносфери та спостереження за штучними космічними об'єктами, зокрема в контексті зростаючої актуальності прогнозування космічної погоди та проблем космічного сміття.

Список літератури:

1. Xue D., Yang J., Liu Z. Potential impact of GNSS positioning errors on the satellite-navigation-based air traffic management // Space Weather. 2022. Vol. 20, is. 7, e2022SW003144. https://doi.org/10.1029/2022SW003144.

2. Reznychenko M. O., Kotov D. V., Richards P. G., Bogomaz O. V., Reznychenko A. I., Goncharenko L. P., et al. The thermosphere was poorly predictable not only during but also before and after the Starlink Storm on 3–4 February 2022 // Geophys. Res. Lett. 2025. Vol. 52. e2024GL112620. https://doi.org/10.1029/2024GL112620.

3. Evans J. V. Theory and practice of ionosphere study by Thomson scatter radar // Proceedings of the IEEE. Vol. 57, no 4. 1969. P. 496–530. https://doi.org/10.1109/PROC.1969.7005.

4. Sato T., Ito A., Oliver W. L., Fukao S., Tsuda T., Kato S., Kimura I. Ionospheric incoherent scatter measurements with the middle and upper ionosphere radar. Techniques and capability // Radio Sci. 1989. Vol. 24. P. 85–98. https://doi.org/10.1029/RS024i001p00085.

5. Holt J. M., Erickson P. J., Gorczyca A. M., Grydeland T. MIDAS-W: a workstation-based incoherent scatter radar data acquisition system // Ann. Geophys. 2000. Vol. 18, no. 9. P. 1231–1241. https://doi.org/10.1007/s00585-000-1231-3.

6. Woodman R. F., Farley D. T., Balsley B. B., Milla M. A. The early history of the Jicamarca radio observatory and the incoherent scatter technique // Hist. Geo Space. Sci. 2019. Vol. 10. P. 245–266. https://doi.org/10.5194/hgss-10-245-2019.

7. Yue X., Wan W., Ning B., et al. Development of the Sanya incoherent scatter radar and preliminary results // J. Geophys. Res.: Space Physics. 2022. Vol. 127, e2022JA030451. https://doi.org/10.1029/2022JA030451.

8. Lehtinen M., Markkanen J., Väänänen A., Huuskonen A., Damtie B., Nygrén T., Rahkola J. A new incoherent scatter technique in the EISCAT Svalbard Radar // Radio Sci. 2002. Vol. 37, no. 4. https://doi.org/10.1029/2001RS002518.

9. Wannberg U. G. et al. EISCAT_3D: A Next-Generation European Radar System for Upper-Atmosphere and Geospace Research // The Radio Science Bulletin. 2010. No 332. P 75–88.

10. Ding Z., Wu J., Xu Z., Xu B., Dai L. The Qujing incoherent scatter radar: system description and preliminary measurements // Earth Planets Space. 2018. Vol. 70, no. 1. https://doi.org/10.1186/s40623-018-0859-8.

11. Grydeland T., Lind F. D., Erickson P. J., Holt J. M. Software Radar signal processing // Ann. Geophys. 2005. Vol. 23, no. 1., P. 109–121. https://doi.org/10.5194/angeo-23-109-2005.

12. Damtie B. New incoherent scatter radar measurement techniques and data analysis methods // Department of Physical Sciences, University of Oulu, Finland, 2004, Report №29.

13. Emelyanov L. Ya., Zhivolup T. G. History of the development of IS radars and founding of the Institute of Ionosphere in Ukraine // Hist. Geo Space. Sci. 2013. Vol. 4. P. 7–17. https://doi.org/10.5194/hgss-4-7-2013.

14. Domnin I. F., Chepurnyy Ya. M., Emelyanov L. Ya., Chernyaev S. V., Kononenko A. F., Kotov D. V., Bogomaz O. V., Iskra D. A. Kharkiv Incoherent Scatter Facility // Bulletin of the National Technical University "Kharkiv Politechnic Institute": scientific papers. Issue: Radiophysics and ionosphere. Kharkiv : NTU "KhPI". 2014. No. 47 (1089). P. 28–42. http://nbuv.gov.ua/UJRN/vcpiri_2014_47_7.

15. Stone M. L.; Banner G. P. Radars for the Detection and Tracking of Ballistic Missiles, Satellites, and Planets // Linc. Lab. J. 2000. Vol. 12, no. 2. P. 217–244.
https://archive.ll.mit.edu/publications/journal/pdf/vol12_no2/12_2detectsatellitiesplanets.pdf.

16. Boley A. C., Wright E., Lawler S., Hickson P., Balam, D. Plaskett 1.8 m Observations of Starlink Satellites // The Astronomical Journal. 2022. Vol. 163 (5). P. 199. https://doi.org/10.48550/arXiv.2109.12494.

17. Graham M. J., Kulkarni S. R., Bellm E. C., et al. The zwicky transient facility: science objectives // Publications of the Astronomical Society of the Pacific. 2019. 131:078001 (23 pages). https://doi.org/10.1088/1538-3873/ab006c.

18. Parham J. B., Li J., Dickson M., Ginet G., Erickson P. J., Lind F. D. Debris plasma density perturbations as seen through a modern collective Thomson scatter radar processing chain // 2024 International Conference on Electromagnetics in Advanced Applications (ICEAA), 02–06 September 2024.
https://doi.org/10.1109/ICEAA61917.2024.10701844.

19. Stamm J., Vierinen J., Gustavsson B., Spicher A. A technique for volumetric incoherent scatter radar analysis // Ann. Geophys. 2023. Vol. 41, is. 1. P. 55–67. https://doi.org/10.5194/angeo-41-55-2023.

20. Skolnik M. I. Introduction to Radar Systems. New York, McGraw-Hill, 3rd Edition, 2001.

21. Zakharenko V., Konovalenko A., Zarka P. et al. Digital receivers for low-frequency radio telescopes UTR-2, URAN, GURT // Journal of Astronomical Instrumentation. 2016. Vol. 5, no. 4. 1641010 (19 pages). https://doi.org/10.1142/S2251171716410105164010-1.

22. Thompson A. R., Moran J. M., Swenson Jr. G. W. Interferometry and Synthesis in Radio Astronomy. Cham.: Springer Nature. 2017. https://doi.org/10.1007/978-3-319-44431-4.

23. Bostan S. M., Urbina J., Mathews J. D., Bilén S. G., Breakall J. K. An HF software-defined radar to study the ionosphere // Radio Science. 2019. Vol. 54. P. 839–849. https://doi.org/10.1029/2018RS006773.

24. Kalita B. R., Nath S. J., Bhuyan P. K., Khandare A., Kulkarni A. SAMEERDU—digital ionosonde: Brief system description and initial results from a low-latitude location Dibrugarh // Radio Science. 2019. Vol. 54. P. 1142–1155. https://doi.org/10.1029/2019RS006813.

25. Barona Mendoza J. J., Quiroga Ruiz C. F., Pinedo Jaramillo C. R. Implementation of an electronic ionosonde to monitor the Earth's ionosphere via a projected column through USRP // Sensors. 2017. Vol. 17, no. 5. 946 p. https://doi.org/10.3390/s17050946.

26. Rejfek L., Kouba D., Mošna Z., Knížová P. K., Tran P. T., Dong C. S. T. Passive ionospheric radar builds with USRP N210 // Journal of Electrical Engineering. 2019. Vol. 70, is. 2. P. 159–164. https://doi.org/10.2478/jee-2019-0023.

27. Li Y., Yuan K., Yao M., Deng X. The prototype incoherent scatter radar system of Nanchang University // IEEE Geoscience and Remote Sensing Letters. 2020. P. 1184–1188. https://doi.org/10.1109/LGRS.2020.2994082.

28. Thayaparan T., Dupont D., Ibrahim Y., Riddolls R. High-frequency ionospheric monitoring system for Over-the-Horizon radar in Canada // IEEE Transactions on Geoscience and Remote Sensing. 2019. Vol. 57, is. 9. P. 6372–6384. https://doi.org/10.1109/TGRS.2019.2905757.

29. Holdsworth D. A., Spargo A. J., Reid I. M., Adami C. Low Earth Orbit object observations using the Buckland Park VHF radar // Radio Sci. 2020. Vol. 55, is. 2. e2019RS006873. https://doi.org/10.1029/2019RS006873.

30. Kenington P. B. RF and baseband techniques for software defined radio. Artech House Mobile Communicat. 2005. 332 p. https://picture.iczhiku.com/resource/eetop/WHIWPiloyEJraxXB.pdf.

31. Farley D. T. Faraday rotation measurements using incoherent scatter // Radio Sci. 1969. Vol. 4, no 2. P. 143–152. https://doi.org/10.1029/RS004i002p00143.

32. Skvortsov T. A., Yemelyanov L. Ya., Fisun A.V. Radar measurements of geomagnetic field in the ionosphere // Telecommunications and Radio Engineering. 2015. Vol. 74, No 10. P. 921–931. https://doi.org/10.1615/TelecomRadEng.v74.i10.80.

33. Emelyanov L., Chepurnyy Y., Bogomaz O. Simultaneous sounding of the ionosphere in the vertical and oblique directions using incoherent scatter radar // 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO). 2018. P. 458–463. https://doi.org/10.1109/ELNANO.2018.8477456.

34. Yemelyanov L. Ya. Development of principles and instrumentation for generation of test and control signals of the incoherent scatter radar // Telecommunications and Radio Engineering. 2017. Vol. 76, is. 14. P. 1259–1271. https://doi.org/10.1615/TelecomRadEng.v76.i14.50.

35. Kraus J. D. Radio Astronomy, 2nd ed. Cygnus-Quasar Books, Powell, OH. 1986.

36. Emelyanov L., Miroshnikov A. Development of methodological, hardware, and software of the incoherent scatter radar of Institute of Ionosphere (Kharkiv, Ukraine) // International Journal of Electronics and Telecommunications. 2023. Vol. 69, no. 3. P. 579–586. https://doi.org/10.24425/ijet.2023.146510.

37. Emelyanov L., Rogozhkin E., Podyachiy Y. Method for sampling narrowband radio signals with known center frequency of spectrum // 2024 IEEE 42nd International Conference on Electronics and Nanotechnology (ELNANO). May 13–16, 2024, Kyiv, Ukraine, Conference Proceedings. 2024. P. 544–548. https://doi.org/10.1109/ELNANO63394.2024.10756842.

38. Bogomaz O., Barabash V., Iskra D. Some aspects of developing a multipurpose radio system for monitoring the geospace // Advanced Applied Energy and Information Technologies 2021. Proceedings of the International Conference (Ternopil, 15–17 of December 2021.) Ternopil : TNTU, 2021. P. 114–119. https://elartu.tntu.edu.ua/bitstream/lib/36934/2/ICAAEIT_2021_Bogomaz_O-Some_aspects_of_developing_114-119.pdf.

Надійшла до редколегії 17.05.2025

Відомості про авторів:

Ємельянов Леонід Якович – кандидат фізико-математичних наук, старший науковий співробітник, Науково-дослідний інститут Іоносфери Національного технічного університету "Харківський політехнічний інститут", провідний науковий співробітник; Україна; e-mail: leonid.ya.emelyanov@gmail.com, ORCID: https://orcid.org/0000-0002-2117-2675.

Богомаз Олександр Вікторович – кандидат технічних наук, Науково-дослідний інститут Іоносфери Національного технічного університету "Харківський політехнічний інститут", старший науковий співробітник; Україна; e-mail: o.v.bogomaz1985@gmail.com, ORCID: https://orcid.org/0000-0001-6824-2346.

Под'ячий Юрій Іванович – кандидат фізико-математичних наук, доцент, Національний технічний університет "Харківський політехнічний інститут", професор кафедри мікро- та наноелектроніки; Україна; e-mail: uipodyachiy@gmail.com, ORCID: https://orcid.org/0009-0007-8252-1553.

Мірошніков Артем Євгенійович – Науково-дослідний інститут Іоносфери Національного технічного університету "Харківський політехнічний інститут", молодший науковий співробітник; Україна; e-mail: moneytu@gmail.com, ORCID: https://orcid.org/0000-0002-2473-4370.

УДК 621.396 **DOI:10.30837/rt.2025.3.222.11**

С.С. ЖИЛА, д-р техн. наук, О.В. ОДОКІЄНКО, канд. техн. наук, Д.І. КОВАЛЬЧУК, К.О. ЩЕРБИНА, канд. техн. наук, Я.Д. СИДОРОВ

СТАТИСТИЧНА ОПТИМІЗАЦІЯ ТА АНАЛІЗ МЕТОДУ ФОРМУВАННЯ РАДІОЛОКАЦІЙНИХ ЗОБРАЖЕНЬ У ЧАСОВІЙ ТА ЧАСТОТНІЙ ОБЛАСТЯХ

Вступ

У сучасних радіолокаційних системах дедалі більшої актуальності набуває завдання підвищення просторової та контрастної роздільної здатності зображень, що формуються на основі скатерометричних даних. Проте традиційні методи формування радіолокаційних зображень часто не забезпечують необхідної якості через обмеженість моделей, недостатній урахунок статистичних характеристик сигналів або використання евристичних підходів до фільтрації й інтерпретації даних. Особливо це стосується задач виявлення, ідентифікації та класифікації об’єктів у складних умовах дії перешкод, шумів і неоднорідного фону.

З огляду на це постає проблема розробки статистично обґрунтованого підходу до формування скатерометричного радіолокаційного зображення, який дозволяв би враховувати властивості шумів і сигналів, а також забезпечував би оптимальні характеристики за заданим критерієм ефективності – зокрема, мінімізації середньоквадратичної помилки.

Метою роботи є статистична оптимізація методу формування радіолокаційних зображень та аналіз оптимального алгоритму побудови скатерометричного зображення у часовій та частотній областях.

Основні операції оброблення передбачають квадратурне детектування прийнятих сигналів, виділення прямої та квадратурної компонент на проміжній частоті, перетворення Фур’є за координатою дальності, подальшу частотну декореляцію в спектральній області, побудову оцінки коефіцієнта віддзеркалення поверхні та згортку результату з декорельованим опорним сигналом.

Моделі сигналів, шумів та рівняння спостереження

Радіозображення, сформоване в РСА з ЛЧМ-сигналами, будується в прямокутній системі координат: азимут – дальність. Вісь дальності відображає відстань до цілі в напрямку лінії візування радара та відповідає моменту надходження відбитого сигналу (тобто часовій затримці), що перетворюється в дальність після обробки. Вісь азимута – це напрямок уздовж траєкторії руху носія, який визначається за змінами допплерівської частоти або зміщенням фази сигналу в часі. Початок відліку по дальності прив'язується до моменту реєстрації першого відбиття, а по азимуту – до центрального положення апертури, синтезованої під час проходження поверхні. Така система координат дозволяє точно визначити положення об’єкта на місцевості в просторі. Отримане зображення може бути представлене у вигляді інтенсивності відбитого сигналу або його радіометричних характеристик у заданій площині. Такий підхід забезпечує точне позиціювання об’єктів і високий рівень деталізації поверхні

Розвиток моделей віддзеркалених сигналів полягає у врахуванні просторово-неоднорідної структури коефіцієнту відбиття електромагнітних хвиль реальних поверхонь, що докладно розглянуто в попередньому розділі. Модель буде записана з урахуванням феноменологічного підходу до визначення розсіяних полів в антенних решітках [1]. Зондуючий сигнал представляє собою безперервний сигналі з лінійною частотною модуляцією, що змінює значення частоти за періодичною пилкоподібною функцією. Геометрія, що відповідає процесу вимірювання розсіяних сигналів РСА в строго боковому напрямку, показана на рис. 1.

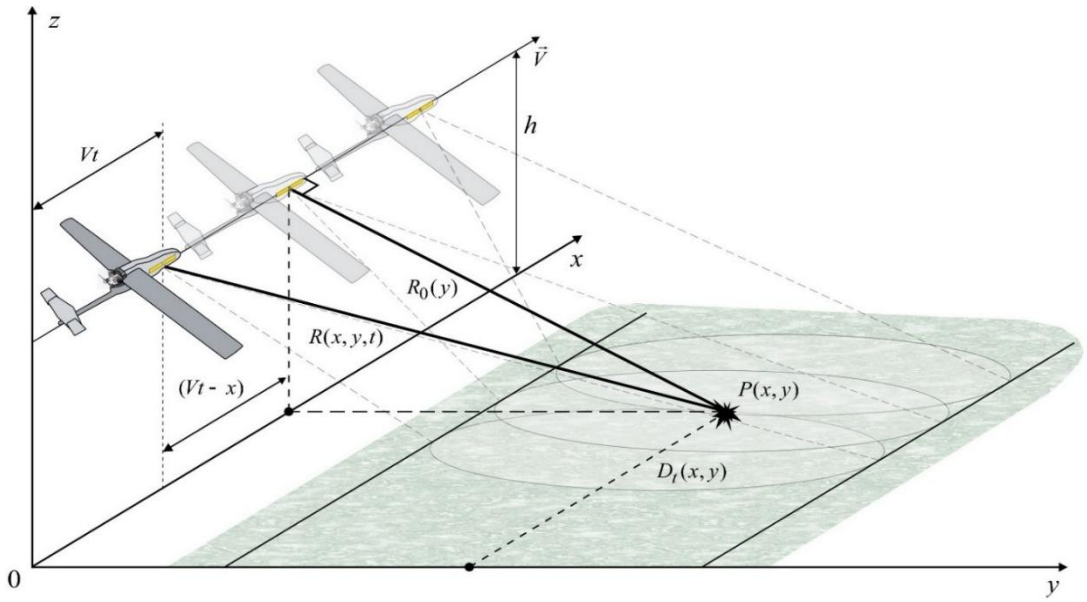


Рис. 1. Геометрія зондування поверхні з борта літального апарату

Відповідно до рис. 1 вважаємо що літальний апарат рухається прямолінійно вздовж осі x, на висоті h і з постійною швидкістю V. В передавачі формується зондуючий безперервний ЛЧМ сигнал, що подається на вхід передавальної антени. Передавальна антена перетворює сигнал на електромагнітні коливання та спрямовує їх в строго боковому напрямку від траєкторії польоту на поверхню землі. Просторово-часове поле у безпосередній близькості до поверхні у межах ділянки, що опромінена діаграмою спрямованості, запишемо наступним чином:

$$s_t(t,x,y) = D_t(x,y)\,A\sum_{i=0}^{\infty}\Pi_P(t-iT_P)\times\sum_{k=0}^{K}\Pi_s(t-iT_P-k\Delta t)\cos\left(2\pi f_0 t+2\pi\left(\alpha k\Delta t\right)\left(t-iT_P\right)+\varphi_0\right), \qquad (1)$$

де $D_t(x,y)$ – діаграма спрямованості передавальної антени, що представлена у координатах поверхні, або пляма діаграми спрямованості передавальної антени на поверхні, A – постійна амплітуда зондуючого сигналу, $\Pi_P(t-iT_P)$ – прямокутний імпульс, що визначає довжину одного періоду ЛЧМ сигналу, T_P – період пилкоподібної функції зміни частоти, i – номер періоду модуляції, $\Pi_s(t-iT_P-k\Delta t)$ – прямокутний імпульс, що визначає тривалість частини ЛЧМ сигналу у межах одного періоду і має постійне значення частоти, Δt – інтервал часу, на якому частота сигналу не змінюється, k – номер імпульсу $\Pi_s(\cdot)$, f_0 – початкова частота, φ_0 – початкова фаза, $\alpha=\frac{(F_{\max}-f_0)}{T_p}$ – нахил пилкоподібної функції зміни частоти, $(F_{\max}-f_0)$ – девіація частоти, $F_{\max}$ – максимальна частота, яку може набувати ЛЧМ сигнал.

Приклад сигналу (1), коли $f_0=10\ Гц$, $F_{\max}=100\ Гц$, $x=0\,м$, $y=0\,м$, $G_t(x,y)=1$, $T_P=1c$, $T_s=200\,мс$, представлено на рис. 2.

Досягнувши поверхні землі, сигнал (1) відбивається від кожної її точки з урахуванням комплексного коефіцієнту відбиття цієї точки $\dot{F}(x,y)$. В області спостереження, в приймальній антені радару, буде спостерігатися сигнал з урахуванням феноменологічного підходу до визначення розсіяних полів в антенних системах [1, 2]:

$$s_r(t) = \mathrm{Re}\left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} D_r(x,y)\, D_t(x,y)\, \dot{F}(x,y)\, \dot{s}_0(t,x,y)\, dxdy \right\}, \tag{2}$$

$$\dot{s}_0(t,x,y) = A\mathrm{e}^{j\varphi_0} \sum_{i=0}^{\infty} \Pi_P(t - iT_P - t_{del}(x,y)) \times \sum_{k=0}^{K} \Pi_s(t - iT_P - k\Delta t - t_{del}(x,y)) \times \\ \times \mathrm{e}^{j\left(2\pi f_0 (t - t_{del}(x,y)) + 2\pi(\alpha k \Delta t)(t - iT_P - t_{del}(x,y))\right)} \tag{3}$$

– одиничний сигнал, який би був прийнятий радаром від однієї точки поверхні $P(x,y)$ при $\dot{F}(x,y)=1$, $D_r(x,y)$ – діаграма спрямованості приймальної антени, що представлена у координатах поверхні, або пляма діаграми спрямованості приймальної антени на поверхні,

$$t_{del}(x,y) = \frac{2R(x,y,t)}{c} \tag{4}$$

– час затримки на розповсюдження сигналу від передавальної антени до кожної точки поверхні і в зворотному напрямку в приймальну антену, c – швидкість розповсюдження електромагнітних хвиль, $2R(x,y)$ – шлях, що проходять електромагнітні хвилі.

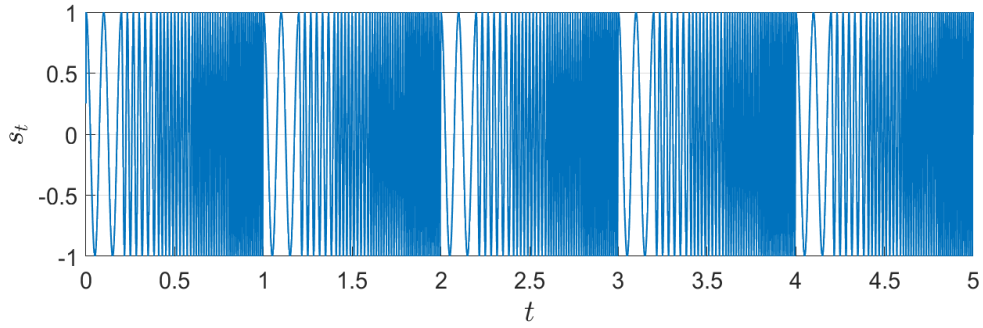


Рис. 2. Приклад моделювання сигналу (1) в одній точці на початку просторової системи координат

З урахуванням наведеної на рис. 1 геометрії запишемо дальність $R(x,y)$ до кожної точки поверхні від літального апарату наступним чином:

$$R(x,y,t) = \sqrt{(Vt - x)^2 + y^2 + h^2}\,, \tag{5}$$

де Vt – поточна координата положення літального апарату, що рухається вздовж осі x або шлях, що пролетів літальний апарат від початку системи координат.

Винесемо в (5) з-під знаку кореня дальність $R_0(y) = \sqrt{y^2 + h^2}$ до точки $P(x,y)$ при її спостереженні під прямим кутом азимуту

$$R(x,y,t) = R_0(y)\sqrt{1 + \frac{(Vt - x)^2}{R_0^2(y)}}\,. \tag{6}$$

Величина під знаком кореня $\frac{(Vt - x)^2}{R_0^2(y)} << 1$ при практичних вимірюваннях, адже максимальне значення $(Vt - x)$ – це половина поперечного розміру плями діаграми спрямованості за координатою x, а $R_0(y)$ – похила дальність. Дальність, зазвичай, в сотні та тисячі разів більша за пляму діаграми спрямованості. В такому випадку корінь може бути розкладений в ряд Тейлора наступним чином:

$$\sqrt{1+\frac{(Vt-x)^2}{R_0^2(y)}}=1+\frac{1}{2}\frac{(Vt-x)^2}{R_0^2(y)}-\ldots \tag{7}$$

Розкладання в (7) на більшу кількість членів ряду не має сенсу, так як нові складові для представленої геометрії не інформативні. Вираз (6) з урахуванням (7) набуде вигляду

$$R(x,y,t)=R_0(y)+\frac{(Vt-x)^2}{2R_0(y)}. \tag{8}$$

Підставляючи (8) і (4) в одиничний сигнал (3), отримаємо

$$\begin{gathered}\dot{s}_0(t,x,y)=\dot{A}\sum_{i=0}^{\infty}\Pi_P(t-iT_P-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})\\ \times\sum_{k=0}^{K}\Pi_s(t-iT_P-k\Delta t-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})\,\mathrm{e}^{j2\pi f_0 t}\times\\ \times\mathrm{e}^{-j2\pi f_0 2R_0(y)c^{-1}}\,\mathrm{e}^{-j2\pi f_0(Vt-x)^2R_0^{-1}(y)c^{-1}}\times\mathrm{e}^{j2\pi(\alpha k\Delta t)(t-iT_P-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})},\end{gathered} \tag{9}$$

де $\dot{A}=A\mathrm{e}^{j\varphi_0}$ – комплексна огинаюча зондуючого сигналу.

Винесемо в отриманому виразі (9) за знаки суми всі складові, що не залежать від індексів i і k та отримаємо остаточний вигляд одиничного сигналу

$$\begin{gathered}\dot{s}_0(t,x,y)=\dot{A}\mathrm{e}^{j2\pi f_0 t}\,\mathrm{e}^{-j2\pi f_0 2R_0(y)c^{-1}}\times\mathrm{e}^{-j2\pi f_0(Vt-x)^2R_0^{-1}(y)c^{-1}}\times\sum_{i=0}^{\infty}\Pi_P(t-iT_P-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})\\ \times\sum_{k=0}^{K}\Pi_s(t-iT_P-k\Delta t-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})\times\mathrm{e}^{j2\pi(\alpha k\Delta t)(t-iT_P-2R_0(y)c^{-1}-(Vt-x)^2R_0^{-1}(y)c^{-1})}.\end{gathered} \tag{10}$$

При реєстрації сигналів приймачем до них додаються внутрішні шуми $n(t)$.

Коливання, що підлягатимуть подальшій оптимальній обробці, будемо називати рівняннями спостереження. Для розглянутих умов рівняння спостереження має вигляд

$$u(t)=s_r(t)+n(t). \tag{11}$$

У виразі (12) випадковими є внутрішні шуми $n(t)$ і комплексний коефіцієнт розсіювання $\dot{F}(x,y)$. Обидва ці процеси є білими шумами, що мають нульове математичне сподівання і наступні кореляційні функції:

$$R_n(t_1,t_2)=\langle n(t_1)n(t_2)\rangle=\frac{1}{2}N_{0n}\delta(t_1-t_2), \tag{12}$$

$$R_{\dot{F}}(x_1,x_2,y_1,y_2)=\langle\dot{F}(x_1,y_1)\dot{F}(x_2,y_2)\rangle=\sigma^0(x_1,y_1)\,\delta(x_1-x_2)\,\delta(y_1-y_2), \tag{13}$$

де $\langle\cdot\rangle$ – знак статистичного усереднення по ансамблю реалізацій (математичне сподівання), $\sigma^0(x_1,y_1)$ – питома ефективна поверхня розсіювання поверхні, що буде підлягати оцінюванні в даному розділі.

З урахуванням (12) і (13) кореляційна функція рівняння спостереження має вигляд

$$\begin{gathered}R_u(t_1,t_2)=\langle u(t_1)u(t_2)\rangle=\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x,y)D_r(x,y)D_t(x,y)\,\dot{s}_0(t_1,x,y)\,\dot{s}_0^*(t_2,x,y)dxdy+\\ +\frac{1}{2}N_{0n}\,\delta(t_1-t_2).\end{gathered} \tag{14}$$

Кореляційна функція (14) містить всю необхідну інформацію для вирішення оптимізаційної задачі. Питома ефективна поверхня розсіювання $\sigma^0(x,y)$ є статистичною характеристикою комплексного коефіцієнту віддзеркалення поверхні і, на відміну від класичних задач радіолокації, потребує оцінки енергетичних параметрів, а саме складових $R_u(t_1,t_2)$. Параметр $\sigma^0(x,y)$ не входить в явному вигляді в рівняння спостереження, але підлягає оцінюванню і буде включений в $R_u(t_1,t_2,\sigma^0(x,y))$ і в функцію $W_u(t_1,t_2,\sigma^0(x,y))$, що їй обернена.

Постановка задачі

За результатами прийому бортовим когерентним радаром, розміщеного на літальному апараті, віддзеркаленого від поверхні корисного сигналу $s_r(t)$, що спостерігається на фоні внутрішніх шумів приймача $n(t)$, необхідно оптимальним чином, в рамках методу максимальної правдоподібності, сформувати радіозображення підстильної поверхні, що представлено у вигляді енергетичного параметру $\sigma^0(x,y)$, як складової статистичних характеристик прийнятих коливань $u(t)$.

Статистична оптимізація методу формування радіозображень

Оптимальний метод радіобачення підстильної поверхні визначимо в рамках методу максимальної правдоподібності, що для корельованих стохастичних процесів потребує знаходження максимуму наступної функції правдоподібності:

$$p[u(t)\,|\,\sigma^0(x,y)] = \kappa[\sigma^0(x,y)]\exp\left\{-\frac{1}{2}\int_0^T\int_0^T [u(t_1)-m_u(t_1)]\times W_u(t_1,t_2,\sigma^0(x,y))[u(t_2)-m_u(t_2)]dt_1dt_2\right\}. \quad (15)$$

У вираз (15) входить деякий множник $\kappa[\sigma^0(x,y)]$, що залежить від енергетичних параметрів, постійна часу спостереження T, обернена кореляційна функція $W_u(t_1,t_2,\sigma^0(x,y))$ і математичне сподівання $m_u(t,\lambda)$ процесу $u(t)$. Математичне сподівання

$$m_u(t) = \langle u(t)\rangle = \langle s_r(t)+n(t)\rangle = \langle s_r(t)\rangle + \langle n(t)\rangle = 0\,, \quad (16)$$

тому (15) може бути переписано наступним чином:

$$p[u(t)\,|\,\sigma^0(x,y)] = \kappa[\sigma^0(x,y)]\times\exp\left\{-\frac{1}{2}\int_0^T\int_0^T u(t_1)W_u(t_1,t_2,\sigma^0(x,y))u(t_2)dt_1dt_2\right\}. \quad (17)$$

Знаходячи математичний вираз, що відповідає точці максимуму (17), отримаємо необхідні математичні операції для формування оптимальних оцінок радіозображення підстильної поверхні $\hat{\sigma}^0(x,y)$, де $\hat{(\cdot)}$ – символ позначення оцінки. Оцінка буде відрізнятися від істинного значення на величину граничної похибки вимірювань.

Функція $W(t_1,t_2,\sigma^0(x,y))$ визначається через інтегральне рівняння

$$\int_0^T R_u(t_1,t_3,\sigma^0(x,y))W_u(t_3,t_2,\sigma^0(x,y))dt_3 = \delta(t_1-t_2)\,. \quad (18)$$

Максимум (17) можна було б знайти у результаті розрахунку похідної та прирівнювання її до нуля, але оцінюванню в поставленій задачі підлягає питома ефективна поверхня розсіювання поверхні, що є функцією просторових координат, і звичайна похідна не може бути

взята по функції. В такому випадку необхідно застосовувати математичний апарат варіаційної похідної. Представимо $\overset{0}{\sigma}(x,y)$ у вигляді двох доданків

$$\overset{0}{\sigma}(x,y)=\sigma^0(x,y)+\delta\sigma^0(x,y), \tag{19}$$

де $\sigma^0(x,y)$ – істинне значення, $\delta\sigma^0(x,y)$ – варіація оцінки, що визначається деяким відхилення від істинного значення,

$$\delta\sigma^0(x,y)=\chi\gamma(x,y), \tag{20}$$

$\gamma(x,y)$ – довільна недетермінована функція відхилення оцінки від істинного параметру з одиничною амплітудою, χ – величина відхилення. Зрозуміло, що при $\chi=0$ буде отримане істинне значення радіозображення поверхні. Таким чином, замість пошуку мінімуму функції (17) по функції $\sigma^0(x,y)$ можна визначити мінімум (17) взявши звичайну похідну по χ в точці $\chi=0$. В результаті критерій оптимізації набуде вигляду

$$\left.\frac{d\ p[u(t)\,|\,\sigma^0(x,y)]}{d\chi}\right|_{\chi=0}=0. \tag{21}$$

З виразу (17) також випливає, що функція правдоподібності є експоненціальною функцією з деяким коефіцієнтом. Застосування до (17) натуральний логарифм точка максимуму функції правдоподібності не зміниться. В такому випадку (21) можна переписати:

$$\left.\frac{d\ln p[u(t)\,|\,\sigma^0(x,y)]}{d\chi}\right|_{\chi=0}=0. \tag{22}$$

Результат диференціювання і прирівнювання (22) має назву рівняння правдоподібності [1] і для визначених умов задачі має вигляд

$$-\int_0^T\int_0^T\frac{d\ R_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\ \chi}W_u\left[t_1,t_2,\sigma^0(x,y)\right]dt_1dt_2=$$

$$=\int_0^T\int_0^T u(t_1)\frac{d\ W_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\ \chi}u(t_2)dt_1dt_2. \tag{23}$$

Для подальшого вирішення рівняння правдоподібності визначимо всі його складові. Обернена кореляційна функція може бути представлена через згортку з дельта-функцією

$$W_u\left[t_1,t_2,\sigma^0(x,y)\right]=\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\delta(t_2-t_3)\,dt_3\ . \tag{24}$$

І, підставляючи (18), отримаємо її узагальнений вигляд:

$$W_u\left[t_1,t_2,\sigma^0(x,y)\right]=\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]R_u\left[t_3,t_4,\sigma^0(x,y)\right]\times W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4. \tag{25}$$

Вираз (25) конкретизуємо для визначеної форми кореляційної функції (14)

$$W_u\left[t_1,t_2,\sigma^0(x,y)\right]=\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\times\left(\begin{array}{c}\frac{1}{2}\mathrm{Re}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\left[\sigma^0(x,y)\right]D_r(x,y)D_t(x,y)\dot{s}_0(t_1,x,y)\dot{s}_0^*(t_2,x,y)dxdy+\\ +\frac{1}{2}N_{0n}\,\delta(t_1-t_2)\end{array}\right)\times \tag{26}$$
$$\times W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4$$

або

$$W_u\left[t_1,t_2,\sigma^0(x,y)\right]=\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\sigma^0(x,y)\right]D_r(x,y)D_t(x,y)\times$$
$$\times\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\dot{s}_0(t_3,x,y)\dot{s}_0^*(t_4,x,y)W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4\,dx\,dy+$$
$$+\frac{1}{2}\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]N_{0n}\,\delta(t_3-t_4)W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4. \tag{27}$$

Визначивши загальний вигляд оберненої кореляційної функції, попередньо необхідно визначити похідну виразу (25) також у загальному вигляді

$$\frac{d\,W_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}=-\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\frac{d\,R_u\left[t_3,t_4,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}\times \tag{28}$$
$$\times W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4.$$

З виразу (28) випливає, що для подальшої його конкретизації необхідно визначити похідну (14) по параметру χ, що матиме вигляд

$$\frac{d\,R_u\left[t_3,t_4,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}=\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\dot{s}_0(t_3,x,y)\dot{s}_0^*(t_4,x,y)dxdy. \tag{29}$$

Отриманий вираз для похідної кореляційної функції підставляємо в вираз (28)

$$\frac{d\,W_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}=-\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\times \tag{30}$$
$$\times\dot{s}_0(t_3,x,y)\,\dot{s}_0^*(t_4,x,y)W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4\;dxdy$$

або

$$\frac{d\,W_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}=-\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\times \tag{31}$$
$$\times\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\dot{s}_0(t_3,x,y)dt_3\int_0^T W_u\left[t_4,t_2,\sigma^0(x,y)\right]\dot{s}_0^*(t_4,x,y)\,dt_4\;dxdy.$$

Інтеграли в (31), що описують згортку кореляційної функції з одиничним сигналом, є новим типом сигналів, що необхідно застосовувати при оптимальному оцінюванні статистичних характеристик випадкового коефіцієнту віддзеркалення поверхні. На відміну від відомих операцій ці згортки потребують додаткової інверсної фільтрації в фільтрі з частотною характеристикою, що дорівнює спектру оберненої кореляційної функції $W_u(\cdot)$. Інверсні філь-

три в теорії побудови радіосистем ще мають назву вибілюючими фільтрами, основна задача яких підвищити точність оцінювання параметрів випадкових процесів, в граничному випадку білого шуму. Позначимо ці згортки наступним чином:

$$\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]=\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\dot{s}_0(t_3,x,y)dt_3, \tag{32}$$

$$\dot{s}^*_{0W}\left[t_2,\sigma^0(x,y)\right]=\int_0^T W_u\left[t_4,t_2,\sigma^0(x,y)\right]\dot{s}^*_0(t_4,x,y)\,dt_4\ . \tag{33}$$

Підставляючи (32) і (33) в (31), отримаємо

$$\frac{d\,W_u\left[t_1,t_2,\sigma^0(x,y)+\chi\gamma(x,y)\right]}{d\,\chi}=-\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\,\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]\dot{s}^*_{0W}\left[t_2,\sigma^0(x,y)\right]\,dxdy. \tag{34}$$

Визначивши всі складові рівняння спостереження в (29) і (34), запишемо його наступним чином:

$$\begin{aligned}&\int_0^T\int_0^T\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\,\dot{s}_0(t_3,x,y)\,\dot{s}^*_0(t_4,x,y)dxdy\times W_u\left[t_1,t_2,\sigma^0(x,y)\right]dt_1dt_2-\\&-\int_0^T\int_0^T u(t_1)\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\times\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]\dot{s}^*_{0W}\left[t_2,\sigma^0(x,y)\right]\,dxdy\;u(t_2)dt_1dt_2=0\end{aligned} \tag{35}$$

або з урахуванням (27)

$$\begin{aligned}&\int_0^T\int_0^T\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\,\dot{s}_0(t_3,x,y)\,\dot{s}^*_0(t_4,x,y)dxdy\times\\&\times\left(\begin{array}{l}\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x,y)D_r(x,y)D_t(x,y)\times\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]\dot{s}_0(t_3,x,y)\times\\\qquad\times\dot{s}^*_0(t_4,x,y)W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4\,dx\,dy+\\+\frac{1}{2}\int_0^T\int_0^T W_u\left[t_1,t_3,\sigma^0(x,y)\right]N_{0n}\,\delta(t_3-t_4)W_u\left[t_4,t_2,\sigma^0(x,y)\right]dt_3dt_4\end{array}\right)dt_1dt_2-\\&-\int_0^T\int_0^T u(t_1)\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\times\\&\times\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]\dot{s}^*_{0W}\left[t_2,\sigma^0(x,y)\right]\,dxdy\;u(t_2)dt_1dt_2=0.\end{aligned} \tag{36}$$

Подальше спрощення можливе в результаті застосування в (36) методу комплексних огинаючих, що передбачає

$$\mathrm{Re}\,\dot{x}(t)\,\mathrm{Re}\,\dot{y}(t)=\frac{1}{2}\mathrm{Re}\,\dot{x}(t)\dot{y}(t)+\frac{1}{2}\mathrm{Re}\,\dot{x}(t)\dot{y}^*(t), \tag{37}$$

де $(\cdot)^*$ – знак комплексного спряження.

З урахуванням виразів (37) рівняння правдоподібності прийме вигляд

$$\operatorname{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\Bigg(\frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x_1,y_1)D_r(x_1,y_1)D_t(x_1,y_1)\times$$
$$\times\int_0^T \dot{s}_0(t_1,x,y)\dot{s}^*_{0W}(t_1,x_1,y_1)dt_1\int_0^T \dot{s}^*_0(t_2,x,y)\dot{s}_{0W}(t_2,x_1,y_1)dt_2\quad dx_1dy_1+ \tag{38}$$
$$+\frac{1}{2}\frac{N_{0n}}{2}\int_0^T\left|\dot{s}_{0W}\left[t_3,\sigma^0(x,y)\right]\right|^2 dt_3-\frac{1}{2}\int_0^T u(t_1)\,\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]dt_1\int_0^T u(t_2)\,\dot{s}^*_{0W}\left[t_2,\sigma^0(x,y)\right]\,dt_2\Bigg)dxdy=0.$$

В отриманому виразі всі математичні операції мають фізичний сенс та відповідають відомим в теорії оптимізації обробки сигналів функціям, тому перепишемо його наступним чином:

$$\operatorname{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\gamma(x,y)\,D_r(x,y)D_t(x,y)\times\Bigg(\frac{1}{4}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x_1,y_1)D_r(x_1,y_1)D_t(x_1,y_1)\left|\dot{\Psi}_W(x,y,x_1,y_1)\right|^2 dx_1dy_1+$$
$$+\frac{N_{0n}}{2}E_W(x,y)-\frac{1}{2}\left|\dot{Y}(x,y)\right|^2\Bigg)dxdy=0\,, \tag{39}$$

де

$$\dot{\Psi}_W(x,y,x_1,y_1)=\int_0^T \dot{s}_0(t_1,x,y)\dot{s}^*_{0W}(t_1,x_1,y_1)dt_1 \tag{40}$$

– функція розузгодженості за просторовими координатами підстильної поверхні для радара з обробкою безперервних ЛЧМ сигналів. Функцію (40) ще називають функцією невизначеності або реакцією системи формування зображення на точкове джерело випромінювання. Ця функція визначає роздільну здатність радару, що реалізує нову запропоновану обробку з декореляцією.

В (40) також введено енергію одиничного сигналу з урахуванням декореляції

$$E_W(x,y)=\frac{1}{2}\int_0^T\left|\dot{s}_{0W}\left[t_3,\sigma^0(x,y)\right]\right|^2 dt_3\,. \tag{41}$$

Вираз

$$\dot{Y}(x,y)=\int_0^T u(t_1)\,\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]dt_1 \tag{42}$$

є ключовою складовою (39), адже він описує оптимальну обробку прийнятих коливань – рівняння спостереження $u(t)$. Сутність обробки полягає в розрахунку кореляційного інтегралу або узгодженій фільтрації $u(t)$ в фільтрі з імпульсною характеристикою, що повторює одиничний сигнал. На відміну від існуючих рішень, одиничний сигнал включає операцію декореляції та дозволяє збільшити кількість некорельованих відліків в вихідному ефекті $\dot{Y}(x,y)$. Одиничний сигнал в такому випадку застосування також називають опорним сигналом і в задачах відновлення радіозображень він може бути факторизований на опорний сигнал стиснення отриманих коливань за азимутом та за дальністю.

Достатньою умовою для виконання рівняння (39) є наступна рівність:

$$\left|\dot{Y}(x,y)\right|^2=\frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x_1,y_1)D_r(x_1,y_1)D_t(x_1,y_1)\times\left|\dot{\Psi}_W(x,y,x_1,y_1)\right|^2 dx_1dy_1+N_{0n}\,E_W(x,y)\,. \tag{43}$$

Досягти рівності (43) дуже важко, адже в правій стороні залишились детерміновані функції часу та просторових координат, а в лівій стороні – результат оптимальної обробки, що

характеризується деякою потенційною точністю. Кореляційний інтеграл або фільтрація в узгодженому фільтрі корисного сигналу з шумами ніколи не дасть детермінованої функції, завжди залишиться недоусереднений залишок. Цей недоусереднений залишок і не дає досягнути рівності, а його дисперсія визначає похибку оцінки радіозображення поверхні. З практики формування радіолокаційних зображень [1] відомо, що зображення отримані алгоритмом (42) має плямисту структуру, що називається сепекл-шумом. Цей спекл-шум і є недоусередненим залишком впливу шумів та випадкової структури комплексного коефіцієнту віддзеркалення реальних поверхонь. Таким чином, доцільно в (43) говорити про прирівнювання лівої та правої частини. Повна рівність досягається лише при статистичному усередненні лівої частини за ансамблем реалізацій

$$\left\langle \left|\dot{Y}(x,y)\right|^2 \right\rangle = \frac{1}{2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sigma^0(x_1,y_1) D_r(x_1,y_1) D_t(x_1,y_1) \times \left|\dot{\Psi}_W(x,y,x_1,y_1)\right|^2 dx_1 dy_1 + N_{0n} E_W(x,y) . \quad (44)$$

Аналіз оптимального алгоритму формування скатерометричних радіолокаційних зображень у часовій області

Для аналізу отриманого алгоритму (42) послідовно підставимо в отриманий вираз формулу для одиничного сигналу з декореляцією (33)

$$\dot{Y}(x,y) = \int_0^T u(t) \int_0^T W_u\left[t, t_1, \sigma^0(x,y)\right] \dot{s}_0(t_1, x, y) dt_1 dt \quad (45)$$

і одиничний сигнал (9)

$$\begin{aligned}\dot{Y}(x,y) = \int_0^T u(t) \int_0^T W_u\left[t, t_1, \sigma^0(x,y)\right] \times \dot{A} \sum_{i=0}^{\infty} \Pi_P(t_1 - iT_P - 2R_0(y)c^{-1} - (Vt_1 - x)^2 R_0^{-1}(y)c^{-1}) \\ \times \sum_{k=0}^{K} \Pi_s(t_1 - iT_P - k\Delta t - 2R_0(y)c^{-1} - (Vt_1 - x)^2 R_0^{-1}(y)c^{-1}) \mathrm{e}^{j2\pi f_0 t_1} \times \\ \times \mathrm{e}^{-j2\pi f_0 2R_0(y)c^{-1}} \mathrm{e}^{-j2\pi f_0 (Vt_1 - x)^2 R_0^{-1}(y)c^{-1}} \times \mathrm{e}^{j2\pi(\alpha k\Delta t)(t_1 - iT_P - 2R_0(y)c^{-1} - (Vt_1 - x)^2 R_0^{-1}(y)c^{-1})} dt_1\, dt .\end{aligned} \quad (46)$$

Одразу проаналізувати фізичну сутність алгоритму фільтрації прийнятих коливань в фільтрі, імпульсна характеристика якого співпадає з сумою затриманих у часі декорельованих високочастотних імпульсів достатньо складно. Доцільно спочатку визначити основні операції в часовій області без декореляції, вважаючи, що $W_u\left[t, t_1, \sigma^0(x,y)\right] = \delta(t - t_1)$. В такому випадку вираз (46) має вигляд

$$\begin{aligned}\dot{Y}(x,y) = \mathrm{e}^{-j2\pi f_0 2R_0(y)c^{-1}} \times \int_0^T dt\, \mathrm{e}^{-\frac{j2\pi f_0 (Vt - x)^2}{R_0(y)c}} \sum_{i=0}^{\infty} \Pi_P(t - iT_P - 2R_0(y)c^{-1} - (Vt - x)^2 R_0^{-1}(y)c^{-1}) \times \\ \times \sum_{k=0}^{K} \left[u(t) \left[\begin{array}{c} \Pi_s(t - iT_P - k\Delta t - 2R_0(y)c^{-1} - (Vt - x)^2 R_0^{-1}(y)c^{-1}) \times \\ \times \dot{A}\, \mathrm{e}^{j(2\pi f_0 t + 2\pi(\alpha k\Delta t)(t - iT_P))} \end{array} \right] \times \mathrm{e}^{-j2\pi(\alpha k\Delta t)\frac{(Vt - x)^2}{R_0(y)c}} \right] \times \mathrm{e}^{-j2\pi(\alpha k\Delta t)\frac{2R_0(y)}{c}}\end{aligned} \quad (47)$$

Розглянемо аргументи деяких функцій, що входять в (47). Нехай висота польоту літального апарату дорівнює 1500 м, максимальний кут візування від надиру 45º, тоді максимальна дальність $R_{0\max} = 2121{,}3$ *м*, час затримки $\frac{2R_{0\max}}{c} = 14{,}1$ *мкс*. Максимальне відхилення координати x визначається половиною ширини діаграми спрямованості вдовж координати азимута і у випадку використання бортової антени розміром 30 см, проведення вимірювань на частоті 10 ГГц дорівнює 106 м. Нехай період пилкоподібної функції зміни частоти більше

ніж в два рази перевищує максимальний час затримки сигналу і дорівнює $T_P = K\Delta t = 30 мкс$. Девіація частоти будемо вважати складає 1 % від центральної частоти і дорівнює 100 МГц, тоді $\alpha = \frac{(F_{\max} - f_0)}{T_p} = 333 \cdot 10^{10}$. Максимальний час спостереження T визначається шириною плями діаграми спрямованості на поверхні вздовж координати азимута і швидкістю рух літального апарату, $T = \frac{\Delta G(x)}{V}$. Для швидкості руху 90 км/год максимальний час спостереження буде приблизно дорівнювати $T = 8{,}5\, c$. Для наведених величин було проведене імітаційне моделювання функцій $\exp\left(-j2\pi(\alpha k\Delta t)\frac{(Vt-x)^2}{R_0(y)c}\right)$ і. $\exp\left(-\frac{j2\pi f_0(Vt-x)^2}{R_0(y)c}\right)$, результат якого наведено на рис. 3.

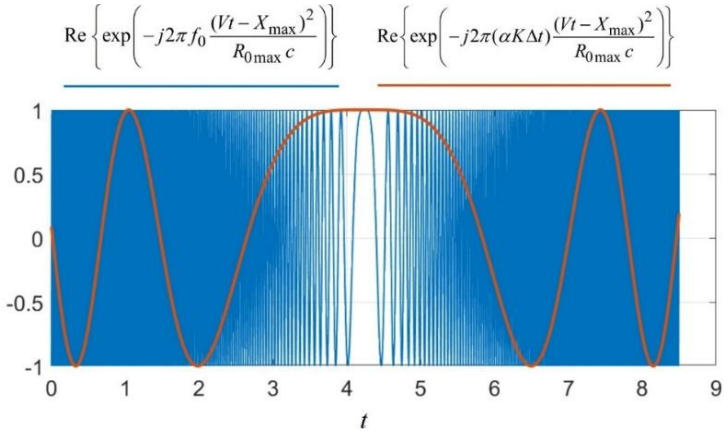


Рис. 3. Імітаційне моделювання опорних функцій одиничного сигналу

Функція $\exp\left(-\frac{j2\pi f_0(Vt-x)^2}{R_0(y)c}\right)$ відома з класичних алгоритмів формування радіозображень з викою роздільною здатністю, коли використовується імпульсний режим роботи радара. Ця функція є базисною для відновлення дрібних деталей на первинному зображенні або іншими словами – для синтезування апертури антени. В порівнянні з $\exp\left(-\frac{j2\pi f_0(Vt-x)^2}{R_0(y)c}\right)$, функція $\exp\left(-j2\pi(\alpha k\Delta t)\frac{(Vt-x)^2}{R_0(y)c}\right)$ є більш низькочастотною та не дозволить відновити малорозмірні деталі на радіозображеннях, тому в подальшому може не враховуватися. Чим менше буде відношення девіації частоти зондуючого сигналу до несучої, тим менше буде вплив множника $\exp\left(-j2\pi(\alpha k\Delta t)\frac{(Vt-x)^2}{R_0(y)c}\right)$.

Також необхідно сказати, що максимальна затримка обумовлена розширення діаграми спрямованості за координатою азимута $\frac{(Vt-x)^2}{R_0(y)c}$ складає 70 нс, що в 200 разів менше за максимальну затримку по дальності. Таким чином затримкою прямокутних імпульсів на величину $\frac{(Vt-x)^2}{R_0(y)c}$ можна знехтувати, так само, як і залежністю $\frac{2R_0(y)}{c}$ від y. Дійсно

підібрати затримку прямокутних імпульсів під кожну дальність неможливо, тому будемо використовувати деяке $R_{0\min}$. З урахуванням введених припущень, алгоритм (47) набуде наступного вигляду:

$$\dot{Y}_1(x,y)=\int_0^T dt\left(\sum_{i=0}^{\infty}\Pi_P\left(t-iT_P\right)\dot{U}_{\times}(t-iT_P,y)\right)\mathrm{e}^{-\frac{j2\pi f_0(Vt-x)^2}{R_0(y)c}}, \tag{48}$$

де

$$\dot{U}_{\times}(t-iT_P,y)=\sum_{k=0}^{K}\dot{u}_{\times}(t-iT_P,k\Delta t)\,\mathrm{e}^{-j2\pi(\alpha k\Delta t)\frac{2R_0(y)}{c}} \tag{49}$$

– дискретне перетворення Фур'є від результату перемноження прийнятого рівняння спостереження на всі можливі імпульси з частотами $f_0+\alpha k\Delta t$, що визначається виразом

$$\dot{u}_{\times}(t-iT_P,k\Delta t)=u(t)\,\Pi_s\left(t-iT_P-k\Delta t-\frac{2R_{0\min}}{c}\right)\dot{A}\,\mathrm{e}^{j\left(2\pi f_0 t+2\pi(\alpha k\Delta t)(t-iT_P)\right)}. \tag{50}$$

Точно визначити час приходу віддзеркалених від поверхні сигналів неможливо при використанні неперервних сигналів, тому доцільно у якості множника $\Pi_s\left(t-iT_P-k\Delta t-\frac{2R_{0\min}}{c}\right)\dot{A}\,\mathrm{e}^{j(2\pi f_0 t+2\pi(\alpha k\Delta t)(t-iT_P))}$ в (50) використовувати копію зондуючого сигналу і його квадратурну складову.

Аналіз оптимального алгоритму формування скатерометричних радіолокаційних зображень у частотній області

Проаналізувавши основні оптимальні операції, що визначаються одиничним сигналом, доцільно розглянути вплив декорелюючого фільтра. Визначити аналітичний вигляд $W_u\left[t,t_1,\sigma^0(x,y)\right]$ з (18) достатньо складно, тому пропонується визначити форму декорелюючого фільтра спектральній області. Для цього спочатку представимо кореляційну функцію у вигляді суми

$$R_u(t_1,t_3,\sigma^0(x,y))=R_s(t_1,t_3,\sigma^0(x,y))+R_n(t_1-t_3), \tag{51}$$

де $R_s(t_1,t_3)=\frac{1}{2}\mathrm{Re}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sigma^0(x,y)D_r(x,y)D_t(x,y)\,\dot{s}_0(t_1,x,y)\,\dot{s}_0^*(t_3,x,y)dxdy$,

$R_n(t_1-t_3)=\langle u(t_1)u(t_3)\rangle=\frac{1}{2}N_{0n}\,\delta(t_1-t_3).$

Отриману суму (51) підставляємо в (18)

$$\int_0^T R_s(t_1,t_3,\sigma^0(x,y))+R_n(t_1-t_3)W_u(t_3,t_2,\sigma^0(x,y))dt_3=\delta(t_1-t_2) \tag{52}$$

або

$$\int_0^T R_s(t_1,t_3,\sigma^0(x,y))W_u(t_3,t_2,\sigma^0(x,y))dt_3+\frac{1}{2}N_{0n}W_u(t_1,t_2,\sigma^0(x,y))=\delta(t_1-t_2). \tag{53}$$

Рівняння (53) є інтегральним рівняннями Фредгольма першого роду, для вирішення якого необхідно перейти в спектральну область, визначити спектр $W_u(\cdot)$ і застосувати обернене

перетворення Фур'є. Процеси $u(t)$, що підлягають обробці, та, зокрема, корисний сигнал є нестаціонарними випадковими процесами. Про їх нестаціонарність свідчить кореляційна функція (14), що залежить не лише від різниці $(t_1 - t_2)$, а і від часу t_1. В такому випадку для визначення спектральної щільності нестаціонарного випадкового процесу необхідно використовувати подвійне перетворення Фур'є за змінними t_1 і t_2.

Перетворення Фур'є за змінною t_1 від лівої та правої частей (53) запишемо наступним чином:

$$\int_{-\infty}^{\infty}\left(\begin{array}{c}\int_0^T R_s(t_1,t_3,\sigma^0(x,y))W_u(t_3,t_2,\sigma^0(x,y))dt_3 + \\ \\ \frac{1}{2}N_{0n}W_u(t_1,t_2,\sigma^0(x,y))\end{array}\right)e^{-j2\pi f_1 t_1}dt_1 = \int_{-\infty}^{\infty}\delta(t_1-t_2)e^{-j2\pi f_1 t_1}dt_1 \tag{54}$$

або, змінивши порядок інтегрування,

$$\int_0^T\left(\int_{-\infty}^{\infty}R_s(t_1,t_3,\sigma^0(x,y))e^{-j2\pi f_1 t_1}dt_1\right)W_u(t_3,t_2,\sigma^0(x,y))dt_3 + \frac{1}{2}N_{0n}\int_{-\infty}^{\infty}W_u(t_1,t_2,\sigma^0(x,y))e^{-j2\pi f_1 t_1}dt_1 = e^{-j2\pi f_1 t_2}. \tag{55}$$

Вираз в дужках є спектральною щільністю потужності нестаціонарного випадкового процесу і буде позначатися $G_{R_s}(f_1,t_3,\sigma^0(x,y))$, а перетворення Фур'є від $W_u(t_1,t_2,\sigma^0(x,y)$ – $G_W(f_1,t_2,\sigma^0(x,y))$,

$$\int_0^T G_{R_s}(f_1,t_3,\sigma^0(x,y))W_u(t_3,t_2,\sigma^0(x,y))dt_3 + \frac{N_{0n}}{2}G_W(f_1,t_2,\sigma^0(x,y)) = e^{-j2\pi f_1 t_2}. \tag{56}$$

Тепер застосуємо перетворення Фур'є за змінною t_2

$$\int_0^T G_{R_s}(f_1,t_3,\sigma^0(x,y))\left(\int_{-\infty}^{\infty}W_u(t_3,t_2,\sigma^0(x,y))e^{-j2\pi f_2 t_2}dt_2\right)dt_3 + \frac{N_{0n}}{2}\int_{-\infty}^{\infty}G_W(f_1,t_2,\sigma^0(x,y))e^{-j2\pi f_2 t_2}dt_2 = \int_{-\infty}^{\infty}e^{-j2\pi(f_1+f_2)t_2}\,dt_2. \tag{57}$$

Права сторона (57) дорівнює дельта-функції, а в лівій введемо функцію

$$G_W(t_3,f_2,\sigma^0(x,y)) = \int_{-\infty}^{\infty}W_u(t_3,t_2,\sigma^0(x,y))e^{-j2\pi f_2 t_2}dt_2. \tag{58}$$

Після розрахунків отримаємо новий вираз для (57)

$$\int_0^T G_{R_s}(f_1,t_3,\sigma^0(x,y))G_W(t_3,f_2,\sigma^0(x,y))\,dt_3 + \frac{N_{0n}}{2}G_W(f_1,f_2,\sigma^0(x,y)) = \int_{-\infty}^{\infty}e^{-j2\pi(f_1+f_2)t_2}\,dt_2. \tag{59}$$

Для подальшого визначення $G_W(\cdot)$ застосуємо перетворення Фур'є до обох частин (59)

$$\int_{-\infty}^{\infty}\left(\begin{array}{c}\int_0^T G_{R_s}(f_1,t_3,\sigma^0(x,y))G_W(t_3,f_2,\sigma^0(x,y))\,dt_3 + \\ \\ +\frac{N_{0n}}{2}G_W(f_1,f_2,\sigma^0(x,y))\end{array}\right)e^{-j2\pi f_3 t_3}dt_3 = \delta(f_1+f_2)\int_{-\infty}^{\infty}e^{-j2\pi f_3 t_3}\,dt_3. \tag{60}$$

В лівій частині перетворення Фур'є від згортки двох функцій по змінній t_3 є добутком узагальнених спектральних щільностей потужності, а в правій – дельта-функція,

$$\left(G_{R_s}(f_1, f_2, \sigma^0(x,y)) + \frac{N_0}{2} \right) G_W(f_1, f_2, \sigma^0(x,y)) = \delta(f_1 + f_2). \tag{61}$$

З (61) випливає, що для знаходження узагальненої спектральної щільності потужності $G_W(f_1, f_2, \sigma^0(x,y))$ необхідно знайти $G_R^{-1}(f_1, f_2, \sigma^0(x,y))$ або, більш точно,

$$G_W(f_1, f_2, \sigma^0(x,y)) = \frac{\delta(f_1 + f_2)}{G_{R_s}(f_1, f_2, \sigma^0(x,y)) + \frac{N_0}{2}}. \tag{62}$$

Для подальшого використання (62) необхідно представити в спектральній області декорельований одиничний сигнал з (45). Застосуємо спочатку перетворення Фур'є до цього сигналу за змінною t

$$\begin{aligned} \int_{-\infty}^{\infty} \left(\int_0^T W_u\left[t, t_1, \sigma^0(x,y) \right] \dot{s}_0(t_1, x, y) dt_1 \right) e^{-j2\pi ft} dt &= \int_0^T \dot{s}_0(t_1, x, y) \left(\int_{-\infty}^{\infty} W_u\left[t, t_1, \sigma^0(x,y) \right] e^{-j2\pi ft} dt \right) dt_1 = \\ &= \int_0^T \dot{s}_0(t_1, x, y) G_W\left(f, t_1, \sigma^0(x,y) \right) dt_1. \end{aligned} \tag{63}$$

Тепер представимо інтеграл від добутку функцій в спектрі

$$\begin{aligned} \int_0^T \dot{s}_0(t_1, x, y) \left(\int_{-\infty}^{\infty} G_W\left(f, f_1, \sigma^0(x,y) \right) e^{j2\pi f_1 t_1} df_1 \right) dt_1 &= \left(\int_{-\infty}^{\infty} G_W\left(f, f_1, \sigma^0(x,y) \right) \int_0^T \dot{s}_0(t_1, x, y)\, e^{j2\pi f_1 t_1} dt_1\, df_1 \right) = \\ &= \int_{-\infty}^{\infty} G_W\left(f, f_1, \sigma^0(x,y) \right) \dot{S}_0(f_1, x, y)\, df_1. \end{aligned} \tag{64}$$

Підставляючи (62) в (64), отримаємо

$$\int_{-\infty}^{\infty} \frac{\delta(f + f_1)}{G_{R_s}(f, f_1, \sigma^0(x,y)) + \frac{N_0}{2}} \dot{S}_0(f_1, x, y)\, df_1 = \frac{\dot{S}_0(-f, x, y)}{G_{R_s}(f, -f, \sigma^0(x,y)) + \frac{N_0}{2}}. \tag{65}$$

Для остаточного визначення фізичного впливу операції декореляції визначимо узагальнену спектральною щільністю потужності нестаціонарного випадкового корисного сигналу $s_r(t)$. Для цього застосуємо послідовно перетворення Фур'є від (14) за змінними t_1 і t_2

$$\begin{aligned} G_{R_s}(f_1, f_2, \sigma^0(x,y)) &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} R_s(t_1, t_2) e^{-j2\pi f_1 t_1} dt_1 \right) e^{-j2\pi f_2 t_2} dt_2 = \frac{1}{2} \mathrm{Re} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^0(x,y) D_r(x,y) D_t(x,y) \times \\ &\times \int_{-\infty}^{\infty} \dot{s}_0(t_1, x, y) e^{-j2\pi f_1 t_1} dt_1 \int_{-\infty}^{\infty} \dot{s}_0^*(t_2, x, y) e^{-j2\pi f_2 t_2} dt_2\, dxdy = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sigma^0(x,y) D_r(x,y) D_t(x,y) \frac{1}{2} \mathrm{Re}\left(\dot{S}_0(f_1, x, y) \dot{S}_0^*(f_2, x, y) \right) dxdy. \end{aligned} \tag{66}$$

Підставляючи (66) в (65), отримуємо одиничний сигнал, що декорельований у спектральній області

$$\frac{\dot{S}_0^*(f,x,y)}{\frac{1}{2}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\sigma^0(x,y)D_r(x,y)D_t(x,y)\,\mathrm{Re}\left\{\dot{S}_0^2(f,x,y)\right\}dxdy+\frac{N_0}{2}}. \quad (67)$$

Необхідно підкреслити, що в (67) одиничний сигнал в чисельнику є комплексно спряженим до випроміненого. Така обробка повністю відповідає узгодженій обробці з класичної теорії оптимального оцінювання параметрів сигналу. Новою операцією є декореляція, що сприяє підвищенню кількості некорельованих відліків при обробці стохастичних сигналів.

Для визначення $\dot{s}_{0W}\left[t_1,\sigma^0(x,y)\right]$ в новому оптимальному алгоритмі (41) необхідно застосувати обернене перетворення Фур'є до (67).

На рис. 4 наведено приклад амплітудного спектру вихідного сигналу, інверсного фільтру та одиничного сигналу з декореляцією при $\frac{1}{\frac{1}{2}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\sigma^0(x,y)D_r(x,y)D_t(x,y)\,\mathrm{Re}\left\{\dot{S}_0^2(f,x,y)\right\}dxdy+\frac{N_0}{2}}$ співвідношенні сигнал-завада 10 дБ.

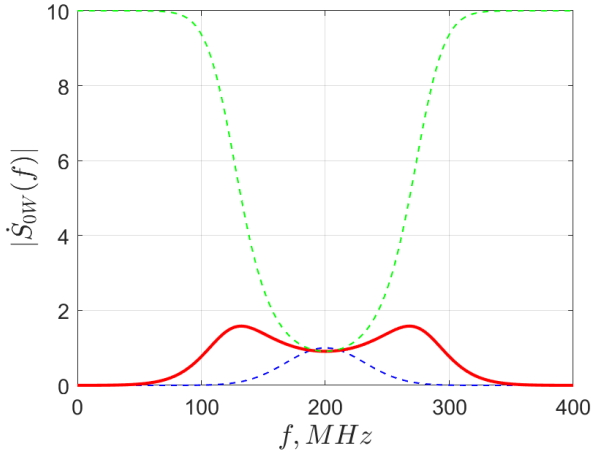


Рис. 4. Амплітудні спектри: синя штрихова лінія – випроміненого сигналу, зелена штрихова лінія – інверсного фільтру, червона суцільна лінія – декорельованого одиничного сигналу

З отриманих графіків випливає, що при перевищенні сигналу над шумами опорний сигнал має більш розширений спектр та збільшені амплітуди в областях згасання спектру зондуючого сигналу. З огляду на випадковість віддзеркалених сигналів і їх широкий спектр одиничний сигнал адаптується під прийняті коливання, розширюючи полосу пропускання. В той самий час розширення до нескінченності не відбувається, ступінь декореляції та адаптація під варіацію амплітуди і фази віддзеркаленого від поверхні сигналу визначається співвідношенням сигнал-завада.

Висновки

Запропоновано статистично обґрунтований метод формування скатерометричних радіолокаційних зображень, що базується на оптимальному обробленні стохастичних сигналів. Розвинуті моделі зондуючих і відбитих сигналів дозволяють враховувати фізичні та статистичні властивості віддзеркалення від поверхонь. Алгоритм реалізує операції детектування, перетворення Фур'є, декореляції та оцінювання коефіцієнтів віддзеркалення, що забезпечує підвищену точність зображення. Застосування критерію мінімізації середньоквадратичної похибки дозволило оцінити ефективність методу через межу Крамера–Рао. Порівняльний аналіз у часовій та частотній областях підтвердив доцільність такого підходу для бортових РСА з ЛЧМ сигналами. Результати можуть бути використані в

подальших практичних розробках систем дистанційного зондування та аерокосмічного радіобачення.

Список літератури:

1. Wang Z., Bovik A. C., Sheikh H. R., & Simoncelli E. P. Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE Transactions on Image Processing. 2004. Vol. 13, No. 4.
2. Damera-Venkata N., Kite T. D., Geisler W. S., Evans B. L., & Bovik A. C. Image Quality Assessment Based on a Degradation Model // IEEE Transactions on Image Processing. 2000. Vol. 9, No. 4.
3. Wang Z. et al. Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE Transactions on Image Processing. 2004.
4. Eskicioglu A. M., & Fisher P. S. Image Quality Measures and Their Performance // IEEE Transactions on Communications. 1995.
5. Shnayderman A., Gusev A., & Eskicioglu A. M. An SVD-Based Grayscale Image Quality Measure for Local and Global Assessment // IEEE Transactions on Image Processing. 2006. Vol. 15, No. 2.
6. Sheikh H. R., & Bovik A. C. Image Information and Visual Quality // IEEE Transactions on Image Processing. 2006

Надійшла до редколегії 20.06.2025

Відомості про авторів:

Жила Семен Сергійович – доктор технічних наук, доцент, Національний аерокосмічний університет «Харківський авіаційний інститут», завідувач кафедри аерокосмічних радіоелектронних систем; Україна; email: s.zhyla@khai.edu; ORCID: https://orcid.org/0000-0003-2989-8988

Одокієнко Олексій Володимирович – кандидат технічних наук, Національний аерокосмічний університет «Харківський авіаційний інститут», декан факультету радіоелектроніки, комп'ютерних систем та інфокомунікацій; Україна; email: o.odokiienko@khai.edu; ORCID: https://orcid.org/0000-0002-5227-1000

Ковальчук Даниїл Іванович – Національний аерокосмічний університет «Харківський авіаційний інститут; асистент кафедри аерокосмічних радіоелектронних систем»; Україна; email:d.i.kovaljchuk@khai.edu; ORCID: https://orcid.org/0009-0007-6847-6610

Щербина Ксенія Олександрівна – Національний аерокосмічний університет «Харківський авіаційний інститут», старший науковий співробітник кафедри аерокосмічних радіоелектронних систем; Україна; email: k.shcherbyna@khai.edu; ORCID: https://orcid.org/0009-0005-7870-3675

Сидоров Ярослав Дмитрович – Національний аерокосмічний університет «Харківський авіаційний інститут», інженер кафедри аерокосмічних радіоелектронних систем, Україна; email: y.d.sydorov@khai.edu; ORCID: https://orcid.org/0009-0002-4088-2127

УДК 004.8:681.5 **DOI:10.30837/rt.2025.3.222.12**

О.В. ЗУБКОВ, канд. техн. наук, Н.В. БОЙКО, Т.С. МАЧОНІС

ДОСЛІДЖЕННЯ РОЗПІЗНАВАННЯ ДРОНІВ ЗА ЇХ АКУСТИЧНИМ ВИПРОМІНЮВАННЯМ З ВИКОРИСТАННЯМ ПОВНОЗВ'ЯЗНИХ НЕЙРОННИХ МЕРЕЖ

Вступ

У сучасних умовах широкого розповсюдження безпілотних літальних апаратів (БПЛА), або дронів, питання їх своєчасного виявлення та розпізнавання набуває особливої актуальності. Дрони використовуються як у цивільних, так і у військових цілях, а в окремих випадках становлять загрозу безпеці критичної інфраструктури та громадської безпеки [1–3]. Традиційно для виявлення дронів застосовуються радіолокаційні та оптичні методи [4–6]. Проте ці підходи мають низку суттєвих обмежень, які знижують їхню ефективність в умовах реального середовища.

Радіолокаційні системи виявлення демонструють високу точність, проте малорозмірні дрони з малими ефективними поверхнями розсіювання часто залишаються поза межами їхньої чутливості, особливо у складних рельєфних або міських умовах [4, 5]. Крім того, дрони можуть використовувати матеріали та конструкції, що знижують радіолокаційну помітність. Оптичні методи, зокрема відеоспостереження та інфрачервона зйомка, залежать від погодних умов, освітлення та наявності прямої видимості, що також значно обмежує їх застосування у нічний час або за несприятливої погоди [6].

На цьому фоні акустичні методи виявлення та розпізнавання дронів виділяються рядом переваг. Акустичне випромінювання, що виникає під час роботи електродвигунів та обертання гвинтів, є унікальним для кожного типу дрона і може бути виявлене навіть при відсутності прямої видимості. Такі сигнали менш залежні від погодних умов і можуть застосовуватись у міських або лісистих умовах, де інші методи втрачають ефективність [7–9].

З огляду на складність та варіативність акустичних сигналів, особливо важливою є розробка ефективних методів їх автоматичного аналізу, які можуть ефективно виділяти та аналізувати загальні ознаки таких сигналів в умовах дії зовнішніх шумів [10]. В цьому контексті нейронні мережі, зокрема повнозв’язні (fully connected) моделі, демонструють високу здатність до розпізнавання складних шаблонів у спектральних представленнях звуку. Тому дослідження застосування повнозв’язних нейронних мереж для задач розпізнавання дронів за їх акустичним випромінюванням є актуальним напрямом сучасної науково-технічної діяльності.

Метою цього дослідження були: аналіз ефективності використання повнозв’язних нейронних мереж для виявлення дронів на різних відстанях та оцінка апаратної складності реалізації алгоритму розпізнавання на сучасних мікропроцесорних платформах.

Аналіз спектральних особливостей власного акустичного випромінювання дронів

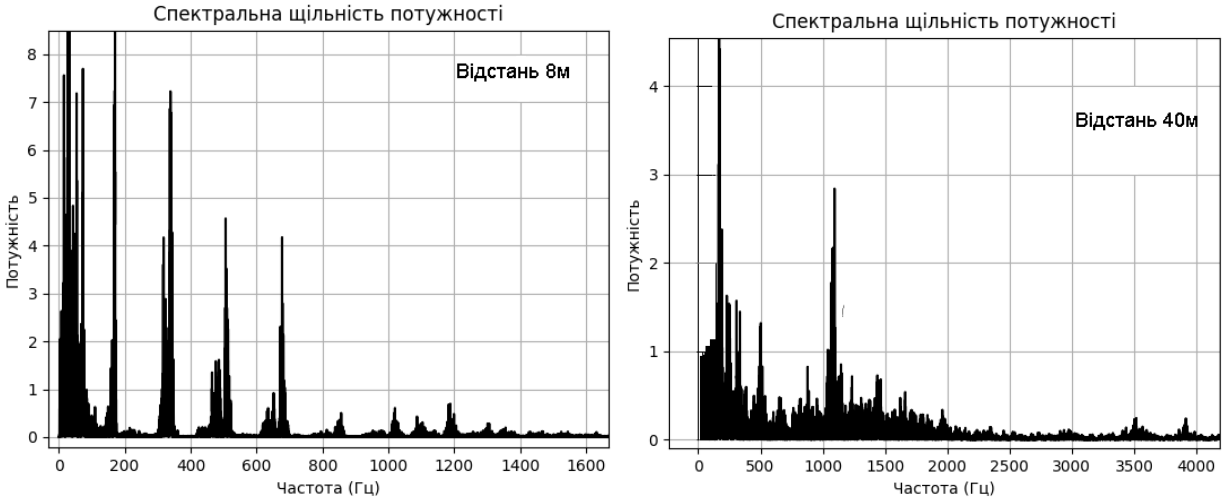
Існує багато моделей дронів за конструктивними особливостями. Але всі вони мають зазвичай від 1 до 8 гвинтів, які саме і створюють акустичне випромінювання.

При побудові методів ефективного виявлення дронів за їх власним акустичним випромінюванням аналізуються часові або спектральні особливості такого випромінювання. З точки зору конструкції дронів їх акустичні випромінювання представляють собою періодичні коливання, пов'язані з обертанням пропелерів та роботою двигунів. Наряду з періодичними складовими спостерігаються імпульсні складові, зумовлені турбулентними потоками повітря та взаємодією конструктивних елементів дрона з навколишнім середовищем [6, 11].

