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 1 0

Kharkiv
Kharkiv National
University of Radio Electronics
2022

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: 171 – Electronics; 172 – Telecommunications and Radio Engineering; 173 – Avionics; 125 – Cybersecurity; 151 – Automation and Computer-Integrated Technologies; 152 – Metrology and Information-Measuring Equipment; 153 – Micro- and Nanosystem Technology; 163 – Biomedical Engineering; 105 – Applied Physics and Nanomaterials.

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: *I.V. Svyd, 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 № 8 dated 28.09.2022.

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

Journal "Radiotekhnika" is included in the Catalog of subscription editions of Ukraine, subscription index **08391**.

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 1 0

Харків
Харківський національний
університет радіоелектроніки
2022

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: 171 – Електроніка; 172 – Телекомунікації та радіотехніка; 173 – Авіоніка; 125 – Кібербезпека; 151 – Автоматизація та комп'ютерно-інтегровані технології; 152 – Метрологія та інформаційно-вимірювальна техніка; 153 – Мікро- та наносистемна техніка; 163 – Біомедична інженерія; 105 – Прикладна фізика та наноматеріали.

Сайт: rt.nure.ua

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

За зміст статті відповідальні автори.

Відповідальні за випуск: *І.В. Свид, 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 № 8 від 28.09.2022.

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

Збірник «Радіотехніка» включено до Каталогу передплатних видань України, передплатний індекс **08391**.

CONTENT

METHODS, ALGORITHMS AND TOOLS FOR CRYPTOGRAPHIC PROTECTION OF INFORMATION

RADIOLOCATION AND RADIONAVIGATION

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS

INFORMATION METHODS OF RADIO ENGINEERING

MEANS OF TELECOMMUNICATIONS

ЗМІСТ

МЕТОДИ, АЛГОРИТМИ ТА ЗАСОБИ КРИПТОГРАФІЧНОГО ЗАХИСТУ ІНФОРМАЦІЇ

РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

ІНФОРМАЦІЙНІ МЕТОДИ РАДІОТЕХНІКИ

ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ

УДК 004.056.5 **DOI:10.30837/rt.2022.3.210.01**

І.Д. ГОРБЕНКО, д-р техн. наук, О.Г. КАЧКО, канд. техн. наук,
М.В. ЄСІНА, канд. техн. наук, В.А. ПОНОМАР, канд. техн. наук

ПОРІВНЯЛЬНА ХАРАКТЕРИСТИКА АЛГОРИТМІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ CRYSTALS-KYBER ТА СКЕЛЯ (ДСТУ 8961-2019)

Вступ

Алгоритми інкапсуляції ключів (Key-establishment Algorithms, KEA) формують загальний секрет – ключ для симетричного алгоритму шифрування.

В роботі розглянуто два KEA алгоритмів, які застосовують алгебраїчні решітки: один з фіналістів 3-го раунду Crystals-Kyber [1] (далі Kyber) та алгоритм Скеля (ДСТУ 8961-2019) [2] (далі Скеля).

Алгоритм Kyber спочатку виконує несиметричне шифрування повідомлення завдовжки 32 байти, а потім виконується формування загального секрету.

Алгоритм Скеля виконує ті ж дії, але для несиметричного шифрування застосовує повідомлення будь-якої довжини, яка не перевищує максимально можливої. Ось чому останній алгоритм можна застосовувати не тільки як KEA алгоритм, а і як алгоритм несиметричного шифрування.

Згідно з NIST Security level алгоритм Kyber забезпечує криптографічну стійкість 1, 3 та 5 рівнів, а алгоритм Скеля забезпечує криптографічну стійкість 3, 5 та 7 рівнів. Криптографічна стійкість, яка забезпечується, для обох алгоритмів визначається набором параметрів.

Далі розглянуто деталі реалізації кожного з алгоритмів, виконано порівняння алгоритмів генерації ключів, інкапсуляції та декапсуляції для алгоритмів Kyber та Скеля з боку довжин ключових даних, і результату інкапсуляції та обчислювальної складності обох алгоритмів.

Алгоритми застосовують випадкові послідовності. Позначимо функцію для генерації випадкової послідовності gen_rand, яка приймає в якості вхідних даних довжину в байтах і повертає випадковий рядок.

1. Алгоритм Kyber

1.1. Параметри

Набір параметрів залежить від криптографічної стійкості. Його наведено в табл. 1.

Для кожного набору параметрів є два типи алгоритмів:

- алгоритм за замовчуванням;
- алгоритм, позначений як KYBER_90S.

У залежності від типу алгоритму застосовують різні алгоритми для гешування та генерації псевдовипадкових послідовностей (потокове шифрування).

Таблиця 1

Параметри для алгоритму Kyber

Позначення	Призначення	Значення
λ	Рівень криптостійкості	1, 3, 5
n	Степінь полінома x^n+1	256
q	Модуль	3329
k	Розмір вектора, компонентами якого є поліном	2 *для* $\lambda = 1$ 3 *для* $\lambda = 3$ 4 *для* $\lambda = 5$
η_1	Визначає граничні значення для коефіцієнтів поліномів s, e	3 *для* $\lambda = 1$ 2 *для* $\lambda = 3$ 2 *для* $\lambda = 5$
η_2	Визначає граничні значення для коефіцієнтів поліномів e_1, e_2	2 для $\lambda = 1$ 2 для $\lambda = 3$ 2 для $\lambda = 5$
(du, dv)	du – кількість бітів для коефіцієнту полінома при перетворенні вектора поліномів в рядок байтів dv – кількість бітів для елементу при перетворенні полінома в рядок байтів	10,4 для $\lambda = 1$ $10,4$ для $\lambda = 3$ $11,5$ для $\lambda = 5$
m_len	Довжина повідомлення для шифрування	32

Різниця між цими алгоритмами наведена у табл. 2.

Таблиця 2

Алгоритм Kyber – різниця між алгоритмом
за замовченням та алгоритмом KYBER_90S

Тип алгоритму	Гешування	Потокове шифрування
За замовченням	H – SHA3-256 G – SHA3-512	XOF – SHAKE-128 KDF – SHAKE-256 PRF – SHAKE-256
KYBER_90S	H – SHA-256 G – SHA-512	XOF – AES-256 в режимі CTR KDF – SHAKE-256 PRF – AES-256 в режимі CTR

Довжини гешів: H_len=32, G_len=64 байтів.

Як показали експериментальні дослідження для обох типів алгоритмів, обчислювальна складність алгоритму за замовчуванням менше, тому в подальшому передбачається застосування саме цього режиму.

1.2. Алгоритми Parse

Застосовуються при генерації поліномів з невід'ємними коефіцієнтами в інтервалі [0, q–1] (Parse^q) та при генерації малих поліномів з цілими коефіцієнтами, діапазон для яких задається параметрами η_1, η_2. (Parse$^\eta$). У тестових прикладах автори застосовують LITTLE ENDIAN представлення даних в пам'яті. Нумерація бітів в байті починається справа, тобто біт 0 – самий правий біт байту.

1.2.1. Алгоритм Parse^q для матриці

Алгоритм застосовує квадратну матрицю розміром k рядків. Елементом матриці є поліном степені n, коефіцієнти полінома – невід'ємні значення, задані в діапазоні [0, q–1]. Для інкапсуляції та декапсуляції застосовують транспоновану матрицю.

Для генерації псевдовипадкової послідовності застосовують XOF алгоритм. Ініціалізація виконується для кожного елементу матриці, тобто для кожного полінома. Для ініціалізації генератора застосовують *seed*, номер рядка і номер колонки матриці. У разі відновлення транспонованої матриці номер рядка і номер колонки міняються місцями. Довжина псевдовипадкового рядка buf_len$\approx\lceil log_2 q * n/(bytelen * p)\rceil$, де $bytelen$ – кількість бітів в байті

(зазвичай, 8), p – імовірність, що число завдовжки $log_2 q$ біт буде менше, ніж q. Для Kyber q=3329, n=256, $log_2 q$=12, $bytelen$=8, $p=q/2^{12}$=0.8127, buf_len≈473 (байтів). В тестових прикладах послідовність генерується блоками, розмір блоку для XOF алгоритму дорівнює 64 байти, тому buf_len=512 байтів.

Для кожного з n коефіцієнтів полінома в рядку псевдовипадкових байтів функція виділяє наступні $log_2 q$ бітів, перевіряє, що відповідне значення менше, ніж q та приймає це значення в якості наступного коефіцієнту полінома матриці. Якщо значення не менше, ніж q, тоді таке значення ігнорується.

Якщо вичерпано усю псевдовипадкову послідовність, генерується додаткова послідовність завдовжки один блок.

П р и к л а д . Хай буфер з псевдовипадковими даними починається з байтів:

0x4d 0x50 0xe8.

В пам'яті ці дані розташовані наступним чином:

Номер байту 2 1 0.

Значення байту 0xe8 0x50 0x4d.

Молодшим 12 бітам відповідає значення 0x04d=77<q – приймається, перший коефіцієнт полінома дорівнює 77.

Наступним 12 бітам відповідає значення 0xe85=7434≥q – не приймається.

1.2.2. Алгоритми Parse$^{\eta}$ для малих поліномів

Для генерації псевдовипадкових даних застосовують алгоритм PRF, для ініціалізації – відповідний *seed* та значення *nonce*, яке збільшується від 0 на 1 при кожному наступному застосуванні. Довжина псевдовипадкового рядка buf_len=2*η*n/$bytelen$. Для Kyber n=256, η=2 або 3, $bytelen$=8, $buf\_len = \begin{cases} 128 \text{ для } \eta = 2 \\ 192 \text{ для } \eta = 3 \end{cases}$.

Отриманий рядок байтів ділиться на бітові послідовності завдовжки η біт та обчислюється сума бітів кожної послідовності. Отримані суми розглядаються парами. Позначимо перший елемент пари, який відповідає молодшим бітам як a, другий елемент пари – b, тоді результат дорівнює $a–b$.

П р и к л а д 1 (η = 2) . Хай перші 4 байти байтової послідовності дорівнюють 0xE0 0x33 0x56 0x15

0xE0	0x33	0x56	0x15
1110 0000	0011 0011	0101 0110	0001 0101
Порції	Порції	Порції	Порції
11 10 00 00	00 11 00 11	01 01 01 10	00 01 01 01
Суми	Суми	Суми	Суми
a0 = 0 + 0 = 0;	a2 = 1 + 1 = 2;	a4 = 0 + 1 = 1	a6 = 1 + 0 = 1
b0 = 0 + 0 = 0;	b2 = 0 + 0 = 0;	b4 = 1 + 0 = 1	b6 = 1 + 0 = 1
a1 = 0 + 1 = 1;	a3 = 1 + 1 = 2;	a5=1 + 0 = 1	a7 = 1 + 0 = 1
b1 = 1 + 1 = 2;	b3 = 0 + 0 = 0;	b5= 1 + 0 = 1	b7 = 0 + 0 = 0
r[0] = a0 – b0 = 0	r[2] = a2 – b2=2	r[4] = a4 – b4 = 0	r[6] = a6 – b6 = 0
r[1] = a1 – b1=–1	r[3] = a3 – b3=2	r[5] = a5 – b5 = 0	r[7] = a7 – b7 = 0

П р и к л а д 2 (η = 3) . Хай перші 3 байти байтової послідовності дорівнюють 0xE0 0x33 0x56, тобто

Номер байту	2		1		0	
Значення байту	0x56		0x33		0xE0	
	01010110		0011 0011		1110 0000	
Порції						
B3	A3	B2	A2	B1	A1	B0	A0
010	101	100	011	001	111	100	000
Суми							

A0 = 0 B0 = 1 r[0] = -1
A1 = 3 B1 = 1 r[1] = 2
A2 = 2 B2 = 1 r[2] = 1
A3 = 2 B3 = 1 r[3] = 1

1.3. Алгоритми пакування та розпакування

Алгоритми пакування $pack^{q}$, $pack^{du}$, $pack^{dv}$ застосовують для перетворення поліномів та векторів у відповідні рядки байтів. Алгоритм $pack^{q}$ застосовують для запису компонентів секретного та відкритого ключа, алгоритми $pack^{du}$, $pack^{dv}$ – для пакування вектора ($pack^{du}$) та полінома ($pack^{dv}$), які є результатом шифрування.

Алгоритми розпакування ($unpack^{q}$, $unpack^{du}$, $unpack^{dv}$) виконують зворотні операції.

1.3.1. Алгоритми $pack^{q}$, $unpack^{q}$ для полінома (pol_pack^{q}, pol_unpack^{q})

1. Усі коефіцієнти полінома приводяться до інтервалу [0.. q–1].

2. Пара коефіцієнтів записується в наступні $\log_2 q$=12 бітів, перший коефіцієнт пари займає молодші, наступний – старші 12 бітів.

У результаті пакування для полінома отримуємо рядок байтів завдовжки poly_len=n*$\log_2 q$/bits_in_byte=384 (байт).

Функція pol_unpack^{q} є зворотньою до функції pol_pack^{q}.

1.3.2. Алгоритми $pack^{du}$ та $pack^{dv}$ для полінома (pol_pack^{du}, pol_unpack^{du}, pol_pack^{dv}, pol_unpack^{dv})

1. Усі коефіцієнти полінома приводяться до інтервалу [0.. q–1].

2. Усі коефіцієнти полінома обчислюються за формулою

$$coef_i := (coef_i * 2^d + q/2)/q,$$

де $d=du$ для pol_pack^{du} та $d=dv$ для pol_pack^{du}.

3. Кожний коефіцієнт записується в пам'ять в наступні d біт, починаючи з молодших байтів.

Алгоритм $pack^{dv}$ застосовують для перетворення поліномів, а $pack^{du}$ – для перетворення елементів вектора.

Функції pol_unpack^{du}, pol_unpack^{dv} є зворотними до функцій pol_pack^{du}, pol_pack^{dv}.

У результаті пакування для полінома отримуємо рядок байтів завдовжки $poly\_len^{dv}=n*dv/bits\_in\_byte=\begin{cases}128 & dv=4\\160 & dv=5\end{cases}$ (байт).

1.3.3. Алгоритми $pack^{q}$, $pack^{du}$ для вектора (vec_pack^{q}, vec_unpack^{q}, vec_pack^{du}, vec_unpack^{du})

У разі застосовування векторів поліномів кожний поліном, починаючи з початку, пакується окремо (алгоритм pol_pack^{q} або pol_pack^{du}) та записується як наступний блок пам'яті. В результаті отримуємо блок пам’яті завдовжки

$$vec\_len = k * poly\_len.$$

Функції vec_unpack^{q}, vec_unpack^{du} є зворотними до функцій vec_pack^{q}, vec_pack^{du}.

Для параметрів алгоритму

$$vec\_len^{q}=\begin{cases}768 & k=2\\1152 & k=3\\1536 & k=4\end{cases} \qquad vec\_len^{du}=\begin{cases}640 & k=2, du=10\\960 & k=3, du=10\\1408 & k=4\ du=11\end{cases}.$$

1.4. Перетворення повідомлення в поліном та зворотне перетворення. Функції pol_from_msg, pol_to_msg

В алгоритмі повідомлення завдовжки 32 байті (256 бітів) і поліном містить 256 коефіцієнтів, тобто один біт повідомлення відповідає одному коефіцієнту полінома та навпаки.

Біт 0 відповідає нульовому коефіцієнту. Біт 1 – коефіцієнту (q+1)/2.

1.5. Генерація ключів

Вихідними даними функції генерації ключів є рядки байтів *pk*, *sk* завдовжки pk_len, sk_len для відкритого та секретного ключів відповідно.

1. Формується випадковий рядок байтів завдовжки seed_len

buf:=gen_rand (seed_len).

2. Обчислюється геш (функція G) для цього рядка завдовжки 2 seed_len, перша половина обчисленого значення застосовується як псевдовипадковий рядок public_seed для створення матриці, а друга половина – в якості псевдовипадкового рядка (noise_seed) для генерації малих поліномів.

3. Генерується квадратна матриця (A), яка має *k* рядків та колонок (алгоритм $Parse^{q}$).

4. Генеруються вектори vec_sk, vec_e завдовжки *k* елементів кожний. Елементом вектора є поліном, коефіцієнти якого цілі числа (алгоритм $Parse^{\eta 1}$).

5. Обчислюється вектор vec_pk=A*vec_sk+vec_e.

6. *sk*, sk_len_internal:=vec_pack^{q}(vec_sk, *k*).

7. *pk*, pk_len:=vec_pack^{q}(vec_pk, k)||public_seed.

8. Генерація випадкового рядка *seed* завдовжки seed_len.

9. *sk*, sk_len:=*sk*||*pk*||*H*(*pk*)||*seed*

Довжина відкритого ключа: pk_len=vec_len+seed_len,

$$pk\_len = \begin{cases} 800 & k = 2 \\ 1184 & k = 3 \\ 1568 & k = 4 \end{cases} \text{ (байтів).}$$

Довжина секретного ключа

sk_len_internal=vec_len (без відкритого ключа),

sk_len=vec_len+pk_len+H_len+seed_len,

$$sk\_len = \begin{cases} 1632 & k = 2 \\ 2400 & k = 3 \\ 3168 & k = 4 \end{cases} \text{ (байтів).}$$

1.6. Інкапсуляція

Виконується несиметричне шифрування обраного повідомлення завдовжки 32 байти за допомогою відкритого ключа отримувача (функція encrypt).