Для аналізу сигналів у часовій області, наприклад, застосовуються рекурентні нейронні мережі (RNN), які дозволяють ефективно розпізнавати акустичні сигнали, навіть в умовах

впливу шумів, але вимагають для практичної реалізації значних обчислювальних ресурсів та великих об'ємів пам'яті [12]. Такі вимоги пов'язані з великою кількістю часових відліків акустичних сигналів, що значно навантажує обчислювальне ядро.

У спектральній площині присутні чітко виражені основні частоти, що відповідають обертанню двигунів, а також їх вищі гармоніки. Акустичне випромінювання дронів у спектральній площині можна охарактезувати, як широкосмугове, що охоплює частоти від десятків герць до 8–10 кГц [6–8], хоча найчастіше цей діапазон обмежений частотами до 2–5 кГц. Приклади результатів спектрального аналізу записів власних акустичних випромінювань дронів DJI Phantom 3 та Syma X5SW, що мають по 4 гвинта, на відстанях 8 та 40 м наведено на рис. 1.



а *б*

Рис. 1. Приклади спектральної щільності потужності власного акустичного випромінювання дронів на відстанях 8 та 40 м

Аналіз прикладів на рис. 1 показує, що коли дрон висить на постійній висоті і гвинти обертаються не дуже швидко, то спектральна щільність потужності обмежується частотою 1,6 кГц. При польоті на великих швидкостях частота обертання двигунів проаналізованих дронів збільшується і максимальна частоти гармонік досягають 8 кГц. На невеличких відстанях (8 м) відношення корисний сигнал/зовнішні шуми достатньо високе і у спектрі можна побачити порядка 9–10 гармонік основної частоти 185 Гц. Із зростанням відстані до 40 м відношення сигнал/шум значно зменшується, тобто рівень шумових спектральних компонентів стає близьким до рівню гармонічних складових. Кількість гармонічних складових, що суттєво перевищують рівень шумів зменшується 6–7. Таке зменшення можна пояснити більшим затуханням високочастотних складових спектра у атмосфері у порівнянні з низькочастотними. Отже власні дослідження спектрального складу та особливостей спектрів акустичних випромінювань дронів співпадають з проаналізованими науковими роботами.

Перехід у спектральну площину дозволяє значно зменшити об'єм даних, що треба обробляти, що особливо актуально для подальшого використання розпізнавання цієї інформації з використанням нейронних мереж. При реалізації такого переходу найпоширенішими та найефективнішими є перетворення Фур'є та Вейвлет перетворення [13]. Вейвлет перетворення забезпечує ефективний аналіз у складних завадових обставинах, дуже ефективне при аналізі нестаціонарних сигналів, дозволяє робити одночасно аналіз у частотній та часовій площинах. Однак складність обирання базових функцій розкладання, велика кількість обчислень значно обмежують практичне використання. Тому для сигналів з періодичним характером більш якісніше використання перетворення Фур'є. Після виконання перетворення Фур'є кількість спектральних компонентів у діапазоні частот від 0 Гц до 5 кГц залишається великою для безпосередньої обробки. Тому на практиці існують методи вторинної обробки

спектрів, що дозволяють суттєво зменшити кількість значень у інформаційному векторі, що описує спектральні характеристики акустичного сигналу.

Найефективнішим методом формування вектора ознак за результатами обчислення спектральної щільності потужності є метод Mel-Frequency Cepstral Coefficients (MFCC) [14, 15]. В основу цього метода покладено модель чутливості вуха людини, що сприймає звуки нелінійно у частотній області. Обробка сигнала починається з розбиття вхідного сигнал *x(n)* на короткі сегменти довжиною *N* із перекриттям *M*, що дорівнює 25–50 % ширини вікна. Для аналізу мовних сигналів довжина вікна складає 20–40 мс, але з точки зору накопичення енергії корисного слабкого сигнала довжину вікна треба обирати значно більшою. З метою зменшення спектральних витіків виконується згладження одним із вікон *w(n)*: Ганна, Хемінга і т.і. Далі обчислюється швидке перетворення Фур'є (ШПФ) і спектр потужності *P(k)*:

$$P\ k\ = \left| \sum_{n=0}^{N-1} x\ n\ w\ n\ e^{-j2\pi kn/N} \right|^2 ,k = 0,1,\ldots,N-1 \tag{1}$$

Далі обчислюється шкала частот Mel, що відповідає частотній характеристиці вуха людини:

$$f_{mel} = 2595 \cdot log_{10}\left(1+\frac{f}{100}\right) \tag{2}$$

Пілся цього задається набір трикутних фільтрів, рівномірно розташованих у шкалі Mel. Вихід кожного фільтра M обчислюється як зважена сума потужності спектральних складових:

$$S_m = \sum_{k=f_{m-1}}^{f_{m+1}} P\ k\ H_m\ k \tag{3}$$

де $H_m(k)$ – коефіцієнти трикутного фільтра, а f_{m-1} і f_{m+1} – межі фільтра.

Далі виконується обчислення кепстральних коефіцієнтів. До отриманих значень S_m застосовується дискретне косинусне перетворення (DCT), що дозволяє зменшити кореляцію між коефіцієнтами:

$$C_n = \sum_{m=1}^{M} \log\ S_m\ cos\left[\frac{\pi n\ m-0.5}{M}\right], n = 1,2,\ldots,L, \tag{4}$$

де L – кількість коефіцієнтів MFCC.

За результатами досліджень різних вчених кількість коефіцієнтів, що потрібна для класифікації типу акустичного сигналу значно відрізняється. Так для аналізу мовних сигналів зазначається, що потрібно 8–14 коефіцієнтів, в інших випадках обирають до 20 коефіцієнтів. Однак для задач виявлення дронів дослідники не обґрунтували кількості коефіцієнтів MFCC, яка б була оптимальною з точки зору максимізації імовірності вірного виявлення та мінімізації похибкового прийняття інших звуків, як звуків дронів. Також для розпізнавання акустичних випромінювань дронів на фоні шумів важливим є накопичення енергії сигнала при прийомі цього сигнала з великих відстаней. Тому можна зробити припущення, що часові вікна, що обробляються, треба значно збільшувати, але таке збільшення повинно бути допустимим з точки зору часових затримок при обробці сигналів, а також не потребувати занадто великих вимог до апаратної платформи при практичній реалізації.

Розробка алгоритму досліджень та нейронної мережі

Відомі Інтернет платформи, такі як Kaggle, Roboflow надають для досліджень готові датасети із значеннями MFCC коефіцієнтів для різних моделей дронів та напрямів руху, що значно спрощує процес навчання нейронних мереж [16, 17]. Але суттєвий недолік таких дата сетів – це відсутність інформації о тривалостях часового вікна, вікна згладження, відстанях між мікрофоном та дроном, частоти дискретизації та параметрів мікрофонів. Все це не

дозволяє точно визначити вплив цих факторів на результат навчання та подальшого розпізнавання. Для вирішення виявлених вище та проаналізованих проблем було вирішено створити власну базу даних записів акустичних випромінювань дронів DJI Phantom 3 та Syma X5SW при польотах на різних відстанях в межах від 5 до 100 м. Для запису акустичного випромінювання цих дронів використовувались: конденсаторний мікрофон Superlux ECM-999 та зовнішня звукова карта U-Phoria UM2. Частота дискретизації була обрана 44100 Гц, а формат зберігання аудіоданих – wav, щоб уникнути втрат інформації. Далі була створена база даних обчислених коефіцієнтів MFCC, що відповідають кількості відліків ШПФ від 2048 до 16384, а в часовій області вікнам від 46 до 372 мс. Для кожного часового вікна обчислюється 32 MFCC коефіцієнта за формулами (1) – (4). На вхід нейронної мережі подається 8, 12, 16, 20, 24, 28 або 32 коефіцієнти, що забезпечує пошук їх оптимальної кількості при тестуванні нейронної мережі.

При навчанні нейронних мереж результат навчання описується сукупністю параметрів [6, 18]: TP – кількість вірно розпізнаних об’єктів, TN – кількість нерозпізнаних об’єктів; FP – кількість помилкових позитивних передбачень, FN – вірних нерозпізнавань. Тобто якісно навчена нейронна мережа повинна забезпечити не тільки високу ймовірність вірного розпізнавання, але і мінімально можливі ймовірності прийняття фонових звуків за акустичні випромінювання дрона. Для цього була підготована база аудіозаписів фонових звуків Ground (вулиць міста, природи), а також джерел звуків, які за часовими характеристиками мають близький характер, тобто теж містять періодичні складові: літак, автобус, автомобіль, гелікоптер, мотоцикл, поїзд, фура. Всього було підготовано 853 аудіофайла з тривалістю не менше 10 с. З цих файлів для навчання нейронної мережі було обрано 70 %, для валідації – 10 %, а для тестування навченої нейронної мережі – 20 %.

Для досліджень були спроектовані 3 архітектури нейронних мереж з використанням фреймворку Tensorflow [19]. Кожна з них має від 8 до 32 входів, що відповідає кількості коефіцієнтів MFCC, а також 9 виходів, що відповідають 9 класам об’єктів розпізнавання (дрон, Ground, літак, автобус, автомобіль, гелікоптер, мотоцикл, поїзд, фура). В першій архітектурі є 3 проміжних шари з кількістю нейронів в шарах 1024, 265, 64. В другій архітектурі є 4 проміжних шари з кількістю нейронів в шарах 1024, 512, 128, 32. Такі архітектури є простими з точки зору реалізації, але мають проблеми з процесом навчання. Тому в третю архітектуру були додані додаткові шари BatchNormalization, що нормалізує активації нейронів для прискорення навчання та стабілізації градієнтів, та Dropout шари, що випадково вимикає частину нейронів під час навчання для запобігання перенавчанню. В трьох проміжних шарах цієї архітектури міститься 512, 128 та 32 нейрона. Остання архітектура наведена на рис. 2.

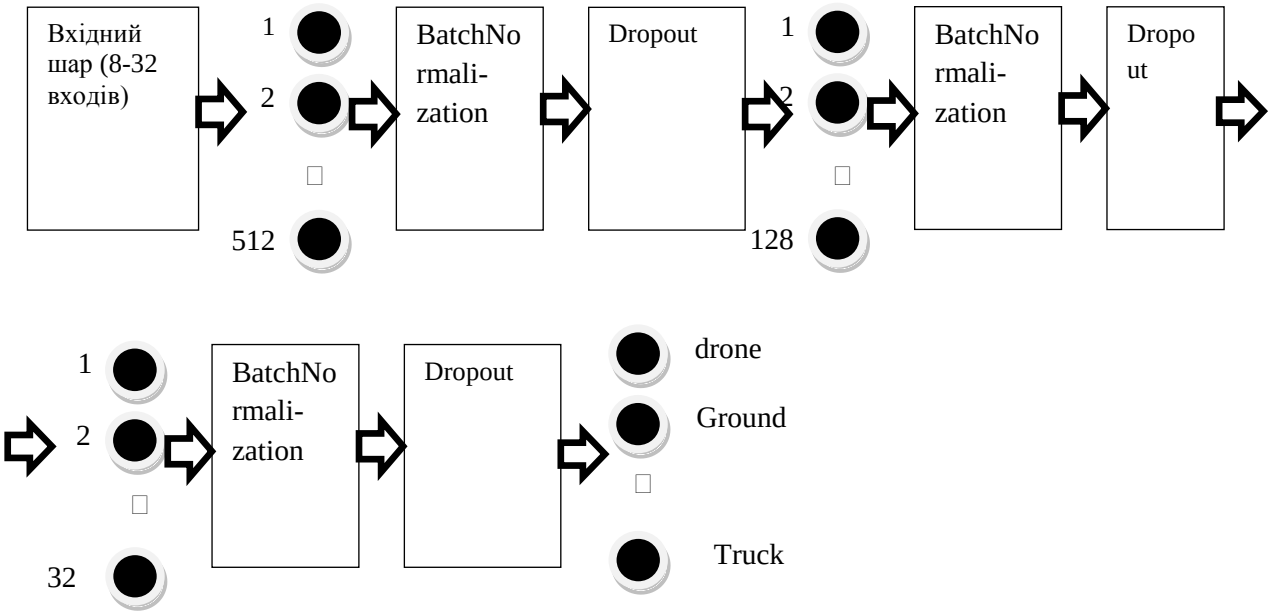


Рис. 2. Архітектура нейронної мережі із захистом від перенавчання

В усіх нейронах проміжних шарів використовується функція активації ReLU, що найбільш відповідає біологічному аналогу математичного нейрона та потребує невеликого процесорного навантаження, а в вихідному шарі використовується функція активації softmax. Навчання нейронних мереж виконувалось з використанням двох найбільш ефективних оптимізаторів Adam та RMSprop [20]. RMSprop працює за рахунок збереження експоненційного середнього квадратів градієнтів (дисперсії) та поділу градієнта на квадратний корінь цієї величини, що дозволяє уникнути проблем із занадто великими або малими оновленнями ваг. Він добре працює для рекурентних нейронних мереж і задач, де спостерігаються шумні або нестаціонарні градієнти. Adam, на відміну від RMSprop, поєднує два механізми: експоненційне згладжування середніх значень градієнтів (моменту) як у методу Momentum, і експоненційне згладжування квадратів градієнтів, як у RMSprop. Це дозволяє Adam швидко збігатися, добре працювати при нестабільних градієнтах і мінімізувати ручне налаштування гіперпараметрів. Завдяки поєднанню переваг двох підходів, Adam зазвичай забезпечує швидшу та стабільнішу збіжність порівняно з RMSprop, особливо в задачах з великими наборами даних або складними функціями втрат.

Навчання архітектур нейронних мереж та аналіз результатів навчання

Для навчання розроблених архітектур нейронних мереж використовувалась бібліотека tensorflow. Навчання проводилось впродовж 150 епох, але з функцією автоматичного припинення навчання, коли впродовж 8 епох немає зменшення втрат. Розмір одного батча 400 векторів MFCC. Навчання проводилось багаторазово для різної кількості коефіцієнтів MFCC, різних архітектур та оптимізаторів. За результатами кожного навчання було отримано матрицю помилок, приклад якої наведено на рис. 3.

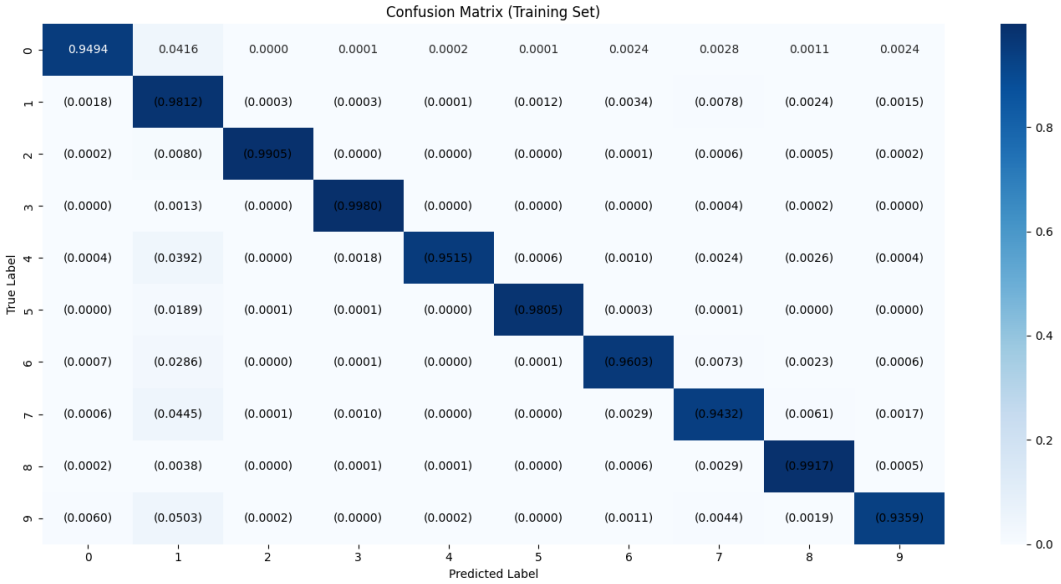


Рис. 3. Матриця помилок для архітектури №3, оптимізатора Adam, кількість часових відліків 8192 та 28 кількості коефіцієнтів MFCC

Як можна побачити, з матриці помилок навчання було виконано достатньо якісно, тобто ймовірність вірного розпізнавання знаходиться в межах від 93,5 до 99,8 %, ймовірність помилкового розпізнавання інших класів сигналів, як дронів до 0,6 %. Аналіз аналогічних матриць для різної кількості MFCC коефіцієнтів в діапазоні від 8 до 36 показує, що значення ймовірності вірного розпізнавання може суттєво змінюватись. Наприклад, при довжині вікна 8192 відліки та використанні оптимізатора Adam діапазон ймовірностей вірного розпізнавання знаходиться в межах від 0,81 до 0,987, тобто складає 17,7 %.

Після навчання проводилось тестування кожної навченої нейронної мережі на незалежній сукупності аудіозаписів. За результатами цього тестування було отримано декілька груп залежностей: 1) залежності ймовірності вірного розпізнавання від кількості MFCC коефіцієнтів; 2) залежності ймовірності вірного розпізнавання від довжини часового вікна та оптимізатора; 3) залежності ймовірності вірного розпізнавання від архітектури нейронної мережі. На рис. 4 наведено залежності ймовірності вірного розпізнавання від кількості MFCC коефіцієнтів.

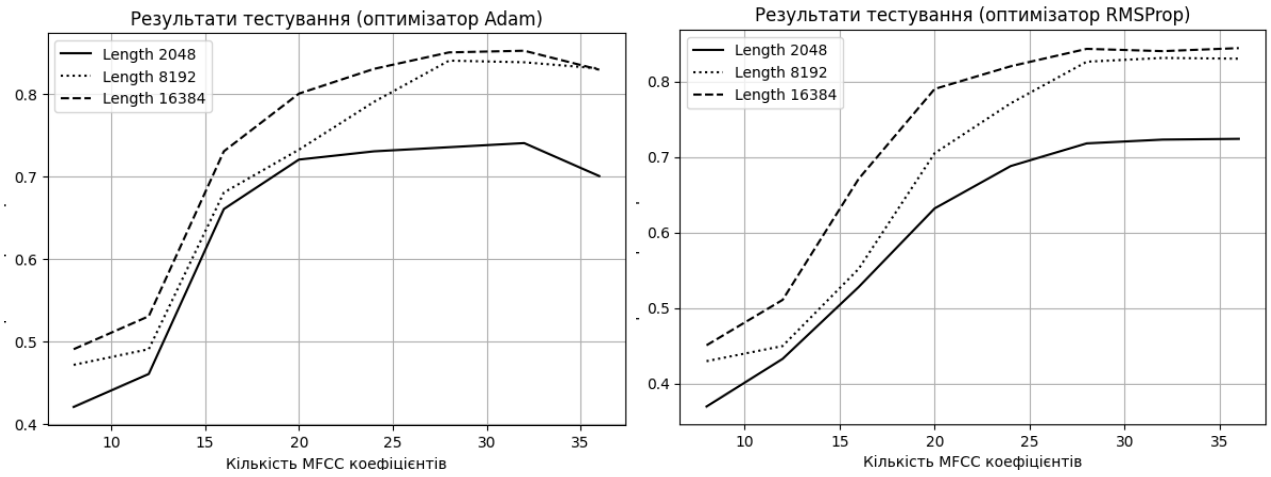


Рис. 4. Залежності ймовірності вірного розпізнавання
від кількості MFCC коефіцієнтів

Аналізуючи залежності, наведені на рис. 4, можна зробити висновки, що існує оптимальне значенні кількості коефіцієнтів MFCC для подальшої обробки нейронними повнозв'язними мережами, що забезпечує максимізацію ймовірності вірного розпізнавання. Для часових вікон з кількістю відліків 2048 оптимальним є використання 32 коефіцієнтів MFCC, із зростанням довжини вікна до 16384 відліків достатньо використовувати 28 коефіцієнтів MFCC, бо подальше збільшення їх числа не дозволяє збільшувати ймовірність вірного розпізнавання. Використання малого числа коефіцієнтів 8–12 поступається використанню оптимального числа коефіцієнтів до 38 %. Якщо порівнювати між собою залежності, наведені на рис. 4, *а*, *б*, то можна побачити, що оптимізатор Adam забезпечує на 4,3 % більшу ймовірність вірного розпізнавання при малій кількості коефіцієнтів MFCC (8–10) та на 0,9–1,7 % більшу ймовірність при оптимальних значеннях числа коефіцієнтів MFCC. Отже, оптимізатор Adam можна рекомендувати до практичного використання. Також порівнюючи використання різних довжин часових вікон можна зробити висновок, що малі часові вікна довжиною 2048 відліків значно програють більш довгим (8192 та 16384 відліки) за ймовірністю вірного розпізнавання. Цей програш становить 10,8 % використанню вікна довжиною 8192 відліки. В той самий час збільшення довжини вікна з 8192 до 16384 відліків покращує ймовірність вірного розпізнавання до 1 %. Тобто, якщо кінцевою метою системи розпізнавання є максимізація ймовірності вірного розпізнавання, то слід використовувати часові вікна довжиною 16384 відліки, а при реалізації компромісу між максимізацією ймовірності вірного розпізнавання та апаратними вимогами слід використовувати часові вікна 8192 відліки.

На рис. 5 показані залежності сумарних ймовірностей похибкового розпізнавання різноманітних акустичних джерел як випромінювання дрона. Аналіз цих залежностей показує, що оптимальною кількістю коефіцієнтів MFCC є 28–32 коефіцієнта, коли забезпечується мінімум ймовірностей похибкового розпізнавання. Оптимальною довжиною часової виборки є 8192 відліки. Використання більших виборок навіть збільшує ймовірностей похибкового розпізнавання. Найкращі результати знов забезпечує використання оптимізатора Adam.

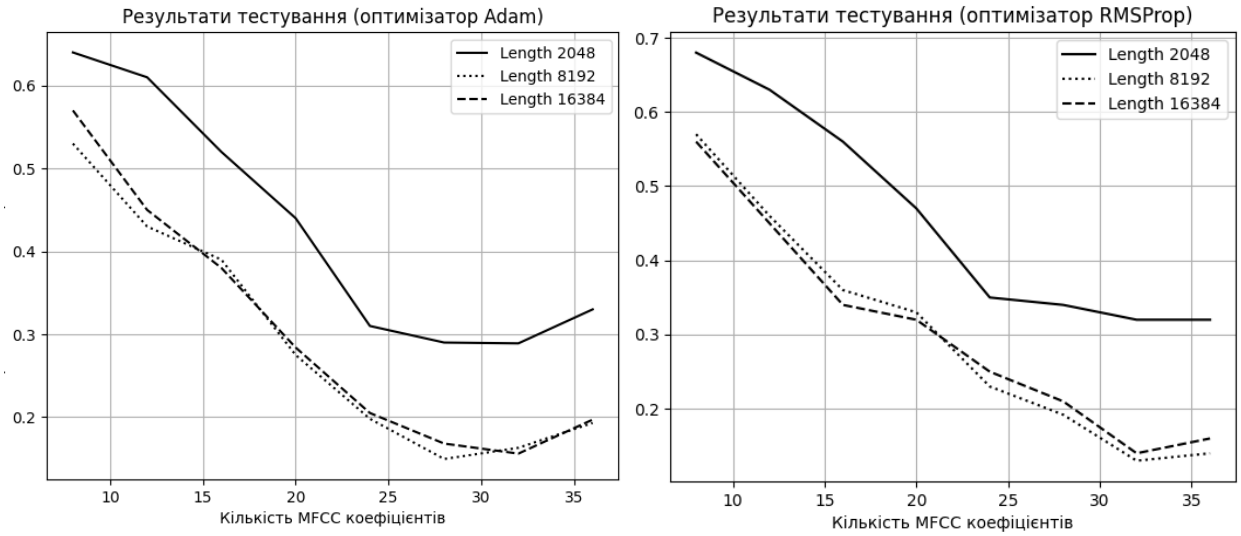


Рис. 5. Залежності сумарних ймовірностей похибкового розпізнавання
8 акустичних джерел як випромінювання дрона

Аналогічні дослідження були проведені для першої та другої архітектур нейронних мереж. Результати цих досліджень для часового вікна з кількістю відліків 16384 та оптимізатора Adam наведені у табл. 1.

Таблиця 1

Номер архітектури	Ймовірність вірного розпізнавання TP		Ймовірность похибкового розпізнавання FP	
	Adam	RMSprop	Adam	RMSprop
1	0,734	0,732	0,1773	0,1884
2	0,772	0,758	0,1628	0,1716
3	0,882	0,879	0,1495	0,1523

Аналіз даних табл. 1 показує, що архітектури нейронних мереж без використання шарів, що запобігають перенавчанню та покращують навчання програють архітектурі з такими шарами до 14,8–11 %.

Для архітектури нейронної мережі №3, оптимізатора Adam та кількості відліків у часовому вікні 16384 було проведено тестування ефективності роботи нейронної мережі в діапазоні польотів дрону від 5 до 100 м з GPS трекінгом польоту дронів та подальшої сумісної обробки трекінгу та акустичного випромінювання. Приклад такого розпізнавання показаний на рис. 6.

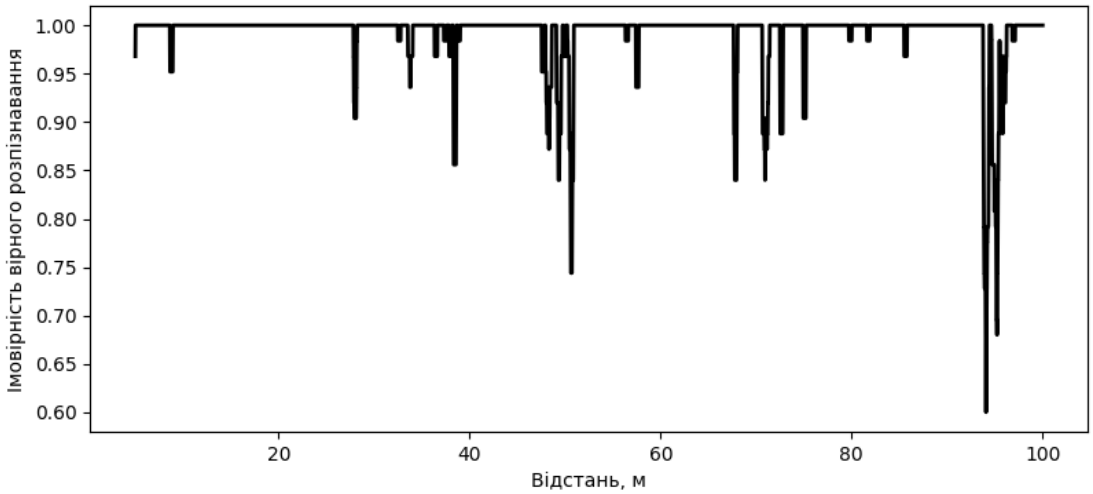


Рис. 6. Приклад розпізнавання дрону на відстанях від 5 до 100 м

Як можна побачити з даних рис. 6 та аналогічних обробок польотів дронів, якісне розпізнавання (практично без помилок) можливе на відстанях до 50 м. Із зростанням дальності зростає кількість помилок і на відстанях порядка 100 м вона складає до 23 %, що потребує використання додаткових вторинних методів обробки, до яких можна віднести методи трекінгу та прийняття рішення по сукупності результатів обробки впродовж інтервалів часу, що відповідають тривалості 3–5 часових вікон.

Крім аналізу якості роботи розроблених нейронних мереж було проаналізовано кількість параметрів нейронної мережі №3 та кількість потрібної оперативної пам'яті на реалізацію роботи навченої нейронної мережі на спеціалізованих процесорах для вбудування нейронних мереж, таких як STM32MP157. При використанні 16-розрядного вбудованого в мікропроцесор аналого-цифрового перетворювача, канала DMA з подвійним буфером для зберігання одразу 32768 відліків акустичного сигналу, а також обробки 32 коефіцієнтів потрібно не менше 762 Кбайт оперативної пам'яті. Тобто використання повнозв'язних нейронних мереж дозволяє значно зменшити вимоги до апаратної платформи при реалізації розпізнавання акустичних випромінювань дронів і зменшити вартість подібних пристроїв.

Висновки

1. Проведено аналіз сучасних методів розпізнавання акустичних випромінювань дронів в часовій області та частотній. Обґрунтовано ефективність використання MFCC для подальшого розпізнавання векторів цих коефіцієнтів повнозв'язними нейронними мережами.

2. Розроблено 3 архітектури нейронних мереж, виконано акустичні записи польотів дронів на різних відстанях, а також же 8 різних джерел фонових звуків та звуків з періодичними часовими складовими, проведено навчання розроблених архітектур за допомогою створеного датасету акустичних записів, а також тестування навчених архітектур на незалежній частині датасету.

3. Доказано, що існує оптимальне значення кількості коефіцієнтів MFCC для розпізнавання акустичних випромінювань дронів. Воно дорівнює 28–32, при якому забезпечується максимізація ймовірності вірного розпізнавання до 0,882 та мінімізація похибкового розпізнавання до 0,1495. Оптимальною довжиною часового вікна при розрахунку коефіцієнтів MFCC є 372 мс. Використання оптимізатора Adam дає виграш 0,9–1,7 % в ймовірності вірного розпізнавання у порівнянні з оптимізатором RMSProp. Використання шарів BatchNormalization та Dropout при створенні архітектури нейронної мережі дозволяє збільшити ймовірність вірного розпізнавання на 11–14,8 %.

4. На відстанях до 50 м реалізується розпізнавання акустичних випромінювань дронів з високою ймовірністю вірного розпізнавання. На відстанях порядка 100 м похибки розпізнавання складають порядка 23 %. Для зменшення впливу цих похибок розпізнавання дронів, а також похибкового розпізнавання інших джерел звуку, як дронів, можливо використання методів вторинної обробки, що є предметом подальших наукових досліджень

5. Використання повнозв'язних нейронних мереж при розпізнаванні акустичних випромінювань дронів можливо реалізувати на спеціалізованих процесорних платформах типу STM32MP157, що значно зменшую собівартість таких рішень завдяки невеликим вимогам до оперативної пам'яті порядка 1 Мбайт.

Список літератури:

1. Agapiou A. Drones in Construction: A Comparative International Review of the Legal and Regulatory Landscape // Management Procurement and Law. 2020. Vol. 174, No.3. P.1–8. DOI:10.1680/jmapl.19.00041

2. Pyrgies J. The UAVs threat to airport security: risk analysis and mitigation // Journal of Airline and Airport Management. 2019. Terrassa. Vol. 9, Iss.2. P.63–96.

3. Mowafaq SA, Muhyeeddin A, Al-Batah MS. AI in the Sky: Developing Real-Time UAV Recognition Systems to Enhance Military Security // Data and Metadata. Nazarre. 2024. Vol. 3. P.1–19. https://doi.org/10.56294/dm2024.417

4. Aouladhadj D., Kpre E., Deniau V., Kharchouf A., Gransart C., Gaquière C. Drone Detection andTracking Using RF IdentificationSignals // Sensors. Basel. 2023. Vol. 23. P.1–24. https://doi.org/10.3390/s23177650

5. Yousaf J., Zia H., Alhalabi M., Yaghi M., Basmaji T., Shehhi E.A., Gad A., Alkhedher M., Ghazal M. Drone and Controller Detection and Localization: Trends and Challenges// Appl. Sci. Basel. 2022. Vol.12, Iss.24, P.1–22. https://doi.org/10.3390/app122412612

6. Zubkov O.V., Sheiko S.O., Oleynikov V.M., Kartashov V.M., Babkin S.I. Investigation Of The Yolov5 Algorithm Efficiency For Drone Recognization // Telecommunications and Radio Engineering. Danbury. 2024. Vol.83, Iss.1. P. 65–79. DOI: 10.1615/TelecomRadEng.2023048987

7. Oleynikov V., Zubkov O., Kartashov V., Korytsev I., Babkin S., Sheiko S. Investigation of the efficiency of detection and recognition of small-sized unmanned aerial vehicles by their acoustic radiation. Radiotekhnika. 2018. Vol.4, No.195. P.209–217. https://doi.org/10.30837/rt.2018.4.195.21

8. Paszkowski W., Gola A., Świć A. Acoustic-Based Drone Detection Using Neural Networks – A Comprehensive Analysis// Advances in Science and Technology Research Journal. Lublin. 2024. Vol.18, Iss.1. P.36–47. https://doi.org/10.12913/22998624/175863

9. Tejera-Berengue D., Zhu-Zhou F., Utrilla-Manso M., Gil-Pita R., Rosa-Zurera M. Analysisof Distance and EnvironmentalImpact on UAV Acoustic Detection // Electronics. Basel. 2024. Vol. 13. https://doi.org/10.3390/electronics13030643

10.Othman E., Cibilić I., Poslončec-Petrić V., Saadallah D. Investigating Noise Mapping in Cities to Associate Noise Levels with Sources of Noise Using Crowdsourcing Applications// Urban Sci. 2024. Vol.8, Iss.1. https://doi.org/10.3390/urbansci8010013

11.Dombrovschi M., Deaconu M., Cristea L., Frigioescu T.F., Cican G., Badea ., Totu G. Acoustic Analysis of a Hybrid Propulsion System for Drone Applications// Acoustics. Basel. 2024. Vol. 6, Iss.3. P. 698–712. https://doi.org/10.3390/acoustics6030038

12.Islam M., Haque M., Islam S., Mia Z.A., Rahman M. DCNN-LSTM Based Audio ClassificationCombining Multiple Feature Engineeringand Data Augmentation Techniques Intelligent // Computing & Optimization Springer. Cham. 2021. Vol. 371. P.227–236. DOI:10.1007/978-3-030-93247-3_2 3

13.Wei N., Gu J.X., Gu F., Chen Z., Li G., Wang T., Ball A.D. An Investigation into the Acoustic Emissions of Internal Combustion Engines with Modelling and Wavelet Package Analysis for Monitoring Lubrication Conditions // Energies. Basel. 2019. Vol.12, Iss.4. P. 1–14. https://doi.org/10.3390/en12040640

14. Zrar K.A., Abdulbasit K.A. Mel Frequency Cepstral Coefficient and its Applications: A Review // IEEE Access. Piscataway. 2022. Vol.10. P.122136–122158. DOI:10.1109/ACCESS.2022.3223444

15. Sithara A., Abraham T., Mathew D. Study of MFCC and IHC Feature Extraction Methods With Probabilistic Acoustic Models for Speaker Biometric Applications // Procedia Computer Science. Amsterdam. 2018. Vol. 143. P. 267–276. DOI:10.1016/j.procs.2018.10.395

16. Fazal M.A., Baig M.A., Manj W.A., Faraz Z., Mallah G.A. Implementation of Deep Learning for Acoustic Classification // Journal of Xidian University. Christchurch. 2023. Vol. 17, Iss. 8. P.1653–1673. DOI:10.37896/jxu17.8/138

17. Pham L., Ngo D., Salovic D.c, Jalali A., Schindler A., Nguyen P. X., Tran K., Vu H. Lightweight deep neural networks for acoustic scene classification and an effective visualization for presenting sound scene contexts // Applied Acoustics. Basel. 2023. Vol. 211. P.723–731. DOI:10.1016/j.apacoust.2023.109489

18.Jung H.K., Choi G.S. Improved YOLOv5: Efficient Object Detection Using Drone Images under Various Conditions // Appl. Sci. Basel. 2022. Vol. 12. P.1–16. DOI: 10.3390/app12147255

19.Pang B., Nijkamp E., Nian Y.W. Deep Learning With TensorFlow: A Review // Journal of Educational and Behavioral Statistics. 2019. Vol. 45, Iss. 2. P.415–421. https://doi.org/10.3102/107699861987276

20. Soujanya B., Sitamahalakshmi T. Optimization with ADAM and RMSprop in Convolution neural Network (CNN): A Case study for Telugu Handwritten Characters // International Journal of Emerging Trends in Engineering Research. Pawan. 2020. Vol. 8. No. 9. P.5116–5121.

Надійшла до редколегії 12.06.2025

Відомості про авторів:

Зубков Олег Вікторович – канд. техн. наук, професор, Харківський національний університет радіоелектроніки, завідуючий кафедрою мікропроцесорних технологій і систем, Україна; e-mail: oleh.zubkov@nure.ua; ORCID: https://orcid.org/0000-0002-8528-6540

Бойко Наталія Вікторівна – Харківський національний університет радіоелектроніки, асистент кафедри мікропроцесорних технологій і систем, Україна; e-mail: natalia.boiko@nure.ua

Мачоніс Тадас Сігітасович – Харківський національний університет радіоелектроніки, аспірант кафедри мікропроцесорних технологій і систем, Україна; e-mail: tadas.machonis@nure.ua

УДК 621.396.96 **DOI:10.30837/rt.2025.3.222.13**

В.М. ОЛЕЙНІКОВ, канд. техн. наук

ОСОБЛИВОСТІ ВИЯВЛЕННЯ МАЛОРОЗМІРНИХ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ МЕТОДОМ РАДІОАКУСТИЧНОЇ ЛОКАЦІЇ

Вступ

Малорозмірні безпілотні літальні апарати (МБПЛА) завдяки своїй багатофункціональності знаходять широке застосування в різних сферах людської діяльності. Вони ефективно використовуються у цивільному секторі, а також стали невід'ємною складовою збройних конфліктів останнього часу [1].

Сучасні МБПЛА характеризуються високою маневреністю, низькою помітністю та здатністю функціонувати в складних умовах експлуатації, включаючи несприятливі зовнішні впливи. Їх виявлення становить надзвичайно складну задачу, що обумовлена компактними габаритами, зниженою радіолокаційною, оптичною, інфрачервоною та акустичною помітністю, а також високою швидкістю і маневреністю. Особливу складність становлять МБПЛА з оптоволоконними каналами зв'язку та МБПЛА з автономним управлінням на базі штучного інтелекту, у яких відсутні активні випромінювання бортового обладнання [2].

Для виявлення, ідентифікації та визначення координат МБПЛА використовуються радіолокаційні, оптичні, інфрачервоні та акустичні методи [3–7]. Однак у складних завадових умовах або на значних відстанях засоби, створені на основі зазначених методів, часто не забезпечують належної оперативності та достовірності виявлення МБПЛА, що свідчить про обмеженість традиційних технологій і потребу в розробці нових підходів до вирішення цієї проблеми.

Одним із перспективних напрямів підвищення ефективності виявлення малопомітних МБПЛА є використання явища розсіювання електромагнітних хвиль на акустичних збуреннях, які створюються цими апаратами в навколишньому повітряному середовищі [8–11].

Метою роботи є розгляд структурних та просторових особливостей акустичного випромінювання МБПЛА, а також визначення методом імітаційного моделювання умов формування розсіяного сигналу, що виникає внаслідок дифракції електромагнітних хвиль на неоднорідностях діелектричної проникності атмосфери, спричинених поширенням акустичних хвиль від МБПЛА.

Радіоакустичний метод зондування атмосфери

Можливість відбиття електромагнітних хвиль (ЕМХ) від періодичної структури, утвореної звуковою хвилею в атмосфері була передбачена теоретично, а потім це явище було продемонстровано в експерименті з дифракції світла на ультразвукових хвилях у рідинах та кристалах [12].

В атмосфері відбиття радіохвиль може походити від об'ємних неоднорідностей діелектричної проникності середовища розміром $l=\lambda/2$. Значне збільшення рівня відбитого сигналу відбувається при відбитті радіохвиль від дифракційної решітки, утвореної періодичними неоднорідностями середовища, яке створюють штучно. При радіоакустичному методі зондування (РАЗ) в атмосферу випромінюються короткі звукові імпульси, створюють періодичну зміну діелектричної проникності газів атмосфери. У разі виконання умови Брегга забезпечується когерентне складання розсіяних ЕМХ. Умова Брегга [13, 14] полягає у виконанні наступного співвідношення:

$$\frac{\lambda_e}{\lambda_a} = 2\sin\frac{\theta}{2}, \tag{1}$$

де λ_e – довжина ЕМХ; λ_a – довжина звукової хвилі; θ – кут розсіювання ЕМХ.

На основі теорії поширення хвиль у шаристонеоднорідних середовищах, для випадку радіоакустичного зондування атмосфери [15], отримано розрахункове співвідношення для коефіцієнта відбиття ЕМХ від звукової посилки:

$$K_{\text{в}}^2 = \frac{P_{\text{пр}}}{P_{\text{пер}}} = B\frac{N}{R^2}^2 \left[\frac{\sin N\pi(\frac{2\lambda_{\text{a}}}{\lambda_{\text{e}}}-1)}{\sin N\pi(\frac{2\lambda_{\text{a}}}{\lambda_{\text{e}}}-1)}\right]^2 exp\left[-\int_0^{R_{max}} \alpha(R)dR\right], \quad (2)$$

де N – число довжин хвиль в акустичному пакеті; $P_{пер}$ – потужність випромінюваного передавачем радіосигналу; $P_{пер}$ – потужність радіосигналу, що приймається; R – висота зондування; B – коефіцієнт, що враховує втрати в тракті прийому та обробки; α – коефіцієнт поглинання звуку атмосфері.

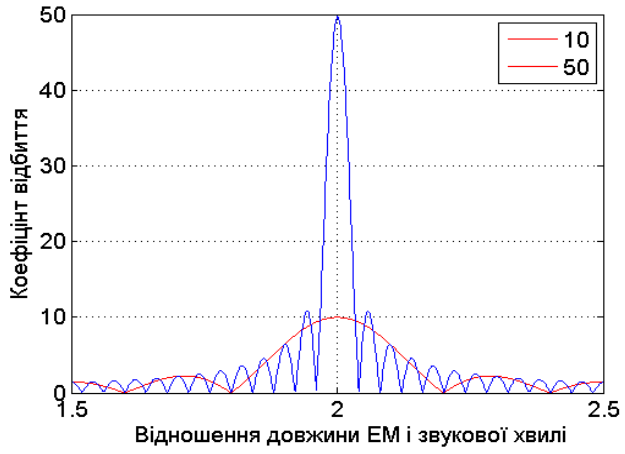


Рис. 1. Залежність коефіцієнта відбиття від співвідношення довжин звукових та ЕМХ при довжині акустичного пакета *N=10* і *N=50*

Під час зміни частоти акустичних коливань потужність розсіяного електромагнітного випромінювання, а отже й коефіцієнт відбиття електромагнітних хвиль від звукового імпульсу змінюються відповідно до функції $(sin\, x\, /x)^2$.

Залежність коефіцієнта відбиття має пелюстковий характер і періодичну структуру. На рис. 1 наведено залежність коефіцієнта відбиття від співвідношення довжин звукових та ЕМХ при довжині акустичного пакета *N=10* і *N=50*. Відбивна здатність звукової посилки пропорційна квадрату довжини акустичного пакета. Вважається, що оптимальна довжина акустичного пакета лежить у межах 35–50.

Особливості акустичного поля малорозмірного безпілотного літального апарата

ЕМ поле розсіяне на неоднорідностях атмосфери, обумовлене акустичним випромінюванням (АВ) МБПЛА, цілком визначається характеристиками його акустичного поля. АВ МБПЛА має ряд характерних особливостей, які необхідно враховувати при розробці системи радіоакустичної локації. Спектр АВ гвинтомоторної групи МБПЛА має гармонійні складові частоти обертання ротора та гармоніки лопатевої частоти [16, 17]. При обертанні повітряного гвинта виникають коливання тиску повітря, що відбуваються за рахунок витіснення повітря, об'єм якого дорівнює об'єму лопаті повітряного гвинта. Це призводить до появи шуму витіснення. Амплітуди гармонік АВ МБПЛА зменшуються зі збільшенням частоти. Дискретні складові спектру АВ, пов'язані з шумом обертання та взаємодії, як правило, мають на 15–20 дБ вищі рівні ніж широкосмуговий шум обтікання лопаті, кількість гармонік лопатевої частоти до декількох десятків. Частота проходження лопатей сучасних МБПЛА в залежності від режиму роботи і моделі знаходиться у межах від 75 до 1000 Гц.

Для МБПЛА мультироторного типу спектральні лінії гармонік АВ окремих гвинтів згруповані в широкі багатопикові делянки спектральної щільності. При збільшенні кількості гвинтів до шести і більше спектр мультироторного сигналу МБПЛА стає шумоподібним.

Формування акустичного поля залежить від характеристик спрямованості випромінювання гвінтомоторної групи МБПЛА. Характеристика спрямованості АВ МБПЛА визначає розподіл випромінюваної акустичної енергії у просторі. На рис. 2 представлено характеристики спрямованості окремих гармонічних складових АВ гвинтомоторної групи квадрокоптера DJI Phantom 3, отримані в натурному експерименті.

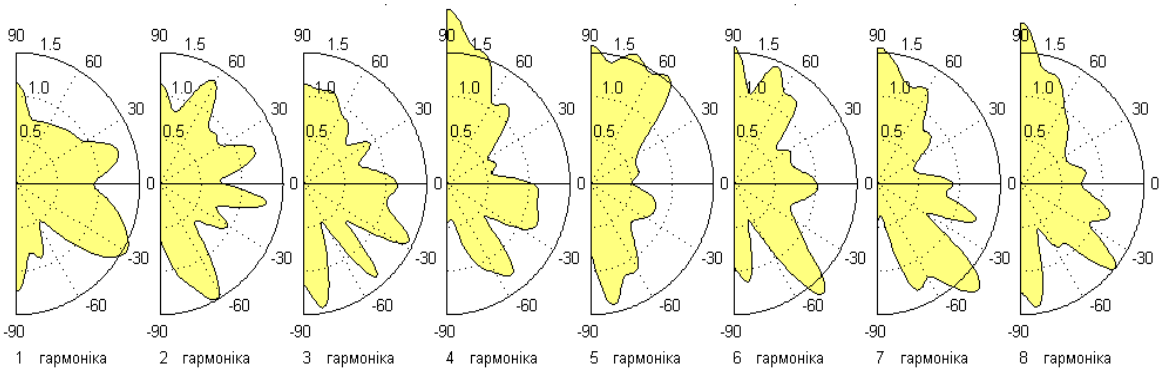


Рис. 2. Переріз 3D характеристики спрямованості окремих гармонійних складових АВ гвинтомоторної групи квадрокоптера DJI Phantom 3

АВ гвинтомоторної групи має складну залежність просторової спрямованості. З підвищенням номера гармоніки АВ спостерігається ускладнення форми характеристики спрямованості – вона стає більш порізаною, з великою глибиною провалів, ширина пелюстків зменшується, відбувається зміна напрямку домінуючої пелюстки АВ. При маневруванні МБПЛА змінюється орієнтація щодо точки спостереження та спрямованість АВ окремих гармонік, що призводить до зміни інтенсивності акустичного поля у точці прийому і форми його спектру.

У процесі польоту МБПЛА положення спектральних ліній АВ на частотній осі є нестабільними та постійно змінюються. Це зумовлено тим, що польотний контролер постійно регулює швидкість обертання кожного двигуна з метою балансування апарата і підтримання стабільного польоту. При цьому частота обертання кожного двигуна змінюється незалежно. Процес регулювання обертів двигунів супроводжується модифікацією спектру АВ у часі, відносна зміна частоти основного тону АВ протягом одиниць секунд може становить десятки відсотків.

Умови отримання розсіяного сигналу при радіоакустичній локації АВ МБПЛА

Порівняймо особливості реалізації радіоакустичного зондування (РАЗ) атмосфери та радіоакустичної локації АВ МБПЛА. Обидві технології ґрунтуються на явищі дифракції електромагнітних хвиль на неоднорідностях діелектричної проникності атмосфери, спричинених поширенням акустичних хвиль.

На рис. 3 наведено геометричну схему просторового розподілу неоднорідностей діелектричної проникності повітря, зумовлених впливом акустичних хвиль: для системи РАЗ – рис. 3, *а* та для умов радіоакустичної локації МБПЛА – рис. 3, *б, в*. Також на рисунках показано орієнтацію діаграми спрямованості антен РЛС.

У моностатичному варіанті реалізації системи РАЗ хвильові поверхні окремих періодів АВ майже повністю збігаються з хвильовими поверхнями електромагнітного випромінювання. Це сприяє фокусуванню електромагнітного сигналу та синфазному додаванню електромагнітних коливань, відбитих від окремих хвильових поверхонь акустичного пакета.

Коли антена РЛС орієнтована на джерело АВ, яким э МБПЛА, спостерігається розфокусування ЕМ-хвилі на опуклих і увігнутих сферичних поверхнях акустичної хвилі. Оскільки радіуси кривизни акустичних і електромагнітних хвиль не збігаються, синфазне складання розсіяних ЕМ-хвиль можливе лише для обмежених ділянок хвильової поверхні в межах діаграми спрямованості антени РЛС, вісь якої орієнтована на МБПЛА. У такому випадку в межах діаграми спрямованості можна виокремити дві зони хвильових поверхонь (рис. 3, *б*): перед МБПЛА – у напрямку до РЛС, і позаду МБПЛА – у напрямку від РЛС. Взаємодія електромагнітних полів, відбитих від акустичних хвиль, що рухаються у протилежних напрямках, призводить до виникнення сигналу, параметри огинаючої якого залежать від швидкості звуку та частоти АВ. Глибина модуляції сигналу на вході приймача РЛС, що виникає внаслі-

док інтерференції, визначається співвідношенням потужностей сигналів, сформованих ділянками, розташованими до і після МБПЛА в межах діаграми спрямованості антени.

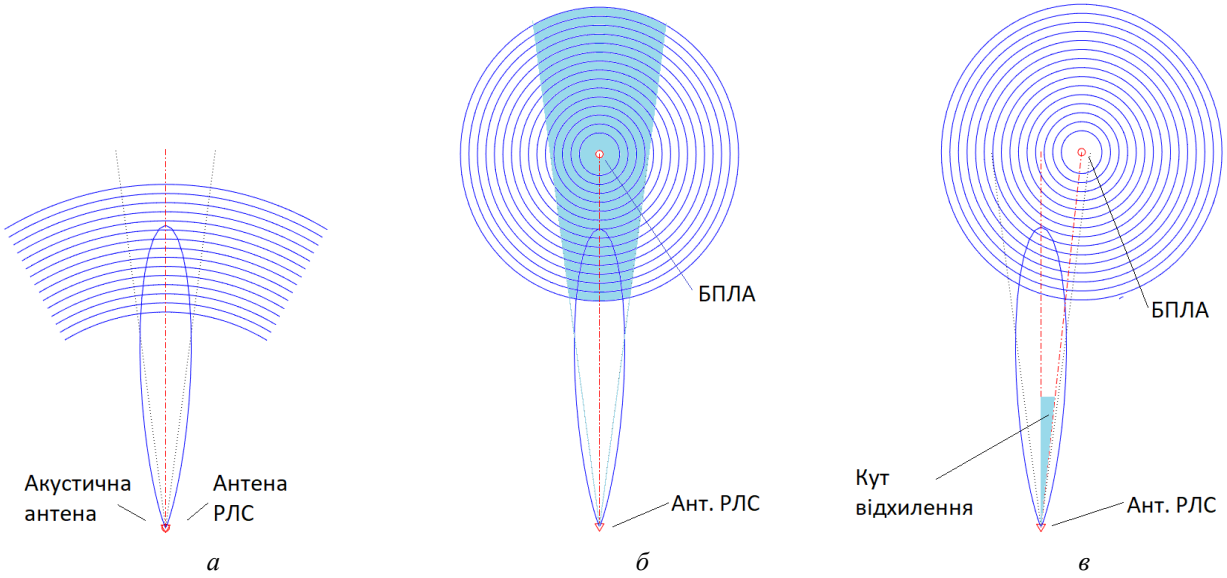


Рис. 3. Орієнтація антен РЛС і геометрична схема просторової структури неоднорідностей діелектричної проникності повітря, що виникає під впливом акустичних хвиль для умов системи РАЗ (*а*) та умов радіоакустичної локації МБПЛА (*б*, *в*)

Для виявлення сигналів, розсіяних на акустичному випромінюванні МБПЛА, важливу роль відіграють його енергетичні характеристики та параметри спектральних складових АВ. На рис. 4 представлено графік залежності частки середньої потужності АВ квадрокоптера DJI Phantom 3 [17] від ширини смуги використаної ділянки спектра сигналу, частота першої лопатевої гармоніки становить f_a=*170 Гц*. Домінуючими за внеском в загальний енергетичний баланс є три перші гармоніки – загалом до 70 % (28, 30 і 12 % відповідно). Частота електромагнітних коливань РЛС, що відповідає умові Брегга для частоти першої лопатевої гармоніки, дорівнює f_e=*75 МГц*. Спрямована антена РЛС метрового діапазону хвиль має досить великі габаритні розміри. Для практичного застосування доцільніше використовувати РЛС з вищою робочою частотою, що дозволяє зменшити габаритні розміри антен. Наприклад, на частоті 1,2 ГГц умова Брегга виконується для 16-ї лопатевої гармоніки (f_a=*2,72 кГц*) АВ квадрокоптера DJI Phantom 3. Частка акустичної потужності цієї гармоніки в загальному спектрі випромінювання МБПЛА є незначною і становить близько 1 %. Відповідно, рівень відбитих електромагнітних хвиль від таких неоднорідностей буде низьким.

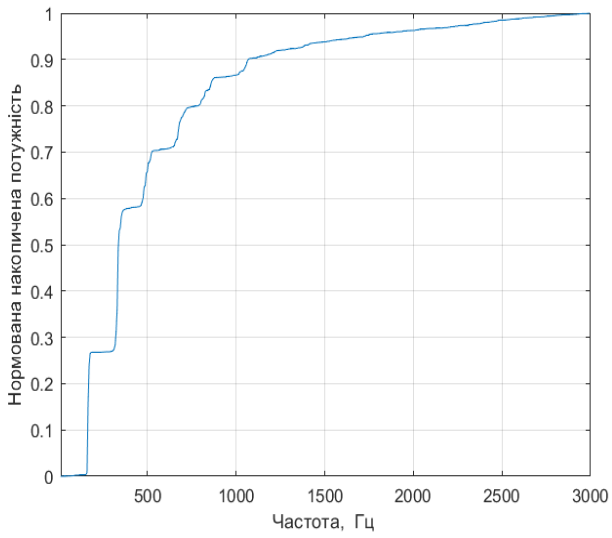


Рис. 4. Частка повної середньої потужності АВ квадрокоптера DJI Phantom 3 в залежності від смуги спектра сигналу, що використовується

Виконання умови Брегга можливе і тоді, коли період неоднорідності, утворений акустичною хвилею, дорівнює кратному числу напівхвиль електромагнітного коливання:

$$\lambda_a = m \cdot \lambda_e / 2 \,, \tag{3}$$

де m – порядок дифракції.

При цьому необхідно, щоб вектори акустичної та електромагнітної хвиль збігалися або мали протилежний напрямок. В такому випадку на частоті зондуючого сигналу можливе відбиття, зумовлене дифракцією вищого порядку від перших, найпотужніших лопатевих гармонік АВ МБПЛА (для робочої частоти РЛС 1,2 ГГц частота другої лопатевої гармоніки АВ квадрокоптера DJI Phantom 3 становить $f_a = 340$ *Гц,* порядок дифракції $m=8$).

При прийомі відбитого від періодичної акустичної неоднорідності спектральна щільність потужності прийнятого сигналу $S_v(f)$ являє собою згортку спектра сигналу АВ МБПЛА $S(f)$ та залежності коефіцієнта відбиття $K_в(f)$:

$$S_v(f) = S(f) \cdot K_в(f). \tag{4}$$

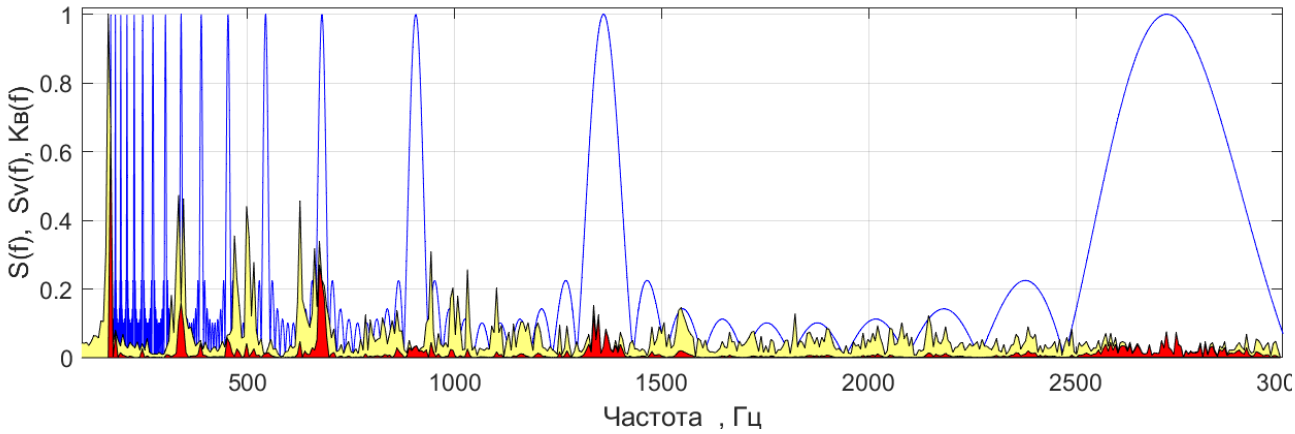


Рис. 5. Спектральна щільність потужності прийнятого сигналу $S_v(f)$ – червоний колір лінії та заливки, спектр сигналу АВ МБПЛА $S(f)$ – жовтий колір лінії та заливки, частотна характеристика коефіцієнта відбиття $K_в(f)$ – синій колір лінії

Як видно з рис. 5, у разі радіоакустичної локації з використанням дифракції вищого порядку резонансне відбиття спостерігається лише на окремих ділянках спектра АВ МБПЛА. Максимальне значення спектральної щільності потужності прийнятого сигналу $S_v(f)$ досягається за умови збігу частот домінуючих спектральних компонентів АВ МБПЛА з частотами пелюсток залежності коефіцієнта відбиття $K_в(f)$, яка має періодичну структуру.

Для забезпечення умов резонансного відбиття від спектральних складових звукового сигналу необхідно змінювати частоту зондувального електромагнітного сигналу, тобто реалізувати адаптацію локаційної системи до зміни умови Брегга. Це дає змогу аналізувати повний частотний діапазон, у якому акустичні хвилі, які формуються сучасними МБПЛА, ефективно відбивають електромагнітні хвилі в резонансному режимі розсіювання [13, 14].

Результати моделювання

У середовищі MATLAB розроблено імітаційну модель процесу формування сигналу при відбитті електромагнітної хвилі від неоднорідностей, зумовлених акустичним випромінюванням МБПЛА. У моделі задаються координати, діаграми спрямованості та орієнтації передавальної й приймальної електромагнітних антен, а також положення МБПЛА у просторі. Основні фактори, що впливають на параметри сигналу: зміна кута між напрямком діаграм спрямованості антен і напрямком на центр АВ МБПЛА, варіації залежності коефіцієнта відбиття та інтенсивності електромагнітного поля, а також зміна відстані.

Неоднорідності діелектричної проникності, зумовлені акустичним випромінюванням МБПЛА, моделювались у вигляді набору блискучих точок (БТ), розташованих на концентричних сферах – хвильових поверхнях акустичних хвиль, які формуються з кроком, що відповідає довжинам хвиль відповідної лопатевої гармоніки випромінювання МБПЛА. З часом радіус окремих сфер збільшується відповідно до швидкості звуку в повітрі.

Для спрощення розрахунків далі розглядається двовимірна просторова модель. У ній акустичні випромінювання МБПЛА подані у вигляді набору БТ, розташованих на концентричних колах з кроком, що відповідає довжинам акустичних хвиль.

У процесі моделювання розраховуються затримки електромагнітного сигналу для кожної БТ на хвильових поверхнях, визначається довжина променів падаючих і відбитих хвиль. Фаза сигналу визначається за зміною відстані. Також задається необхідне співвідношення сигнал/шум для сигналів, що обробляються.

Рівень розсіяного електромагнітного сигналу на вході приймального пристрою визначається шляхом геометричного складання векторів напруженості поля БТ, рівномірно розподілених на хвильових фронтах кожного періоду АВ з урахуванням згасання електромагнітних та акустичних сигналів, коефіцієнта відбиття, форми діаграм спрямованості електромагнітних передавальної та приймальної антен, а також форми діаграми спрямованості АВ МБПЛА.

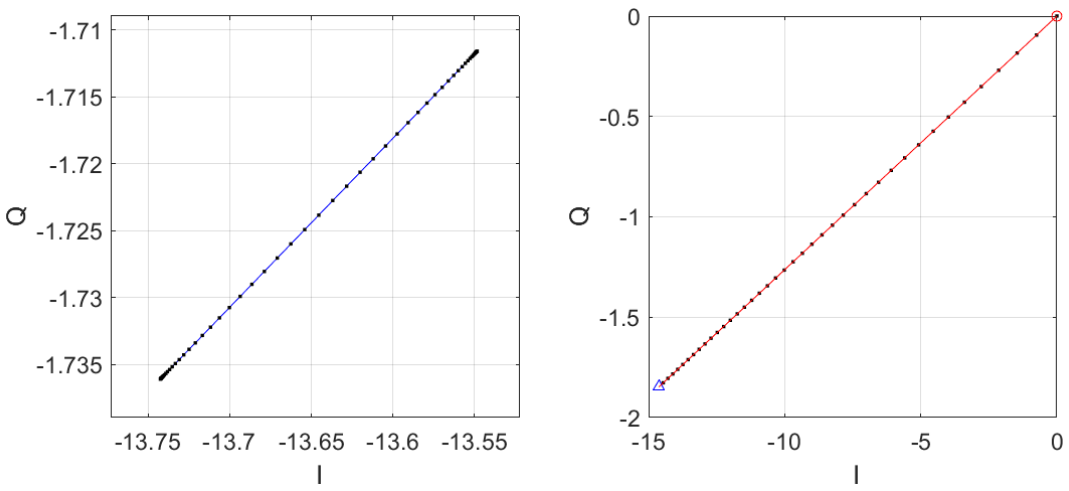
Годограф напруженості електричного поля розсіяного сигналу, побудований на комплексній площині, дозволяє візуалізувати зміни фази, амплітуди сигналу, оцінити ступінь синфазності. При моделюванні БТ окремих періодів акустичного коливання формують хвилі, що приходять у приймальну антену. При зменшенні інтервалу, на якому розташовані БТ в межах одного періоду АВ, фази сигналів сусідніх БТ змінюються плавно і ламана лінія годографа напруженості електричного поля звертається в плавну криву.

Заумов Брегга амплітуди векторів напруженості електричного поля, створюваних окремими хвильовими поверхнями, мало відрізняються між собою. Фаза результуючого вектора, сформованого окремою хвильовою поверхнею, залежить від взаємного розташування БТ, а також від положення передавальної і приймальної антен. Вектори результуючих напруженостей поля, створених сусідніми хвильовими поверхнями, за умови дотримання умови Брегга спрямовані в один бік, тобто їхні фази збігаються, а амплітуди визначаються коефіцієнтом відбиття $K_в$.

На рис. 6 подано годографи напруженості електричного поля розсіяного сигналу у системі РАЗ від однієї хвильової поверхні (*а*) і від усіх хвильових поверхонь пакета (*б*).

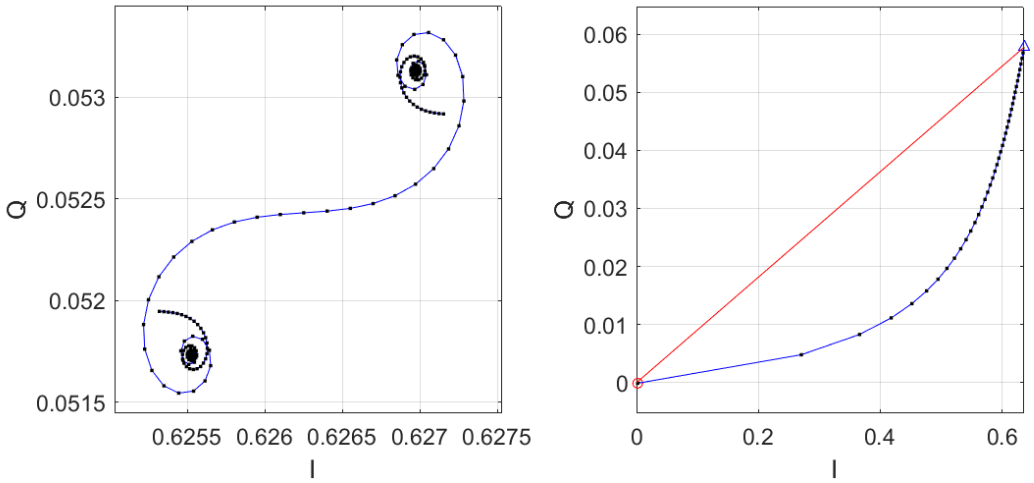
На кінцях годографа сигналу від однієї хвильової спостерігається зниження рівня сигналу, зумовлене впливом діаграми спрямованості антени. З урахуванням незначного ослаблення акустичного сигналу, внаслідок збільшення відстані (рис. 6, *б*), для годографа від усіх хвильових поверхонь пакета фіксується векторне складання синфазних складових, що свідчить про сприятливі умови для формування відбитого сигналу.

Розглянемо особливості формування розсіяного сигналу в умовах дифракції електромагнітних хвиль на звукових коливаннях АВ МБПЛА. Будемо вважати, що частота другої лопатевої гармоніки АВ МБПЛА становить $f_а$*=340 Гц,* робоча частота РЛС *1,2 ГГц,* порядок дифракції *m=8*, ширина діаграми спрямованості антени 6° та відстань до МБПЛА 100 м. За умови орієнтації антени РЛС у напрямку на МБПЛА, на рис. 7 представлено годографи напруженості електричного поля розсіяного сигналу від однієї хвильової поверхні (*а*) і від усіх хвильових поверхонь АВ (*б*).



а *б*

Рис. 6. Годографи відбитого сигналу у системі РАЗ від однієї хвильової поверхні (*а*) і від усіх хвильових поверхонь пакета (*б*)



а *б*

Рис. 7. Годографи відбитого сигналу за умов орієнтації антени РЛС на МБПЛА від однієї хвильової поверхні (*а*) і від усіх хвильових поверхонь АВ (*б*)

Основний внесок у результуюче поле дають БТ, розташовані на хвильовій поверхні, в межах якої різниця ходу променів не перевищує половини довжини хвилі тобто, першої зони Френеля. Поля, створювані БТ на краях хвильових поверхонь, де фаза не відповідає умовам синфазності, взаємно компенсуються, на годографі напруженості електричного поля взаємодія ціх складових представлена спіралями, що закручуються.

За розглянутих умов формування розсіяного сигналу основна частина енергії прийнятого електромагнітного сигналу формується внаслідок розсіювання на перших 20 періодах акустичних коливань. Подальші коливання мають значно менший вплив на рівень сигналу. На відміну від системи РАЗ, у випадку радіоакустичної локації зростання напруженості електромагнітного поля не є пропорційним кількості хвильових поверхонь, оскільки внесок наступних хвильових поверхонь суттєво зменшується через згасання рівня акустичного сигналу.

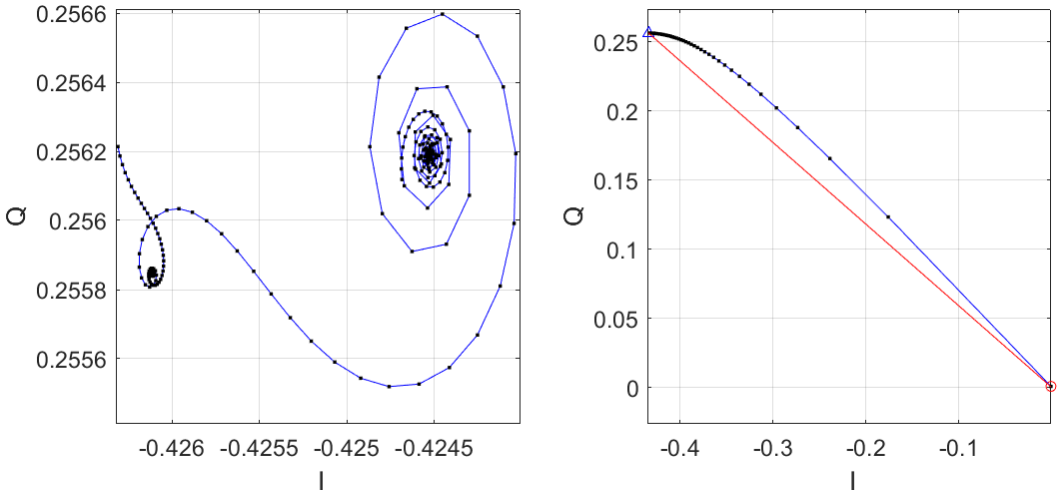
На рис. 8 показано годографи напруженості електричного поля розсіяного сигналу у випадку відхилення антени РЛС від напряму на МБПЛА.

Для оцінки ефективності сумування окремих складових поля розсіяного сигналу введемо нормований коефіцієнт відбиття:

$$K_{вн}=L_c/\ L_{cc},$$

де L_c – довжина сумарного вектора, яку одержали у модельному експерименті; L_{cc} – довжина сумарного вектора, за умови повної синфазності окремих складових поля розсіяного сигналу.

За умови орієнтації антени РЛС у напрямку на МБПЛА було проведено дослідження залежності нормованого коефіцієнта відбиття $K_{вн}$ від ширини діаграми спрямованості антени РЛС для кількох фіксованих відстаней до МБПЛА (75, 150, 225 м). Високі значення коефіцієнта відбиття $K_{вн}$ досягаються лише за використання високоспрямованих антен із шириною діаграми спрямованості менше ніж 4–6°, що зумовлено необхідністю забезпечення просторової селекції синфазно відбивних ділянок хвильових поверхонь АВ МБПЛА.



а *б*

Рис. 8. Годографи відбитого сигналу за умов відхилення антени від напряму на МБПЛА від однієї хвильової поверхні (*а*) і від усіх хвильових поверхонь АВ (*б*)

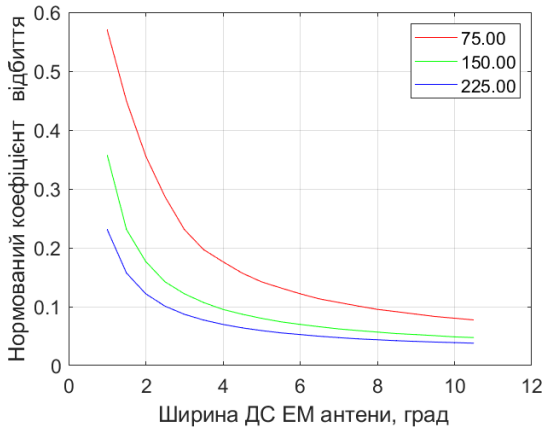


Рис. 9. Залежність нормованого коефіцієнта відбиття від ширини діаграми спрямованості антени при різних відстанях (75, 150, 225 м)

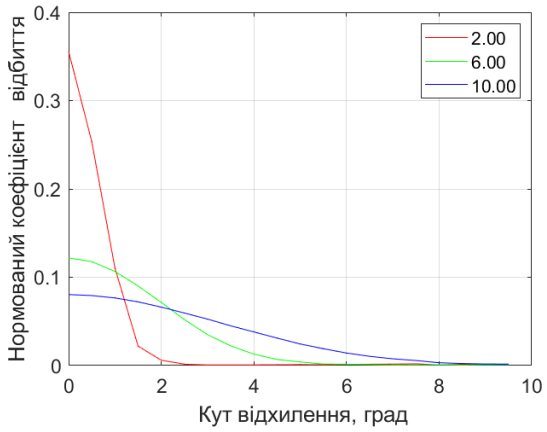


Рис. 10. Залежність нормованого коефіцієнта відбиття від кута відхилення антени від МБПЛА при різній ширині діаграми спрямованості (2°, 6°, 10°)

На рис. 10 показано залежність нормованого коефіцієнта відбиття від кута відхилення антени від МБПЛА за різної ширини діаграми спрямованості (2°, 6°, 10°) у випадку відхилення антени РЛС від напрямку на МБПЛА. Відхилення антени РЛС від напрямку на МБПЛА руйнує умови, за яких забезпечується синфазне складання розсіяних сигналів від хвильових поверхонь акустичного поля МБПЛА.

Висновки

1. Визначається, що сучасні технічні засоби виявлення малопомітних МБПЛА у деяких випадках не забезпечують належної оперативності та достовірності їх виявлення. Одним із перспективних напрямів підвищення ефективності виявлення малопомітних МБПЛА є використання явища розсіювання електромагнітних хвиль на акустичних збуреннях, які створюються цими апаратами в навколишньому повітряному середовищі

2. Спектр акустичного випромінювання МБПЛА містить гармонійні складові частоти обертання ротора та гармоніки лопатевої частоти. Його просторова спрямованість є складною і під час маневрування впливає на інтенсивність і форму спектра прийнятого сигналу.

3. У разі радіоакустичної локації МБПЛА сигнал на вході приймача є результатом інтерференції електромагнітних хвиль, відбитих від акустичних хвиль, що рухаються у протилежних напрямках. Параметри огинаючої цього сигналу залежать від швидкості звуку, частоти акустичного випромінювання, а також від співвідношення потужностей сигналів, сформованих ділянками до і після МБПЛА в межах діаграми спрямованості антени.

4. Робоча смуга частот локаційної системи повинна охоплювати діапазон, у якому забезпечується ефективне резонансне відбиття електромагнітних хвиль. Для зменшення габаритів високоспрямованих антен слід обирати робочу частоту РЛС таким чином, щоб забезпечувалося виконання умов Брегга з використанням дифракції вищих порядків.

5. Методом імітаційного моделювання досліджено особливості дифракції вищих порядків електромагнітних хвиль на акустичному випромінюванні МБПЛА. Встановлено, що основна частина енергії розсіяного сигналу формується на акустичних коливаннях у безпосередній близькості до МБПЛА, тоді як внесок коливань на великих відстанях є незначним через їх згасання.

6. Під час радіоакустичної локації акустичного випромінювання МБПЛА високі значення коефіцієнта відбиття досягаються лише за використання вузькоспрямованих антен із шириною діаграми спрямованості менше ніж 4–6°, що пов’язано з необхідністю просторової селекції ділянок хвильових поверхонь акустичного випромінювання МБПЛА, що забезпечують синфазне складання розсіяних сигналів.

7. Відхилення антени РЛС від напрямку на МБПЛА руйнує умови, за яких забезпечується синфазне складання розсіяних сигналів від хвильових поверхонь акустичного поля МБПЛА.

Список літератури:

1. Медведєв, В.К., Коренівська, І.С., Хажанець, Ю.А., Салов, А.О. Безпілотні літальні апарати та їхній вплив на перебіг російсько-української війни // Наука і оборона. 2023. №2. С. 52–59.
2. Чому оптоволоконні безпілотники можуть схилити шальки терезів у війні між Україною та Росією. URL: https://vgi.com.ua/en/ukraine-russian-war-fiber-optics/?utm_source=chatgpt.com.
3. Kartashov V.M., Oleynikov V.N, Sheyko S.A., Koryttsev I.V., Babkin S.I., Zubkov O.V. Peculiarities of small unmanned aerial vehicles detection and recognition // Telecommunications and Radio Engineering. New York. 2019. Vol. 78, Issue 9. P. 771–781.
4. Олейников В.М., Зубков О.В., Карташов В.М., Коритцев І.В., Бабкин С.І., Шейко С.О. Дослідження ефективності виявлення і розпізнавання малорозмірних безпілотних літальних апаратів по їх акустичному випромінюванню // Радіотехніка. 2018. Вип. 195. С. 209–217.
5. Карташов В.М., Олейніков В.Н., Рябуха В.П., Бабкін С.И., Воронін В.В., Капуста А.И., Селєзнев И.С Методи комплексної обробки та інтерпретації радіолокаційних, акустичних, оптичних і інфрачервоних сигналів безпілотних літальних апаратів // Радіотехніка. 2020. Вип. 202. С. 173–182.
6. Карташов В.М., Олейніков В.М., Шейко С.О., Бабкін С.І., Коритцев І.В.., Зубков О.В. Особливості виявлення та розпізнавання малих безпілотних літальних апаратів // Радіотехніка. 2018. Вип. 195. С. 235–243.
7. Карташов В.М., Посошенко В.О., Воронін В.В., Колесник В.І., Капуста А.І., Рибников М.В., Першин Є.В. Методи виявлення-розпізнавання радіолокаційних, акустичних, оптичних і інфрачервоних сигналів безпілотних літальних апаратів // Радіотехніка. 2021. Вип. 205. С. 138–153.
8. Пат. 127007 Україна, МПК G 01 S 13/00 G 01 S 17/00. Радіоакустичний спосіб виявлення малопомітних безпілотних літальних апаратів / В.В. Семенець та ін. № а 202004704 ; заявл. 24.07.2020 ; опубл. 08.03.2023 ; Бюл. № 10. 9 с.

9. Карташов В.М., Харченко О.І., Посошенко В.О., Колісник В.І., Єгоров А.Б., Тимошенко Л.П., Капуста А.І. Виявлення безпілотних літальних апаратів з використанням розсіювання радіохвиль на акустичних обуреннях середовища, що створюються літальними апаратами // Радіотехніка. 2021. Вип. 206. С. 122–130.

10. Карташов В.М., Посошенко В.О., Колісник В.І. та ін. Виявлення радіолокаційних сигналів, розсіяних на акустичних збуреннях, створюваних БПЛА // Радіотехніка. 2021. Вип. 207. С. 113–122.

11. Карташов В.М., Посошенко В.О., Колісник К.В., Колісник В.І., Бобнєв Р.О., Капуста А.И. Алгоритм оцінювання розподілу енергії радіолокаційних сигналів, які розсіюються на акустичних збуреннях, створених МБПЛА // Радіотехніка. 2022. Вип. 211. С. 16–25.

12. Борн М., Вольф Е. Основи оптики. 1973. 719 с.

13. Каллистратова М.А., Кон А.И. Радиоакустическое зондирование атмосферы. 1985. 196 с.

14. Дистанционные методы и средства исследования процессов в атмосфере Земли ; под ред. Б.Л. Канцева, Е.Г. Прошкина, М.Ф. Лагутина. Разд. 2. Радиоакустическое зондирование пограничного слоя атмосферы. Харьков : Коллегиум, 2002. С. 44–98.

15. Прошкин Б. Г. Определение основных метеорологических величин в пограничном слое атмосферы методом радиоакустического зондирования // Радиотехника. 1996. Вип. 100. С. 196–204.

16. Карташов В.М., Олейников В.Н., Шейко С.А., Бабкин С.И., Корытцев И.В., Зубков О.В., Анохин М.А. Информационные характеристики звукового излучения малых беспилотных летательных аппаратов // Радиотехника. 2017. Вип. 191. С. 181–187.

17. Oleynikov V.N., Kartashov, V.M., Babkin, S. I., Zubkov, O.V., Korytsev I.V., Sheiko, S.A. Seleznov I.S. Structure and Parameter Unmanned Aerial Vehicles Sound Fields // Telecommunications and Radio Engineering. New York. 2020. Vol. 79, №17. P.1539–1550.

Надійшла до редколегії 28.05.2025

Відомості про автора:

Олейніков Володимир Миколайович – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; e-mail: vladimir.oleinikov@nure.ua; ORCID: https://orcid.org/0000-0002-3358-5987

УДК 621.311.6:006.83 **DOI:10.30837/rt.2025.3.222.14**

О.В. ВОРГУЛЬ, канд. техн. наук, І.В. ІГНАТЮК, Т.В. МАЧОНІС, О.Д. ШУНІБОРОВ

БЕЗДРОТОВА ПЕРЕДАЧА ЕНЕРГІЇ (БПЕ): АНАЛІЗ СТАНДАРТІВ, КОМЕРЦІЙНИХ ТЕХНОЛОГІЙ ТА ПЕРСПЕКТИВ

Вступ

Дана робота є спробою скласти структурований звіт про бездротову передачу енергії (БПЕ) на основі двох стандартів від МСЕ [2303, 2392] і фактичні комерційні результати. По-перше, авторів цікавить цілепокладання цього напряму. На нашу думку, використання кабелів робить електроніку безпечнішою. Тоді, з якою ж метою робити БПЕ, коли вона ефективніша та інша? По-друге, ознайомившись зі звітами МСЕ, ми знаємо, що за дальністю роботи БПЭ діляться на групи – контактні (зарядні станції різної потужності) – середня зона – далека зона. Остання має на увазі віддалення на кілометри, можливо, навіть ближній космос. Нас цікавить опис кожної групи, практичні приклади проєктів, основні вимоги та параметри. По-третє, пошуки на цю тематику приводять нас до комерційних рішень, таких, як Qi, Air Fuel Resonant та інші. Наше завдання – збирання та структурування короткої інформації, можливо, історії з динамікою, коротко – параметри та практичне застосування.

1. Цілепокладання БПЕ: Навіщо її використовувати, якщо є кабелі?

Кабелі дійсно забезпечують найвищу ефективність, надійність та безпеку для стаціонарних пристроїв, хоча їх використання не завжди можливе і не завжди зручне. Проте БПЕ вирішує принципово інші завдання:

Зручність та спрощення: усуває необхідність фізичного підключення/відключення (заряджання смартфонів, навушників, зубних щіток, інструментів). Особливо критично для пристроїв із частим використанням/підзарядкою.

Підвищена надійність у специфічних умовах.

Усунення роз'ємів:

- знижує зношування (механічні пошкодження контактів);
- виключає корозію контактів від вологи, пилу, хімікатів (промисловість, вуличні пристрої);
- дозволяє створювати повністю герметичні корпуси (медичні імплантати, підводні датчики).

Мобільність і автономність: заряджання пристроїв у русі (електромобілі на світлофорах/зупинках, AGV/роботи на складах, дрони).

Доступність і безпека: заряджання пристроїв у труднодоступних, небезпечних або стерильних середовищах (біомедичні датчики всередині тіла, пристрої в чистих кімнатах, обладнання у вибухонебезпечних зонах) без порушення цілісності середовища.

Динамічне електроживлення: передача енергії до об'єктів, що постійно рухаються (крани на виробництві, поїзди, конвеєрні системи).

Резервне або альтернативне харчування: бездротова доставка енергії в зони лиха або віддалені локації (супутники, IoT-сенсори в полях/лісах).

Естетика та дизайн: приховані зарядні поверхні в меблях, автомобілях, громадських місцях.

Висновок з цілепокладання: БПЕ не прагне повністю замінити кабелі там, де вони оптимальні. Вона створює нові можливості та вирішує специфічні проблеми, де фізичне з'єднання неможливе, незручне, небезпечне або знижує надійність/мобільність.

2. Класифікація БПЕ за дальністю (на основі підходів МСЕ)

МСЕ (ITU) визначає кілька класів, здебільшого можна розглядати три зони:

2.1. Контактна/ближня зона (Near-Field / Contact-based):

Опис: передача на відстані значно меншій за розміром передавальної/прийомної котушки (сантиметри). Засновано на *магнітній індукції* (Qi) або *магнітному резонансі* (ефективніший на трохи більшій відстані/при розбіжності осей).

Фізика: сильний зв'язок котушок. Енергія передається через змінне магнітне поле.

Приклади проєктів чи застосування:

- Qi (WPC): смартфони, навушники, розумний годинник, мишки, клавіатури, медичні інструменти (швидка стерилізація без роз'ємів);
- кухонна техніка: індукційні плити (дуже висока потужність);
- промисловість: заряджання безпілотних AGV (Automatic Guided Vehicles) на станціях;
- електромобілі: стаціонарна зарядка паркування (SAE J2954).

Основні параметри та вимоги:

- дальність: міліметри – сантиметри (до ~5–10 см для резонансних);
- потужність: мілівати (датчики) – кіловати (індукційні плити, швидка зарядка EV);
- ККД: високий (70–95 %), сильно залежить від відстані, поєднання котушок та навантаження;
- безпека: стандарти (Qi, SAE J2954) суворо регламентують рівні ЕМП, контроль температури сторонніх об'єктів (FOD), зв'язок між передавачем та приймачем;

Ключові вимоги: точне позиціонування (особливо для індукції), контроль якості зв'язку, керування потужністю.

2.2. Середня зона (Mid-Field):

Опис: передача на відстані, порівнянному або трохи більшому за розмір антен (десятки сантиметрів – одиниці метрів). Використовує *магнітний резонанс* або *радіочастотні (РЧ) методи* (спрямовані антени).

Фізика: ослаблений зв'язок котушок (резонанс) або спрямоване випромінювання РЧ-хвиль. Вимагає більш складного управління.

Приклади проєктів/застосування:

- побут: заряджання на столі/поверхні без точного позиціонування (AirFuel Resonant);
- медицина*:* зарядка імплантатів (кардіостимулятори, нейростимулятори) через шкіру;
- робототехніка*:* зарядка сервісних роботів при заїзді на базу;
- споживча електроніка*:* "зарядка в кімнаті" (концепти, що розвиваються);
- IoT: живлення/зарядка датчиків у межах кімнати або невеликого приміщення.

Основні параметри/вимоги:

- дальність: ~0,5 – 5 м (умовно);
- потужність**:** мілівати – сотні ватт (зазвичай, десятки ватів для електроніки);
- ККД: помірний (10–50 %), різко падає з відстанню;
- безпека**:** суворий контроль ЕМП, особливо для РЧ-методів. Гарантування безпеки людей та тварин у зоні дії. Регламенти FCC, ICNIRP;

Ключові вимоги: управління променем/фокусування (для РЧ), компенсація розладу резонансу, складна електроніка, питання безпеки та регулювання.

2.3. Далека зона (Far-Field):

Опис: передача на відстанях, що значно перевищують розмір антен (метри – кілометри). Використовує *спрямовані електромагнітні хвилі* (лазери, НВЧ-випромінювання).

Фізика: перетворення енергії у вузьконаправлений промінь (лазер або НВЧ), уловлювання променя приймальної антени (фотовольтаїка для лазера, ректену – для НВЧ).

Приклади проєктів/застосування:

- космос: передача енергії з орбітальних СБС (Сонячні космічні електростанції – концепти JAXA, Caltech, ESA) на Землю; живлення супутників/зондів від материнського корабля;
- БПЛА: підзарядка дронів у польоті від наземної станції чи іншого дрона (PowerLight Technologies, DARPA проєкти);
- віддалені об'єкти*:* живлення IoT-сенсорів, метеостанцій у важкодоступних місцях;
- військові: бездротова передача енергії на полі бою.

Основні параметри/вимоги:

- дальність: метри – кілометри (теоретично до орбіти і далі);
- потужність: може бути дуже високою (кіловати-мегавати для СБС);
- ККД: низький (одиниці – десятки відсотків на *всьому* ланцюжку "генерація – передача – прийом – перетворення"). Великі втрати на розбіжності променя та атмосферні ефекти;
- безпека: критичний чинник**.** Небезпека високоінтенсивного променя (термічне ураження, ушкодження очей для лазерів, вплив НВЧ на живі організми). Вимагає виняткових заходів безпеки (захищені коридори, миттєве вимкнення при порушенні променя).

Ключові вимоги: надточне наведення та відстеження променя, найвища ефективність перетворювачів, вирішення проблем атмосферної інтерференції (особливо для лазерів), подолання нормативно-правових бар'єрів безпеки.

3. Комерційні рішення та стандарти

3.1. Qi (Wireless Power Consortium – WPC):

Опис: домінуючий світовий стандарт для *індуктивної* та *резонансної* (базова версія) зарядки малопотужних пристроїв (<15Вт для базового, до 50–100Вт+ у розробці/специфічних реалізаціях).

Історія/Динаміка: заснований у 2008 р. Став де-факто стандартом для смартфонів (з Apple iPhone 8/X у 2017 р.). Постійний розвиток: збільшення потужності, покращення позиціонування (EPP), додавання резонансного режиму.

Параметри:

- потужність: 5 Вт (Baseline), 15Вт (Extended Power Profile), до 50Вт + (спеціальні профілі);
- частота: ~100–205 кГц (низькочастотний діапазон);
- дальність: міліметри – сантиметри (індукція), до ~4-5 см (резонанс у специфікації);
- ККД: 70–80 % у добрих умовах.

Практичне застосування: повсюдне. Зарядні станції в аеропортах, кафе, автомобілях, меблях. Пристрої: смартфони, навушники, годинники, гаджети. Головна перевага – універсальність та сумісність.

Ресурси: https://www.wirelesspowerconsortium.com/

3.2. AirFuel Alliance (Об'єднання AirFuel Resonant та PMA):

Опис: консорціум, що просуває магнітно-резонансну (Magnetic Resonance) та радіочастотну (RF) технології для більшої гнучкості (відстань, позиціонування, одночасне заряджання декількох пристроїв).

Історія/динаміка: утворено в 2015 р. злиттям Alliance for Wireless Power (A4WP, резонанс) та Power Matters Alliance (PMA, індукція). Фокус змістився на резонанс і РЧ як альтернативу/доповнення Qi. Комерційне проникнення значно поступається Qi, але є нішеві застосування (громадські зарядки в Starbucks, деякі ноутбуки Dell / Lenovo, медичні/промислові рішення).

Параметри (резонанс):

- потужність: до 65Вт (і вище у специфікаціях);

- частота: 6,78 МГц (ISM band);

- дальність: до 5 см (стандарт), реальні реалізації можуть працювати на 10–20 см зі зниженим ККД;

- ККД: порівняємо з Qi при контакті, краще при несуміщенні/великій відстані.

Параметри (RF):

- потужність: міллівати – одиниці Ватт (для зарядки);

- частота: зазвичай 900 МГц, 2.4 ГГц, 5.8 ГГц;

- дальність: метрии;

- ККД: дуже низький (одиниці відсотків), підходить тільки для малопотужних пристроїв (IoT-датчики, слухові апарати).

Практичне застосування:

- резонанс: публічні зарядні зони (кафе, аеропорти), заряджання ноутбуків/планшетів, промислові/медичні програми (що вимагають гнучкості позиціонування);

- RF: бездротові датчики, мітки, малопотужні пристрої (де заміна батарейки складна), слухові апарати. *Не* підходить для швидкого заряджання смартфонів.

Ресурси: https://www.airfuel.org/

3.3. WiTricity (Технологія та компанія):

Опис: піонер у комерціалізації *сильнозв'язаного магнітного резонансу* (Higly Resonant Wireless Power Transfer). Засновано на роботах MIT (2007).

Історія/динаміка: засновано у 2007 р. для ліцензування технології. Ліцензіати включають виробників автокомпонентів (для EV), медичного обладнання, споживчої електроніки. Технологія лежить в основі резонансного режиму AirFuel і частково вплинула на розвиток резонансу Qi. Фокус на додатках з вимогою більшої дистанції/гнучкості, ніж індукція (EV, роботи, медичні імплантати).

Параметри:

- аналогічні AirFuel Resonant (частота ~100 кГц – 10 МГц;

- дальність до ~10–20 см (при хорошому ККД для потужних систем).

Практичне застосування: системи бездротової зарядки електромобілів (SAE J2954 заснований частково на цій технології), промислові AGV, медичні імплантати, спеціалізовані споживчі та промислові рішення.

Ресурси: https://witricity.com/

4. Ключові ресурси для глибокого занурення

Міжнародний союз електрозв'язку (МСЕ/ITU):

- сектор радіозв'язку (ITU-R): вивчайте звіти та рекомендації робочої групи 5B (WP 5B) "Wireless power transmission technologies and systems". Вони охоплюють усі аспекти, включаючи безпеку, сумісність, методи. https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5b/Pages/default.aspx;

- шукайте конкретні звіти серії "ITU-R SM." (Spectrum Management) та "ITU-R M." (Mobile, radiodetermination, amateur and related satellite services), що стосуються БПЕ.

IEEE:

- журнал IEEE Transactions on Power Electronics: Ключові наукові публікації за схемами, методами управління, проєктування котушок;

- IEEE Journal of Electromagnetics, RF і Microwaves in Medicine and Biology: для медичних додатків БПЕ;

- IEEE Microwave Magazine: оглядові статті з сучасних технологій, включаючи БПЕ далекої зони (НВЧ, лазери);

- конференції: IEEE Wireless Power Transfer Conference (WPTC), IEEE Energy Conversion Congress and Exposition (ECCE), IEEE International Microwave Симпозіум (IMS).

Консорціуми:

- Wireless Power Consortium (WPC – Qi): https://www.wireless power consortium. com / (Специфікації, списки сертифікованих пристроїв);
- AirFuel Alliance: https://www.airfuel.org/ (Специфікації, новини, кейси).

Книги (фундаментальні):

- Kurs, A. et al. (2007). Wireless Power Transfer via Strongly Coupled Magnetic Resonances. Science (основна стаття MIT);
- Sample AP, Meyer DT, & Smith JR (Eds.). (2015). Wireless Power Transfer: Principles and Engineering Explorations. CRC Press. (Гарний огляд різних технологій);
- Bosshard R., & Kolar JW (2016). Inductive Power Transfer: System Analysis and Optimization. Springer. (Поглиблено в індукцію);
- Shinohara N. (2014). Wireless Power Transfer via Radiowaves. Wiley**.** (Фокус на дальній зоні, НВЧ).

Оглядові статті (IEEE, Nature Electronics):

- шукайте огляди по "wireless power transfer", "wireless charging", "microwave power transmission", "laser power beaming", "wireless power for implants / IoT /EV".

Новинки та аналітичні портали:

- Wireless Power Technology (https://www.wirelesspowertransfertech.com/);
- Charged EVs (розділ про бездротову зарядку EV) (https://chargedevs.com/);
- IEEE Spectrum (https://spectrum.ieee.org/). – Пошук за ключовими словами.

Висновок

БПЕ – область, що активно розвивається, де технологія визначається застосуванням. Контактна/ближня зона (Qi) стала масовим ринком завдяки зручності. Стандарти (WPC *Qi*, AirFuel) і дослідження (під егідою ITU, IEEE) критично важливі для забезпечення сумісності, ефективності та безпеки.

Список літератури:

1. Офіційний сайт консорціуму WPC: Переваги технології Qi [Електронний ресурс]. URL: https://www.wirelesspowerconsortium.com/benefits/
2. IEEE Journal of Electromagnetics, RF і Microwaves in Medicine and Biology [Електронний ресурс]. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=7507313
3. Sample A.P. Wireless Power Transfer : Principles and Engineering Explorations / A.P. Sample, D.T. Meyer, J.RSmith (eds.). Boca Raton : CRC Press, 2015. 358 p.
4. Офіційний сайт WiTricity : Промислова автоматизація [Електронний ресурс]. URL: https://witricity.com/applications/industrial-automation/
5. Офіційний сайт AirFuel Alliance : Програми. Небезпечні середовища [Електронний ресурс]. URL: https://www.airfuel.org/technology/applications/
6. Офіційний сайт Bombardier Primove [Електронний ресурс]. URL: https://primove.bombardier.com/
7. Caltech Space Solar Power Project [Електронний ресурс]. URL: https://www.spacesolar.caltech.edu/ .
8. Рік. MCE- R SM.2303-1 (05/2021). Бездротова передача енергії з використанням засобів радіозв'язку = Wireless power transmission using radiocommunication means [Електронний ресурс]. URL: https://www.itu.int/rec/R-REC-SM.2303
9. Робоча програма MCE-R 5B: Робота з бездротовою передачею енергії (WPT) = ITU-R WP 5B Work on WPT [Електронний ресурс]. URL: https://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5b/Pages/default.aspx
10. Рік. MCE-R SM.2392-0 (10/2022). Методи вимірювання систем бездротової передачі енергії для застосування в радіозв'язку = Measurement techniques of wireless power transmission systems for radiocommunication applications [Електронний ресурс]. URL: https://www.itu.int/rec/R-REC-SM.2392 (дата звернення: 30.05.2024).
11. Shinohara N. Wireless Power Transfer via Radiowaves. Chichester : Wiley, 2014. 232 p.
12. Офіційний сайт Консорціуму бездротового живлення (WPC – Qi) [Електронний ресурс]. URL: https://www.wirelesspowerconsortium.com/
13. Офіційний сайт AirFuel Alliance [Електронний ресурс]. URL: https://www.airfuel.org/
14. Офіційний сайт WiTricity [Електронний ресурс]. URL: https://witricity.com/
15. SAE J2954: 202010. Бездротова передача енергії для легкових електромобілів = Wireless Power Transfer for Light-Duty Electric Vehicles [Електронний ресурс]. URL: https://www.sae.org/standards/content/j2954_202010/

16. Офіційний сайт Drayson Technologies : Технологія Freevolt [Електронний ресурс]. URL: https://www.draysontechnologies.com/technology/freevolt
17. Офіційний сайт ElectReon : Динамічний бездротовий заряд електромобілів [Електронний ресурс]. URL: https://electreon.com/
18. IEEE Transactions on Biomedical Circuits and Systems (TBioCAS) [Електронний ресурс]. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=4157303.
19. Офіційний сайт JAXA: Системи космічних сонячних електростанцій (SSPS) = JAXA: Space Solar Power Systems [Електронний ресурс] URL: https://www.kenkai.jaxa.jp/eng/research/ssps/ssps.html
20. Офіційний сайт ЕКА: Ініціатива SOLARIS = ESA: SOLARIS Initiative [Електронний ресурс]. URL: https://www.esa.int/Enabling_Support/Space_Engineering_Technology/SOLARIS
21. Офіційний сайт PowerLight Technologies (Лазерна передача енергії) = PowerLight Technologies (Laser Power Beaming) [Електронний ресурс]. URL: https://powerlighttech.com/
22. Офіційний сайт Emrod : Бездротова передача енергії на великі відстані = Emrod : Long-Range WPT [Електронний ресурс]. URL: https://emrod.energy/
23. Офіційний сайт uBeam (Ультразвукова передача енергії) = uBeam Ultrasound WPT) [Електронний ресурс]. URL: https://ubeam.com/
24. IEEE Transactions on Antennas and Propagation (TAP) [Електронний ресурс]. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=8
25. ICNIRP Guidelines. Guidelines for Limiting Exposure to Electromagnetic Fields (100 kHz to 300 GHz) [Електронний ресурс] // Міжнародна комісія із захисту від неіонізуючих випромінювань (ICNIRP). URL: https://www.icnirp.org/en/publications/index.html
26. IEEE Transactions on Power Electronics [Електронний ресурс]. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=63
27. IEEE Microwave Magazine [Електронний ресурс]. URL: https://ieeexplore.ieee.org/xpl/RecentIssue.jsp?punumber=6668

Надійшла до редколегії 19.07.2025

Відомості про авторів:

Воргуль Олександр Васильович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри мікропроцесорних технологій і систем; Україна; e-mail: oleksandr.vorgul@nure.ua; ORCID: https://orcid.org/0000-0002-7659-8796

Ігнатюк Іван Валентинович – Харківський національний університет радіоелектроніки, магістр кафедри мікропроцесорних технологій і систем; Україна; e-mail: ivan.ihnatiuk@nure.ua

Мачоніс Томас Володимирович – Харківський національний університет радіоелектроніки, магістр кафедри мікропроцесорних технологій і систем; Україна; e-mail: tomas.machonis@nure.ua

Шуніборов Олег Дмитрович – Харківський національний університет радіоелектроніки, магістр кафедри мікропроцесорних технологій і систем; Україна; e-mail: oleh.shuniborov@nure.ua

ELECTRONIC COMMUNICATIONS
ЕЛЕКТРОННІ КОМУНІКАЦІЇ

УДК 621.396.2 **DOI:10.30837/rt.2025.3.222.15**

В.В. ДОВГИЙ, канд. техн. наук, В.М. ГРИГА, канд. техн. наук,
Б.С. ДЗУНДЗА, д-р техн. наук, І.В. СВИД, канд. техн. наук,
А.І. ТЕРЛЕЦЬКИЙ, канд. фіз.-мат. наук, М.Ф. ПАВЛЮК, канд. фіз.-мат. наук

АНАЛІЗ ЦИФРОВИХ ІНТЕРФЕЙСІВ ПЕРЕДАЧІ ДАНИХ У КАНАЛАХ ЗВ'ЯЗКУ КОМП'ЮТЕРНИХ СИСТЕМ

Вступ

На даний час область використання комп’ютерів та Інтернету є настільки великою, що безперебійна робота комп’ютерних систем та мереж є критично важливою для більшості галузей економіки. Важливо, щоб обмін даними був швидким і без помилок, але ще більш важливо забезпечити стабільність зв’язку, зокрема і при аварійних відключеннях електрики. З кожним днем технології, що використовуються, стають кращими завдяки науково-технічному прогресу та розвитку елементної бази. Зокрема, з появою оптоволоконних каналів зв’язку досягнуто надвисоких швидкостей передачі даних [1 – 3]. Оптичні канали зв’язку забезпечують надійність зв’язку тільки у випадку, якщо оптичний кабель заведено напряму до абонента, де забезпечено резервне живлення [4, 5]. Але до тепер у використанні залишається велика кількість проводових каналів зв’язку [6].

Технологія DSL (цифрова абонентська лінія) також використовується для швидкої та безпечної передачі даних. Технологія DSL успішно виконує дистанційне підключення та передачу великих об’ємів даних з можливістю роботи від резервного живлення.

Авторами дослідження [7] технологія DSL розглядається з точки зору співвідношення зони використання, відстані та швидкості. Кожен тип xDSL аналізується та порівнюється зі швидкістю передачі даних, структурою (симетрія, асиметрія) і відстанню передачі. Надано пропозиції щодо того, яка технологія xDSL зручніша для конкретної задачі. Також наведено можливі переваги, які може надати технологія DSL. Зазначено, що технологія DSL може працювати з різними типами інтегрованих програм і ефективність може бути збільшена.

У роботі [8] розглянуто побудову інформаційних систем підвищеної надійності. Показано необхідність використання декількох каналів доступу в Інтернет та перспективу використанням ADSL як резервного каналу зв’язку. Запропонований варіант дозволяє використовувати програмне й апаратне резервування.

Авторами роботи [9] аналізується практична реалізація віртуальної приватної мережі (VPN) на основі xDSL. Представлено типи реалізації VPN. Обговорено соціально-економічні переваги використання технології xDSL.

Отже, незважаючи на наявність більш швидкісних технологій xDSL і зважаючи на наявність великої інфраструктури мереж та особливості роботи, xDSL досі залишається актуальною, особливо в якості резервного підключення.

Технології управління каналами передачі цифрових даних

Для того щоб цифровий обмін даними був ефективний, необхідно разом з контролем проводити і управління обміном. В роботі детально розглянуто пересилання даних у каналах зв'язку. Для якісного управління фізичним інтерфейсом передачі даних необхідно ввести рівень логіки, яка називається – управління каналом передачі даних. Щоб краще зрозуміти необхідність в управлінні каналом зв'язку, зупинимось на вимогах і умовах, які є необхідними для отримання ефективного зв'язку між комп'ютерними системами:

- синхронізація кадрів, яка зменшить помилковість;

- управління потоком цифрових даних та каналом зв'язку;
- захист від помилок, причому помилкові біти повинні бути виправлені;
- адресація як необхідна умова оптимальної передачі даних;
- передача сигналів управління і даних повинна здійснюватися по одному каналу зв'язку.

Проста форма управління потоком даних, відома як управління потоком із зупинками, працює наступним чином. Об'єкт джерела передає кадр адресату. Після того як об'єкт призначення приймає цей кадр, він демонструє свою готовність прийняти наступний кадр, для чого надсилає підтвердження приймання отриманого кадру. Перш, ніж джерело передає наступний кадр, необхідно отримати підтвердження приймання попереднього. Відповідно, адресат може зупинити потік даних, просто утримавшись від відправки підтвердження приймання. Така процедура працює надійно і її не можна покращити при передачі даних великими блоками.

Другий варіант управління полягає в тому, що в кожний момент в каналі передачі може знаходитись тільки один кадр. Якщо двійкова довжина каналу зв'язку перевищує довжину кадру, то канал використовується дуже неефективно. Суттєво підвищити ефективність передачі можна, якщо дозволити знаходження в каналі передачі декількох кадрів. Розглянемо як буде працювати дана технологія для двох станцій А і В, зв'язаних між собою дуплексним каналом передачі даних. Розмір буфера станції B складає W кадрів. Інакше В станція може прийняти W кадрів без факту підтвердження, щоб можна було слідкувати за тим, як підтверджується отримання кадрів, то останні отримують послідовні номера. Станція В підтверджує приймання кадру, відправляє підтвердження, що містить порядковий номер наступного очікуваного кадру. Це підтвердження також неявно вказує, що станція В готова прийняти наступні W кадрів, починаючи з кадру, що має вказаний номер. Описану систему можна використати і для підтвердження приймання декількох кадрів. Станція В може, наприклад, прийняти кадри 2, 3 і 4, але затримати підтвердження їх приймання до прибуття кадру 4. Повернувши підтвердження з порядковим номером 5, станція В одночасно підтверджує приймання кадрів 2, 3 та 4. Станція А зберігає список порядкових номерів кадрів, які вона може посилати, а станція В зберігає список порядкових номерів, які вона готова прийняти. Кожний із цих списків можна розглядати як вікно кадрів. Така схема називається управлінням потоком методом динамічних вікон.

Даний метод є більш ефективним, ніж попередній. Причина полягає в тому, що при використанні методу динамічних вікон канал передачі даних розглядається як конвеєр, який може перенаповнюватись кадрами, які знаходяться на шляху передачі. Метод із зупинками, навпаки, допускає в кожний момент часу наявність в каналі тільки одного кадру.

Розглянемо технологію виявлення помилок в каналі передачі. Для цього визначимо ймовірності, які відносяться до помилок в кадрах, що передаються у каналі зв'язку, а саме:

$P_в$ – ймовірність виникнення окремого помилкового біта, або інакше швидкість виникнення помилкового біта (Bit Error Rate – BER);

P_1 – ймовірність того, що в отриманому кадрі немає помилки;

P_2 – ймовірність того, що в отриманому кадрі є одна або декілька невиявлених помилок.

Тоді зв'язок між цими ймовірностями для кадру розміром F-бітів можна описати виразом

$$\begin{aligned} P_1 &= (1 - P_в)^F, \\ P_2 &= 1 - P_1, \end{aligned} \tag{1}$$

а це дозволяє сформулювати наступний висновок: при рості ймовірності виникнення помилкового біта зменшується ймовірність того, що кадр прийде без помилки, та із збільшенням розміру кадру ймовірність того, що він прийде без помилки, зменшується.

Проста схема виявлення помилки полягає в додаванні у кінці інформаційного блоку біта парності (рис. 1). Типовий приклад: передавання символів, при якій до кожного семибітового символу IRA додається біт парності. Значення цього біта вибирають таким чином, щоб зага-

льна кількість одиниць у символі була парною (додатна парність) і непарною (від'ємна парність). Тому, наприклад, при передаванні символу *G* в кодуванні IRA (1110001) і при використанні від'ємної парності передавач додає до блоку і передає вже символ як 11100011. Приймач вивчає отриманий код символу і при непарній кількості одиниць робить висновок про безпомилкову його передачу. Якщо в процесі передавання один біт (або будь-яка непарна кількість бітів) був помилково інвертованим (наприклад 1100001), то приймач виявить цю помилку. Додатна парність традиційно використовується при синхронній передачі, а від'ємна – при асинхронній передачі.

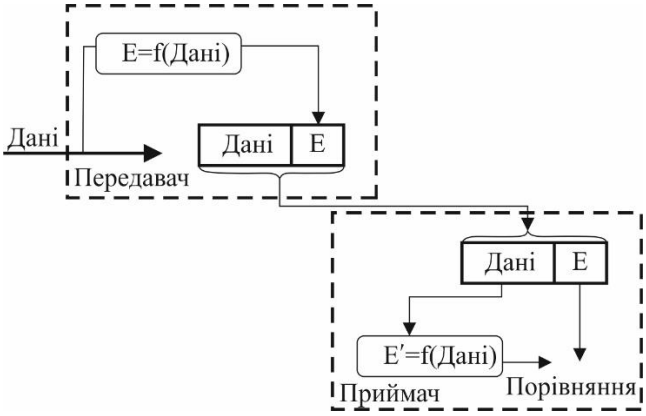


Рис. 1. Структурна схема виявлення помилок
(E,E′ – коди виявлення помилок f – функція коду виявлення помилок)

Одним із найбільш розповсюджених і найбільш потужних кодів виявлення помилок є циклічна перевірка парності з надлишковістю (Ciclyc Redundancy Check – CRC). Для даного R-бітового інформаційного блоку або повідомлення передавач генерує n-бітову послідовність, яка називається контрольною послідовністю кадру (Frame Check Sequence – FCS). Отриманий в результаті (R+n) – бітовий кадр ділиться без залишку на наперед визначене число і, у випадку відсутності залишку, робить висновок про безпомилкову передачу. Така процедура може здійснюється трьома способами: через арифметичну дію за модулем; дією з поліномами; за допомогою цифрової логіки.

Далі розглянемо механізми захисту від помилок, які з’являються в процесі передавання кадрів. Зокрема, виявлення помилок; позитивне підтвердження для успішно прийнятих, безпомилкових бітів; повторна передача після визначеного часу очікування; від'ємне підтвердження кадрів, в яких були виявлені помилки.

Сукупність таких механізмів утворює автоматичний захист повторної передачі ARQ (Automatic Repeat Request). Стандартизовані версії ARQ передбачають наступні варіанти роботи: 1) запит ARQ із зупинками; 2) запит з поверненням; 3) вибірково-відмовний запит ARQ.

Дуже важливим протоколом управління каналом зв'язку є протокол HDLC (High-level Data Link Control), який підтримує стандарт ISO 3009 [10]. Для того щоб цей протокол міг широко використовуватися, в ньому визначено три типи станцій (головна, підлегла і комбінована); дві конфігурації каналу (несиметрична і симетрична) та три режими передавання даних: 1) нормальний режим, який використовується із несиметричною конфігурацією; 2) асинхронний симетричний режим; 3) режим відгуку, який використовується з несиметричною конфігурацією.

В протоколі HDLC використовується синхронна передача. Вся інформація передається у формі кадрів і є достатньою для передачі всіх типів даних та обміну управляючою інформацією. На рис. 2 подана структура кадру протоколу HDLC. Прапор, адрес і управляючі поля, що додаються до інформаційного поля, називаються заголовком кадра. А поле FSC коду

помилок і прапор, що йде за інформаційним полем, називають контрольною послідовністю кадру. FSC використовується для виявлення помилок і містить залишок від ділення полінома, утвореного вмістом поля даних, на утворюючий поліном CRC-16 і/або CRC- CCITT.

В протоколі HDLC визначено три типи кадрів, кожний з яких має власний формат управляючого поля. Інформаційні кадри переносять дані, які необхідно передати користувачу, а також каскадне підтвердження прийому інформації та захисту від помилок механізмом ARQ. Кадри управління забезпечують реалізацію механізму ARQ, у якого каскадне підтвердження не використовується. Ненумеровані кадри забезпечують додаткові функції управління каналом зв'язку. Перші 1-2 біти управляючого поля визначають тип кадру. Всі формати управляючих полів містять біт запиту-запізнення P/F (pol/final). Функціонування протоколу HDLC полягає в обміні I-кадрами, S-кадрами та U-кадрами між двома станціями. Різні види команд і відгуки для цих типів кадрів перераховані в табл. 1.

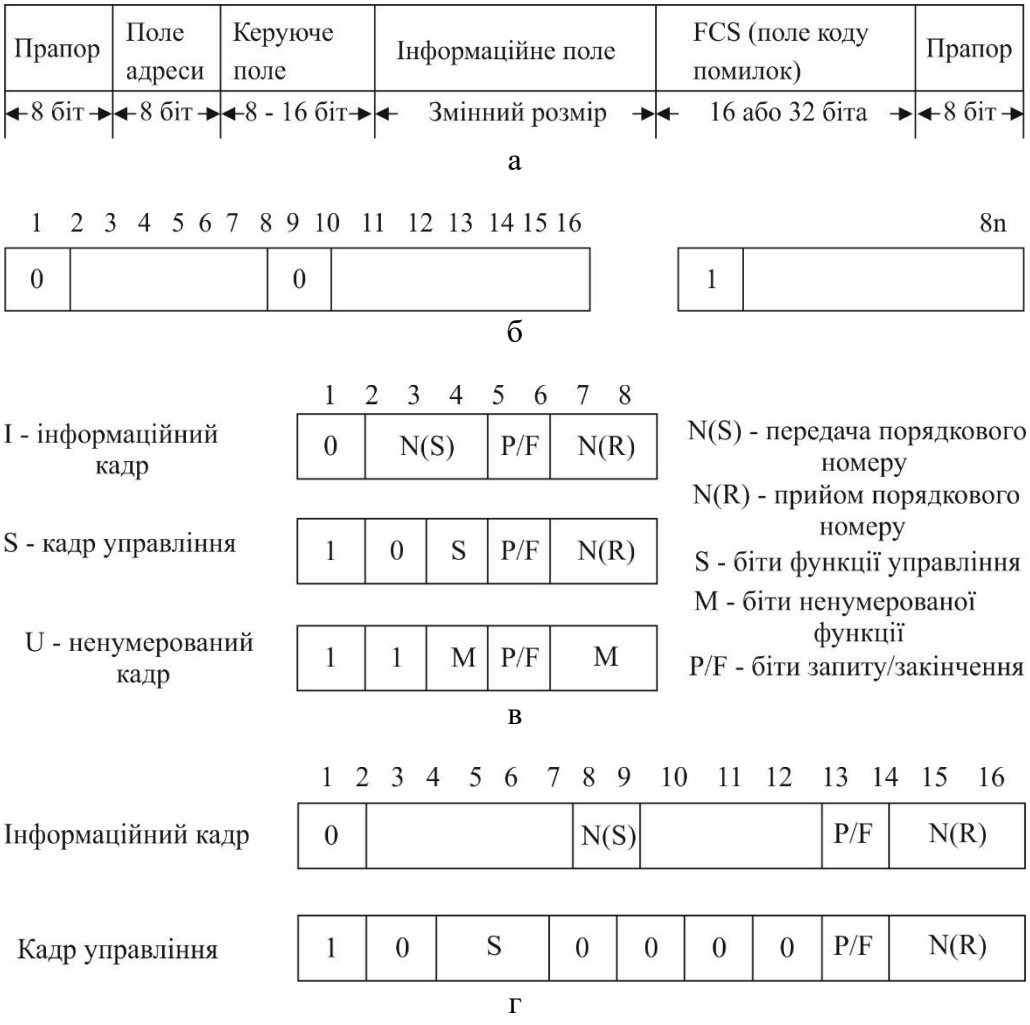


Рис. 2. Структура кадру протоколу HDLC:
а – формат кадру; *б* – розширене поле адреса; *в* – формат управляючого поля із 8 біт;
г – формат управляючого поля із 16 біт

Робота протоколу HDLC проходить в три етапи. Спочатку одна із сторін ініціалізує канал передачі даних, щоб забезпечити впорядкований обмін кадрами. Під час даної фази узгоджуються всі параметри, які будуть використані в процесі передавання. Після ініціаліза-

ції обидві сторони обмінюються даними і управляючою інформацією з метою здійснення управління потоком і захисту від помилок. На завершальному етапі одна із сторін повідомляє про зупинення роботи.

Таблиця 1

Команди і відгуки протоколу HDLC

Номер п/п	Назва	Команда/ відгук	Опис
1	Інформаційний кадр I.	К/В	Обмін даними користувачів.
2	Кадр управління S:		
	- готовий до приймання (RR);	К/В	Додатне підтвердження; готовність до приймання I – кадру.
	- не готовий до приймання (RNR);	К/В	Додатне підтвердження; неготовність до приймання I – кадру.
	- відмова (REJ);	К/В	Від'ємне підтвердження; повернення.
	- вибіркова відмова (SREJ).	К/В	Від'ємне підтвердження; вибіркова відмова.
3	Ненумерований кадр (U):		
	- вибір звичайного / розширеного режиму відгуку(SARM/SARME);	К	Вибір режиму і розширений – семибітові порядкові номера.
	- вибір звичайного / розширеного асинхронного режиму відгуку (SABM/SABME);	К	Вибір режиму і розширений – семибітові порядкові номера.
	- вибір звичайного / розширеного асинхронного збалансованого режиму відгуку (SABM/SABME);	К	Вибір режиму і розширений – семибітові порядкові номера.
	- встановлення режиму ініціалізації (SIM)К	К	Ініціалізація функцій зв'язку на станції – адресат.
	- розрив з'єднання (DISC);	К	Розрив логічного з'єднання.
	- ненумероване підтвердження (UA);	В	Підтвердження однієї з команд вибору режиму.
	- режим розриву з'єднання (DM);	В	Станція – адресат знаходиться в режимі розриву з'єднання.
	- запит на розрив з'єднання RD)	В	Запит на команду DISC.
	- запит на встановлення режиму ініціалізації (RIM);	В	Необхідна ініціалізація; запит на команду SIM.
	- ненумерована інформація (UI);	К/В	Використовується для обміну управляючої інформації.
	- ненумероване очікування (UP);	К	Використовується для запиту управляючої інформації.
	- скидання (RSET);	К	Використовується для відновлення; відновлює значення N(R) і N(S).
	- ідентифікація обміну (XID);	К/В	Використовується запиту/ звіту про стан.
	- тестування (TEST);	К/В	Обмін ідентичними інформаційними полями з метою тестування.
	- відпилення кадра (FRMR).	В	Звіт про отримання неприйнятого кадру.

Окрім протоколу HDLC, існують й інші протоколи управління каналом передавання даних:

- протокол LAPB (Link Access Procedure, Balanced – збалансована процедура доступу до каналу.) Цей протокол є частиною протоколу HDLC, що забезпечує тільки асинхронний симетричний режим (ABM);

- протокол LAPD (Link Access Procedure D-channel – процедура доступу до каналу D), який забезпечує управління каналом зв'язку через D канал, логічний канал інтерфейсу між користувачем і мережею ISDN;

- протокол LLC (Logical Link Control) – управління логічним каналом і відрізняється форматом кадрів з двома рівнями;

- технологія Frame Relay (ретрансляція кадрів) – це засіб управління каналом передачі даних у високошвидкісній мережі з комутацією пакетів;

- технологія ATM (Asynchronous Transfer Mode – асинхронний режим передачі) – засіб передачі даних по високошвидкісній мережі, базується не на протоколі HDLC, а на певному форматі кадра у вигляді комірки, що мінімізує обробку даних.

Для ефективного використання високошвидкісних ліній зв'язку застосовується ущільнення, яке дозволяє декільком передавальним джерелам використовувати більшу пропускну здатність каналу. Може використовуватися дві форми управління: 1) з частотним розділенням каналів; 2) з часовим розділенням каналів.

Управління з частотним розділенням каналів можна використовувати для аналогових сигналів. Кожній із групи сигналів виділяється певна смуга частот, і всі ці сигнали одночасно передаються через одне і теж середовище. Для об'єднання сигналів в необхідну смугу частот використовуються модулюючі пристрої, а для об'єднання отриманих сигналів різної частоти використовують ущільнюючі – мультиплексори. Система з частотним ущільненням подана на рис. 3.

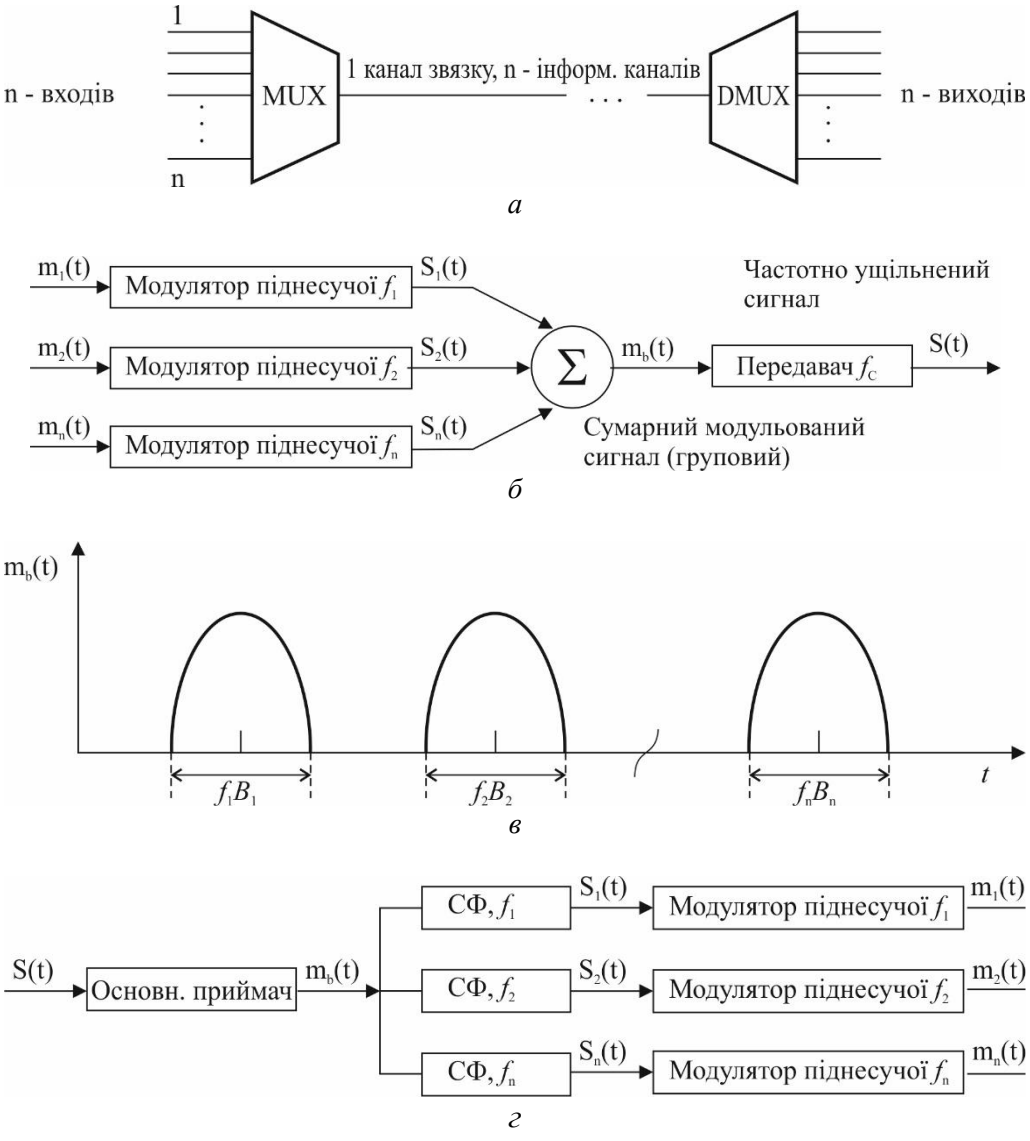


Рис. 3. Система передачі з частотним ущільненням:
а – пристрій ущільнення; *б* – передавач; *в* – спектр групового сигналу; *г* – приймач

Синхронне управління з часовим розділенням використовується з цифровими і з аналоговими сигналами, які переносять цифрові дані. При такому ущільненні дані від різних джерел передаються у вигляді періодично повторюваних кадрів. Кожний такий кадр складається

з набору часових інтервалів, а кожному джерелу виділяється один або декілька інтервалів в кадрі. Результатом такого ущільнення є чергування інформаційних бітів від різних джерел. Чергування може бути побітовим або ж можуть чергуватись цілі блоки. Система синхронного ущільнення з частотним розділенням наведена на рис. 4.

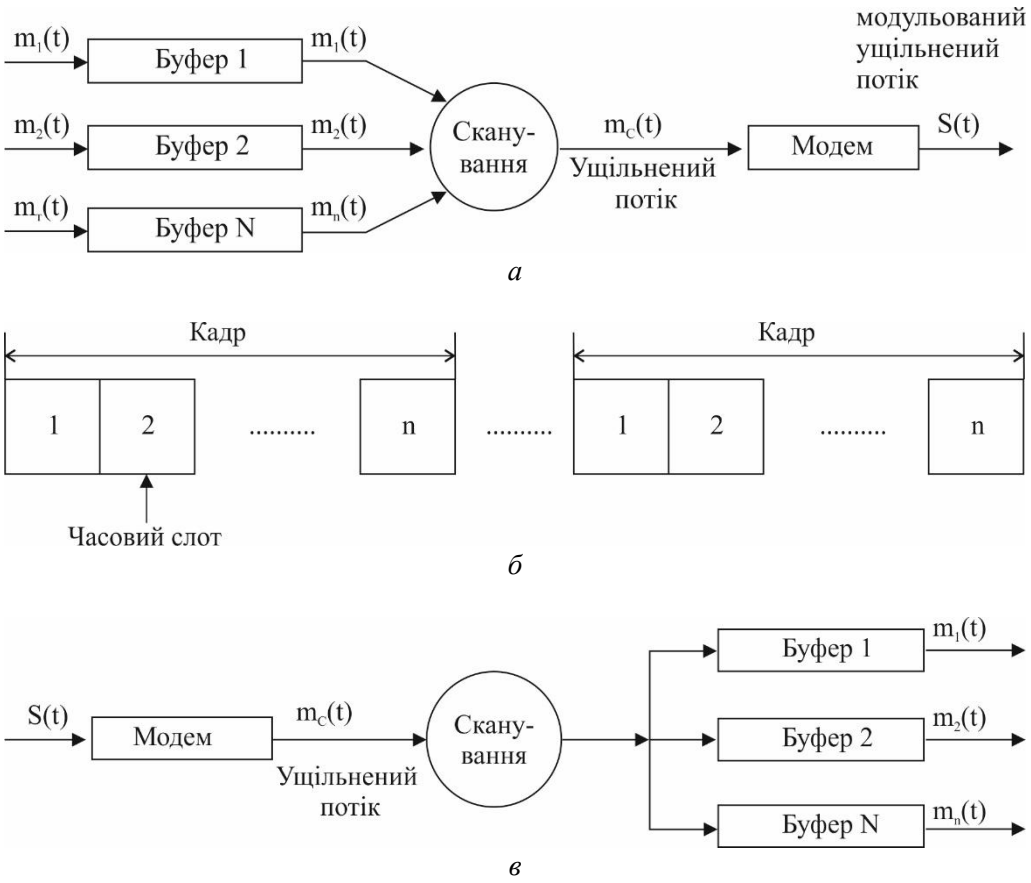


Рис. 4. Система передачі із синхронним часовим ущільненням:
а – передавач; *б* – кадри ТДМ; *в* – приймач

Статистичне ущільнення з часовим розділенням є більш ефективним засобом підтримки термінального обладнання, ніж синхронні TDM. При статистичному TDM розподіл часових інтервалів за джерелами інформації вже не виконується попередньо. Замість цього дані користувача заносяться в буфер і передаються у разі появи доступних часових інтервалів. Якщо статистичний і синхронний мультиплексори використовують канали з однаковими швидкостями передачі даних, то статистичний мультиплексор може забезпечити підтримку більшої кількості пристроїв.

В реалізації високошвидкісних глобальних цифрових мереж більша частина каналу знаходиться між абонентом і мережею, яку називають цифровою абонентською лінією. Сучасна технологія ADSL є однією із модемних технологій, розроблених спеціально для високошвидкісної передачі інформації за звичайними телефонними лініями.

Асиметрична лінія ADSL має високу пропускну здатність в напрямку основного інформаційного потоку, ніж у зворотному напрямку. В ADSL по-іншому застосовується ущільнення із частотним розділенням, що дозволяє підняти пропускну здатність витої пари до 1 МГц. Технологія ADSL включає в себе три елементи, показані на рис. 5:

- резервування самих низьких частот до 25 Гц для передачі мови – POTS (Plain Old Telephone Service – проста телефонна мережа). Мова передається в смузі частот 0 - 4 кГц, а решта смуги використовується для запобігання перехресних завад між каналами;

- використання ехоподавлення або частотного ущільнення для виділення двох смуг;

- використання частотного ущільнення в обох інформаційних смугах. В цьому випадку єдиний потік бітів розчеплюється на паралельні потоки, кожний із яких передається в окремій смузі.

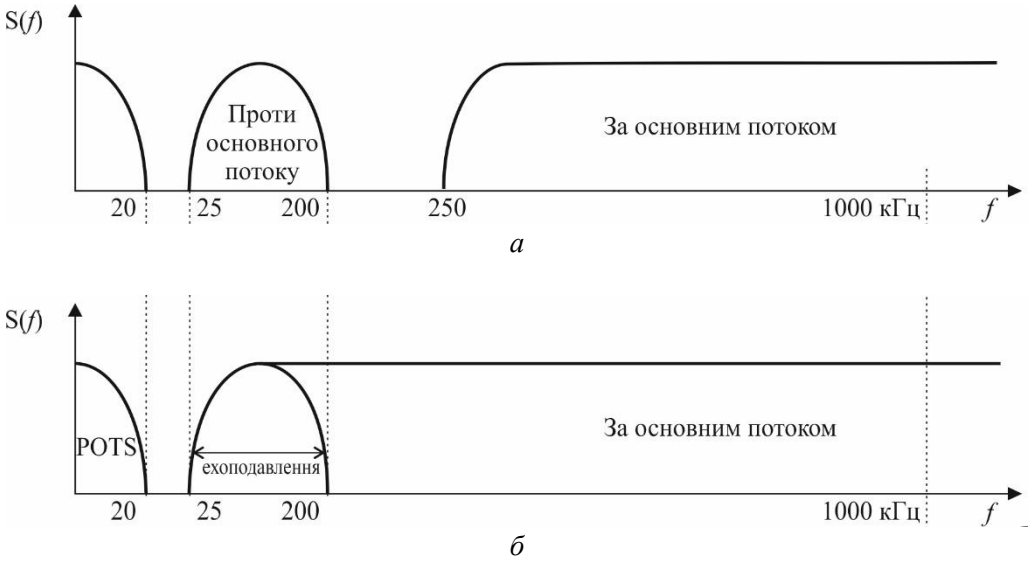


Рис. 5. Конфігурація каналу лінії ADSL:
а – ущільнення з частотним розділенням; *б* – ехоподавлення

Технологія дискретної багатоканальної посилки (DMT – Discrete Multi Tone) використовує декілька несійних сигналів на різних частотах, передаючи деякі біти за ехо-каналами. DMT-технологія передбачає перетворення двійкового послідовного потоку в паралельний (рис. 6).

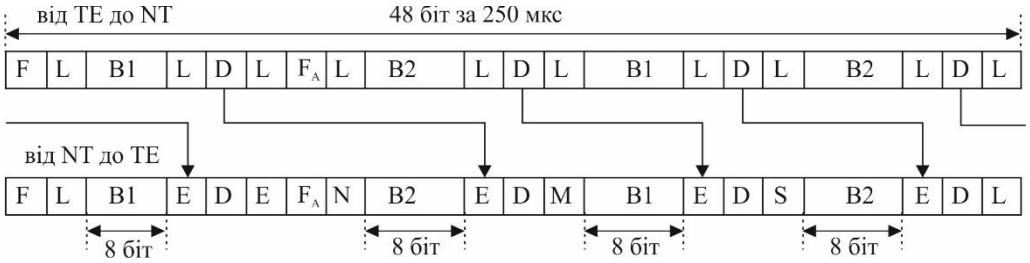


Рис. 6. Структура кадру номінального доступу до ISDN:
F – біт кадрування; L – біт компенсації; E – ехо-біт D – канала; A – біт активації;
FA – допоміжний біт кадрування; N – значення протилежне FA; M – багатокадровий біт;
B1 – біти B-каналу (16 на кадр); B2 – біти B- каналу (16 на кадр);
D – біти D-каналу (4 на кадр); S – резервні біти; E – біт розділення конфлікту

Схеми ADSL/DMT – технології, що використовують 256 каналів передачі за напрямом основного потоку. Теоретично, якщо у кожному чотиригерцовому підканалі буде вестись передача із швидкістю 60 кбіт/с, то загальна швидкість передачі вже може досягти величини 60x256=15,36 Мбіт/с. Міжканальні спотворення і завади дещо знижують таку швидкість, яка складає 9 Мбіт/с.

Схеми високошвидкісної цифрової передачі даних на основі xDSL

Схема xDSL – це одна із схем, що визначає високошвидкісну цифрову передачу по абонентській лінії. В табл. 2 узагальнюються і порівнюються схеми, які мають загальну назву xDSL: ADSL, HDSL (High Date Rate DSL – високошвидкісна цифрова абонентська лінія),

SDSL (Single Line DSL – одно канальна цифрова абонентська ліня), VDSL (Very High Data Rate DSL – супервисокошвидкісна цифрова абонентська лінія).

Таблиця 2

Порівняння параметрів технологій xDSL

Номер п/п	Вид технологій xDSL			
	ADSL	HDSL	SDSL	VDSL
1. R біт/с	1,5 – 9 Мбіт/с (за напрямом потоку) 10 – 640 кбіт/с (проти потоку)	1,544 або 2,048 Мбіт/с	1,544 або 2,048 Мбіт/с	13 – 52 Мбіт/с (за напрямом потоку) 1,5 – 2,3 Мбіт/с (проти потоку)
2. Режим	Асиметричний	Симетричний	Симетричний	Асиметричний
3. Кількість мідних провідників	1	2	1	1
4. Діапазон (довжина) витої пари	3,7 – 5,5 км	3,7 км	3,0 км	1,4 км
5. Передача сигналів	Аналогова	Цифрова	Цифрова	Аналогова
6. Кодіровка каналу зв’язку	CAP/DMT	2B/Q	2B/Q	DMT
7. Частота	1 – 5 МГц	196 кГц	196 кГц	10 МГц
8. Біт/цикл	змінна	4	4	змінна

Технологія xDSL пропонує окремі смуги частот для різних послуг (рис. 5). Цей розподіл виглядає наступним чином: POTS: 0-4 кГц; ISDN: 4-80 кГц; проти основного потоку 300 – 700 кГц; за основним потоком 1МГц і вище.

Технологія ISDN дозволяє ущільнити трафік від кількох пристроїв користувача, підключених однієї лінії, в цифрову мережу з інтеграцією послуг. Можуть використовуватися два інтерфейси: номінальний і базовий.

Структура номінального доступу складається з двох B-каналів на 64 кбіт/с і одного каналу на 16 кбіт/с. Ці канали, які підтримують навантаження в 144 кбіт/с, ущільнюються в інтерфейсі між користувачем і мережею на швидкості 192 кбіт/с. Залишок пропускної спроможності використовується для різних задач синхронізації та керування. Передача при номінальному режимі доступу розбивається на періодично повторювані кадри фіксованого розміру. В даному випадку розмір кадра є рівним 48 біт, і при швидкості передачі 192 кбіт/с кадри повинні повторюватися із швидкістю 1 кадр в 250 мкс. Структура кадру приведена на рис. 6, верхній кадр передається термінальним обладнанням (Terminal Equipment – TE) абонента в мережу (network – NT); нижній кадр передається в зворотному напрямку.

Другий базовий інтерфейс, як і номінальний інтерфейс, працює з ущільненням багатьох каналів в єдиному середовищі передачі. У випадку базового інтерфейсу дозволяється виключно двоточкова конфігурація. Як правило, даний інтерфейс підтримує відомчі або внутрішні телефонні мережі. Для базового інтерфейсу визначені дві швидкості передавання даних 1,544 і 2,048 Мбіт/с. Формат кадру базового доступу до ISDN приведений на рис 7.

Вимірювання параметрів швидкісних систем передачі цифрових даних

Із аналізу видно, що достовірне вимірювання параметрів швидкісних систем передачі цифрових даних, є актуальною задачею.

Для вимірювання параметрів швидкісних цифрових систем пропонується метод комп'ютерного вимірювання, оснований на модульному принципі. На такому принципі побудований, наприклад, універсальний модульний аналізатор мереж доступу типу SunSet MTT [11].

Модульна універсальна портативна платформа SunSet MTT із набором вимірювальних модулів, призначених для: проведення вимірювань в процесі інсталяції систем, вимірювання параметрів, технічного обслуговування та пошуку несправностей в мережах доступу, на

лініях xDSL, волоконно-оптичних систем зв'язку, а також в транспортних цифрових мережах EI, мережах передачі даних, Gigabit Ethernet. Понад тридцять вимірювальних методів можуть задовольнити всі потреби як операторів, так і проведення досліджень у мережах. Всі ці методи встановлюються в базову комп'ютерну вимірювальну систему SunSet MTT, яка забезпечує весь комплекс вимірювань.

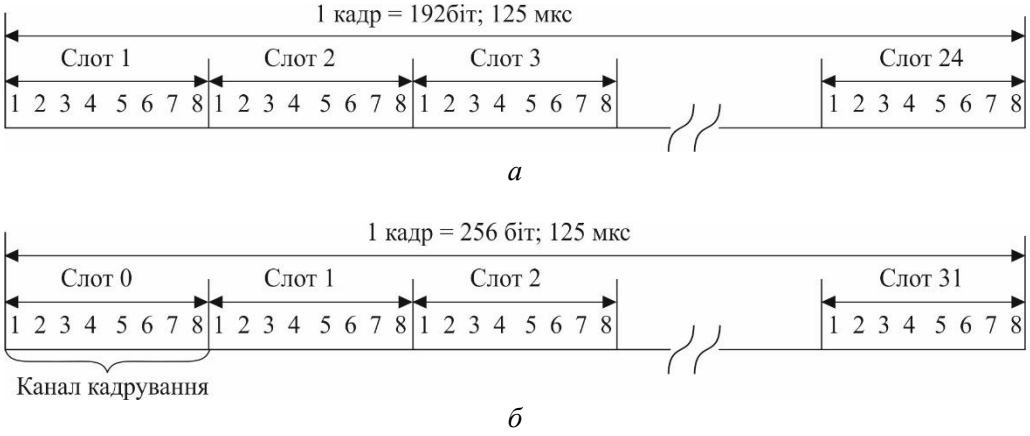


Рис. 7. Структура кадру базового доступу до ISDN:
а – на швидкості 1,54 Мбіт/с; *б* – на швидкості 2,048 Мбіт/с

Дана система забезпечує: тестування середовища передачі цифрової абонентської лінії xDSL: ADSL, GSHDSL, IDSL, SDSL, VDSL; тестування послуг доступу до Інтернету: ATM BERT, IP, IP PING, вимірювання пропускної здатності ATM і IP; тестування регіональних мереж Ethernet, Gigabit Ethernet, Fibre Channel; тестування мереж PDH/SDH; вимірювання параметрів та діагностика в оптичних мережах: оптичний рефлектометр, тестування оптичних каналів, вимірювання оптичних параметрів; тестування мереж доступу: мережі передачі даних, DDS, EI, ISDN, ретрансляція кадрів; вимірювання і оцінка параметрів витої пари, зокрема, аналіз спектральної густини потужності (PSD) шуму, широкосмугового вимірювання рівня, вимірювання асиметрії в кабелі, відношення сигнал/шум, перехідних завад, режим цифрового мультиметра (DMM), рефлектометра (TDR).

Висновки

Проаналізовано технології управління каналом передачі цифрових даних з ущільненням з використанням швидкісних ліній передачі типу xDSL та приведена методологія вимірювання їх параметрів, що базується на модульному принципі з застосуванням універсальних модульних аналізаторів мереж. Показано, що на основі технології xDSL можна будувати високошвидкісні мережі передачі даних, перевагою яких є використання наявної інфраструктури та можливість роботи від резервного живлення що дозволяє будувати мережі підвищеної надійності. Все це робить xDSL технологію актуальною, особливо в якості резервного каналу доступу до Інтернету.

Список літератури:

1. L. Gilli, G. Cossu, E. Ciaramella. New Prospects of Optical Wireless Communication Systems Exploiting VCSEL-Based Transmitters // Journal of Lightwave Technology. 2025. Vol.43, no.4. P. 1615–1624. https://doi.org/10.1109/JLT.2025.3530819

2. Ahmed Al-Kinani, Cheng-Xiang Wang, Li Zhou, Wensheng Zhang Optical Wireless Communication Channel Measurements and Models // IEEE Communications Surveys & Tutorials. 2018. Vol. 20, I. 3. P. 1939–1962. https://doi.org/10.1109/COMST.2018.2838096

3. Безрук В.М., Бідний Ю.М., Колтун Ю.М., Астраханцев А.А., Свид І.В., Ширяєв А.В., Харченко Н.А. Інформаційні мережі зв'язку. Ч. 2. Телекомунікаційні технології стаціонарних мереж зв'язку: навч. посіб. Харків : ХНУРЕ, 2011. 492 с.

4. Ситников В. Д., Колесник І. С., Черняк О. І. Комп'ютерна мережа із застосуванням оптичної лінії зв’язку із автономним живленням // Матеріали LIII наук.-техн. конф. підрозділів ВНТУ, Вінниця, 20–22 березня 2024 р. https://conferences.vntu.edu.ua/index.php/all-fitki/all-fitki-2024/paper/view/19714.

5. Romanov O. I., Svyd I. V., Korniienko N. I., Romanov A. O. Optical Network Management by ONOS-Based SDN Controller // Radiotekhnika. № 210. P. 188–196, верес. 2022. https://doi.org/10.30837/rt.2022.3.210.16

6. Singh S.V., Khursheed A., Alam Z. Wired Communication Technologies and Networks for Smart Grid-A Review // Studies in Computational Intelligence. 2022. Vol 1007. P. 183–195. https://doi.org/10.1007/978-981-16-8012-0_15

7. Esra Söğüt, Saadin Oyucu, O. Ayhan Erdem, Hüseyin POLAT Recommendations for xDSL Technologies and Applications // ICENS International Conference on Engineering and Natural Science, 3-7 May 2017, Budapest, Hungary https://www.researchgate.net/publication/333561999

8. Octávio Pereira, Paulo Neves, Ricardo Mostardinha, Eduardo Valente, Vasco Soares, Paulo Alves, Fernando Silva, Hélder Couteiro. A Network Project Case Study Leveraging xDSL Technology // IADIS International Conference Applied Computing 2007. P. 781–783.

9. О.І. Чумаченко, В.В. Цілицький, М.О. Білий Побудова інформаційної системи підвищеної надійності одного класу // Електроніка та системи управління. 2011. №4(30). С. 127–134.

10. ISO 3009:2003. ISO. https://www.iso.org/standard/37415.html

11. Методика вимірювань параметрів телекомунікаційних мереж. Офіційний вебпортал парламенту України. https://zakon.rada.gov.ua/laws/show/z0582-21

Надійшла до редколегії 06.06.2025

Відомості про авторів:

Довгий Віктор Володимирович – канд. техн. наук, Карпатський національний університет імені Василя Стефаника, старший викладач кафедри комп’ютерної інженерії та електроніки, Україна; e-mail: viktor.dovhyi@pnu.edu.ua, ORCID: https://orcid.org/0009-0009-7158-6938

Грига Володимир Михайлович – канд. техн. наук, доцент, Карпатський національний університет імені Василя Стефаника, доцент кафедри комп’ютерної інженерії та електроніки, Україна; e-mail: volodymyr.gryga@pnu.edu.ua, ORCID: https://orcid.org/0000-0001-5458-525X

Дзундза Богдан Степанович – д-р техн. наук, с.н.с., Карпатський національний університет імені Василя Стефаника, професор кафедри комп’ютерної інженерії та електроніки, Україна; e-mail: bohdan.dzundza@pnu.edu.ua, ORCID: https://orcid.org/0000-0002-6657-5347

Свид Ірина Вікторівна – канд. техн. наук, доцент, Карпатський національний університет імені Василя Стефаника, професор кафедри комп’ютерної інженерії та електроніки; Харківський національний університет Повітряних Сил імені Івана Кожедуба, доцент кафедри авіаційних радіотехнічних систем навігації та посадки, Україна; e-mail: iryna.svyd@pnu.edu.ua, ORCID: https://orcid.org/0000-0002-4635-6542

Терлецький Андрій Іванович – канд. фіз-мат. наук, доцент, Карпатський національний університет імені Василя Стефаника, доцент кафедри комп’ютерної інженерії та електроніки, Україна; e-mail: andrii.terletskyi@pnu.edu.ua, ORCID: https://orcid.org/0000-0002-1667-3467

Павлюк Мирослав Федорович – канд. фіз-мат. наук, доцент, Карпатський національний університет імені Василя Стефаника, доцент кафедри комп’ютерної інженерії та електроніки, Україна, e-mail: myroslav.pavlyuk@pnu.edu.ua, ORCID: https://orcid.org/0000-0002-5663-2918

УДК 621.372(075) **DOI:10.30837/rt.2025.3.222.16**

Д.Г. ФОКІН, М.О. ЄВДОКИМЕНКО, д-р техн. наук

АНАЛІЗ МЕТОДІВ ПРОТОКОЛЬНОЇ СТЕГАНОГРАФІЇ В ПРОГРАМНО-КОНФІГУРОВАНИХ МЕРЕЖАХ

Вступ

Приховані канали зв’язку, що забезпечують передачу даних шляхом маскування їх під легітимний мережний трафік, становлять серйозну загрозу для цілісності, конфіденційності та контролю інформаційних систем. На відміну від традиційних методів витоку інформації, ці канали функціонують без потреби в ескалації привілеїв або ініціації користувачем підозрілої активності, оскільки ґрунтуються на легітимній поведінці мережних протоколів або експлуатації часових характеристик передачі даних.

З огляду на стрімке впровадження програмно-конфігурованих мереж (Software-Defined Network, SDN), проблема прихованих каналів набуває нових вимірів. Характерні ознаки SDN – централізоване керування, програмованість логіки маршрутизації та чітке розділення площини управління й площини передачі – створюють не лише додаткові вектори загроз, але й потенційні засоби для протидії. З одного боку, гнучкість SDN-архітектури може бути використана зловмисниками для побудови стійких і малопомітних прихованих каналів, зокрема шляхом маніпуляції заголовками протоколів, міжпротокольної взаємодії або зміни порядку пакетів. З іншого боку, глобальна видимість трафіку, притаманна SDN-контролерам, відкриває перспективи для впровадження ефективних механізмів централізованого моніторингу, виявлення аномалій і адаптивного реагування.

Мережна стеганографія в контексті SDN розглядається як потужний інструмент для реалізації прихованого управління зловмисним програмним забезпеченням, ексфільтрації конфіденційних даних, обходу систем контролю трафіку та здійснення координації дій у межах складних атак. Її застосування дозволяє підтримувати зв'язок із скомпрометованими вузлами навіть у сегментованих або ізольованих середовищах, мінімізуючи ризик виявлення завдяки використанню дозволених протоколів і портів.

Актуальність дослідження методів протокольної стеганографії в SDN-середовищах визначається як еволюцією атакуючих технік, що адаптуються до сучасних архітектур, так і обмеженнями традиційних засобів стегоаналізу, які здебільшого орієнтовані на класичні мережеві моделі. Відсутність комплексних підходів до виявлення багатошарових і динамічних прихованих каналів у SDN залишається суттєвою прогалиною у сфері інформаційної безпеки.

У роботі проведено систематизацію сучасних підходів до реалізації протокольної стеганографії в SDN, а також здійснено порівняльний аналіз методів їх виявлення й нейтралізації з урахуванням особливостей архітектури. Такий підхід дозволяє не лише узагальнити наявні дослідження, а й виокремити вразливі елементи інфраструктури, що потребують посиленої уваги при проєктуванні захищених мереж на базі SDN.

Аналіз методів стеганографії в програмно-конфігурованих мережах

Традиційно приховані канали класифікуються на основі принципу їхньої роботи – як канали на основі зберігання (storage-based channels) та канали на основі синхронізації (timing-based channels). Такий поділ відображає механізм передачі стеганограм між відправником і отримувачем: перші маніпулюють значеннями полів протоколів, тоді як інші використовують часові характеристики трафіку, тобто замість того, щоб вбудувати дані в заголовки пакетів або корисне навантаження, кодують інформацію за допомогою затримки пакетів або швидкості передачі пакетів. Проте з погляду практичної реалізації в мережах доцільнішим є підхід, що враховує рівень мережної моделі, для якого розроблено конкретний метод.

У цьому огляді класифікацію методів стеганографії здійснено відповідно до мережних рівнів, у межах яких вони реалізуються (канальний, мережний, транспортний, тощо).

Між непомітністю та пропускною здатністю прихованого каналу, як правило, існує обернена залежність: високопродуктивні канали частіше генерують аномальний трафік, що підвищує ймовірність виявлення, тоді як малошвидкісні канали забезпечують вищий рівень маскування у фоновому трафіку. У середовищі програмно-конфігурованих мереж (SDN) захисники можуть використовувати глобальне бачення мережі, яким володіє контролер, для ідентифікації нетипових шаблонів у поведінці трафіку. Наприклад, OpenFlow-додатки здатні виявляти аномальні значення Flow Label або присутність ненульових зарезервованих полів, які зазвичай не використовуються у легітимному трафіку.

Водночас централізований характер SDN створює нові можливості і для зловмисників. Завдяки централізованому контролю можливо координувати складні сценарії прихованого зв'язку які дозволяють обходити обмеження традиційного фільтрування та сегментації мережі.

Особливий інтерес становить міжпротокольна стеганографія (inter-protocol steganography) [1], що використовує взаємодію між кількома протоколами для організації стійких і важковиявних каналів. У таких підходах одночасно можуть застосовуватись різні типи каналів – наприклад, канал на основі синхронізації у періоди низького навантаження на мережу, і канал на основі заголовків у моменти інтенсивного трафіку, що забезпечує ефективну адаптацію до змін умов. Централізоване управління в SDN може слугувати як інструментом для виявлення подібної активності, так і, у випадку компрометації контролера, засобом для реалізації високорівневої координації між прихованими компонентами атак.

З огляду на динамічну природу сучасних мереж і широкі можливості SDN, дослідження методів стеганографії повинні враховувати не лише окремі протоколи, а й їхню взаємодію, змінність поведінки трафіку, а також потенціал як для приховування, так і для виявлення таких каналів у масштабах всієї мережної інфраструктури.