Виконується обчислення загального секрету визначеної довжини.

В х і д .

pk – відкритий ключ отримувача, масив байтів завдовжки pk_len.

В и х і д .

ct – інкапсульований ключ (масив байтів завдовжки crypt_len).

ss – загальний секрет (масив байтів завдовжки).

1. Генерація випадкового рядка *seed* завдовжки seed_len та обчислення його гешу за допомогою функції *H* (саме це значення буде застосовуватись в якості повідомлення для зашифрування):

m:=*H*(*seed*, seed_len).

2. Обчислення гешу від відкритого ключа за допомогою функції *H* buf2:=*H*(*pk*, pk_len).

3. Обчислення гешу за допомогою функції *G*

kr:=*G*(*m*||buf2, 2*seed_len).

4. Обчислення *kr*1 (ліва половина) та *kr*2 (права половина), такі, що

kr:=*kr*1||*kr*2 (*kr*1, *kr*2 завдовжки seed_len).

5. Шифрування *m* за допомогою відкритого ключа отримувача (*pk*) та псевдовипадкового даного *kr*2

ct:=encrypt (*m*, *pk*, *kr*2).

6. Обчислення гешу для *ct* за допомогою функції *H*

$$kr2:=H(ct, \text{crypt\_bytes}).$$

7. Обчислення *ss* за допомогою функції KDF

$$ss:=\text{KDF}(kr1||kr2, 2*\text{seed\_len}).$$

1.6.1. Функція шифрування (encrypt)

В х і д .
m – повідомлення для шифрування, рядок байтів завдовжки seed_len;
pk – відкритий ключ отримувача, рядок байтів завдовжки pk_len;
seed – псевдовипадкове дане завдовжки seed_len.
В и х і д .
ct – результат шифрування повідомлення *m* (рядок байтів завдовжки crypto_len).
1. Ініціалізація *nonce*:=0.
2. Розпаковка відкритого ключа: vec_pk, seed_pk:=vec_unpack^{q}(*pk*).
3. Обчислення полінома pol_p:=pol_from_msg(*m*).
4. Відновлення матриці згідно з seed_pk в транспонованому вигляді (алгоритм Parse^{q})

$$At:=\text{Parse}^{q}(\text{seed\_pk}).$$

5. Генерація вектора *s* згідно з *seed* та поточним значенням nonce (алгоритм Parse$^{\eta 1}$)

$$\text{vec\_s, nonce:= Parse}^{\eta 1}(seed, nonce).$$

6. Генерація вектора vec_e згідно з *seed* та поточним значенням nonce (алгоритм Parse$^{\eta 2}$)

$$\text{vec\_e, } nonce:=\text{Parse}^{\eta 2}(seed, nonce).$$

7. Генерація полінома *ep* згідно з *seed* та поточним значенням nonce (функція Parse$^{\eta 2}$)

$$\text{pol\_ep, } nonce:= \text{Parse}^{\eta 2}(seed, nonce).$$

8. Обчислення:

$$\text{vec\_u}:=At*\text{vec\_s}+\text{vec\_e}; \text{ pol\_v}:=\text{vec\_pk}*\text{vec\_s}+\text{pol\_ep}+\text{pol\_p}.$$

9. Формування інкапсульованого ключа *cc*

$$cc:=\text{vec\_pack}^{du}(\text{vec\_u})||\text{pol\_pack}^{dv}(\text{pol\_v}).$$

Довжина інкапсульованого ключа:

$$\text{cc\_len}=n*(k*du+dv)/\text{bits\_in\_byte};$$

$$\text{cc\_}len=\begin{cases} 768, & k=2, du=10, dv=4 \\ 1088, & k=3, du=10, dv=4 \\ 1568, & k=4, du=11, dv=5 \end{cases}.$$

1.7. Декапсуляція

Генерація загального секрету по зашифрованому тексту *cc* та секретному ключу *sk*.
В х і д .
cc – зашифрований текст;
sk – секретний ключ.
В и х і д .
ss – загальний секрет завдовжки crypto_len.
1. Виділення відкритого ключа та *seed* з секретного

$$pk:=\text{subst}(sk, \text{sk\_len\_internal}, \text{pk\_len}),$$
$$seed:= \text{subst}(sk, \text{sk\_len\_internal}–\text{seed\_len}, \text{seed\_len}).$$

2. Розшифрування повідомлення (функція indcpa_dec)

$$\text{m\_calc:=indcpa\_dec}(ct, sk).$$

3. *buf*:=m_calc||*pk*.
4. Обчислення гешу для *buf*

$$kr:=G(buf).$$

5. *coins*:=*subst*(*kr*, seed_len, seed_len)
6. Зашифрування отриманої послідовності

$$\text{ct\_cmp:=indcpa\_enc}(buf, pk, coins).$$

7. Перевірка коректності.

success:=OK;

if *ct*≠ct_cmp then

success:=ERROR;

end if

8. Обчислення гешу зашифрованого повідомлення (функція *H*) та запис його в другу половину *kr*

kr:=subst(*kr*, 0, seed_len)||*H*(*ct*, crypto_len).

9. Коригування *kr* в залежності від успішності операції розшифрування

if success=ERROR then

kr:=*seed*;

end if

10. Формування загального секрету (функція KDF)

ss:= KDF (*kr*, 2*seed_len).

1.7.1. Розшифрування. Функція indcpa_dec

В х і д .

C – зашифроване повідомлення;

sk – секретний ключ.

В и х і д .

m – повідомлення.

1. Розпаковка *C*. Виділення вектора vec_u та полінома pol_v (функції vec_unpack^{du}, pol_unpack^{dv}).

2. Розпаковка секретного ключа. Обчислення вектора поліномів vec_s (функція vec_unpack^{q}).

3. Обчислення *mp*=pol_v–vec_s*vec_u.

4. Обчислення повідомлення *m*=poly_to_msg (mp).

2. Алгоритм Скеля

Алгоритм є NTRU подібним алгоритмом [4] з полем, визначеним в [5].

2.1. Параметри

Загальні параметри алгоритму наведено в табл. 3.

Таблиця 3

Загальні параметри алгоритму Скеля

Позначення	Призначення	Формула або значення
λ	Рівень криптостійкості	3, 5, 7
n	Степінь полінома. Визначає кількість його коефіцієнтів. Просте число, для якого поліном $x^n–x–1$ є незвідним	$\begin{cases} 881 & \lambda = 3 \\ 1201 & \lambda = 5 \\ 1471 & \lambda = 7 \end{cases}$
p	Менший модуль	*p*=3
q	Більший модуль, просте число, за яким зводять усі коефіцієнти полінома *R*/*q*	$\begin{cases} 7673 & \lambda = 3 \\ 9221 & \lambda = 5 \\ 12251 & \lambda = 7 \end{cases}$
t	Натуральне число, кількість ненульових елементів поліномадорівнює 2*t*	$\begin{cases} 159 & \lambda = 3 \\ 192 & \lambda = 5 \\ 255 & \lambda = 7 \end{cases}$
seed_len	Довжина *seed* (байтів)	$\begin{cases} 32 & \lambda = 3 \\ 48 & \lambda = 5 \\ 64 & \lambda = 7 \end{cases}$

У стандарті застосовують поліноми, позначені $R/3$, R/q. Поліном $R/3$ в якості коефіцієнтів застосовує значення {0, 1, –1}. Поліном R/q в якості коефіцієнтів застосовує цілі значення в інтервалі [0, q–1].

В якості алгоритмів гешування та потокового шифрування можна застосовувати довільні дозволені функції. Для обчислення тестових векторів в якості алгоритму гешування застосовується алгоритм Kupyna (ДСТУ 7564:2014), алгоритму потокового шифрування – Струмок (ДСТУ 8845:2019). У подальшому функції гешування позначені: H256, H512, H. H256 – результат гешування 256 біт, H512 – результат гешування 512 біт, H – результат гешування 256 (для λ=3) та 512 біт (для λ=5, 7).

2.2. Ідентифікатор алгоритму

Алгоритм застосовує ідентифікатор – рядок байтів завдовжки 3 байти. Позначимо цей ідентифікатор OID. Усі функції алгоритму застосовують один ідентифікатор.

2.3. Допоміжні функці

2.3.1. Генерація випадкових та псевдовипадкових послідовностей

Для генерації послідовностей застосовують довільні дозволені алгоритми. Для обчислення тестових векторів застосовують алгоритм ДСТУ 8845:2019 [3]. Перед застосуванням алгоритму треба ініціалізувати генератор випадковими ключем та вектором ініціалізації (функція RandomInit). Довжина ключа визначається рівнем криптостійкості і дорівнює 256 бітів для λ=3 та 512 – для λ=5, 7. Вектор ініціалізації завжди має довжину 256 бітів. Для обчислення ключа (*key*) та вектора ініціалізації (*iv*) для тестових векторів застосовуються два 64-бітних числа *val*1, *val*2. Стан генератора визначається параметром *ctx*, який змінюється після кожного застосування генератора:

$$key := H(val1, 8); \qquad iv := H256(val2, 8).$$

Замість значень ключа та вектора ініціалізації для ініціалізації генератора можна застосовувати рядок байтів (*b*) зазначеної довжини (b_len) (функція PsevdoRandomInit). У цьому випадку

$$\text{Ключ} = H(b, \text{b\_len}).$$

Для генерації вектора ініціалізації застосовують: *OID* (3 байти), останній байт в *b* (*blast*, самий правий байт), рядок з 28 байтів – нулів (zero28)

$$\text{Вектор ініціалізації} = H256(OID||blast||\text{zero28}).$$

В подальшому функція для генерації наступної послідовності позначена gen_posl. Функція gen_posl в якості параметрів приймає *ctx* та довжину послідовності в байтах та повертає в якості результату *ctx* та сформовану послідовність.

2.3.2. Алгоритми пакування

Алгоритми пакування застосовують для перетворення поліномів різних типів в рядок октетів, зворотніх перетворень та для перетворення рядка октетів в цілі невід'ємні числа при заданій кількості бітів для одного числа.

При формуванні бітового рядка передбачається, що біт 0 – самий лівий біт байту (в алгоритмі Kyber біт 0 – самий правий біт байту).

2.3.2.1. Пакування R/3 поліномів з заданою кількістю ненульових елементів (pack^3, unpack^3)

1. Виділяється бітовий рядок *bs* завдовжки $\left\lceil \frac{2t}{8} \right\rceil * 8 + \left\lceil \frac{n}{16} \right\rceil * 16$.

2. В перші $\left\lceil \frac{2t}{8} \right\rceil * 8$ бітів рядка *bs* записується 0, якщо наступний ненульовий коефіцієнт дорівнює –1, та 1, якщо 1.

3. В бітовий рядок *bs*1 записується 0, якщо відповідний коефіцієнт дорівнює 0, та 1 інакше.

Довжина рядка байтів для *R*/3 полінома дорівнює $\left\lceil \frac{2t}{bits\_in\_byte} \right\rceil + 2 * \left\lceil \frac{n}{2*bits\_in\_byte} \right\rceil$.

Для параметрів алгоритму $pol3\_len = \begin{cases} 152 & \lambda = 3 \\ 200 & \lambda = 5. \\ 248 & \lambda = 7 \end{cases}$

2.3.2.2. Пакування R/3 поліномів (pack^{2_3}, unpack^{2_3}) в рядок байтів для довільної кількості ненульових елементів

Цей алгоритм перетворює два сусідніх коефіцієнти полінома в три біта і навпаки відповідно до табл. 4.

Таблиця 4

Кодування коефіцієнтів R/3 полінома

Бітовий рядок	Коефіцієнти полінома	Бітовий рядок	Коефіцієнти полінома
000	0, 0	100	1, 1
001	0, 1	101	1, -1
010	0, -1	110	-1, 0
011	1, 0	111	-1, 1

Довжина рядка байтів для *R*/3 полінома дорівнює $\left\lceil \frac{3(n+1)}{2* bits\_in\_byte} \right\rceil$.

Для параметрів алгоритму $pol3\_len = \begin{cases} 166 & \lambda = 3 \\ 226 & \lambda = 5\cdot \\ 276 & \lambda = 7 \end{cases}$

При перетворенні полінома в рядок байтів (функція pack^{2_3}) можлива помилка, якщо два сусідніх елементи полінома, починаючи з елементу з парним номером, дорівнюють -1, тобто, знайдеться i, для якого $c[2i]$=-1, $c[2i+1]$=-1.

2.3.2.3. Пакування R/q поліномів (pack^{q}, unpack^{q})

Коефіцієнти полінома записуються в бітовий рядок один за одним. Під кожний коефіцієнт відводиться $\log_2 q$ біт.

Довжина рядка байтів для *R*/q полінома дорівнює $\left\lceil \frac{n*\log_2 q}{bits\_in\_byte} \right\rceil$.

Для параметрів алгоритму $polq\_len = \begin{cases} 1432 & \lambda = 3 \\ 2102 & \lambda = 5. \\ 2575 & \lambda = 7 \end{cases}$

2.3.2.4. Пакування R/q поліномів по модулю 4 (pack^{4}, unpack^{4})

Коефіцієнти полінома беруться по модулю 4 та записуються в бітовий рядок один за одним. Під кожний коефіцієнт відводиться 2 біт.

Довжина рядка байтів для *R*/q полінома дорівнює $\left\lceil \frac{2n}{bits\_in\_byte} \right\rceil$.

Для параметрів алгоритму $pol4\_len = \begin{cases} 221 & \lambda = 3 \\ 301 & \lambda = 5\cdot \\ 368 & \lambda = 7 \end{cases}$

2.3.2.5. Розпакування рядка байтів (unpack^{1}, unpack^{w})

Кількість бітів для кожного числа дорівнює 1 або визначається значенням параметра n і приймає значення

$$w = \begin{cases} 11 & \lambda = 3\ (n = 881) \\ 12 & \lambda = 5\ (n = 1201). \\ 12 & \lambda = 7\ (n = 1471) \end{cases}$$

Із заданого рядка байтів обираються наступні 1 або w бітів, починаючи з самого лівого (біту 0). Ця бітова послідовність розглядається як ціле невід'ємне число, яке і є наступним елементом масиву.

2.3.3. Генерація поліномів

2.3.3.1. Генерація R/3 полінома із заданою кількістю ненульових елементів. Функція gen_r3t_pol

Нехай кількість ненульових елементів дорівнює *T*.

Функція генерує псевдовипадковий рядок, в якому спочатку розташовані біти для визначення знаків ненульових елементів (всього $\left\lceil \frac{T}{bits\_in\_byte} \right\rceil$ байтів), а потім – для індексів ненульових елементів, всього потрібно *T* індексів. Для кожного індексу відводиться $w=\lceil \log_2 n \rceil+1$ біт. Для забезпечення однакової імовірності усіх індексів значення поточного числа для *i*-го індексу, яке перевищує $2^w-(2^w \bmod i)$, ігнорується. Ось чому довжина послідовності для генерації індексів перевищує необхідну. Для генерації випадкової перестановки застосовується алгоритм Рональда Фішера (Ronald Fisher) і Франка Йетса (Frank Yates) [6].

В х і д .

ctx – контекст генератора;

T – визначає кількість ненульових елементів.

В и х і д .

ctx – контекст генератора;

r3_pol – *R*/3 поліном.

1. Визначення довжини псевдовипадкового рядка для визначення індексів ненульових елементів та їх знаків. Для індексів виділяється подвійна пам'ять

$$\text{len\_sign}:=\left\lfloor \frac{T}{bits\_in\_byte} \right\rfloor + 1;\ \text{len\_ind}:=2 * \left\lceil \frac{\log_2 n*T}{bits\_in\_byte} \right\rceil.$$

2. Генерація псевдовипадкового рядка для знаків та індексів

buf:=gen_posl (len_sign+len_ind)

sign_buf:=subst (Buf, 0, len_sign);

ind_buf:=subst (Buf, len_sign, len_ind);

3. Перетворення sign_buf в масив чисел {0 або 1}

sign, count_sign:= unpack[1] (sign_buf, len_sign);

4. Створення випадкової послідовності

r3_pol=0 {*i* = 0..*n*–1} Початкова ініціалізація

k=0, *i*=*n*–*T*, *j*=0, *w*=$\log_2 n$+1, max_ind=2^w

while *k*<*T* do

ind, count_ind:=unpack^w (ind_buf, len_ind);

while *i*<*n* та *j*<count_ind та *k*<*T* do

ind=ind[*j*];

j=*j*+1;

if ind<max_ind–max_ind mod *i* then

ind = ind mod *i*;

r3_pol [*i*] = r3_pol [ind]; r3_pol [ind] = 1

if sign [*k*] == 0 then

r3_pol [ind] = -1

end if

k=*k*+1

i=*i*+1

end if

end while

if *k*!= T then Обробка закінчення послідовності ind

len_ind := $\left\lceil \frac{2 * (T - k) * w}{bits\_in\_byte} \right\rceil$

ind_buf, *ctx*:=gen_posl (*ctx*, len_ind)

end if

end while

2.3.3.2. Генерація R/3 полінома з довільною кількістю ненульових елементів. Функція gen_r3_pol

Спочатку генерують псевдовипадковий рядок байтів для обчислення коефіцієнтів полінома. Якщо значення наступного байту ≥243 (3^5), цей октет ігнорують, в іншому разі його застосовують для формування значень п'яти коефіцієнтів, а саме – значення октету переводять в трирічну систему числення, тобто формують відповідні значення $b_4b_3b_2b_1b_0$, які записують як відповідні коефіцієнти полінома в форматі {-1, 0, 1}.