Зважаючи на гетерогенність мережевих середовищ та варіативність підходів до прихованої передачі даних, доцільним є розгляд методів стеганографії в контексті окремих рівнів моделі OSI. Такий підхід дозволяє точніше визначити вразливості, пов'язані з кожним рівнем, оцінити вплив прихованого каналу на легітимний трафік, а також підібрати релевантні методи детектування та протидії. Нижче подано структурований огляд методів стеганографії, згрупованих за рівнями OSI-моделі, що дозволяє краще зрозуміти їхню специфіку, можливості застосування та ризики для безпеки SDN-інфраструктур.

Фізичний та канальний рівні

На фізичному та канальному рівнях моделі OSI стеганографія переважно використовує можливості протоколів локальних мереж, зокрема Ethernet та IEEE 802.11. Такі методи є особливо актуальними в локальних сегментах інфраструктур SDN, які здебільшого функціонують на базі Ethernet-комутаторів.

У роботі [1] описано метод PadSteg – міжпротокольний стеганографічний підхід, що використовує поле доповнення в Ethernet-кадрах для прихованої передачі інформації. За стандартом Ethernet кожен кадр повинен мати щонайменше 64 байти, з яких 46 – це мінімальний розмір поля даних. Якщо передаються менші обсяги даних, кадр автоматично доповнюється нульовими байтами. Однак уразливість під назвою Etherleak полягає в тому, що не всі реалізації драйверів коректно обнуляють ці додаткові байти – вони можуть містити залишкову пам'ять. PadSteg цілеспрямовано створює короткі Ethernet-кадри, у доповнення яких вбудовуються приховані дані, при цьому використовується протокол ARP разом з транспортними протоколами, як-от TCP або ICMP. У найкращому випадку метод забезпечує до 45 байт на кадр, однак для збереження непомітності використовується лише природне доповнення. В умовах реального навантаження пропускна здатність методу становить у середньому 32 біт/с, що є досить ефективним показником для стеганографії.

У статті [2] представлено прихований канал на основі зберігання, який не передбачає безпосереднього трафіку між учасниками. Один з вузлів мережі використовується як так званий "Dead Drop" – проміжний носій повідомлення. Відправник вбудовує дані в ARP-відповіді, які потрапляють до кешу, адрес цього вузла, а отримувач зчитує їх за допомогою SNMP. Через відсутність прямої комунікації канал дуже важко виявити, однак його пропускна здатність вкрай низька – у межах кількох бітів на хвилину.

В роботі [3] описано реалізацію внутрішньосмугового (in-band) прихованого каналу шляхом маніпуляції протоколом LLDP, який використовується в SDN для визначення топології мережі. Зловмисники підробляють LLDP-пакети, змушуючи SDN-контролер помилково вважати, що два віддалені OpenFlow-комутатори безпосередньо з'єднані. У такий спосіб створюється логічний "тунель" між вузлами, який використовується для передавання прихованих повідомлень через площину управління. Пропускна здатність такого каналу обмежена – зазвичай лише кілька байтів за один обмін LLDP, однак його цінність полягає у здатності функціонувати поза площиною даних.

У статті [4] описано метод StegoFrameOrder, який реалізує прихований канал на основі синхронізації у безпроводових мережах Wi-Fi. Він ґрунтується на маніпуляції порядком надходження кадрів у мережі з доступом за схемою CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance). Стеганограма кодується через черговість передачі кадрів між різними вузлами: наприклад, першим передає "0", другим – "1". Перевага методу полягає у відсутності змін до самих кадрів; використовується лише часовий аспект. Середня пропускна здатність становить до 10 біт/с у типовому середовищі IEEE 802.11.

Іншим важливим прикладом є метод HICCUPS (Hidden Communication System for Corrupted Networks), описаний у [5]. Він реалізує прихований канал шляхом наміру пошкоджувати CRC у кадрах. У звичайних умовах такі кадри відкидаються, однак у HICCUPS попередньо налаштовані вузли вважають їх сигналом прихованого повідомлення, інтерпретуючи тип помилки як біт "0" або "1". Оскільки в безпроводових мережах кадри часто зазнають ушкоджень через перешкоди, невелике зростання частоти "помилок" може бути непоміченим. Проте через високу ймовірність втрати та необхідність обмеження кількості навмисно пошкоджених кадрів, пропускна здатність становить лише кілька біт/с.

Усі згадані методи демонструють, що навіть на нижчих рівнях OSI-моделі – фізичному та канальному – існує широкий спектр технік прихованої передачі даних, які можуть бути реалізовані в умовах SDN. Особливої уваги потребують протоколи Ethernet, ARP, LLDP, а також механізми доступу до середовища у Wi-Fi, що можуть бути використані як носії прихованих повідомлень без зміни змісту легітимного трафіку.

Мережний рівень

Методи мережного рівня стеганографії приховують дані у полях заголовків протоколів IP, зокрема IPv4 та IPv6, або використовують службові протоколи цього рівня, як-от ICMP. Ці методи особливо актуальні в площині даних SDN, зокрема при передачі між маршрутизаторами та комутаторами, що оперують IP-трафіком.

У класичній роботі [6] запропоновано один із перших відомих методів стеганографії в стеку TCP/IP, що використовує 16-бітне поле Identification в заголовку IPv4. Це поле, призначене для фрагментації, може залишатися незмінним або містити довільне значення, якщо пакет не фрагментується. Таким чином, зловмисник може вбудовувати повідомлення у вигляді бітів у кожен пакет. Пропускна здатність методу теоретично досягає 1600 біт/с за умов високої інтенсивності трафіку. Простота реалізації та висока продуктивність роблять його актуальним і сьогодні, хоча відстеження повторюваних значень поля може дати змогу виявити канал за допомогою систем виявлення вторгнень.

Інша класична праця [7] описує використання зарезервованих або необов'язкових бітів у протоколах IPv4. Наприклад, стеганографічні біти можуть бути закодовані у зарезервованому прапорці або у полі ToS/DSCP, яке часто не використовується в типових налаштуваннях

QoS. Пропускна здатність тут становить 1–3 біт на пакет, але такі канали мають високу непомітність, якщо політики QoS не змінюють значення полів при транзиті.

Опції IPv4, такі як Record Route та Timestamp, також демонструють потенціал для прихованої передачі інформації. У роботі [8] автори кодують байти у фальсифікованих IP-адресах маршруту, а в [9] – у полі переповнення опції Timestamp. Обидва методи мають обмеження у вигляді максимального розміру опцій IPv4 (до 40 байт), але забезпечують десятки або навіть сотні біт/пакет при правильній реалізації.

ICMP-повідомлення, зокрема Echo-запити (ping), давно використовуються як транспортний механізм для стеганографії. Утиліта Loki (1996) і сучасні приклади, як Pingback (2021) [10], демонструють вбудовування повідомлень у поля даних, Identifier або Sequence. Частота ping-пакетів визначає пропускну здатність, яка може досягати 800 біт/с при помірному трафіку.

Для IPv6 характерні нові поля, що відкривають додаткові можливості. 8-бітне поле Traffic Class є аналогом ToS/DSCP в IPv4 і може кодувати біти, якщо мережа не застосовує активні політики QoS. Поле Flow Label (20 біт), як зазначено у [11], прямо визначене як потенційне джерело прихованих каналів. У роботі [12] показано, що за умови достатнього обсягу трафіку і належної рандомізації значень, можна досягти тисяч біт/с без очевидного порушення легітимності трафіку.

Також у [12] розглядається використання опцій-доповнень в розширених заголовках IPv6. Пропускна здатність таких каналів теоретично сягає 2048 біт/пакет, хоча на практиці застосовуються значно менші обсяги, щоби уникнути виявлення засобами глибокого аналізу пакетів.

У роботі [13] описано стеганографічний метод на основі варіативності IP-адреси. Змінюючи хост-частину IPv6-адреси, можна кодувати дані, які зчитуються за допомогою аналізу відправлених адрес.

Крім цього, на мережному рівні можливе створення прихованих каналів на основі синхронізації. Наприклад, у [14] Він модулює секретні дані в трафік VoNR (Voice over New Radio), змінюючи міжпакетні затримки. Ці модифікації ретельно розроблені, щоб імітувати статистичну поведінку (зокрема, кумулятивну функцію розподілу) природних тремтінь IPD у реальному трафіку VoNR.

Сукупно методи мережного рівня охоплюють широкий спектр векторів стеганографії – від модифікації зарезервованих бітів до складних схем зміни IP-адрес і затримок між пакетами. У контексті SDN їхня ефективність або вразливість визначається архітектурною здатністю до централізованого моніторингу, обробки потоків та динамічної зміни політик обробки трафіку.

Транспортний рівень

Методи транспортного рівня зосереджуються на протоколах TCP, UDP, SCTP та інших. Стеганографія на цьому рівні реалізується шляхом модифікації полів заголовків транспортних протоколів або шляхом маніпулювання характерними особливостями їх функціонування. У SDN, зокрема в площині управління, протокол OpenFlow функціонує поверх TCP, а трафік площини даних зазвичай реалізується через TCP або UDP. Таким чином, приховані канали цього рівня найчастіше проявляються у трафіку площини даних.

У класичній роботі [6] запропоновано використання поля Initial Sequence Number (ISN) в TCP як носія прихованого повідомлення. При встановленні TCP-з’єднання обидві сторони обмінюються 32-бітовими випадковими початковими номерами. Прихований відправник може ініціювати TCP-сеанс, задавши ISN таким чином, щоб він кодував певну послідовність бітів. Одержувач, який контролює трафік SYN (synchronize), вилучає значення ISN і реконструює приховане повідомлення. Пропускна здатність методу становить до 16 біт на з’єднання, що за умови одного з’єднання на секунду дає лише 16 біт/с. Цей підхід ефективний за умови великої кількості короткоживучих з’єднань.

У тій же роботі розглянуто використання поля Acknowledgement Number TCP. Зловмисник може кодувати дані в ACK-пакетах, змінюючи 32-бітове значення підтвердження відповідно до бітової послідовності повідомлення. Такий метод вимагає контролю над формуванням трафіку та скоординованості між відправником і одержувачем, і найкраще підходить для штучно створених з'єднань. Теоретично, кожен пакет може нести до 32 біт прихованої інформації, але на практиці це може викликати підозру при перехресній перевірці з очікуваними значеннями підтвердження.

Методи, що використовують опції TCP, також заслуговують на увагу. Наприклад, у 32-бітному полі опції Timestamp можливо кодувати дані в молодших бітах значення часової мітки. Інший підхід – використання опції No-Operation (NOP), яка зазвичай застосовується для вирівнювання, як прихованого бітового сигналу. У роботі [15] продемонстровано, що комбінація змін у кількох полях TCP та IP, включаючи опції, може забезпечити приховану передачу до 31 біта на пакет. Однак це граничний випадок, що значно ускладнює непомітність каналу.

У статті [16] запропоновано метод приховання даних у змінюваній довжині UDP-пакетів. Оскільки довжина UDP-пакету визначається як сума 8 байт заголовка і довжини навантаження, контрольований відправник може кодувати біти, трохи змінюючи довжину пакета (на 1 байт для представлення "0" або "1"). В експериментальних умовах досягнуто пропускної здатності до 456 біт/с. Цей метод залишається малопомітним, якщо варіації довжини інтегруються у природний розподіл розмірів пакетів і використовуються з випадковими інтервалами.

Протокол SCTP, хоча й не широко розповсюджений, має низку унікальних характеристик, що роблять його перспективним для стеганографії. У роботі [17] розглянуто використання полів Initiate Tag, Payload Protocol Identifier та Heartbeat Info для прихованої передачі. Крім того, можливості multistreaming (багатопотоковості), multihoming (підтримки кількох маршрутів) та часткової надійності в SCTP дозволяють створювати складні сценарії кодування, включно з використанням повторних передач або зміни маршруту як сигналу. У випадку SDN ці можливості є більш теоретичними, оскільки більшість середовищ не використовує SCTP, і будь-яка його поява може одразу викликати підозру, за винятком телекомунікаційних систем, де протокол є стандартом.

Таким чином, методи стеганографії транспортного рівня охоплюють як внутрішньопротокольні канали в TCP, так і канали на основі особливостей UDP або специфічних властивостей SCTP. Їх пропускна здатність значно варіюється залежно від протоколу та методики, але всі вони мають спільну ознаку – вони потребують ретельного налаштування для уникнення виявлення, зокрема у контрольованих середовищах SDN, де доступний централізований моніторинг трафіку.

Сеансовий рівень

Сеансовий рівень відповідає за встановлення, підтримку та контроль сеансів між кінцевими точками. На практиці багато функцій сеансового рівня обробляються протоколами вищого рівня, але деякі протоколи та механізми можна розглядати як такі, що працюють на даному рівні. У цьому контексті інформація може бути прихована у протоколах ініціації сеансу та сигнальних повідомленнях, а також у маніпулюванні поведінкою сеансу.

Session Initiation Protocol (SIP) і супровідний Session Description Protocol (SDP) є класичними прикладами протоколів сеансового рівня, які можуть передавати приховані дані. У статті [18] продемонстрували, що SIP-повідомлення, які використовуються для встановлення VoIP-дзвінків, можуть бути використані для створення прихованих каналів. Вони показали, що багато полів, маркерів і параметрів у SIP/SDP мають гнучкі або невизначені формати, які можна використовувати для вбудовування інформації. Наприклад, необов'язкові маркери, як-от ідентифікатори SIP «tag» і «branch», які мають бути випадковими; поля заголовка, які не мають суворих обмежень по довжині, можуть бути розширені. Такі

методи забезпечують пропускну здатність каналу від низької до помірної – в експериментах від кількох десятків до кількох сотень біт на секунду.

Окрім конкретних протоколів, сеансовий рівень охоплює керування з'єднаннями між кінцевими точками. Приховані канали можна побудувати, використовуючи початок, синхронізацію та завершення сеансу. Наприклад, шляхом навмисного ініціювання та розривання з'єднань у шаблоні (або введення певних затримок між подіями сеансу) можна передати повідомлення. В статті [19] показали, що навіть без будь-якого контролю над контролером або комутаторами SDN хости можуть створити прихований канал синхронізації, ретельно організовуючи пакетні спалахи (packet bursts), які запускають певну поведінку контролера. По суті, сама взаємодія SDN з площиною керування стає частиною прихованого каналу – наприклад, час, потрібний контролеру для встановлення правила потоку або відповіді на нове з'єднання, може кодувати біти. Ця техніка використовувала те що SDN контролер бачить нові запити на сеанс, щоб створити канал зв'язку, невидимий для звичайного моніторінгу трафіка.

Хоча методи стеганографії на сеансовому рівні є порівняно менш дослідженими у порівнянні з мережним або транспортним рівнями, існуючі дослідження формують концептуальну основу того, що будь-який протокол, який керує сеансом або діалогом може бути використаний для передачі інформації. Таким чином, сеансовий рівень пропонує приховані канали через метадані та динаміку сеансів – від метаданих викликів VoIP до синхронізації налаштувань потоку SDN.

Рівень представлення

На цьому рівні відбувається шифрування та перетворення форматів даних (наприклад, TLS забезпечує конфіденційність і цілісність переданої інформації). У статті [20] автори проаналізували сім методів прихованого каналу в TLS, зокрема приховану передачу через довжину запису, вектор ініціалізації та поле типу вмісту. Експерименти підтвердили ефективність цих підходів: найвищу пропускну здатність показав метод маніпулювання вектором ініціалізації. Водночас стандартні системи виявлення вторгнень (Snort, Bro/Zeek, Suricata) не здатні «з коробки» розпізнати такі TLS-канали, що свідчить про їхню непомітність [20].

Робота [21] демонструють розширення стеганографії рівня представлення на сучасні протоколи. Так, у протоколі QUIC (що лежить в основі HTTP/3) зловмисники можуть використовувати його поля для укриття даних. Зокрема, показано можливість використання необов'язкового біта spin bit у заголовку QUIC для сигналізації прихованих біт між клієнтом і сервером. Хоча пропускна здатність такого каналу невелика (передається лише 1 біт інформації на кожен пакет, що дає десятки біт/с), його трафік майже не відрізнити від легітимного HTTP/3-обміну, що ускладнює виявлення. Інший підхід використовує механізм міграції з'єднання в QUIC: функцію Server Preferred Address можна експлуатувати для прихованої передачі даних під виглядом перемикання сервером адреси підключення. Дослідники реалізували такий канал (інструмент QUIC-Exfil) і показали, що сучасні міжмережеві екрани не здатні відрізнити фальшиву міграцію від легітимної. Перевага методів на основі QUIC – використання зашифрованих заголовків і динаміки протоколу, що робить їх малопомітними для засобів DPI. Недоліком є складність реалізації та залежність від підтримки протоколом певних опцій (наприклад, наявності spin bit) і можливість контрзаходів у майбутніх версіях протоколів.

Прикладний рівень та площина управління SDN

Починаючи з прикладного рівня моделі OSI відкриваються широкі можливості для реалізації прихованих каналів через зловживання функціональністю високорівневих протоколів. У середовищі SDN до цього рівня також відносяться протоколи управління мережею, зокрема OpenFlow. Стеганографічні методи на прикладному рівні охоплюють як маніпуляції структурованими протоколами на кшталт HTTP або DNS, так і специфічні для SDN підходи, пов'язані з контролером і комутаторами.

В роботі [22] описано вразливість у процедурі встановлення з'єднання між OpenFlow-комутатором та контролером, яка дозволяє створити прихований канал у повідомленні Features Reply. Зокрема, Datapath ID (унікальний 64-бітовий ідентифікатор пристрою) або інші поля відповіді, як-от MAC-адреси портів чи імена інтерфейсів, можуть бути використані для кодування повідомлень. Оскільки контролер OpenFlow за замовчуванням довіряє отриманим значенням без обов'язкової перевірки їх достовірності, ці поля можуть бути використані зловмисником-комутатором для прихованої передачі інформації контролеру. Пропускна здатність обмежується обсягом інформації, що вміщується в одному Features Reply (кілька десятків байтів), але цього достатньо для передачі конфіденційних параметрів або тригерів.

Дослідження [23] представляє прихований канал синхронізації, реалізований через архітектуру SDN. Автори демонструють, як два ізольовані OpenFlow-комутатори можуть обмінюватися даними через централізований контролер без прямого канального з'єднання. Зловмисник модифікує поведінку одного з комутаторів таким чином, щоб викликати у контролера певну відповідь (наприклад, переналаштування маршруту або відправлення Barrier Request). Інший комутатор, що спостерігає за інтенсивністю або часом реакції контролера, розшифровує приховані біти. У прототипі, реалізованому за допомогою операційної системи для SDN-контролера Open vSwitch та ONOS (Open Network Operating System), досягнуто пропускної здатності до 20 біт/с. Цей канал є типовим каналом синхронізації, де не вміст, а порядок і час повідомлень використовуються як носії.

Ще один значний напрямок прикладного рівня – HTTP та DNS-стеганографія. Протокол HTTP, широко використовуваний у інтерфейсі прикладного рівня SDN-контролерів Northbound API, може бути модифікований для прихованої передачі інформації. В роботі [24] розглянуто приховування даних в полях заголовків, використовуючи допустимі але непомітні модифікації. Наприклад, варіювання пропусків (табуляція/пробіл) у заголовку може кодувати біти 0/1, так само як і зміна регістру літер. У [25] розглядаються методи кодування секретних бітів шляхом вставлення даних у параметри URL та інші поля HTTP-запитів. Крім того, запропоновано метод прихованої передачі інформації через контрольовані часові інтервали між HTTP-запитами, що утворює канал синхронізації.

Щодо DNS, в [26] продемонстровано кодування повідомлення в іменах доменів, що запитуються. Секретні дані розбиваються на фрагменти та вставляються у вигляді довгих піддоменів при DNS-запитах. Також автори розглядають вбудовування інформації в поля ресурсних записів – наприклад, TXT-записи здатні нести до ~64 КБ текстових даних, що дозволяє передавати великі повідомлення у відповідях DNS. Також можуть застосовуватися поля адресних записів. Поле Time-To-Live у DNS-відповіді, що зазвичай визначає час кешування, теж може слугувати каналом. Прихований відправник може варіювати значення TTL у відповіді за певною схемою, щоб закодувати біти [27]. В статті [28] автори розглядають канал синхронізації. Попередньо домовившись про інтерпретацію порядку відправник і отримувач здатні передавати повідомлення, просто варіюючи черговість DNS-запитів.

Отже, стеганографія прикладного рівня в SDN є ефективною завдяки широкому спектру доступних механізмів і слабкому моніторингу окремих елементів протоколів, однак якісний аналіз трафіку та машинне навчання здатні значно ускладнити приховану передачу даних.

Порівняльний аналіз методів стеганографії

З метою систематизації аналізованих методів протокольної стеганографії в табл. 1 продемонстровано ключові характеристики, зокрема пропускну здатність, рівень непомітності, технічні особливості впровадження та основні недоліки кожного методу. Такий підхід дозволяє не лише порівняти методи за ефективністю, а й виявити найбільш уразливі рівні мережної архітектури до прихованих каналів у середовищі SDN. Це особливо важливо для вибору пріоритетних напрямів подальших досліджень у сфері виявлення й нейтралізації стеганографічних загроз.

Таблиця 1

Порівняльні характеристики методів стеганографії в SDN за рівнями моделі OSI

Метод	Пропускна здатність	Особливість	Недолік	Непомітність
Фізичний/Канальний рівень				
StegoFrameOrder	~10 біт/с	Немає зміни кадрів	Потрібна координація між вузлами	висока
HICCUPS	кілька біт/с	Маскування під звичайні помилки	Висока ймовірність втрати кадрів	висока
PadSteg	~32 біт/с (в середньому)	Використання Etherleak доповнення	Потребує вразливості NIC Etherleak	висока
Dead Drop (ARP)	лічені біти/хв	Немає прямої взаємодії між відправником і отримувачем	Потрібні контроль над проміжним вузлом і узгоджені дії сторін	дуже висока
LLDP Tunnel	кілька байтів/обмін	Канал на основі площини управління	Маніпуляція топологією	середня
Мережний рівень				
IPv4 ID Field	до ~1600 біт/с	Поле заголовка IPv4 високої ємності	Повторювані значення ID	середня
IPv4 Reserved/ToS Bits	~1–3 біт на пакет	Вбудовування у невикористовувані біти заголовків	Політики QoS можуть змінювати або очищати біти з інформацією	висока
IPv4 Record Route Option	десятки–сотні біт/пакет	Опція рідко використовується	Великий заголовок запускає DPI	середня
IPv4 Timestamp Overflow	десятки–сотні біт/пакет	Вбудовування в поле переповнення	Обмежений розмір опціонального поля	середня
ICMP Echo (Ping)	до ~800 біт/с	Використання частих ping-пакетів	Вразливість до аналізу ping-трафіку	середня
IPv6 Traffic Class	8 біт/пакет	Непомітний якщо не використовується QoS	Вразливість до повторного маркування	висока
IPv6 Flow Label	до кількох тисяч біт/с	Використання 20-бітового поля Flow Label	Потрібен значний трафік і ретельна рандомізація	висока
IPv6 Extension Headers	до ~2048 біт/пакет	Приховування великих обсягів інформації	Великий розмір опцій легко помічається	середня
IPv6 Address Variation	кілька біт/с (залежно від частоти)	Кодування через хост-частину	Потреба часто змінювати IP-адресу	середня
VoNR IPD Timing	низька (залежить від трафіку)	Маскування природніх коливань затримок VoNR-трафіку	Вимагає точної імітації статистичних характеристик VoNR-трафіку	висока
Транспортний рівень				
TCP ISN	~16 біт/с (при 1 SYN/с)	Використовує ISN під час встановлення з’єднання	Вимагає великої кількості коротких з’єднань	середня
TCP ACK Number	до 32 біт/пакет	Велика ємність на пакет	Дуже помітний, неправильний ACK руйнує сеанс	низька
TCP Timestamp	кілька біт/пакет	Вбудовування в Timestamp	Аномальна динаміка Timestamp викликає підозру	середня
TCP NOP	~1 біт/пакет	Використання опції No-Operation як носія біта	Додаткові NOP у заголовку роблять структуру опцій TCP незвичною	середня
UDP Length Modulation	до ~456 біт/с	Кодування бітів шляхом зміни загальної довжини UDP-пакета на 1 байт	Потребує повного контролю над формуванням UDP-пакетів	висока
SCTP (Initiate Tag, PPI, Heartbeat тощо)	низька (обмежена випадками використання SCTP)	Використання унікальних полів і можливостей SCTP	SCTP трафік є підозрілим сам по собі	низька

Сеансовий рівень				
SIP/SDP Fields	десятки–сотні біт/с	Гнучкі невикористані поля заголовка	Можливе втручання SIP-інфраструктури	середня
Session Timing (SDN)	кілька біт/с	На основі факту початку/завершення сеансу	Необхідна точна синхронізація та попередня домовленість про інтерпретацію часових інтервалів	висока
Рівень представлення				
TLS Record Length	до сотень біт/с	Зміна довжини TLS-записів	Потребує контролю над фрагментацією записів	висока
TLS Initialization Vector	найвища серед TLS-методів (до сотень біт/с)	Маніпуляція вектором ініціалізації	Вимагає втручання в криптографічний процес TLS	висока
TLS Content Type Field	~1 біт/запис	Використання поля Content Type як носія	Невірне або неочікуване значення Content Type може не підтримуватися отримувачем	середня
QUIC Spin Bit	десятки біт/с	Передача даних через необов'язковий Spin Bit у заголовку QUIC (HTTP/3)	Залежність від реалізації QUIC (наявність та незаблокованість Spin Bit у конкретній мережі)	дуже висока
QUIC Connection Migration	кілька байтів на сеанс	Кодування в Server Preferred Address	Складність реалізації та залежність від підтримки цієї функції	висока
Прикладний рівень та площина управління SDN				
OpenFlow Features Reply	десятки байт на з'єднання	Комутатор кодує дані в полях повідомлення Features Reply	Одноразовий канал вимагає компрометації мережевого обладнання	висока
SDN Inter-switch Timing Channel	~20 біт/с	Два ізольовані OpenFlow-комутатори обмінюються інформацією через спільний SDN-контролер	Покладається на передбачувану реакцію контролера, яка може змінюватися	середня
HTTP Headers	кілька біт на запит	Зміни пробілів та регістру кодують біти	Загроза фільтрації/нормалізації проміжними проксі	висока
HTTP Parameters	кілька байт на запит	Вбудовування у поля запитів або URL	Незвично закодовані параметри можуть виділятися	середня
HTTP Timing Channel	~1–2 біт/с (дуже низька)	Кодування бітів в інтервалах між запитами	Потрібна точна синхронізація між сторонами	висока
DNS TXT Records	до ~64 КБ/відповідь (дуже висока)	Вбудовування в TXT ресурс	Простий для виявлення	низька
DNS TTL	кілька біт/запит	Вбудовування в значення TTL	Повторювані TTL зміни можуть виділятися	середня
DNS Query Ordering	біт/хвилини (дуже низька)	Кодування бітів в порядок виконання DNS-запитів	Для передачі даних потрібна значна кількість упорядкованих запитів	висока

Порівняльний аналіз методів протокольної стеганографії, узагальнений у табл. 1, дозволяє зробити низку важливих висновків. Зокрема, на фізичному та канальному рівнях методи прихованої передачі завдяки мінімальному впливу на легітимний трафік забезпечують дуже високий рівень непомітності, проте мають надзвичайно низьку пропускну здатність і часто вимагають специфічних умов.

Натомість на мережному рівні вбудовування даних у поля заголовків IP або службові протоколи дозволяє суттєво підвищити пропускну здатність прихованого каналу при збереженні помірної чи навіть високої непомітності за обережної реалізації. Водночас нетипові

шаблони (повторювані ідентифікатори, надмірні опції тощо) на цьому рівні можуть привертати увагу засобів глибокого аналізу трафіку або бути зміненими мережними політиками, особливо з огляду на централізований контроль у SDN.

Методи транспортного рівня здатні забезпечувати як низьку, так і значну пропускну здатність, але всі вони потребують ретельного налаштування, щоб зміни не порушили нормальний перебіг з'єднань і не привертали уваги систем моніторингу. Вбудовування даних у керівні поля TCP (номери послідовності/підтвердження, опції) або незначна модифікація довжини UDP-пакетів залишається непомітною в межах типової динаміки трафіку, проте грубе втручання (некоректні ACK, надлишкові опції тощо) легко викривається; більш того, у SDN централізований аналіз трафіку полегшує виявлення таких аномалій.

Сеансовий рівень також може слугувати для прихованої передачі даних – через метадані з'єднань і синхронізацію подій. Такі канали характеризуються низькою пропускною здатністю, але забезпечують високий рівень непомітності, маскуючись під нормальну поведінку сеансів; утім, вони вимагають точної синхронізації, і будь-які нетипові затримки чи незвичні службові поля можуть бути помічені системами керування сеансами (наприклад, SIP-проксі або SDN-контролером).

На рівні представлення (TLS, QUIC) приховані канали використовують переваги шифрування та гнучких полів протоколів, забезпечуючи дуже високий рівень непомітності при значній пропускній здатності. Наприклад, маніпуляція довжиною TLS-записів або вектором ініціалізації, а також використання опціонального біта spin-bit у QUIC дозволяють передавати сотні біт/с непомітно для стандартних IDS. Втім, реалізація цих методів потребує складного втручання в криптографічні процеси і залежить від специфічних функцій протоколів, що ускладнює впровадження та може бути нейтралізовано оновленнями протоколів.

Нарешті, на прикладному рівні та в площині управління SDN спостерігається найширший спектр стеганографічних технік – від майже непомітних, але низької пропускної здатності (мінімальні варіації заголовків чи часових інтервалів HTTP-запитів) до високопродуктивних, проте легко виявлюваних (приховування значних обсягів даних у DNS або службових повідомленнях SDN). Різноманіття високорівневих протоколів робить такі канали ефективними та часто непоміченими традиційним моніторингом, однак ретельний аналіз трафіку й використання методів машинного навчання дозволяють викрити аномальні шаблони та суттєво обмежити їх.

Висновки

Здійснено комплексний огляд протокольних методів стеганографії з позицій їх релевантності до сучасних мережних архітектур, зокрема SDN. Запропонована класифікація методів за рівнями моделі OSI дозволила систематизувати наявні підходи та встановити залежності між рівнем реалізації протоколу, можливостями прихованої передачі, рівнем непомітності та операційними обмеженнями. Виявлено, що високопродуктивні приховані канали зазвичай генерують аномальні шаблони трафіку, які полегшують їх виявлення, тоді як низькошвидкісні канали забезпечують значно вищий рівень маскування у фоновому потоці. Проведений порівняльний аналіз продемонстрував, які саме рівні мережної моделі є найбільш уразливими до стеганографічних каналів у SDN-середовищі, що важливо для визначення пріоритетів захисту мережі.

Окрему увагу приділено особливостям архітектури SDN, які впливають на розвиток прихованих каналів і методи протидії їм. Глобальна видимість трафіку, яку забезпечує SDN-контролер, відкриває широкі можливості для централізованого моніторингу та виявлення аномалій (нетипового використання полів протоколів, нестандартної динаміки сеансів тощо). Водночас централізований характер управління може бути використаний зловмисниками: у разі компрометації контролера або змови мережевих вузлів стає можливим координувати складні багатокрокові приховані обміни даними, що обходять традиційні механізми фільтрації та сегментації трафіку. Продемонстровано перспективність міжпротокольної стеганогра-

фії, яка завдяки взаємодії кількох протоколів забезпечує підвищену стійкість прихованого каналу та адаптацію до змін мережевих умов. Таким чином, навіть у сегментованих або ізольованих середовищах SDN приховані канали здатні підтримувати нелегітимний зв'язок, використовуючи дозволені протоколи й порти та мінімізуючи ризик виявлення.

Отримані результати вказують на необхідність подальших досліджень, спрямованих на розвиток спеціалізованих засобів стегоаналізу для SDN. Зокрема, перспективним є впровадження централізованих систем виявлення прихованих каналів на рівні контролера, що використовують глобальне бачення мережі й методи машинного навчання для розпізнавання малопомітних аномалій у багатошаровому трафіку. Важливим напрямом є також дослідження міжпротокольних та багатошарових каналів з метою розробки ефективних методів їх нейтралізації до того, як такі техніки набудуть поширення у зловмисників. Таким чином, комплексний підхід, який поєднує врахування вразливостей усіх рівнів мережі, можливості SDN для глобального моніторингу і адаптивний аналіз трафіку, є ключовим для випередження розвитку стеганографічних загроз у майбутніх мережах.

Список літератури:

1. B. Jankowski, W. Mazurczyk, and K. Szczypiorski. PadSteg: Introducing inter-protocol steganography // Telecommunication Systems, preprint, 2011. doi: 10.1007/s11235-011-9616-z.
2. T. Schmidbauer et al. Introducing dead drops to network steganography using ARP-caches and SNMP-walks // Proc. 14th Int. Conf. Availability, Reliability and Security (ARES). 2019. P. 1–10. doi: 10.1145/3339252.3341488.
3. J. Hua, Z. Zhou, and S. Zhong. Flow misleading: Worm-hole attack in software-defined networking via building in-band Covert Channel // IEEE Trans. Inf. Forensics Security. 2021. Vol. 16. P. 1029–1043. doi: 10.1109/TIFS.2020.3013093.
4. K. Sawicki, G. Bieszczad, and Z. Piotrowski. StegoFrameOrder-AC layer covert network channel for wireless IEEE 802.11 networks // Sensors. 2021. Vol. 21, no. 18. P. 6268. doi: 10.3390/s21186268.
5. K. Szczypiorski. HICCUPS: Hidden Communication System for Corrupted Networks // Proceedings of the 10th International Multi-Conference on Advanced Computer Systems (ACS), Międzyzdroje, Poland, Oct. 2003, pp. 31–40.
6. C. H. Rowland. Covert channels in the TCP/IP protocol suite // First Monday, preprint, 1997. doi: 10.5210/fm.v2i5.528.
7. M. Wolf. Covert channels in LAN protocols // Lect. Notes Comput. Sci., 1989, pp. 89–101. doi: 10.1007/3-540-51754-5_33.
8. Z. Trabelsi and I. Jawhar. Covert File Transfer Protocol Based on the IP Record Route Option // Journal of Information Assurance and Security. 2010. Vol. 5, no. 1. P. 64–73.
9. P. Bedi and A. Dua. Network steganography using the overflow field of timestamp option in an IPv4 packet // Procedia Comput. Sci. 2020. Vol. 171. P. 1810–1818. doi: 10.1016/j.procs.2020.04.194.
10. A. Sharma. New Windows 'Pingback' malware uses ICMP for covert communication // BleepingComputer, May 4, 2021. Available: https://www.bleepingcomputer.com/news/security/new-windows-pingback-malware-uses-icmp-for-covert-communication/
11. S. Amante, B. Carpenter, S. Jiang, and J. Rajahalme. IPv6 Flow Label Specification, IETF, RFC 6437, Nov. 2011. Available: https://datatracker.ietf.org/doc/html/rfc6437
12. N. B. Lucena, G. Lewandowski, and S. Chapin. Covert channels in IPv6 // Privacy Enhancing Technologies, G. Danezis and D. Martin, Eds. Heidelberg : Springer, 2006. P. 147–166. doi: 10.1007/11767831_10.
13. M. Bobade and A. Sagar. Survey and design approach of protocol steganography in IPv6 // Int. J. Comput. Appl. 2013. Vol. 69, no. 7. P. 31–34. doi: 10.5120/11856-7623.
14. M. Wang, S. Cao, and Y. Wang.VoNR-IPD: A Novel Timing-Based Network Steganography for Industrial Internet // Security and Communication Networks. 2020. Vol. 2020, Article ID 8846230, 14 p., Jun. 2020. doi: 10.1155/2020/8846230.
15. G. Fisk, M. Fisk, C. Papadopoulos, and J. Neil. Eliminating steganography in internet traffic with active wardens // Information Hiding, F. A. P. Petitcolas, Ed. Heidelberg: Springer, 2003. P. 18–35. doi: 10.1007/3-540-36415-3_2.
16. A. S. Nair, A. Kumar, A. Sur, and S. Nandi. Length based network steganography using UDP protocol // Proc. 2011 IEEE 3rd Int. Conf. Commun. Softw. Netw., Xi'an, China, May 2011. P. 588–592. doi: 10.1109/ICCSN.2011.6014994.
17. W. Fraczek, W. Mazurczyk, and K. Szczypiorski. Stream control transmission protocol steganography // Proc. 2010 Int. Conf. Multimedia Inf. Netw. Secur., Nanjing, China, Nov. 2010. P. 829–834. doi: 10.1109/MINES.2010.176.

18. W. Mazurczyk and K. Szczypiorski. Covert channels in SIP for VoIP signalling // Multimedia Communications, Services and Security, R. Choras, Ed. Heidelberg: Springer, 2008, pp. 65–76. doi: 10.1007/978-3-540-88873-4_6.

19. Y. Ji, Y. Wang, and Y. Zhang. Constructing SDN covert timing channels between hosts with unprivileged attackers // IEEE/ACM Trans. Netw. 2024. Vol. 32, no. 1. P. 1–14. doi: 10.1109/TNET.2024.3496997.

20. C. Heinz, M. Zuppelli, and L. Caviglione. Covert Channels in Transport Layer Security: Performance and Security Assessment // Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications. 2021. Vol. 12, no. 4. P. 22–36. doi: 10.22667/JOWUA.2021.12.31.022.

21. T. Grübl, W. Niu, J. von der Assen, and B. Stiller. QUIC-Exfil: Exploiting QUIC's Server Preferred Address Feature to Perform Data Exfiltration Attacks // Proc. ACM Asia Conf. on Computer and Communications Security (AsiaCCS). Hanoi, Vietnam, 2025. doi: 10.1145/3708821.3733872.

22. K. Thimmaraju, R. Krösche, L. Schiff, and S. Schmid. CVE-2018-1000155: Denial of service, improper authentication and authorization, and covert channel in the OpenFlow 1.0+ handshake. oss-sec mailing list, May 9, 2018. Available: https://seclists.org/oss-sec/2018/q2/99

23. R. Krösche, K. Thimmaraju, and S. Schmid. I DPID it my way! A covert timing channel in software-defined networks // Proc. 2018 IFIP Netw. Conf., Zurich, Switzerland, May 2018, pp. 1–9. doi: 10.23919/IFIPNetworking.2018.8696597.

24. S. Bistarelli, M. Ceccarelli, C. Luchini, I. Mercanti, and F. Santini. A Preliminary Study on the Creation of a Covert Channel with HTTP Headers // Proceedings of the Italian Conference on CyberSecurity (ITASEC 2024), Ancona, Italy, Jan. 2024. Available: https://ceur-ws.org/Vol-3731/paper34.pdf

25. E. Brown, B. Yuan, D. Johnson, and P. Lutz. Covert Channels in the HTTP Network Protocol: Channel Characterization and Detecting Man-in-the-Middle Attacks // Journal of Information Warfare. 2010. Vol. 9, no. 3. P. 1–12. Available: https://repository.rit.edu/other/781/

26. S. Smendowski. Hidden Communication Using Covert Channels // GitHub repository, 2021. Available: https://github.com/Smendowski/hidden-communication-using-covert-channels

27. W. A. Dimitrova and G. S. Panayotova. The Impacts of DNS Protocol Security Weaknesses // Journal of Communications. 2020. Vol. 15, no. 5. P. 1–9. Available: https://www.jocm.us/show-245-1596-1.html

28. M. Hildebrandt, R. Altschaffel, K. Lamshöft, M. Lange, M. Szemkus, T. Neubert, C. Vielhauer, Y. Ding, and J. Dittmann. Threat Analysis of Steganographic and Covert Communication in Nuclear I&C Systems // presented at the International Conference on Nuclear Security: Sustaining and Strengthening Efforts (ICONS 2020), Vienna, Austria, Feb. 2020. Available: https://conferences.iaea.org/event/181/contributions/15608/attachments/8569/11404/CN278_478-stealth_v006.pdf

Надійшла до редколегії 15.06.2025

Відомості про авторів:

Фокін Денис Геннадійович – Харківський національний університет радіоелектроніки, аспірант кафедри інфокомунікаційної інженерії ім. В.В. Поповського, Україна; e-mail: denys.fokin@nure.ua; ORCID: https://orcid.org/0009-0002-3282-842X

Євдокименко Марина Олександрівна – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри інфокомунікаційної інженерії ім. В.В. Поповського, Україна; e-mail: maryna.yevdokymenko@ieee.org; ORCID: https://orcid.org/0000-0002-7391-3068

УДК 621.396 (075) **DOI:10.30837/rt.2025.3.222.17**

О.Й. КАДАЦЬКА, канд. техн. наук, С.О. САБУРОВА

МАТЕМАТИЧНА МОДЕЛЬ ЗАТРИМКИ ПЕРЕДАЧІ ДАНИХ В SDN-КЕРОВАНІЙ 5G МЕРЕЖІ

Вступ

Мережі мобільного зв'язку 5G нового покоління повинні забезпечити підтримку різноманітних сервісів, які можна згрупувати [1]:

1) покращений мобільний широкосмуговий доступ (enhanced Mobile Broadband, eMBB);

2) масове підключення та обмін даних між машинними терміналами (massive Machine Type Communications, mMTC);

3) наднадійний зв'язок із низькими затримками (Ultra-Reliable and Low Latency Communications, URLLC).

Для реалізації таких різноманітних сценаріїв доставки послуг 5G мережі повинні відповідати основним вимогам:

1) пікова швидкість передачі даних на нізсхідному каналі зв'язку – 20 біт/с та 10 Гбіт/с на висхідному каналі зв'язку;

2) пікова спектральна ефективність на нізсхідному каналі зв'язку – 30 біт/с/Гц та 15 біт/с/Гц на висхідному каналі зв'язку;

3) мінімальна затримка в підсистемі радіодоступу 0,5 мс для сервісів унаднадійного зв'язку URLLC та 4 мс – високошвидкісного зв'язку eMBB;

4) автономна робота машинних терміналів без підзарядки акумулятора протягом щонайменше 10-ти років (і бажано до 15 років);

5) максимальна щільність підключених до мережі пристроїв 1000000 пристроїв/км^2 у міських умовах;

6) функції мобільності повинні підтримуватись при максимальній швидкості руху об'єктів 500 км/год.

Програмно визначена мережа SDN (Software-Defined Networking) – підхід, що забезпечує інтелектуальні та гнучкі програмовані 5G мережі, здатні конфігурувати та контролювати додатки/сервіси більш детально та в масштабах усієї мережі. SDN створює віртуалізовану площину управління, яка здатна застосовувати інтелектуальні управлінські рішення в мережних функціях. Це усуває розрив між наданням і управлінням якістю послуг (Quality of Experience, QoE) в 5G мережах [2]. SDN може забезпечувати контекстно-залежне управління QoE та відповідати ключовим показникам продуктивності (Key Performance Indicator, KPI) 5G мережі, зберігаючи її цілісність, надійність, а також зменшуючи затримку для чутливих до затримок мультимедійних програм.

Основні методи конфігурування контролерів SDN Open Flow при взаємодії з центром обробки даних 5G мережі

Операційна система мережі (Network Operating System, NOS) використовується для маршрутизації пакетів та є контролером в парадигмі SDN, центральним компонентом, де реалізується об’єднання роз’єднаних інтелектуальних даних звичайних мережних пристроїв. Існують різні контролери SDN, які побудовані на стеку різних мов програмування.

OpenFlow є основним протоколом архітектури SDN. Контролери використовують протокол OpenFlow для зв'язку з комутаторами (пристроями пересилання). OpenFlow не є специфічним протоколом постачальника, який передбачає, що контролер може спілкуватися з кожним комутатором незалежно від постачальника. Комутатори SDN відрізняються від звичайних комутаторів, тому комутатори SDN – пристрої пересилання, оскільки вони оснащені лише площинами даних. Вони можуть бути апаратними або програмними: OVS (Open

vSwitch) – це більш популярний віртуальний комутатор, який використовується в парадигмі SDN для підключення кінцевих пристроїв.

Інструмент моделювання SDN Mininet – симулятор з відкритим кодом для тестування мережі SDN OpenFlow, що дозволяє всю мережу створити як віртуальну машину, яку можна завантажувати, запускати, перевіряти та змінювати. Віртуальні комутатори в Mininet (Open vSwitch) є різновидом програмних комутаторів OpenFlow.

За допомогою SDN у центрі обробки даних кожна програма може розпізнавати конфігурацію 5G мережі на основі того, як для неї виглядає мережне обладнання, але лише швидкість, доступні місця та ресурси такі, як сховище. Площина управління надсилає сигнали на площину даних для переналаштування комутаторів або іншого обладнання для задоволення (або кращого задоволення) усіх потреб додатків, що логічно, але не фізично змінює конфігурацію мережі так, що нова мережа виглядає як інша фізична мережа для програм, включаючи мережні ресурси. За допомогою такого типу віртуалізації потоки даних програми через мережу прискорюються і вона має доступ до більшої пам'яті.

Математична модель затримки в SDN-орієнтованому управлінні ресурсами 5G-мережі

Модель взаємодії між кожним користувачем і радіодоступом RAP на основі SDN платформі з OpenFlow протоколом для 5G мережі включає елементи:

- мережні контролери;
- сервери радіобазових станцій;
- комутатори (switch OpenFlow);
- радіобазові станції;
- транспортні ресурси 5G мережі;
- термінали 5G споживачів.

Розробка методики розрахунку, аналізу та оцінки параметрів якості в системи управління елементами контрольованого об'єкта (КО) на основі моделі взаємодії між кожним користувачем і RAP з підтримкою протоколу OpenFlow SDN платформи для 5G мережі є метою гарантії достовірної оцінки за вимогами виконання норм показників якості QoS:

- затримки передачі пакетів з кінця в кінець;
- вірогідності втрачених пакетів;
- характеристик короткочасних перерв сигналів;
- захисту апаратного та програмного забезпечення елементів;
- інших збоїв.

Алгоритм обробки пакетів у протоколі OpenFlow на SDN платформі представлено на рис. 1. Контролер SDN забезпечує гнучке управління та вибір маршрутизатора для підключення всіх мереж радіодоступу (Radio Access Network, RAN) до основної 5G мережі, де базова мережа контролера SDLL складається з двох основних частин, як об'єкт єдиного управління (Unified Control Entity, UCE) та єдиний шлюз даних (Unified Data Gateway, UDW). Роль UCE полягає у визначенні правил управління в 5G мережі, таких як:

- об'єкт управління мобільністю (Mobility Management Entity, MME);
- площина управління шлюзом обслуговування (SGW-C);
- площина управління шлюзом пакетної передачі даних (Gateway Control Plane, PGW-C).

У цьому випадку UDW визначає правила для пересилання даних 5G мережі:

- площина даних шлюзу обслуговування (Service Gateway Data Plane, SGW-D);
- площина даних шлюзу пакетної передачі даних (Gateway Data Plane, PGW-D).

Для забезпечення гарантії стабільного з'єднання між кожним користувачем і RAP розробимо модель управління елементами на основі SDLL платформи для 5G мережі, яка буде забезпечувати необхідний взаємозв'язок між площиною даних і площиною управління.

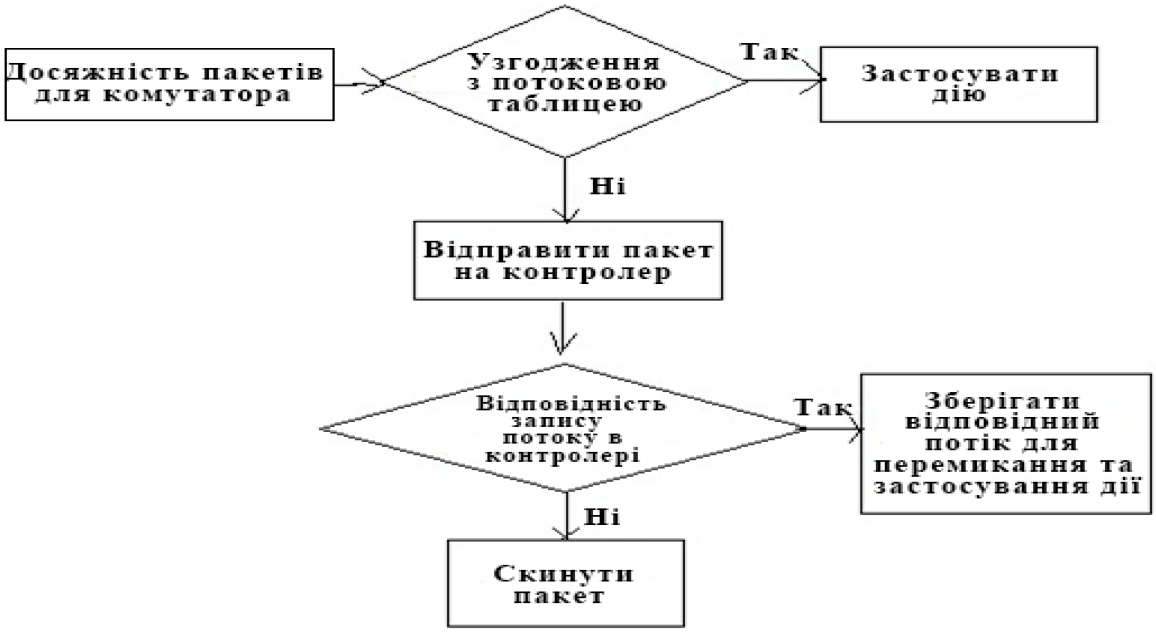


Рис. 1. Алгоритм обробки пакетів у протоколі OpenFlow на SDN платформі

Базуючись на головних вимогах системи якості QoS і SDN платформи запропоновано модель взаємодії між кожним користувачем і кожної точки радіодоступу (Radio Access Points, RAP) таким чином, що контролер SDN повинен динамічно регулювати смугу пропускання для RAP до модуля основної смуги (Baseband Unit, BBU) 5G мережі (рис. 2).

У запропонованій моделі площина управління, яка управляється мережними серверами, надає інструменти управління та оптимізації для площини даних, якою управляє базова 5G мережа. Вона складається з програмно визначених базових станцій (Software-Defined Base Stations, SD-BS) у RAN та програмно-визначених комутаторів (Software-Defined Switches, SD-switches). Контролер обслуговує фізичні, MAC (Medium Access Control) і мережні рівні функцій на комп'ютерах і віддалених центрах обробки даних. В свою чергу протокол OpenFlow підтримує три різні типи повідомлень. Повідомлення від контролера до комутатора є асинхронними та симетричними, ініціюються контролером і використовуються для перевірки стану та стану таблиці потоків комутаторів. Асинхронні повідомлення надсилаються від комутатора до контролера, що є подією та позначає зміну у стані комутатора або мережі.

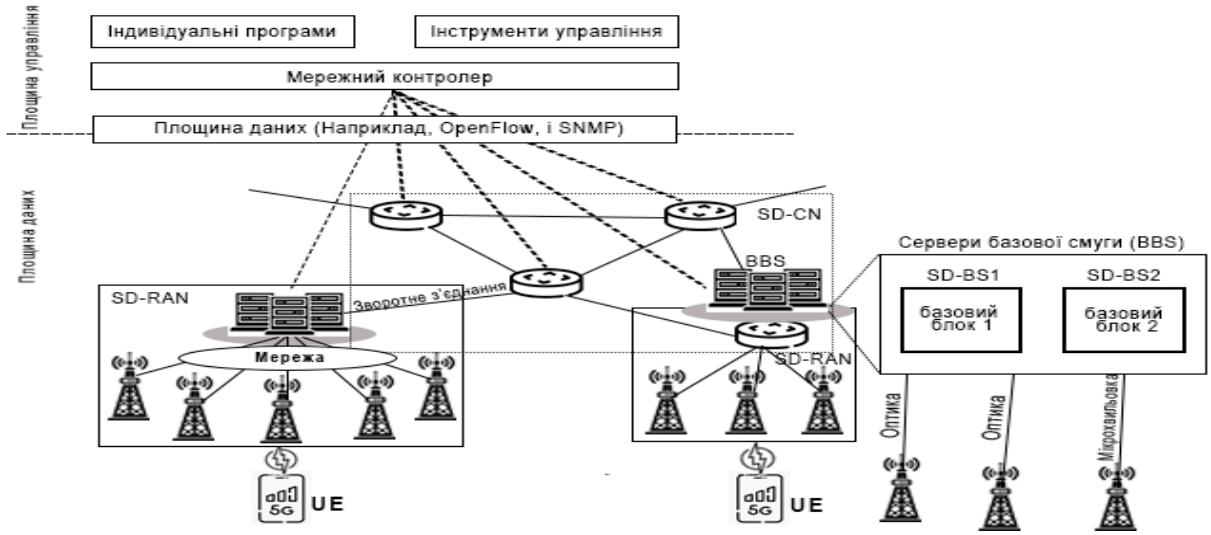


Рис. 2. Модель взаємодії користувачів і RAP на основі SDN платформи для 5G мережі

Серед багатьох подій дія Packet-in має важливе значення. Ця подія виникає, коли пакет не має відповідного запису в таблиці потоку. Повідомлення Packet-in надсилається до контролера, який приймає рішення про встановлення потоку для пакета, симетричні повідомлення надсилаються у будь якому напряму та використовуються для перевірки активності контролера. Коли пристрій із підтримкою Openflow намагається налаштувати себе вперше, він спочатку надсилає повідомлення синхронізації TCP на контролері TCP-порта за замовчуванням. Повідомлення підтвердження надсилається з обох сторін під час реалізації протоколу TCP між контролером і комутатором. Повідомлення реалізації протоколу TCP використовуються в процесі встановлення з'єднання. Після встановлення з'єднання між комутатором і контролером хости можуть спілкуватися з OpenFlow мережею.

Розрахунки показників якості, аналіз ймовірнісної-часової характеристики обробки інформації проводяться, як правило, в центрі обробки результатів даних розподіленої системі управління контрольованими об'єктами (КО) на основі моделі взаємодії користувачів і RAP з підтрикою протоколу OpenFlow SDN платформи для 5G мережі. Ліквідація проблем збоїв в роботі контролерів SDN платформи реалізуються прийняттям оперативних рішень для забезпечення норм надійності і безпеки функціонування об'єктів, що контролюються (КО). В центрі обробки результатів даних розподіленої системі управління контрольованими об'єктами OpenFlow SDN платформи для 5G мережі основними задачами являються проведення розрахунків нижче вказаних показників математичної моделі:

– математичного очікування, де T – це час обробки інформаційного повідомлення від вбудованого контролера, починаючи з моменту його ініціалізації в елементі КО до завершення циклу контроля та обробки його інформації в БД сервера, у т.ч. час передачі результатів обробки в БД протоколу OpenFlow SDN платформи;

– циклічного опитування K_i контролерів, які забезпечують контроль кожного елемента КО з запитом про стан показників якості при порівнянні з показниками норм;

– незалежність випадкової довжини черги – A_i в вихідному буфері K_i-го контролера в моменти опитування від випадкових процесів в інших об'єктах системи, побічно визначаються в сукупності випадковим середнім часом циклу опитування – $T_{\text{сер}}$.

Повідомлення від кожного і-го контролера надходять через випадковий час T_{A_i} з функцією розподілу *Ai(t)* та інтенсивністю λ в системі управління ресурсами 5G мережі;

$$A_i(t) = P\{T_{A_i} \le t\}, \tag{1}$$

$$\lambda_i = \frac{1}{E[T_{A_i}]}. \tag{2}$$

Тоді інтенсивність сумарного потоку:

$$\lambda_0 = \sum_{i=1}^{K} \lambda_i. \tag{3}$$

Тривалість обслуговування повідомлення від контролерів є випадковими величинами T_{B_i} з довільними функціями розподілу:

$$B_i(t) = P\{T_{B_i} \le t\}. \tag{4}$$

$$E[T_{B_i}] = \int_0^{\infty} [1 - B_i(t)]dt = bi \le \infty. \tag{5}$$

Опитування контролерів проводиться циклично, при чому при кожному зверненні обслуговується не більше одного повідомлення з буфера вихідного контролера. Оскільки буфером використовується пам'ять ОЗП, а від кожного контролера трафік відомий, то розмір буфера вибирається таким, що його можна вважати необмеженим. Процедура опитування контролера вимагає випадкового часу організації діалогу T_{D_i} з функцією розподілу:

$$D_i = E[T_{D_i}]. \tag{6}$$

У разі надходження інформаційного повідомлення від будь-якого контролера після його обробки потрібно передати результат його обробки в головний сервер протягом випадкового часу T_v і математичним сподіванням (середній час, який витрачається на обробку 1 заявки):

$$v = E[T_v]. \tag{7}$$

Нехай TR_i – час між двома послідовними зверненнями до вихідного буфера i-го контролера має функцію розподілу $R_i(t) = P\{TR_i \le t\}$. Тоді математичне сподівання середнього часу циклу опитування:

$$r = d_0 \sum_{i=1}^{K} [(\lambda_i \cdot r) \cdot b_i + (\lambda_i \cdot r)v], \tag{8}$$

$$r = d_0 + r \cdot \sum_{i=1}^{K} [(\lambda_i \cdot b_i + \lambda_i \cdot v],$$

$$r - r \cdot \sum_{i=1}^{K} [(\lambda_i \cdot b_i + \lambda_i \cdot v] = d_0,$$

$$r \cdot \left(1 - \Sigma_{i=1}^{K} [(\lambda_i \cdot b_i + \lambda_i \cdot v)]\right) = d_0$$

$$r = \frac{d_0}{1 - \sum_{i=1}^{K}[(\lambda_i \cdot b_i + \lambda_i \cdot v)]} = \frac{d_0}{1 - (\sum_{i=1}^{K} \lambda_i \cdot b_i + \sum_{i=1}^{K} \lambda_i \cdot v)} =$$

$$= \frac{d_0}{1 - (\sum_{i=1}^{K} \rho_i + v \cdot \sum_{i=1}^{K} \lambda_i)},$$

$$r = \frac{d_0}{1 - (\rho_0 + v\lambda_0)} \ , \tag{9}$$

де d_0 – сумарний час опитування всіх буферів, ч; λ_i – інтенсивність навантаження мережі; $v = \mathrm{T}$ – математичне сподівання опитування 1 контролера; b_i – номер контролера; K – загальна кількість контролерів SDN; $\rho_0 = \sum_{i=1}^{K} \rho_i$; $\rho_i = \lambda \cdot b_i, v = b_i T$; λ_0 – інтенсивність сумарного потоку.

Для аналізу та оцінки параметрів якості функціонування елементів КО в системі управління об'єктами запропонованої моделі проведемо розрахунки з використанням MATLAB (табл. 1).

Таблиця 1

Число контролерів b_i,	Інтенсивність сумарного потоку λ	Час передачі повідомлень d, с	Час циклу опитування Тс
від 1 до 200 (з кроком 10)	від 0,1 до 1,0 (з кроком 0,1)	0,0012	0,001 для 1 контролера

Графіки залежності часу затримки циклу опитування від заданої кількості контролерів мережних елементів опитування.

I сценарій: при заданому циклу опитування, $\mathrm{T} = 0{,}001$с для 1 контролера і інтенсивності 0,1 сумарний час опитування в результаті зростає лінійно в умовах збільшення КО та кількості K. Мінімальне значення $\mathrm{r}(\mathrm{к}) = 0 \ldots 0{,}05$ с фіксується для K від 1 до 40. Зріст навантаження викликає зріст часу затримки для циклу опитування. Тому максимального значення математичне сподівання досягає при K=200 для $\mathrm{r}(\mathrm{к}) = 0{,}2 \ldots .0{,}35$ с в умовах залежності від інтенсивності надходження потоків. Мінімальне значення r (к) краще вибирати при К≤ 20 , що забезпечує час затримки, r (к)=0.03с. Графіки залежності циклу середнього часу опитування від кількості мережних елементів представлені на рис. 3.

II сценарій: при $\mathrm{T} = 0{,}002$ с з інтенсивністю 0,1 сумарний час опитування при зростанні кількості елементів та контролерів також має лінійне зростання та при K=140 досягає 0,3 с.

Мінімальне значення математичного сподівання фіксується при K від 1 до 40 для $r(к) = 0{,}00001 \ldots 0{,}1$ с. Зріст навантаження на контролери викликає збільшення часу затримки для циклу опитування. Тому максимального значення математичне сподівання досягає для r (к) =0,4 ... 1,05 с при K =200 в умовах залежності від інтенсивності надходження потоків. Мінімальне значення r(к) краще вибирати при К≤ 20 , що забезпечує час затримки r(к)= 0,005 с. За результатами розрахунків представлено графіки залежності циклу середнього часу опитування від кількості мережних елементів на рис. 4.

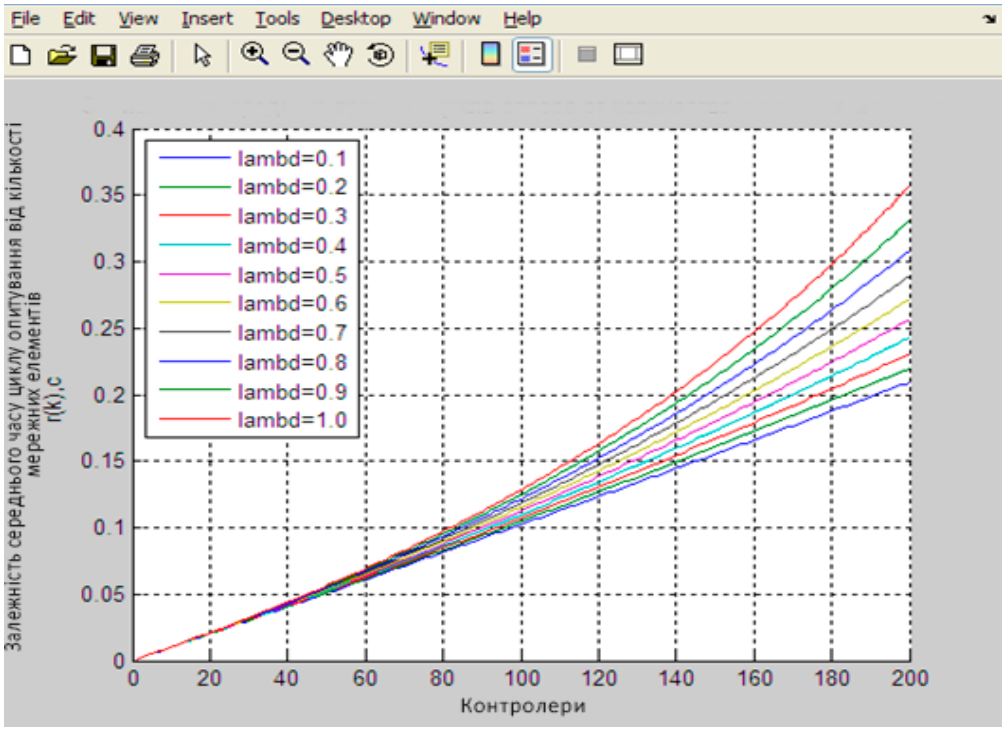


Рис. 3. Залежність числа K для середнього часу циклу опитування від кількості мережних елементів при $r(к) = 0{,}001 \ldots 0{,}35$ $с$

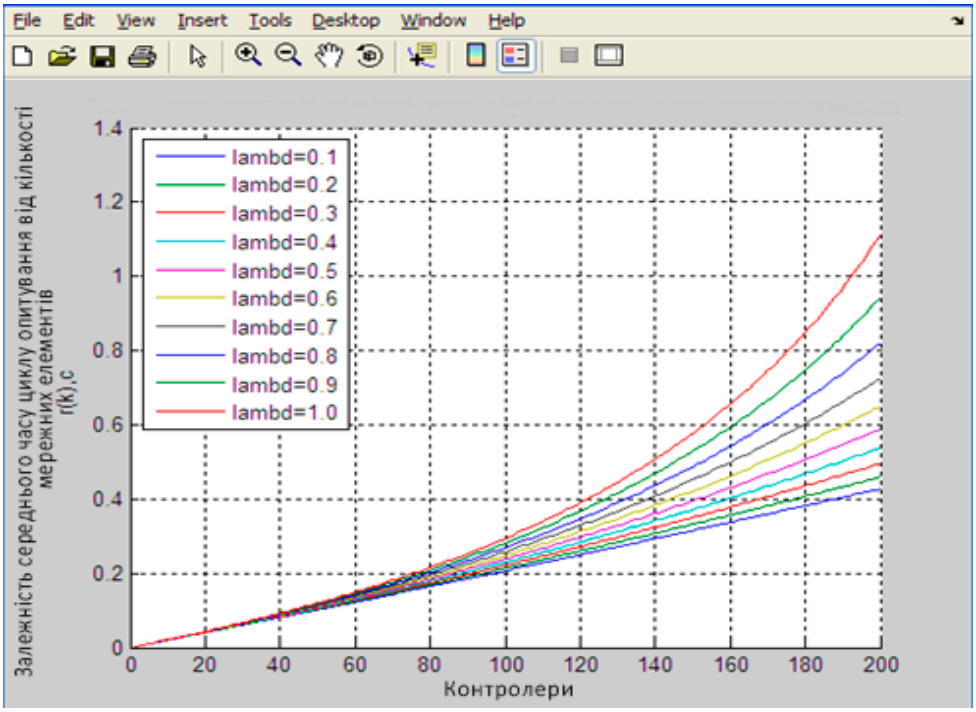


Рис. 4. Залежність числа K для середнього часу циклу опитування від кількості мережних елементів при $r\ (к)\ =\ 0{,}4 \ldots 1{,}05$ $с$

III сценарій: при T = 0,003 с з інтенсивністю 0,1 сумарний час опитування при збільшенні кількості контролерів також має лінійне зростання, для К= 140 сумарний час опитування досягає 0,4с. Оптимальне значення r(к) краще вибирати при середній кількості контролерів К від 140 до 180 для часу затримки 1 с, мінімальне значення r (к) = 0 ...0,2 с для К від 1 до 40. Зріст навантаження викликає зростання часу затримки для циклу опитування. Тому максимальне значення, r (к) = 0,7 ... 3,7 с досягає при *K*=200 в умовах залежності від інтенсивності надходження потоків. Мінімальне значення r(к) краще вибирати при К≤ 20, що забезпечує час затримки r(к)= 0,1 с. За результатами розрахунків представлені графіки залежності циклу середнього часу опитування від кількості мережних елементів на рис. 5.

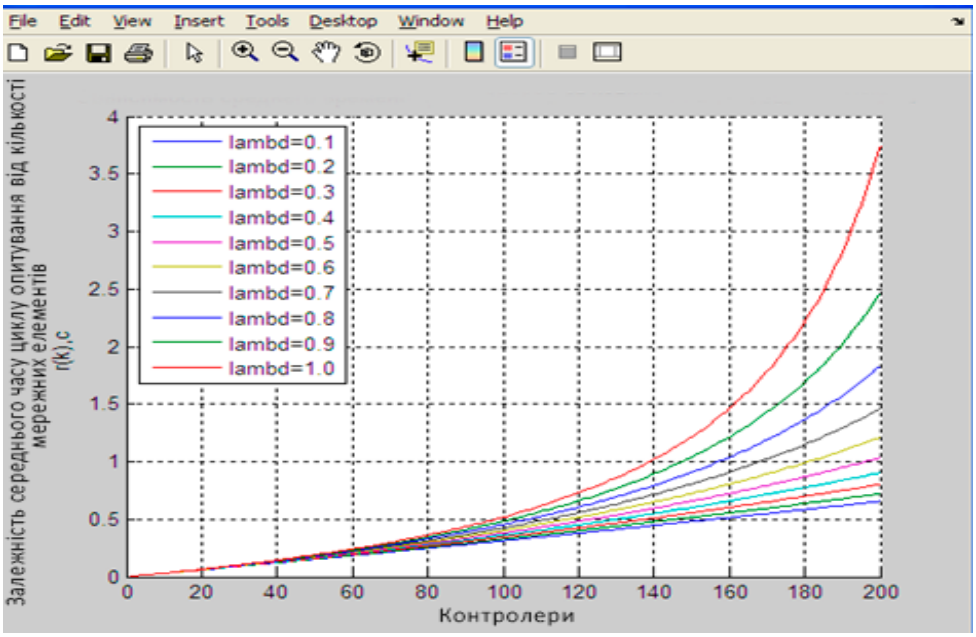


Рис. 5. Залежність числа *K* для середнього часу циклу опитування від кількості мережних елементів при r (к) = 0,7 ... 3,7 с

В архітектурі SDN типові та поширені проблеми безпеки забезпечуються на рівнях SDN. Рівень додатків, також відомий як рівень управління, є найвищим рівнем в архітектурі SDN, всі бізнес-додатки та програми безпеки виконуються на цьому рівні. Додатки, контрольовані цим рівнем, включають реалізацію брандмауера, контроль доступу,балансування навантаження, систему запобігання вторгненням (Intrusion Prevention System, IPS), систему виявлення вторгнень (Intrusion Detection System, IDS) та віртуалізацію мережі.

Висновки

Результати математичного моделювання, розрахунки, аналіз графіків та оцінка даних показали ефективність системи управління на SDN платформі з протоколом OpenFlow. Для різних сценаріїв проведено моделювання з використанням пакету математичного моделювання MATLAB та визнано кращий сценарій конфігурації, а саме при T=0,001 с з інтенсивністю 0,1 сумарний час опитування при збільшенні кількості контролерів має лінійне зростання. Середній час в циклі опитування набуває мінімального значення при кількості контролерів не вище 20, що забезпечує час затримки r (к) =0,005 с. Такі значення рекомендовано вибирати за результатами моделювання і вони відповідають 1 класу обслуговування 3GPP TS 23.501:5QI для QoS [3].

Таким чином, повне використання переваг системи управління повинно відповідати об’єкту моделі взаємодії між кожним користувачем і RAP на основі SDN платформи з використанням OpenFlow протоколу для 5G мережі. Такий підхід дозволяє системі управління в реальному часі контролювати використання і безпеку загальних ресурсів з нормованими параметрами надійності доступу до них.

Список літератури:

1. Introduction to 3GPP and 3GPP 5G Releases 15, 16 and 17. URL: https://5g.security/miot-5g/5g-3gpp-releases-15-16-17.

2. Barakabitze A., Barman N., hmad A., Zadtootaghaj S., Sun L., G. Martini M., Atzori L. QoE management of multimedia streaming services in future networks: A tutorial and survey // IEEE Commun. Surv. Tutor. 22(1) (2019). P.526-565.

3. ETSI TS 123 501 V16.6.0 (2020-10) URL: https://www.etsi.org/deliver /etsi ts/123500 123599 /123501 /16.06.00 60/ts 123501v160600p.pdf.

Надійшла до редколегії 10.05.2025

Відомості про авторів:

Кадацька Ольга Йосипівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В. В. Поповського, Україна; e-mail: olga.kadatska@nure.ua; ORCID: https://orcid.org/0000-0002-5331-4324

Сабурова Світлана Олександрівна – Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В.В. Поповського, Україна; e-mail: svitlana.saburova@nure.ua; ORCID: https://orcid.org/0000-0002-4001-1594

УДК 621.372(075) **DOI:10.30837/rt.2025.2.222.18**

O.J. KADATSKAYA, PhD, C.O. SABUROVA

CONTROL OF CONTACT CENTER MODEL FUNCTIONAL PARAMETERS TO AGENTS LOAD REDUCTION

Introduction

Today, a call center is an integral part of any telecommunications operator. When establishing a contact center, it is essential to choose a platform that supports workforce optimization, centralized call routing, task handling, and interaction history retention. Such a platform should contribute to achieving goals across all service areas, including self-service, agent-assisted service, and proactive customer engagement by the organization [1]. A call center achieves maximum efficiency through the use of digital telecommunications technologies, balanced operator workload distribution, and partial automation of call handling. The level of customer service and call volume handled by the same number of agents varies depending on the applied load distribution management methods. Customer inquiry statistics indicate that most users reach out to centers via alternative multimedia channels. A cloud-based call center structure allows for customer interaction through any channel or device, at any time of day. These processes are integrated into a virtual cloud network and scale rapidly.

In the future, chat will become the primary mode of customer interaction, while social media will serve as a communication channel for the entire organization. A modern contact center must be equipped with multiple customer communication channels, each capable of resolving issues upon first contact. Contact center workflows are integrated with control systems, utilize real-time customer data, interaction history, and resource planning. Operators have access to analytical customer information regardless of the interaction channel. This requires a platform integrated with customer databases, CRM systems, and other components that support the company's business processes.

Evaluation of the contact center model work

The main factors influencing customer satisfaction with a contact center are the speed of request handling and the convenience of accessing necessary information. One of the key objectives of a contact center is to maintain a minimum required number of agents without dropping calls. Each contact center evaluates performance according to its own standards.

To support further analysis and description of a mathematical model for network-based contact center management, we identify key performance indicators (KPIs) for operator efficiency. The primary metrics for assessing contact center performance include:

– Average Speed of Answer (ASA): measures the average time a customer waits before being connected;

– Abandon Rate (AR): the percentage of calls terminated by callers while waiting;

– Service Level (SL): the percentage of calls answered within a predefined time frame (typically not exceeding 3–4 %).

The center is structured so that 80 % of calls are connected within 20 seconds. In cases where an IVR system is deployed, this waiting period may extend to up to 2 minutes.

From a technical perspective, a typical contact center is a software-hardware integrated complex [2, 3]. The equipment handles inbound and outbound calls, records conversations, logs voice traffic, automatically recognizes caller IDs, and populates customer databases. These services are executed on servers running specialized software, often integrated with a CRM system in modern deployments. In a call center, integration with CRM (Customer Relationship Management) software empowers agents to manage customer data efficiently and automate workflows (fig.1). A CRM-enabled call center offers numerous advantages, including real-time access to client profiles, workflow automation, better analytics, omnichannel support, enhanced collaboration [4].

Fig. 1. Call center CRM features

Contact center servers host applications that support extended features. A cloud-based contact center structure consists of the following four components:

– Infrastructure as a Service (IaaS);
– Platform as a Service (PaaS);
– Software as a Service (SaaS);
– Application Integration (enterprise applications).

Traditional contact centers perform three core functions: data exchange, call distribution, and business application management. These are unified within a single operational environment. The CRM system is a pivotal component – it maintains statistical records, manages customer contact logs, and supports customer relationship management.

Statistical parameters affecting contact center performance are aggregated within the Call Management System (CMS), including:

– call type;
– number of calls within a defined time period;
– average queue length;
– average call duration;
– ratio of clients served by IVR versus live agents;
– time during which all lines were busy;
– average operator occupancy time;
– average number of operators active over a time period;
– average time;
– average duration between the end of one call and the start of the next.

Reports based on these statistics are generated at various points across the contact center model. Different vendors provide varying sets of reports, but these levels are common to most systems. Typically, the entry point to a contact center is a virtual extension number of the switchboard, not physically tied to specific hardware. It can be accessed via any method designated for internal extensions. Reporting includes agent performance, agent group efficiency, and queue statistics at the system entry points and trunk lines.

Contact center of operational management mathematical model

One of the core tasks in operating a contact center within a large infocommunications company is establishing a system for real-time resource distribution control, optimal use of network equipment, and efficient coordination between agents and support staff.