П р и к л а д .

Нехай потрібно визначити коефіцієнти полінома, які відповідають октету 0x56. Шістнадцятковому значенню 56 відповідає десяткове значення 86, що дорівнює 10012 в трирічній системі. Отримуємо відповідні значення коефіцієнтів полінома a_4=1, a_3=0, a_2=0, a_1=1, a_0=-1.

Імовірність появи значень ≥243 за рівноімовірного генератора становить (256–243)/256, тобто менше 6 %. Для забезпечення достатньої кількості сформованих значень рекомендують формувати кількість октетів в два рази більше, ніж потрібно.

2.3.4. Обчислення загального секрету (ss) та частини інкапсульованого ключа (C)

В х і д .
r – *R*/3 поліном;
ss_len – довжина *ss*.
В и х і д .
C – частина інкапсульованого ключа, рядок байтів;
C_len – довжина *C*.
ss – загальний секрет (рядок байтів).
1. Пакування полінома *r*
temp, temp_len := pack^3(*r*)
2. Обчислення *C*, *ss*
if λ=3 then
temp:=H512 (temp, temp_len)
C_len:=32;
C:=subst (temp, 0, seed_len);
ss:= subst (temp, seed_len, ss_len)
else
*temp*1:=*temp*||1; *temp*2:=*temp*||2
C:= H512 (*temp*1, temp_len+1)
C_len := 64;
*temp*2 := H512 (*temp*2, temp_len+1)
ss:= subst (*temp*2, 0, ss_len)
end if

2.4. Генерація ключів (gen_keys)

В х і д .
Немає.
В и х і д .
sk – секретний ключ завдовжки $\left\lceil \frac{2t+n}{bit\_byte} \right\rceil$,
pk – відкритий ключ завдовжки $\left\lceil \frac{n*\log_2 q}{bit\_byte} \right\rceil$.
1. Ініціалізація генератора випадкових чисел
ctx:=RandomInit()
2. Генерація полінома *G*
G, *ctx*:= gen_r3t_pol (*ctx*, 2**n*/3+1)

3. success:=ERROR
4. while success=ERROR do Цикл генерації полінома *f*
Генерація полінома *F*
F, ctx:= gen_r3t_pol (*ctx*, 2**t*)
Обчислення полінома *f*
f:=(1+*pF*) mod *q*
Обчислення інверсії для *f* в полі $x^n–x–1$
*f*_1, success := inverse (*f*)
end while
5. Генерація відкритого ключа
h:=*g***f*_1 в полі $x^n–x–1$
6. Формування *pk*, *sk*
pk:=pack^q (*h*); *sk* := pack^3 (*F*)

$$\text{sk\_len} = \begin{cases} 152 & \lambda = 3 \\ 200 & \lambda = 5 \\ 248 & \lambda = 7 \end{cases} \qquad \text{pk\_len} = \begin{cases} 1432 & \lambda = 3 \\ 2102 & \lambda = 5 \\ 2575 & \lambda = 7 \end{cases}.$$

2.5. Алгоритм інкапсуляції

У процесі виконання алгоритму формується випадкове повідомлення завдовжки не менше 32 байтів і не більше, ніж pol3_len–seed_len–1 та дорівнює $\max\_msg\_len = \begin{cases} 133 & \lambda = 3 \\ 177 & \lambda = 5 \\ 211 & \lambda = 7 \end{cases}$. Для усіх параметрів max_msg_len не перевищує 255, тому для запису довжини повідомлення відводиться один байт.

В х і д .
pk – відкритий ключ одержувача;
ss_len – загальний секрет, довжина, байтів.
В и х і д .
success – успішність операції, OK – для успішної операції, ERROR – навпаки;
ss – загальний секрет – рядок байтів, довжина якого не перевищує seed_len;
cc – інкапсульований ключ (рядок байтів).
1. Перевірка параметра ss_len
if ss_len <= seed_len
success := OK
else
success := ERROR
end if
2. Продовження, якщо *success*=OK
if *success* = OK then
3. Розпакування відкритого ключа та генерація випадкового повідомлення
h:=unpack^q (*pk*); msg, msg_len := gen_msg ();
success:=ERROR
4. while *success*=ERROR do Цикл шифрування
5. Формування рядка *M*
M:=*seed*||msg_len||*msg*||*zero*
де:
seed – псевдовипадковий рядок завдовжки seed_len;
msg, msg_len – повідомлення для шифрування (*msg*) та його довжина (seed_len≤msg_len≤max_msg_len);
zero – рядок байтів с кодом 0 завдовжки max_msg_len–msg_len.
6. Перетворення рядка *M* в *R*/3 поліном

MTrin:=unpack$^{2-3}$(*M*)

7. Формування рядка *s* та його довжини

s_len:=3+m_len+2*seed_len

s:=*OID*||*m*||*seed*||*subst*(*pk*, 0, seed_len)

8. Формування *R*/3 полінома з 2*t* ненульовими елементами згідно з рядком *s*

ctx:=PsevdoInitRandom (*ctx*, *s*, s_len)

r:=gen_r3t_pol (*ctx*, 2*t*)

9. Обчислення

R:=*r***h* в полі *R*/*q*; *R*4:= pack^{4} (*R*);

mask:= gen_r3_pol (*R*4);

*pol*3:=(*MTrin*+*mask*) mod *p*

10. Перевірка умови за кількістю (ненульових (*k*1), нульових (*n*–*k*1)) елементів полінома *pol*3

if *k*1>=2**t* та *n*–*k*1>=*t* then

success=OK

end if

11. end while

12. Результат шифрування

c:=(*R*+*pol*3) mod *q*; *cc*:=pack^{q} (*c*)

13. Обчислення загального секрету та частини інкапсульованого ключа

ss, *C*, C_len:=gen_C_ss (*r*)

14. Обчислення інкапсульованого ключа

cc:=*C*||*cc*; cc_len:=C_len+polq_len

end if

2.6. Алгоритм декапсуляції

В х і д .

cc – інкапсульований ключ;

sk – секретний ключ отримувача;

pk – відкритий ключ отримувача.

ss_len – довжина загального секрету (байтів).

В и х і д .

success – ознака успішності;

ss – загальний секрет

1. Перевірка коректності довжини загального секрету ss_len.

success:=ERROR

if success ≤seed_len then

2. Розпакування секретного ключа, відкритого ключа та окремих компонентів *cc*.

F:=unpack^{3}(*sk*); *h*:=unpack^{q}(*pk*); *f*=*p***F*+1

3. Розпакування окремих компонентів *cc*.

if λ=3 then C_len:=32; else C_len:=64 end if

C'=subst (*cc*, 0, C_len); *c*'=subst (*cc*, C_len, polq_len);

e:=unpack^{q}(*c*')

4. Розшифрування

a'=*e***f* в полі *R*/*q* *pol*3:=*a*'mod *p*

5. Перевірка умови за кількістю (ненульових (*k*1), нульових (*n*–*k*1)) елементів полінома *pol*3.

success:=ERROR;

if *k*1>=2**t* та *n*–*k*1>=*t* then

success=OK

end if

6. Відновлення рядка *M*.
if *success*=OK then
R:=(*e*–*pol*3) mod *q*; *R*4:= pack^{4} (*R*);
mask:=gen_r3_pol (*R*4);
MTrin:=*pol*3–*mask*; *M*, *success*:=pack$^{2-3}$(*MTrin*)
end if
7. В разі успішного завершення відновлення рядка байтів *M* та перевірка коректності усіх полів.
if *success*=OK then
success=ERROR;
seed:=subst (*M*, 0, seed_len);
msg_len:=subst (*M*, seed_len, 1);
if msg_len≥*seed* та msg_len≤max_msg_len then
msg:= subst (*M*, seed_len+1, msg_len);
zero_count := max_msg_len–msg_len;
zero':=subst (*M*, seed_len+1+msg_len, zero_count);
if zero_count байтів *zero*' дорівнюють 0 then
success=OK;
end if
endif
8. У разі успішного завершення відновлення рядка байтів *s* та полінома *r*
if *success*=OK then
s:=*OID*||*msg*||*seed*||subst (*pk*, 0, seed_len)
s_len:=3+msg_len+seed_len+seed_len
ctx:=PsevdoInitRandom (*ctx*, *s*, s_len)
r:=gen_r3t_pol (*ctx*, 2*t*)
9. Обчислення полінома *R*' та порівняння його з *R*.
R'=*r***h*
if *R*'≠*R* then
success:= ERROR;
end if
end if
10. У разі успішного завершення обчислення *ss*
if *success* = OK then
ss, *C*, C_len:=gen_C_ss (*r*)
if *C*'≠*C* then
success:=ERROR;
end if
end if

3. Порівняння алгоритмів

3.1. Порівняння розмірів ключів та інкапсульованих даних

Для порівняння застосовується загальний режим при λ=3, 5.

Таблиця 5

Порівняння розмірів ключів та інкапсульованих даних

Алгоритм	λ=3			λ=5		
	pk_len	sk_len	Cc_len	pk_len	sk_len	Cc_len
Kyber	1184	2400	1088	1568	3168	1568
Скеля	1432	152	1464	2102	200	2166

Порівняння розмірів показує, що розмір секретного ключа для алгоритму Скеля суттєво менший за розмір секретного ключа для алгоритму Kyber, ніж у випадку, якщо разом

з секретним ключом зберігати відкритий, як це робиться в алгоритмі Kyber. Розміри відкритого та інкапсульованого ключа для алгоритму Kyber менші.

3.2. Порівняння швидкодії

Порівняння швидкодії виконується для криптостійкостей, які підтримуються для обох алгоритмів, тобто λ=3, 5.

Для визначення швидкодії алгоритму Kyber застосовується реалізація авторів алгоритму Kyber (round 3, Optimized_Implementation).

Для визначення швидкодії алгоритму Скеля застосовується реалізація авторів статті.

Характеристика обладнання для експерименту:

11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHz.

Microsoft Visual C++ 2019 00435-00000-00000 AA349.

Mode=64 bit.

Таблиця 6

Порівняння швидкодії

Алгоритм	λ=3			λ=5		
	GenKeys	Incaps	Decaps	GenKeys	Incaps	Decaps
Kyber	127018	139615	160693	216923	232294	234565
Скеля	605320	59532	75849	1048405	82969	105700

Час генерації ключів для Скелі перевищує час генерації ключів для Kyber, але у той же час у алгоритмі Kyber відсутня операція інверсії, а у Скелі вона є, і вона є досить ресурсоємною.

Висновок

При збільшенні довжини відкритого ключа та результату інкапсуляції алгоритм Скеля більш ефективний, ніж алгоритм Kyber для інкапсуляції та декапсуляції.

Список літератури:

1. Post-Quantum Cryptography. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/.

2. ДСТУ 8961:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритми асиметричного шифрування та інкапсуляції ключів. Режим доступу: http://online.budstandart.com/ua/catalog/doc-page.html?id_doc=88056.

3. DSTU8845. [Електронний ресурс]. Режим доступу: https://github.com/outspace/dstu8845.

4. NTRU. A submission to the NIST post-quantum standardization effort. [Електронний ресурс]. Режим доступу: https://ntru.org/.

5. NTRU Prime. [Електронний ресурс]. Режим доступу: https://ntruprime.cr.yp.to/.

6. D. Knuth The Art of Computer Programming vol. 2. 3rd. Boston : Addison–Wesley, 1998. P. 145–146.

Надійшла до редколегії 05.08.2022

Відомості про авторів:

Горбенко Іван Дмитрович – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ "Інститут Інформаційних Технологій", головний конструктор, Україна; e-mail: gorbenkoi@iit.kharkov.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», начальник відділу програмування; Україна; e-mail: iit@iit.kharkov.ua; ORCID: https://orcid.org/0000-0001-9249-0497

Єсіна Марина Віталіївна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, доцент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; науковий співробітник-консультант АТ «Інститут Інформаційних технологій»; Україна; e-mail: m.v.yesina@karazin.ua; ORCID: https://orcid.org/0000-0002-1252-7606

Пономар Володимир Андрійович – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, науковий співробітник кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: Laedaa@gmail.com; ORCID: https://orcid.org/0000-0001-5271-2251

УДК 004.056.5 **DOI:10.30837/rt.2022.3.210.02**

Ю.І. ГОРБЕНКО, канд. техн. наук, С.О. КАНДІЙ

ПОРІВНЯННЯ АРГУМЕНТІВ БЕЗПЕКИ ПЕРСПЕКТИВНИХ МЕХАНІЗМІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ

Вступ

Постквантова криптографія є напрямом досліджень, що направлені на розробку та стандартизацію асиметричних криптографічних перетворень, які суттєво будуть захищені від квантових та класичних атак [1 – 3]. У 2016 році NIST США оголосили про початок конкурсу NIST PQC [1], метою якого є створення нових постквантових стандартів криптографічних перетворень. Наразі завершився третій етап цього конкурсу. Спеціалісти NIST обрали механізм інкапсуляції ключів CRYSTALS-Kyber [2] для стандартизації і ряд інших механізмів для подальшого вивчення у четвертому етапі. У той же час в Україні вже стандартизовано механізм інкапсуляції ключів ДСТУ 8961:2019 “Скеля” [3], а в європейській спільноті доволі популярним механізмом є FrodoKEM [4]. Таке різноманіття KEM робить актуальною проблему порівняння та аналізу безпеки цих механізмів.

Метою статті є порівняння теоретичних та практичних аргументів безпеки постквантових механізмів інкапсуляції ключів CRYSTALS-Kyber, FrodoKEM та ДСТУ 8961:2019 “Скеля”, а також розробка рекомендацій з їх використання у світовій на національній практиці.

1. Формальні визначення безпеки

Особливістю сучасної криптографії є те, що безпека кожної криптографічної схеми або протоколу підкріплюється формальними математичними доказами, які, звісно ж, не гарантують, що не існує атак взагалі, проте дозволяють гарантувати, що не існує атак певного виду, які формально визначені через модель атак [5]. Для практичних задач корисними є наступні моделі:

- *Модель атак на основі обраних відкритих текстів (CPA)* [6]. Супротивник обирає відкритий текст, а потім отримує відповідний зашифрований текст. Супротивник використовує отриману інформацію, щоб відновити відповідний відкритий текст для шифротексту, який раніше не бачив. Схеми шифрування з відкритим ключем є прикладом, коли супротивник може зашифрувати будь-яке повідомлення за своїм вибором під відкритим ключем жертви. Модель атак з адаптивно обраним відкритим текстом (CPA2) – це атака CPA, у якій вибір відкритого тексту супротивником може залежати від зашифрованого тексту, створеного під час попередніх шифрувань.
- *Модель атак на основі обраного шифротексту (CCA)* [7]. Під час атаки за допомогою обраного шифротексту супротивник може розшифрувати довільні зашифровані тексти, наприклад, за допомогою доступу до обладнання для дешифрування з надійно вбудованим ключем дешифрування. Мета полягає в тому, щоб вивести відкритий текст із раніше не баченого шифротексту. CCA має два спеціальні варіанти: у неадаптивній атаці на основі шифротексту (CCA1), яку також називають «атакою в обідній час» або «опівночі», зловмисник може мати доступ до системи лише протягом обмеженого часу або обмеженої кількості пар відкритий текст-зашифрований текст. Атаку називають неадаптивною, оскільки зловмисник не може адаптувати свої запити до оракула дешифрування відповідно до зашифрованого тексту виклику. У CCA1 зашифрований текст виклику надається після закінчення терміну дії можливості супротивника здійснювати вибрані запити зашифрованого тексту. Однак зловмисник може зробити адаптивні запити обраного зашифрованого тексту до того, як буде надано зашифрований текст виклику. В адаптивній атаці обраного зашифрованого тексту (CCA2), яка є сильнішою, ніж CCA1, зловмисник має доступ до оракула дешифрування

навіть після отримання зашифрованого тексту. У CCA2 запити супротивника до оракула дешифрування можуть залежати від зашифрованого тексту, але супротивник може не запитувати розшифрування самого зашифрованого тексту виклику.

Щоб застосовувати моделі атак для реальних схем та протоколів, необхідно ввести додаткові поняття, які дозволяють формалізувати визначення моделей CPA, CCA. Ідея поняття нерозрізнювальності (indistinguishability) [8] шифротекстів полягає у тому, що аналітик не може на практиці дізнатися будь-яку значну інформацію про відкритий текст, що лежить в основі зашифрованого тексту. Вважається, що схема має властивість нерозрізнювальності шифротекстів, якщо ймовірність того, що аналітик зможе дізнатися інформацію про відкритий текст, є незначною. Під «незначною» мається на увазі, що ймовірність зменшується зі збільшенням певного параметра безпеки λ швидше (починаючи від деякого λ_0), ніж будь-яка функція вигляду $|1/f(\lambda)|$, де $f(\lambda)$ – будь-який поліном.

На основі цього можливо визначити IND-CPA, IND-CCA1, IND-CCA2 безпеку для схем шифрування з відкритим ключем наступним чином [5]:

Нехай $П = (Gen, Enc, Dec)$ позначає схему шифрування з відкритим ключем, а $A = (A_1, A_2)$ позначає супротивника з двома підалгоритмами. Для атаки $atk \in \{cpa, cca1, cca2\}$ і параметра безпеки λ ймовірність успіху супротивника визначається як $Adv_{A,П}^{ind-atk}(n) = |\Pr[Exp_{A,П}^{ind-atk-1}(n) = 1] - \Pr[Exp_{A,П}^{ind-atk-0}(n) = 1]|$, для $b \in \{0,1\}$, експеримент $Exp_{A,П}^{ind-atk-b}(n) = b'$ визначається як

$$
\begin{array}{ccc}
(pk, sk) & \leftarrow & Gen(1^n) \\
(m_0, m_1, s) & \rightarrow & Gen(1^n) \\
b & \in_R & \{0,1\} \\
c & \leftarrow & Enc_{pk}(m_b) \\
b' & \leftarrow & A_2^{O_2}(m_0, m_1, s, c) \\
Повернути & b' &
\end{array}
$$

де для

$$
\begin{aligned}
&atk = cpa \Rightarrow O_1(\cdot) = \varepsilon, O_2(\cdot) = \varepsilon \\
&atk = cca1 \Rightarrow O_1(\cdot) = O_{Dec}(\cdot), O_2(\cdot) = \varepsilon \\
&atk = cca2 \Rightarrow O_1(\cdot) = O_{Dec}(\cdot), O_2(\cdot) = O_{Dec}(\cdot)
\end{aligned}
$$

Схема шифрування безпечна в сенсі IND-ATK, якщо $Adv_{A,П}^{ind-atk}(\cdot)$ є незначним у λ [9].

2. Ідеалізовані моделі безпеки

Довести безпеку криптографічних схем або протоколів доволі важко у IND-CPA/IND-CCA моделях через використання криптографічних геш-функцій. У тому, як геш-функція взаємодіє з повідомленням бо іншими змінними, можуть бути вразливості, навіть якщо сама геш-функція є криптографічно безпечною [10]. На практиці лише до невеликої кількості геш-функцій можливо застосувати необхідний аналіз. Для інших випадків застосовуються ідеалізовані моделі безпеки [5].

Популярним вибором є модель випадкового оракула (ROM) [10]. У цій моделі геш-функції замінюються на їх ідеалізований варіант – випадкових оракулів. Випадковий

оракул це функція, яка на кожен запит повертає істинно випадкове значення з рівномірного розподілу. Значення для кожного аргументу не повторюються. Звичайно, у моделі ROM, можливо довести захист не від всіх атак, проте доказ у моделі ROM є гарним аргументом безпеки.

Модель квантового випадкового оракула (QROM) [11] є аналогом ROM для квантових комп'ютерів. Квантовий супротивник, знаючи як реалізувати деяку функцію f у вигляді послідовності геймів, може реалізувати унітарний оператор, що асоційований до f і дозволяє робити запити до суперпозиції станів $f:\sum_x \alpha_x |x\rangle \mapsto \sum_x \alpha_x |x\rangle |f(x)\rangle$. У моделі QROM супротивник може реалізувати деякий невідомий оператор U_H, який визначений як

$$U_H : |x\rangle |y\rangle \mapsto |x\rangle |y \oplus H(x)\rangle, \quad (1)$$

де $H(x)$ – випадковий оракул. Ця модель є корисною, оскільки існують квантові атаки, що не використовують властивостей геш функцій [12]. Наприклад, багато класичних атак можливо прискорити за допомогою алгоритма Гровера. Проте, варто зауважити, що модель квантового оракула передбачає, що обчислення ведуться на стандартній квантовій машині Тьюринга і не враховує випадки адіабатичних обчислень та інших нестандартних моделей. За необхідності, звичайно ж, можна адаптувати для них QROM, проте цей напрям досліджень наразі майже не розвивається.

3. Складні проблеми у криптографії на решітках

Більшість схем шифрування з асиметричним ключем побудовані на нерозв'язності деяких складних проблем [13]. Під складністю проблеми у теоретико-числовому сенсі зазвичай розуміється складність вирішення найгіршого випадку. Для того щоб проблему можна було використовувати в криптографії, необхідна не тільки складність у найгіршому випадку, але й складність у середньому. Необхідно, щоб ймовірність того, що випадковий екземпляр проблеми можна швидко вирішити, була незначною [14]. У криптографії на решітках використовуються здебільшого проблеми, які мають редукції "від найгіршого до середнього" з проблемами теорії решіток, що є унікальною властивістю, яка з теоретичної точки зору, значно збільшує безпеку схем. Вперше така проблема, що має редукції "від найгіршого до середнього" з проблемами теорії решіток, була запропонована у 2005 році – проблема навчання з помилками (LWE) [15]. Для її формального визначення введемо поняття LWE-розподілу.

Нехай $s \in \mathbf{Z}_q^n$, χ-розподіл ймовірностей для помилок. LWE-розподіл $A_{s,\chi}$ над $\mathbf{Z}_q^n \times \mathbf{Z}_q$ формується через вибір $a \in \mathbf{Z}_q^n$ з рівномірного розподілу, $e \leftarrow \chi$ над $\mathbf{Z}$ та вихідною є пара $(a,b), b = \langle s, a \rangle + e \bmod q$.

Проблема $Search - LWE_{n,m,q,B,\chi}$ формулюється наступним чином. Нехай $s \in \mathbf{Z}_q^n$ обрано з деякого розподілу ймовірностей B. Дано m незалежно отриманих екземплярів $(a_1,b_1),...,(a_m,b_m) \in \mathbf{Z}_q^n \times \mathbf{Z}_q$ з LWE-розподілу $A_{s,\chi}$. Потрібно знайти s.

Проблема $Decision - LWE_{n,m,q,B,\chi}$ формулюється наступним чином. Нехай $s \in \mathbf{Z}_q^n$ обрано з деякого розподілу ймовірностей B. Дано m незалежно отриманих екземплярів $(a_1,b_1),...,(a_m,b_m) \in \mathbf{Z}_q^n \times \mathbf{Z}_q$ з LWE-розподілу $A_{s,\chi}$ або з рівномірного розподілу. Потрібно визначити, з якого саме розподілу були отримані екземпляри $(a_1,b_1),...,(a_m,b_m) \in \mathbf{Z}_q^n \times \mathbf{Z}_q$.

У $Search - LWE_{n,m,q,B,\chi}$ без втрати загальності можна вважати, що розподіли B, χ є нормальньним розподілом з параметром α (розподіл можливо звести до нормального за

допомогою саморедукції [16]). Надалі позначатимемо таку версію $Search-LWE$ як $Search-LWE_{n,m,q,\alpha}$

Існує як квантова, так і класична редукція від проблеми $GapSVP_\gamma$ до $Search-LWE_{n,m,q,\alpha}$. Проблема $GapSVP_\gamma$ визначається наступним чином. Нехай задано решітку Λ розмірності n та деякий вектор v. Необхідно визначити чи знаходиться найближчий до v вектор на решітці на відстані $(1,\gamma]$. Теоретико-числова складність $GapSVP_\gamma$ залежить від значення параметра γ. У табл. 1 зведені відомі результати про складність $GapSVP_\gamma$ в залежності від γ.

Таблиця 1

Складність $GapSVP_\gamma$

Значення параметра γ	Клас складності
$2^{(\log n)^{1-\varepsilon}}$	NP-складна
$\sqrt{n}$	$NP \cap coNP$
$poly(n)$	Невідомо. Найкращі алгоритми (BKZ і т.д.) дають експоненційний час роботи.
$2^{\sim n}$	P

Для $Search-LWE_{n,m,q,\alpha}$ відома наступна редукція. Нехай $n,q \geq 1$ – цілі числа, $\alpha \in (0,1)$, для якого виконується $\alpha q \geq 2\sqrt{n}$. Тоді існує квантова редукція від найгірших випадків n-мірного $GapSVP_{O(n/\alpha)}$ до $Search-LWE_{n,q,\alpha}$. Якщо $q \geq 2^{n/2}$, то існує також квантова редукція.

Ця редукція актуальна саме для тих випадків, що використовуються в криптографії. До LWE існує дуальна проблема – SIS, яка визначається наступним чином.

Нехай $n,m,q>0$ є цілими числами, $\beta>0$ – дійсне число. Дана матриця $A \in \mathbf{Z}_q^{n\times m}$ з рівномірного розподілу, необхідно знайти ненульовий вектор $z \in \mathbf{Z}^m$ з евклідовою нормою $\|z\| \leq \beta$, для якого виконується $Az = 0 \in \mathbf{Z}_q^n$.

Для $SIS_{n,m,q,\beta}$ також існує редукція від найгіршого до середнього з $GapSVP$, проте більш важлива редукція $SIS_{n,m,q,\beta}$ до іншої складної проблеми з теорії решіток – $SIVP_\gamma$, яка полягає у знаходженні n лінійно незалежних векторів на решітці, для яких виконується $\|b_i\| \leq \gamma\lambda_i(\Lambda)$. Для будь-яких поліноміально обмежених m,β та простого $q \geq \beta\omega(\sqrt{n\log n})$ проблема $SIS_{m,m,q,\beta}$ така ж складна, як і $SIVP_\gamma$ з фактором $\gamma = \beta O(\sqrt{n})$.

Існує багато поліноміальних редукцій для різноманітних проблем на решітках. Вони утворюють складний ландшафт. Базові проблеми та деякі відомі редукції [17] зображено на рис. 1.

Популярною модифікацією є LWE та SIS на структурованих решітках [18]. Структурованість додається завдяки використанню елементів певного поля $R_q = \mathbf{Z}[X]/(f(x))$. За визначенням решітка є дискретною абелевою групою [19]. То ж, будь-яка підструктура у полі R_q, що є дискретною абелевою групою, може бути розглянута як решітка. Наприклад, розглянемо ідеал $\langle a \rangle$ для елемента $a \in R_q$. Кожен елемент в цьому ідеалі має вигляд $a \cdot s, s \in R_q$. Цей ідеал може бути вкладений у Евклідовий простір за допомогою звичайного коефіцієнтного вкладення:

$$Vec(a_0 + a_1 x + ... + a_{n-1}x^{n-1}) \mapsto (a_0, a_1, ..., a_{n-1}). \quad (2)$$

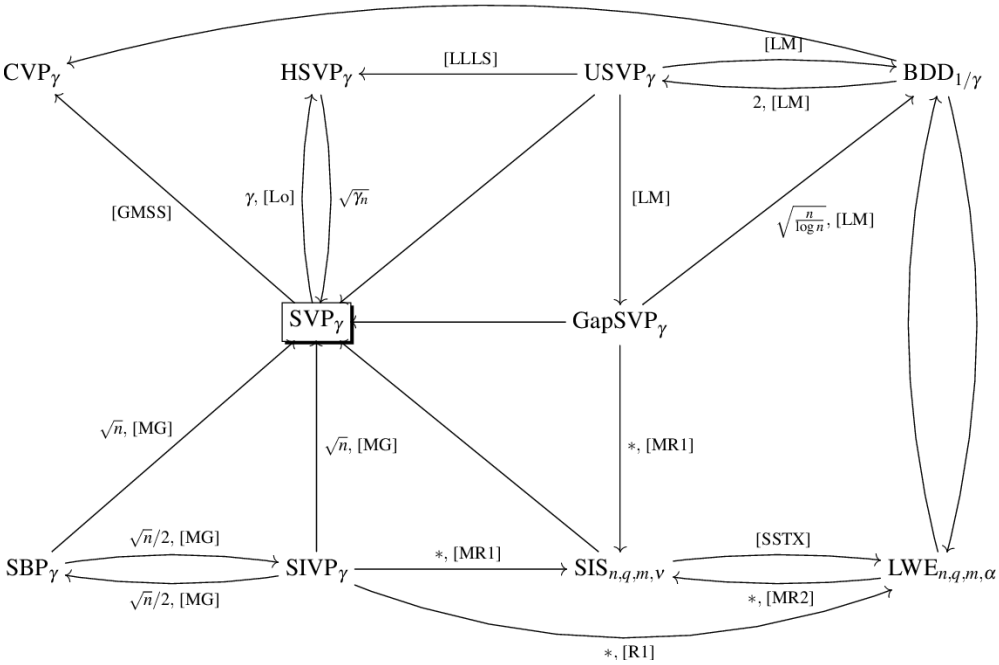


Рис. 1. Поліноміальні редукції між складними проблемами в теорії решіток

Тобто кожному поліному відповідає вектор його коефіцієнтів. Якщо правильно ввести операції додавання та множення на множині усіх векторів $(a_0, a_1, ..., a_{n-1})$, то зберігатиметься структура кільця. Операцію додавання можливо ввести покоординатно:

$$(b_0, b_1, ..., b_{n-1}) + (b_0, b_1, ..., b_{n-1}) = (a_0 + b_0 \bmod q, a_1 + b_1 \bmod q, ..., a_{n-1} + b_{n-1} \bmod q)$$

Множення можливо ввести наступним чином:

$$rot(a) \cdot (b_0, b_1, ..., b_{n-1}) = (c_0 \bmod q, c_1 \bmod q, ..., c_{n-1} \bmod q), \tag{3}$$

де $rot(a)$ є матрицею $n \times n$, яка задається як

$$rot(a) = \begin{pmatrix} Vec(a \bmod f(x)) \\ Vec(a \cdot x \bmod f(x)) \\ \cdots \\ Vec(a \cdot x^{n-1} \bmod f(x)) \end{pmatrix} \tag{4}$$

Фактично $rot(a)$ задає систему з n лінійно незалежних векторів [19]. Оскільки усі елементи вектора $(b_0, b_1, ..., b_{n-1})$ є цілими числами, то маємо звичайну решітку у евклідовому просторі, яка є структурованою у тому сенсі, що базис $rot(a)$ має додаткову структуру ідеалу. Решітки з базисом вигляду

$$\begin{pmatrix} rot(a_1) \\ rot(a_2) \\ \cdots \\ rot(a_m) \end{pmatrix}, a_1, a_2 ..., a_m \in R_q \tag{5}$$

мають назву “ідеальні решітки”. Відповідні складні проблеми з теорії решіток позначаються з префіксом “Ideal-“: $Ideal - SIVP, Ideal - GapSVP, Ideal - SVP$, тощо. Відповідні версії LWE

та SIS мають назву Ring-LWE та Ring-SIS, які також мають відповідні редукції від $Ideal-SIVP, Ideal-GapSVP$. На жаль, виявилося, що $Ideal-SVP$ на квантовому комп'ютері може бути вирішений за поліноміальний час [20].

Наразі популярним узагальненням є Module-LWE (Module-SIS) [22], яке використовує структуру R_q-модуля. Така решітка має вигляд

$$\begin{pmatrix} rot(a_{1,1}) & rot(a_{1,2}) & \dots & rot(a_{1,m}) \\ rot(a_{2,1}) & rot(a_{2,2}) & \dots & rot(a_{2,m}) \\ \vdots & \vdots & \ddots & \vdots \\ rot(a_{m,1}) & rot(a_{m,2}) & \dots & rot(a_{m,m}) \end{pmatrix}, a_{i,j} \in R_q \quad (6)$$

Формальне визначення $Module-SIS_{R,m,k,q,\beta}$ наступне:

Нехай дано m векторів поліномів $a_1,...,a_m \in R_q^k$, що обрані з рівномірного розподілу. Розглянемо їх як стовпці матриці $A \in R_q^{m\times k}$. Необхідно знайти ненульовий вектор поліномів $z \in R_q^k$ з нормою $Az = 0$.

Відповідно визначення $Module-LWE_{R,m,k,q,B,\chi}$ наступне:

Нехай $s \in R_q^k$ – вектор поліномів, що обрано з деякого розподілу B. Дано m екземплярів $(a_1,b_1),...,(a_m,b_m) \in R_q^{m\times k}$. Необхідно визначити чи взяті вони з рівномірного розподілу, або з розподілу LWE $A_{R_q,s,\chi}$ над полем R_q.

Module-LWE також має редукції з $Module-SIVP, Module-GapSVP$ [23]. Наразі невідомо про безпосередні атаки на Module-LWE чи відповідні складні проблеми [2], проте існує поліноміальна редукція від Module-LWE до Ring-LWE, що фактично означає існування квантового поліноміального алгоритму для Module-LWE за певних параметрів [24]. Проте, за тих значень, за яких працює редукція, для криптографічних застосувань ця атака не є актуальною. Хоча сама наявність такої редукції є поганим знаком, проте для Module-LWE невідомо ефективних атак, що враховували модульну структуру решітки.

Існують також інші модифікації, які вносять додаткову структурованість: MP-LWE, Order-LWE, Poly-LWE, Cyclic-LWE, тощо. Гарний огляд існуючих варіантів LWE на структурованих решітках наведено в роботі [18]. Основною проблемою структурованих варіантів є те, що є невідомою складність проблем на таких структурованих решітках. У деякому сенсі, серед зазначених проблем на структурованих решітках найкращі докази має MP-LWE. Є багато редукцій різних проблем до MP-LWE, що є доволі сильним аргументом безпеки. Проте, на практиці, вона працює значно повільніше, то ж не набула популярності.