CRM platforms are vital. Their wide functionality allows the center to:

– keeping records of customer interactions;
– maintain interaction histories;
– generate reports swiftly.

This feature set gives CRM systems a competitive advantage [5]. While CRM solutions focus on customer relationship processes, ERP systems support the internal organizational structure. Together, they form a synergistic automation system.

A contact center consists of subsystems or functional blocks, which can later be supplemented with new components. The system consolidates all customer interactions, enables selection of optimal request handling algorithms, often relies on complex logic circuits, requiring significant development of programmer effort even to create the simplest scenarios a call center workflow.

To improve data transmission efficiency and terminal management, it is proposed to augment the CRM components with new functionalities aimed at solving tasks such as:

– collecting data on the functioning of operators' terminal, customer profiles, and operator profiles;

– remote resolution of emerging tasks to reduce customer servicing requests times due to the current significant increase in request volume;

– collecting data on the functioning of operators' terminal, customer profiles, and operator profiles;

– remote resolution of tasks, especially under high request volumes;

– centralized updates from the server, considering multiple contextual parameters;

– real-time transmission of terminal parameters for bandwidth optimization and call center channels overload prevention.

Operator terminal management involves their rapid, centralized, remote connection to update real-time agent occupancy data based on call profiles. A vital tool for system monitoring and operational diagnostics is the collection of statistical data and the implementation of user feedback mechanisms.

Different contact centers evaluate their performance based on the specific criteria relevant to their operational context. Nevertheless, during peak hours, performance typically degrades, and maintaining adequate service levels becomes a key performance metric. Incoming call traffic is uneven throughout the day, leading to peak load periods and, consequently, critical agent shortages – an issue common across various centers.

Overloads in a contact center result in queue length increases and call losses during processing. Therefore, it is essential to manage contact center operations proactively to prevent service degradation. Each operator group may face queues of several hundred calls, but such queue lengths should never lead to excessive customer waiting times. Call redistribution among operators is often handled manually, which contributes to queue delays. A more effective approach is a self-adjusting system that dynamically selects the optimal call redistribution algorithm based on total service time estimation.

Let us consider a mathematical model for managing a call center network that includes operator terminals and agent groups. During peak load conditions, the management system receives real-time information and compares the workload of multiple agents or operator groups based on this parameter. It then routes the call to the one with the shortest handling time. It is important to examine the service level, maximum response delays, and the call profile. The key principle of queue organization: maximize the number of processed calls with minimal staff engagement – without compromising service quality or overloading personnel.

In the model, call processing points are grouped into separate states and positioned at designated locations within the center’s hardware-software infrastructure.

The key factors to be considered in the model include:

– the number of operators involved in handling calls during the given time period;

– the volume of statistical data received from the source;

– the intensity of incoming data, which are treated as random variables.

Waiting time can be determined both at the level of each individual call and at the level of a segmented operator group.

To characterize the quality of service, the system's performance indicators include:

– the average number of requests in the system;

– the average number of requests in the queue;

– the average time a request spends on the system;

– the average waiting time before service begins.

To describe the operation of the service system, it is necessary to consider:

– the average number of devices or channels occupied by request processing;

– the utilization factor of servicing devices;

– the idle time factor of servicing devices.

Let there be m operator devices in the network, each associated with a set of parameters B defining call profiles. All devices are grouped into K operator groups.

We define:

– the assignment of parameter i to a block, where i = 1,...,k;

– the assignment of parameter j to a device, where j = 1,...,m.

The resulting call profile state vector and operator profiles are transmitted to a central server. The polling cycle time for each parameter of the j-th reference terminal equipment (TE) depends on the number of parameters involved. When the sever interacts with either a single j-th device or all m devices, and data flows randomly between them.

A mathematical model is used to describe such a hypothetical data stream. To describe a network model with such properties is a stochastic Bernoulli flow [6]. Collection of information about the parameters of control objects carried out for m control objects (CO). Matrix $B = [nxm]$ characterizes the values of all measured parameters of all control objects, i.e. for all agents, including agent profile, call profile. Each parameter $b_{ij}=m_j x n_i$, the number n_i-number of parameters in the i-th CO of the operator's workplace n, the number m_j - number of parameters in the m-th CO operator's workplace information collection systems for the control node of the network monitoring system.

Each i-th control object is characterized by a set of $B_i=\{b_{ij}\}$ parameters, each element of which is a random variable. The distribution function of each parameter $b_{ij,}$ known. The information collection system (ICS) polls each i-th object cyclically. Each request is processed in a random processing time to(i) measurement time of one object parameters, transferred to the MS and evaluated. Each request has a really fixed volume and connects resources for a certain time equal to the average value of the busy request. The data received on demand is transmitted over the network at a certain time, which consists of 2 parts, namely, the time for transmitting a constant value for this process and the time for transmitting changed data. In the general case, the delay of messages in the network [7] is

$$Ti = Twaiting + Tforwarding + Tprocessing \quad (1)$$

Then the characteristic of the delay time of messages in the network is the sum of the time of transmission of the message through the channels of the network Tt, the processing time in the switching nodes Tp and the waiting time in the queue Tq (fig. 2).

The time for which this data is processed consists of the processing time of constant and changed parameter values. Then the total random time for obtaining the value of the b_{ij} parameter from the moment the request received is

$$tpar(i,j) = to(i) + ttr(i,j) + tpr(i,j), \quad (2)$$

where to(i) – request time; ttr(i,j) – request transmission time tpr(i,j) – request processing time.

The average value and variance of the time to obtain the value of the parameter b_{ij} can be determined as follows:

$$M(tr) = M(to) + M(ttr) + M(tpr)\ , \quad (3)$$

$$(tr) = D(to) + D(ttr) + D(tpr)\ . \quad (4)$$

Messages from each i-th device of the operator arrive at random time with the distribution function [8]:

$$A_{ij}(b_{ij}) = F\{b_{ij} \leq t\} \quad (5)$$

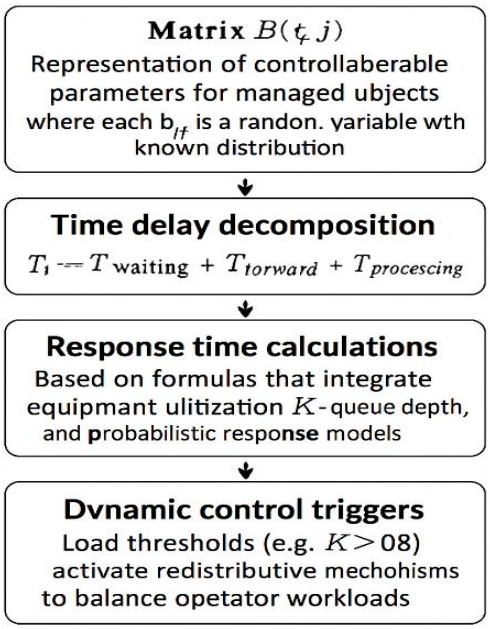


Fig. 2. Flowchart of proposed mathematical model operational control

For the parameters b_{ij} of the control objects, the exponential distribution function is characteristic $A_{ij}(b_{ij}) = 1 - exp(-qt)$. Probability of obtaining a parameter during the poll

$$P(b_{ij}) = 1 - exp[to(i) + t1 + t2] . \quad (6)$$

In the model under consideration, a random data stream comes from the server of the control node to all m devices of the call center. For a mathematical description of the data flow in the call center network configuration using the description of the stochastic Bernoulli flow, the probability of receipt in the time interval Δt of the number of responses y to send requests will be:

$$P_y(td) = C_B^y K^y (1 - K)^{B-y}, \quad (7)$$

where C_B^y – the number of combinations from the number of answers received according to the total number of parameters B: $C_B^y = B!/(y!(B - y))$ [9].

Equipment utilization factor K= tco / td , time tco - time during which the equipment was utilized for processing requests, and time td – the total time during which the equipment was available for operation (the duration of the device's operational cycle within the network). The model is a queuing system and all incoming responses that are not processed by the server are buffered [9]. Using the intensity of responses from devices in the network, we determine the average data processing time in the queuing system for the network, i.e. average processing time for device parameters of m control objects (CO):

$$rpr(i) = n \cdot to + \Sigma_{i=1}^{n}[t1(i) + t2(i)] \quad (8)$$

The total time to receive the value of all parameters b_{ij} for m=1,…,j and n =1,…,i or the time to receive a response

$$tr = (t \cdot rpr/K) \Sigma_{j=1}^{B}[j \cdot C_B^y \cdot K^y (1 - K)^{B-y}] \quad (9)$$

Average load on the host server for one polling cycle

$$Lavr = \Sigma_{j=1}^{B} P_j(td) \cdot j \quad (10)$$

Knowing the network performance parameters, it is possible to determine the time for generating a general server response based on the results of polling all support of devices.

$$K^y(1-K)^{B-y} = T \cdot t_{ob} / \Sigma_{j=1}^{B} j \cdot C_B^y \quad (11)$$

where t_{ob} – the time it takes to process the response; T – total time observation TE.

To verify the performance of the proposed model and visually demonstrate how total service time is formed during each polling cycle, we present a computational example for two matrices B, depending on the equipment utilization coefficient K, for a call center network with ten operators (fig. 3).

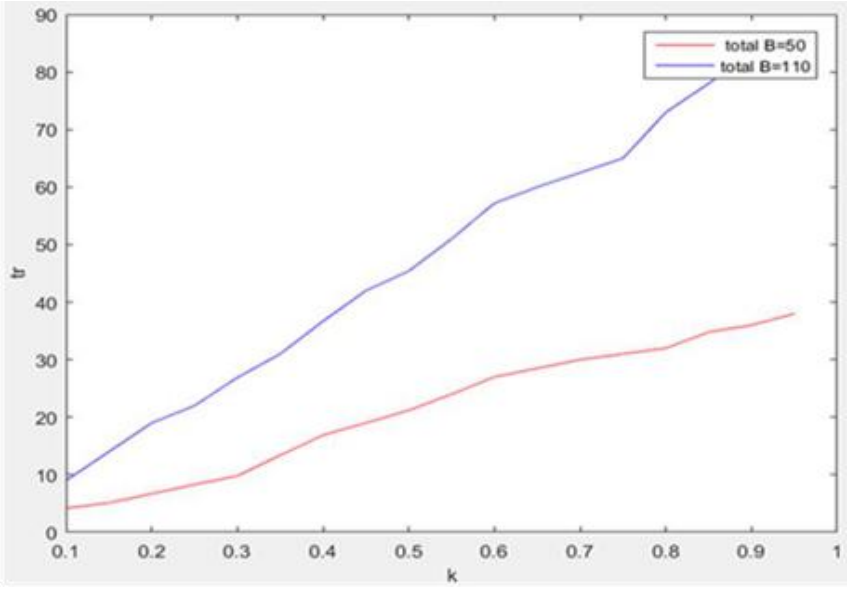


Fig. 3. The total time for obtaining the value of controlled parameters for the control model, depending on the equipment utilization factor

From the figure, it is evident that the total time required to retrieve the values of monitored parameters in the control model increases uniformly as the equipment utilization coefficient K grows. Starting from approximately 80–85 % network load, the growth in response time becomes noticeably more pronounced. This leads to an increase in queue size and necessitates redistribution of operator workloads.

Redistributing agent workload in an overloaded call center isn't merely organizational adjustment – it's a systems-level engineering solution and implementation approaches include:

– redirecting incoming calls to less-loaded agent groups or shifts based on real-time traffic and resource availability.

– reassigning calls on the fly to agents with shorter projected response times.

– shifting agents between tasks (inbound/outbound) or increasing active headcount during peak intervals, using traffic statistics like TD (Traffic Density) and Lavr (Average Waiting Time).

– rerouting calls to backup agents or voice bots when wait time or queue depth exceeds predefined thresholds.

– allocating resources ahead of time using proposed forecast models, based on expected system load K(t).

Conclusions

A contact center is a composite of subsystems that can be configured into a desired operational model. The system is designed to be supplemented with new functions or components based on statistical analysis, the performance of the call redistribution control system, and queue management models. Effective queue length control is achieved through strategic resource planning and the implementation of efficient call handling algorithms, especially under overload conditions.

The call center operator (or group of operators) represents the most resource-intensive element, which makes reducing call handling time via load balancing – through the control system – a critically important task within the model. To obtain the full set of controllable parameters, terminal equipment (TE) at the operator's workstation is used. From the moment a call is received, the management system collects the maximum amount of relevant data and interfaces with all connected information systems. This enables real-time operator workload monitoring and responsive corrective actions. Predictive load balancing use response prediction models and expected load to allocate capacity before overload uses curs.

By incorporating models and tical model for calculating the core probabilistic-temporal performance metrics – supplemented by numerical examples – the proposed framework can extend the functionality of the call processing module (CPM) without complex configurations. Ultimately, this leads to the development of a key performance indicator (KPI) system for the contact center.

References:

1. Call Center Representative Job Description: Top Duties and Qualifications 6, 2022.
2. TechNet Magazine: System Center Operations Manager 2012: Ease of expanding monitoring capabilities. URL: http://technet.microsoft.com (accessed 03 May 2021).
3. How to write a Call Centre Representative Job Description: Top Duties and Qualifications. 2025. 16 p.
4. CRM Architecture https://www.scribd.com/document/225621058/CRM-Architecture
5. CRM Design That Improves Customer Relationships. https://www.eleken.co/blog-posts/how-to-design-a-crm-system-all-you-need-to-know-about-custom-crm
6. H. Altoum, A. Ettaieb, H. Rguigui, Generalized Bernoulli–Wick differential equation, Infin. Dimens. Anal. Quantum Probab. and Relat. Top., 24, Issue 01, 2021. DOI: https://doi.org/10.1142/S0219025721500089
7. Wallace R. et al. Models of Network Delay. In: Einbeck, J., Maeng, H., Ogundimu, E., Perrakis, K. (eds) Developments in Statistical Modelling. IWSM 2024. Contributions to Statistics. Springer, Cham. https://doi.org/10.1007/978-3-031-65723-8_36.
8. Analytic Methods in Applied Probability . EditorsYu.M. Suhov .Providence, RI American Mathematical Society Pages25-36. ISBN 0-8218-3306-5. 2002 .American Mathematical Society Translations, Series 2. Vol. 207
9. S.W. Fuhrmann A. Note on the M/G/1 Queue with Server Vacations12.1984. URL: https://doi.org/10.1287/opre.32.6.1368.

Received 11.06.2025

Відомості про авторів:

Кадацька Ольга Йосипівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В.В. Поповського; Україна; e-mail: olga.kadatska@nure.ua; ORCID: https://orcid.org/0000-0002-5331-4324

Сабурова Світлана Олександрівна – Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В.В. Поповського; Україна; e-mail: svitlana.saburova@nure.ua; ORCID: https://orcid.org/0000-0003-2214-2440

УДК 537.874.46 **DOI:10.30837/rt.2025.3.222.19**

Є.М. ОДАРЕНКО, д-р фіз.-мат. наук, С.О. ЮХНО, Є.В. СУЛІМА,
О.С. ГНАТЕНКО, канд. фіз.-мат. наук

ХАРАКТЕРИСТИКИ ВЛАСНИХ РЕЖИМІВ ФОТОННО-КРИСТАЛІЧНОГО ХВИЛЕВОДУ З РЕШІТКОЮ КАГОМЕ

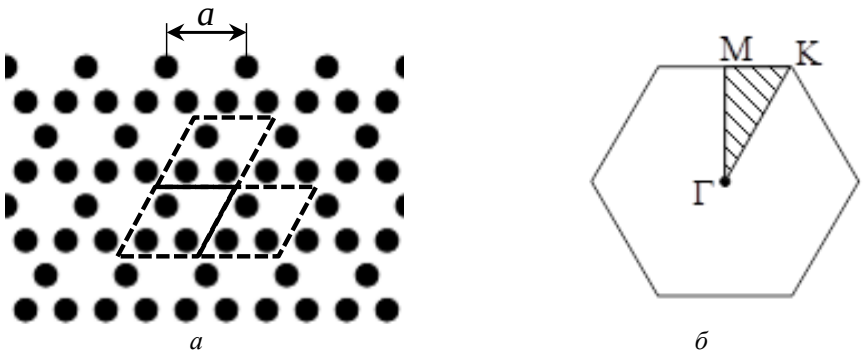
Вступ

Фотонно-кристалічні хвилеводи та резонатори складають основу для побудови різноманітних пристроїв оптичного, терагерцового та мікрохвильового діапазонів. Функціональність цих пристроїв забезпечується механізмами локалізації електромагнітної енергії в локальних або лінійних дефектах періодичності фотонних кристалів. Зазвичай виділяються два основних механізми: повне внутрішнє відбиття та наявність фотонних заборонених зон [1, 2]. Важливою особливістю фотонно-кристалічних хвилеводів є можливість використання пустотілого хвилеводного каналу, що дозволяє реалізовувати інтенсивну взаємодію електромагнітних хвиль з різноманітними речовинами, в тому числі активними та нелінійними, а також транспортувати лазерне випромінювання. В цьому випадку хвилеводний режим забезпечується лише механізмом фотонної забороненої зони. Характеристики забороненої зони визначаються розмірами елементів, які формують фотонний кристал, їх матеріальними параметрами та конфігурацією решітки, у вузлах якої ці елементи розташовані. Одним із перспективних варіантів конфігурації решітки фотонних кристалів є решітка Кагоме, яка переважно застосовується в оболонках фотонно-кристалічних волоконних хвилеводів [3, 4]. Завдяки унікальним електродинамічним характеристикам, такі хвилеводи застосовуються в системах формування суперконтинууму [5], дозволяють отримувати кероване ультрафіолетове випромінювання [6], забезпечують широкосмугову передачу даних з низьким рівнем втрат [7], використовуються для побудови сучасних волоконних гіроскопів [8, 9]. Характерною відмінністю фотонно-кристалічних хвилеводів з решіткою Кагоме є наявність специфічного для них механізму локалізації енергії, який пов'язаний із слабкою взаємодією мод пустотілого хвилеводного каналу та мод фотонно-кристалічної оболонки [7, 10, 11]. Саме цей механізм обумовлює широкосмуговість таких волоконних хвилеводів в умовах, коли фотонні заборонені зони відсутні. В той же час залишається відкритим питання існування такого специфічного механізму локалізації енергії в умовах наявності фотонних заборонених зон.

В роботі досліджуються механізми реалізації власних режимів фотонно-кристалічного хвилевода з решіткою Кагоме у випадку, коли оболонка пустотілого хвилеводного каналу має фотонні заборонені зони. Розглянуто структуру оболонки, сформована решіткою з діелектричних циліндрів.

Дисперсійні характеристики фотонного кристалу з решіткою Кагоме

Розглянемо нескінченний двовимірний фотонний кристал, який сформований діелектричними циліндрами, розташованими у вузлах решітки Кагоме (рис. 1, *а*). Нормований радіус циліндрів r/a = 0,15 (a – період структури). Діелектрична проникність матеріалу циліндрів ε = 11,5. Штриховим контуром на рис. 1, *а* позначено три елементарні комірки періодичної структури. Трансляція комірки здійснюється уздовж двох напрямків, кут між якими становить $\pi/3$, з періодом a. Розрахунки дисперсійних характеристик фотонного кристалу проводилися з використанням програмного пакету MIT Photonic Bands, що вільно розповсюджується [12]. В цьому пакеті використовується метод розкладання по плоским хвилям, який добре апробований в задачах дослідження різноманітних періодичних структур. Розрахунки проводяться в межах першої зони Бріллюена досліджуваної решітки, яка представлена на рис. 1, *б*. Буквами на рисунку позначено точки високої симетрії, які обмежують так звану незвідну зону Бріллюена (irreducible Brillouin zone).



а *б*

Рис. 1. *а* – схема фотонного кристалу з решіткою Кагоме; *б* – перша зона Бріллюена з позначеними точками високої симетрії

На рис. 2 представлено результати чисельних розрахунків дисперсійної діаграми двовимірного фотонного кристалу з решіткою Кагоме. Уздовж осі ординат відкладена нормована частота, яка обчислюється як $\omega a/(2\pi c)$, де c – це швидкість світла у вакуумі. Фактично нормована частота представляє собою відношення періоду структури до довжини хвилі випромінювання у вакуумі. Дисперсійна діаграма на рис. 2 побудована для ТМ поляризації випромінювання. В цьому випадку існує лише одна координатна компонента електричного поля, спрямована уздовж діелектричних циліндрів. Компоненти магнітного поля знаходяться у площині рисунку. Для ТЕ поляризації випромінювання (з одною компонентою магнітного поля) фотонні заборонені зони для досліджуваної структури практично відсутні.

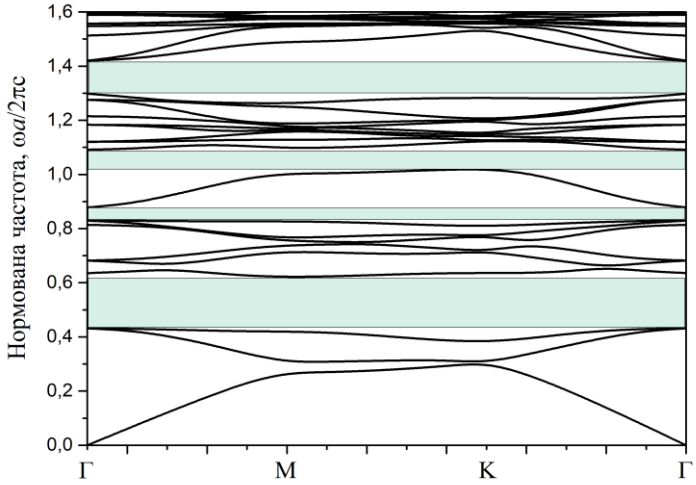


Рис. 2. Дисперсійна діаграма фотонного кристалу з решіткою Кагоме

Горизонтальними смугами на рис. 2 позначено фотонні заборонені зони, тобто інтервали значень нормованої частоти, в яких неможливе розповсюдження електромагнітних хвиль в фотонному кристалі в будь-якому напрямку на площині зони Бріллюена. В даному випадку реалізується чотири заборонені зони, три з яких є відносно широкими. Крім того, з рис. 2 видно, що для даного фотонного кристалу існують заборонені зони при значеннях нормованої частоти, більших за одиницю. Це означає, що в даному випадку довжина хвилі випромінювання є меншою за період структури, але залишається більшою за відстань між найближчими циліндрами. Слід також відзначити, що дисперсійні властивості фотонного кристалу суттєво залежать від нормованого радіусу циліндрів та діелектричної проникності матеріалу, з якого вони виготовлені. Це типова ситуація для таких структур [13]. Зокрема, при збільшенні значення параметру r/a кількість фотонних заборонених зон та їх ширина зменшуються.

Оскільки практично всі фотонно-кристалічні хвилеводні та резонансні пристрої працюють в межах фотонних заборонених зон, то застосування решітки Кагоме для формування таких пристроїв може розширити їх функціональні можливості.

Фотонно-кристалічний хвилевід

Зазвичай фотонно-кристалічні хвилеводи формуються шляхом внесення дефекту періодичності в структуру. На основі двовимірних фотонних кристалів існують два різновиди хвилеводів: волоконні (фотонно-кристалічні волоконні хвилеводи), в яких хвилі розповсюджуються в напрямку, в якому структура є регулярною, і лінійні фотонно-кристалічні хвилеводи, які формуються лінійним дефектом періодичності. Цей лінійний дефект виконує функцію хвилеводного каналу. В роботі розглядається другий тип хвилеводу, схема якого представлена на рис. 3. Зрозуміло, що для решітки Кагоме існують кілька варіантів внесення дефекту періодичності в структуру. Результати дослідження одного з таких варіантів показали незвичайні електродинамічні властивості такого хвилеводу, які дозволяють отримувати надзвичайно малі значення групової швидкості хвиль, і, відповідно, реалізовувати вузькосмугову фільтрацію [14].

З рис. 3 видно, що в даному випадку дефект періодичності фотонного кристалу реалізується видаленням одного горизонтального ряду діелектричних циліндрів. В результаті формується пустотілий хвилеводний канал з періодичними границями, сформованими двома рядами циліндрів. Відстань між центрами цих граничних циліндрів дорівнює $a/2$. Таким чином, досліджуваний хвилевід відрізняється від фотонно-кристалічних хвилеводів з квадратними, трикутними та шестикутними решітками, де період елементів на границях хвилеводного каналу зазвичай співпадає з періодом фотонного кристалу.

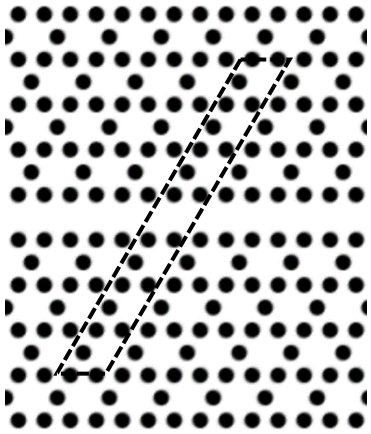


Рис. 3. Схема фотонно-кристалічного хвилеводу з позначеною надкоміркою

Слід відзначити, що наявність дефекту періодичності порушує правила застосування методу розкладання по плоским хвилям. Структура перестає бути періодичною в одному напрямку. Тому для моделювання таких структур застосовується метод так званої надкомірки (supercell) [15, 16]. Тобто формується нова періодична структура з більшою коміркою, для якої можливе застосування методу розкладання по плоским хвилям. Ця надкомірка показана на рис. 3 пунктирним контуром. В результаті отримуємо періодичну структуру, яка містить нескінченну кількість періодично розташованих хвилеводних каналів, відстань між якими становить шість періодів базового фотонного кристалу. Тепер при розрахунку власних режимів нового фотонного кристалу необхідно враховувати той факт, що вірогідні результати будуть отримуватися лише для режимів, які характеризуються високим ступенем електромагнітної ізоляції між сусідніми дефектами періодичності фотонного кристалу.

Результати розрахунків дисперсійної діаграми фотонно-кристалічного хвилеводу з решіткою Кагоме представлені на рис. 4. Уздовж осі абсцис відкладені значення нормованого поздовжнього хвильового числа $\beta a/2\pi$ в межах першої зони Бріллюена. Порівняння з рис. 2 показує, що в кожній забороненій зоні фотонного кристалу, який формує оболонку хвилеводного каналу, реалізується одна хвилеводна мода. Дисперсійні криві, які відповідають цим модам, позначені червоними пунктирними кривими. В даному випадку одномодовий режим роботи хвилеводу є цілком очікуваним, оскільки ширина хвилеводного каналу мінімальна для даної конфігурації. Крім того, цей канал є пустотілим, що виключає з розгляду механізм локалізації електромагнітної енергії, обумовлений повним внутрішнім відбиттям.

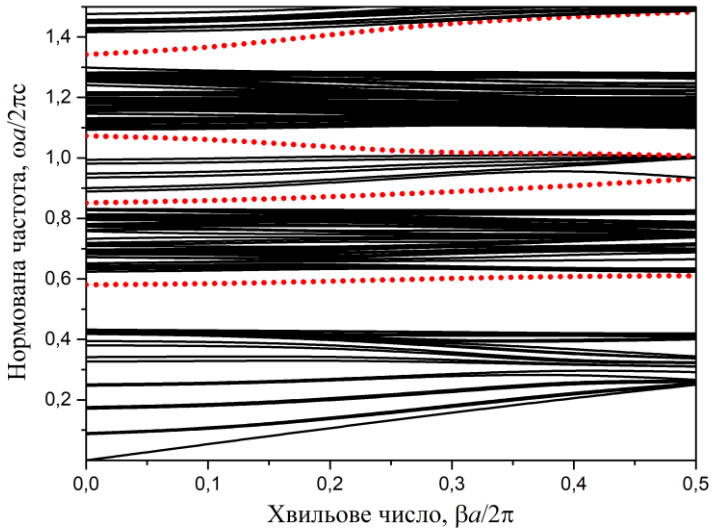


Рис. 4. Дисперсійна діаграма фотонно-кристалічного хвилеводу

На рис. 5 представлено результати розрахунків просторових розподілів електричного поля, які відповідають наведеному вище критерію вірогідності власних режимів структури, з відповідними значеннями власних частот. Тобто для цих розподілів характерною є локалізація енергії електромагнітного поля в хвилеводних каналах з відповідною відсутністю зв'язку між сусідніми каналами. Результати на рис. 5 отримані для значення нормованого поздовжнього хвильового числа $\beta a/2\pi = 0{,}2$. Видно, що в досліджуваному діапазоні частот реалізуються п'ять власних режимів фотонно-кристалічного хвилеводу. Чотири з них відповідають фотонним забороненим зонам оболонки хвилеводного каналу. Але один з визначених режимів (рис. 5, *г*) має власну частоту $\omega a/(2\pi c) = 1{,}212$, яка знаходиться за межами фотонних заборонених зон. Тому можна зробити припущення, що в цьому випадку реалізується інший

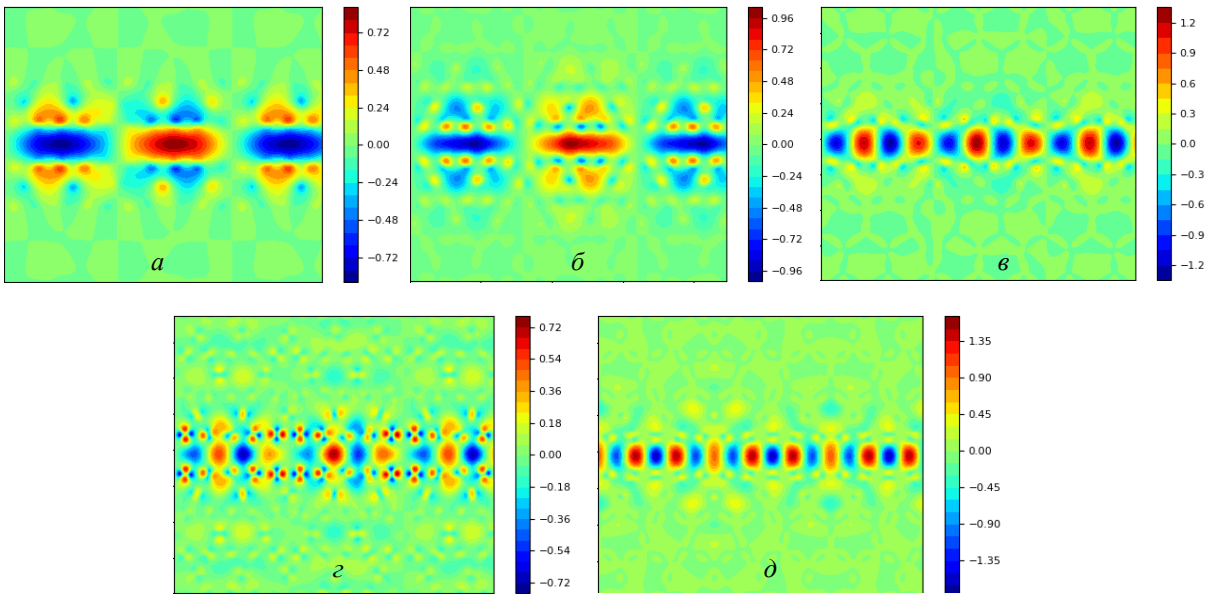


Рис. 5. Просторові розподіли електричного поля для хвилеводних мод фотонно-кристалічної структури; $\omega a/(2\pi c) = 0{,}592$; 0,871; 1,036; 1,212; 1,408 для *а* – *д* відповідно

механізм локалізації енергії в хвилеводному каналі, який відсутній для інших конфігурацій фотонних кристалів. Дійсно, при дослідженні фотонно-кристалічних волоконних хвилеводів з решіткою Кагоме було виявлено додатковий механізм, який отримав назву “пригнічений зв'язок” (inhibited coupling) [7, 10, 11]. Для цього механізму локалізації енергії характерне суттєве послаблення зв'язку між модами хвилеводного каналу (осердя волокна) та модами оболонки. Це послаблення відбувається через порушення узгодження між поперечними компонентами полів мод каналу та оболонки, яке обумовлене швидкими фазовими осциляціями полів мод оболонки. Слід враховувати, що означений механізм пригніченого зв'язку був реалізований для решітки Кагоме, сформованої тонкими прямолінійними перетинками між вузлами решітки. Така структура є характерною для волоконних хвилеводів. В цьому випадку поле мод оболонки було сконцентровано саме в цих перетинках і швидко загасало при віддаленні від них у вільний простір. Якщо решітка Кагоме сформована з окремих діелектричних циліндрів, розташованих у вузлах цієї решітки, то поле мод оболонки не концентрується переважно в циліндрах (рис. 5, *г*) і на даному етапі досліджень можна говорити лише про вплив загальної симетрії фотонного кристалу на його унікальні електродинамічні властивості. Дійсно, в фотонно-кристалічних волоконних хвилеводах зі звичайною трикутною решіткою, яка формується тонкими перетинками, реалізуються лише два механізми локалізації енергії – повне внутрішнє відбиття та фотонна заборонена зона. Отже наявність механізму пригніченого зв'язку обумовлена не формою елементів, які формують фотонний кристал, а саме конфігурацією решітки з цих елементів.

На рис. 6 представлено відповідні до рис. 5 просторові розподіли нормованої інтенсивності електричного поля хвилеводних мод в перетині, площина якого перпендикулярна поздовжній осі фотонно-кристалічного хвилеводу. Вертикальними штриховими лініями позначені границі пустотілого хвилеводного каналу, тобто місце розташування граничних рядків діелектричних циліндрів. Уздовж осі абсцис відкладена відстань від центру каналу *D*,

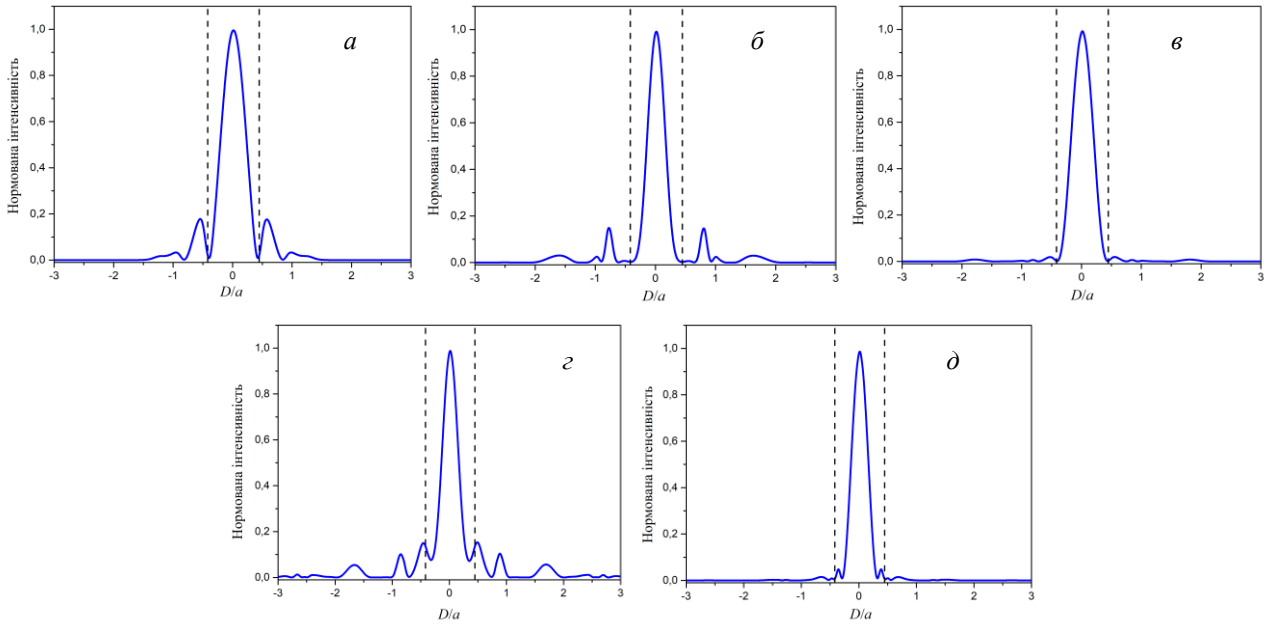


Рис. 6. Просторові розподіли нормованої інтенсивності електричного поля

нормована на період фотонного кристалу *a*. Графіки на рис. 6, *а*, *б*, *в*, *д* відповідають модам, розташованим у межах фотонних заборонених зон оболонки хвилеводу. В цих випадках реалізується практично експоненційне загасання інтенсивності поля при заглибленні в оболонку хвилеводного каналу. Це звичайний результат для механізму фотонної забороненої зони.

В цих зонах Блохівське хвильове число має уявну частину, яка обумовлює експоненційне загасання поля хвилі при розповсюдженні в фотонному кристалі. В результаті на границях надкомірки інтенсивність поля практично дорівнює нулю. На рис. 6, *г* спостерігається інша ситуація. Окрім також доволі суттєвого загасання поля при віддаленні від хвилеводного каналу спостерігаються невеликі осциляції інтенсивності поля в межах надкомірки фотонного кристалу. Результати додаткових розрахунків показують, що ці осциляції зберігаються при збільшенні відстані між сусідніми хвилеводами. Це свідчить про існування в цьому випадку окремих мод оболонки хвилеводу подібно до фотонно-кристалічних волоконних хвилеводів з решіткою Кагоме.

Висновки

Електродинамічні характеристики лінійного фотонно-кристалічного хвилеводу чисельно розраховані з використанням методу розкладання по плоским хвилям, який імплементовано в пакет MIT Photonic Bands. Оболонка пустотілого хвилеводного каналу сформована діелектричними циліндрами, розташованими у вузлах решітки Кагоме. Дисперсійні характеристики хвилеводу демонструють існування хвилеводних мод в фотонних заборонених зонах оболонки хвилеводу, а також за їх межами, що дозволяє зробити висновок про реалізацію додаткового механізму локалізації електромагнітної енергії в пустотілому хвилеводному каналі. Виходячи з результатів розрахунків просторових розподілів електричного поля власних режимів фотонно-кристалічного хвилеводу, зроблено припущення про спорідненість цього додаткового механізму та відомого для фотонно-кристалічних волокон з решіткою Кагоме механізму послабленого електродинамічного зв'язку між модами пустотілого осердя волокна та фотонно-кристалічної оболонки. Отримані результати дозволяють розробляти функціональні пристрої на основі фотонно-кристалічних хвилеводів з решіткою Кагоме, зокрема, для вузькосмугової фільтрації сигналів в сучасних оптичних телекомунікаційних системах.

Список літератури:

1. Joannopoulos J.D., Villeneuve P.R., and Fan S. Photonic crystals: putting a new twist on light // Nature. 1997. 386. P. 143–149. doi:10.1038/386143a0.

2. Soukoulis C.M. Photonic Crystals and Light Localization in the 21st Century. Springer Dordrecht, 2001. doi:10.1007/978-94-010-0738-2.

3. Saraceno C. J. et al. Kagome-type hollow-core photonic crystal fibers for beam delivery and pulse compression of high-power ultrafast lasers // Proc. SPIE. 2015. 9346; Components and Packaging for Laser Systems, 93460Z. doi:10.1117/12.2080749.

4. Couny F., Benabid F., and Light P.S. Large pitch kagome-structured hollow-core photonic crystal fiber // Optics Letters. 2006. Vol. 31. P. 3574–3576. doi:10.1364/OL.31.003574.

5. Ferreira M.F.S. Supercontinuum Generation // Solitons in Optical Fiber Systems. 2022. Wiley. P. 337–368. doi: 10.1002/9781119506669.ch17.

6. Rodrigues S.M.G., Facão M., and Ferreira M.F.S. Ultraviolet light generation in gas-filled Kagome photonic crystal fiber // Fiber and Integrated Optics. 2015. Vol. 34(1–2). P. 76–89. doi:10.1080/01468030.2014.1001092.

7. Debord B., et al. Ultralow transmission loss in inhibited-coupling guiding hollow fibers // Optica. 2017. Vol.4. P. 209–217. doi:10.1364/OPTICA.4.000209.

8. Suo X., Yu H., Li J., and Wu, X. Transmissive resonant fiber-optic gyroscope employing Kagome hollow-core photonic crystal fiber resonator // Optics Letters. 2020. Vol.45. P. 2227–2230. doi:10.1364/OL.388274.

9. Odarenko E.N., and Hnatenko O.S. Photonic crystal fibers with triangular and Kagome structures for fiber optic gyroscopes // Journal of Nano- and Electronic Physics. 2024. Vol.16(6). P. 06029. doi:10.21272/jnep.16(6).06029.

10. Couny F., et al. Generation and photonic guidance of multi-octave optical-frequency combs // Science. 2007. Vol. 318. P. 1118. doi:10.1126/science.1149091.

11. Wang Y.Y., et al. Low loss broadband transmission in hypocycloid-core Kagome hollow-core photonic crystal fiber // Optics Letters. 2011. Vol. 36. P. 669–671. doi:10.1364/OL.36.000669.

12. Johnson S.G. and Joannopoulos J. D. Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis // Optics Express. 2001. Vol. 8(3). P. 173–190. doi:10.1364/OE.8.000173.

13. Joannopoulos J. D. et al. Photonic Crystals: molding the flow of light. Princeton University Press, 2008.

14. Schulz S.A., Upham J., O'Faolain L., and Boyd R.W. Photonic crystal slow light waveguides in a kagome lattice // Optics Letters. 2017. Vol. 42(16). P. 3243–3246. doi:10.1364/OL.42.003243.

15. Zhi W., Guobin R., Shuqin L., and Shuisheng J. Supercell lattice method for photonic crystal fibers // Optics Express. 2003. Vol. 11(9). P. 980–991. doi:10.1364/OE.11.000980.

16. Cerjan A., and Fan Sh. Complete photonic bandgaps in supercell photonic crystals // Physical Review A. 2017. Vol. 96(5). P. 051802. doi:10.1103/PhysRevA.96.051802.

Надійшла до редколегії 17.07.2025

Відомості про авторів:

Одаренко Євген Миколайович – д-р фіз.-мат. наук, Харківський національний університет радіоелектроніки, професор кафедри фізичних основ електронної техніки, Україна; e-mail: yevhen.odarenko@nure.ua, ORCID: http://orcid.org/0000-0001-7656-0440

Юхно Сергій Олександрович – Харківський національний університет радіоелектроніки, аспірант кафедри фізичних основ електронної техніки, Україна; e-mail: serhii.iukhno@nure.ua; ORCID: https://orcid.org/0009-0004-7930-9816

Суліма Євген Валерійович – Харківський національний університет радіоелектроніки, аспірант кафедри фізичних основ електронної техніки, Україна; e-mail: yevhen.sulima@nure.ua; ORCID: https://orcid.org/0009-0008-3233-2188

Гнатенко Олександр Сергійович – канд. фіз.-мат. наук, Харківський національний університет радіоелектроніки, завідувач кафедри фізичних основ електронної техніки, Україна; e-mail: oleksandr.hnatenko@nure.ua; ORCID: https://orcid.org/0000-0001-7722-0923

УДК 621. 373.826 **DOI:10.30837/rt.2025.3.222.20**

S.S. HERASYMOV, O.S. HNATENKO, Ph.D.

SILVER FILM AND DISTRIBUTED BRAGG REFLECTOR MICROCAVITY: MULTILAYERED LASER MODEL THRESHOLD ANALYSIS

Introduction

With the growing demand for miniaturized and integrated optical systems, traditional solid-state lasers are being increasingly replaced by microlasers. These compact devices offer superior monochromaticity, higher beam quality, lower noise, and seamless integration with micro- and nanoelectronics systems, unlocking new perspectives in technologies [1–3]. A significant category of microlasers includes the VCSEL (Vertical-Cavity Surface-Emitting Laser) and photonic crystal surface-emitting laser (PCSEL) [4]. VCSEL and PCSEL emit visible or infrared light perpendicularly to the surface of the fabricated wafer. They are commonly used in data communication and sensing applications and are becoming increasingly popular in light detection and ranging (LiDAR) applications due to their energy efficiency and high precision [5–7].

Rare-earth (RE) ions-activated micro- and nanomaterials have emerged as promising candidates for next-generation microlasers, offering advantages such as cost-effective fabrication, high environmental stability, and broad spectral coverage from ultraviolet to mid-infrared [8]. Their high photoluminescence quantum yield (PLQY) and low surface defect density enhance their optical performance. Recent advancements in RE ions-activated luminescent materials, coupled with the development of novel micro- and nanocavities for optical feedback, have significantly improved laser efficiency and stability [9–11]. These innovations have driven substantial progress in the field, paving the way for highly efficient, miniaturized laser sources with applications in optical communication, sensing, and integrated photonics [6, 10, 12]. The rare-earth doped yttrium aluminum garnet (YAG) crystal is one of most popular, and reliable lasing gain material in conventional solid-state bulk lasers. By doping YAG with different rare-earth elements, it enables laser emission over a broad wavelength range. For example, the Nd:YAG crystal, featuring a four-level laser system, offers an exceptionally low lasing threshold, making it ideal for on-chip applications, nonlinear optics, and biosensing [12]. It has been widely explored as a waveguide-based on-chip light source.

Layered microlasers are an emerging class of compact laser devices that use multiple stacked layers of optical and gain materials to enhance performance and tunability. By carefully designing these layers, researchers achieve improved photon confinement, lower lasing thresholds, and broader spectral control, making them highly efficient and versatile. Common structures include distributed Bragg reflectors (DBRs), dielectric-metal-dielectric (DMD) stacks, and heterostructures incorporating quantum wells or 2D materials. These microlasers are also promising for applications in on-chip photonics, optical communication, LiDAR, and biosensing, as they can be seamlessly integrated into micro- and nanoscale systems [13, 14]. Recent advancements, such as perovskite-based microlasers, rare-earth-doped structures, and graphene-enhanced designs, have further expanded their potential by enabling tunable, high-quality lasing across ultraviolet to infrared wavelengths. Their small size, low power consumption, and compatibility with silicon-based platforms position layered microlasers as key components in next-generation photonic and optoelectronic technologies [8].

In microlaser structures, silver is commonly employed due to its superior plasmonic properties, which enhance light confinement and amplification at the nanoscale. For instance, integrating silver nanorings with silica microcavities has been shown to achieve high-quality hybrid plasmonic modes, facilitating efficient lasing in compact designs [15]. The pure silver has the highest reflectivity of all noble metals, particularly in the visible and near-infrared regions [16]. Additionally, the use of atomically smooth epitaxial silver films in plasmonic nanocavities has demonstrated reduced optical losses, leading to improved performance in nanolaser applications [17]. The noble metal, including silver, cavities are widely investigated and have some drawbacks as sizable ohmic losses.

Alternative reflector could be dielectric DBR, consist of high and low refractive indices layers. Their high reflectivity helps reduce radiation losses, thereby lowering the lasing threshold. These structures are commonly fabricated using the reliable method of metal-organic chemical vapor deposition (CVD), which enables precise control of both layers' composition and thickness. DBRs offer a compact and flexible design, allowing for tailored reflectivity and wavelength properties to suit specific applications [13]. In general, modern microlaser designs use distributed Bragg reflectors and noble metal cavity components to improve their efficiency and overall performance.

Understanding the lasing threshold conditions is crucial in laser design and development. Since laser emission is influenced by multiple physical mechanisms, full modeling presents a complex theoretical challenge. By neglecting all non-electromagnetic effects and focusing on electromagnetic fields, we can simplify the analysis and use a source-free linear Maxwell's equations, along with boundary conditions at edges and the radiation condition at infinity. The key aspect is considering the presence of the active region. This method, known as the Lasing Eigenvalue Problem (LEP), is described in detail in [18, 19]. This approach allows examination of specific modes' threshold values of the gain index along with its emission frequencies. These two parameters are treated as ordered, mode-specific two-component eigenvalues, which include the wavelength and the threshold gain index. Over the past two decades, the LEP has been applied to various material and shape micro- and nanolaser models to analyze their threshold conditions [20–23]. Recently, we have used such approach to investigate the host gain materials and their thickness impact in the laser microcavity [24] and the noble metal-walled cavity microlaser configuration [25].

In this paper, we use the LEP approach to analyze the thresholds conditions for the modes of the active cavity, sandwiched between a finite-thickness silver film and substrate made of full dielectric DBR.

1-D lasing eigenvalue problem formulation and basic equations

Laser modes on the stationary emission threshold, in the LEP approach, can be considered as open cavity's natural modes with natural frequencies, which are real-valued. As is known from Poynting theorem, purely real natural frequencies are forbidden in passive open cavities, however, the presence of active regions permits such frequencies. Then, we seek these laser modes as solutions of the source-free, time-harmonic Maxwell equations defined over an infinite spatial domain, subject to appropriate boundary and radiation conditions.

Fig. 1 illustrates the one-dimensional microlaser configuration under consideration, comprising an active layer filled with gain material, a silver superstrate film, and a substrate DBR composed of alternating dielectric layer. The time dependence is presented as $e^{-i\omega t}$, where $\omega > 0$, positive cyclic frequency. Then, we assume that the gain material has a complex relative dielectric permittivity $\varepsilon_c = \varepsilon' + i\varepsilon''$ featuring a negative imaginary part, which is responsible for gain effect in this consideration. The silver film dielectric permittivity is denoted as $\varepsilon_f(\lambda)$ while all materials are considered nonmagnetic. In laser science, the gain material is traditionally characterize1d020. equivalently by its refractive index $\nu = \sqrt{\varepsilon_c} = \alpha_c - i\gamma$. Here, the cavity is supposed uniformly active, so that $\alpha_c = \mathrm{Re}\,\nu > 0$ is the known chosen material refractive index, and $\gamma = -\mathrm{Im}\,\nu > 0$, is unknown threshold gain index value. The silver film refractive index $\alpha_f(\lambda) = \sqrt{\varepsilon_f}$ has positive imaginary part responsible for the ohmic losses. We obtain these values from the experimental data presented in [26], using the Akima spline interpolation method for smooth estimation. It is well established that the bulk refractive index can be applied when the size of the metal particles or the film thickness, in our case, exceeds the electron path without collision in the metal, which is approximately 3□nm. For modelling, we assume a frequency-independent gain index for wider versatility.

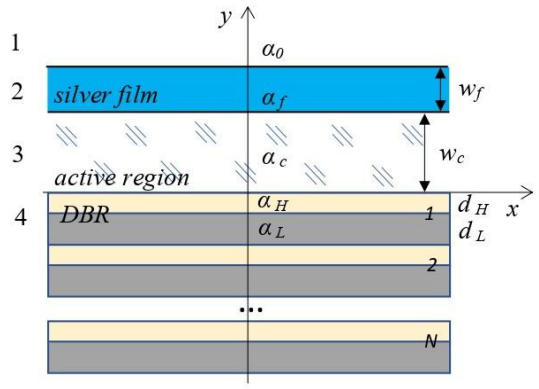


Fig. 1. Schematic laser model composed of a fully active region, covered by a finite-thickness silver film and supported by a finite distributed Bragg reflector (DBR)

The field function off the layer's boundaries must satisfy the Helmholtz equation as presented in (1). There, as *E*-field does not vary along x and z-axes, we have to look for a scalar function *E*(*y*), which represents the E_z field component.

$$\left(\Delta + k_D^2\right) E(y) = 0, \quad y_D < y < y_{D+1}, \tag{1}$$

where $D = 1, \ldots Q + 1$ is the domain number counted from the air halfspace above the silver film and y_D is the boundary coordinate, so that $Q = 4+2N$ for the cavity with DBR, N is the number of layer pairs. Besides, $k_1 = k = \omega / c$ is the free space wavenumber, $k_2 = k_f = k\alpha_f$, $k_3 = k_c = k\nu$ in the cavities. For DBR, the wavenumber values k_n, where $n \geq 4$, alternate between the higher and lower reflective index layers. In this paper, the adjacent to the cavity layer of DBR has higher refractive index, therefore, $k_{4+2n} = k_H = k\alpha_H$, $k_{5+2n} = k_L = k\alpha_L$, $n = 1,...,N-1$, and the DBR opens to the free halfspace, i.e. $k_{2N+4} = k$.

The electromagnetic field tangential components must remain continuous across the interfaces of material.

$$E^{(\mathrm{D})}(y) = E^{(\mathrm{D}+1)}(y)\Big|_{y=y_D}, \qquad \frac{\partial}{\partial y}\left(\frac{1}{\mu_D} E^{(\mathrm{D})}(y) - \frac{1}{\mu_{D+1}} E^{(\mathrm{D}+1)}(y)\right)\Bigg|_{y=y_D} = 0 \quad , \tag{2}$$

where $y_1 = w_c + w_f$, $y_2 = w_c$, $y_3 = 0$, and additional boundaries, y_n with $n = 4, \ldots, 3+2N$, correspond to the interfaces between the DBR layers, and $y_{4+2N} = -(w_L + w_H)N$. Additionally, the field function must exhibit outgoing wave behaviour in accordance with the radiation condition, as follows

$$E(y) = C\, e^{ik|y|}, \quad y \to \pm\infty \tag{3}$$

To analyse the threshold conditions using the LEP, we seek the eigenvalues of equations (1) – (3) as pairs of real numbers (λ_m, γ_m): the first representing the emission wavelengths, and the second corresponding to the gain index threshold values of laser modes.

By substituting the relevant electric field expressions into the boundary conditions at the interfaces and performing some algebraic manipulations, the LEP simplifies to the following transcendental equation:

$$\Phi(k,\gamma) = e^{-i2k(\alpha_c - i\gamma)w_c} - \frac{R_{DBR}\left(e^{i2k\alpha_M w_f} R_{12} - R_{32}\right)}{e^{i2k\alpha_M w_f} R_{12}R_{32} - 1} = 0\,, \tag{4}$$

where

$$R_{jp} = \left(\alpha_j - \alpha_p\right)\left(\alpha_j + \alpha_p\right)^{-1},\ j \neq p\ , \tag{5}$$

are the reflection coefficients from the interfaces between the corresponding domains.

For deriving the reflection coefficient of DBR, R_{DBR}, we apply the Transfer Matrix Method (TMM). As detailed in [19], the corresponding transfer matrix, $\mathbf{M}$, for L-layer dielectric structure can be obtained as a product of transmittance matrices between the neighbouring layers, $\mathbf{T}(\alpha_{i+1}/\alpha_i)$, and propagation matrices in each layer $\mathbf{P}(\alpha_i k d_i)$, namely

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = \mathbf{T}\left(\frac{\alpha_1}{\alpha_c}\right) \cdot \prod_{i=1}^{L-1} \mathbf{P}(\alpha_i k d_i)\mathbf{T}\left(\frac{\alpha_{i+1}}{\alpha_i}\right) \cdot \mathbf{P}(\alpha_L k d_L)\mathbf{T}\left(\frac{\alpha_{air}}{\alpha_L}\right), \tag{6}$$

where $\mathbf{T}(z) = \frac{1}{2}\begin{pmatrix} 1+z & 1-z \\ 1-z & 1+z \end{pmatrix}$, $\mathbf{P}(x) = \begin{pmatrix} e^{-ix} & 0 \\ 0 & e^{ix} \end{pmatrix}$

Afterward, the reflection and transmission coefficients of DBR as follows

$$R_{DBR} = \frac{m_{21}}{m_{11}},\ T_{DBR} = \frac{1}{m_{11}} \tag{7}$$

Therefore, we seek the LEP eigenvalues, which correspond to the roots of the equation (4). To determine their precise locations, a gradient-based iterative search algorithm is used.

Active microcavity featuring silver film and distributed Bragg reflector

As mentioned previously, we studied the silver-walled microcavity laser structure in [25], it consisted of a top silver film, a gain material slab, and a bottom silver substrate. To enhance microcavity efficiency and performance, we compare the reflection coefficients of silver half-spaces and a DBR of various number of alternating layer pairs used as a substrate reflector.

$$\alpha_H d_H = \alpha_L d_L = \frac{\lambda_0}{4}, \tag{8}$$

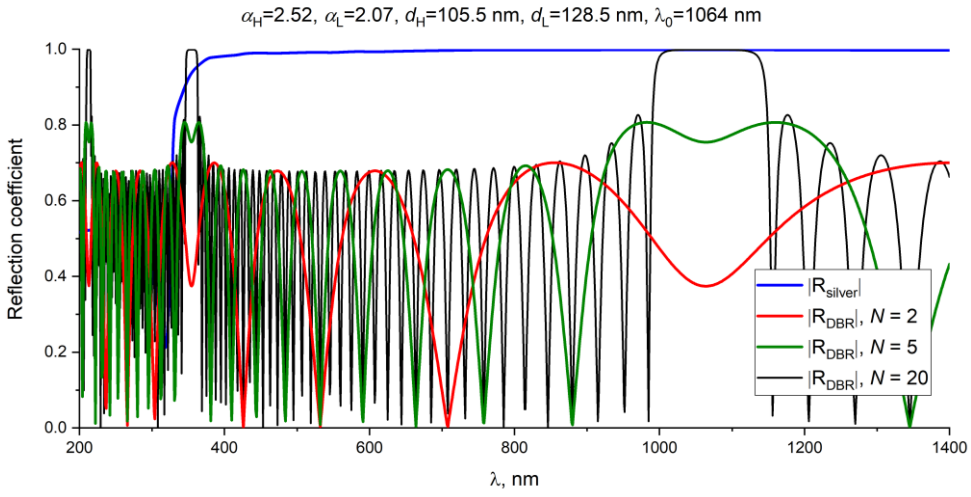


Fig. 2. The absolute values of reflection coefficient of several cavity substrates: silver halfspace and DBR with N = 2, 5, and 20 layer pairs

Their reflection coefficients comparison in the visible and near infra-red ranges is shown in Fig. 2. The DBR is designed for the wavelength of around 1064 nm. The results are presented for 2, 5, and 20 pairs of dielectric layers (e.g. SiO and TiO_2) with α_L = 2.07 and α_H = 2.52, each layer thickness is a quarter of that wavelength, as suggest (8), in the corresponding material, so that w_L = 128.5 nm and w_H = 105.5 nm.

A 2-pair and 5-pair DBR are clearly improper for good reflection, as they yield a reflection coefficient of only under 0.8. As one can see, the widest 20-pairs DBR band gap is located between 1000 and 1150 nm, which corresponds to a high-reflection band. The other much narrower band gaps can be seen between 340 and 360 nm and around 220 nm. The reflection coefficient of silver halfspace at 1064 nm give $|R|_{\text{Ag}}$ = 0.998. In contrast 20-pair DBR has reflection around $|R|_{\text{DBR}}$ = 0.9999 at this wavelength, making it more effective for providing optical feedback.

For the selected configuration, we compute the locations of the LEP eigenvalues and display them on colour maps of the function $|\Phi(\lambda,\gamma)|$, for the $\alpha_c = 1.81$, w_c = 230 nm cavity with 10-nm and 100-nm thick silver films for DBR substrate of 2 and 20 pairs of alternating TiO_2 and SiO layers (Fig. 3).

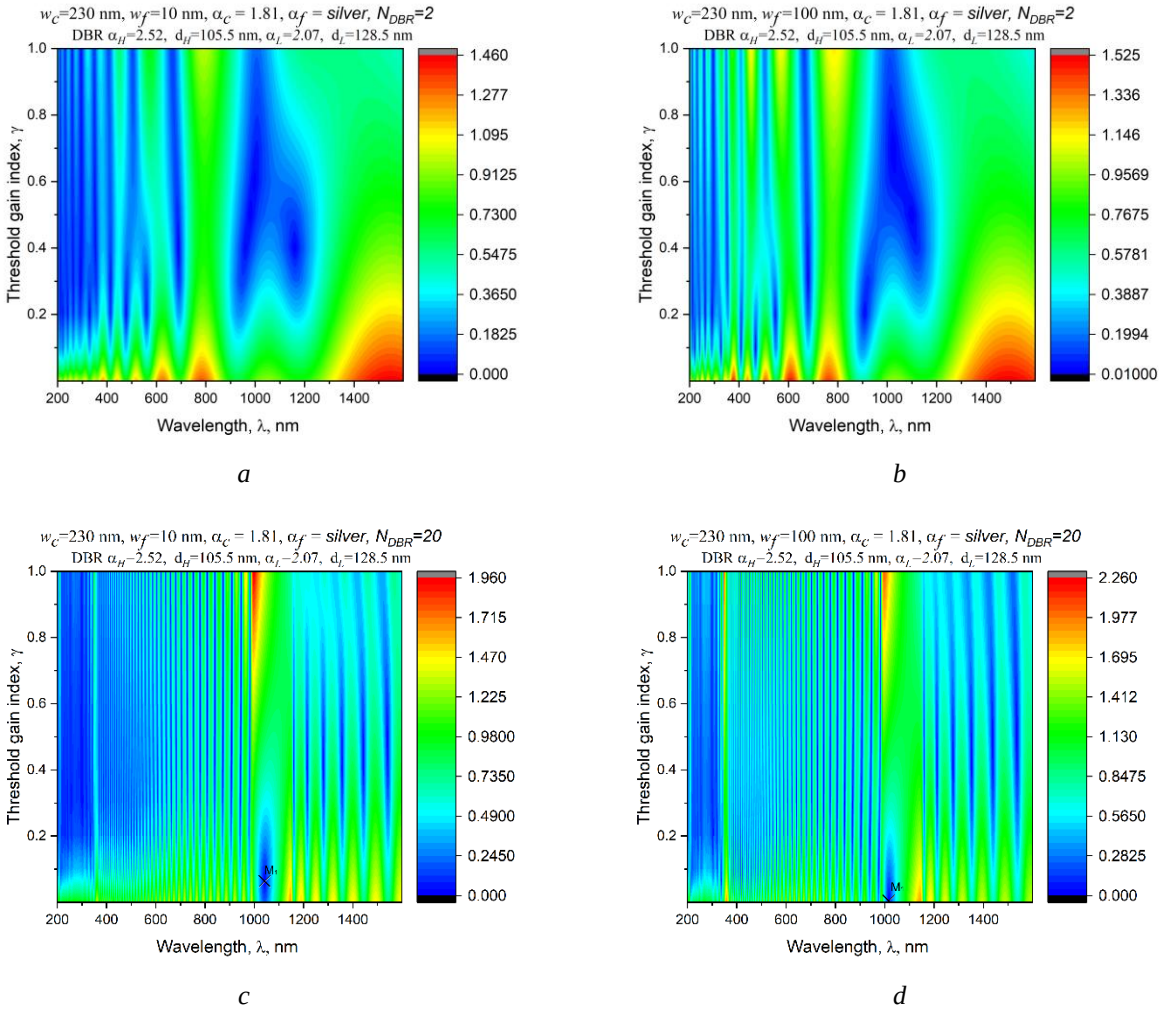


Fig. 3. Colour maps of the absolute value of (4) on the plane of threshold gain index and wavelength for the 2 (*a*), (*b*) and 20 (*c*), (*d*) pairs DBR substrate active microcavity of refractive index α_c = 1.81 with 10-nm (*a*), (*c*) and 230-nm (*b*), (*d*) bulk silver films.

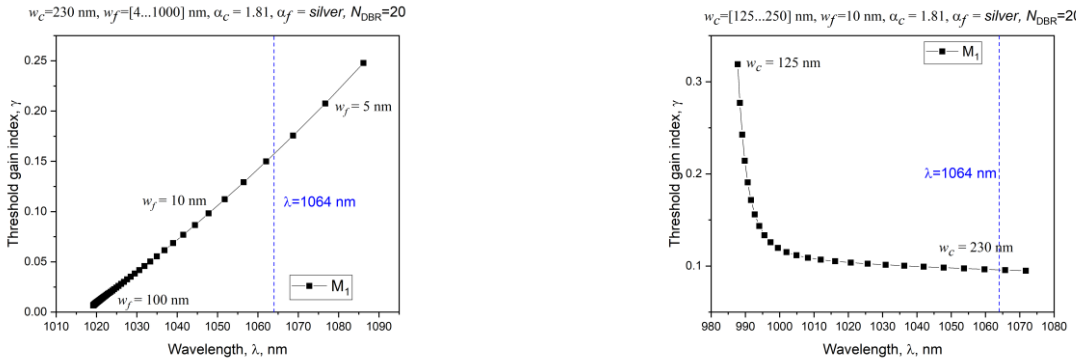
As well visible in Fig. 3, *a*, *b*, there are prominent blue spots, represent the lasing eigenvalue modes. Among them, some exhibit significantly lower threshold gain index values than the others. Here, the influence of the DBR is less pronounced; however, as the thickness of the silver metal

film increases, the threshold value of the mode decreases noticeably. This effect is well visible for M_1, marked with cross, for DBR with 20 pair of alternating layers in Fig. 3, c, *d*. In contrast to our previous model of the same cavity embedded in a silver or even gold environment in [25], the addition of large quantity of DBR layers as substrate in the current configuration shows new modes within the visible and near-infrared ranges. These modes appear in a form of many blue dips on the colour maps. Moreover, increasing the number of layer pairs, *N*, results in a greater number of emerging modes. That means they can be identified as "parasitic" modes of every DBR sub-cavity.

Fig. 3, *c* illustrates that mode M_1, with a wavelength around 1050 nm, lies within the primary band gap, and it has a lower γ, than for less pair quantity DBR configuration. This effect arises because, within its band gaps, the DBR effectively suppresses field leakage into the free halfspace. As a result, the emission threshold is reduced for cavity modes whose frequencies fall within the band gaps. In contrast, all other "parasitic" modes exhibit higher thresholds compared to the "matched" mode mentioned earlier – in the case of mode M_1, which has the lowest threshold. This behaviour can be attributed to the weak overlap between the electric fields of the DBR modes and the active region [18]. The impact of thicker silver film superstrate is similar to 2-pair DBR and shown in Fig. 2, *d*.

Further, we present the first mode M_1 trajectories for the considered microcavity with 20-pair DBR substrate and silver superstrate film under the variation of two thicknesses w_f and w_c in Fig. 4, *a*, *b*, respectively. The DBR's photonic band gap centred around 1064 nm, for example for Nd:YAG active crystal, indicated by the dashed line in the figures. As shown, careful adjustment of either the silver film (Fig. 4, *a*) or the active layer (Fig. 4, *b*) thickness allows more precise tuning of the target emission wavelength, which is of particular importance in lasers.

Note that the threshold gain index value for variation of silver film thickness linearly decreases, while, for variation of active layer thickness stops dropping dramatically from some value of w_c.



a *b*

Fig. 4. Trajectories of the modes M_1 of the active microcavity covered with silver film and lying on 20-pairs DBR under the variation of the silver film thickness w_f from 4 to 1000 nm (*a*) and active layer thickness w_c from 125 to 250 nm (*b*)

Conclusion

The 1-D mathematical model of layered microlaser threshold condition analysis has been presented. The study has examined how variations in the active layer and silver film thickness, along with the structure of the DBR, specifically the number of alternating dielectric layers, affect the emission wavelength and threshold gain index of the operating modes. For parametric analysis of the lasing eigenvalues the Lasing Eigenvalue Problem approach and Transfer Matrix Method have been used. The numerical results have demonstrated that both the emission frequency and the threshold gain index can be effectively tuned by adjusting the thickness of the silver superstrate film and the number of DBR layer pairs. Moreover, the influence of the DBR extends beyond the formation of photonic band gaps; it also leads to the emergence of numerous "parasitic" modes, which appear to originate from the DBR layers acting as unintended resonant cavities. The number of

these modes scales with the number of alternating high- and low-refractive-index layers in the DBR. Compared to the primary operating mode, these modes have exhibited considerably higher threshold gain values due to their radiation losses are suppressed by the DBR structure. Consequently, the properties of the band gap could be used to isolate the working mode by selecting the parameters of the DBR, thereby eliminating the influence of parasitic modes.

References:

1. Wang X., et al. Beam Scanning and Capture of Micro Laser Communication Terminal Based on MEMS Micromirrors // Micromachines. 2017. Vol. 14, №7. P.1917. doi: 10.3390/mi14071317

2. Feng J., et al. Random Organic Nanolaser Arrays for Cryptographic Primitives // Advanced Materials. 2019. Vol. 31, № 36. art.no. 1807880. doi: 10.1002/adma.201807880

3. Tsunekane M., Taira T. Long Time Operation of Composite Ceramic Nd:YAG/Cr:YAG Micro-chip Lasers for Ignition // Laser Ignition Conference. 2015. p. T4A.3. doi: 10.1364/lic.2015.t4a.3

4. Xu J., McCulloch D., Charlton M. D. B. Modeling full PCSELs and VCSELs using modified rigorous coupled-wave analysis // Optics Express. 2024. Vol. 32, № 13. P. 22169. doi: 10.1364/oe.522484

5. Moon S. et al. High Performance Thin Film VCSELs Integrated with a Copper Plated Heatsink // Advanced Materials Interfaces. 2023. Vol. 10, № 18. art.no. 2300191. doi: 10.1002/admi.202300191

6. Yu H., Wang L., Xu J., Chiang P. Y. A dToF Ranging Sensor with Accurate Photon Detector Measurements for LiDAR Applications // Sensors. 2023. Vol. 23, № 6. P. 3011. doi: 10.3390/s23063011

7. De Zoysa M., et al. Non-mechanical three-dimensional LiDAR system based on flash and beam-scanning dually modulated photonic crystal lasers // Optica. 2023. Vol. 10, № 2. P. 264. doi: 10.1364/optica.472327

8. Chen Z., Dong G., Barillaro G., Qiu J., Yang Z. Emerging and perspectives in microlasers based on rare-earth ions activated micro-/nanomaterials // Progress in Materials Science. 2021. Vol. 121. P. 100814. doi: 10.1016/j.pmatsci.2021.100814

9. Lin J., et al. Low-threshold whispering-gallery-mode microlasers fabricated in a Nd:glass substrate by three-dimensional femtosecond laser micromachining // Optics Letters. 2013. Vol. 38, № 9. P. 1458. doi: 10.1364/ol.38.001458

10. Lin H., et al. Diode-pumped tape casting planar waveguide YAG/Nd:YAG/YAG ceramic laser // Optics Express. 2015. Vol. 23, № 6. P. 8104. doi: 10.1364/oe.23.008104

11. Frigenti G., et al. Rare earth-doped glass whispering gallery mode micro-lasers // The European Physical Journal Plus. 2023. Vol. 138, № 8. art. no. 679. doi: 10.1140/epjp/s13360-023-04275-9

12. Li H., Wang Z., Wang L., Tan Y., Chen F. Optically pumped Milliwatt Whispering-Gallery microcavity laser // Light: Science & Applications. 2023. Vol. 12, № 1. art.no. 223. doi: 10.1038/s41377-023-01264-6

13. Tsutsumi N., Ishibashi T. Organic dye lasers with distributed Bragg reflector grating and distributed feedback resonator // Optics Express. 2009. Vol. 17, № 24. P. 21698. doi: 10.1364/oe.17.021698

14. Arshavsky-Graham S., Massad-Ivanir N., Segal E., Weiss S. Porous Silicon-Based Photonic Biosensors: Current Status and Emerging Applications // Analytical Chemistry. 2018. Vol. 91, № 1. P. 441–467. doi: 10.1021/acs.analchem.8b05028

15. Lu Q., Chen D., Wu G., Peng B., Xu J. A hybrid plasmonic microresonator with high quality factor and small mode volume // Journal of Optics. 2012. Vol. 14, № 12. P. 125503. doi: 10.1088/2040-8978/14/12/125503

16. Yevtushenko D. O., Dukhopelnykov S. V. Visible light from modulated electron beam moving between twin circular silver nanowires forming plasmonic photonic molecule // Journal of Optics. 2020. Vol. 22, № 2. P. 025002. doi: 10.1088/2040-8986/ab65d8

17. Lu Y.-J., et al. Plasmonic Nanolaser Using Epitaxially Grown Silver Film// Conference on Lasers and Electro-Optics. 2012. P. CTh5C.7. doi: 10.1364/cleo_si.2012.cth5c.7

18. Smotrova E. I., Byelobrov V. O., Benson T. M., Ctyroky J., Sauleau R., Nosich A. I. Optical Theorem Helps Understand Thresholds of Lasing in Microcavities With Active Regions // IEEE Journal of Quantum Electronics. 2011. Vol. 47, № 1. P. 20–30. doi: 10.1109/jqe.2010.2055836

19. Byelobrov V. O., Nosich A. I. Mathematical analysis of the lasing eigenvalue problem for the optical modes in a layered dielectric cavity with a quantum well and distributed Bragg reflectors // Optical and Quantum Electronics. 2007. Vol. 39, № 10–11. P. 927–937. doi: 10.1007/s11082-007-9159-4

20. Shapoval O. V., Kobayashi K., Nosich A. I. Electromagnetic Engineering of a Single-Mode Nanolaser on a Metal Plasmonic Strip Placed into a Circular Quantum Wire // IEEE Journal of Selected Topics in Quantum Electronics. 2017. Vol. 23, № 6. P. 1–9. doi: 10.1109/jstqe.2017.2718658

21. Natarov D. M., Benson T. M., Nosich A. I. Electromagnetic analysis of the lasing thresholds of hybrid plasmon modes of a silver tube nanolaser with active core and active shell // Beilstein Journal of Nanotechnology. 2019. Vol. 10. P. 294–304. doi: 10.3762/bjnano.10.28

22. Herasymova D. O., Dukhopelnykov S. V., Natarov D. M., Zinenko T. L., Lucido M., Nosich A. I. Threshold conditions for transversal modes of tunable plasmonic nanolasers shaped as single and twin graphene-covered circular quantum wires // Nanotechnology. 2022. Vol. 33, № 49. P. 495001. doi: 10.1088/1361-6528/ac8e0c

23. Kaliberda M. E., Pogarsky S. A., Kostenko O. V., Nosych O. I., Zinenko T. L. Circular quantum wire symmetrically loaded with a graphene strip as the plasmonic micro/nano laser: threshold conditions analysis // Optics Express. 2024. Vol. 32, № 7. P. 12213. doi: 10.1364/oe.514643

24. Herasymov S.S., Hnatenko O.S., et al. Threshold Conditions for 1-D Model of Laser with Partial Active Region // Journal of Nano- and Electronic Physics. 2024. Vol. 16, № 4. P. 040331-040338. doi: 10.21272/jnep.16(4).04033

25. Herasymov S.S., Hnatenko O.S. Threshold Conditions for 1-D Model of Laser Cavity Covered with Noble Metal Film // Proc. IEEE 5th KhPI Week on Advanced Technology (KhPIWeek 2024). 2024. P. 1–4.

26. Johnson, P. B., Christy, R. W. Optical constants of the noble metals // Physical Review B. 1972. Vol. 6. P. 4370–4379.

Received 05.07.2025

Information about the authors:

Oleksandr S. Hnatenko – Ph.D., Head of Department of Physical Fundamentals of Electronic Engineering, Kharkiv National University of Radio Electronics, Ukraine, email: oleksandr.hnatenko@nure.ua; ORCID: https://orcid.org/0000-0001-7722-0923

Serhii S. Herasymov – student, Department of Physical Fundamentals of Electronic Engineering, Kharkiv National University of Radio Electronics, Ukraine, email: serhii.herasymov@nure.ua; ORCID: https://orcid.org/0009-0005-1127-7544

УДК 535.4 **DOI:10.30837/rt.2025.3.222.21**

А.В. БЕЗУГЛИЙ, канд. фіз.-мат. наук

ДИФРАКЦІЯ СВІТЛА НА ОДНІЙ ТА ДВОХ НЕСКІНЧЕННО ВУЗЬКИХ ЩІЛИНАХ В ЕКРАНІ

Вступ

Перші результати експериментів з дифракції світла на одній окремій щілині та системі двох паралельних шілин, прорізаних в непрозорому екрані, датуються 1927 р. Однак і сьогодні пояснення результатів цих експериментів не задовольняють фізиків. Знов і знов з‘являються спроби пояснити їх.