Варто окремо зазначити проблему Continuous-LWE [25], яка є нещодавньою розробкою та притягла до себе увагу багатьох дослідників. Вона має доволі сильні докази складності. Детальний огляд цієї проблеми виходить за межі цієї роботи, аналіз можливо знайти в роботі [25]. Окремою її особливістю є доказова захищеність від атак з використанням машинного навчання.

Дещо окремо від інших криптографічних проблем на решітках стоїть проблема NTRU [26]. Фактично, NTRU була першою проблемою на решітках, що дозволила створювати дійсно практичні криптографічні схеми та протоколи і має наступне формальне визначення:

Нехай q є цілим додатнім числом. Дано елемент $h \in R_q$ з деякого розподілу D, для якого виконується $h \cdot f = g \bmod q, (f, g) \in R^2, \|f\|, \|g\| \le \sqrt{q} / \gamma$.

У проблемі $Search - NTRU_{R,q,D,\gamma}$ необхідно знайти пару $(f, g) \in R^2$.

Щоб зрозуміти відношення проблеми NTRU до криптографії на решітках, розглянемо матриці

$$\begin{bmatrix} rot(1) & rot(h) \\ rot(0) & rot(q) \end{bmatrix}, \begin{bmatrix} rot(f) & rot(g) \\ rot(F) & rot(G) \end{bmatrix}, \tag{7}$$

де $F, G \in R_q$ задовольняють рівнянню $fG - gF = q$ у полі R_q. Ці матриці задають базис тієї ж самої решітки. Перший базис має великі значення елементів векторів, другий базис має малі значення, тому його можливо використовувати для вирішення складних задач на решітці за поліноміальний час, на відміну від першого.

Одним з недоліків проблеми NTRU є те, що вона, на відміну від LWE та SIS, не має редукцій від найгіршого до середнього від складних проблем з теорії решіток [1]. Нещодавно була знайдена редукція від найгіршого до найгіршого від Module-uSVP до NTRU [35]. Це є слабкішим результатом, проте той факт, що за більш ніж двадцять років NTRU не змогли вирішити за поліноміальний час у загальному випадку є практичним свідченням складності проблеми NTRU.

Криптографічні проблеми в теорії решіток на цьому не обмежуються. Це окремий напрямок досліджень у теоретичній криптографії, який має складний ландшафт. Це є величезною перевагою криптографії на решітках, оскільки інші напрямки досліджень не мають такого різноманіття та динамічності розвитку.

З теоретичної точки зору, захист проблем Module-LWE та NTRU обумовлений тим, що невідомий шлях до узагальнення технік, що використовувалися для вирішення Ideal-SVP через модульну структуру останніх. Більш того, ці техніки ґрунтуються на циклотомічних полях, що означає, що вибір нециклотомічного для цих проблем зробить цей клас квантових атак нерелевантним.

4. Механізми інкапсуляції ключів

Існують різні підходи до побудови механізмів інкапсуляції ключей. Найбільш поширеним є наступний алгоритм [1]:

- Обирається деяка складна проблема.
- Будується CPA-безпечна схема асиметричного шифрування на основі обраної проблеми.
- Застосовується перетворення, що доказово робить з CPA-безпечної схеми CCA-безпечний механізм інкапсуляції ключів.

CRYSTALS-Kyber, FrodoKEM та ДСТУ 8961:2019 “Скеля” побудовані саме за цим принципом. В табл. 2 наведені складні проблеми, на яких ґрунтуються ці схеми.

Таблиця 2

Порівняння складних проблем у перспективних КЕМ

	CRYSTALS-Kyber	FrodoKEM	ДСТУ 8961:2019 “Скеля”
Проблема	Module-LWE	LWE	NTRU
Поле	$\mathbf{Z}_q[X]/(X^n+1)$	-	$\mathbf{Z}_q[X]/(X^n - X - 1)$

ДСТУ 8961:2019 “Скеля” має доволі незвичайний вибір поля. Річ у тому, що на проблему NTRU в полі $\mathbf{Z}_q[X]/(X^n+1)$ існує серія атак, яка дозволяє вирішити проблему NTRU за поліноміальний час, використовуючи структуру підкілець [28]. Зауважимо, що зазначені атаки можливі тільки у випадку $q > O(n^3)$, який не є релевантним для ДСТУ 8961:2019 “Скеля”. Такі ситуації зустрічаються здебільшого тільки у схемах гомоморфного шифрування та більш екзотичних конструкціях. Поле $\mathbf{Z}_q[X]/(X^n - X - 1)$, яке використовується у ДСТУ 8961:2019, не має нетривіальних підкілець, що робить подібні атаки неможливими, навіть у випадку, якщо будуть знайдені розширення цих атак на випадок тих параметрів, що використовуються у ДСТУ 8961:2019. Додатково, вибір цього поля захищає від потенційного узагальнення технік, що використовувалися для вирішення Ideal-SVP за поліноміальний час. Можна сказати, що вибір поля $\mathbf{Z}_q[X]/(X^n - X - 1)$ є додатковим аргументом безпеки для випадку алгебраїчних та квантових атак.

FrodoKEM ґрунтується на проблемі LWE, що дає гарні докази безпеки, оскільки існують редукції до складних проблем в теорії решіток. Проте, з іншої сторони, це також означає наявність множення великих матриць у реалізації, що робить FrodoKEM повільним.

CRYSTALS-Kyber грунтується на проблемі Module-LWE. При цьому використовується поле $\mathbf{Z}_q[X]/(X^n+1)$. Такий вибір поля є стандартним для криптографії на решітках, оскільки воно дозволяє використовувати теоретико-числове перетворення для множення/ділення поліномів [2] і має гарні теоретичні властивості [22] у тому сенсі, що докази безпеки значно спрощуються. Багато доказів безпеки (здебільшого редукцій до складних проблем на модульних решітках) використовують властивості цього поля [16, 19, 22, 24]. CRYSTALS-Kyber має не такі сильні докази, як FrodoKEM, оскільки складність Module-* проблем, від яких Module-LWE має редукції, невідома. Проте, CRYSTALS-Kyber працює швидше, ніж FrodoKEM, що робить його гарним вибором для багатьох практичних застосувань, для яких FrodoKEM є відносно повільним.

Для того щоб з CPA безпечних схем асиметричного шифрування отримати CCA безпечні KEM, використовуються різновиди перетворення Фуджісакі – Окамото. У табл. 3 зібрана інформація щодо перетворень, що застосовуються у CRYSTALS-Kyber, FrodoKEM та ДСТУ 8961:2019 “Скеля”.

Таблиця 3

Порівняння CCA перетворень у перспективних KEM

	CRYSTALS-Kyber	FrodoKEM	ДСТУ 8961:2019 “Скеля”
Перетворення	Власна модифікація перетворення Фуджісакі – Окамото	Фуджісакі – Окамото з прямим відхиленням	Фуджісакі – Окамото з відхиленням
Докази у ROM	+ (сильний доказ)	+ (сильний доказ)	+/-
Докази у QROM	+	+	+/-

Усі KEM, що розглядаються, мають певні докази у QROM та ROM [2, 3, 4, 27], які можна віднести до них прямо чи непрямо, проте для ДСТУ 8961:2019 ситуація є дещо невизначеною у цьому сенсі. Публікацій, що аналізують саме ДСТУ 8961:2019 у моделях ROM чи QROM, не існує. З іншої сторони, ДСТУ 8961:2019 є модифікацією стандарта ANSI X9.98 [29] і відрізняється тільки іншим полем та алгоритмом формування малих поліномів [30]. У цілому, усі докази у QROM або ROM для ANSI X9.98, що не використовують структуру поля, є релевантними і для ДСТУ 8961:2019. Проте, комплексного аналізу стандарту x9.98 також є доволі мало. Особливо це стосується випадку QROM. Хоча окремі результати відомі, проте для ДСТУ 8961:2019 не вистачає комплексного вивчення у моделях ROM та QROM.

5. Атаки на перспективні КЕМ

Оскільки ДСТУ 8961:2019 ґрунтується на проблемі NTRU, то питання криптоаналізу так чи інакше пов'язане з редукцією відкритого базису:

$$\begin{bmatrix} rot(1) & rot(h) \\ rot(0) & rot(q) \end{bmatrix}. \tag{8}$$

Найкращою відомою атакою для ДСТУ 8961:2019 є гібридна атака [31], яка використовує той факт, що малі поліноми у ДСТУ 8961:2019 мають коефіцієнти з множини {0,1,-1}. Ідея полягає у тому, щоб проводити редукцію не усієї решітки, а лише певної підрешітки з r векторів, потім інші вектори знаходити комбінаторним перебором. Час атаки буде мінімізуватися, якщо час роботи редукції решітки та комбінаторного пошуку буде приблизно однаковим, що можливо зробити, якщо підібрати значення параметра r.

В [32] з використанням евристичних міркувань отримано формулу для трудомісткості комбінаторного етапу описаної атаки:

$$T_2(\delta, r) = \frac{2^{15} r!}{c_{-1}! c_1! (r - c_{-1} - c_1)!} \left(\binom{2c_{-1}}{c_{-1}} \binom{2c_1}{c_1} p^{|S|} \right)^{-1/2} \frac{1}{\tilde{p}}, \tag{9}$$

де

$$p = \prod_{i=1}^{2n-r+1} \left(1 - \frac{1}{r_i B\left(\frac{2n-r}{2}, \frac{1}{2}\right)} \int_{-r_i}^{-r_i} \int_{\max\{-1, z-r_i\}}^{z+r_i} (1-t^2)^{\frac{2n-r-2}{2}} \, dt \, dz \right), \tag{10}$$

$$|S| = 2 + 2(n-t-1) p_S, \tag{11}$$

$$p_S = \frac{p_{\text{NP}} \, 2^{-4c_1} r!}{(2c_{-1})! (2c_1)! (r - 2c_{-1} - 2c_1)!} \binom{n-r}{4t-4c_1} \binom{n}{2t}^{-1}, \tag{12}$$

$$p_{\text{NP}} = \prod_{i=1}^{2n-r+1} \left(1 - \frac{2}{r_i B\left(\frac{2n-r}{2}, \frac{1}{2}\right)} \int_{-1}^{\max\{-r_i, -1\}} (1-t^2)^{\frac{2n-r-2}{2}} \, dt \right), \tag{13}$$

$$\tilde{p} = 1 - (1 - p_S)^{n-t}. \tag{14}$$

У формулах (10), (13) $B(\cdot, \cdot)$ позначає бета-функцію Ойлера, а числа r_i визначаються за формулами

$$r_i = \frac{R_i(\delta)}{2l}, \; i \in \overline{1, 2n-r+1}, \tag{15}$$

де

$$R_i(\delta) = q, \text{ якщо } 1 \le i \le 2n - r + 1 - \mu;$$

$$R_i(\delta) = q^{-2(i-(2n-r+1-\mu)-1)+\mu} q^{\frac{\mu-(n-r)}{\mu}}, \text{ якщо } 2n - r + 1 - \mu < i \le 2n - r + 1,$$

$$\mu = \min\left\{2n - r + 1, \left\lfloor \sqrt{\frac{n-r}{\log_q \delta}} \right\rfloor\right\},\ \delta > 1.$$

Оцінка часу етапу редукції решітки буде розглянута в наступних розділах.

Для FrodoKEM найкращою відомою атакою є безпосередня редукція LWE решітки. LWE задається кортежом $(A,c) = (A, A*s+e) \in \mathbf{Z}_q^{m\times n} \times \mathbf{Z}_q^{m\times 1}$, з яким можна асоціювати решітку $L = \{Ax \bmod q : x \in \mathbf{Z}_q^m\}$. З означення решітки випливає, що вектор s належить решітці та є найближчим до вектора $As+e$ за умови, якщо вектор помилки має достатньо малі значення. Поширеним методом пошуку рішення задачі є алгоритм Бабаї [19]. Для вирішення проблеми алгоритм потребує поліноміальну кількість кроків, проте від того, наскільки якісно редукований базис решітки, залежить ймовірність знаходження правильного рішення. У випадку проблеми LWE відома наступна оцінка ймовірності [33]:

$$\prod_{i=0}^{m-1} erf\left(\frac{\|b_i^*\|\sqrt{\pi}}{2\alpha q}\right), \tag{16}$$

де $\|b_i^*\|$ є $l2$ нормою i-го базисного вектора решітки після процедури ортогоналізації за Граммом – Шмідтом; α, q є параметрами LWE, $erf(\cdot)$ – функція помилок.

Для максимізації ймовірності коректної роботи алгоритму необхідно мінімізувати значення $\|b_i^*\|$, що фактично означає редукцію базису решітки.

Атаку можна модифікувати. Замість пошуку вектора s на оригінальній решітці можна побудувати таку решітку, яка буде містити вектор $(s,e,1)$, і він буде найменшим унікальним вектором (задача uSVP) згідно з [2, 33]. Можливим варіантом такої решітки є

$$\Lambda = \{x \in \mathbf{Z}^{m+n+1} : (A \mid I_m \mid -b) * x = 0 \bmod q\}. \tag{17}$$

Аналогічно до попередньої атаки можна застосувати алгоритм редукції решітки. Для оцінки фактора Ерміта можливо скористатися співвідношенням

$$\log \delta_0 = \frac{1}{4n^2 \ln^2 q}\left(W\left((-2n\ln q) * (\sqrt{n \log q}) * \frac{(\tau\alpha)^2}{2\pi}\right)\right)^2, \tag{18}$$

де τ – допустима ймовірність, з якою алгоритм повинен завершити роботу успішно; α, n, q – параметри розподілу LWE; $W(\bullet)$ є W-функція Ламберта, яка визначається як розв'язок функціонального рівняння

$$z = W(z)e^{W(z)}\ . \tag{19}$$

Варто зауважити, що ця функція не може бути представлена через елементарні функції.

Іншим вектором атаки є використання дуальної решітки. Побудуємо дуальну решітку як $\hat{L} = \{x \in \mathbf{Z}_q^m \mid A*x = 0 \bmod q\}$. Редукція такої решітки фактично є вирішенням проблеми SIS. Якщо відомий вектор, що задовольняє умовам проблеми SIS, то задачу Decision-LWE можливо легко вирішити. Розглянемо детальніше редукцію Decision-LWE до SIS. Нехай задано m кортежів векторів вигляду $(A,c) = (A, A*s+e) \in \mathbf{Z}_q^{m\times n} \times \mathbf{Z}_q^{m\times 1}$. Знайдемо скалярний добуток $\langle x, c\rangle$:

$$\langle x, c\rangle = x*a*s + x*e = 0*s + x*e = x*e = \langle x, e\rangle. \tag{20}$$

Так як значення вектора $x \in \mathbf{Z}^n$ є заданим, то значення вектора помилок знаходиться перебором всіх можливих варіантів, оскільки простір пошуку значно зменшується. В роботі [33] показано, що норма вектора x є не більшою за

$$\|x\|_2 = \frac{1}{\alpha} * \sqrt{\frac{\ln(\frac{1}{\varepsilon})}{\pi}}, \tag{21}$$

де ε – допустима ймовірність, з якою алгоритм повинен завершити роботу успішно; α – параметр розподілу SIS.

Можливо з як завгодно близькою ймовірністю до 1 знайти значення вектору помилок, що фактично означає вирішення проблеми навчання з помилками. При цьому знадобиться $\frac{1}{\varepsilon^2}$ запусків вирішувача проблеми SIS. Так як для вирішення задачі потрібно знайти достатньо малий вектор на решітці, то рішення зводиться до задачі SVP. У роботі [33] була надана оцінка необхідного значення фактора Ерміта δ_0 при редукції решітки:

$$\log \delta_0 = \frac{\log^2(\frac{1}{\alpha}\sqrt{\frac{\ln(\frac{1}{\varepsilon})}{\pi}})}{4 * n \log q}, \tag{22}$$

де ε – допустима ймовірність, з якою алгоритм повинен завершити роботу успішно; α, n, q – параметри розподілу SIS.

Сімейство атак з використанням дуальних решіток має назву Dual Attack. При знаходженні кількісних оцінок атаки потрібно враховувати, що існують різні методи вирішення проблеми SIS. Класичним підходом є використання редукції решіток.

Щодо проблеми MLWE – наразі невідомо атак, що використовували би її алгебраїчну структуру, то ж можливо розглядати MLWE як частковий випадок LWE. Багато останніх робіт пропонують нові квантові алгоритми на Ideal-SVP [20, 21], тобто вирішення проблеми знаходження найменшого вектора на ідеальних решітках. Робота [21] присвячена квантовій атаці на Ring-LWE з використанням нової техніки, але використання її до Module-LWE створює численні труднощі. У роботі [24] показана редукція від MLWE до RLWE, тобто на основі поліноміального алгоритму, що може вирішити RLWE з правильними параметрами, можливо побудувати поліноміальний алгоритм, що може вирішити MLWE. Проте, на практиці ця атака веде себе значно гірше, ніж відомі атаки, оскільки зростає розмірність решітки. Це означає, що зростанням розмірності модуля можливо довести безпеку до значень, на які важко реалізувати атаку. Якщо через цю редукцію атакувати CRYSTALS-Kyber, то це призведе до вирішення RLWE з дуже великим модулем та помилками $(q' = q^3, \zeta' > q^2\zeta)$, тому вимагатиме від криптоаналітика більше, ніж 1 семпл.