Отже, метою даної роботи є теоретичний аналіз явища дифракції світла на вказаних структурах.

Дифракція фотонів на одній щілині

Хай однорідний монохроматичний потік фотонів одиничної інтенсивності (рис. 1) падає нормально з боку негативних значень осі *Y* на екран із необмеженою вздовж осі *Z* нескінченно вузькою щілиною. В цьому випадку щілину можна представити як лінійне джерело вторинних хвиль [1]. Отже, фотону, що пройшов через нескінченно вузьку щілину, можна співставити, згідно з гіпотезою де-Бройля, амплітуду ймовірності

$$\psi(\vec{k}\vec{r}) = \psi(kr) = J_0(kr)\,, \tag{1}$$

де $J_0(kr)$ – функція Бесселя нульового порядку [2]; k=2π/λ; λ – довжина хвилі де-Бройля*;* *r* – радіус-вектор, що з’єднує будь-яку точку щілини з будь-якою точкою екрана спостереження в площині XOY

Часовий множник опускаємо, розглядаємо стаціонарний процес.

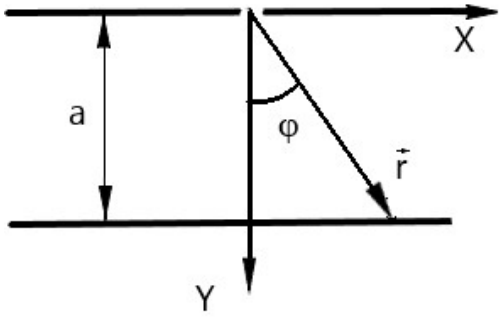


Рис. 1. Екран з однією нескінченно вузькою щілиною

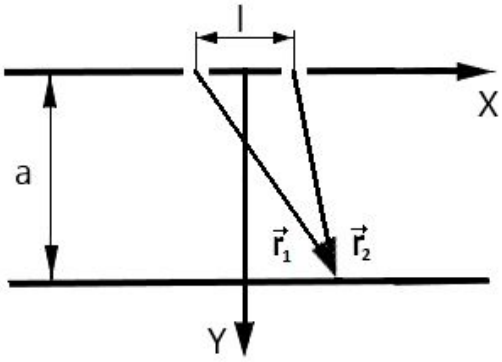


Рис. 2. Екран з двома нескінченно вузькими щілинами

Будемо вважати, що екран спостереження знаходиться на великій відстані від екрана із щілиною, тобто *kr >> 1.* Тоді, скориставшись асимптотикою функції Бесселя [2], отримаємо

$$\psi(kr) = \sqrt{\frac{2}{\pi kr}}\cos(kr - \frac{\pi}{4})\ ,. \tag{2}$$

Після нескладних тригонометричних перетворень остаточно маємо

$$\psi(kr) = \frac{1}{\sqrt{\pi kr}} \tag{3}$$

або для ймовірності

$$|\psi(kr)|^2 = \frac{1}{\pi kr} = \frac{\lambda}{2\pi^2\sqrt{x^2+a^2}}. \quad (4)$$

Таким чином, маємо неперервний розподіл імовірності потрапляння фотона в ту чи іншу точку екрана (або інтенсивності потоку фотонів). Якщо щілина розташована в точці з координатою x = 0 матимемо в точці з координатами x = 0, y = a. Якщо щілина прорізана на відстані $x = l/2$ або x = - $l/2$, відповідно максимум буде здвиинутий на таку ж саму відстань, але при цьому вирази (3), (4) практично не зміняться внаслідок того, що r >> $l/2$.

Дифракція фотонів на двох щілинах

Маємо дві нескінченно вузькі паралельні щілини, розташовані симетрично відносно початку координат на відстані l одна від одної (рис. 2). В цьому випадку для амплітуди ймовірності фотона, що пройшов через щілину будемо мати наступний вираз:

$$\psi(kr) = J_0(kr_1) + J_0(kr_2) \quad (5)$$

де – $r_1 = \sqrt{(x+\frac{l}{2})^2 + y^2}$, $r_2 = \sqrt{(x-\frac{l}{2})^2 + y^2}$.

Покладаючи $kr_1 >> 1$, $kr_2 >> 1$ та користуючись асимптотикою функції Бесселя, отримаємо

$$\psi(kr) = \sqrt{\frac{2}{\pi kr_1}}\cos(kr_1 - \frac{\pi}{4}) \quad + \sqrt{\frac{2}{\pi kr_2}}\cos(kr_2 - \frac{\pi}{4}). \quad (6)$$

Будемо вважати також, що відстань між щілинами.значно менша відстані між екранами l << a. Тоді, застосовуючи біном Ньютона та нехтуючи квадратичним членом по l, $\frac{l^2}{r^2}$, запишемо r_1, r_2 у вигляді

$$r_1 = \sqrt{1+\frac{lx}{r^2}} \quad , \quad r_2 = \sqrt{1-\frac{lx}{r^2}} \, . \quad (7)$$

Підставляючи вирази (7) в (6), після незначних тригонометричних перетворень отримаємо для $\psi-$ функції наступний вираз:

$$\psi(kr) = \frac{\sqrt{2}}{\sqrt{\pi kr}}\cos\frac{kxl}{2r} = \sqrt{\frac{2\lambda}{\pi^2\sqrt{x^2+a^2}}}\cos(\frac{\pi x}{\sqrt{x^2+a^2}}\frac{l}{\lambda}) \, . \quad (8)$$

Для $|\psi|^2$ відповідно маємо

$$|\psi|^2 = \frac{2\lambda}{\pi^2\sqrt{x^2}+a^2}\cos^2\left[\frac{\pi x}{\sqrt{x^2+a^2}}(\frac{l}{\lambda})\right]. \quad (9)$$

У випадку, коли відношення відстані між щілинами до довжини хвилі l/λ <<1, із (9) отримаємо

$$|\psi(kr)|^2 = \frac{2\lambda}{\pi^2\sqrt{x^2+a^2}} \quad . \quad (10)$$

Порівнюючи співвідношення (4) і (10), можна констатувати, що у випадку двох щілин характер дифракційної картини залишається таким же, як і для однієї щілини, що не узгоджується з відомим результатом експерименту.

Вираз (9), отриманий для $|\psi|^2$, справедливий для будь-якого співвідношення між відстанню між щілинами і довжиною хвилі. Із його аналізу видно, що при певних значеннях координати *x,* які визначаються умовою

$$\frac{\pi x}{\sqrt{x^2+a^2}}\frac{l}{\lambda}=n\pi, \tag{11}$$

в дифракційній картині будуть спостерігатися максимуми, при

$$\frac{\pi x}{\sqrt{x^2+a^2}}\frac{l}{\lambda}=(n+\frac{1}{2})\pi \tag{12}$$

– мінімуми. Із отриманих рівнянь знаходимо координати максимумів і мінімумів:

$$(x_n)_{\max}=\frac{na}{\sqrt{(\frac{l}{\lambda})^2-n^2}} \quad , \quad (x_n)_{\min[}=\frac{(n+\frac{1}{2})a}{\sqrt{(\frac{l}{\lambda})^2-(n+\frac{1}{2})^2}}\ , \tag{13}$$

де n = ±1,±2 . ±…

Із виразів (13) випливає, що кількість максимумів обмежена. Вона залежить від відношення між відстанню між щілинами l і довжиною хвилі λ. При $l/\lambda<1$ маємо монотонно спадаючу криву з максимумом при $x = 0$. При $l/\lambda > n$ будемо мати $2n+1$ максимумів в дифракційній картині. В цьому плані результати збігаються з результатами, отриманими в роботах [3, 4] при розв'язанні задач дифракції фотонів на гратці з періодом l. На нашу думку, доцільно ввести поняття характеристичного або ефективного розміру перешкоди, оскільки, як ми бачимо, для двох щілин, розташованих на відстані l, і для нескінченної гратки з періодом l при нормальному падінні характеристичний або ефективний розмір є однаковим і відношення $l/\lambda=1$ є пороговим. При $l/\lambda<1$ маємо неперервний спектр.

Якшо взяти відношення $\frac{(x_n)_{\max}}{a}$, отримаємо тангенс кута дифракції, під яким спостерігається максимум:

$$tg\varphi_n=\frac{n}{\sqrt{(\frac{l}{\lambda})^2-n^2}}, \tag{14}$$

а після незначних перетворень – добре відоме рівняння дифракційної гратки

$$l\sin\varphi_n=n\lambda\ . \tag{15}$$

З іншого боку,

$$tg\varphi_n = \frac{p_x}{p_y} = \frac{n}{\sqrt{(\frac{l}{\lambda})^2 - n^2}} = \frac{\frac{2\pi\hbar}{l}n}{\sqrt{p^2 - (\frac{2\pi\hbar}{l})^2 -}}, \quad (16)$$

де $p_x = \frac{2\pi\hbar}{l}n$ – *x*-складова імпульсу фотона, $p = \frac{2\pi\hbar}{\lambda}$ повний імпульс фотона;

$\frac{2\pi\hbar}{l} -$ – величина кванта імпульсу, що неодноразово відмічалось в роботах [3, 4].

Аналіз отриманих результатів

Аналізуючи отримані результати, можна стверджувати, що в результаті пружної взаємодії з перешкодою фотон отримує *x*-складову імпульсу. При цьому величина імпульсу фотона не змінюється, змінюється лише його напрямок. Тут слушна нагода зауважити, що термін „хвиля огинає перешкоду" (даючи визначення явищу дифракції) є некоректним. Фотон чи хвиля не огинає перешкоду вздовж якоїсь кривої, а різко змінює свій напрямок руху.

Варто також зауважити, що в явищі дифракції яскраво проявляється двоїста природа світла: з одного боку фотон – це частинка (неділима), яка згідно з представленням де-Бройля має певне значення енергії і імпульсу, з другого – це хвиля, але не фізична хвиля. Що здійснив де-Бройль, сформулювавши свою геніальну гіпотезу, яка багаторазово підтверджена практикою (знову ж таки дивно, що до цього часу вона залишається „гіпотезою")? Він об'єднав те, що, здавалося, неможливо об'єднати. До нього існували дві окремі діаметрально протилежні фізичні моделі: матеріальної точки і гармонічної хвилі. Ймовірність знайти матеріальну точку у просторі – в усіх точках – нуль, крім однієї, в якій імовірність дорівнює одиниці. Ймовірність знайти гармонічну хвилю у просторі в будь-якій точці дорівнює одиниці.

В одновимірному випадку хвильова функція ψ, що визначає стан частинки, залежить від просторової змінної *x* та часу *t*. Величина $\int |\psi|^2 dx$ дорівнює ймовірності знаходження частинки в околі точки *x* у момент часу *t* (М. Борн, 1926 р.). Якщо проінтеґрувати цю величину за всіма можливими значеннями *x*, отримаємо одиницю, тобто частинка знаходиться десь в просторі [5].

Отже, хвильові властивості фотона проявляються в тому, як визначається ймовірність його знаходження в певній точці простору. Таким чином, на нашу думку, твердження, що частинка одночасно знаходиться в усіх можливих станах і тільки при колапсі набуває певного значення, базується на тому, що, забуваючи про корпускулярні властивості, вважають її 100 %-ною (природною, фізичною) хвилею.

Висновки

Встановлено, що картина дифракції світла визначається величиною відношення характеристичного розміру перешкоди до довжини хвилі. Існує порогове значення $l/\lambda=1$, що розмежовує дві її форми: при $l/\lambda<1$ маємо монотонну криву розподілу ймовірності (інтенсивності) із максимумом на лінії симетрії системи, що відповідає неперервному спектру значень ψ-функції; при $l/\lambda>1$ – криву з чередуванням максимумів і мінімумів, що відповідає дискретному спектру значень ψ-функції.

Явище дифракції виникає в результаті пружної взаємодії фотонів із перешкодою, внаслідок якої фотони отримують складову імпульсу у напрямку, перпендикулярному до первинного. При $l/\lambda<1$ маємо неперервний спектр складової імпульсів, при $l/\lambda>1$ – дискретний.

Число максимумів в дифракційній картині залежить від відношення l/λ. Це число дорівнює $2n+1$, де $n = l/\lambda$.

Причиною виникнення відомого парадоксу при тлумаченні результатів експериментів з дифракції світла на одній та двох щілинах, на нашу думку, є те, що експеримент на одній щілині було проведено для випадку, коли довжина хвилі більша ширини щілини, а на двох щілинах – для випадку, коли довжина хвилі значно менша відстані між щілинами $l/\lambda>>1$.

Список літератури:

1. Born M.and Volf E. Prinsiples of optics. M. : Nauka, 1973. 371 p.

2. Янке Е., Эмде Ф., Леш Ф. Directmedia. 343 p.

3. Bezugly A.V., Petchenko O.M., Petchenko G.O. Photon Flux Density in the Diffraction Pattern During Scattering of H-polarized Photons by the Infinite Grating of Metallic Strips // Journal of Nano and Electronic Physics. 2021. Vol. 13, №1. 01002 (4 p).

4. Bezugly A.V. Petchenko O.M., Petchenko G.O., Dulfan H.Y. Dfraction of E-polarized Photons on Periodic Grating of Metal Strips // Journal of Nano and Electronic Physics. 2022. Vol/14, No 3. 03032 (5 p.).

5. Вакарчук І. О. Квантова механіка : підручник. 4-те вид., доп. Львів : ЛНУ ім. Івана Франка, 2012. 872 с.

Надійшла до редколегії 03.08.2025

Відомості про автора:

Безуглий Анатолій Васильович – кандидат фізико-математичних наук, доцент, Харківський національний університет міського господарства; Україна; e-mail: anatoliy.bezugliy@kname.edu.ua; ORCID: https://orcid.org/0009-0009-7546-0670

УДК 621.382.323 **DOI:10.30837/rt.2025.3.222.22**

В.М. ГРИГА, канд. техн. наук., В.М. ВІНТОНЯК, В.С. ГУЛА

МОДЕЛЮВАННЯ MOSFET-ТРАНЗИСТОРІВ З УРАХУВАННЯМ ПАРАЗИТНИХ ОПОРІВ ВИТОКУ ТА СТОКУ

Вступ

Сучасні MOSFET транзистори, особливо в умовах високих струмів або малої довжини каналу, суттєво піддаються впливу паразитних опорів витоку (R_S) та стоку (R_D). Класична модель Шічмана–Ходжеса (SPICE Level 1), що закладає основу для багатьох симуляторів (зокрема, початкових версій SPICE), не враховує прямо цих паразитних ефектів [1, 2]. Натомість реальні інтегральні MOSFET мають скінченні омічні опори в області підкладки до контакту витоку та стоку, що призводить до падіння напруги на виході та зменшення струму витоку. Для точнішого моделювання характеристик необхідно враховувати R_S і R_D навіть у спрощених моделях. Наприклад, практичне застосування SPICE Level 1 для моделювання комерційних MOSFET-масивів показує, що включення паразитних опорів дозволяє значно підвищити точність симуляцій при збереженні простоти моделі [3].

У більшості базових підходів до схемотехнічного моделювання приймається ідеалізований транзистор без паразитних опорів [4]. Це спрощує аналіз, але може давати розбіжності між розрахунковими та реальними характеристиками MOSFET. З іншого боку, сучасні компактні моделі, такі як BSIM, враховують ці ефекти численними параметрами, проте їх використання виходить за рамки локальних задач (наприклад, навчальних або первинних інженерних оцінок). Таким чином, виникає необхідність у проміжному підході: як включити вплив R_S та R_D в просту модель SPICE Level 1, зберігаючи її аналітичну прозорість та швидкодію.

У роботі запропоновано та досліджено метод модифікації моделі MOSFET рівня 1 (Shichman–Hodges) для врахування паразитних опорів витоку та стоку. Даний метод є достатньо простим та при цьому забезпечує помітно кращу відповідність реальним ВАХ транзистора у порівнянні зі стандартною моделлю без R_S і R_D.

Методика дозволяє інженерам і студентам швидко оцінювати вплив паразитних опорів без переходу на складні моделі, а також використовувати метод у навчальному процесі для демонстрації паразитних ефектів.

Еволюція моделей MOSFET і тенденції їх розвитку

Перший компактний опис статичних характеристик MOSFET запропонований Х. Шічманом і Д. Ходжесом у 1968 р. [1]. У цій моделі, відомій як модель рівня 1 SPICE або модель Шічмана–Ходжеса, передбачається довгий канал і квадратична залежність струму стоку від напруги затвор–витік (при напрузі вище порогового значення). Формули для розрахунку струму стоку I_D у моделі рівня 1 SPICE наведені в літературі [5] і широко використовуються для опису лінійного та насиченого режимів роботи транзистора. Зокрема, в лінійному режимі (транзистор відкритий, але не насичений) при $V_{DS} < V_{GS} - V_{TH}$:

$$I_D = \mu_n C_{ox} \frac{W}{L}\left[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2\right](1 + \lambda V_{DS}) \quad (1)$$

а у режимі насичення ($V_{DS} \geq V_{GS} - V_{TH}$):

$$I_D = \frac{1}{2}\mu_n C_{ox} \frac{W}{L}(V_{GS} - V_{TH})^2(1 + \lambda V_{DS}) \quad (2)$$

де μ_n – рухливість носіїв, C_{ox} – питома ємність оксиду, $\frac{W}{L}$ – відношення ширини каналу до довжини, V_{TH} – порогова напруга, а λ – коефіцієнт модуляції довжини каналу (враховує невелику залежність насиченого струму від V_{DS}).

Наведені рівняння відповідають реалізації моделі рівня 1 у SPICE і використовуються як базові у багатьох джерелах [2, 5]. Варто зазначити, що в цих формулах не враховуються паразитні опори – модель ідеалізує транзистор, припускаючи ідеальний електричний контакт до внутрішніх ділянок стоку та витоку. Подальші моделі рівня 2 і 3 доповнили базову модель напівемпіричним врахуванням короткоканальних ефектів, проте концепція зовнішніх паразитних опорів залишилася незмінною: у SPICE можна явно задавати R_S і R_D безпосередньо в картці моделі або через еквівалентну кількість квадратів дифузії (параметри RSHRSH, NRDNRD, NRSNRS) [6]. Наприклад, довідник до HSPICE зазначає, що «паразитні послідовні опори стоку та витоку можуть бути задані як R_D і R_S (Ом) або як питомий опір RSH з зазначенням кількості квадратів NRD, NRS» [6]. Таким чином, можливість моделювати вплив R_S, R_D у симуляціях існує вже тривалий час; проте аналітичний розгляд їхнього впливу в контексті найпростішої моделі MOSFET часто опускається.

У літературі можна виокремити декілька підходів до врахування паразитних опорів:

- явне додавання резисторів у схемі. Найпростіший спосіб – включити R_S та R_D як окремі елементи в схемі при моделюванні (тобто під'єднати невеликі резистори послідовно до витоку і стоку транзистора). Цей практичний прийом активно використовується схемотехніками та підтримується усіма SPICE-симуляторами. Його недолік полягає в тому, що виникає потреба в ітераційному вирішенні на рівні симулятора, а аналітичне вираження для I_D вже не може бути записане в замкненій формі;
- розширені компактні моделі. Сучасні моделі (BSIM3, BSIM4, BSIM CMG для FinFET, тощо) мають параметри для складних паразитних ефектів, включно із залежними від напруги опорами. Наприклад, у моделі BSIM4 враховано погіршення мобільності від поля, ефекти підняття опору при сильному інверсному зміщенні підкладки тощо [9]. Втім, ці моделі суттєво складніші та потребують численних параметрів, отриманих експериментально;
- аналітичні наближення**.** У низці наукових праць пропонуються спрощені аналітичні моделі, що враховують вплив паразитних опорів джерела (R_S) та стоку (R_D). Для довгоканальних MOSFET-транзисторів часто застосовується підхід ефективної напруги, згідно з яким внутрішня напруга на виводі джерела вища за зовнішню ($V_{S,int} = I_D \cdot R_S$), а на виводі стоку – нижча ($V_{D,int} = V_D - I_D \cdot R_D$). У такому випадку рівняння для внутрішнього транзистора записуються зі зміненими ефективними напругами: $V_{GS}^{eff} = V_G - V_{S,int}$ та $V_{DS}^{eff} = V_{D,int} - V_{S,int}$. Цей підхід часто згадується у літературі, однак не завжди супроводжується виведенням явної формули для струму стоку I_D, оскільки обчислення вимагає одночасного знаходження I_D, $V_{S,int}$ і $V_{D,int}$, що призводить до трансцендентного рівняння. Для розв'язання таких рівнянь пропонуються числові методи [3]. У деяких оглядових дослідженнях описано десятки методик екстракції паразитних опорів із вольтамперних характеристик та способи врахування їх у моделях MOSFET [7, 8]. Сучасні підходи передбачають використання лише одного тестового зразка для визначення асиметричних значень R_S та RD, що значно спрощує параметризацію навіть простих моделей, зокрема SPICE рівня 1. Запропоновано методи, які дозволяють визначити паразитні опори на основі DC-вимірювань у звичайному та інверсному режимах роботи, використовуючи криві $I_D(V_{GS})$ або $I_D(V_{DS})$ [9]. Ці методи вирізняються простотою реалізації та придатністю до аналітичного моделювання базових моделей, що забезпечує їхню високу практичну цінність;
- у силовій електроніці параметр $R_{DS(on)}$ (опір у відкритому стані) є одним із ключових для оцінки втрат у комутаційних режимах [4]. У сучасних дослідженнях запропоновано метод моделювання цього опору шляхом модифікації напруги затвора у поведінкових моделях з метою досягнення заданого значення $R_{DS(on)}$ [4]. Такий підхід фактично є альтернативою явному введенню паразитного опору джерела: замість того, щоб прямо враховувати R_S, модифікується ефективний поріг або керуюча напруга, що дозволяє симуляції точно відтворити виміряний опір каналу у відкритому стані. Цей метод особливо корисний для точного моделювання втрат у силових ланцюгах, проте він менш придатний для аналізу впливу окремих паразитних елементів, таких як R_S і R_D, оскільки не передбачає їх окрему інтерпретацію;

• в останні роки спостерігається зростання інтересу до використання нейронних мереж для побудови компактних моделей MOSFET. У деяких роботах пропонується гібридний підхід, при якому штучна нейронна мережа виконує обчислення поверхневого потенціалу, на основі якого далі формується аналітична модель транзистора [10]. Попри те, що цей напрям виходить за межі безпосередніх цілей даного дослідження, він ілюструє загальну тенденцію до підвищення точності відтворення паразитних ефектів. Завдяки здатності до апроксимації складних нелінійних залежностей нейронні мережі дозволяють автоматично враховувати зміну ефективного опору області стоку та витоку в різних режимах роботи. Водночас класичні аналітичні моделі зберігають актуальність завдяки своїй прозорості, фізичній інтерпретованості та простоті реалізації в стандартних інструментах моделювання.

Теоретична основа моделі з урахуванням R_S і R_D

Як зазначено вище, базові рівняння для струму стоку MOSFET без паразитних опорів описуються квадратичною залежністю між напругами і струмом за умови довгого каналу та сталої рухливості носіїв заряду [5, 11]. У контексті подальшого розширення моделі введемо позначення: параметр $K = \mu_n C_{ox} \frac{W}{L}$ питома крутизна, що є пропорційною геометричним і технологічним характеристикам транзистора). Тоді наведені вище формули можна переписати у більш компактній формі:

$$I_D = K\left[(V_{GS} - V_{TH})V_{DS} - \frac{1}{2}V_{DS}^2\right] \text{ (якщо } V_{DS} < V_{GS} - V_{TH}\text{)}, \quad (3)$$

$$I_D = \frac{1}{2}K(V_{GS} - V_{TH})^2 \text{ (якщо } V_{DS} \geq V_{GS} - V_{TH}\text{)}. \quad (4)$$

Для спрощення поки знехтуємо канал-модуляцією ($\lambda = 0$) – її можна врахувати додатковим множником $(1 + \lambda V_{DS})$ на пізнішому етапі без концептуальних змін методики.

Нехай на зовнішні виводи транзистора подані напруги V_G (затвор відносно витоку), V_D (стік відносно витоку) і V_S (витік, зазвичай заземлений або рівний нулю). У присутності паразитних опорів еквівалентна схема містить опори R_S між зовнішнім контактом витоку і внутрішнім вузлом витоку транзистора, та R_D між внутрішнім вузлом стоку і зовнішнім контактом стоку [5, 11]. Таким чином, внутрішні напруги на транзисторі відрізняються від зовнішніх. Введемо такі позначення:

- $V_{S,int}$ – напруга на внутрішньому витоці (відносно зовнішньої опорної точки, зазвичай заземлення). Через омічне падіння на R_S маємо $V_{S,int} = V_S + I_D R_S$. Якщо витік заземлено зовні ($V_S = 0$), то $V_{S,int} = I_D R_S$;

- $V_{D,int}$ – напруга на внутрішньому стоку (відносно заземлення). Вона пов'язана із зовнішньою напругою на стоку як $V_{D,int} = V_D - I_D R_D$, де падіння напруги на R_D пропорційне струму стоку I_D, який спрямований від стоку до витоку для NMOS-транзистора.

Тоді ефективні напруги зміщення транзистора у внутрішній схемі визначаються так:

$$V_{GS}^{eff} = V_G - I_D R_S, \quad V_{DS}^{eff} = V_D - I_D(R_S + R_D) \quad (5)$$

За цих умов сам транзистор (внутрішній MOSFET) можна змоделювати все тими ж рівняннями рівня 1, але вже для ефективних напруг V_{GS}^{eff} та V_{DS}^{eff}. Таким чином, отримуємо неявне рівняння для струму, оскільки $V_{S,int}$ і $V_{D,int}$ містять I_D. Підставивши визначення внутрішніх напруг, отримаємо, наприклад, для насиченого режиму ($V_{DS}^{eff} \geq V_{GS}^{eff} - V_{TH}$)

$$I_D = \frac{1}{2}K(V_G - I_D R_S - V_{TH})^2, \quad (6)$$

якщо на даному етапі знехтувати λ і вважати, що транзистор перейшов у насичений режим для даного I_D, тоді отримане рівняння є квадратичним відносно $\sqrt{I_D}$ або ж трансцендентним (через наявність I_D під квадратом і поза ним). Аналогічно можна записати рівняння і для лінійного режиму:

$$I_D = K\left[(V_G - I_D R_S - V_{TH})\left(V_D - I_D(R_S + R_D)\right) - \frac{1}{2}\left(V_D - I_D(R_S + R_D)\right)^2\right], \qquad (7)$$

де явно проявляється нелінійність відносно I_D. Аналітичного розв'язку для I_D із цих виразів у загальному випадку не існує, оскільки рівняння має п'ятий або вищий ступінь. Проте такий підхід дозволяє сформувати алгоритм для чисельного розв'язання. Замість безпосередньої спроби аналітичного розв'язання наведених формул застосуємо ітераційний метод. Для зручності спочатку перепишемо постановку задачі у компактнішій формі та введемо необхідні позначення: $f\left(V_{GS}^{eff}, V_{DS}^{eff}\right)$ – функція, що описує модельний струм транзистора рівня 1 без урахування паразитних опорів, визначається частинами залежно від режиму роботи (лінійного або насиченого). Наприклад, у насиченому режимі: $f = \frac{1}{2}K\left(V_{GS}^{eff} - V_{TH}\right)^2$, в лінійній області $f = K\left[\left(V_{GS}^{eff} - V_{TH}\right)V_{DS}^{eff} - \frac{1}{2}\left(V_{DS}^{eff}\right)^2\right]$. Тоді розширений модельний струм I_D має задовольняти: $I_D = f(V_{GS} - I_D R_S V_{DS} - I_D(R_S + R_D))$. Це основне рівняння, яке необхідно розв'язати відносно I_D.

Аналіз впливу R_S і R_D на граничні режими показує, як саме паразитні опори змінюють характеристики роботи транзистора в різних режимах. Варто відзначити якісні особливості впливу R_S і R_D на характеристики транзистора**.** У режимах малих струмів (коли $I_D \to 0$) спостерігається, що $V_{GS}^{eff} \to V_{GS}$, $V_{DS}^{eff} \to V_{DS}$, тобто вплив опорів зникає – це очікувано, оскільки паразитні опори не впливають на статичну роботу при нульовому струмі.

У режимі насичення, коли транзистор обмежує струм, введення R_S зменшує ефективну напругу затвора, отже насичений струм зменшується. Введення R_D зменшує напругу на внутрішньому стоку; це не впливає на сам насичений струм (який визначається переважно $V_{GS} - V_{TH}$), однак впливає на перехід з лінійного режиму у насичення: точка, де $V_{DS}^{eff} = V_{GS}^{eff} - V_{TH}$, буде досягнута при більшому зовнішнього значення V_{DS}. Інакше кажучи, точка насичення зміститься праворуч на графіку $I_D(V_{DS})$. Це означає, що транзистор виглядатиме більш "резистивним" на виході: досягнувши внутрішнього насичення, зовнішнє збільшення V_{DS} в основному падає на R_D, не збільшуючи суттєво I_D. Таким чином, ефект R_D проявляється як додатковий нахил ВАХ $I_D(V_{DS})$ навіть у насиченому режимі, що означає збільшення вихідної провідності.

Дослідження впливу паразитних опорів за допомогою запропонованої моделі

Для дослідження впливу R_S та R_D застосовано чисельний метод із використанням Python-бібліотеки PySpice [12], що надає інтерфейс до інструмента симуляції схем Ngspice. Обрано такий шлях з двох причин: 1) він дозволяє гнучко реалізувати ітераційний алгоритм поза самим SPICE, водночас використовуючи перевірені алгоритми для перевірки результатів; 2) легко автоматизувати серію прогонів для різних значень R_S, R_D та параметрів транзистора.

Розглянуто умовний n-канальний MOSFET з наступними параметрами моделі рівня 1 (наближені до технології 0,5 мкм):

- V_{TH} = 1.0 В (порогова напруга);
- $K = \mu_n C_{ox} W/L = 1 \times 10^{-3}\ А/В^2$. Цей параметр вибрано так, щоб для V_{GS} = 5 В насичений струм був близько кількох міліампер (для наочності графіків). Умовно це може відповідати, наприклад, $\mu_n = 400\ см^2/В \cdot с$, $C_{ox} = 7\ мФ/м^2$, $W/L \approx 100$ (що цілком реально для транзистора);
- коефіцієнт λ спочатку прийнято нульовим, щоб сфокусуватися на впливі саме R_S, R_D. Надалі в обговоренні врахуємо, як ненульова λ модифікує результати.

Діапазон симуляцій охоплює вимірювання вихідних характеристик $I_D(V_{DS})$ при різних значеннях опорів. Зокрема, фіксується напруга затвора V_{GS}, а V_{DS} змінюється від 0 до 5 В. Такий режим відповідає типовому зняттю серії вихідних вольтамперних характеристик при змінному транзистора при різних V_{GS}. Для наочності та спрощення аналізу, в роботі основна увага приділяється одній кривій при фіксованому V_{GS} (близькому до режиму сильного наси-

чення струму), а саме V_{GS} = 5 В. Це вище порогу на 4В, отже транзистор за заданих параметрів без паразитних опорів входить в насичення приблизно при $V_{DS} \approx 4$ В і досягає струму $I_D \approx 8$ мА.

Розглянуто наступні випадки: 1) ідеальний транзистор: $R_S = 0$, $R_D = 0$ (базова крива для порівняння); 2) симетричні опори: $R_S = R_D$ з декількома значеннями (наприклад, 10 Ω, 50 Ω, 100 Ω). Така симетрія спрощує аналіз, хоч у реальних процесах часто $R_S \approx R_D$. Значення в десятки Ом відповідають, для порівняння, силовому транзистору на кристалі із опором контакту та провідників; для інтегральних MOSFET на кристалі опори можуть бути меншими (≈1 Ω або менше), але у даному дослідженні взято дещо завищені значення для більш виразного ефекту на графіках; 3) несиметричний випадок: за потреби можна задати R_S і R_D різними, але в даній роботі основний акцент зроблено на симетричний випадок, де обидва присутні. Відомо, що R_S (опір витоку) впливає дещо сильніше на струм, адже зменшує ефективну напругу затвор–витік, тоді як R_D більше впливає на вихідну характеристику після насичення.

Ітераційний алгоритм полягає у тому, що для кожної комбінації заданих V_{GS}, V_{DS} і значень R_S, R_D:

1. Обчислюється початкове наближення для I_D. Найпростіший варіант – взяти $I_{D0} = f(V_{GS}, V_{DS})$, тобто поки що знехтувати паразитними опорами;

2. Підставляється I_{D0} в ефективні напруги: $V_{GS}^{\text{eff}} = V_{GS} - I_{D0}R_S$, $V_{DS}^{\text{eff}} = V_{DS} - I_{D0}(R_S + R_D)$;

3. Обчислюється нове значення струму: $I_{D1} = f\left(V_{GS}^{\text{eff}}, V_{DS}^{\text{eff}}\right)$ за модельними рівняннями рівня 1;

4. Якщо I_{D1} досить близьке до I_{D0} (за обраним критерієм збіжності, наприклад, різниця < 0,1 % або < 10^{-6} А), то алгоритм завершується, приймаючи I_{D0} як рішення. Якщо ні – підкладаємо $I_{D0} \leftarrow I_{D1}$ і повертаємось до кроку 2;

5. Як правило, за кілька ітерацій (3–5) процес сходиться, бо функція f є монотонно зростаючою по кожному аргументу і задача добре обумовлена (для фізично прийнятних значень параметрів).

Такий алгоритм реалізовано на Python. Додатково, для перевірки, побудовано SPICE-схему, де транзистор моделюється стандартною моделлю рівня 1, а R_S, R_D підключено як резистори поза транзистором. PySpice дозволив отримати результати обома способами і переконатися, це ітераційне рішення збігається з SPICE-симуляцією схемного рівня (що очікувано, але важливо з точки зору верифікації).

Код на Python використовує бібліотеки sympy (для формул) та mpmath (для чисельного рішення рівнянь), а PySpice – для формування й аналізу схем. Втім, оскільки ітераційний алгоритм виявився простим, основні результати отримано "ручним" обчисленням функції f та застосуванням функції пошуку кореня для рівняння $I_D - f(V_{GS} - I_D R_S, V_{DS} - I_D(R_S + R_D)) = 0$. Такий підхід еквівалентний описаному ітераційному процесу.

Вихідні дані для аналізу формуються після отримання збіжного рішення I_D для кожної точки: будуються криві I_D від V_{DS}. Також розраховуються таблиці зі значеннями I_D при вибраних фіксованих режимах для різних значень R_S, R_D – це потрібно для кількісного оцінювання впливу (наприклад, у відсотках). Зокрема, визначено:

- $I_D^{(0)}$ – струм без паразитних опорів (базове значення);
- $I_D^{(x)}$ – струм при деякому $R_S = R_D = x$;
- відносне відхилення $\Delta I_D(\%) = \frac{I_D^{(x)} - I_D^{(0)}}{I_D^{(0)}} \times 100\%$ (буде від'ємним, бо струм зменшується відносно ідеалу);
- зміщення точки насичення – порівнюється значення V_{DS}, при якому I_D досягає 98 % від свого насиченого значення, з і без опорів. Це дає інтуїтивну оцінку, наскільки далі по осі V_{DS} "відсувається" насичення через падіння напруги на R_D.

У табл. 1 узагальнено результати для кількох значень R_S=R_D при фіксованих напругах V_{GS} і V_{DS}. Тут вибрано дві операційні точки: точка близька до глибокого насичення

(V_{GS} = 5В, V_{DS} = 5В) та точка у лінійному режимі (V_{GS} = 4В, V_{DS} = 1,5В) у цьому випадку V_{DS} менший за $V_{GS} – V_{TH}$ = 2В, тобто транзистор точно не насичений).

Таблиця 1

Вплив $R_S = R_D$ на струм транзистора (результати моделі рівня 1; λ = 0)

$R_S = R_D$ (Ом)	I_D(V_{GS}=5В,V_{DS}=5В) (мА)	ΔI_D (%)	I_D(V_{GS}=3В,V_{DS}=1.5В) (мА)	ΔI_D (%)
0	7,999	–	1,875	–
10	7.695	–3,8	1,821	–2,9
50	6,714	–16,1	1,659	–11,5
100	5,836	–27,0	1,500	–20,0

Як показано у табл. 1, навіть відносно невеликий опір 10 Ом призводить до зниження струму приблизно на 4 % при високих напругах. При R_S+R_D=50 Ом спостерігається близько 16 % падіння струму в точці (V_{GS}, V_{DS})=(5 В, 5 В). Ці значення можна інтерпретувати як похибку, що виникає при моделюванні без урахування паразитних опорів. Наприклад, при сумарному опорі 100 Ом (витік + стік) нехтування цим впливом призводить до переоцінки струму приблизно на 27 %.

У лінійному режимі (при нижчих значеннях V_{DS}) вплив паразитних опорів також спостерігається, хоча й менш виражений. Наприклад, при 50 Ом падіння струму становить близько 11,5 %. Це закономірно, оскільки в лінійному режимі сам транзистор працює як резистор із опором у декілька кОм, і додаткові 100 Ом істотно не змінюють загальний еквівалентний опір між витоком і стоком.

У насиченому режимі внутрішній транзистор наближається до джерела струму (в ідеалі – з нескінченним вихідним опором), тому весь струм визначається балансом між внутрішньою характеристикою та зовнішніми паразитними опорами. Через це навіть десятки Ом дають помітний ефект на вихідний струм, спричиняючи суттєві похибки у моделюванні без їх урахування.

Також було оцінено зміщення точки насичення. Формально можна визначити напругу $V_{DS,\mathrm{sat}}^{\mathrm{ext}}$, при якій внутрішній транзистор досягає насичення. Це відбувається, коли $V_{DS}^{\mathrm{eff}} = V_{GS}^{\mathrm{eff}} - V_{TH}$. Підставивши $V_{DS}^{\mathrm{eff}} = V_{DS} - I_D(R_S + R_D)$ і $V_{GS}^{\mathrm{eff}} = V_{GS} - I_D R_S$, отримаємо умову

$$V_{DS} - I_D(R_S + R_D) = V_{GS} - I_D R_S - V_{TH}\,, \tag{8}$$

яку можна розв’язати щодо V_{DS}, використовуючи вже знайдений I_D в насиченому режимі. Для випадку V_{GS}=5В, R_S=R_D=50Ω знайдений $I_D \approx$ 6.714 мА. Підставляючи, одержимо $V_{DS,\mathrm{sat}}^{\mathrm{ext}} \approx 4{,}34$ В, тоді як для ідеального транзистора (I_D=8мА) ($V_{DS,sat}$=4В). Різниця ~0,34 В, що добре узгоджується з графіком рис. 1 (де червона крива виходить на плато трохи пізніше). Таким чином, паразитні опори збільшують ефективну напругу насичення. З точки зору схеми, це означає, що для досягнення заданого струму потрібна більша різниця потенціалів між стоком і витоком, щоб компенсувати падіння на резисторах. Практично це може знизити коливання напруги на виході підсилювачів на MOSFET або зменшити допустимий динамічний діапазон.

На рис. 1 показано вихідні характеристики I_D(V_{DS}), отримані для V_{GS} = 5В без паразитних опорів (жовта крива) та з доданими опорами $R_S = R_D$ = 50 Ом (червона крива).

Як випливає із результатів R_S, R_D призводить до двох помітних ефектів:

1. Максимальний (насичений) струм зменшується: на рис. 1 бачимо, що помаранчева крива (з опорами) виходить на плато приблизно на рівні 6,7 мА, тоді як жовта (ідеальна) – ~8 мА. Це приблизно –16 % відносного зменшення, що узгоджується з кількісними оцінками вище;

2. Крива з опорами переходить в насичення за більшої напруги V_{DS}. Замість різкого вигину біля $V_{DS} \approx$ 4В (жовта крива), у червоній кривій вигин згладжений і фактично насичення

досягається лише біля ~4,3 В. Після насичення, крива з R_S, R_D має помітний нахил вгору (бо частина напруги падає на R_D, трохи збільшуючи струм далі). Це еквівалентно введенню додаткового опору в канал і відповідає підвищенню вихідної провідності транзистора.

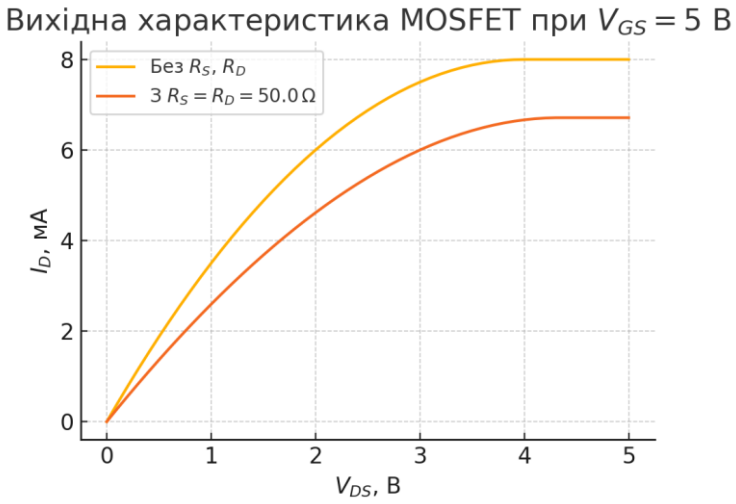


Рис. 1. Вихідні характеристики транзистора

Якщо врахувати ефект модуляції довжини каналу (λ>0), то внутрішній транзистор вже не має абсолютно плоского плато струму – навіть без R_D струм дещо ростиме із V_{DS}. Було перевірено, що при невеликому λ результати кількісно майже не змінюються, оскільки домінуючу роль у нахилі вихідної кривої все одно відіграє R_D. Наприклад, при $\lambda=0{,}02В^{-1}$ і $R_S=R_D=50$ Ω насичений струм зріс на ~3 % (з 6,71 мА до ~6,9 мА), а без опорів – з 8 мА до ~8,2 мА. Тобто відносне відхилення струму залишилося близьким (≈15 % замість 16 %). Різниця у $V_{DS,sat}$ також незначно згладилася. Отже, включення λ у базову модель просто додає паралельний вплив і не ускладнює сам метод врахування R_S, R_D.

Додатково було побудовано аналогічні сімейства характеристик для різних V_{GS} (не наведені тут задля стислості). Загальна тенденція: при меншому V_{GS} (ближчому до порогу) абсолютний вплив опорів на струм менший, але відносний вплив залишається помітним. При дуже великому V_{GS} паразитні опори ще сильніше обмежують струм, тож відсоткове зниження може зростати.

На основі аналізу можна сформулювати наступні практичні рекомендації:

- при моделюванні цифрових КМОН-схем, у яких транзистори працюють у ключовому режимі, врахування паразитних опорів R_S і R_D навіть на рівні 5–10 Ом може суттєво підвищити точність оцінки швидкодії та енергоспоживання. Такі значення дають більш реалістичні результати для імпульсних струмів, що, у свою чергу, впливають на затримки сигналів;
- у аналогових схемах (дзеркала струму, підсилювачі), де MOSFET працюють у режимі насичення, варто враховувати, що паразитні опори зменшують вихідний опір каскаду. Тому реальний коефіцієнт підсилення може бути меншим, ніж прогнозує ідеалізована модель. Запропонований метод дозволяє швидко оцінити, наскільки меншим шляхом моделювання, як зроблено вище, для типових значень R_S, R_D певного процесу;
- при екстраполяції на інші моделі: описаний підхід можна пристосувати і до моделей рівня 2 чи 3, де формули $f(V_{GS}, V_{DS})$ складніші, але загальна ідея збігається: розв'язати рівняння $I = f(V_{GS} - IR_S, V_{DS} - I(R_S + R_D))$. Звісно, можна скористатися числовим рішенням, як показано, або впровадити цю залежність прямо в SPICE через підсхему з джерелом керованого струму і зовнішніми резисторами.

Варто підкреслити, що в рамках моделі рівня 1 не було враховано:

- залежності R_S, R_D від режиму (наприклад, від напруги підкладка-витік або від того, чи перебуває транзистор у глибокому насиченні). У реальних пристроях ефективний опір витоку

може дещо зростати при певних умовах через зміни контактного опору або ефектів нерівномірності струму. Проте у перших наближеннях R_S, R_D вважаються сталими;

- температурні ефекти: з підвищенням температури рухливість падає, а опір провідників зростає, тож вплив R_S, R_D відносно може навіть зростати. Врахувати це можна, використовуючи температурні коефіцієнти в моделі (SPICE дозволяє задавати температурну залежність для резисторів). У нашому аналізі температура фіксована (25°C);
- ефект підкладки: якщо витік піднятий над підкладкою ($V_{SB,int} = V_{S,int} – V_B$), то ефективна порогова напруга збільшується на величину ефекту підкладки. У моделі рівня 1 це задається параметром γ та формулою $V_{TH}(V_{SB}) = V_{TH}(0) + \gamma\left(\sqrt{|2\phi_F + V_{SB}|} - \sqrt{|2\phi_F|}\right)$. де φ_F – потенціал Фермі. У даних розрахунках припускається, що підкладка зовнішньо з'єднана з джерелом (0 В), отже $V_{SB,ext=0}$. Однак, якщо I_DR_S набуває значних значень, внутрішня напруга $V_{S,int}$ стає додатною, а підкладка залишається на потенціалі 0 В, тому $V_{SB,int}=I_DR_S$. Це призводить до деякого збільшення внутрішньої порогової напруги V_{TH}. Описаний вище ітераційний алгоритм це не враховував (було використано фіксовану V_{TH}). Тим не менше, оцінити вплив можна: підставимо, наприклад, I_D = 6,7 мА, R_S = 50 Ω, тоді $V_{SB,int}$ = 0,335 В. Для типової $\gamma \approx 0{,}4 В^{1/2}$ і $\phi_F \approx 0{,}3 В$ отримаємо збільшення порогу на ~0,02–0,03 В. Це еквівалентно зменшенню струму ще на ~1–2 %. Тобто ефект підкладки другого порядку і може бути додатково включений в алгоритм (додаванням ще одного рівня ітерації: оновлювати оцінку V_{TH} при кожному обчисленні f). У рамках даної точності цим було знехтувано, але згадується, щоб підкреслити: для коротких каналів або високих R_S ефект підкладки від R_S може стати помітним.

Висновки

Запропоновано чисельний метод модифікації моделі MOSFET рівня 1 для врахування паразитних опорів витоку та стоку (R_S, R_D). Розв'язання нелінійного рівняння для струму виконано ітераційно поза SPICE, із використанням Python та бібліотеки PySpice – це дозволило гнучко контролювати параметри, автоматизувати обчислення. Метод базується на ітеративному розв'язанні рівняння струму з урахуванням ефективних напруг, які включають падіння на R_S, R_D. Такий підхід еквівалентний спрощеній ітерації Н'ютона–Рафсона і реалізується як поза SPICE, так і в його схемному описі.

Перевагою запропонованого підходу є можливість оцінити вплив паразитних опорів на характеристики транзистора без використання складних моделей типу BSIM. Отримані результати свідчать, що навіть невеликі значення $R_S=R_D=10\ \Omega$ знижують насичений струм на ≈4 %, а при 100 Ω похибка сягає ≈27 %. Паразитні опори також спричиняють зміщення точки насичення та зниження вихідного опору.

Метод особливо корисний:

- для аналізу цифрових КМОН-схем, де R_S, R_D впливають на імпульсні струми і затримки;
- аналогових схем (дзеркала струму, підсилювачі), де зменшується коефіцієнт підсилення;
- навчального процесу – як демонстрація впливу неідеальностей.

Метод може бути розширений для врахування температурних ефектів, ефекту підкладки та залежності R_S, R_D від режиму роботи.

Список літератури:

1. Baker R. J. CMOS: Circuit Design, Layout, and Simulation. 4th ed. Hoboken, NJ : Wiley-IEEE Press, 2019. 1280 p. ISBN 978-1-119-48151-5.

2. Sahrling M. Analog Circuit Simulators for Integrated Circuit Designers: Numerical Recipes in Python. Cham: Springer, 2021. XV. 404 p. ISBN 978-3-030-64205-1.

3. J. E. Meza et al. SPICE models for electrical simulation of commercial MOSFET arrays ALD1105/06 Aguilar /07 // ResearchGate. Jun. 2022. doi: 10.13140/RG.2.2.36661.06886.

4. Guran I.-C. A Novel ON-State Resistance Modeling Technique for MOSFET Power Switches / I.-C. Guran, A. Florescu, L. A. Perișoară // Mathematics. 2023. Vol. 11, iss. 1. P. 72. doi: https://doi.org/10.3390/math11010072.
5. Sedra A. S. Microelectronic Circuits / A. S. Sedra, K. C. Smith. 7th ed. Oxford : Oxford University Press, 2015. 1488 p.
6. Jaeger R. C. Microelectronic Circuit Design / R. C. Jaeger, T. N. Blalock. 5th ed. New York : McGraw-Hill Education, 2016. 1360 p.
7. Ortiz-Conde A. A review of DC extraction methods for MOSFET series resistance and mobility degradation model parameters / A. Ortiz-Conde, A. Sucre-González, F. Zárate-Rincón, R. Torres-Torres, R. S. Murphy-Arteaga, J. J. Liou, F. J. García-Sánchez // Microelectronics Reliability. 2017. P. 1–16. doi: https://doi.org/10.1016/j.microrel.2016.12.016.
8. A. Ortiz-Conde et al. A review of recent MOSFET source and drain resistances extraction methods using a single test device // IEEE Trans. Electron Devices. Vol. 68, no. 4. P. 1234–1240, Apr. 2021. doi: 10.1109/TED.2021.3056789.
9. R. A. Rodriguez-Davila et al. On the DC extraction of the asymmetric parasitic source and drain resistances for MOSFETs // Solid-State Electron. Vol. 170, 107837. Aug. 2020. doi: 10.1016/j.sse.2020.107837.
10. Huang S. MOSFET Physics-Based Compact Model Mass-Produced: An Artificial Neural Network Approach / S. Huang, L. Wang // Micromachines. 2023. Vol. 14, iss. 2. P. 386. doi: https://doi.org/10.3390/mi14020386.
11. Altair. HyperSpice: MOSFET Model Parameters [Електронний ресурс] / Altair. 2021. Режим доступу: https://2021.help.altair.com/2021.0.1/activate/business/en_us/block_reference_guide/_mo/_lib/HyperSpice/HTML/mos_t.html.
12. Salvaire F. PySpice: Circuit Simulation Library [Електронний ресурс] / Fabrice Salvaire. 2021. Режим доступу: https://pyspice.fabrice-salvaire.fr/releases/v1.5/.

27.07.2025

Відомості про авторів:
Грига Володимир Михайлович – канд. техн. наук, доцент, Карпатський національний технічний університет імені Василя Стефаника, доцент кафедри комп'ютерної інженерії та електронік; Україна; e-mail: volodymyr.gryga@pnu.edu.ua; ORCID: https://orcid.org/0000-0001-5458-525X
Вінтоняк Віталій Мирославович – Карпатський національний технічний університет імені Василя Стефаника, аспірант; Україна; email vitalii.vintoniak.23a@pnu.edu.ua; ORCID: https://orcid.org/0009-0002-1538-1881
Гула Вадим Сергійович – Карпатський національний технічний університет імені Василя Стефаника, аспірант, Україна; email: vadym.hula.22@pnu.edu.ua; ORCID: https://orcid.org/0009-0007-3336-8644

BIOMEDICAL ENGINEERING
БІОМЕДИЧНА ІНЖЕНЕРІЯ

УДК 638.14:614.894:535.372 **DOI:10.30837/rt.2025.3.222.23**

А.Ю. РУДЕНКО, В.А. МАРДЗЯВКО, Л.В. ВАХОНІНА, канд. фіз.-мат. наук,
М.П. КУНДЕНКО, д-р техн. наук

ПРОЄКТУВАННЯ ЗНЕЗАРАЖЕННЯ УЛЬТРАФІОЛЕТОМ З ОПТИМІЗАЦІЄЮ ДОЗУВАННЯ ОПРОМІНЕННЯ ЗАСОБАМИ ВИМІРЮВАННЯ ТА КОНТРОЛЮ ПАРАМЕТРІВ УФ-ВИПРОМІНЕННЯ

Вступ

Умови сьогодення для ведення будь-якого виробництва продукції тісно повязані з використанням хімічних або синтизовано отриманих ферментів. Даний напрям розвитку господарства хвилює не тільки споживачів але й всі цивілізовані країни. Основну увагу зосереджено на бджільництві, галузі, яка страждає найбільше від хімії, що застосовують фермерські господарства. Бджільництво також тісно пов'язане з застосуванням ферментів та препаратів для захисту та знезараження. Ефективність захисту бджіл як елемента екосистеми в майбутньому полягає в розробці та впровадженні електрофізичних методів. Зокрема, як альтернатива розглядаються методи опромінення та знезараження.

Україна традиційно входить до п'ятірки лідерів у світі за кількістю бджолосімей та виробництва меду. Але в останнє десятиліття зафіксовано масові втрати бджіл – у деяких регіонах щорічно зникає до 30–50 % бджолосімей. Цей процес має як локальні, так і глобальні причини [1, 2].

Скорочення чисельності бджолосімей в Україні є серйозною екологічною та аграрною проблемою. Цей процес має комплексний характер і зумовлений низкою взаємопов'язаних чинників.

Однією з головних причин є отруєння бджіл пестицидами, особливо неонікотиноїдами, які широко застосовуються в сільському господарстві. Часто аграрії обробляють поля інсектицидами в період цвітіння медоносних культур, не попереджаючи пасічників. Це призводить до масової загибелі бджіл після контакту з отруєним пилком або нектаром.

Не менш небезпечною є поширеність хвороб і паразитів, зокрема вароатозу, що викликається кліщем Varroa destructor. Цей паразит послаблює імунітет бджіл і сприяє розвитку вірусних, бактеріальних та грибкових інфекцій. Також загрозу становлять хвороби розплоду, такі як американський і європейський гнилець, а також грибкові ураження – аспергільоз і аскомікоз.

Ще один чинник – зміна клімату: нестабільна зима з різкими перепадами температур, ранні відлиги або затяжні дощі негативно впливають на стан бджолосімей. Такі погодні умови сприяють підвищенню вологості у вуликах, що створює сприятливі умови для розвитку грибкових інфекцій.

Зменшення природної кормової бази також значно впливає на виживання бджіл. Масове вирощування монокультур, вирубка лісосмуг, знищення дикорослих медоносів та розширення урбанізованих територій призводять до дефіциту нектару та пилку. У результаті бджолам не вистачає ресурсів для розвитку сім'ї та підготовки до зими [3].

Важливим аспектом є також людський фактор і недоліки в організації бджільництва. Відсутність належного ветеринарного контролю, брак досвіду у нових пасічників, недосконаліть законодавства у сфері захисту бджіл і недостатня комунікація між аграріями та бджолярами створюють додаткові ризики для пасік.

Таблиця 1

Статистичні дані (2020–2024)

Рік	Орієнтовна кількість бджолосімей, млн	Втрати (за оцінками пасічників)
2020	~3,5	до 25 %
2021	~3,2	до 30 %
2022	~2,9	понад 35 % (особливо на сході)
2023	~2,8	до 50 % в окремих регіонах
2024	~2,6	тенденція до скорочення зберігається

Отже, скорочення бджолосімей в Україні спричинене поєднанням хімічного, біологічного, кліматичного та соціального впливів. Для збереження популяції бджіл необхідні комплексні заходи: обмеження використання шкідливих агрохімікатів, посилення ветеринарного контролю, підтримка пасічників і розвиток екологічно сталого землеробства.

У зв'язку зі зростаючими загрозами для бджільництва особливої актуальності набуває пошук нових, екологічно безпечних підходів до профілактики та захисту бджолосімей. Одним із перспективних напрямів є використання ультрафіолетового (УФ) випромінення, яке має виражену бактерицидну дію.

Мета роботи – дослідити ефективність використання УФ-опромінення як механізму боротьби з грибковими та вірусними захворюваннями поверхні вулика.

Аналіз попередніх досліджень та літературних джерел

Пошук та альтернатива сучасним методам полягає в реалізації електротехнологій в бджільництві, які є більш ефективними для застосування. Використання УФ-опромінення є методом, який уособлює високу ефективність та в той же час простоту використання.

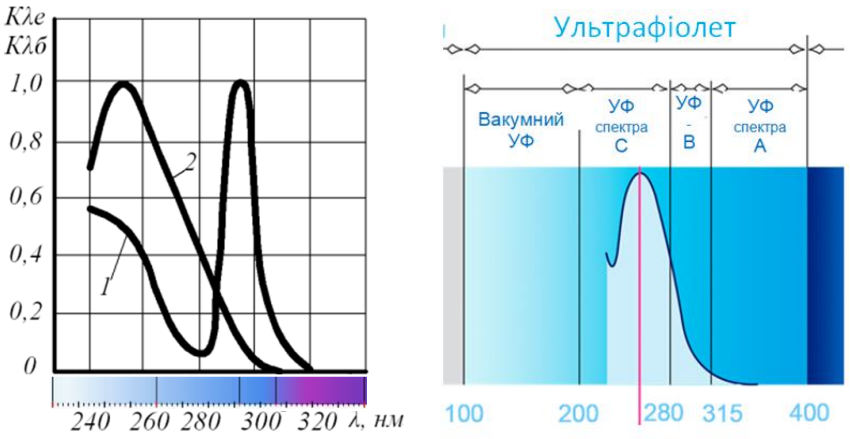


Рис. 1. Спектр дії УФ-випромінювання:1 – спектр еритемної дії УФ-випромінювання; 2 – спектр бактерицидної дії УФ-випромінювання

Застосування УФ-випромінювання у діапазоні довжин хвиль до 300 нм, відоме як UVC-діапазон, належить до так званого агресивного або бактерицидного спектру (рис. 1). Саме цей вузький спектр має виражену фотохімічну активність, яка дає змогу ефективно знищувати широкий спектр патогенних мікроорганізмів – включаючи бактерії, віруси, грибки та їхні спори. Основним механізмом дії UVC-випромінювання є руйнування нуклеїнових кислот – ДНК і РНК. УФ-промені поглинаються азотистими основами ДНК, зокрема тиміном, що призводить до утворення тимінових димерів, які перешкоджають реплікації й транскрипції. Внаслідок цього мікроорганізм втрачає здатність до розмноження і гине [4].

Оскільки UVC-випромінювання не залишає хімічних залишків і не утворює побічних токсичних продуктів, воно набуває дедалі більшого поширення як альтернатива традиційним дезінфекційним засобам. Застосування УФ-опромінення є особливо актуальним у сферах, де

критично важливо забезпечити високий ступінь стерильності при мінімальному втручанні у фізико-хімічні властивості матеріалів чи біооб'єктів.

Зокрема, у ветеринарії та бджільництві ультрафіолет використовується для знезараження вуликів, рамок, інвентарю та поверхонь, заражених патогенними мікроорганізмами, включаючи спори грибів (Ascosphaera apis, Aspergillus spp., Candida spp.), що є збудниками небезпечних захворювань розплоду. УФ-опромінення дозволяє ефективно пригнічувати розвиток хвороб без застосування агресивних дезінфектантів, які можуть вплинути на якість меду чи здоров'я бджолосімей.

В аграрній техніці УФ використовується для стерилізації насіння, обробки тари, пакувальних матеріалів, систем зберігання продуктів. У медицині UVC-промені застосовуються в знезаражувальних камерах, операційних, лабораторіях, вентиляційних системах та поверхневій обробці інструментів і приміщень.

Універсальність UVC-опромінення, його миттєва дія, простота застосування та відсутність хімічних реагентів дозволяють вважати його одним із найперспективніших засобів безконтактного знезараження, особливо в умовах підвищеної біологічної небезпеки або стійкості патогенів до традиційних антисептиків [4].

Для кількісного оцінювання ефективності УФ-обробки використовується поняття дози, потужності опромінення, щільність опромінення, порогова доза, яка розраховується за формулами:

$$D = E \times t, \tag{1}$$

D – доза опромінення (Дж/см²); E – потужність випромінювання (інтенсивність), Вт/см²; t – час опромінення, с.

Потужність випромінення УФ-С:

$$P_{el} \approx \frac{P_A}{\eta_{UVC} k_{meop}}; \tag{2}$$

В дослідженнях національного університету біоресурсів і природокористування (2022) підтверджено, що застосування УФ-С ламп для періодичного знезараження вуликів призводить до зниження рецидиву грибкових інфекцій на 70 % у порівнянні з традиційними методами обробки (пар, оцтова кислота). Метод виявився ефективним у міжсезонний період.

Реалізація застосування УФ-опромінювача в пасічному господарстві та практична реалізація

Для аналізу та переконання в практичній ефективності дії опромінюючої установки було виготовлено макет дослідного зразка. Установка має енергоефективну конструкцію, яка складається з УФ-опромінювача бактерицидного діапазону, гілки живлення у вигляді компактної автономної сонячної установки, а також однофазного реле часу, що забезпечує контроль тривалості опромінення (рис. 2).

Така конфігурація дозволяє здійснювати санітарну обробку вуликів або елементів пасічного обладнання без прив'язки до стаціонарної електромережі, що є критично важливим для використання в польових умовах [5, 6].

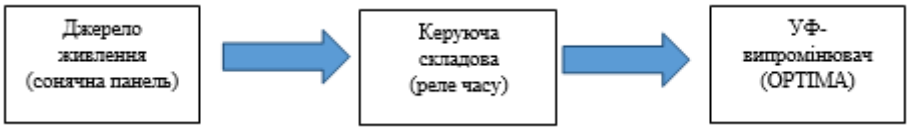


Рис. 2. Схема установки знезараження на базі УФ-опромінювача

Дана установка базується на використанні агресивних хвиль. Порогова біодоза для бджоли становить 25 МВт·с·см² (250 Дж/м²), а при повторному опроміненні – 45 МВт·с·см². На практиці зазвичай застосовують 0,125–0,667 біодози, при перевищенні якої бджола, кліщ або інший біоорганізм не виживають. Волосковий покрив бджоли зменшує інтенсивність

УФ-променів на 12,5–15 %. Дозування ускладнюється індивідуальними особливостями бджіл і кліщів: породою, віком, фізичним станом, кольором та хітиновим покриттям.

Діапазон застосування до 300 нм. Даний світильник забезпечує якісне виконання функції знезараження в бактерицидному спектрі. Основою теорії дослідження слугують дослідження (наприклад, Інституту ветеринарної медицини НААН), опромінення ультрафіолетом з довжиною хвилі 254 нм. Дослідження обрання правильної довжини хвилі – запорука якісної обробки пасічного інвентарю та елементів вулика.

Рис. 3. Світильник з лампою УФ

Відповідно до мети роботи проведено експериментальні дослідження для оцінки ефективності УФ-С випромінення (254 нм) у знезараженні дерев'яних елементів вулика, інфікованих грибковими та вірусними патогенами, шляхом порівняння ступеня зниження інфекційного навантаження при різній тривалості опромінення [6, 7].

Для дослідження ефективності знезараження використовували дерев’яні вулики 0,6 х 0,5 м, інокульовані спорами. Як джерело опромінення застосовували УФ-випромінювач (рис. 3) з довжинами хвиль 210, 222, 254, 265 та 280 нм. Тривалість опромінення становила 10 хв при фіксованій інтенсивності випромінювання $E = 0{,}4 Вт/м^2$. Відстань від джерела випромінювання до поверхні зразків – 50 см. Ефективність оцінювалась шляхом підрахунку колонієутворюючих одиниць (CFU) та визначення відсотка їх зниження порівняно з контрольним зразком, що не піддавався опроміненню.

Для забезпечення точності та повторюваності експерименту використано контрольно-вимірювальні прилади, що дозволяють фіксувати основні параметри УФ-опромінення та контролювати умови знезараження [8]:

- люксметр/радіометр УФ-діапазону (UVC): застосовувався для визначення інтенсивності випромінювання у заданій точці експозиції. Прилад був відкалібрований на довжину хвилі 254 нм і забезпечував вимірювання у межах 0,1–10 мВт/см² з точністю ±5 %;
- таймер-реле: використовувався для точного дозування часу опромінення. Пристрій забезпечував стабільну тривалість експозиції з похибкою не більше ±1 с на кожні 10 хв роботи. Це дозволяло виключити людський фактор при регулюванні тривалості впливу;
- термометр цифровий: контролював температуру в зоні опромінення для запобігання впливу теплового чинника на результати експерименту. Температурні коливання не перевищували ±1 °C;
- лабораторна установка з опромінювачем містила змінні джерела УФ-випромінення з можливістю швидкої заміни ламп різної довжини хвилі. Конструкція забезпечувала сталу відстань 50 см між джерелом випромінення і зразком.

Застосування наведених приладів забезпечило надійний контроль за основними параметрами опромінення та точну інтерпретацію експериментальних результатів. Контрольні вимірювання проводилися перед кожною серією опромінення для верифікації стабільності випромінювача та відтворюваності умов досліду [9].

Результати експерименту зведено у табл. 2.

Таблиця 2

Результати експерименту

Довжина хвилі (нм)	Кількість колоній CFU після опромінення	Ефективність знезараження, %
210	80	68
222	50	80
254	8	96,8
265	12	94,4
280	45	82
Контроль (0 хв)	250	0

Для якісного аналізу виконання знезараження запропоновано проведення знезараження пасічного інвентарю за різними діапазоном випромінення та часом обробки у відповідності до формули розрахунку потужності випромінення (1), рис. 4.

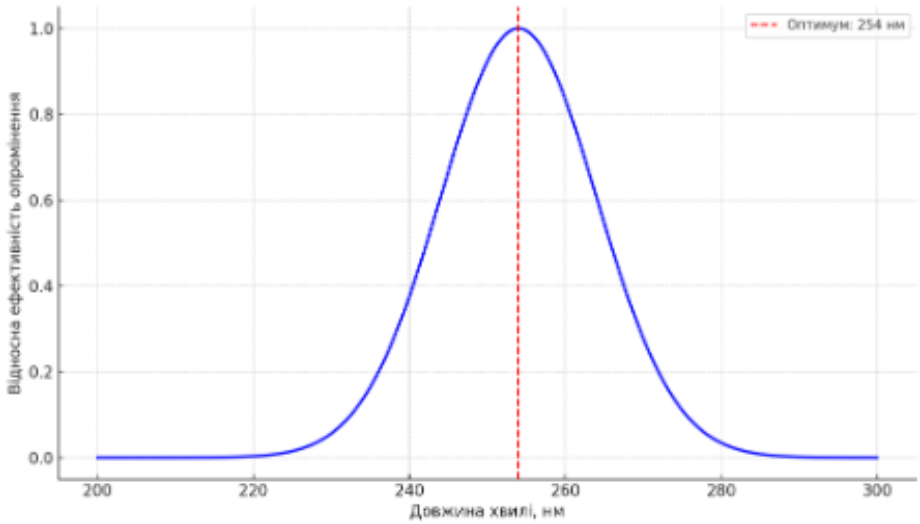


Рис. 4. Графік експериментальних данних за ефективністю знезараження від довжини хвилі УФ-випромінювання

З рис. 4 видно, що оптимальна довжина хвилі становить 254 нм, при якій спостерігається максимальна ефективність – саме в цьому діапазоні УФ-С випромінення найкраще деструкує ДНК мікроорганізмів (спори, грибки, бактерії), але визначено, що діапазон за часом для ефективної обробки повинен сягати 12,5 хв [8, 9]. Хоча дана система показала себе якісно як метод для заміни стандартних методів обробки інвентарю в пасічному господарстві, але дана система знезараження якісно бориться з найпростішими, при цьому використання унеможливлюється там, де є бджоли. Агресивне випромінення може руйнувати зорову активність комах, що деструктує їх орієнтацію в просторі з можливістю повернення до норми або цілковитою втратою та може знищити колонію або особину повністю. Недоліком є потреба в переселенні колонії в новий вулик при потребі обробки.

Для розуміння правильності застосування методу проведено комплексний економічний аналіз методів в табл. 3, що дає змогу побачити правильність обрання методу з економічної точки зору.

Таблиця 3

Затрати за методами знезараження

Показник	УФ-знезараження	Парова обробка	Оцтова кислота	Термообробка (газова/вогнева)
Первинна вартість обладнання	1500–3000 грн (УФ-камера)	1000–2000 грн (парогенератор)	200–400 грн (ємності + кислота)	800–1200 грн (газова горілка)
Експлуатаційні витрати (на рік)	200–300 грн (електрика + лампа)	600–900 грн (газ/вода)	400–600 грн (кислота + засоби)	700–1000 грн (газ)

Тривалість процедури (1 вулик)	10–15 хв	30–45 хв	24 год (випаровування)	10–15 хв
Ризик пошкодження дерева	Мінімальний	Помірний	Високий (волога + кислота)	Високий (перегрів)
Ресурс/термін служби	2000 год (лампа)	1–2 роки	Кожен раз нове	Витратний газ
Захист працівника	Захисні окуляри/одяг	Перчатки, вентиляція	Респіратор, витяжка	Вогнестійкий одяг

Таблиця 4

Ефективність протигрибкових спор (Ascosphaera, Aspergillus)

Метод	Ефективність, %	Додаткові переваги
УФ (254 нм)	до 99,5	Без хімії, швидка обробка
Парова обробка	80–85	Глибоке проникнення, доступність
Оцтова кислота	до 70	Дешева, але довга дія, шкідлива пара
Вогнева/термообробка	до 90	Швидко, але небезпечно і пошкоджує воск

Таблиця 5

Умовний розрахунок на 100 вуликів / рік

Метод	Річна вартість (грн)	Примітки
УФ	~300–500	За умови ресурсу лампи 2000 годин
Парова обробка	~800–1200	Витрати палива + вода
Оцтова кислота	~900–1300	Регулярні покупки кислот
Вогнева обробка	~1000–1500	Газ + висока витрата часу

Отже, бачимо, що система УФ-знезараження економічно вигідна вже після першого сезону, при заживленні її від сонячного джерела вона стає ще вигіднішою. Метод є безпечним для деревини, не залишає залишків, не потребує витратних матеріалів після встановлення.

Висновки

Основні причини скорочення бджолосімей полягають у поєднанні біотичних, техногенних та організаційних чинників. Серед них найвагомішими є отруєння бджіл пестицидами, неправильне використання хімічних препаратів у сільському господарстві, поширення хвороб, зокрема вароозу та нозематозу, а також грибкових інфекцій, до яких бджоли вкрай вразливі в умовах послабленого імунітету. Важливу роль відіграє і кліматичний чинник – різкі зміни погоди, підвищення середньорічних температур, зниження вологості та тривалі періоди спеки. До цього додається зменшення кормової бази для бджіл внаслідок скорочення площ медоносної рослинності та вирубування лісосмуг. Окрему загрозу становлять віруси, які передаються через паразитів і слабкі сім'ї, а також недотримання ветеринарних вимог щодо профілактики та карантину. Організаційні проблеми – відсутність належного контролю, несвоєчасне проведення дезінфекції, а також нехтування сучасними знезаражувальними методами – лише посилюють кризу. З огляду на це, постає необхідність у пошуку ефективних, екологічно безпечних та економічно вигідних рішень, серед яких важливе місце займає використання ультрафіолетового випромінювання як методу знезараження вуликів та пасічного інвентарю. Такий підхід дозволяє мінімізувати ризик інфекційного ураження бджіл без застосування хімічних речовин, що робить його перспективним засобом профілактики та збереження бджолиних родин.

Список літератури:

1. Агроновини. Пчелосемей в Украине стало меньше почти на 6 % / AgroPortal.ua. AgroPortal.ua. URL: https://agroportal.ua/ru/news/zhivotnovodstvo/bdzholosimey-v-ukrajini-pomenshalo-mayzhe-na-6.
2. Interfax-Ukraine. Смертність бджіл в Україні у 2024 р. становила 20–25 %, а втрати галузі через війну сягають 30 % / Інститут бджільництва. Інтерфакс-Україна.
URL: https://interfax.com.ua/news/general/1038912.html.
3. Fedoriak M. RESULTS OF MONITORING OF HONEY BEE COLONY LOSSES IN UKRAINE AFTER THE WINTER OF THE FIRST YEAR OF THE WAR (2021–2022) // Biolohichni systemy. 2024. Vol. 16, no. 3. URL: https://doi.org/10.31861/biosystems2024.03.300
4. Санін Ю. К. Методи та засоби впливу УФ випромінюванням для знезараження варроатозу бджіл // Thesis. 2020. URL: http://repository.kpi.kharkov.ua/handle/KhPI-Press/48427.
5. Центр громадського здоров'я України | МОЗ.
URL: https://phc.org.ua/sites/default/files/uploads/documents/files/32354b7fa7425199a4b4c67ce5a89f53.pdf
6. Semenov A., Semenova K. Ultraviolet disinfection of water in recirculating aquaculture system: a case study at sturgeon caviar fish farm // Acta agriculturae Slovenica. 2022. Vol. 118, no. 3. P. 1. URL: https://doi.org/10.14720/aas.2022.118.3.2488.
7. Blau K., Gallert C. Efficacy of UV-C 254 nm Light and a Sporicidal Surface Disinfectant in Inactivating Spores from Clostridioides difficile Ribotypes In Vitro Pathogens. 2024. Vol. 13, no. 11. P. 965. URL: https://doi.org/10.3390/pathogens13110965.
8. Chen H., Moraru C. I., Protasenko V. V. Maximizing the disinfection effectiveness of 254 nm UV-C light with a special design unit: simulation and experimental approaches // Frontiers in Food Science and Technology. 2023. Vol. 3. URL: https://doi.org/10.3389/frfst.2023.1223829 .
9. Ultraviolet germicidal irradiation disinfection of Stachybotrys chartarum / C. F. Green et al // Canadian Journal of Microbiology. 2005. Vol. 51, no. 9. P. 801–804. URL: https://doi.org/10.1139/w05-061.
10. Glover K. K., Nunayon S., Zhong L. Ultraviolet germicidal irradiation: Advances in viral inactivation and vaccine development // Indoor Environments. 2025. P. 100099. URL: https://doi.org/10.1016/j.indenv.2025.100099.