6. Оцінка часу редукції решіток

У попередньому розділі було показано, що найкращі атаки на механізми KEM зводяться до редукції базису решіток без використання їх алгебраїчної структури. Найкращими відомими алгоритмами редукції решіток є BKZ та його численні модифікації [34]. BKZ є рандомізованим варіантом LLL. На кожному кроці знаходиться найменший вектор на проєктивній решітці розмірності β (розмір блоку редукції), додається до базису і проводиться LLL редукція. У роботі [34] було показано, що BKZ робить

$$O\left(\frac{n^2}{\beta^2}\log n\right) \tag{23}$$

викликів до процедури пошуку малого вектора на решітці розміру β. Від значення параметра β залежить якість редукції базису. Вектор з необхідною нормою можливо знайти за умови [2, 4]:

$$\sigma\sqrt{\beta} \le \delta_0^{2\beta-d-1} q^{\frac{n}{d}}, \tag{24}$$

де σ – середньоквадратичне відхилення для розподілу, з якого отримано коефіцієнти секретного вектора; d – розмірність усієї решітки; δ_0 – максимальне значення фактора Ерміта, за якого решітка стає достатньо редукованою.

Значення δ_0 також залежить від β і виражається [2, 19] як

$$\delta_0 \approx ((\pi B)^{\frac{1}{B}} * \beta / 2\pi e)^{1/2(\beta-1)}. \tag{25}$$

Ці міркування працюють як для LWE, та і для NTRU. Час роботи процедури пошуку малого вектора на решітці розміру β залежить від алгоритму, що використовується. Існує два класичних підходи – комбінаторні методи та методи на основі решета [19]. Комбінаторні методи роблять повний перебір, зменшуючи простір пошуку за допомогою різних евристик та використання інформації про базис решітки. Вони мають час роботи $O(2^\beta \log 2^\beta)$. Методи на основі решета генерують експоненційно велику множину векторів на решітці і на кожній ітерації "просіюють" її, доки не буде отримано достатньо малого вектора. Кількість таких ітерацій є поліноміальною. Такі алгоритми працюють за $O(2^\beta)$. Найкращий відомий класичний алгоритм працює за $2^{0.292\beta}$, а найкращий квантовий алгоритм – за $2^{0.265\beta}$.

На практиці для оцінки безпеки криптографічних систем на решітках, як правило, використовується модель core-SVP, яка полягає у тому, що вартість редукції визначається часом роботи алгоритма пошуку малого вектора. Поліноміальна кількість викликів ігнорується. Модель є доволі старою, проте вона є загальноприйнятим стандартом в криптографії на решітках. Сьогодні у науковій спільноті тривають дискусії про створення інших моделей, що прив'язані до певних технічних чи фізичних характеристик, проте, ці моделі доволі сирі. У табл. 4 зведені дані про конкретні оцінки безпеки КЕМ. На рис. 2 наведено порівняння конкретних оцінок безпеки КЕМ у вигляді діаграми.

Таблиця 4

Конкретні оцінки безпеки перспективних КЕМ

Рівень безпеки	FrodoKEM	CRYSTALS-Kyber	ДСТУ 8961:2019 "Скеля"
128	145/132	118/107	-
192	210/191	183/166	-
256	275/250	255/232	265/248
384	-	-	424/408
512	-	-	562/534
Найкраща атака	Редукція решітки	Редукція решітки	Гібридна атака

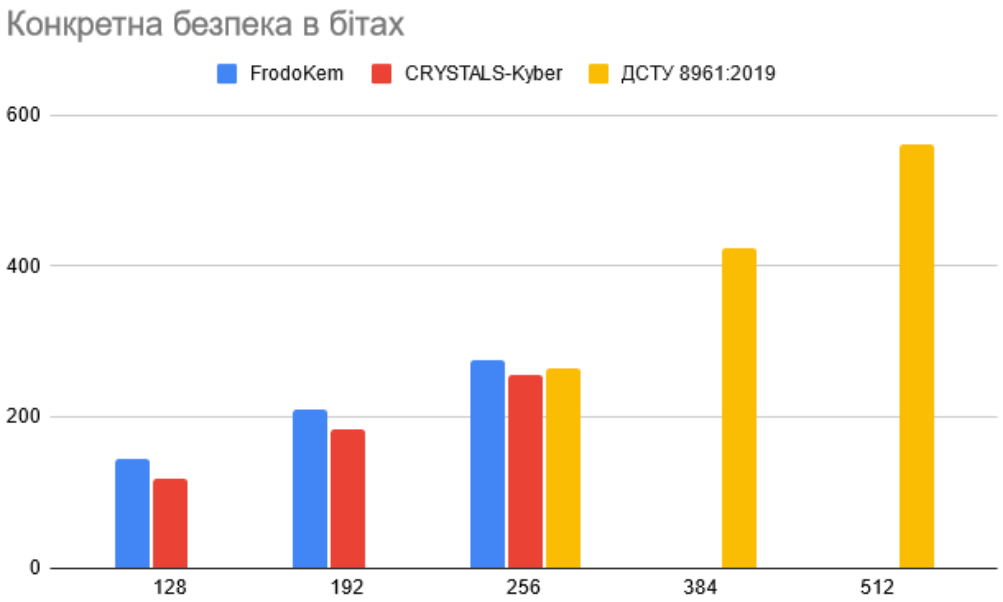


Рис. 2. Порівняння конкретної безпеки в бітах для класичних атак

Висновки

1. ДСТУ 8961:2019 “Скеля” ґрунтується на проблемі NTRU, яка має довгу історію вивчення, що є практичним аргументом безпеки. Структура, яку мають NTRU решітки, добре відома, проте редукцій від найгіршого до середнього від складних проблем теорії решіток (GapSVP, SIVP, тощо) до варіанту NTRU, що використовується у ДСТУ 8961:2019 “Скеля”, невідомо, хоча нещодавно була знайдена редукція від найгіршого до найгіршого для Module-uSVP. То ж, можливо сказати, що у той час, як практичний аспект безпеки NTRU є доволі вивченим, теоретичні питання досі потребують досліджень. Використання поля з твірним поліномом $X^n - X - 1$ у ДСТУ 8961:2019 “Скеля” є нетиповим вибором для криптосистем на решітках. З однієї сторони це дає захист від ряду алгебраїчних та деяких потенційних квантових атак, оскільки не існує нетривіальних підполів, проте, з іншої сторони, властивості цього полінома не так добре вивчені, як $X^n + 1$, який є стандартним вибором у криптографії на решітках. Хоча поява алгебраїчних атак на поле з твірним поліномом $X^n - X - 1$ є малоймовірною подією, проте це питання також потребує додаткових досліджень.

2. FrodoKEM ґрунтується на проблемі LWE. Широкі дослідження цієї проблеми почалися не так давно, як NTRU, проте вона має редукції від найгіршого до середнього від складних проблем теорії решіток, що є сильним теоретичним аргументом, якого не мають більшість розділів криптографії. Серед розглянутих варіантів FrodoKEM має найкращі теоретичні докази безпеки, проте через необхідність виконання операцій з матрицями реалізація FrodoKEM є помітно повільнішою за інші схеми, що фактично стало причиною виключення її з конкурсу NIST PQC після третього етапу.

3. CRYSTALS-Kyber ґрунтується на проблемі Module-LWE, яка є структурованим варіантом LWE. Ця структурованість дозволяє перейти до операцій в полі поліномів, що робить її надзвичайно швидкою. Module-LWE має редукції, аналогічно до LWE, від складних проблем на Module-LWE решітках, проте складність цих проблем є під питанням. Відомо, коли для складних проблем на структурованих решітках знаходили поліноміальні алгоритми вирішення, що використовують цю структурованість. Наразі для Module-LWE невідомо атак, які б ефективно використовували структурованість, проте це питання потребує детальнішого вивчення.

4. Для FrodoKEM та CRYSTALS-Kyber існують докази у моделях ROM та QROM, що є значною перевагою. Для ДСТУ 8961:2019 “Скеля” ситуація є дещо складнішою. Досліджень безпеки ДСТУ 8961:2019 у моделях ROM та QROM, на жаль, немає. Проте, зважаючи, що ДСТУ 8961:2019 є модифікованою версією стандарта ANSI X9.98, який є версією криптосистеми NTRUEncrypt, то відомі для цієї криптосистеми результати можуть бути адаптовані для ДСТУ 8961:2019. Питання доказів безпеки ДСТУ 8961:2019 “Скеля” у моделях ROM та QROM потребує додаткових досліджень.

5. ДСТУ 8961:2019 “Скеля”, на відміну від інших KEM, має параметри безпеки від 256 до 512 біт. З однієї сторони це робить її безпечнішою у деякому сенсі, при появі нових атак, що знижують безпеку на експоненційний фактор, проте, з іншої сторони, це означає, що реалізація буде дещо повільнішою, ніж FrodoKEM та CRYSTALS-Kyber (можливо, окрім набору параметрів для 256 біт безпеки), що робить її застосування доречним переважно для тих випадків, коли стійкість до появи нових малоймовірних атак важливіша за швидкодію.

6. Криптографія на решітках має величезне різноманіття складних проблем. Разом з редукціями це дає дуже різноманітний ландшафт для вивчення. Це є унікальною властивістю криптографії на решітках, що непритаманна іншим напрямкам досліджень, як у доквантовій, так і у постквантовій криптографії. При проєктуванні криптографічних систем це, безсумнівно, має враховуватись. Розглянуті KEM є рішеннями, які можливо впровадити “тут і зараз”. Проте, у майбутньому вірогідна поява інших криптографічних систем на решітках, які зможуть перевершити показники існуючих. Необхідне вивчення ландшафту проблем теорії решіток та застосування структурованості решіток.

Список літератури:

1. NIST Post-Quantum Cryptography Standartization Project : веб сайт. URL: https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization
2. CRYSTALS Kyber: a CCA-secure module-lattice-based KEM / Leo Ducas., and other. // URL: https://eprint.iacr.org/2017/634.pdf
3. ДСТУ 8961:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритми асиметричного шифрування та інкапсуляції ключів // URL: http://online.budstandart.com/ua/catalog/doc-page.html?id_doc=88056
4. FrodoKEM. Learning With Errors Key Encapsulation Algorithm Specifications. And Supporting Documentation. / Leo Ducas., and other // URL: https://frodokem.org/files/FrodoKEM-specification-20210604.pdf
5. M. Toorani Security Protocols in a Nutshell // URL: https://arxiv.org/abs/1605.09771
6. Katz, Jonathan; Lindell, Yehuda (2007). Introduction to Modern Cryptography: Principles and Protocols.
7. Chakraborty, Debrup; Rodríguez-Henríquez., Francisco (2008). Çetin Kaya Koç (ed.). Cryptographic Engineering. p. 340. ISBN 9780387718170.
8. Möller, Bodo (2004). A Public-Key Encryption Scheme with Pseudo-random Ciphertexts // Computer Security – ESORICS 2004. Lecture Notes in Computer Science. Vol. 3193. pp. 335–351.
9. M. Bellare, A. Desai, D. Pointcheval, and P. Rogaway. Relations among notions of security for public-key encryption schemes. // Advances in Cryptology – CRYPTO’98, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1998, vol. 1462, pp. 26–45. [Online]. Available: http://dx.doi.org/10.1007/BFb0055718
10. Canetti R., Goldreich O., Halevi S. The random oracle methodology, revisited // 30th symposium on theory of computing. STOC, 1998. P. 209–218.
11. Boneh D., Dagdelen O., Fischlin M., Lehmann A., Schaffner C., Zhandry M (2011). Random oracles in a quantum world. Advances in Cryptology – ASIACRYPT 2011, eds Lee DH, Wang X (Springer Berlin Heidelberg, Berlin, Heidelberg), pp 41–69.
12. P. W. Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer // URL: https://arxiv.org/abs/quant-ph/9508027
13. Горбенко І. Д., Горбенко Ю. І. Прикладна криптологія. Теорія. Практика. Застосування : монографія. Харків : Форт, 2012. 880 с.
14. Goldreich O. Foundations of Cryptography : Vol. 1. Cambridge University Press, 2000. 392 p.
15. O. Regev. On lattices, learning with errors, random linear codes, and cryptography// ACM, 56(6):1–40, 2009. Preliminary version in STOC 2005.
16. Classical Hardness of Learning with Errors / Chris Peikert, Oded Regev, and other // URL: http://perso.ens-lyon.fr/damien.stehle/downloads/LWE.pdf
17. Thijs Laarhoven, J. V. D. Pol, B. D. Weger Solving Hard Lattice Problems and the Security of Lattice-Based Cryptosystems // URL: http://deweger.xs4all.nl/papers/%5B51%5DLvdPdW-Kolkata%5B2012%5D.pdf

18. Обзор LWE на структурированых решетках.
19. Vaikuntanathan V. Advanced Topics in Cryptography: Lattices : веб сайт. URL: https://people.csail.mit.edu/vinodv/6876-Fall2015/
20. Kirsten Eisenträger, Sean Hallgren, Alexei Kitaev, and Fang Song. A quantum algorithm for computing the unit group of an arbitrary degree number field. In STOC '14 Proceedings of the forty-sixth annual ACM symposium on Theory of computing, pages 293–302. ACM, 2014. http://www.personal.psu. edu/kxe8/unitgroup.pdf. 31
21. Ronald Cramer, Léo Ducas, and Benjamin Wesolowski. Short Stickelberger class relations and application to Ideal-SVP. In Jean-Sébastien Coron and Jesper Buus Nielsen, editors, Advances in Cryptology – EUROCRYPT 2017, volume 10210 of LNCS, pages 324–348. Springer, 2017. https: //eprint.iacr.org/2016/885. 31, 35
22. Yang Wang, Mingqiang Wang Module-LWE versus Ring-LWE, Revisited // URL: https://eprint.iacr.org/2019/930
23. A. Langlois, D. Stehlé. Worst-Case to Average-Case Reductions for Module Lattices // URL: https://perso.ens-lyon.fr/damien.stehle/downloads/MSIS.pdf
24. Martin R. Albrecht, A. Deo. Large Modulus Ring-LWE ≥ Module-LWE // URL: https://eprint.iacr.org/2017/612.pdf
25. Joan Bruna, Oded Regev, Min Jae Song, Yi Tang. Continuous LWE // URL: https://arxiv.org/abs/2005.09595
26. Hoffstein J., Pipher J., Silverman J.H. NTRU: a ring based public key cryptosystem // Algorithmic Nuber Theory, Third International Symposium, Portland, Oregon, USA, June 21 – 25, 1998. Proceedings. Springer, 1998. P. 267 – 288.
27. Proovable NTRU.
28. Micheli G., Heninger N., Shani B. Characterizing overstretched NTRU attacks Journal of Mathematical Cryptology. 2020. Vol 14, Is 1. P. 110-119.
29. American National Standard X9.98-2010. Lattice-based polynomial public key encryption algorithm, Part 1: key establishment, Part 2: data encryption. 2010.
30. I.D. Gorbenko. Calculation of general parameters for NTRU Prime Ukraine of 6-7 levels of stability / I. D. Gorbenko, A. N. Alekseychuk, O. Kachko ,M. Yesina, I. V. Stelnik, S. Kandy,V. A. Bobukh // Telecommunications and Radio Engineering. 78(4):327-340.
31. I.D. Gorbenko. Methods of building general parameters and keys for ntru prime ukraine of 5th-7th levels of stability. product form / I. D. Gorbenko, Yu. I. Gorbenko, O. Kachko ,M. Yesina, I. V. Stelnik, S. Kandy // Telecommunications and Radio Engineering. 78(7):579-594.
32. Wunderer Th. Revising the hibrid attack: improved analysis and refined security estimates // URL: http://eprint.iacr.org/2016/733.
33. Player R. Parameter selectionin lattice-based cryptography. URL: https://pure.royalholloway.ac.uk/portal/files/29983580/2018playerrphd.pdf
34. Jianwei Li, Phong Q. Nguyen. A Complete Analysis of the BKZ Lattice Reduction Algorithm // URL: https://eprint.iacr.org/2020/1237.pdf
35. J.Felderhoff, A. Pellet-Mary, D.Stehl. On Module Unique-SVP and NTRU // URL: https://eprint.iacr.org/2022/1203.pdf.

Надійшла до редколегії 12.09.2022

Відомості про авторів:

Горбенко Юрій Іванович – канд. техн. наук, АТ "Інститут Інформаційних Технологій", перший заступник головного конструктора, Україна; e-mail: gorbenkou@iit.kharkov.ua; ORCID: https://orcid.org/0000-0003-0073-9107

Кандій Сергій Олегович – Харківський національний університет імені В. Н. Каразіна; аспірант кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ "Інститут Інформаційних Технологій", технік-конструктор; Україна; e-mail: sergeykandy@gmail.com , ORCID: https://orcid.org/0000-0003-0552-8341

УДК 004.056.5 **DOI:10.30837/rt.2022.3.210.03**

Я.А. ДЕРЕВ'ЯНКО, І.Д. ГОРБЕНКО, д-р техн. наук

ВРАЗЛИВІСТЬ ЕП FALCON ДО СПЕЦІАЛЬНИХ АТАК ТА ЙОГО ЗАХИЩЕНІСТЬ

Вступ

Прогрес обчислювальних технологій суттєво впливає на злам існуючих стандартизованих асиметричних криптографічних перетворень. Подібно до того, як Bombe, перший електромеханічний комп'ютер зламав сумнозвісний шифр Enigma, перший практичний квантовий комп'ютер може зламати сучасні схеми асиметричного шифрування. Дійсно, добре відомо, що квантові алгоритми пропонують експоненціальне прискорення при розв'язанні завдань цілочисельної факторизації [1] та (еліптичної кривої) дискретного логарифму [2], на які покладаються існуючі системи з відкритим ключем. Для їх захисту розробляються та стандартизуються альтернативні рішення – постквантові стандарти асиметричного шифрування, протоколи інкапсуляції ключів та електронні підписи. Вони зі значною ймовірністю можуть протистояти квантовому криптоаналізу. Зростаюче занепокоєння квантовою загрозою спонукало Національний інститут стандартів і технологій (NIST) запросити та оцінити заявки на стандарт постквантової криптографії, який є постійним процесом, який планується завершити до 2023 року.

Одним із кандидатів на міжнародний стандарт електронного підпису є Falcon – алгоритм електронного підпису, що заснований на математиці алгебраїчних решіток. Недоліком проєкту стандарту електронного підпису є мала кількість досліджень стійкості проти спеціальних атак, а також атак побічними каналами [3 – 7].

У статті розглядаються та аналізуються існуючі атаки на реалізацію Falcon, а також оцінюється швидкодія при застосуванні контрзаходів, які б перешкоджали таким атакам. Незважаючи на те, що схема Falcon, а також певні математичні перетворення, все ж є вразливими до атак [8] (що в свою чергу дозволяє отримати приватний ключ), ефективність компонентів та математики даного алгоритму електронного підпису сприяє тому, що він здатен конкурувати з іншими схемами, навіть з контрзаходами проти спеціальних атак. Виявлено, що для схеми Falcon контрзаходи мають, в середньому, лише до 5 % впливу на ефективність.

1. Часові атаки на схему Falcon

Певні компоненти схеми ЕП Falcon мають вразливість до атак, що базуються на аналізі часових показників. Це пояснюється складнощами в реалізації даних компонентів постійними у часі. В цьому пункті коротко описуються вразливі місця та пропонуються контрзаходи, які здатні перешкоджати часовим атакам на них.

1.1. Вразливості алгоритму Falcon

Відбірник Гауса [9, 10]

Відбірники Гауса – одне зі слабких місць, яке є схильним до часових атак на схеми, які використовують в якості основи алгебраїчні решітки. Відбірник Гауса у схемі Falcon застосовується для відбору коротких векторів у решітках.

Теоретичне перетворення числа (NTT) [11]

Через використання значної кількості модульної арифметики NTT є потенційно вразливим до часових атак, оскільки така арифметика є складною для реалізації за постійний час. Застосування NTT для схеми Falcon необхідне з метою пришвидшення множення кільцевих многочленів [12].

HashtoPoint

Функція гешування потенційно може бути вразливою до атак на аналізі часових показників, оскільки процес гешування може бути складним для реалізації постійним у часі.

1.2. Контрзаходи проти часових атак

Одним з ефективних на даний момент заходів проти часових атак є алгоритм BlindVector, запроваджений Saarinen у 2017 році. Він є розширеним та вдосконаленим алгоритмом перемішування Фішера – Йейтса, який, як правило, використовується для ефективної випадкової перестановки коефіцієнтів вектору, розподіленого за гаусовим розподілом. BlindVector отримав покращення випадкових перемішувань. Саме це, а також постійна за часом реалізація, і дає змогу протистояти атакам побічними каналами [13, 14]. Цикли алгоритму не залежать від даних, а операції завжди виконуються, незалежно від того, замінено якесь значення чи ні.

Серед контрзаходів також можна відмітити процес відкидання зразка. Основна особливість його роботи полягає у зчитуванні з випадкових адрес додаткового кешу. Це відбувається для спотворення статистики для SCA. Після цього зайві зчитування відкидаються. Діапазон норми відкидання обрано на рівні 6,25, 12,5 та 25 % [6].

До нововведень відноситься більш ефективна конструкція гаусової вибірки (CDT) з постійним часом. Її перевагами є чітка кількість операцій пошуку-зчитування, а також застосування однакової арифметики, яка не залежить від гілки, яку було обрано згідно з CDT алгоритмом [6]. Максимальна кількість необхідних пошукових операцій оцінюється як $\lceil \log_2 N \rceil$, тому кожен виклик пробовідбірника доповнюється до найближчого ступеня двійки для того, щоб займати стільки ж тактових циклів. Побудований таким чином пробовідбірник здатен забезпечити кращу швидкодію, ніж пробовідбірник Knuth-Yao, дискретний відбірник Ziggurat'a та відбірник Бернуллі, за умови, що він виконується за усталений час.

Функції NTT та FFT отримали постійну за часом реалізацію, що також надає можливість протидії часовим атакам. Це досягається шляхом обробки всіх необхідних змінних гілки та відмови від логіки для більш тривалих арифметичних операцій. Застосування векторизації SIMD та «ледачого» скорочення (lazy reduction) дозволяє значно зменшити вплив цих контрзаходів на продуктивність [6].

Можливість множення двох кілець у домені NTT [13] забезпечується шляхом застосування поточкового модульного множення. Завдяки цьому кожен елемент обчислюється незалежно. Про усталеність у часі даних операцій свідчить їх послідовність та безумовність. Важливою особливістю є також застосування вдосконаленого NTT алгоритму Cooley-Tukey. Вдосконалення стосуються покращення продуктивності та автоматичної векторизації. Також гарантується, що умовні операції розгалуження не будуть використовуватися [6]. Автоматична векторизація виконується окремо від модульного скорочення з метою покращення продуктивності. До того ж, для модульного скорочення застосовується спрощене обмеження діапазону, що дає змогу перейти від логічних до арифметичних операцій, які є усталеними за часом в більшості випадків [6].

2. Злам схеми постквантового підпису Falcon шляхом атаки побічними каналами

Хоча алгоритми можуть бути математично обґрунтованими проти класичного або квантового криптоаналізу, їх реалізація може призвести до витоку секретної інформації через побічні канали [3]. Ці атаки знаходять кореляцію між секретними значеннями та поведінкою реалізації, як-от час виконання, енергоспоживання та електромагнітне (ЕМ) випромінювання. Серед цих атак фізичні побічні канали (наприклад, витік ЕМ) мають особливе значення, оскільки вони існують не через неправильний вибір дизайну як такого, а через фізику залежної від даних активності CMOS. Ці атаки можуть бути успішними за допомогою всього лише кількох тестів, застосованих до фізичного пристрою, і без потреби в будь-якому функціональному квантовому комп'ютері. Атаки побічними каналами також важливі для NIST, оскільки вони є критерієм для визначення кінцевого стандарту [4].

У роботі Emre Karabulut та Aydin Aysu [15] пропонується перша атака побічними каналами на Falcon. Така атака є атакою з відомим відкритим текстом, яка використовує електромагнітні вимірювання пристрою для отримання секретних ключів підпису, які потім можна

використовувати для підробки підписів у довільних повідомленнях. Запропонована атака націлена на унікальне множення з плаваючою комою в рамках швидкого перетворення Фур'є алгоритму Falcon за допомогою нової стратегії розширення та скорочення, яка отримує змінні знака, мантиси та експоненти без помилкових спрацьовувань. Потім отримані значення з плаваючою комою відображаються назад у коефіцієнти секретного ключа. Подібна атака, зокрема, не вимагає попередньої характеристики профілю потужності цільового пристрою або створення спеціальних вхідних даних. Натомість статистичні відмінності на отриманих даних достатні для успішного виконання запропонованого диференційного електромагнітного аналізу. Результати на ARM-Cortex-M4, що працює з довідковим програмним забезпеченням FALCON NIST, показують, що приблизно 10 тисяч вимірювань достатньо, щоб отримати весь ключ.

У розділі:

• Пропонується перша атака побічними каналами на FALCON. Проводиться аналіз алгоритму, виявлення вразливих обчислень, через які може статися витік інформації, і цей витік може спричинити підробку підписів у довільних повідомленнях.

• Показано, що пряма атака на цільові обчислення зазнає невдачі через помилкові спрацьовування при множенні, і представляємо нову атаку, яка може вирішувати помилкові припущення за допомогою стратегії розширення та скорочення.

• Відбувається застосування запропонованої атаки на еталонне програмне забезпечення FALCON, взяте з веб-сайту NIST, і демонструємо, що запропонована атака може отримати цілі ключі підпису за допомогою кількох тисяч вимірювань, коли FALCON працює на мікроконтролері ARM-Cortex-M4.

2.1. Попередні відомості

У цьому розділі надається довідкова інформація про алгоритм цифрового підпису Falcon і модель загрози.

А. Модель загрози (зловмисника)

Робота дотримується стандартних припущень щодо атак побічним EM-каналом, коли зловмисник має фізичний доступ до пристрою та фіксує EM-вимірювання, поки виконуються обчислення, що залежать від ключових елементів. Дві помітні переваги нашої атаки в порівнянні з деякими нещодавніми роботами з побічних каналів решітчастої криптографії полягає в тому, що не потрібно створювати спеціальні вхідні дані або інше джерело вразливості, таке як побічний канал синхронізації для вилучення секретної інформації.

Б. Схема електронного підпису Falcon [7]

Falcon – це постквантова схема підпису, заснована на решітках [4]. Falcon складається з процедур генерації ключів, підписання та перевірки підпису. В роботі детально розглядаються перші дві процедури, оскільки вони мають вирішальне значення для розуміння атаки.

В наступному алгоритмі показано процес генерації.

Алгоритм 1: Keygen(ϕ, q): вхідні дані – одиничний поліном $\phi \in \mathbb{Z}[x]$, модуль q; вихідні дані – приватний ключ sk та відкритий ключ pk.

$f, g, F, G \leftarrow \mathsf{NTRUGen}(\phi, q)$
$\mathbf{B} \leftarrow \begin{bmatrix} g & -f \\ G & -F \end{bmatrix}$
$\hat{\mathbf{B}} \leftarrow \mathsf{FFT}(\mathbf{B})$
$\mathbf{G} \leftarrow \hat{\mathbf{B}} \times \hat{\mathbf{B}}^{\star}$
$\mathsf{T} \leftarrow \mathsf{ffLDL}^{\star}(\mathbf{G})$
for each leaf leaf of T do
    leaf.value $\leftarrow \sigma / \sqrt{\text{leaf.value}}$
$\mathsf{sk} \leftarrow (\hat{\mathbf{B}}, \mathsf{T})$
$h \leftarrow g f^{-1} \bmod q$
$\mathsf{pk} \leftarrow h$
return sk, pk

Рис. 1. Алгоритм Keygen

Рис. 1 показує процедуру генерації ключа, яка створює секретний ключ sk для створення та відкритий ключ h для перевірки підписів. Вхідними даними алгоритму є параметри ϕ і q: усі операції відбуваються над монічним поліномом ϕ ступеня n, який має вигляд x^n+1 для двійкового випадку та $x^n - x^{n/2}+1$ для трійкового випадку, тоді як q є простим модулем. Цей алгоритм спочатку випадковим чином відбирає коефіцієнти поліномів f і $g \in \mathbb{Z}[x]$ із дискретного розподілу Гауса, а потім обчислює F і $G \in \mathbb{Z}[x]$, які задовольняють рівнянню NTRU $fG - gF = q \bmod \varphi$. Багаточлени f, g, F і G називають приватними елементами. Потім ці поліноми об'єднуються, пропускаються через FFT і перетворюються на повнорангову матрицю Грама G. Щоб обчислити бінарне дерево T, Falcon застосовує LDL розклад на G. Алгоритм генерації ключа повертає відкритий ключ h, який задовольняє рівняння $gf^{-1} = h$ та секретний ключ sk містить два компоненти $\hat{B}$ і T, які походять від чотирьох поліномів $f, g, F, G \in \mathbb{Z}[x]$. Таким чином, стійкість генерації ключів базується на квантовостійкій проблемі NTRU, яка спирається на складність відновлення поліномів f і g, заданих поліноміальним кільцевим елементом h. Коефіцієнти цих приватних поліномів f і g мають діапазон від -127 до +127.

Якщо дано приватний ключ sk та повідомлення m, підписувач використовує sk для підпису m таким чином [4]:

- випадковий модифікатор входу геш-функції ($salt$) r генерується рівномірно в множині $\{0,1\}^{320}$. Після цього об'єднаний рядок ($r \| m$) гешується до точки $c \in \mathbb{Z}_q[x]/(\phi)$.
- обчислюється прообраз t для c і потім подається як вхідні дані для використання алгоритмом швидкої вибірки Фур'є. На виході отримаємо два коротких поліноми $s_1, s_2 \in \mathbb{Z}[x]/(\phi)$ (у FFT представленні) такі, що $s_1 + s_2 h = c \bmod q$.
- s_2 кодується (стискається) до рядку бітів s.
- підписом є пара (r, s).

Алгоритм 2: Sign($m, sk, \lfloor \beta^2 \rfloor$): вхідні дані – повідомлення m, приватний ключ sk, межа $\lfloor \beta^2 \rfloor$; вихідні дані – підпис sig для m (рис. 2).

1: $\mathbf{r} \leftarrow \{0,1\}^{320}$ uniformly
2: $c \leftarrow \mathsf{HashToPoint}(\mathbf{r} \| \mathbf{m}, q, n)$
3: $\mathbf{t} \leftarrow \left(-\frac{1}{q}\mathsf{FFT}(c) \odot \mathsf{FFT}(F), \frac{1}{q}\mathsf{FFT}(c) \odot \mathsf{FFT}(f)\right)$
4: do
5: do
6: $\mathbf{z} \leftarrow \mathsf{ffSampling}_n(\mathbf{t}, \mathbf{T})$
7: $\mathbf{s} = (\mathbf{t} - \mathbf{z})\hat{\mathbf{B}}$
8: while $\|\mathbf{s}\|^2 > \lfloor \beta^2 \rfloor$
9: $(s_1, s_2) \leftarrow \mathsf{invFFT}(\mathbf{s})$
10: $\mathbf{s} \leftarrow \mathsf{Compress}(s_2, 8 \cdot \mathsf{sbytelen} - 328)$
11: while $(\mathbf{s} = \perp)$
12: return $\mathsf{sig} = (\mathbf{r}, \mathbf{s})$

Рис. 2. Алгоритм Sign

B. FFT над числами з плаваючою крапкою

Falcon використовує односторонній відбірник секрету і тому повинен працювати з арифметикою з плаваючою комою та FFT, а не з цілочисельною арифметикою та NTT [4]. Falcon округлює (наближує) числа з плаваючою комою, які використовуються в арифметичних операціях, подібно до стандарту IEEE 754 з плаваючою комою (подвійна точність). Це наближення представляє число з плаваючою комою 64 бітами, де MSB є знаковим бітом, наступні

11 бітів є експонентою, а решта 52 біти є мантисою. Falcon вимагає арифметики з плаваючою комою під час підписання та етапів генерації ключів (алгоритми 1 і 2).

Falcon прискорює множення кільцевих поліномів за допомогою FFT, яке працює над кільцем $\mathbb{Z}_q / \varphi(x)$, де $\varphi(x)$ є монічним скорочувальним поліномом. FFT зменшує часову складність шляхом перетворення поліномів з $\mathbb{Z}_q / \varphi(x)$ в іншу область, де поліноміальне множення стає покоефіцієнтним (скалярним) множенням. Процедура підпису перетворює гешоване повідомлення та коефіцієнти елементів закритого ключа (f, g, F, G) у числа з плаваючою комою, а потім застосовує до них перетворення домену FFT (Алгоритм 2, рядок 3). Таким чином, алгоритм FFT перетворює 8-розрядні цілі коефіцієнти елементів закритого ключа на 64-розрядні коефіцієнти з плаваючою крапкою. Алгоритм FFT застосовує операції додавання, віднімання та множення з плаваючою крапкою між коефіцієнтами вхідного полінома. Після переходу до домену FFT алгоритм підписання Falcon виконує скалярне множення з плаваючою комою між елементами приватного ключа f і F і гешованим повідомленням c.

2.2. Запропонована атака побічними каналами

У цьому розділі представлено запропоновану атаку на алгоритм цифрового підпису Falcon і пов'язані з цим проблеми. По-перше, опишемо проміжні обчислення і чому такий спосіб дозволяє відновлювати секретні ключі та підробляти підписи. Потім ми покажемо проблеми виконання атаки побічними каналами та те, як ми вирішували ці проблеми.

A. Цільова операція $FFT(c) \odot FFT(f)$ та обґрунтування

Запропонована атака спрямована на множення з плаваючою комою в рамках обчислень $FFT(c) \odot FFT(f)$ (Алгоритм 2, рядок 3). Стверджується, що можлива атака за допомогою відомого відкритого тексту на ці обчислення та що захоплення коефіцієнтів $FFT(f)$ за допомогою атаки побічними каналами дозволяє зловмиснику підписувати довільні повідомлення.

Секретні елементи Falcon складаються з поліномів f, g, F і G. Ці поліноми використовуються для обчислення компонентів закритого ключа підпису $\hat{B}$ і T (Алгоритм 2). Поліноми F і G утворюють рівняння NTRU разом з f і g; отже, якщо зловмисник знає поліноми f і g, він може обчислити F і G, вивести весь секретний ключ і успішно підписати довільні повідомлення. Оскільки відкритий ключ h також є добутком gf^{-1}, зловмиснику необхідно отримати або поліном f, або g, щоб здійснити успішну атаку.