Надійшла до редколегії 10.07.2025

Відомості про авторів:

Руденко Андрій Юрійович – асистент кафедри електроенергетики, електротехніки та електромеханіки, Миколаївський національний аграрний університет, Україна; e-mail: rudenkoau@mnau.edu.ua; ORCID: https://orcid.org/0000-0002-5103-6412

Мардзявко Віталій Анатолійович – асистент кафедри електроенергетики, електротехніки та електромеханіки, Миколаївський національний аграрний університет, Україна; e-mail: mardzyavko@mnau.edu.ua; ORCID: https://orcid.org/0000-0001-7327-9215

Вахоніна Лариса Володимирівна – канд. фіз.-мат. наук, доцент, Миколаївський національний аграрний університет, доцент кафедри електроенергетики, електротехніки та електромеханіки, Україна; e-mail: vakhoninalv@mnau.edu.ua; ORCID: https://orcid.org/0000-0002-1668-2275

Кунденко Микола Петрович – д-р техн. наук, професор, Національний технічний університет «Харківський політехнічний інститут», завідувач кафедри теплотехніки та енергоефективних технологій, Україна; e-mail: mykola.kundenko@khpi.edu; ORCID: https://orcid.org/0000-0002-5841-4367

УДК 577.3:621.3.049:004.942 **DOI:10.30837/rt.2025.3.222.24**

А.Ю. РУДЕНКО, В.А. МАРДЗЯВКО, В.О. МАРТИНЕНКО, канд. техн. наук,
М.П. КУНДЕНКО, д-р техн. наук

ДОСЛІДЖЕННЯ ВПЛИВУ ЕЛЕКТРОМАГНІТНОГО ПОЛЯ НА ІОННІ КАНАЛИ КЛІТИНИ З ВИКОРИСТАННЯМ СИСТЕМ МОДЕЛЮВАННЯ ТА ВИМІРЮВАННЯ

Вступ

Актуальність дослідження зумовлена сучасними тенденціями розвитку біомедичних технологій, де ключову роль відіграють механізми міжклітинної взаємодії та регуляція іонного транспорту. Іонні канали, зокрема для іонів Na^+, Ca^{2+} і K^+, є основними провідниками імпульсних сигналів, що лежать в основі багатьох життєво важливих процесів, таких як активація ферментів, скорочення м'язової тканини та апоптоз. Дослідження пульсацій клітини, які поєднують електричні, механічні й метаболічні параметри, дозволяє глибше зрозуміти механізми регулювання клітинної активності та обміну інформацією між клітинами.

Особливий інтерес становить вплив зовнішніх електромагнітних полів на електрофізіологічні властивості мембрани та іонні канали. Моделювання процесів проходження іонів Na^+ через мембрану під дією електромагнітного поля сприяє розкриттю механізмів зміни мембранного потенціалу, визначенню умов деполяризації або гіперполяризації, а також оцінці можливості цілеспрямованої стимуляції чи пригнічення клітинної активності [1, 2].

Метою даного дослідження є розробка та аналіз математичних моделей, що описують динаміку іонного транспорту, пульсацій клітини та зміну мембранного потенціалу під дією зовнішніх електромагнітних впливів. Отримані результати спрямовані на поглиблення розуміння фундаментальних механізмів електрофізіологічної регуляції та можуть стати основою для розробки нових методів контролю клітинних функцій у біомедичній практиці.

Аналіз попередніх досліджень біологічних аспектів міжклітинні зв'язки

Сучасні дослідження продовжують вивчати біологічні аспекти міжклітинні зв'язки для подальшого розвитку сфери управління та передачі іонів транспорту [1]. Виходячи з теорії обміну інформації, розуміємо, що передача відбувається за рахунок імпульсів через іонні зв'язки транспорту (Na++; Ca++; K+). Насамперед, осциляційні іони (Na++; Ca++; K+) відповідають за передачу імпульсних сигналів, чим керують основними функціями, а саме активація ферментів, скорочення м'язової тканини, апоптоз.

Можна звернути увагу на цитоскелетні пульсації клітини, що забезпечуються рухом, зміну форми та мітоз [2].

Пульсації клітини є фундаментальними явищами, що є сукупністю електричних коливань, соматичних, механічних та металогічних переметрів. Одним з елементів транспорту є Na+. Натрій бере участь в формуванні мембранного потенціалу клітини, електричних імпульсах, транспорті речовин. Пульсації клітин можуть проявлятися у вигляді змін електрофізичного складу мембрани клітини в момент часу осциляцій потенціалу клітини .

Вираження пульсацій клітини можливо через коливання всього об'єму клітини в життєвому моменті зміни внутрішнього іонного складу і безпосередньо в механічних коливаннях. Дані активності відповідають за регулювання іонних каналів та транспорті іонів зокрема Na+. В табл. 1 висвітлено основні транспортні механізми активної та пасивної фази [3].

Таблиця 1

Основні транспортні механізми

Активні	Пасивні
Na+/ K + – АТФаза основна, що підтримує низький рівень Na+	Напругозалежність Na+каналу
Na+/ Ca ++ – бере участь в кальцієвому гомеостазі	Ліганд-регульовані канали (відповідь на нейромедіатори)
Na+/ H+ важливий для регуляції pH	

Пояснення процесу обміну іонами в міжклітинному середовищі через математичну модель

Для дослідження обрано процеси клітини за залученням іону Na+. Для аналітики використовують: Patch-clmp – визначення електричної активності іонних каналів; іонну мікроскопію – візуальне дослідження іонного складу клітин; SBFI – динамічне визначення вмісту Na+.

Побудова Matlab та математичних моделей. Зазвичай в процесі моделювання пульсацій та механічних скорочень побудови математичних моделей використовують рівняння Навє– Стокса для рідини в середині клітини [4]:

$$R_{(t)} = R_0 + A\sin(\omega t) \tag{1}$$

де $R_{(t)}$ – радіус клітини в момент часу *t*; R_0 – середній радіус клітини; A – амплітуда пульсації клітини; ω – частота пульсації клітини.

Можливе використання також і математичних моделей для визначення пульсацій клітини, зокрема модель

$$\frac{dV}{dt} = k\left(P_{in} - P_{out}\right). \tag{2}$$

Дана модель поєднує прості параметричні складові, а саме V – об'єм та P – тиск, що дає деяку картину скорочень клітини, але має деякий умовний коефіцієнт *k* – коефіцієнт проникності або коефіцієнту зміни об'єму клітини. Для аналізу міжклітинних коливань поєднують взаємодії клітини в певній популяції, яка до того ж буде враховувати зовнішні фактори збурення.

Для аналізу виконано моделювання клітинної системи на основі електрофізичної моделі.

Моделювання проводиться за допомогою використання моделі клітинної взаємодії між собою. Така модель актуальна для дослідження та моделювання поширення сигналу в тканинах або колоніях одноклітинних, за умови, що кожна клітина має власну фазу осциляції $\theta_i(t)$ та взамодією з сусідніми клітинами. Застосуємо модель Курамато [5]:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{i=1}^{n}\sin(\theta_j - \theta_i) \tag{3}$$

Дана модель показує що є певний синхронізм клітини при певних значеннях параметру K, що дає можливість клітинам осцилювати синхронно. Можна спостерігати картину, яка не є повною, слід враховувати певний фактор зовнішнього збурення зовнішнього впливу. Це дає можливість модифікувати модель Курамати з A-амплітудою та f-частотою зовнішнього синхронізму:

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{N}\sum_{i=1}^{n}\sin(\theta_j - \theta_i) + A\sin(f(t) - \theta_i), \tag{4}$$

отже введення додаткового члена $A\sin(f(t) - \theta_i)$ до функції моделі Курамати дає можливість розуміти, що явище синхронізму можливе й при зовнішніх впливах на систему [6]. Дане явище називають ентрайнмент та представлено на рис. 1.

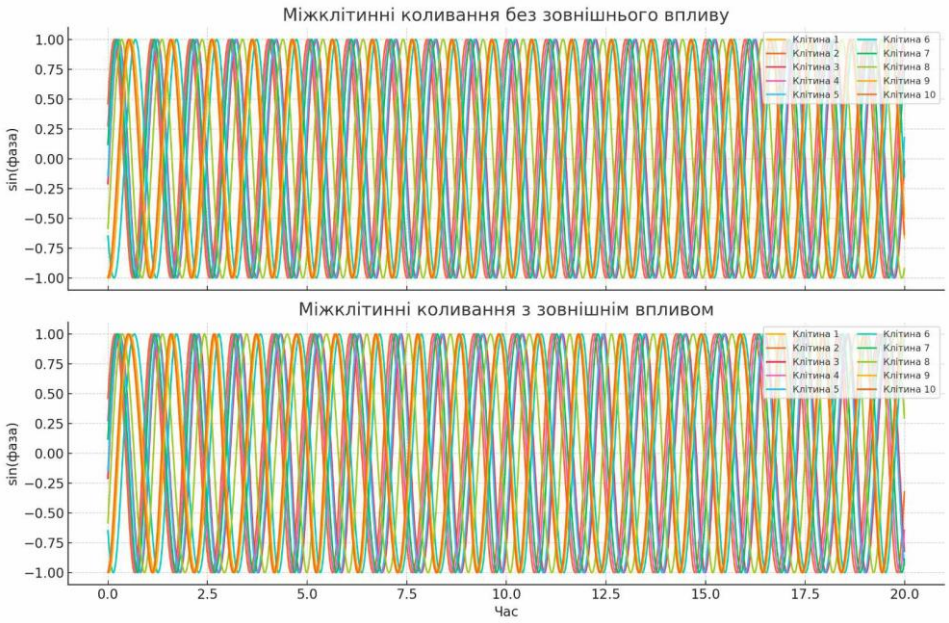


Рис. 1. Порівняння міжклітинних коливань без та під дією зовнішнього електромагнітного впливу

Таблиця 2

Основні обмеження та припущення моделі

№	Обмеження/припущення	Опис
1	Ізольована клітина	Розглядається як окрема система без врахування впливу сусідніх клітин чи тканинного середовища.
2	Гомогенність середовища	Внутрішньоклітинне середовище передбачається однорідним, без локальних градієнтів концентрацій іонів.
3	Стаціонарність параметрів	Параметри проникності мембрани, властивості каналів та інші фізіологічні характеристики вважаються сталими протягом моделювання.
4	Спрощена модель каналів	Іонні канали описуються як ідеалізовані провідникові елементи без детального урахування складної кінетики та модулюючих білків.
5	Спрощений вплив електромагнітного поля	Зовнішній вплив моделюється як гармонічне поле постійної амплітуди та частоти, без просторових варіацій.
6	Обмежений часовий інтервал	Аналізуються процеси протягом короткого проміжку часу, не враховуються довготривала адаптація чи зміни метаболічних параметрів.
7	Відсутність метаболічних ефектів	Не враховано вплив метаболічних процесів на активність каналів і динаміку іонного складу.

Перенесення моделювання клітинного процесу на електричну модель дослідження

Для подальшого моделювання та аналізу використаємо електричну модель, що є базисною (рис. 2).

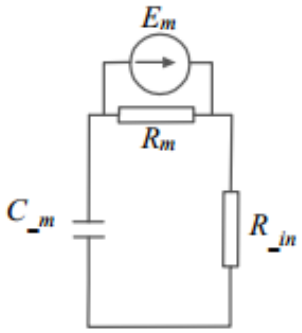


Рис. 2. Класична модель Герца

Для точності розширення опису електричних властивостей мембрани враховують різні іонні канали, які являють собою окремі кола з резисторами та джерелом ЕРС [7]. Для цього використовують модель Голдмана–Ходжикіна–Каца (рис. 3).

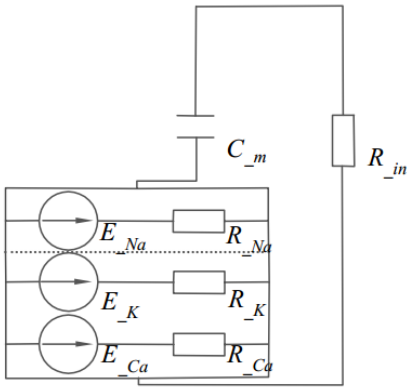


Рис. 3. Модель Голдмана–Ходжикіна–Каца

Уточнення та конкретизація параметрів клітинної системи необхідні для уточнення параметрів рішення. Побудуємо рівняння балансу струмів для розуміння перехідного процесу в клітині [8, 9]. Основою рівняння балансу є параметр сумарної мембранної провідності та канальної провідності.

Рівняння сумарного струму, що проходить через мембрану:

$$I_{\Sigma} = C_m \frac{dV_m}{dt} + \sum I_{ion}, \tag{5}$$

$$I_{ion} = \frac{V_m - E_{ion}}{R_{ion}}.$$

Загальне рівняння має вигляд

$$I_{\Sigma} = C_m \frac{dV_m}{dt} + \sum \frac{V_m - E_{ion}}{R_{ion}} - I_{evt} \tag{6}$$

Наступним кроком буде врахування впливу навколишнього електромагнітного поля на клітину (6):

$$I_{evt} = I_{\Sigma} + I_{EM(t)}$$

Отже, рівняння для чисельного моделювання буде мати вигляд

$$C_m \frac{dV_m}{dt} + \sum \frac{V_m - E_{ion}}{R_{ion}} - I_{\Sigma} + I_{EM(t)}, \tag{7}$$

диференціальне рівняння з часозалежними коефіцієнтами, яке можна розв’язати чисельно. На основі рівняння побудовано модель в MATLAB/Simulink графік представлено на рис. 4.

Рис. 4. Статичний графік, який показує, як змінюється мембранний потенціал клітини під впливом електромагнітного поля

На графіку показано зміну мембранного потенціалу клітини під впливом електромагнітного поля з частотою 50 Гц та амплітудою змінного струму 100 нА. Початкове значення потенціалу становить −70 мВ, що відповідає стану спокою клітини. Під дією поля мембранний потенціал поступово підвищується, досягаючи значення близько −56 мВ, що свідчить про деполяризацію мембрани. Така зміна супроводжується хвилеподібними коливаннями, період яких відповідає частоті зовнішнього впливу. Замість пригнічення (гіперполяризації) спостерігається посилення збудливості клітини, що вказує на стимулюючий ефект електромагнітного поля [9, 10], а, отже, пригнічення проходження іонів Na^+ через клітинну мембрану.

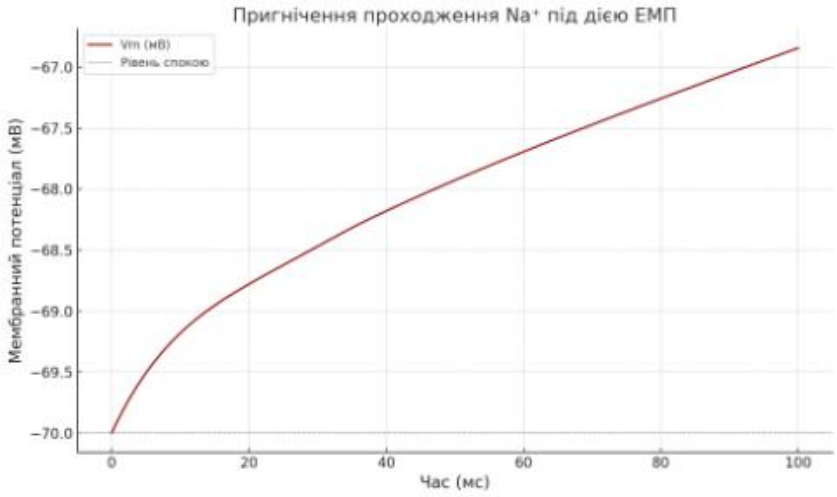


Рис. 5. Пригнічення проходження іонів Na^+ через клітинну мембрану під дією електромагнітного поля

На графіку продемонстровано процес пригнічення проходження іонів Na^+ через клітинну мембрану під дією електромагнітного поля. Помітно, що з часом мембранний потенціал стабілізується на більш негативному рівні порівняно з умовами нормальної провідності, що вказує на розвиток гіперполяризації. Такий ефект є наслідком зростання опору натрієвих каналів, що призводить до зменшення струму Na^+ та імітує інгібуючий вплив електромагнітного поля на збудливість клітини [11, 12].

Отримані дані перебувають у добрій відповідності з результатами інших досліджень, де було зафіксовано аналогічне зниження проникності натрієвих каналів під впливом змінних електромагнітних полів. Це дає підстави вважати запропоновану математичну модель адекватною та науково обґрунтованою.

Висновки

Підтверджено ключову роль іонів натрію (Na^+) у регуляції мембранного потенціалу клітини, генерації електричних імпульсів та забезпеченні критичних життєвих функцій, таких як скорочення м'язової тканини, активація ферментів і запуск апоптозу. Пульсації клітини, що є поєднанням електричних, механічних та метаболічних процесів, розглянуто як важливий показник функціонального стану та міжклітинної взаємодії.

Розроблені математичні моделі – зокрема модель Нав'є–Стокса для опису рідинних коливань усередині клітини, об'ємно-транспортні моделі та модифікована модель Курамото – дозволили дослідити динаміку іонного транспорту та механізми синхронізації клітин під впливом зовнішніх факторів. Увагу приділено впливу електромагнітного поля частотою 50 Гц і амплітудою 100 нА, яке, згідно з результатами моделювання в MATLAB/Simulink, здатне як стимулювати деполяризацію мембрани (підвищення збудливості), так і спричиняти гіперполяризацію (пригнічення) залежно від режиму впливу.

Встановлено, що тривале або цілеспрямоване електромагнітне навантаження може призводити до зниження проникності натрієвих каналів, що імітує інгібуючий вплив на клітину та потенційно дозволяє регулювати її активність. Такий висновок узгоджується з літературними даними, де спостерігалося аналогічне зменшення струму Na^+ під впливом змінних електромагнітних полів.

Розроблена модель є науково обґрунтованою та має високу практичну значущість. Вона може слугувати основою для створення нових методів контролю клітинної активності в біомедичній інженерії, зокрема при розробці технологій електромагнітної терапії, біоелектронних імплантів та систем спрямованого впливу на іонні канали.

Отримані результати поглиблюють розуміння фундаментальних механізмів електрофізіологічної регуляції клітин і відкривають перспективи для подальших експериментальних і прикладних досліджень у сфері біомедичних технологій.

Список літератури:

1. Ушакова Г. О., Недзвецький В. С., Кириченко С. В. Молекулярні механізми міжклітинної комунікації : моногр. Дніпро : ЛІРА, 2018. 216 с.
2. Bergtrom G. Basic cell and molecular biology. 4e. 2020. 628 p.
3. A mechano-osmotic feedback couples cell volume to the rate of cell deformation / L. Venkova et al. // BioRxiv. 2021. P. 1–26.
4. Simulation of microtubule-cytoplasm interaction revealed the importance of fluid dynamics in determining the organization of microtubules / M. Mohammad et al. // Plant direct. 2023. Vol. 7, no. 7. P. 1–18.
5. Glimm T., Gruszka D. Modeling the interplay of oscillatory synchronization and aggregation via cell-cell adhesion. Analysis of PDEs. 2023. P. 17–25.
6. Costa G. S., Novaes M., de Aguiar M. A. M. Bifurcations in the Kuramoto model with external forcing and higher-order interactions // Chaos: an interdisciplinary journal of nonlinear science. 2024. Vol. 34, no. 12.
7. Dexuan X. An extension of goldman-hodgkin-katz equations by charges from ionic solution and ion channel protein // Quantitative biology. 2022. P. 22–30.
8. Hussain Z. Electrophysiology of membrane potentials: mathematical physiology and mathematical medicine // Int. j. biol. biotech. 2022. Vol. 19, no. 2. P. 161–170.
9. Kundenko M., Rudenko A., Mardziavko V. Research on the Method of Improving Fuel Quality for Heat Generators // 2023 IEEE 5th International Conference on Modern Electrical and Energy System (MEES), Kremenchuk, Ukraine, 27–30 September 2023. 2023. URL: https://doi.org/10.1109/mees61502.2023.10402419
10. Hayashi S., Kakikawa M. Exposure to 60 Hz magnetic field can affect membrane proteins and membrane potential in human cancer cells // Electromagnetic biology and medicine. 2021. P. 1–8.
11. Development of a Model of Cell Functioning to Measure the Interaction of Low-Energy EMF / M. Kundenko et al. // 2022 XXXII International Scientific Symposium Metrology and Metrology Assurance (MMA), Sozopol, Bulgaria, 7–11 September 2022.
12. Effects of modulation on sodium and potassium channel currents by extremely low frequency electromagnetic fields stimulation on hippocampal CA1 pyramidal cells / Y. Zheng et al. // Electromagnetic biology and medicine. 2021. Vol. 40, no. 2. P. 274–285.

Надійшла до редколегії 11.09.2025

Відомості про авторів:

Руденко Андрій Юрійович – асистент кафедри електроенергетики, електротехніки та електромеханіки; Миколаївський національний аграрний університет, Україна; e-mail: rudenkoau@mnau.edu.ua; ORCID: https://orcid.org/0000-0002-5103-6412

Мардзявко Віталій Анатолійович – асистент кафедри електроенергетики, електротехніки та електромеханіки; Миколаївський національний аграрний університет, Україна; e-mail: mardzyavko@mnau.edu.ua; ORCID: https://orcid.org/0000-0001-7327-9215

Мартиненко Володимир Олександрович – канд. техн. наук, доцент, Миколаївський національний аграрний університет, доцент кафедри електроенергетики, електротехніки та електромеханіки; Україна; e-mail: martynenko@mnau.edu.ua; ORCID: https://orcid.org/0000-0003-4067-3640

Кунденко Микола Петрович – д-р техн. наук, професор, Національний технічний університет «Харківський політехнічний інститут», завідувач кафедри теплотехніки та енергоефективних технологій; Україна; e-mail: mykola.kundenko@khpi.edu; ORCID: https://orcid.org/0000-0002-5841-4367

УДК 004.94:616.8-073.75 **DOI:10.30837/rt.2025.3.222.25**

Т.В. ЖЕМЧУЖКІНА, канд. техн. наук

КЛАСИФІКАЦІЯ ЕЛЕКТРОМІОГРАФІЧНИХ СИГНАЛІВ ЗА ЇХ ЕНТРОПІЙНИМИ ХАРАКТЕРИСТИКАМИ ДЛЯ ДИФЕРЕНЦІАЛЬНОЇ ДІАГНОСТИКИ БОЛЮ У ПОПЕРЕКУ МЕТОДОМ ВИПАДКОВОГО ЛІСУ

Вступ

Біль у попереку (БП) є однією з провідних причин інвалідності у світі та становить 7,75 % від загальної кількості років життя з інвалідністю. В Україні БП становить 13,22 % від загальної кількості років життя з інвалідністю [1]. Однією з фізичних характеристик осіб із хронічним болем у попереку є підвищена втома м’язів-розгиначів спини під час тесту на витривалість у попереку. Діагностика та лікування болю в попереку є надзвичайно складним завданням, оскільки у більшості пацієнтів недостатньо органічних ознак захворювання, які можна виявити за допомогою рентгенографії чи магнітно-резонансної томографії хребта. В результаті у понад 85 % людей, які страждають від болю в попереку, не вдається виявити структурну патологію, і таким пацієнтам часто ставлять діагноз «періодичний або хронічний біль у попереку». Об’єктивна оцінка БП є актуальним завданням для визначення типу болю у пацієнтів з метою застосування правильної терапевтичної тактики та оцінки реабілітаційних заходів [2–5].

Зазвичай первинна діагностика болю в спині включає рентгенологічне дослідження. З його допомогою можна детально проаналізувати стан хребта. Магнітно-резонансна томографія (МРТ) також використовується для діагностики хребта. Під час процедури МРТ більш чітко видно органи з великим вмістом рідини, приховані за кістками скелета.

Більш універсальним методом є електроміографія (ЕМГ), яка реєструє біологічні потенціали м'язів для оцінки стану периферичних нервів. В ЕМГ використовуються поверхневі та голчасті електроди. Поверхнева електроміографія – безпечний та інформативний метод неінвазивної діагностики болю в спині, що дозволяє записувати сигнали під час м'язового тесту з поверхневих електродів. Зареєстрований сигнал електричної активності м’язів називається електроміограмою (ЕМГ) або електроміографічним сигналом.

Більшість випадкових сигналів на практиці, включаючи поверхневий ЕМГ-сигнал, як правило, мають нестаціонарний характер. Саме тому було запропоновано використовувати методи нелінійної динаміки для аналізу ЕМГ-сигналів [6].

Методи. Ентропія

Для задачі обробки ЕМГ-сигналів застосування теорії динамічних систем, яку також називають теорією хаосу або нелінійною динамікою, може бути більш придатним, ніж використання традиційних лінійних методів аналізу. Методи нелінійної динаміки дозволяють оцінювати властивості сигналу, пов’язані з його складністю, варіабельністю та хаотичністю, що створює додаткові можливості для виявлення прихованих закономірностей та диференціації між нормальними і патологічними станами. Наразі активно використовується така галузь математичної фізики, як аналіз часових рядів, яка включає методи теорії динамічних систем та фрактального аналізу. Такий аналіз застосовується для обробки біомедичної інформації. Це пов'язано з тим, що через складність біологічних процесів прогнозування їхньої динаміки надзвичайно складне. Однак це типова ситуація для хаотичних динамічних систем. Експоненціальний розкид близьких траєкторій для класу хаотичних атракторів визначається найвищим показником Ляпунова, а швидкість «розкиду» елемента фазового об'єму визначається ентропією. Обчислюючи будь-яке з цих значень, можна отримати значення горизонту прогнозу для заданого часового ряду та/або поведінку системи в цілому. Крім того, методи теорії динамічних систем, застосовані до біомедичних часових рядів, дозволяють знайти розмі-

рність вкладення, яка визначає, скільки вимірів потрібно для правильної реконструкції багатовимірного фазового простору.

Одновимірний часовий ряд – це набір з N чисел, що представляють значення деякої змінної $x(t)$, виміряної в моменти часу t_i з інтервалом дискретизації t: $x_i = x(t_i)$, $i = 1,\ldots, N$.

У теорії часових рядів можливі два різні випадки, пов'язані з доступністю інформації про систему, яка їх генерує. Якщо математичні моделі, що описують систему, відомі, то обробка часових рядів передбачає знаходження параметрів, що входять до відомих нам рівнянь. У другому випадку виникають принципові труднощі, оскільки математичний об'єкт, який слід пов'язати з отриманими даними (наприклад, біомедичний часовий ряд), невідомий. Таким чином, існують дві основні проблеми аналізу часових рядів: проблема ідентифікації та проблема прогнозування.

Завдання ідентифікації полягає у визначенні параметрів системи, яка сформувала цей часовий ряд – розмірності вбудовування, розмірності кореляції, ентропії тощо.

Розмірність вкладення – це мінімальна кількість динамічних змінних, які однозначно описують спостережуваний процес. Це значення дуже важливе для складання прогнозу. Кореляційна розмірність є оцінкою фрактальної розмірності системного атрактора та окремим випадком узагальненої ймовірнісної розмірності.

Поняття ентропії пов'язане з передбачуваністю значень ряду та всієї системи. Передбачуваність можна розуміти у двох сенсах: 1) об'ємному, тобто як збільшується фазовий об'єм початкової похибки, та 2) лінійному, тобто як збільшується різниця між істинною та збуреною траєкторією. Ентропія K пов'язана з першою інтерпретацією передбачуваності та визначає час передбачуваності для динамічної системи:

$$T_p \sim \frac{1}{K} \log \frac{1}{\varepsilon}, \tag{1}$$

де ε – відносна похибка даних про її стан.

Друга інтерпретація пов'язана з концепцією показників Ляпунова λ_i. У цьому випадку ми також можемо представити час передбачуваності:

$$T_\lambda \sim \frac{1}{\lambda_{\max}} \log \frac{1}{\varepsilon}. \tag{2}$$

За порядком величини обидві ці оцінки збігаються. Однак на практиці значення часу передбачуваності розраховується за спрощеними формулами: $T_p \sim K^{-1}$ та $T_\lambda \sim \lambda_{\max}^{-1}$.

Ентропія є мірою впорядкованості системи. Наразі для опису цієї властивості в тому чи іншому контексті використовується велика кількість різних характеристик, які також називають ентропією. Так, в теорії інформації ентропія H вводиться для систем, які можуть перебувати в певних станах x_i з певними ймовірностями $p_i = p(x_i)$, використовуючи формулу Шеннона:

$$H = -\sum_{i=1}^{M} p_i \ln(p_i), \tag{3}$$

де p_i – відносна частота i-ї події, M – кількість можливих подій.

Спектральна ентропія – це міра розкиду спектральної щільності потужності сигналу. Вона використовується для часових рядів та обробки сигналів у різних галузях, таких як обробка мови або біомедичних сигналів. Спектральна ентропія характеризує рівномірність розподілу потужності сигналу в частотній області та базується на концепції ентропії Шеннона з теорії інформації. Високе значення спектральної ентропії вказує на широкий та рівномірний спектр, характерний для випадкових або шумних сигналів; низьке значення спектральної ентропії вказує на вузький та концентрований спектр, характерний для періодичних або детермінованих сигналів.

Розрахунки спектральної ентропії сигналу $x(n)$ виконуються наступним чином:

- спектр сигналу $X(m)$ розраховується за допомогою перетворення Фур'є;
- спектральна щільність потужності (СЩП) розраховується так:

$$S(m) = \left| X(m) \right|^2; \tag{4}$$

- ймовірність розподілу спектральної потужності за частотами розраховується шляхом нормалізації СЩП таким чином, щоб загальна спектральна щільність потужності дорівнювала 1:

$$P(m) = \frac{S(m)}{\sum_i S(i)}; \tag{5}$$

- для результуючого розподілу $P(m)$ обчислюється ентропія Шеннона. Це буде спектральна ентропія:

$$H = -\sum_{m=1}^{N} P(m) \log_2 P(m); \tag{6}$$

- нормалізована спектральна ентропія (НСЕ) використовується для порівняння сигналів різної довжини або з різними характеристиками. Максимально можлива ентропія для частотних компонентів дорівнює $\log_2 N$. Нормалізована спектральна ентропія розраховується як відношення спектральної ентропії до її максимального значення:

$$H_n = -\frac{H}{\log_2 N}. \tag{7}$$

Значення НСЕ варіюються від 0 до 1: $H_n = 0$ вказує на те, що сигнал має повністю детермінований спектр (вся потужність зосереджена на одній частоті); $H_n = 1$ вказує на те, що спектр сигналу рівномірно розподілений по всіх частотах, що є типовим для білого шуму.

Миттєва спектральна ентропія (або короткочасна спектральна ентропія) – це міра спектральної ентропії, розрахована за короткі часові вікна в сигналі, що дозволяє аналізувати зміни ентропії з часом. Цей метод корисний для сигналів, які мають характеристики, що змінюються в часі, таких як мова, музика або біомедичні сигнали.

Миттєва спектральна ентропія в момент часу t:

$$H(t) = -\sum_{m=1}^{N} P(t,m) \log_2 P(t,m), \tag{8}$$

де розподіл ймовірностей $P(t,m)$ у момент часу t розраховується на основі відомої спектрограми потужності $S(t,f)$ у вигляді частотно-часової залежності:

$$P(t,m) = \frac{S(t,m)}{\sum_f S(t,f)}. \tag{9}$$

Таким чином, спектральна ентропія є важливим інструментом для аналізу сигналів, що дозволяє кількісно визначити їхню складність та ступінь випадковості.

Використання ентропійних характеристик біосигналів довело свою ефективність у задачах медичної діагностики. Зокрема, автори у [7] продемонстрували застосування ентропійних ознак часових рядів для розпізнавання патернів при неврологічних розладах, у [8] порівняли різні міри ентропійної швидкості для оцінки складності часових рядів і показали їх придатність у дослідженнях варіабельності серцевого ритму та дихання, аналіз багатомасштабної ентропії також був успішно використаний для виявлення зниження складності серце-

вих динамічних процесів у пацієнтів із гіпертиреозом [9], у [10] встановили, що ентропійні показники, отримані на основі емпіричної модової декомпозиції, можуть бути використані в якості діагностичних ознак для ідентифікації осіб похилого віку з історією падінь.

У [11] автори провели порівняльний аналіз електроміографічних сигналів здорової людини та особи з болем у попереку. Вони вивчали ентропію часових рядів ЕМГ як функцію логарифма часу. При триваліших моментах часу (від 0,01 до 1 с) ентропія демонструє плато-подібну поведінку, що означає наявність довготривалих кореляцій у сигналі. Плато знаходиться на рівні ентропії, який значно нижчий за максимально можливе значення ентропії. Отже, це демонстрація внутрішньої властивості часового ряду. Результати аналізу показують, що для здорової людини плато ентропії має вищий рівень, ніж для особи з БП.

Автори в [11] проаналізували та порівняли результати спектрального аналізу, який широко використовується в ЕМГ, з результатами ентропійного аналізу. Вони показали, що група здорових осіб та осіб з БП значно перекривалася за показниками медіанної частоти (МЧ) та нахилу МЧ, тоді як групи чітко відрізнялися за значеннями ентропії. Мінливість МЧ та нахилу МЧ пояснюється існуванням довгострокових кореляцій у сигналі.

Методи. Випадковий ліс

Випадковий ліс (Random Forest) – це метод ансамблевої класифікації, заснований на побудові набору дерев рішень та об'єднанні їх прогнозів для підвищення точності та стійкості моделі. Метод був запропонований Л. Брейманом та належить до сімейства алгоритмів пакетування (boot-strap aggregating).

Основні етапи класифікації за випадковими лісами включають: формування навчальних підмножин, навчання дерев рішень та голосування (ансамблювання) [12, 13].

На етапі формування навчальних підмножин випадковим чином генеруються N вибірок бутстрепу з вихідного навчального набору. Метод бутстрепу – це статистичний метод, в якому з вихідних навчальних даних випадковим чином формуються кілька нових підмножин. Якщо вихідний навчальний набір містить N об'єктів, то кожна вибірка бутстрепу також містить N об'єктів, але може включати повторювані елементи, оскільки кожен об'єкт вибирається випадковим чином і незалежно з можливістю вибору більше одного разу. Цей метод широко використовується в ансамблевих алгоритмах для підвищення стійкості та зменшення перенавчання.

Для кожної вибірки бутстрепу будується одне дерево рішень. Під час етапу навчання дерев кожне дерево навчається на відповідній йому підмножині бутстрепу. Цей процес вводить різноманітність серед дерев. У методі випадкового лісу різноманітність між деревами ще більше посилюється шляхом застосування підвибірки ознак на кожному розщепленні вузла, на відміну від стандартних дерев рішень, де всі ознаки розглядаються на кожному розщепленні, щоб знайти найкращу. Випадкові ліси розглядають лише випадково вибрану підмножину ознак на кожному вузлі, а потім найкраще розщеплення визначається з цієї підмножини. Цей метод зменшує міждеревну кореляцію та підвищує стійкість ансамблю до перенавчання.

На заключному етапі класифікації прогнозування виконується за допомогою голосування більшості. Процес голосування більшості – це метод агрегації рішень кількох моделей (в ансамблі), за якого остаточний прогноз визначається на основі більшості голосів від окремих моделей.

Нехай $\{h_1(x), h_2(x), ..., h_N(x)\}$ – прогнози окремих дерев для вхідного вектора x, тоді кінцевий класифікатор визначається як

$$H(x) = \text{mode}\left(\{h_i(x)\}_{i=1}^{N}\right), \tag{10}$$

де $\text{mode}(\cdot)$ – функція, яка повертає найчастіше значення (клас).

Перевагами класифікатора випадкового лісу є висока стійкість до перенавчання, ефективність на незбалансованих наборах даних (з використанням вагових коефіцієнтів класів) та можливість оцінки важливості ознак.

Голосування більшості робить остаточне рішення більш стійким до помилок окремих моделей, забезпечуючи взаємну компенсацію помилок по всьому ансамблю. Голосування більшості є фундаментальним принципом ансамблевого навчання, який використовується в таких методах, як випадковий ліс, баггінг та деякі ансамблі нейронних мереж.

У [14] автори продемонстрували потенціал використання методу випадкового лісу для класифікації сигналів ЕКГ з метою підвищення точності діагностики. У [15] було продемонстровано надійну роботу алгоритму випадкового лісу для класифікації пацієнтів з біполярним розладом та здорових осіб на основі даних електроенцефалограми.

Таким чином, метод випадкового лісу був обраний як класифікатор для ентропійних показників ЕМГ сигналів.

Експерименти та результати

Вихідні дані для дослідження складалися з двох наборів електроміографічних сигналів, записаних з довгого м'яза-розгинача тулуба поперекового відділу хребта за однакових умов у різні періоди часу. Дані ЕМГ були отримані у співпраці з відділом патофізіології, анестезіології та функціональної діагностики Інституту патології хребта та суглобів ім. Ситенка Національної академії медичних наук України.

Кожен з двох наборів даних містить ЕМГ-сигнали осіб з трьох діагностичних груп: 1 – здорові особи без скарг на біль у спині; 2 – умовно здорові особи зі скаргами на біль у спині, тобто особи без виявленої органічної патології (так званий дисфункціональний біль); 3 – особи з болем у спині, які мають дегенеративні захворювання хребта (функціональний біль). У першому наборі даних кількість сигналів становила 96 (група 1), 46 (група 2) та 235 (група 3); у другому наборі даних – 33 (група 1), 38 (група 2), 72 (група 3).

Сигнали з першого набору були використані для навчання класифікатора. Отримані моделі тестувалися на другому наборі даних.

ЕМГ-сигнали були записані з поперекового відділу хребта L4-L5 з довгого м'яза-розгинача тулуба. Сигнали не групувалися за статтю та віком. Всі сигнали були попередньо оброблені шляхом фільтрації смуговим фільтром Баттерворта 4-го порядку зі смугою від 10 до 500 Гц. Частоту дискретизації записаних сигналів було зменшено до 1024 Гц.

В [16] проаналізовано значення ентропії за (3) для M=100 для послідовних грубозернистих часових рядів, отриманої методом багатомасштабної ентропії, описаним у [17]. Грубозернистий часовий ряд було побудовано шляхом усереднення послідовно зростаючої кількості точок даних у неперекриваючихся вікнах. Кожен елемент грубозернистого часового ряду розраховується відповідно до рівняння:

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \ , \quad 1 \le j \le \frac{N}{\tau}, \tag{11}$$

де τ – масштабний коефіцієнт ($10^2 \le \tau \le 10^4$) відповідно до інтервалу плато. Для масштабу 1 грубозернистий часовий ряд – це вихідний часовий ряд.

Було розраховано середнє, медіанне та максимальне значення ентропії.

Миттєва спектральна ентропія за (8), її середнє, медіанне та максимальне значення, а також нормалізована спектральна ентропія за (7) були розраховані для всіх сигналів.

Всі розраховані параметри були перевірені на нормальність розподілу за допомогою тесту Ліллієфорса. У цьому тесті нульова гіпотеза полягає в тому, що дані нормально розподілені. Для всіх розрахованих значень результат тесту відхиляє нульову гіпотезу на рівні значущості 5 %. Це означає, що всі розраховані параметри не є нормально розподіленими.

Для перевірки всіх груп на різницю за всіма розрахованими параметрами було використано двосторонній тест рангової суми Вілкоксона, оскільки всі дані не відносяться до нор-

мального розподілу. Критерій суми рангів Вілкоксона перевіряє нульову гіпотезу про те, що медіани двох незалежних вибірок не відрізняються. Гіпотезу було перевірено на рівні значення 5 %.

Результати перевірки гіпотези показали, що групи здорових без болю та здорових з болем, здорових без болю та осіб з функціональним болем; здорових з болем та осіб з функціональним болем не відрізняються за показниками максимальної спектральної ентропії, спектральної ентропії, медіанної ентропії, середньої ентропії, максимальної ентропії.

Два показники (медіанна спектральна ентропія та середня спектральна ентропія) різні для всіх груп, окрім груп здорових людей з болем та осіб з вертеброгічними болями та груп осіб з функціональним болем та осіб зі сколіозом.

Групи осіб з функціональним болем та осіб зі сколіозом та групи осіб з функціональним болем та осіб зі сколіозом відрізняються лише медіанною ентропією та середньою ентропією. Отже, ці два показники можна використовувати разом із медіанною спектральною ентропією та середньою спектральною ентропією для диференціальної діагностики болю в попереку.

Чотири показники, виділені в [16], а саме медіанна ентропія, середня ентропія, медіанна спектральна ентропія та середня спектральна ентропія, було проаналізовано класифікатором на основі випадкового лісу з метою диференціальної діагностики болю в попереку.

Аналізувалися можливості класифікації даних ентропійних показників ЕМГ сигналів з різних комбінацій діагностичних груп: здорові без болю та здорові із скаргами на біль (групи 1 та 2); здорові без болю та хворі на вертеброгічні захворювання (групи 1 та 3); здорові із скаргами на біль та хворі на вертеброгічні захворювання (групи 2 та 3), а також всі три групи разом.

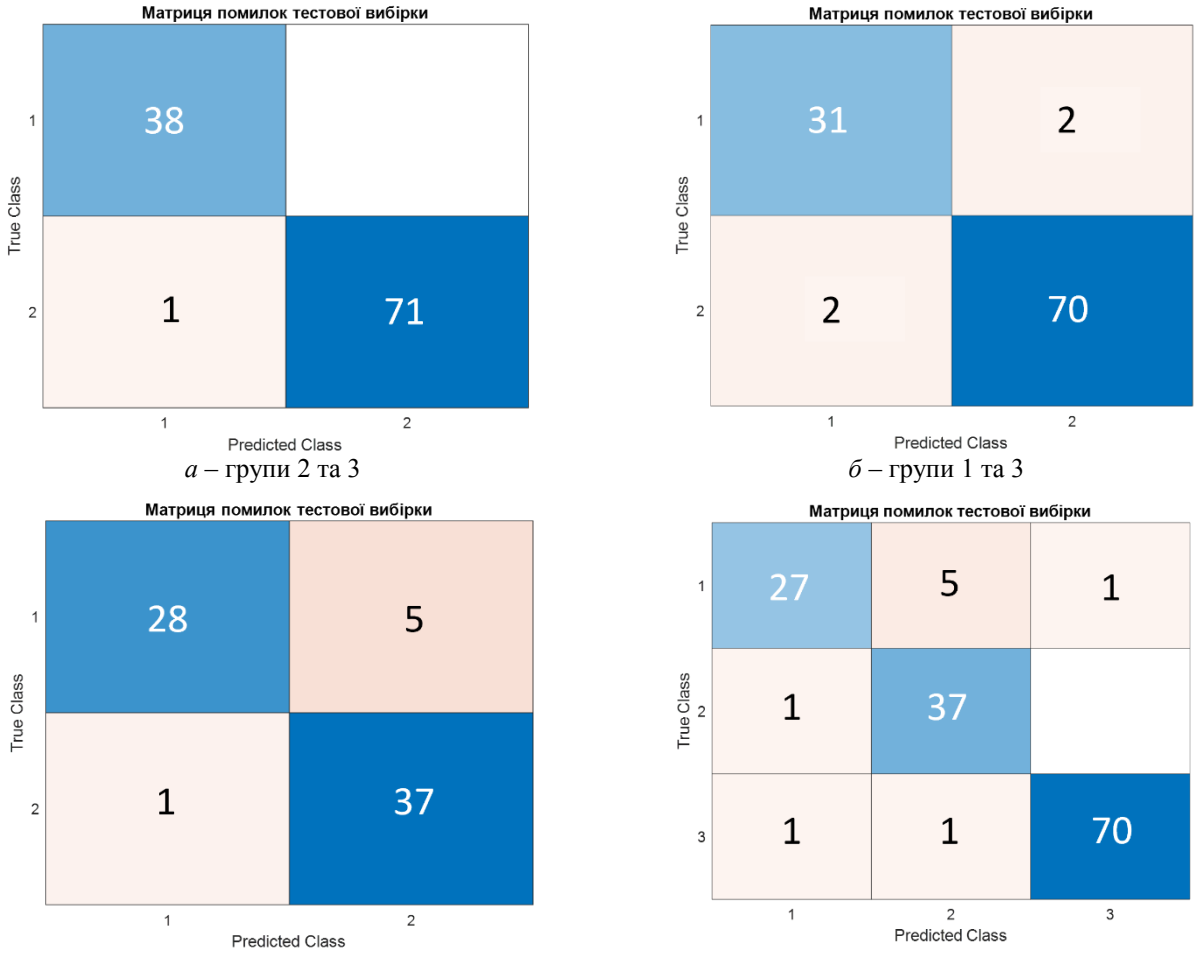
Для дослідження ефективності методу випадкового лісу в задачі класифікації було проведено низку обчислювальних експериментів. Навчання моделей здійснювалося на першому наборі даних, що характеризується вираженим дисбалансом класів. Отримані моделі тестувалися на другому наборі даних, що є незалежною вибіркою, що не перетинається з навчальною.

У ході експерименту варіювалося ключове гіперпараметричне значення класифікатора – кількість дерев в ансамблі. Було розглянуто чотири конфігурації випадкового лісу: 100, 200, 300 та 500 базових класифікаторів. Для кожної конфігурації алгоритм навчався на першій вибірці та оцінювався на другій вибірці.

Особлива увага приділялася впливу дисбалансу класів на якість класифікації. З цією метою досліджувалися дві стратегії навчання: без урахування ваги класів, коли всі об'єкти навчальної вибірки мали рівний внесок у функцію помилки, та зі зважуванням даних, коли ваги об'єктів задавалися обернено пропорційно їх представленості у вибірці. Величини ваг розраховувалися з урахуванням гістограми розподілу даних за класами, що дозволило компенсувати зміщення у бік переважаючого класу.

Таким чином, підсумковий експериментальний план включав вісім сценаріїв (чотири значення числа дерев по два режими зважування класів). Моделі, отримані на навчальній вибірці для кожного зі сценаріїв, оцінювалися на тестовій вибірці за допомогою комплексу показників якості. Зокрема, обчислювалися точність класифікації, F1-міра як збалансований показник точності та повноти, а також площа під ROC-кривою (AUC) для оцінки дискримінативної здатності моделі. Додатково будувалася ROC-крива, що дозволяє візуально проаналізувати компроміс між чутливістю та специфічністю, та формувалася матриця помилок, яка дає можливість детально дослідити характер і структуру помилок класифікації.

Матриці помилок та ROC-криві для найкращих результатів для кожної з комбінацій груп наведено на рис. 1 та 2 відповідно.



а – групи 2 та 3

б – групи 1 та 3

в – групи 1 та 2 (незважені)

г – групи 1 та 2 та 3 (незважені)

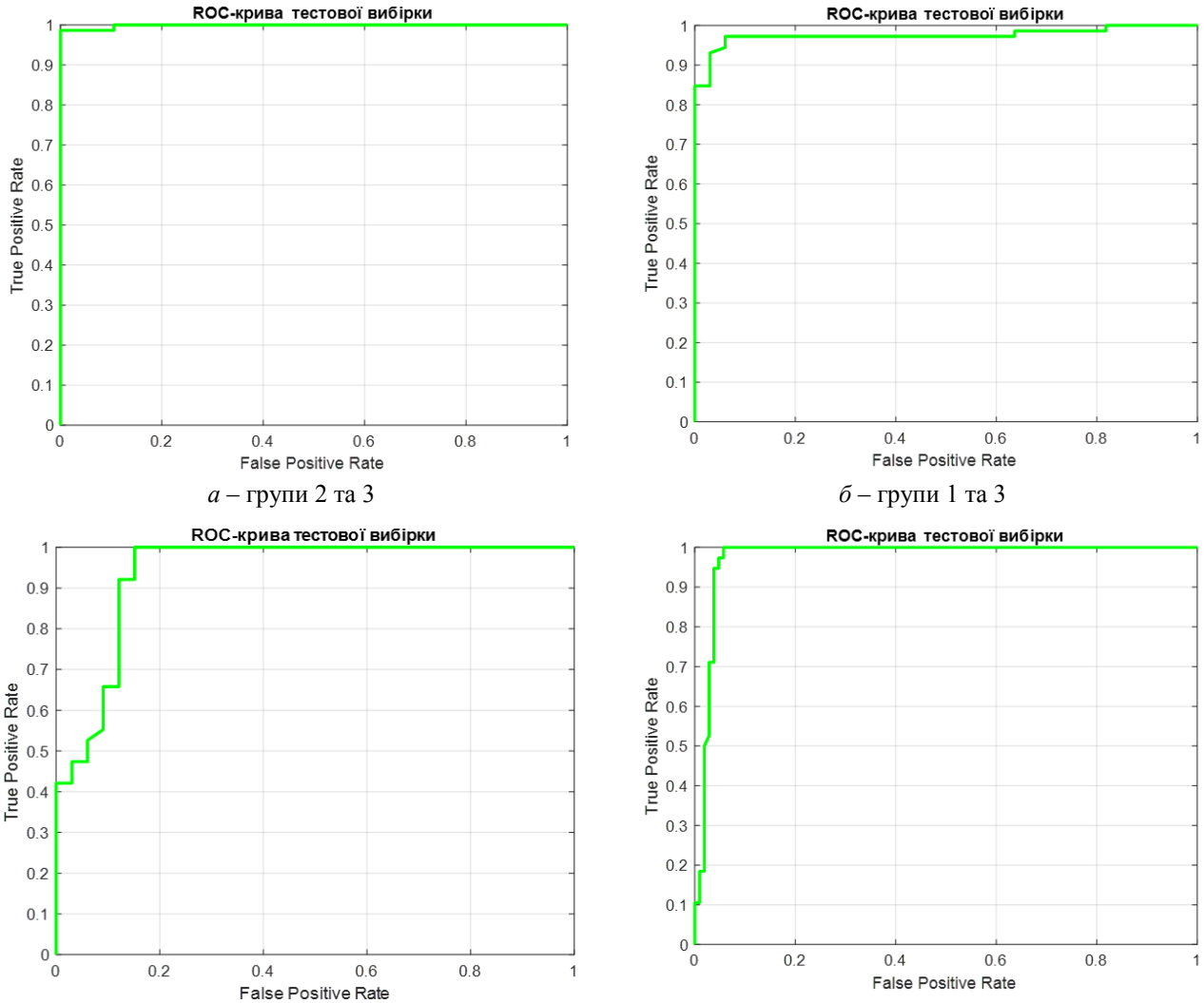
Рис.1. Матриці помилок для найкращих результатів для кожної з комбінацій груп

В табл. 1 наведено якісні показники класифікації, найкращі для кожної комбінації діагностичних груп серед усіх проведених сценаріїв.

Таблиця 1

Групи*	Розмір ліса	Ваги	F1-міра	AUC	Точність, %
2 та 3	100	1:1	0.99	1.00	99.09
1 та 3	100	1:1	0.94	0.97	96.19
1 та 2	100	за гістограмою	0.92	0.92	92.96
	100	1:1	0.90	0.94	91.55
1 та 2 та 3	100	1:1:1	0.90	0.98	93.71
	100	за гістограмою	0.92	0.99	93.71

* 1 – здорові особи без скарг на біль у спині; 2 – умовно здорові особи зі скаргами на біль у спині (дисфункціональний біль); 3 – особи з болем у спині (функціональний біль).



а – групи 2 та 3

б – групи 1 та 3

в – групи 1 та 2 (незважені)

г – групи 1 та 2 та 3 (незважені)

Рис. 2. ROC-криві для найкращих результатів для кожної з комбінацій груп

Виявилося, що збільшення розміру випадкового лісу до 500 не покращувало якість класифікації, крім того, не зважаючи на дисбаланс класів у навчальній виборці, зважування навчальних даних або не покращувало результат, або давало зіставний результат.

Найкраща дискримінативна здатність виявилася у моделі для розділення груп за типом болю: дисфункціональний чи функціональний (групи 2 та 3), що є максимально корисним при діагностиці болю у попереку.

Висновки

У роботі досліджувалася можливість застосування показників ентропії для диференціальної діагностики болю в попереку за допомогою електроміографічних сигналів. Обчислювальні експерименти із застосуванням методу випадкового лісу для класифікації ентропійних показників ЕМГ-сигналів підтвердили ефективність цього підходу в задачах диференціальної діагностики болю у попереку. Оптимальне значення кількості дерев в ансамблі становило 100. Подальше збільшення розміру випадкового лісу до 500 дерев не призводило до істотного підвищення якості класифікації. Врахування дисбалансу класів шляхом зважування навчальних даних не дало суттєвого покращення результатів. У більшості випадків моделі зі зважуванням показували нижчу або зіставну з незваженими результативність. Найвищі показники F1-міри (0.99), точності (99.09 %) та AUC (1.00) отримано при класифікації груп 2 (умовно здорові із скаргами на біль, дисфункціональний біль) та 3 (пацієнти з функціональним болем). Це свідчить про найбільшу дискримінативну здатність моделі саме у задачі

розмежування типів больового синдрому. Результати експериментів демонструють, що використання ентропійних характеристик ЕМГ-сигналів у поєднанні з методом випадкового лісу є перспективним підходом для створення інструментів підтримки клінічних рішень при діагностиці болю у попереку.

Список літератури:

1. Institute for Health Metrics and Evaluation (IHME). Global Burden of Disease 2021: Findings from the GBD 2021 Study. Seattle, WA: IHME, 2024.
2. P. Santos-Moreno, J. A. Sucerquia-Quintero, and R. García-Salinas. Chronic low back pain: Diagnostic Approach for Primary Care // Revista Colombiana de Reumatología (English Edition). Vol. 29, no. 4. P. 303–309, Oct. 2022. doi:10.1016/j.rcreue.2021.02.006
3. Walid Kamal Abdelbasset and Abdelmoneim Sulieman. An Overview on Low Back Pain and Functional Disability: Associated Risk Factors and Management // JDR. 2022. Vol. 1(1):P. 19–22. DOI: 10.57197/JDR-2022-0004.
4. K. Ammer, G. Ebenbichler, T. Bochdansky. Low Back Pain – A Disease or Condition of Impaired Functional Health? Definition-Inherent Consequences for the Comprehensive Care of Back Pain Patients // BioMed. 2022. Vol. 2(2). P.270–281. https://doi.org/10.3390/biomed2020022.
5. T. Zhemchuzhkina, I. Kurochkin. Classification of Functional and Dysfunctional Low Back Pain by Spectral Indicators of Electromyogram // 2023 IEEE 18th International Conference on Computer Science and Information Technologies (CSIT), Lviv, Ukraine, 2023. P. 1–4. doi: 10.1109/CSIT61576.2023.10324058.
6. T. Zhemchuzhkina. Second Order Difference Plot as a Tool for Low Back Pain Differentiation by Electromyographic Signals // 2023 IEEE 4th KhPI Week on Advanced Technology (KhPIWeek), Kharkiv, Ukraine, 2023. P. 1–4. doi: 10.1109/KhPIWeek61412.2023.10312961.
7. Y. Huang, Y. Zhao, A. Capstick, F. Palermo, H. Haddadi, P. Barnaghi. Analyzing entropy features in time-series data for pattern recognition in neurological conditions // Artificial Intelligence in Medicine. 2024. Vol. 150. https://doi.org/10.1016/j.artmed.2024.102821.
8. C. Barà, R. Pernice, C. A. Catania, M. Hilal, A. Porta, A. Humeau-Heurtier, L. Faes. Comparison of entropy rate measures for the evaluation of time series complexity: Simulations and application to heart rate and respiratory variability // Biocybernetics and Biomedical Engineering. 2024. Vol. 44, Iss. 2. P. 380–392, https://doi.org/10.1016/j.bbe.2024.04.004.
9. J.-L. Chen, H.-S. Shen, S.-Y. Peng, H.-M. Wang. Reduced System Complexity of Heart Rate Dynamics in Patients with Hyperthyroidism: A Multiscale Entropy Analysis // Entropy. 2022. Vol. 24(2). P. 258. https://doi.org/10.3390/e24020258.
10. L.-W. Chou, K.-M. Chang, Y.-C. Wei, M.-K. Lu. Empirical Mode Decomposition-Derived Entropy Features Are Beneficial to Distinguish Elderly People with a Falling History on a Force Plate Signal // Entropy. 2021. Vol. 23(4). P. 472. https://doi.org/10.3390/e23040472.
11. P.S. Sung, U. Zurcher, M. Kaufman. Comparison of spectral and entropic measures for surface electromyography time series: a pilot study // J Rehabil Res Dev. 2007. Vol. 44(4). P. 599–609. doi: 10.1682/jrrd.2006.10.0132. PMID: 18247257.
12. H.A. Salman, A. Kalakech, and A. Steiti. Random Forest Algorithm Overview // Babylonian Journal of Machine Learning. 2024. Vol. 2024. P. 69–79. doi: 10.58496/BJML/2024/007.
13. Zhu, Tongtian. Analysis on the Applicability of the Random Forest // Journal of Physics: Conference Series. 2020. 1607. 012123. doi:10.1088/1742-6596/1607/1/012123.
14. S.K. Mohapatra, M.N. Mohanty. Big Data Analysis and Classification of Biomedical Signal Using Random Forest Algorithm // Patnaik S., Ip A., Tavana M., Jain V. (eds). New Paradigm in Decision Science and Management. Advances in Intelligent Systems and Computing. 2020. Vol. 1005. Springer, Singapore. https://doi.org/10.1007/978-981-13-9330-3_20.
15. M. Suárez, A.M. Torres, P. Blasco-Segura, J. Mateo. Application of the Random Forest Algorithm for Accurate Bipolar Disorder Classification // Life. 2025. Vol. 15(3). P.394. https://doi.org/10.3390/life15030394.
16. T. Zhemchuzhkina. Analysis of Entropy Indicators of Electromyographic Signals for Differential Diagnostics of Low Back Pain // 2024 IEEE 5th KhPI Week on Advanced Technology (KhPIWeek). Kharkiv, Ukraine, 2024. P. 1–4. doi: 10.1109/KhPIWeek61434.2024.10878045.
17. M. Costa, A.L. Goldberger, C.K. Peng. Multiscale entropy analysis of complex physiologic time series // Phys Rev Lett. 2002. Aug 5;89(6):068102. doi: 10.1103/PhysRevLett.89.068102. Epub 2002 Jul 19. PMID: 12190613.

Надійшла до редколегії 15.07.2025

Відомості про автора:

Жемчужкіна Тетяна Володимирівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри біомедичної інженерії; Україна; e-mail: tatyana.zhemchuzhkina@nure.ua; ORCID: https://orcid.org/0000-0001-8884-5099

ABSTRACTS РЕФЕРАТИ

SYSTEMS AND METHODS OF INFORMATION PROTECTION
СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

UDC 004.056.5

Optimization of digital signature calculation and verification operations for the FIPS 205 standard. Part 2 / *I.D. Gorbenko, Ye.G. Kachko, Ya.A. Derevianko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 7 – 21.

Currently, hash-based signatures are among the most promising candidates for post-quantum digital signatures. Their advantage is their comprehensible security and robustness and ease of evaluation, as they rely solely on the reliability of cryptographic hash functions.

The previous article discussed and proposed practical improvements to optimize the FIPS 205 algorithm based on the use of parallel computing. This was achieved mainly by optimizing the SHAKE and SHA algorithms. The importance of optimizing hash value computation is related to the fact that hashing is the main operation in FIPS 205. Previous research has shown that common algorithms for key generation, electronic signature creation, and verification consist of sequential steps, each of which applies the result of the previous step, which excludes the possibility of using parallel computing for these algorithms.

This paper discusses optimization methods and results, including those achieved through parallel threads when implementing individual algorithm steps. Optimization through the use of AVX operations is not considered. Basically, the improvement in the performance of individual functions is achieved through more efficient execution of the basic PRF, Tl, H, and F operations, the optimization of which was discussed in the previous article, as well as through the optimization of certain algorithms that are part of other algorithms.

The results obtained show that the implemented improvements allow for acceleration of at least 2 times for all functions and all modes. However, for most functions and modes, the acceleration is more than threefold. The use of parallel computing through the use of multi-core processors significantly increases the performance of functions for WOTS and FORS schemes, as well as the functions that use them. This improvement is very relevant, since the vast majority of modern processors are multi-core.

Key words: post-quantum standards; FIPS 205; parallel computing; optimization; hash functions; SHA; SHAKE.

7 tab. 15 fig. Ref: 3 items.

УДК 004.056.5

Оптимізація операцій обчислення та перевірки цифрового підпису для стандарту FIPS 205. 2 частина / *І.Д. Горбенко, О.Г. Качко, Я.А. Дерев'янко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 7 – 21.

На даний час підписи на основі гешу одні із найперспективніших кандидатів на постквантові цифрові підписи. Їх перевагою є зрозумілість безпеки та стійкості та простота оцінки, оскільки вони покладаються виключно на надійність криптографічних геш функцій.

У попередній статті розглянуто та запропоновано практичні удосконалення з метою оптимізації ЦП для алгоритму FIPS 205 на основі застосування паралельних обчислень. Це було досягнуто в основному за рахунок оптимізації алгоритмів SHAKE та SHA. Важливість оптимізації обчислення геш значень пов'язана з тим, що гешування є основною операцією у FIPS 205. У попередньому дослідженні показано, що загальні алгоритми генерації ключів, вироблення та перевірки електронного підпису складаються з послідовних кроків, кожний з яких застосовує результат попереднього кроку, що виключає можливість застосування паралельних обчислень для цих алгоритмів.

В даній роботі розглядаються засоби та результати оптимізації, в тому числі за рахунок паралельних потоків при реалізації окремих кроків алгоритмів. Оптимізація за рахунок застосування операцій AVX не розглядається. В основному покращення швидкодії окремих функцій досягається за рахунок більш ефективного виконання базових операцій PRF, Tl, H, F, оптимізацію яких розглянуто в попередній статті, а також за рахунок оптимізації певних алгоритмів, які є складовою частиною інших алгоритмів.

Отримані результати показують, що впроваджені удосконалення дозволяють отримати прискорення не менше ніж в два рази для усіх функцій і режимів. Проте, для більшості функцій та режимів прискорення більше ніж в три рази. Застосування паралельних обчислень за рахунок застосування багатоядерних процесорів суттєво збільшує продуктивність функцій для схем WOTS та FORS, а також функцій, які їх застосовують. Таке покращення є дуже актуальним, оскільки переважна більшість сучасних процесорів багатоядерна.

Ключові слова: постквантові стандарти; FIPS 205; паралельні обчислення; оптимізація; функції гешування; SHA; SHAKE.

Табл. 7. Іл. 15. Бібліогр.: 3 назв.

UDC 004.05

Implementation of zero trust architecture based on the proposed model to ensure enterprise cybersecurity / *V.V. Borodavka, V.I. Yesin* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 22 – 54.

Modern enterprises face increasingly complex cybersecurity challenges, requiring a new approach to protecting digital assets that will ensure secure access to corporate resources anytime, anywhere, as well as their effective function wherever they are located. Traditional approaches, such as perimeter-based security models, no longer guarantee an adequate level of security and are unable to effectively counter modern cyber threats, as enterprise infrastructure is undergoing significant changes related to the expansion of the attack surface, the growth in the number of connected devices, the development of artificial intelligence, the use of cloud technologies, and remote access. Given these changes and challenges, more organizations are turning their attention to a new concept and architecture of protection that can satisfy new requirements for information security and cybersecurity. Such a concept is currently the security paradigm known as “zero trust”, which is based on the principle of “never trust, always verify” and is defined as one of the most effective approaches to countering modern cyber threats. Given the dynamics of cyber threats and the architectural complexity of modern IT enterprises, the successful implementation of zero trust principles requires a systematic approach and involves the integration of modern technologies and security mechanisms. However, despite its obvious advantages, the process of implementing zero trust architecture in corporate information systems is accompanied by significant difficulties from both a technical and organizational point of view. The aim of this paper is to develop a model for implementing zero trust architecture to organize and ensure enterprise cybersecurity. The results presented in this paper are intended to help security specialists use the recommendations provided for the practical implementation of zero trust architecture in their IT enterprises in accordance with the proposed model. Specifically, the paper briefly discusses some recommendations for choosing zero trust architecture deployment models for various types of enterprise activities, and the proposed model for implementing zero trust architecture can help to understand the fundamental changes in the approach to organizing cybersecurity, and effectively implement the zero trust concept, considering the technical and organizational capabilities and requirements of a specific IT enterprise. The proposed zero trust architecture implementation model can serve as a guideline for the effective migration to zero trust architecture in modern digital enterprises, and its use opens significant prospects for the development of more reliable and scalable systems for protecting the information resources of IT enterprises, which is extremely relevant in today's digital cyberspace.

Key words: zero trust; zero trust architecture; zero trust architecture deployment models; information security; cybersecurity.

3 tab. 6 fig. Ref: 64 items.

УДК 004.05

Впровадження архітектури нульової довіри на основі запропонованої моделі для забезпечення кібербезпеки підприємства / *В.В. Бородавка, В.І. Єсін* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 22 – 54.

Сучасні підприємства стикаються з дедалі складнішими викликами у сфері кібербезпеки, що потребує нового підходу до захисту цифрових активів, який дозволить забезпечити безпечний доступ у будь-який час і в будь-якому місці до власних корпоративних ресурсів, а також їхнє ефективне функціонування незалежно від того, де вони розташовані. Традиційні підходи, зокрема модель захисту на основі периметра, вже не гарантують належного рівня безпеки та не спроможні ефективно протистояти сучасним кіберзагрозам, оскільки інфраструктура підприємств зазнає значних змін, пов’язаних з розширенням поверхні атак, зростанням обсягу під’єднаних пристроїв, розвитком штучного інтелекту, використанням хмарних технологій та віддаленого доступу. Зважаючи на ці зміни та виклики, все більше організацій звертають увагу на нову концепцію та архітектуру захисту, яка здатна задовольнити нові вимоги до інформаційної безпеки, кібербезпеки. Такою концепцією на даний момент є парадигма безпеки, що отримала назву «нульова довіра», яка ґрунтується на принципі «ніколи не довіряй, завжди перевіряй» і визначається як один з найефективніших підходів для протидії сучасним кіберзагрозам. Враховуючи динаміку розвитку кіберзагроз та архітектурну складність сучасних ІТ-підприємств, успішна реалізація принципів нульової довіри вимагає системного підходу та передбачає впровадження сучасних технологій і механізмів захисту. Проте, попри очевидні переваги, процес впровадження архітектури нульової довіри в корпоративні інформаційні системи супроводжується значними труднощами як з технічної, так і з організаційної точки зору. Метою роботи є розробка моделі впровадження архітектури нульової довіри для організації та забезпечення кібербезпеки підприємства. Представлені в роботі результати покликані допомогти спеціалістам з безпеки використати надані рекомендації щодо практичного впровадження архітектури нульової довіри на своїх ІТ-підприємствах відповідно до запропонованої моделі. А саме, наведені у стислому викладі деякі рекомендації щодо вибору моделей розгортання архітектури нульової довіри для різного роду діяльності підприємств, а також запропонована модель впровадження архітектури нульової довіри можуть допомогти зрозуміти фундаментальні зміни у підході до організації кібербезпеки, а також ефективно впровадити концепцію нульової довіри з урахуванням технічних та організаційних можливостей і вимог конкретного ІТ-підприємства. Запропонована модель впровадження архітектури нульової довіри може слугувати орієнтиром для ефективного переходу до архітектури нульової довіри в сучасних цифрових підприємства, а її використання відкриває значні перспективи для розвитку більш надійних та масштабованих систем захисту інформаційних ресурсів ІТ-підприємства, що є надзвичайно актуальним у сучасному цифровому кіберпросторі.

Ключові слова: нульова довіра; архітектура нульової довіри; моделі впровадження архітектури нульової довіри; інформаційна безпека; кібербезпека.

Табл. 3. Іл. 6. Бібліогр.: 64 назв.

UDC 621.396:002.53

Radio control of radiation from radio electronic devices. Problems and solutions *V.M. Bezruk, Y.M. Holoborodko, V.I. Zabolotnyi, M.S. Skybenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 55 – 61.

This article analyzes the challenges of radio monitoring in the decameter band, which is crucial for ensuring reliable radio communications. It highlights key issues stemming from the propagation characteristics of decameter waves (surface and sky waves), leading to high spectrum congestion and making it difficult to detect target radio emissions (RE). The drawbacks of existing manual and visual-instrumental monitoring methods are examined, characterized by low speed and an inability to effectively detect short-duration signals.

An approach to automating radio monitoring is proposed, utilizing a priori information about radio emissions, specifically their polarization characteristics, carrier frequency, and source location zone. The effectiveness of «thinning» the stream of analyzed signals is substantiated, allowing for a significant reduction in the number of channels requiring detailed analysis. It is shown that automatic selection of far-zone emissions can reduce the scanning time of the band by tens or hundreds of times. The problem of multi-alternative detection of specified signals under conditions of a priori uncertainty is considered, along with the importance of using onboard computing resources for the flexibility and adaptability of radio monitoring systems. The research results confirm the possibility of significantly improving the speed and quality of decameter band radio monitoring through the implementation of the proposed methods.

Key words: radio monitoring; decameter band; short waves; radio emissions; thinning; a priori information; multi-alternative detection; polarization characteristics; automation.

Ref: 14 items.

УДК 621.396:002.53

Радіоконтроль випромінювань радіоелектронних засобів. Проблеми та шляхи вирішення / *В.М. Безрук, Ю.М. Голобородько, В.І. Заболотний, М.С. Скибенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 55 – 61.

Статтю присвячено аналізу проблематики радіоконтролю декаметрового діапазону, що є критично важливим для забезпечення надійності радіозв'язку. Висвітлено основні виклики, пов'язані з особливостями поширення декаметрових хвиль (поверхневі та просторові), що призводить до високого завантаження спектра та ускладнює виявлення цільових радіовипромінювань. Проаналізовано недоліки існуючих ручних та візуально-апаратних методів контролю, які характеризуються низькою швидкістю та нездатністю ефективно виявляти короткочасні сигнали.

Запропоновано підхід до автоматизації радіоконтролю шляхом використання апріорної інформації про радіовипромінювання, зокрема, їхніх поляризаційних характеристик, несучої частоти та зони розміщення джерела. Обґрунтовано ефективність «проріджування» потоку аналізованих сигналів, що дозволяє значно зменшити кількість каналів для детального аналізу. Показано, що автоматична селекція випромінювань дальньої зони може скоротити час огляду діапазону в десятки або сотні разів. Розглянуто завдання багатоальтернативного виявлення заданих сигналів за умов апріорної невизначеності та важливість використання бортових обчислювальних засобів для гнучкості та адаптивності систем радіоконтролю. Результати дослідження підтверджують можливість значного підвищення швидкості та якості радіоконтролю декаметрового діапазону за рахунок впровадження запропонованих методів.

Ключові слова: радіоконтроль; декаметровий діапазон; короткі хвилі; радіовипромінювання; проріджування; апріорна інформація; багатоальтернативне виявлення; поляризаційні характеристики; автоматизація.

Бібліогр.: 14 назв.

UDC 004.056.5

Features of constructing nonlinear transformations of block symmetric ciphers / *I.V. Lysytska, K.E. Lysytskyi, I.M. Haltseva, E.P. Kolovanova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 62 – 68.

The article is devoted to the peculiarities of constructing nonlinear transformations of block symmetric ciphers. The history of the emergence of nonlinear transformations of block symmetric ciphers is briefly reviewed and the evolution of block symmetric ciphers after DES. The paper analyzes the methods of constructing S-blocks (Substitution-boxes), which are the main component that implements nonlinear substitutions and ensures cryptographic stability of the block symmetric cipher. The authors investigate and compare four main types of S-blocks: deterministic (fixed), dynamic (key-dependent), chaotically generated and random ones. The article analyzes in detail the advantages and disadvantages of each approach, as well as the key cryptographic properties ensuring that ensure stable nonlinear transformations, such as high nonlinearity, low differential homogeneity, bijectivity, avalanche effect, balance, high algebraic degree. The influence of post-quantum cryptography conditions is considered, in particular, the opposition to the Grover algorithm. The authors conclude that the choice of the S-block structure is a compromise between security, performance and ease of implementation. In the context of post-quantum symmetric block ciphers, the use of verified, fixed S-blocks combined with a key length sufficient to protect against quantum attacks is preferred. The reasons are the same as in the classical case: the complexity of security analysis, the computational cost of generation, and the problems associated

with reproducibility. However, hybrid approaches combining algebraic and heuristic methods can also be used to achieve optimal results.

Key words: cybersecurity; cryptography; cryptanalysis; nonlinear transformations; S-blocks; quantum computing; Grover's algorithm.

1 tab. Ref: 24 items.

УДК 004.056.5

Особливості побудови нелінійних перетворень блокових симетричних шифрів / *І.В. Лисицька, К.Є. Лисицький, І.М. Гальцева, Є.П. Колованова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 62 – 68.

Статтю присвячено особливостям побудови нелінійних перетворень блокових симетричних шифрів. Стисло розглянута історія виникнення нелінійних перетворень блокових симетричних шифрів та еволюція блокових симетричних шифрів після DES. У роботі аналізуються методи побудови S-блоків (Substitution-boxes), які є основним компонентом, що реалізує нелінійні заміни та забезпечує криптографічну стійкість блокового симетричного шифру. Автори досліджують та порівнюють чотири основні типи S-блоків: детерміновані (фіксовані), динамічні (ключозалежні), хаотично згенеровані та випадкові. Детально проаналізовано переваги та недоліки кожного підходу, а також ключові криптографічні властивості, що забезпечують криптографічно стійке нелінійне перетворення, такі як висока нелінійність, низька диференціальна однорідність, бієктивність, лавинний ефект, збалансованість, високий алгебраїчний ступінь. Розглядається вплив умов постквантової криптографії, зокрема протидія алгоритму Гровера. Автори роблять висновок, що вибір структури S-блоку є компромісом між безпекою, продуктивністю та простотою реалізації. В контексті постквантових симетричних блокових шифрів, перевага надається використанню перевірених, фіксованих S-блоків у поєднанні з достатньою довжиною ключа для захисту від квантових атак. Причини ті ж, що й у класичному випадку: складність аналізу безпеки, обчислювальні витрати на генерацію та проблеми з відтворюваністю. Однак гібридні підходи, що поєднують алгебраїчні та евристичні методи, також можуть бути використані для досягнення оптимальних результатів.

Ключові слова: кібербезпека; криптографія; криптоаналіз; нелінійні перетворення; S-блоки; квантові обчислення; алгоритм Гровера.

Табл. 1. Бібліогр.: 24 назв.

UDC 004.056.55

Zero-knowledge proof protocols: theoretical foundations and applications in modern cryptography / *R.I. Mordvinov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 69 – 73.

The article presents a comprehensive overview of zero-knowledge proof (ZKP) protocols as a fundamental concept of modern cryptography. The historical background of their emergence and the main properties ensuring reliability and confidentiality, i.e., completeness, soundness, and zero-knowledge — are considered. A classification of protocols into interactive and non-interactive ones is provided, with a special focus on modern solutions such as the zk-SNARK and the zk-STARK. The mathematical foundations of ZKPs are described in detail, including discrete logarithm proofs, the use of homomorphic encryption, polynomial commitments, hashing, and elliptic curves. Practical application areas are analyzed, including cryptocurrencies (Zcash, Ethereum), authentication systems, digital identity, and electronic voting. The advantages of using ZKPs are shown, such as enhanced privacy, reduced need for trusted intermediaries, and strengthened security. At the same time, key challenges are outlined, including scalability, implementation complexity, the problem of trusted setup, and potential vulnerability to quantum computing. It is concluded that zero-knowledge proof protocols are a powerful tool for ensuring confidentiality and reliability of digital systems, while further research is aimed at creating more efficient and quantum-resistant solutions.

Key words: zero-knowledge; cryptography; blockchain; authentication; digital privacy; zk-SNARK; zk-STARK; discrete logarithm; homomorphic encryption.

Ref: 9 items.

УДК 004.056.55

Протоколи з нульовим розголошенням: теоретичні основи та застосування в сучасній криптографії / *Р.І. Мордвінов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 69 – 73.

Представлено комплексний огляд протоколів з нульовим розголошенням (Zero-Knowledge Proofs, ZKP) як фундаментальної концепції сучасної криптографії. Розглянуто історичні передумови їх виникнення та основні властивості, що забезпечують достовірність і конфіденційність: повноту, надійність та нульове розголошення. Проведено класифікацію протоколів на інтерактивні та неінтерактивні, з особливим акцентом на сучасні рішення – zk-SNARK та zk-STARK. Детально описано математичні основи ZKP, зокрема доведення знання дискретного логарифма, використання гомоморфного шифрування, поліноміальних зобов'язань, гешування та еліптичних кривих. Проаналізовано практичні напрями застосування, включаючи криптовалюти (Zcash, Ethereum), системи автентифікації, цифрову ідентифікацію та електронне голосування. Показано переваги використання ZKP, серед яких – підвищення приватності, зниження потреби у довірених посередниках та посилення безпеки. Водночас окреслено ключові виклики, пов'язані з масштабованістю, складністю реалізації, проблемою довіреної установки та потенційною вразливістю до квантових обчислень. Зроблено висновок, що протоколи з нульовим розголошенням становлять потужний інструмент для забезпечення конфіденційності й

надійності цифрових систем, а подальші дослідження спрямовані на створення більш ефективних і квантово-стійких рішень.

Ключові слова: нульове розголошення; криптографія; блокчейн; автентифікація; цифрова приватність; zk-SNARK; zk-STARK; дискретний логарифм; гомоморфне шифрування.

Бібліогр.: 9 назв.

UDC 004.056

Development of a typical infrastructure for a quantum random number generator web service / *D.M. Morhul, O.P. Nariezhnii, T.O. Hrinenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 74 – 80.

Quantum Random Number Generators (QRNGs) provide physically unpredictable entropy essential for cryptography, modeling, and scientific research. Local generation is preferred for the highest level of security, as it eliminates the risks associated with transmitting data over a network. At the same time, public QRNG web services accessible via the API serve as a valuable tool in cases where dedicated hardware is unavailable, enabling rapid integration into software prototypes, statistical testing, large-scale simulations, and educational projects. Combining local and remote sources makes it possible to optimize the balance between security, accessibility, and performance.

This article helps to develop and justify a typical QRNG web service infrastructure, which includes functional components, security requirements, access interfaces (API), methods for quality control of randomness, and recommendations for scalability. The proposed infrastructure is intended to serve as a foundation for creating interoperable, secure, and efficient web services of quantum entropy sources.

Key words: quantum entropy source; cybersecurity; extractor; deterministic random bit generator; min-entropy; quantum random number generator; web-service QRNG.

1 fig. Ref: 28 items.

УДК 004.056

Розробка типової інфраструктури для веб-сервісу квантового генератора випадкових чисел / *Д.М. Моргуль, О.П. Нарєжній, Т.О. Гріненко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 74 – 80.

Квантові генератори випадкових чисел (QRNG) забезпечують фізично непередбачувану ентропію, необхідну для криптографії, моделювання та наукових досліджень. Для максимального рівня безпеки перевага надається локальній генерації, що виключає ризики передавання даних через мережу. Водночас публічні веб-сервіси QRNG, доступні через API, є цінним інструментом, у випадку відсутності власного обладнання, для швидкої інтеграції у прототипи програмного забезпечення, проведення статистичних тестів, масштабних симуляцій та навчальних проєктів. Поєднання локальних і віддалених джерел дозволяє оптимізувати баланс між безпекою, доступністю та продуктивністю.

Розроблено та обґрунтовано типову інфраструктуру веб-сервісу QRNG, яка включає функціональні компоненти, вимоги до безпеки, інтерфейси доступу (API), методи контролю якості випадковості та рекомендації щодо масштабування. Запропонована інфраструктура має слугувати основою для створення сумісних, безпечних і продуктивних веб-сервісів джерел квантової ентропії.

Ключові слова: джерело квантової ентропії; кібербезпека; екстрактор; deterministic random bit generator; min-entropy; quantum random number generator; web-service QRNG.

Іл. 1. Бібліогр.: 28 назв.

UDC 004.056.5:005.8

A process model for dynamic analysis and prediction of information security risks for personnel / *T.I. Korobeinikova, A.B. Yamnych* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 81 – 88.

The article addresses the problem of dynamic assessment and forecasting of information security risks driven by the growing role of the human factor amid business-process digitalization, hybrid work models, and a changing access context. The purpose of the study is to improve the accuracy and timeliness of risk management for personnel, enhance access controllability through adaptive policies, and advance audit transparency by integrating an RBAC-blockchain into a closed self-learning loop. The object of the study is the process of dynamic information-security risk analysis in corporate systems with personalized consideration of user behavior. The subject of the study comprises the methods and procedures for constructing a multidimensional feature vector, building a user digital twin, designing risk-adaptive access policies, and implementing system self-learning mechanisms. The authors emphasize that static access-control approaches and periodic audits do not match the dynamics of threats and the contextual nature of resource use. The article analyzes the contemporary components of the process model: (f_1) construction of a multidimensional resource-classification matrix; (f_2 –f_3) collection, unification, and normalization of behavioral and technical data into the feature vector Q; (f_4) forecasting risky events using a user digital twin with access transactions recorded on an RBAC-blockchain; (f_5 –f_6) generation of adaptive countermeasures and delivery of personalized policies and training content; and (f_7 –f_8) Fback feedback collection and self-learning with adjustment of weights, models, and access rules. It is shown that combining statistical methods, machine-learning algorithms, and immutable blockchain logging ensures reproducible auditing, reduces the “risk window,” and supports continuous trust validation in line with Zero Trust princi-

ples. A scheme for triggering countermeasures is proposed based on the probability matrix R and the resource's criticality class. Procedures for fine-tuning and transfer learning are described to keep models current without excessive computational costs. Particular attention is paid to personalized dashboards and multichannel delivery of recommendations that shorten user response time. The importance of qualitative Fback metrics (e.g., user satisfaction and content clarity) is emphasized for revealing elements of security culture. Thus, applying the developed process model establishes an "analysis–forecast–action–feedback–self-correction" cycle that improves the accuracy of risk assessment, enhances response timeliness, and advances transparency in access governance. The results can be integrated into SIEM/UEBA environments, access-management systems, and corporate programs for improving personnel cyber literacy.

Key words: dynamic risk assessment (DRA); digital twin; RBAC-blockchain; UEBA; Zero Trust; feature vector Q; resource-classification matrix; adaptive access policies; insider threats.

1 fig. Ref: 21 items.

УДК 004.056.5:005.8

Процесна модель динамічного аналізу та прогнозування ризиків інформаційної безпеки для персоналу / *Т.І. Коробейнікова, А.Б. Ямнич* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 81 – 88.

Висвітлено проблематику динамічного оцінювання та прогнозування ризиків інформаційної безпеки, зумовлену зростанням ролі людського фактора в умовах цифровізації бізнес-процесів, гібридних моделей роботи та змінного контексту доступу. Метою дослідження є підвищення точності та оперативності ризик-менеджменту для персоналу, покращення керованості доступу через адаптивні політики та вдосконалення прозорості аудиту за рахунок інтеграції RBAC-блокчейну в замкнений цикл самонавчання. Об'єкт дослідження – процес динамічного аналізу ризиків ІБ у корпоративних системах з персоніфікованим урахуванням поведінки користувачів. Предмет дослідження – методи та процедури формування багатовимірного вектора ознак, побудови цифрового двійника користувача, ризик-адаптивних політик доступу та механізмів самонавчання системи. Автори наголошують, що статичні підходи до контролю доступу і періодичні аудити не відповідають динаміці загроз і контексту використання ресурсів. Проаналізовано сучасні компоненти процесної моделі: (f_1) формування багатовимірної матриці класифікації ресурсів; (f_2 –f_3) збір, уніфікацію й нормалізацію поведінкових та технічних даних у вектор ознак Q; (f_4) прогнозування ризикових подій за допомогою цифрового двійника користувача з фіксацією транзакцій доступу в RBAC-блокчейні; (f_5 –f_6) генерацію адаптивних контрзаходів і доставку персоналізованих політик та навчального контенту; (f_7 –f_8) збирання зворотного зв'язку Fback і самонавчання з корекцією ваг, моделей і правил доступу. Доведено, що поєднання статистичних методів, алгоритмів машинного навчання та незмінного журналювання в блокчейні забезпечує відтворюваність аудиту, зменшує «вікно ризику» та підтримує безперервну валідацію довіри відповідно до принципів Zero Trust. Запропоновано схему активування контрзаходів на основі матриці ймовірностей R та класу критичності ресурсу. Описано процедури тонкого налаштування (fine-tuning) і transfer learning для підтримки актуальності моделей без надмірних обчислювальних витрат. Особливу увагу приділено персоналізованим інформаційним панелям і мультиканальній доставці рекомендацій, що скорочують час реакції користувача. Підкреслено важливість якісних метрик Fback (задоволеність, зрозумілість контенту) для виявлення елементів культури безпеки. Таким чином, застосування розробленої процесної моделі забезпечує цикл «аналіз-прогноз-дія-зворотний зв'язок-самокорекція», який підвищує точність оцінювання ризиків, покращує оперативність реагування та вдосконалює прозорість управління доступом. Отримані результати можуть бути інтегровані в SIEM/UEBA-середовища, системи управління доступом і програми підвищення кіберграмотності персоналу у корпоративних мережах.

Ключові слова: динамічне оцінювання ризику (DRA); цифровий двійник; RBAC-блокчейн; UEBA; Zero Trust; вектор ознак; матриця класифікації ресурсів; адаптивні політики доступу; інсайдерські загрози.

Іл. 1. Бібліогр.: 21 назв.

UDC 004.738.5:519.816:004.056.5

Ukrainian internet service providers ranking: multi‑criteria model incorporating cybersecurity / *L.I. Melnikova, A.V. Marchuk, S.V. Shtangei* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 89 – 97.

Internet services have become an integral part of modern life, requiring high standards of quality and security. Choosing an internet provider is a complex task, as it involves many different criteria that often conflict with one another. High demands for speed, stability, and security, along with cost limitations, necessitate an approach that allows for balancing various requirements within a unified system. Despite the growing global attention to cybersecurity, this criterion often remains outside the scope of comprehensive analysis in Ukrainian research. Security is undoubtedly the most important factor when it comes to choosing an internet service provider.

The study presents an analysis of the functioning of internet service providers. It examines organizational principles, identifies key operational characteristics of internet providers, and outlines criteria for evaluating and selecting them. A multi-criteria analysis was conducted using a heuristic procedure, and a mechanism was proposed to ensure the required quality of service for each user when choosing an internet provider.

Key words: internet service providers; multi-criteria optimization; heuristic procedure; operational parameters; cybersecurity; response speed to attack.

7 tab. 3 fig. Ref: 15 items.

УДК 004.738.5:519.816:004.056.5

Оцінка українських інтернет-провайдерів: багатокритеріальна модель з урахуванням кібербезпеки / *Л.І. Мельнікова, А.В. Марчук, С.В. Штангей* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 89 – 97.

Інтернет-послуги стали невід'ємною частиною сучасного життя, що потребує високих стандартів якості та безпеки. Вибір інтернет-провайдера є складною задачею, оскільки включає багато різних критеріїв, які часто мають суперечливий характер. Високі вимоги до швидкості, стабільності та безпеки, а також обмеження щодо вартості вимагають застосування підходу, що дозволяє збалансувати різні вимоги в єдиній системі. Попри зростання уваги до кібербезпеки у світі, в українських дослідженнях цей критерій часто залишається поза межами комплексного аналізу. Безпека – безперечно, найважливіший фактор, коли йдеться про вибір інтернет-провайдера.

Виконано аналіз функціонування інтернет-провайдерів. Розглянуто принципи організації, визначено основні характеристики функціонування інтернет-провайдерів та наведено критерії з урахуванням безпеки, за якими можна оцінювати та обирати інтернет-провайдерів.

Проведено багатокритеріальний аналіз за допомогою евристичої процедури та запропоновано механізм забезпечення необхідної якості обслуговування для кожного користувача при виборі інтернет-провайдера.

Ключові слова: інтернет-провайдери; багатокритеріальна оптимізація; евристична процедура; параметри функціонування; кібербезпека; швидкість реагування на атаку.

Табл. 7. Іл. 3. Бібліогр.: 15 назв.

UDC 621.391:519.2

Evolution of Man-in-the-Middle attacks in 5G telecommunication systems / *Y.V. Kotukh, G.Z. Khalimov, I.Y.Dzhura* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 98 – 107

The rapid deployment of fifth-generation (5G) networks has dramatically transformed telecommunications by enabling ultra-low latency, high bandwidth, and dynamic spectrum allocation. However, these innovations have also expanded the attack surface, introducing unprecedented security vulnerabilities. Among them, Man-in-the-Middle (MITM) attacks have evolved into complex, AI-driven, and persistent threats capable of exploiting 5G's Service-Based Architecture (SBA), virtualized functions, and heterogeneous interoperability with LTE and Wi-Fi. A particularly critical evolution is the rise of Digital Twin attacks, where adversaries replicate devices or network elements with high fidelity, bypassing traditional authentication and maintaining long-term undetectable intrusions. This paper provides a comprehensive analysis of MITM evolution in 5G systems, including vulnerabilities across the OSI model layers, exploitation of NGAP, Diameter, and DSS signaling, and the persistent risks posed by distributed architectures. Special attention is given to quantum-era threats, such as store-now-decrypt-later scenarios and quantum-enhanced MITM attacks undermining 5G-AKA protocols reliant on non-quantum-resistant cryptography. The study emphasizes the need for cryptographic agility, post-quantum authentication, and continuous behavioral validation mechanisms to mitigate persistent and quantum-enhanced MITM exploits. The findings highlight urgent requirements for international standardization and proactive implementation of post-quantum secure protocols in 5G infrastructures.

Key words: 5G telecommunication systems; MITM attacks; Digital Twin; post-quantum cryptography; 5G-AKA; network slicing; quantum computing.

Fig. 3. Ref: 27 items.

УДК 004.056.55

Еволюція атак «людина посередині» у телекомунікаційних системах 5g / *Є.В. Котух, Г.З. Халімов, І.Є. Джура* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 98 – 107.

Швидке розгортання мереж п'ятого покоління (5G) докорінно трансформувало телекомунікаційну сферу, забезпечивши ультранизькі затримки, високу пропускну здатність та динамічний розподіл спектра. Водночас ці інновації суттєво розширили поверхню атак і створили безпрецедентні вразливості безпеки. Серед них особливу загрозу становлять атаки типу «людина посередині» (MITM), які еволюціонували у складні, керовані ШІ та стійкі загрози, здатні експлуатувати сервісно-орієнтовану архітектуру (SBA), віртуалізовані функції та гетерогенну взаємодію 5G з LTE і Wi-Fi. Особливо небезпечним вектором розвитку є атаки «цифрових двійників», коли зловмисник створює точні копії пристроїв або елементів мережі, обходячи класичні механізми автентифікації та підтримуючи довготривалу приховану присутність.

У статті здійснено комплексний аналіз еволюції MITM-атак у системах 5G, включаючи вразливості на різних рівнях моделі OSI, експлуатацію протоколів NGAP, Diameter та DSS, а також ризики децентралізованих архітектур. Окрему увагу приділено загрозам квантової ери, таким як атаки типу «збережи зараз – розшифруй пізніше» та квантово-посилені MITM-атаки, що підривають протокол 5G-AKA, побудований на неквантово-стійкій криптографії. Наголошено на необхідності криптографічної гнучкості, впровадження постквантової автентифікації та механізмів безперервної поведінкової валідації для протидії стійким і квантово-посиленим MITM-загрозам. Результати вказують на нагальну потребу міжнародної стандартизації та проактивного впровадження постквантово захищених протоколів у 5G-інфраструктурах.

Ключові слова: телекомунікаційні системи 5G; атаки MITM; цифровий двійник; постквантова криптографія; 5G-AKA; мережеве нарізання; квантові обчислення.

Рис. 3. Бібліогр.: 27 назв.

RADIO ELECTRONIC SYSTEMS
РАДІОЕЛЕКТРОННІ СИСТЕМИ

UDC 550.388, 621.396.967, 520.8.05

Features and development prospects of the radio receiving system of the incoherent scatter radars of the Institute of Ionosphere, National Technical University "Kharkiv Polytechnic Institute" / *L.Ya. Emelyanov, O.V. Bogomaz, Yu.I. Podyachiy, A.E. Miroshnikov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 108 – 119.

The paper presents the current state of a multichannel radio receiving system developed, implemented, and used for many years as part of the incoherent scatter (IS) radars of the Institute of Ionosphere and is constantly being improved. The technical features and structure of this system, which provides high-precision measurements of ionospheric parameters over a wide altitude range, are described. The requirements for sensitivity, stability, and interference immunity considered of the system under conditions of weak IS signal reception are analyzed. The superheterodyne architecture with triple frequency conversion is detailed, along with principles of signal selection and synchronous detection, and the extraction of the quadrature components of the received signal for subsequent computation of IS signal correlation functions. The implementation of Doppler measurements of the velocity of motion of ionospheric plasma and objects in geospace is based on the coherent structure of the radar complex and the coordinated synchronization of the transmitting and receiving paths.

Modes of operation using composite sounding signals are presented, as well as features of receiving and selecting ionospherically scattered elements of these signals for the study of both the lower and upper ionosphere with sufficiently high altitude and temporal resolution. The mode of simultaneous sounding in the vertical and oblique directions has been implemented to enable three-dimensional analysis of the properties and dynamics of the ionospheric plasma.

The subsystem for receiving, digitizing, and processing signals at the intermediate frequency has been integrated into the radar, allowing for improved accuracy in measuring IS signal parameters by eliminating the influence of a number of instrumental factors and employing digital bandpass filters and processing algorithms adapted to signals corresponding to specific heliogeophysical conditions and investigation altitudes.

The hardware implementation of a programmable geospace monitoring radio system based on SDR (Software Defined Radio) technology is proposed, enabling both active and passive observation modes and significantly expanding the functionality of the radiophysical equipment of the observatory of the Institute of Ionosphere for ionospheric research and space environment monitoring.

The presented technical solutions meet the requirements of modern geospace monitoring systems and are useful for ionospheric research and the observation of artificial space objects, particularly in the context of the growing importance of space weather forecasting and space debris issues.

Key words: incoherent scatter technique; radio receiving systems; radar signals; sounding modes.

4 fig. Ref: 38 items.

УДК 550.388, 621.396.967, 520.8.05

Особливості та перспективи розвитку радіоприймальної системи радарів некогерентного розсіяння НДІ Іоносфери НТУ "ХПІ" / *Л.Я. Ємельянов, О.В. Богомаз, Ю.І. Под'ячий, А.Є. Мірошніков* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 108 – 119.

Наведено сучасний стан багатоканальної радіоприймальної системи, яка була розроблена, впроваджена, використовується протягом багатьох років у складі радарів некогерентного розсіяння (НР) НДІ Іоносфери й постійно вдосконалюється. Наведено технічні особливості та структуру цієї системи, яка забезпечує високоточне вимірювання параметрів іоносфери в широкому висотному діапазоні. Проведено аналіз вимог до чутливості, стабільності та завадостійкості системи в умовах приймання слабких НР сигналів. Детально описано супергетеродинну архітектуру з потрійним перетворенням частоти, принципи селекції сигналів та синхронного детектування, виділення квадратурних складових прийнятого сигналу для подальшого обчислення кореляційних функцій НР сигналу. Реалізація доплерівських вимірювань швидкості руху іоносферної плазми та об'єктів у геокосмосі ґрунтується на когерентній структурі радарного комплексу та узгодженій синхронізації передавальних і приймальних трактів.

Наведено режими використання складених зондувальних сигналів і особливості приймання й селекції розсіяних іоносферою елементів цих сигналів для дослідження нижньої та верхньої іоносфери з достатньо високою роздільною здатністю за висотою та за часом. Реалізовано режим одночасного зондування у вертикальному та похилому напрямках з метою тривимірного аналізу властивостей та динаміки іоносферної плазми.

До складу радара інтегровано підсистему приймання, дискретизації та обробки сигналу на проміжній частоті, що дозволяє підвищити точність вимірювання параметрів НР сигналу завдяки виключенню впливу низки інструментальних факторів, використанню цифрових смугових фільтрів і алгоритмів обробки, що адаптовані до сигналів, відповідних конкретним геліогеофізичним умовам і висотам дослідження.

Запропоновано апаратурну реалізацію програмованої радіосистеми моніторингу геокосмосу на базі SDR (Software Defined Radio), яка дозволяє реалізовувати як активні, так і пасивні режими спостережень, і дає можливість значно розширити функціональність радіофізичного обладнання обсерваторії НДІ Іоносфери для дослідження іоносфери й моніторингу космічного середовища.

Представлені технічні рішення відповідають вимогам сучасних систем моніторингу геокосмосу та є корисними для дослідження іоносфери та спостереження за штучними космічними об'єктами, зокрема в контексті зростаючої актуальності прогнозування космічної погоди та проблем космічного сміття.

Ключові слова: метод некогерентного розсіяння; радіоприймальні системи; радіолокаційні сигнали; режими зондування.

Іл. 4. Бібліогр.: 38 назв.

UDC 621.396

Statistical optimization and analyses of the method of forming radar images in the time and frequency domains / *S.S. Zhyla, O.V. Odokiienko, D.I. Kovaljchuk, K.O. Shcherbyna, Y.D. Sydorov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 120 – 135.

The article presents a statistically grounded approach to the formation of scatterometric radar images based on stochastic signal processing. The developed mathematical model takes into account the spatial structure of the reflecting surface, as well as the physical and statistical characteristics of radar signals. The proposed optimal algorithm combines detection, Fourier transformation, decorrelation filtering, and estimation of surface reflectivity coefficients. It is shown that such an approach ensures high resolution and increased noise immunity of the radar system. The statistical optimization is carried out according to the maximum likelihood criterion with minimization of mean square error, using the Cramér–Rao lower bound. The analysis covers both time and frequency domains, with an emphasis on practical implementation of whitening filters and decorrelation procedures in real signal conditions. Simulation examples confirm the theoretical efficiency of the algorithm and justify its application in airborne radar systems using linear frequency modulated signals for high-precision imaging.

Key words: decorrelation; radar image; optimization; reflectivity; scatterometry; signal detection; statistical estimation; stochastic model; time domain; frequency domain.

4 fig. Ref: 6 items.

УДК 621.396

Статистична оптимізація та аналіз методу формування радіолокаційних зображень у часові та частотній областях / *С.С. Жила,О.В. Одокієнко, Д.І. Ковальчук, К.О. Щербина, Я.Д. Сидоров* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 120 – 135.

Розглянуто статистично обґрунтований підхід до формування скатерометричних радіолокаційних зображень на основі стохастичної обробки сигналів. Запропонована модель враховує просторову структуру поверхні, статистичні властивості сигналів і завад. Розроблено оптимальний алгоритм обробки, який включає детектування, перетворення Фур'є, декореляцію та оцінювання коефіцієнтів відбиття. Доведено, що запропонована методика забезпечує підвищену роздільну здатність і завадостійкість. Оптимізацію проведено за критерієм максимальної правдоподібності з мінімізацією середньоквадратичної похибки на основі межі Крамера–Рао. Виконано порівняльний аналіз роботи алгоритму у часовій і частотній областях із врахуванням особливостей реалізації вибілюючих фільтрів у реальних умовах. Наведені приклади моделювання підтверджують ефективність алгоритму та доцільність його застосування у бортових РЛС з ЛЧМ сигналами.

Ключові слова: декореляція; зображення; оптимізація; радіолокація; розсіювання; сигнал; статистика; стохастика; час; частота.

Іл. 4. Бібліогр.: 6 назв.

UDC 004.8:681.5

Research on drone recognition based on their acoustic emission using fully connected neural networks / *O.V. Zubkov, N.V. Boiko, T.S. Machonis* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 136 – 144.

The relevance of the research on recognizing the self-acoustic emissions of drones using fully connected neural networks and cepstral coefficients has been substantiated. A dataset of acoustic recordings has been created, including self-emissions of various drone models, background sounds, and seven types of sound sources with spectral characteristics similar to drone acoustic emissions. The optimal number of cepstral coefficients has been identified for further recognition by fully connected neural networks in terms of maximizing the probability of correct recognition and minimizing misclassification. The optimal neural network architecture has been determined to ensure the highest probability of correct recognition. The requirements for a microprocessor-based hardware platform for recognizing drone self-acoustic emissions have been calculated.

Key words: drone; recognition; fully connected neural network; cepstral coefficients; acoustic radiation; architecture.

1 tab. 6 fig. Ref: 20 items.

УДК 004.8:681.5

Дослідження розпізнавання дронів за їх акустичним випромінюванням з використанням повнозв'язних нейронних мереж / *О.В. Зубков, Н.В. Бойко, Т.С. Мачоніс* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 136 – 144.

Обгрунтовано актуальність дослідження розпізнавання власного акустичного випромінювання дронів з використанням повнозв'язних нейронних мереж та кепстральних коефіцієнтів. Створено датасет акустичних записів власного випромінювання різних моделей дронів, фонових звуків та семи джерел звукових сигналів близьких до акустичного випромінювання дрона за спектральними характеристиками. Виявлено оптимальні значення кількості кепстральних коефіцієнтів для подальшого розпізнавання повнозв'язними нейронними мережами з точки зору максимізації ймовірності вірного розпізнавання та мінімізації похибкового розпізнавання. Виявлено оптимальну архітектуру нейронної мережі, що забезпечує максимізацію ймовірності вірного розпізнавання. Розраховано вимоги до мікропроцесорної апаратної платформи з розпізнавання власного акустичного випромінювання дронів.

Ключові слова: дрон; розпізнавання; повнозв'язна нейронна мережа; кепстральні коефіцієнти; акустичне випромінювання; архітектура.

Табл. 1. Іл. 6. Бібліогр.: 20 назв.

UDC 621.396.96

Peculiarities of detecting small-size unmanned aerial vehicles using the radioacoustic location method / *V.M. Oleinikov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 145 – 154.

The article states that modern technical means for detecting low-observable small-sized unmanned aerial vehicles (sUAVs) in some cases do not provide adequate responsiveness and reliability. One of the promising directions for improving the efficiency of detecting low-observable sUAVs is the use of electromagnetic wave scattering on acoustic disturbances generated by these vehicles in the surrounding air environment. The article examines the characteristics of the acoustic field of sUAVs. Acoustic radiation from sUAVs has a number of specific features that must be considered when developing a radio-acoustic location system. The acoustic emission spectrum of sUAVs contains harmonics of the rotor rotation frequency and blade-passing frequency. Its spatial directivity is complex and, during maneuvering, affects both the intensity and shape of the acoustic signal spectrum. The operating frequency band of the location system should cover the range in which effective resonant scattering of electromagnetic waves is ensured. To reduce the size of highly directional antennas, the radar frequency is selected in such a way that the Bragg conditions are satisfied using higher-order diffraction. When receiving a signal scattered on inhomogeneities caused by the acoustic emission of sUAVs, the signal at the receiver input is the result of the interference of electromagnetic waves reflected from acoustic waves moving in opposite directions. The parameters of the envelope of this signal depend on the speed of sound, the frequency of the acoustic emission, and the ratio of the powers of signals formed in the sections before and after the sUAV within the antenna’s radiation pattern. The method of simulation modeling was used to study the features of higher-order diffraction of electromagnetic waves on the acoustic emission of sUAVs. It was found that the main part of the energy of the scattered signal is formed by acoustic oscillations in the immediate vicinity of the sUAV, whereas the contribution of oscillations at longer distances is negligible due to their attenuation. Maximum values of the reflection coefficient are achieved only when using antennas with a beamwidth of less than 4–6°, which provides spatial selection of the wavefront segments of the acoustic emission that enable coherent summation of the scattered signals. Deviation of the radar antenna from the direction toward the sUAV disrupts the conditions for coherent summation of the scattered signals from the wavefronts of the sUAV’s acoustic field.

Key words: radio-acoustic location; small UAVs; acoustic emission; acoustic field; atmospheric dielectric permittivity; higher-order diffraction; coherent summation; wavefronts.

10 fig. Ref: 17 items.

УДК 621.396.96

Особливості виявлення малорозмірних безпілотних літальних апаратів методом радіоакустичної локації / *В.М. Олейніков* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 145 – 154.

Визначається, що сучасні технічні засоби виявлення малопомітних малорозмірних безпілотних літальних апаратів **(**МБПЛА) у деяких випадках не забезпечують належної оперативності та достовірності. Одним із перспективних напрямів підвищення ефективності виявлення малопомітних МБПЛА є використання явища розсіювання електромагнітних хвиль на акустичних збуреннях, які створюються цими апаратами в навколишньому повітряному середовищі. Розглядаються особливості акустичного поля МБПЛА. Акустичне випромінювання МБПЛА має ряд характерних особливостей, які необхідно враховувати при розробці системи радіоакустичної локації. Спектр акустичного випромінювання МБПЛА містить гармоніки частоти обертання ротора та лопатевої частоти. Його просторова спрямованість є складною і під час маневрування МБПЛА впливає на інтенсивність і форму спектра акустичного сигналу. Робоча смуга частот локаційної системи повинна охоплювати діапазон, у якому забезпечується ефективне резонансне розсіювання електромагнітних хвиль. Для зменшення габаритних розмірів високоспрямованих антен робоча частота РЛС обирається таким чином, щоб забезпечувалося виконання умов Брегга з використанням дифракції вищих порядків. Під час прийому сигналу, розсіяного на неоднорідностях, зумовлених акустичним випромінюванням МБПЛА, сигнал на вході приймача є результатом інтерференції електромагнітних хвиль, відбитих від акустичних хвиль, що рухаються у протилежних напрямках. Параметри огинаючої цього сигналу залежать від швидкості звуку, частоти акустичного випромінювання, а також від співвідношення потужностей сигналів, сформованих ділянками до і після МБПЛА, в межах діаграми спрямованості антени. Методом імітаційного моделювання досліджено особливості дифракції

вищих порядків електромагнітних хвиль на акустичному випромінюванні МБПЛА. Встановлено, що основна частина енергії розсіяного сигналу формується на акустичних коливаннях у безпосередній близькості до МБПЛА, тоді як внесок коливань на великих відстанях є незначним через їх згасання. Максимальні значення коефіцієнта відбиття досягаються лише за умови використання антен з шириною діаграми спрямованості менш ніж 4-6°, що забезпечує просторову селекцію ділянок хвильових поверхонь акустичного випромінювання які забезпечують синфазне складання розсіяних сигналів. Відхилення антени РЛС від напрямку на МБПЛА руйнує умови, за яких забезпечується синфазне складання розсіяних сигналів від хвильових поверхонь акустичного поля МБПЛА.

Ключові слова: радіоакустична локація; малорозмірні БПЛА; акустичне випромінювання; акустичне поле; діелектрична проникність атмосфери; дифракція вищих порядків; синфазне складання; хвильові поверхні.

Іл. 10. Бібліогр.: 17 назв.

UDC 621.311.6:006.83

Wireless Power Transmission (WPT): Analysis of Standards, Commercial Technologies, and Prospects / *O.V. Vorgul, I.V. Ignatiuk, T.V. Machonis, O.D. Shuniborovв* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 155 – 160.

The article presents the results of a structured analysis of wireless power transmission (WPT) technologies, their purpose, classification by range based on the ITU approaches, and an overview of key commercial standards and solutions. The ITU standards and actual commercial implementations were used as initial data. To describe WPT systems, a classification by three ranges was chosen: short-range (contact), medium-range, and long-range. A systems approach was used as an analysis method, including a comparison of parameters (range, power, efficiency, safety) and practical applicability. The technologies were assessed based on their effectiveness, safety, standardization, and commercial prevalence. The results showed that WPT does not seek to replace cable systems, but solves specific problems where physical connection is impossible, inconvenient, or dangerous. The near field (Qi standard) has received the greatest development and distribution, while far field technologies remain mainly in the area of concepts and specialized applications. Thus, the results confirm that WPT is an actively developing area, where the choice of technology is determined by a specific application, and standardization and research are critical to ensure compatibility, efficiency and safety.

Key words: wireless power transmission; WPT; Qi; AirFuel; magnetic induction; magnetic resonance; transmission range; ITU standards.

Ref: 27 items.

УДК 621.311.6:006.83

Бездротова передача енергії (БПЕ): аналіз стандартів, комерційних технологій та перспектив / *О.В. Воргуль, І.В. Ігнатюк, Т.В. Мачоніс, О.Д. Шуніборов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 155 – 160.

Наведено результати структурованого аналізу технологій бездротової передачі енергії (БПЕ), їх цілепокладання, класифікації за дальністю дії на основі підходів МСЕ (ITU), а також огляд ключових комерційних стандартів та рішень. Як вхідні дані використані стандарти МСЕ та фактичні комерційні реалізації. Для опису систем БПЕ обрано класифікацію за трьома зонами дії: ближня (контактна), середня та дальня. Як метод аналізу застосований системний підхід, що включає порівняння параметрів (дальність, потужність, ККД, безпека) та практичної застосування. Оцінювання технологій проводилося на основі їхньої ефективності, безпеки, стандартизації та комерційної поширеності. Отримані результати показали, що БПЕ прагне замінити кабельні системи, а вирішує специфічні завдання там, де фізичне з'єднання неможливе, незручно чи небезпечно. Найбільшого розвитку та поширення набула ближня зона (стандарт Qi), тоді як технології дальньої зони залишаються переважно у сфері концептів і спеціалізованих застосувань. Таким чином, результати підтверджують, що БПЕ є областю, що активно розвивається, де вибір технології визначається конкретним застосуванням, а стандартизація та дослідження критично важливі для забезпечення сумісності, ефективності та безпеки.

Ключові слова: бездротова передача енергії; БПЕ; Qi; AirFuel; магнітна індукція; магнітний резонанс; дальність передачі; стандарти МСЕ.

Бібліогр.: 27 назв.

ELECTRONIC COMMUNICATIONS
ЕЛЕКТРОННІ КОМУНІКАЦІЇ

UDC 621.396.2

Analysis of the technology for controlling digital data transmission channels with compression in computer systems / *V.V. Dovhij, V.M. Hryha, B.S. Dzundza, I.V. Svyd, A.I. Terletsky, M.F. Pavlyuk* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 161 – 171.

When transmitting digital data, various types of errors may occur, including in the data address, and accordingly, there is a need to regulate the data transfer rate. In this case, simple synchronization technology is no longer enough, and it is necessary to use a data link control protocol that would provide data flow control, error detection and protec-

tion against the latter. For the effective use of high-speed communication lines, signal compression is used, which allows several transmission sources to use the high bandwidth of the channel.

The paper analyzes the technology for controlling digital data transmission channels with compression in computer systems. Two forms of control are considered: frequency-division multiplexing (FDM) and time-division multiplexing (TDM). It is shown that sufficient data transfer speed and bandwidth, low price and the ability to use when operating from autonomous power sources make xDSL technology relevant, especially as a backup Internet access channel.

Key words: data transmission; computer networks; computer systems; compression; frequency; bandwidth.

2 tab. 7 fig. Ref: 11 items.

УДК 621.396.2

Аналіз технології управління каналами передачі цифрових даних з ущільненням в комп'ютерних системах / *В.В. Довгий, В.М. Грига, Б.С. Дзундза, І.В. Свид, А.І. Терлецький, М.Ф. Павлюк* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 161 – 171.

При передачі цифрових даних можливе виникнення різних типів помилок, зокрема, в адресі даних. І, відповідно, існує потреба в регулюванні швидкості передачі даних. При цьому простої технологія синхронізації вже недостатньо, і потрібно застосовувати протокол управління шляхом передачі даних (date link control protocol), який би забезпечив функції управління потоком даних, виявлення помилок і захисту від останніх. Для ефективного використання високошвидкісних ліній зв'язку застосовують ущільнення сигналу, яке дозволяє декільком передавальним джерелам використати високу пропускну здатність каналу.

Проведено аналіз технології управління каналами передачі цифрових даних з ущільненням в комп'ютерних системах. Розглянуто дві форми управління: ущільнення з частотним розділенням (Frequency-Division Multiplexing – FDM) і ущільнення з часовим розділенням (Time-Division Multiplexing – TDM). Показано, що досягається достатня швидкість передачі даних і пропускна здатність, невисока ціна і можливість використання при роботі від автономних джерел живлення. Все це робить xDSL технологію актуальною, особливо в якості резервного каналу доступу в Інтернет.

Ключові слова: передача даних; комп'ютерні мережі; комп'ютерні системи; стиснення; частота; пропускна здатність.

Табл. 2. Іл. 7. Бібліогр.: 11 назв.

UDC 621.372(075)

Analysis of protocol steganography methods in software-defined networks / *D.G. Fokin, M.O. Yevdokymenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 172 – 183.

The paper presents a systematic analysis of protocol steganography methods in the context of Software-Defined Networking (SDN), which constitutes a critical part of modern network infrastructure. The SDN specific characteristics are as follows: the centralized control, programmable routing logic, and separation of control and data planes. They introduce both new threat vectors and opportunities for detecting covert communication channels.

The authors classify the steganographic techniques according to the OSI model layers, namely, from the physical to the application layer, including the session and presentation layers. The study covers both traditional approaches based on header manipulation and timing-based channels, as well as modern methods that exploit specific features of protocols such as OpenFlow, TLS, QUIC, HTTP, and DNS. Special attention is given to the inter-protocol steganography, control plane abuse, and synchronization-based covert channels that enable stealthy coordination of network nodes even in segmented environments.

A comparative analysis is conducted based on key characteristics such as bandwidth, undetectability, operational constraints, and detection vectors. The results are summarized in a comparative table, allowing a reasoned evaluation of risks across OSI layers and the identification of critical areas for steganalysis in SDN. The findings confirm the need for a multi-layered defense approach in SDN infrastructures, incorporating inter-protocol analysis, adaptive anomaly detection systems, and traffic normalization policies. The research provides a scientific foundation for the development of effective countermeasures against hidden communication in next-generation networks.

Key words: steganography; network; security; protocols; covert channels; software-defined networks.

1 tab. Ref: 28 items.

УДК 621.372(075)

Аналіз методів протокольної стеганографії в програмно-конфігурованих мережах / *Д.Г. Фокін, М.О. Євдокименко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 172 – 183.

Проведено системний аналіз методів протокольної стеганографії в контексті програмно-конфігурованих мереж (SDN), що є критичною частиною сучасної мережної інфраструктури. Особливості SDN, зокрема централізоване управління, програмованість маршрутів та відокремлення площини управління й передачі, створюють як нові вектори загроз, так і можливості для виявлення прихованих каналів зв’язку.

Здійснено класифікацію стеганографічних методів за рівнями моделі OSI – від фізичного до прикладного, включно з рівнями сеансу та представлення. Детально розглянуто як традиційні методи, засновані на маніпуляції заголовками TCP/IP та порядку пакетів, так і новітні підходи, що експлуатують специфіку протоколів OpenFlow, TLS, QUIC, HTTP, DNS тощо. Особливу увагу приділено міжпротокольній стеганографії, каналам у площині управління та синхронізаційним механізмам, які дозволяють приховано координувати дії мережних вузлів навіть у сегментованих середовищах.

Проведено порівняльний аналіз методів за критеріями пропускної здатності, рівня непомітності, операційних обмежень та потенційної виявності. Результати систематизовано у таблиці, що дозволяє обґрунтовано оцінити ризики на кожному рівні моделі OSI та визначити найбільш критичні напрямки для стегоаналізу в SDN. Отримані висновки свідчать про необхідність комплексного підходу до захисту SDN-інфраструктур, що передбачає міжпротокольний аналіз, адаптивні засоби виявлення аномалій та політики нормалізації трафіку. Представлене дослідження формує наукову базу для розробки сучасних механізмів протидії прихованій передачі інформації в мережах нового покоління.

Ключові слова: стеганографія; мережа; безпека; протоколи; приховані канали; програмно-конфігуровані мережі.

Табл. 1. Бібліогр.: 28 назв.

UDC 621.396 (075)

Data transmission latency mathematical model in an SDN-controlled 5G network / *O.I. Kadatskaya, S.A. Saburova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №222. P. 184 – 191.

Next-generation services are considered and grouped into the following categories: enhanced mobile broadband access, massive connectivity and machine-type data exchange, and ultra-reliable communication with low latency. To support these diverse service delivery scenarios, 5G networks must meet the key requirements of the Quality of Service (QoS) system.

The technologies and main configuration methods of SDN OpenFlow controllers interacting with the data processing center of the 5G network are examined. OpenFlow is the primary protocol of SDN architecture. Controllers use the OpenFlow protocol to communicate with switches (forwarding devices). OpenFlow is vendor-independent, meaning the controller can interact with any switch regardless of its manufacturer.

Quality indicators calculations and probabilistic–temporal analysis of information processing are typically performed at the data processing center of the distributed control system for controlled objects (CO), based on an interaction model between each user and the RAP, with OpenFlow protocol support via the SDN platform for 5G networks.Failures in the operation of SDN controllers are minimized through timely decision-making aimed at ensuring reliability and security indicators for the functioning of controlled consumer (CO)objects within 5G networks.

At the data processing center of the distributed control system for controlled objects based on the OpenFlow SDN platform in 5G networks, the key tasks involve calculating the specified delay metrics using a dedicated mathematical model. A model has been proposed, followed by computation, analysis, and assessment of data transmission latency during the management of 5G network elements on the SDN platform and recommendations are provided for selecting a control system configuration scenario.

Key words: controllers; configuration; network; services; protocol; index; quality; platform; modeling.

1 tab. 5 fig. Ref: 3 items.

УДК 621.396 (075)

Математична модель затримки передачі даних в SDN-керованій 5G мережі / *О.Й. Кадацька, С.О. Сабурова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 184 – 191.

Розглянуто сервіси нових поколінь, які можна згрупувати: покращений мобільний широкосмуговий доступ, масове підключення та обмін даних між машинними терміналами, наднадійний зв'язок із низькими затримками. Для реалізації таких різноманітних сценаріїв доставки послуг 5G мережі повинні відповідати основним вимогам системи QoS.

Досліджено технологію та основні методи конфігурування контролерів SDN з Open Flow при взаємодії з центром обробки даних 5G мережі. OpenFlow є основним протоколом архітектури SDN. Контролери використовують протокол OpenFlow для зв'язку з комутаторами (пристроями пересилання). OpenFlow не є специфічним протоколом постачальника, який передбачає, що контролер може спілкуватися з кожним комутатором, незалежно від постачальника.

Розрахунки показників якості, аналіз ймовірнісної-часової характеристики обробки інформації проводяться, як правило, в центрі обробки результатів даних розподіленої системі управління контрольованими об’єктами (КО) на основі моделі взаємодії між кожним користувачем і RAP з підтримкою протоколу OpenFlow на SDN платформі для 5G мережі. Збої в роботі SDN-контролера мінімізуються завдяки своєчасному прийняттю рішень, спрямованих на забезпечення показників надійності та безпеки функціонування контрольованих об’єктів (КО) споживачів в 5G мережі в центрі обробки результатів даних розподіленої системі управління контрольованими об’єктами OpenFlow SDN платформи для 5G мережі основними задачами являються проведення розрахунків вказаних показників затримки за допомогою математичної моделі.

Запропоновано математичну модель, проведено розрахунки, аналіз та оцінку затримки передачі даних в процесі управління елементами 5G мережі на SDN платформі та надано рекомендації щодо вибору сценарію конфігурації системи управління.

Ключові слова: контролери; конфігурація; мережа; послуги; протокол; показник; якість; платформа; моделювання.

Табл. 1. Іл. 5. Бібліогр.: 3 назв.

UDC 621.372(075)

Control of contact center model functional parameters to reduce the load on agents / *O.I. Kadatskaya, S.A. Saburova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №222. P. 192 – 198.

The operation of the contact center model has been reviewed, and it has been determined that one of the center's key objectives is to maintain the minimally required number of agents without call interruptions. For further analysis and the development of a mathematical model for managing a networked contact center, we have identified the key performance indicators (KPIs) for agent efficiency, as well as the core metrics for assessing overall contact center performance. From a technical perspective, a typical contact center is an integrated hardware-software system. Services are executed on servers running specialized software integrated with a CRM system**.**

The contact center is conceptualized as a dynamic system composed of multiple interacting subsystems. A mathematical model for its operational control integrates statistical data, system feedback loops, and evolving management techniques to enable resource optimization. This model considers: matrix Bi,j - a representation of controllable parameters for each managed object (e.g. operator profiles, call attributes), where each bij is a random variable with known distribution; time delay decomposition total latency Ti , capturing queueing delay, transfer time, and computational processing per cycle. Has been responsed time calculations based on formulas that integrate equipment utilization K, queue depth, and probabilistic response models Bernoulli-type service distributions.

By incorporating has been proposed mathematic model and tical model for calculating the core probabilistic-temporal performance metrics – supplemented by numerical examples – the proposed framework can extend the functionality of the call processing module CPM without complex configurations. It is obtained that load thresholds K > 0.8 activate redistributive mechanisms to balance operator workloads or route calls to backup services. By analyzing traffic intensity, queue dynamics, and call handling performance across service cycles, the model supports adaptive adjustments to operator assignments and system capacity in near real-time. It also contributes to the formation of Key Performance Indicator (KPI) frameworks without requiring complex reconfiguration, enabling flexible enhancements to call routing modules.

Key words: contact center; mathematic model; CPM; agents; request time; control; overload; response time.

3 fig. Ref: 9 items.

УДК 621.372(075)

Контроль параметрів функціонування моделі контакт-центру для зменшення навантаження на операторів / *О.Й. Кадацька, С.О. Сабурова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 192 – 198.

Проведено аналіз роботи моделі контакт-центру та встановлено, що одним із ключових завдань центру є підтримання мінімально необхідної кількості агентів без переривань викликів. Для подальшого аналізу та розробки математичної моделі управління мережним контакт-центром ми визначили ключові показники ефективності (KPI) ефективності роботи агентів, а також основні метрики для оцінки загальної ефективності контакт-центру. З технічної точки зору типовий контакт-центр є інтегрованою апаратно-програмною системою. Послуги надаються на серверах під керуванням спеціалізованого програмного забезпечення, інтегрованого із CRM-системою.

Контакт-центр концептуалізується як динамічна система, що складається з кількох підсистем, що взаємодіють. Математична модель для його операційного управління поєднує статистичні дані, контури зворотного зв'язку системи та методи управління, що розвиваються, що дозволяють оптимізувати ресурси. Ця модель враховує: матрицю $B_{i,j}$ – представлення керованих параметрів для кожного керованого об'єкта (наприклад, профілі операторів, атрибути дзвінків), де кожна b_{ij} – випадкова величина з відомим розподілом; розкладання загальної затримки за часом T_i, що враховує затримку у черзі, час передачі та обчислювальну обробку за цикл. Було виконано розрахунки часу відгуку на основі формул, що інтегрують коефіцієнт використання обладнання K, глибину черги та імовірнісні моделі відгуку за розподілом послуг типу Бернуллі.

Завдяки включенню запропонованої математичної моделі та технічної моделі для розрахунку основних імовірнісно-часових метрик продуктивності, доповнених числовими прикладами, запропонована структура може розширити функціональність модуля обробки викликів CPM без складного налаштування. Отримано, що граничні значення навантаження K > 0,8 активують механізми перерозподілу для балансування навантаження операторів або маршрутизації викликів на резервні служби. Аналізуючи інтенсивність трафіку, динаміку черги та продуктивність обробки викликів у циклах обслуговування, модель підтримує адаптивне коригування призначень операторів та пропускну спроможність системи практично в режимі реального часу. Вона також сприяє формуванню структур ключових показників ефективності (KPI) без необхідності складного налаштування, що дозволяє гнучко покращувати модулі маршрутизації викликів.

Ключові слова: контакт-центр; математична модель; CPM; агенти; час запиту; контроль; перевантаження; час відгуку.

Іл. 3. Бібліогр.: 9 назв.

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS
ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

UDC 537.874.46

Characteristics of the eigenmodes of a photonic crystal waveguide in a kagome lattice / *Y.M. Odarenko, S.O. Iuhno, Y.V. Sulima, O.S. Hnatenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 199 – 205.

A photonic crystal waveguide with a hollow channel has been theoretically investigated. Waveguide cladding consists of an array of dielectric cylinders placed at the nodes of the kagome lattice. The dispersion diagrams of the photonic crystal cladding and the waveguide have been obtained on the basis of numerical calculations by the plane wave expansion method using the MIT Photonic Bands package. Several photonic band gaps of the cladding have been obtained for the considered system parameters. Waveguide eigenmodes with field energy localization in the hollow channel have been realized in each band gap. An additional waveguide eigenmode of the photonic crystal structure has been identified with the frequency locating outside the photonic band gaps. A physical justification for the implementation of an additional mechanism for field energy localization in the hollow waveguide channel, not involving either total internal reflection or the photonic band gap, has been proposed. This mechanism is similar to the mechanism of the so-called inhibited coupling between the modes of the hollow waveguide channel and the cladding modes, theoretically and experimentally investigated in photonic crystal fiber waveguides with a kagome lattice. The spatial distributions calculations results for the electric field of the photonic crystal structure eigenmodes show the effective localization of the field energy in the hollow channel for all the found waveguide modes. The difference in the spatial distribution of the eigenmode field, corresponding to the mechanism of inhibited coupling, is the presence of weak oscillations of the field in the entire space of the waveguide cladding. This indicates the existence of eigenmodes of the cladding, for which the energy exchange with the modes of the hollow channel is weak. The noted field oscillations are absent for the eigenmodes of the waveguide, the frequencies of which are within the photonic band gap of the waveguide cladding.

Key words: photonic crystal waveguide; kagome lattice; dispersion diagram; photonic band gap; inhibited mode coupling.

6 fig. Ref: 16 items.

УДК 537.874.46

Характеристики власних режимів фотонно-кристалічного хвилеводу з решіткою кагоме / *Є.М. Одаренко, С.О. Юхно, Є.В. Суліма, О.С. Гнатенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 199 – 205.

Теоретично досліджено фотонно-кристалічний хвилевід з пустотілим каналом, оболонки якого складаються з масиву діелектричних циліндрів, розміщених у вузлах решітки кагоме. Дисперсійні діаграми фотонно-кристалічної оболонки та хвилеводу отримані на основі чисельних розрахунків методом розкладання по плоским хвилям із застосуванням пакету MIT Photonic Bands. Для розглянутих параметрів системи отримано кілька фотонних заборонених зон оболонки, в кожній з яких реалізуються хвилеводні власні режими з локалізацією енергії поля в пустотілому каналі. Виявлено додатковий хвилеводний власний режим фотонно-кристалічної структури, частота якого розташована за межами фотонних заборонених зон. Запропоновано фізичне обгрунтування реалізації додаткового механізму локалізації енергії поля в пустотілому хвилеводному каналі, який не стосується ні повного внутрішнього відбиття, ні фотонної забороненої зони. Цей механізм є подібним до механізму так званого пригніченого зв'язку між модами пустотілого хвилеводного каналу та модами оболонки, який теоретично та експериментально досліджений у фотонно-кристалічних волоконних хвилеводах з решіткою кагоме. Результати розрахунків просторових розподілів електричного поля власних режимів фотонно-кристалічної структури показують ефективну локалізацію енергії поля в пустотілому каналі для всіх знайдених хвилеводних мод. Відмінність просторового розподілу поля власного режиму, який відповідає механізму пригніченого зв'язку, полягає в наявності слабких осциляцій поля в усьому просторі оболонки хвилеводу. Це свідчить про існування власних режимів оболонки, для яких енергообмін з модами пустотілого каналу є слабким. Відзначені осциляції поля відсутні для власних хвилеводних режимів, частоти яких знаходяться в межах фотонних заборонених зон оболонки хвилеводу.

Ключові слова: фотонно-кристалічний хвилевід; решітка кагоме; дисперсійна діаграма; фотонна заборонена зона; пригнічений зв'язок мод.

Іл. 6. Бібліогр.: 16 назв.

UDC 621.373.826

Silver film and distributed Bragg reflector microcavity: multilayered laser model threshold analysis / *S.S. Herasymov, O.S. Hnatenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 206 – 213.

The paper is devoted to the theoretical investigation of threshold conditions in layered microlaser structures combining a finite-thickness silver film and a dielectric distributed Bragg reflector (DBR). The research is motivated by the growing demand for compact and efficient coherent light sources that can be integrated into modern photonic and optoelectronic systems. Microlasers based on hybrid metal–dielectric cavities attract considerable attention because they offer reduced size, low power consumption, and the potential for precise spectral control. At the same time, their operation is strongly affected by the reflectivity of cavity boundaries, the thickness of the active medium, and the presence of parasitic resonances. In this context, the present work focuses on analyzing how these structural factors determine the lasing threshold and spectral characteristics of operating modes. The study employs the Lasing Eigenvalue Problem approach, which enables a rigorous description of cavity modes at the emission threshold, and the Transfer Matrix Method, which is widely used for multilayer optical systems. The numerical results demonstrate that increasing the silver

film thickness can noticeably reduce the lasing threshold, while adjustment of the active layer thickness provides efficient wavelength tuning in the visible and near-infrared ranges. A specific example is given for Nd:YAG-based microlasers, where emission near 1064 nm can be accurately controlled by selecting appropriate cavity parameters. A significant part of the analysis is dedicated to the role of the DBR. It is shown that the DBR not only forms photonic band gaps that suppress energy leakage, but also gives rise to a series of additional resonances. These parasitic modes originate from sub-cavities within the multilayer reflector and become more numerous as the number of dielectric pairs increases. However, they are characterized by much higher threshold gain values and weak overlap with the active region, which makes them less favorable for practical lasing. The comparison with a purely metallic cavity demonstrates that the inclusion of a sufficiently wide DBR can improve the performance of the primary mode, lowering its threshold while isolating it from parasitic resonances.

The obtained results highlight the potential of combining noble-metal films with DBR structures in the design of advanced microlasers. By carefully adjusting the reflector parameters, it is possible to optimize mode selection, suppress unwanted resonances, and achieve stable, efficient operation. Such approaches are expected to be valuable for the development of integrated photonic devices, optical communication systems, LiDAR sensors, and compact biosensing platforms. The paper provides useful theoretical guidelines for further research and practical implementation of layered microlasers with tunable properties.

Key words: microlaser; lasing eigenvalue problem; distributed Bragg reflector; silver film; threshold condition.

4 fig. Ref: 26 items.

УДК 621.373.826

Мікрорезонатор зі срібної плівки та розподіленим рефлектором Бреґґа: аналіз порогових умов багатошарової лазерної моделі / *С.С. Герасимов, О.С. Гнатенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 206 – 213.

Розглянуто математичне моделювання порогових умов генерації в шаруватих мікролазерних структурах, що складаються зі срібної плівки скінченної товщини та діелектричного розподіленого рефлектора Бреґґа (РРБ). Актуальність роботи зумовлена зростаючим попитом на мініатюрні лазерні джерела для інтегрованої фотоніки, сенсорики та телекомунікацій, де особливе значення мають зменшені порогові значення, стабільність випромінювання та можливість точного налаштування довжини хвилі. Для аналізу застосовано метод задачі власних значень лазерної структури у поєднанні з методом матриць переносу, що дозволило отримати кількісні залежності між параметрами мікрорезонатора та характеристиками випромінюваних мод. Проведені числові розрахунки показали, що товщина активного шару та срібної плівки істотно впливають як на довжину хвилі випромінювання, так і на пороговий коефіцієнт підсилення. Зокрема, збільшення товщини металевої плівки призводить до помітного зменшення порогу, тоді як варіювання товщини активного шару дозволяє здійснювати точне налаштування резонансної довжини хвилі, наприклад, у діапазоні поблизу 1064 нм для кристалу Nd:YAG. Показано, що РРБ може не лише формувати широкі фотонні заборонені зони, які пригнічують витік енергії в навколишнє середовище, але й породжувати додаткові паразитні моди, кількість яких зростає зі збільшенням числа пар шарів. Ці моди відрізняються підвищеними пороговими значеннями і характеризуються слабким перекриттям з активною областю, що ускладнює їх практичну реалізацію.

Отримані результати свідчать про можливість цілеспрямованого керування спектральними властивостями та пороговими умовами мікролазера завдяки комбінуванню металевих відбивачів і РРБ. Встановлено, що оптимальний вибір кількості пар діелектричних шарів у РРБ дозволяє виділити робочу моду серед паразитних резонансів та забезпечити стабільність випромінювання. Таким чином, запропонована модель демонструє перспективність використання складених відбивачів для зниження втрат, розширення можливостей спектрального налаштування та підвищення ефективності лазерних мікроструктур. Робота робить внесок у розвиток методів проєктування новітніх мікролазерів і може бути корисною для створення високоякісних джерел випромінювання у фотоніці, біосенсориці та системах оптичних комунікацій.

Ключові слова: мікролазер; задача на власні значення для лазерих структур; розподілений бреггівський відбивач; срібна плівка; порогові умови.

Іл. 4. Бібліогр.: 26 назв.

UDC 535.4

Diffraction of light on one and two infinitly narrow slits in a screen / *A.V. Bezugliy* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 214 – 218.

The work addresses the solution of the problems of light diffraction at one and two narrow slits in an opaque screen. The diffraction patterns were analyzed. It was established that their appearance depends on the ratio of the characteristic size of the obstacle to the wavelength of the photon. When this ratio is less than unity, we have a smoothly changing illumination with a maximum at the center of symmetry of the system. If it one is greater than unity, then the pattern has the appearance of light and dark stripes. The numerical value of the specified ratio determines the number of maxima. The obtained results allow us to assume that the reason for the well-known paradox in the interpretation of experiments on the diffraction of light at one and two narrow slits is the next. The experiments at one slit was carried out for the case when the wavelength was greater than the width of the slit, and at two slits - for the case, when the wavelength is much smaller than the distance between slits.

Key words: diffraction; diffraction pattern; material point; harmonic wave; corpuscular; momentum quantum; psi- function; probability.

2 fig. Ref: 5 items.

УДК 535.4

Дифракція світла на одній та двох нескінченно вузьких щілинах в екрані / *А.В. Безуглий* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 214 – 218.

Розв'язано задачі дифракції світла на одній та двох вузьких щілинах в непрозорому екрані. Проведено аналіз дифракційних картин. Встановлено, що їхній вигляд залежить від відношення характеристичного розміру перешкоди до довжини хвилі фотона. Коли це відношення менше одиниці, маємо плавно спадаючу освітленість з максимумом в центрі симетрії системи. Якщо ж воно більше одиниці, то картина має вигляд світлих і темних смуг. Чисельне значення вказаного відношення встановлює число максимумів. Отримані результати дають можливість вважати, що причиною виникнення відомого парадоксу при тлумаченні експериментів з дифракції світла на одній та двох вузьких щілинах є наступне. Очевидно експерименти на одній щілині були проведено для випадку коли довжина хвилі була більша ширини щілини, а на двох щілинах – для випадку, коли довжина хвилі значно менша відстані між щілинами.

Ключові слова: дифракція; дифракційна картина; матеріальна точка; гармонічна хвиля; корпускулярні; квант імпульсу; псі-функія; імовірність.

Іл. 2. Бібліогр.: 5 назв.

UDC 621.382.323

MOSFET transistor modeling including parasitic leakage and drain resistance / *V.M. Hryha, V.M. Vintoniak, V.S. Hula* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 219 – 227.

The aim of the work is to develop and analyze a method for including parasitic source (R_S) and drain (R_D) resistances in the basic SPICE Level 1 MOSFET model. Relevance: even in the simple quadratic MOSFET model (SPICE level 1), ignoring parasitic resistances can lead to significant simulation errors, especially at high currents. This paper provides a literature review of approaches to account for R_S and R_D in compact MOSFET models: from the classic Shichman–Hodges model (SPICE Level 1) to modern works on the extraction of resistances and modeling their impact. Theoretical background describes a modification of the Level 1 model by introducing effective voltages V^{eff}_{GS} and V^{eff}_{DS} that account for voltage drops across R_S and R_D, and analytical equations for the drain current I_D in linear and saturation regions with these resistances. Methodology includes a numerical iterative algorithm implemented in Python/PySpice which solves the implicit equation $I_D(V^{eff}_{GS}, V^{eff}_{DS})$. Results show reduction of current and a shift in saturation point when adding parasitic resistances: for a typical NMOS at V_{GS}=5 V, introducing R_S=R_D=50 Ω reduces I_D by ≈16% and increases the saturation voltage by ≈ 0.3 V. Output characteristics graphs and tables of relative current deviation are presented. The novelty lies in the proposed simple iterative procedure to include R_S, R_D in the SPICE Level 1 model without resorting to more complex models, and the practical value is the applicability of this approach for educational modeling and quick evaluation of parasitic effects on MOSFET behavior. Conclusions: accounting for R_S, R_D significantly improves the accuracy of Level 1 modeling, bringing results closer to real devices with minimal computational complexity, which is useful for engineering practice and further research.

Key words: MOSFET; SPICE Level 1; parasitic resistances; source resistance; drain resistance; modeling; PySpice.

1 tab. 1 fig. Ref: 12 items.

УДК 621.382.323

Моделювання MOSFET-транзисторів з урахуванням паразитних опорів витоку та стоку / *В.М. Грига, В.М. Вінтоняк, В.С. Гула* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 219 – 227.

Метою роботи є розробка та аналіз методу врахування паразитних опорів витоку (R_S) і стоку (R_D) у базовій моделі MOSFET SPICE Level 1. Актуальність зумовлена тим, що у простій квадратичній SPICE моделі MOSFET ігнорування паразитних опорів може призводити до значних похибок при симуляції характеристик транзистора, особливо в режимах великих струмів. У статті проведено огляд літератури щодо підходів до врахування R_S та R_D у компактних моделях MOSFET: від класичної моделі Шічмана–Ходжеса до сучасних робіт з екстракції опорів та моделювання їх впливу. Теоретично описано модифікацію моделі рівня 1: введено ефективні напруги V^{eff}_{GS} і V^{eff}_{DS}, що враховують падіння напруги на R_S, R_D, та отримано аналітичні рівняння для струму I_D у лінійному й насиченому режимах з урахуванням цих опорів. Методика включає чисельний ітераційний алгоритм, реалізований у Python/PySpice, для розв'язання неявного рівняння $I_D(V^{eff}_{GS}, V^{eff}_{DS})$. Результати моделювання демонструють зниження струму та зсув точки насичення при додаванні паразитних опорів: зокрема, для типового NMOS при V_{GS}=5 В введення R_S=R_D=50 Ом зменшує I_D на ≈16% і збільшує напругу насичення на ≈0.3 В. Наведено графіки вихідних характеристик та таблиці відносних відхилень струму. Наукова новизна роботи полягає у запропонованій простій ітераційній процедурі врахування R_S, R_D у моделі SPICE Level 1 без потреби у складніших моделях, а практична цінність – у можливості використовувати цей підхід для навчального моделювання та швидкої оцінки впливу паразитних ефектів на характеристики MOSFET. Висновки: врахування R_S, R_D суттєво покращує точність моделювання на рівні 1, наближаючи результати до реальних при мінімальному ускладненні обчислень, що є корисним для інженерної практики і подальших досліджень.

Ключові слова: MOSFET; SPICE Level 1; паразитні опори; опір витоку; опір стоку; моделювання; PySpice.

Табл. 1. Іл. 1. Бібліогр.: 12 назв.

BIOMEDICAL ENGINEERING
БІОМЕДИЧНА ІНЖЕНЕРІЯ

UDC 638.14:614.894:535.372

Design of ultraviolet disinfection with optimization of irradiation dosage by means of measurement and control of uv radiation parameters / *A.Yu. Rudenko, V.A. Mardzyavko, L.V. Vakhonina, M.P. Kundenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 228 – 234.

Modern beekeeping faces complex threats, among which the pesticide poisoning, spread of diseases and parasites, climate change, decrease in the feed base and organizational shortcomings in the apiary management play a key role. Mass losses of bee colonies in Ukraine, reaching 30–50% per year, create a serious environmental and agricultural problem. In this context, the search for safe and effective disinfection methods that can replace chemicals is relevant.

The aim of the study is to assess the effectiveness of ultraviolet (UV-C) radiation in the range of 210–280 nm for combating fungal and viral infections of the surface of hives. The experiments conducted showed that the most effective is radiation with a wavelength of 254 nm, which provides up to 96.8% disinfection of colonies of microorganisms.

The developed experimental installation based on a bactericidal UV-irradiator with autonomous power supply demonstrated high efficiency, cost-effectiveness and environmental safety in comparison with traditional methods (steam, acid and fire treatment). The results confirm the feasibility of using the UV technologies in beekeeping as a promising direction for disease prevention and preservation of bee colonies.

Key words: beekeeping; disinfection; ultraviolet radiation; UV-C; bee diseases; fungal infections; environmentally friendly technologies.

5 tab. 4 fig. Ref: 10 items.

УДК 638.14:614.894:535.372

Проєктування знезараження ультрофіолетом з оптимізацією дозування опромінення засобами вимірювання та контролю параметрів УФ-випромінення / *А.Ю. Руденко, В.А. Мардзявко, Л.В. Вахоніна, М.П. Кунденко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 228 – 234.

Сучасне бджільництво стикається з комплексними загрозами, серед яких ключову роль відіграють пестицидне отруєння, поширення хвороб та паразитів, кліматичні зміни, зменшення кормової бази й організаційні недоліки у веденні пасік. Масові втрати бджолосімей в Україні, що сягають 30–50 % на рік, створюють серйозну екологічну та аграрну проблему. У цьому контексті актуальним є пошук безпечних і ефективних методів знезараження, здатних замінити хімічні препарати.

Метою дослідження є оцінка ефективності ультрафіолетового (УФ-С) випромінювання в діапазоні 210–280 нм для боротьби з грибковими та вірусними інфекціями поверхні вуликів. Проведені експерименти показали, що найбільш дієвим є випромінювання з довжиною хвилі 254 нм, яке забезпечує до 96,8 % знезараження колоній мікроорганізмів.

Експериментальна установка на базі бактерицидного УФ-опромінювача з автономним живленням продемонструвала високу ефективність, економічність та екологічну безпечність у порівнянні з традиційними методами (парова, кислотна та вогнева обробка). Результати підтверджують доцільність використання УФ-технологій у бджільництві як перспективного напряму профілактики захворювань і збереження бджолосімей.

Ключові слова: бджільництво, знезараження; ультрафіолетове випромінювання; УФ-С; хвороби бджіл; грибкові інфекції; екологічно безпечні технології.

Табл. 5. Іл. 4. Бібліогр.: 10 назв.

UDC 577.3:621.3.049:004.942

Research into the influence of the electromagnetic field on cell ion channels using modeling and measurement systems / *A.Yu. Rudenko, V.A. Mardzyavko, V.O. Martynenko, M.P. Kundenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 235 – 241.

The relevance of the study is due to the growing interest in the regulation of cellular activity under the influence of electromagnetic fields (EMF), which is a promising direction in biomedical engineering. Ion channels, in particular Na^+ , are crucial for maintaining the membrane potential, generating impulses and controlling cell functions. The article proposes the use of computer-aided design systems (CAD) together with control and measurement methods for modeling the influence of EMF on cell ion channels. Mathematical models of ion transport dynamics are presented, in particular the Navier–Stokes equation for describing fluid oscillations, volume transport models and the modified Kuramoto model for studying the synchronization of cellular pulsations. The simulation conducted in MATLAB/Simulink showed that EMF with a frequency of 50 Hz and an amplitude of 100 nA can cause both depolarization (stimulation) and hyperpolarization (inhibition) of the membrane potential depending on the mode of influence. The results demonstrate the possibility of targeted regulation of the permeability of sodium channels, which is confirmed by the data of other studies. The proposed model allows for a deeper understanding of the mechanisms of electrophysiological regulation of the cell and can become the basis for the creation of new methods of therapeutic influence and the development of bioelectronic devices. The results obtained have high practical significance and open up prospects for further experimental and applied research in biomedical technologies.

Key words: ion; channels; membrane; sodium; pulsation; model; cell; EMF; CAD; MATLAB.

2 tab. 5 fig. Ref: 12 items.

УДК 577.3:621.3.049:004.942

Дослідження впливу електромагнітного поля на іонні канали клітини з використанням систем моделювання та вимірювання / *А.Ю. Руденко, В.А. Мардзявко, В.О. Мартиненко, М.П. Кунденко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 235 – 241.

Актуальність дослідження обумовлена зростаючим інтересом до регуляції клітинної активності під впливом електромагнітних полів (ЕМП), що є перспективним напрямом у біомедичній інженерії. Іонні канали, зокрема Na+, мають вирішальне значення для підтримання мембранного потенціалу, генерації імпульсів та управління функціями клітини. Запропоновано використання систем автоматизованого проєктування (САПР) разом із контрольно – вимірювальними методами для моделювання впливу ЕМП на іонні канали клітин. Представлено математичні моделі динаміки іонного транспорту, зокрема рівняння Нав'є–Стокса для опису рідинних коливань, об'ємно-транспортні моделі та модифіковану модель Курамото для дослідження синхронізації клітинних пульсацій. Проведене моделювання в MATLAB/Simulink показало, що ЕМП частотою 50 Гц і амплітудою 100 нА може викликати як деполяризацію (стимуляцію), так і гіперполяризацію (пригнічення) мембранного потенціалу залежно від режиму впливу. Результати демонструють можливість цілеспрямованого регулювання проникності натрієвих каналів, що підтверджено даними інших досліджень. Запропонована модель дозволяє глибше зрозуміти механізми електрофізіологічної регуляції клітини та може стати основою для створення нових методів терапевтичного впливу й розробки біоелектронних пристроїв. Отримані результати мають високу практичну значущість та відкривають перспективи для подальших експериментальних і прикладних досліджень у біомедичних технологіях.

Ключові слова: іонні; канали; мембрана; натрій; пульсація; модель; клітина; ЕМП; САПР; MATLAB.
Табл. 2. Іл. 5. Бібліогр.: 12 назв.

UDC 004.94:616.8-073.75

Classification of electromyographic signals by their entropic characteristics for differential diagnostics of low back pain using the random forest method / *T.V. Zhemchuzhkina* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2025. №222. P. 242 – 250.

The article presents the results of an investigation into the possibilities of using entropic characteristics of electromyographic (EMG) signals for differential diagnostics of pain syndromes in the lumbar spine. Two independent sets of EMG signals recorded in three diagnostic groups were used as initial data: healthy individuals without complaints of back pain, conditionally healthy individuals with complaints of pain (dysfunctional pain), and patients with degenerative diseases of the spine (functional pain). Four entropy indicators were selected to describe the signals – median and mean entropy, as well as median and mean spectral entropy. The random forest algorithm was used as a classification method. Model training was carried out on a data set with a significant class imbalance, and testing was performed using another, independent set. During the study, the number of trees in the ensemble was varied (100, 200, 300 and 500), and the feasibility of taking into account weighting coefficients inversely proportional to the representation of classes in the data was also tested. The quality of the models was assessed based on the classification accuracy, F1-score, area under the ROC curve (AUC), ROC curve and confusion matrix. The results obtained showed that increasing the number of trees above 100 does not provide an increase in the quality of classification, and weighting the training data in most cases does not improve the results compared to unweighted models. The best indicators were achieved when distinguishing the group of patients with dysfunctional pain from the group with functional pain: the F1-score was 0.99, the AUC was 1.00, and the accuracy was 99.09%. Thus, the results confirm the feasibility of using the entropic characteristics of EMG signals in combination with the random forest method to create reliable clinical decision support tools. The greatest diagnostic value is the ability to distinguish the type of pain syndrome (dysfunctional or functional), which can contribute to a more informed choice of treatment tactics in patients with low back pain.

Key words: classification; diagnostics; electromyography; entropy; pain; lower back; random forest; signal.
1 tab. 2 fig. Ref: 17 items.

УДК 004.94:616.8-073.75

Класифікація електроміографічних сигналів за їх ентропійними характеристиками для диференціальної діагностики болю у попереку методом випадкового лісу / *Т.В. Жемчужкіна* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2025. Вип. 222. С. 242 – 250.

Представлено результати дослідження можливостей використання ентропійних характеристик електроміографічних (ЕМГ) сигналів для диференціальної діагностики больових синдромів у поперековому відділі хребта. В якості вихідних даних використовувались два незалежні набори ЕМГ-сигналів, зареєстрованих у трьох діагностичних групах: здорові особи без скарг на біль у спині, умовно здорові зі скаргами на біль (дисфункціональний біль) та пацієнти з дегенеративними захворюваннями хребта (функціональний біль). Для опису сигналів було обрано чотири ентропійні показники – медіанна та середня ентропія, а також медіанна та середня спектральна ентропія. В якості методу класифікації застосовано алгоритм випадкового лісу. Навчання моделей здійснювалося на наборі даних із суттєвим дисбалансом класів, а тестування – на іншому, незалежному наборі. У процесі дослідження варіювалася кількість дерев в ансамблі (100, 200, 300 та 500), а також перевірялася доцільність урахування вагових коефіцієнтів, обернено пропорційних представленості класів у даних. Оцінювання якості моделей проводилося на основі точності класифікації, F1-міри, площі під ROC-кривою (AUC), ROC-кривої та матриці помилок. Отримані результати засвідчили, що збільшення кількості дерев понад 100 не забезпечує зростання якості класифікації, а зважування навчальних даних у більшості випадків не покращує результати порівняно з незваженими моделями. Найкращі показники досягнуті при розмежуванні групи пацієнтів із дисфункціональним болем та групи з функціональним болем: F1-міра становила 0,99, AUC дорівнювала 1,00, а точність – 99,09 %. Таким чином, результати підтверджують доцільність використання ентропійних характеристик ЕМГ-сигналів у поєднанні з методом випадкового лісу для створення надійних інструментів підтримки клінічних рішень. Найбільшу діагностичну цінність має можливість розрізнення типу больового синдрому (дисфункціональний чи функціональний), що може сприяти більш обґрунтованому вибору лікувальної тактики у пацієнтів із болем у попереку.

Ключові слова: біль; випадковий ліс; діагностика; електроміографія; ентропія; класифікація; сигнал.
Табл. 1. Іл. 2. Бібліогр.: 17 назв.

COLLECTION OF SCIENTIFIC PAPERS
RADIOTEKHNIKA
Issue 222
In English and Ukrainian

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 222
Англійською та українською мовами

Коректор Л.І. Сащенко

.

Підп. до друку 25.09.2025. Формат 60x90/8. Папір офсет. Гарнітура Таймс. Друк. ризограф.
Ум. друк. арк**.** 13,3. Обл.-вид. арк. 11,8 Тираж 300 прим. Зам. № 189. Ціна договір.

Харківський національний університет радіоелектроніки (ХНУРЕ)
Просп. Науки, 14, Харків, 61166.

Оригінал-макет підготовлено і збірник надруковано у ПФ „Колегіум”,
Свідоцтво про внесення суб’єкта видавничої діяльності до Державного реєстру видавців.
Сер. ДК №1722 від 23.03.2004.