З цією метою атака має націлюватись на операцію $FFT(c) \odot FFT(f)$ (Алгоритм 2, крок 3), де відбувається поліноміальне множення на основі FFT між гешованим повідомленням c і приватним поліноміальним елементом f. Націлювання на це обчислення за допомогою атаки побічним каналом є здійсненним, оскільки c відоме зловмиснику, і, отже, секрет f можна припустити та перевірити через витік.

На рис. 3 показано деталі цільового множення на основі FFT між гешованим повідомленням c і приватним елементом f:

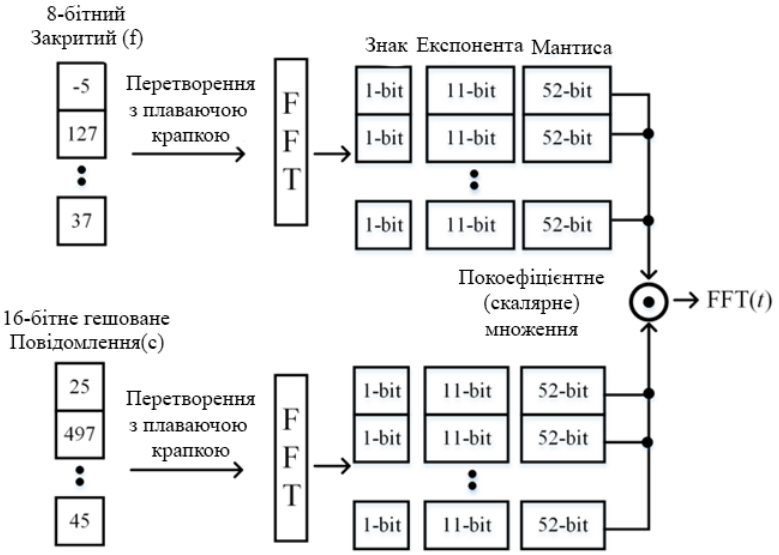


Рис. 3. Множення приватного елемента f і полінома гешованого повідомлення c.
Атака спрямована на скалярні множення з плаваючою крапкою в домені FFT

Еталонна реалізація Falcon спочатку перетворює цілі коефіцієнти поліномів c і f у число з плаваючою крапкою. Потім Falcon застосовує перетворення FFT над цими коефіцієнтами з плаваючою крапкою. Після перетворення FFT Falcon виконує скалярне множення з плаваючою крапкою між 64-бітними коефіцієнтами $FFT(c)$ та $FFT(f)$. Оскільки рівномірно випадкова сіль r і повідомлення m є загальнодоступними, $FFT(c)$ може бути обчислений зловмисником. Таким чином, запропонована атака зосереджена на множенні з плаваючою крапкою між коефіцієнтами відомого $FFT(c)$ і секретного $FFT(f)$. Якщо зловмисник успішно виділяє коефіцієнти полінома $FFT(f)$, він може відновити приватний елемент f, оскільки функція FFT Falcon є оборотною та однозначною.

Б. Отримання $FFT(f)$ шляхом підходу «Розділяй-і-володарюй» [15]

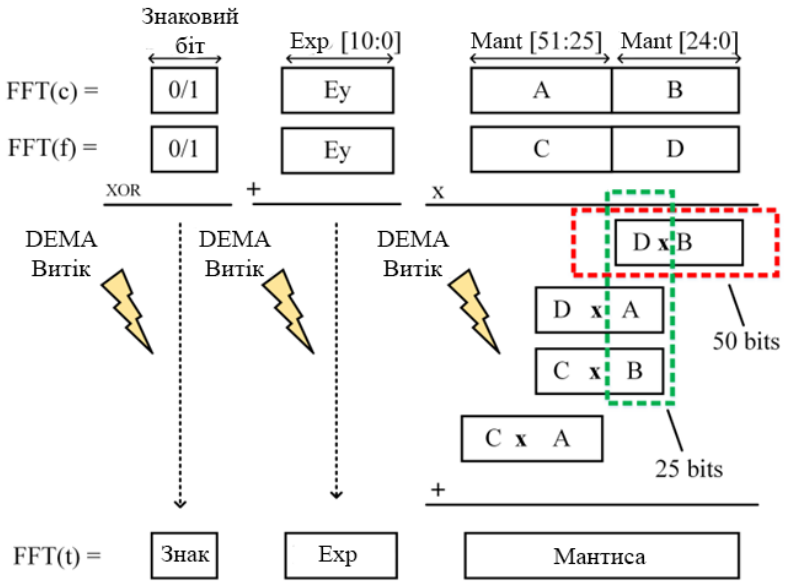


Рис. 4. Пропонована атака на множення FALCON з плаваючою крапкою. Націлювання на множення (показано червоними пунктирними лініями) створює хибні спрацьовування, а націлювання на проміжні додавання (показано зеленими пунктирними лініями) усуває їх

На рис. 4 показано кроки множення з плаваючою крапкою для двох коефіцієнтів. Операція множення приймає два 64-бітні коефіцієнти та генерує 64-бітний вихід. Операція складається з трьох частин: множення мантиси, додавання експоненти та обчислення знакового біта. Множення мантиси має чотири кроки. Перший об'єднує 52-бітну мантису вхідних коефіцієнтів з одним бітом «1», щоб зробити його старшим бітом (MSB). Другий розділяє біти мантиси на 28 біт вищого порядку та 25 біт нижчого порядку. Третій застосовує звичайне множення цілих чисел на розділені біти мантиси двох коефіцієнтів. Така послідовність кроків генерує 106-розрядні добутки. Оскільки кінцевий розмір мантиси має становити 52 біти, нижні біти називаються невикористаними, закріпленими бітами. Таким чином, четвертий крок полягає в тому, щоб округлити добуток множення до 52 бітів, видаливши ці закріплені біти.

Falcon застосовує додавання експоненти до 11-бітових експонент вхідних коефіцієнтів. Додавання експоненти також отримує додатковий біт переносу з результату множення мантиси. Останньою операцією є обчислення знакового біта, який є операцією XOR між бітами MSB вхідних коефіцієнтів.

Для атаки множення з плаваючою крапкою у Falcon пропонується використання стратегії «розділяй-і-володарюй». Запропонований метод атаки окремо відновлює біти мантиси, експоненти та знака та поєднує їх для отримання коефіцієнтів $FFT(f)$. Без такої стратегії проста диференціальна атака потребувала б створення масивних таблиць із 2^{64} записами для кожного коефіцієнта. На рис. 5 показано отриманий ЕМ-трафік цільового множення з плаваючою крапкою між двома коефіцієнтами [15]:

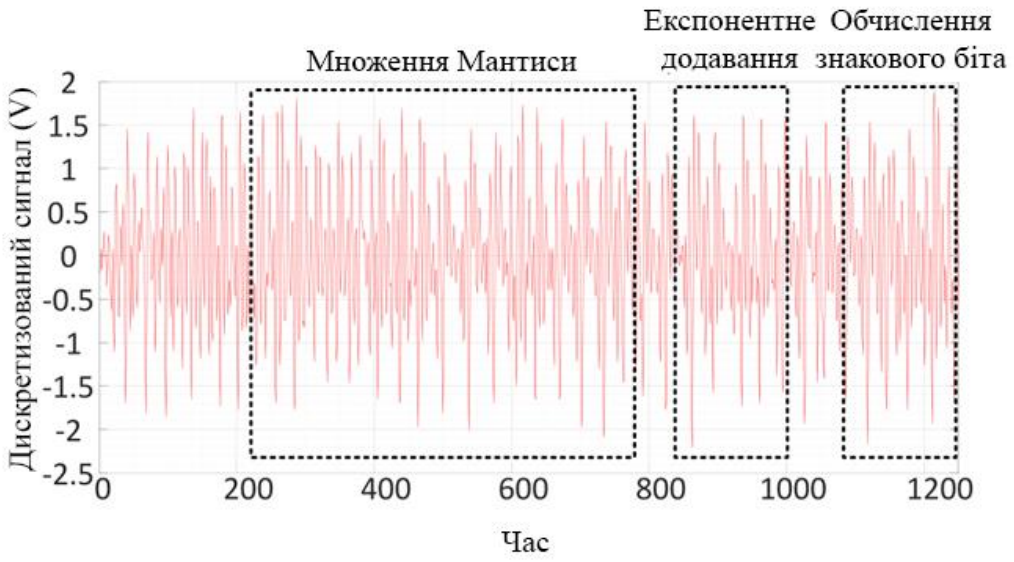


Рис. 5. Приклад даних (сліду) електромагнітних вимірювань з експерименту, що показує відповідні обчислення мантиси, експоненти та знаку

Червона лінія позначає електромагнітний сигнал, а чорні штрихові лінії вказують на те, протягом якого часу виконуються операції з мантисою, експонентою та знаком.

Ключовою проблемою в досягненні цієї атаки є усунення помилкових спрацьовувань, які виникають під час множення мантиси. Дійсно, відомо, що множення дає хибні позитивні результати, оскільки результати корельовані, тобто подібні коефіцієнти дадуть подібний результат множення. Наприклад, коефіцієнт $FFT(f)$ «1» і «2» генеруватиме однакові значення, зміщені на одну двійкову цифру. Це вирішується новою технікою розширення та скорочення [15].

B. «Розширення та скорочення» для видалення хибних спрацьовувань

Ми вирішуємо ключову проблему усунення хибно-позитивних результатів множення за допомогою нової техніки розширення та скорочення. Суть цієї техніки полягає у використанні найкращих припущень, отриманих під час атаки на множення (фаза розширення) та оцінки їх правильності шляхом атаки на наступну операцію (фаза скорочення), яка є додаванням проміжних результатів (рис. 2).

Хоча поліноміальні коефіцієнти f визначаються як цілі числа в діапазоні від -127 до 127, вони перетворюються на числа з плаваючою крапкою, а потім передаються в FFT домен. Оскільки алгоритм FFT змішує та розсіює всі вхідні коефіцієнти та оскільки FFT виконує арифметику з плаваючою крапкою, коефіцієнти домену FFT мають діапазон $[0, 2^{64}]$, навіть якщо вхідні дані мають діапазон $[-127,127]$.

Описана функція «Розширення та скорочення» налаштована під реалізацію арифметики з плаваючою крапкою Falcon. Щоб отримати частину з мантисою, ми спочатку виконується атака 25 біт молодшого порядку множення – на рис. 4 показано 25 біт молодшого порядку як B і D, де B відоме, а D – секретне значення. Запропонована диференціальна атака на множення $D \times B$ виконується за звичайними кроками: створення гіпотетичних припущень щодо секретних 25 біт нижчого порядку (D), обчислення очікуваної активності перемикання для кожного множення (з використанням ваги Хеммінга) і перевірка кореляції між ними та відповідними електромагнітними вимірюваннями. Та сама процедура повторюється і для $D \times A$, де A відоме, а D є секретним значенням. Це розширена фаза атаки, яка, як очікується, призведе до помилкових спрацьовувань.

Другим кроком у атаці є фаза скорочення, де ціллю є додавання $D \times B$ і $D \times A$ для скорочення помилкових позитивних значень і відновлення 25 біт секретної мантиси. На відміну від множення, додавання не призведе до хибних позитивних результатів: наприклад, однакові коефіцієнти «1» проти «2» генерують результати з різними вагами Хеммінга на основі інших вхідних даних додавання. За наявності достатньої кількості тестів секретні коефіцієнти «1» проти «2» (та інші випадки, які дають хибно-позитивний результат множення) можна відрізнити один від одного [15].

Для отримання D при операції додавання також застосовується вищезгаданий порядок дій. Біти вищого порядку множення мантиси виконують ті самі кроки множення та додавання. Тому до старших 27 бітів секретних коефіцієнтів застосовується та сама техніка розширення та скорочення. Зауважте, що ми не обходимо перше множення й безпосередньо атакуємо операцію додавання, оскільки розташування бітів добутку додавання $D \times B$ і $D \times A$ не узгоджуються одне з одним, що призводить до зниження успіху атаки.

Запропонована атака застосовує ту саму процедуру диференціальної ЕМ-атаки (DEMA) для вилучення бітів експоненти та бітів знака. На рис. 4 показано, що в реалізації з плаваючою крапкою відбувається додавання двох експонент поліномів $FFT(c)$ та $FFT(f)$ і застосовує операцію XOR до знакових бітів. Комбінована версія окремо відновлених бітів мантиси, експоненти та знака представляє один повний коефіцієнт [15].

2.3. Результати оцінювання

У дослідженні було використано загальнодоступне довідкове програмне забезпечення Falcon Round 3, яке міститься в пакеті подання до стандартизації NIST. Код було зкомпільовано за допомогою компілятора gcc-arm-none eabi-4_8-2014q1 і з міткою (прапором) -O0, а потім перенесено згенерований виконуваний файл на ARM-Cortex M4. Процесор має тактову частоту 168 МГц, а вимірювання виконуються за допомогою EM Probe LS (низька чутливість) RISC-EMP430LS, який є датчиком ближнього поля для вимірювання до 1 ГГц із чутливістю 20 мВ/1 Т@1 МГц. Шум вимірювань електромагнітного датчика зменшується за допомогою дросельної котушки та відбирається за допомогою осцилографа PicoScope 3206D зі швидкістю 500 Мс/с.

Для запропонованої диференціальної атаки побічним каналом ми використовуємо розрізнювач на основі коефіцієнта кореляції Пірсона на вагових моделях Хеммінга. Цей тест має на меті диференціювати популяції через їх коваріацію, тобто шляхом перевірки того, чи відхилення від середнього відбуваються подібним чином. Кореляційний слід $r_{i,j}$ для припущення i визначається як [15]:

$$r_{i,j} = \frac{\sum_{d=1}^{D}[(h_{d,i} - \bar{h}_i)(t_{d,j} - \bar{t}_j)]}{\sqrt{\sum_{d=1}^{D}(h_{d,i} - \bar{h}_i)^2 \sum_{d=1}^{D}(t_{d,j} - \bar{t}_j)^2}} \quad (1)$$

де D – це кількість слідів (трейсів), кожен з яких має T точок даних, $t_{d,j}$ – це електромагнітний слід з $0 < d \leq D$ і $0 < j \leq T$, $\bar{t}_j$ – середнє значення електромагнітного сліду, $h_{d,i}$ – оцінка витоку в сліді d для припущеного значення i, а $\bar{h}_i$ є середнім значенням цієї оцінки. Результат $r_{i,j}$ повертає кореляційний слід (трейс) зі значеннями в діапазоні $[-1,1]$, який є оцінкою лінійного зв'язку між припущеннями r_i та електромагнітними вимірюванням для кожного припущення i та часу j. Цей слід відображає значення та інформацію про час диференціального електромагнітного витоку [15].

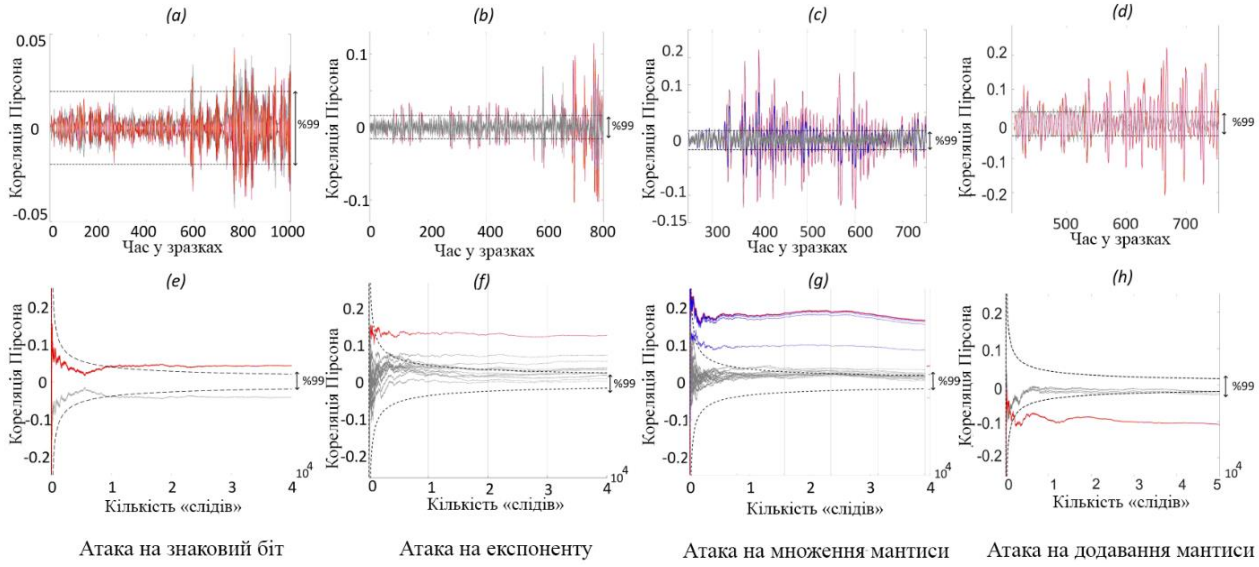


Рис. 6. Результати запропонованої атаки на коефіцієнт з плаваючою крапкою під час підписання за допомогою Falcon. Правильні припущення виділені жирним червоним шрифтом, значні помилкові припущення синім, а пунктирні лінії позначають довірчий інтервал 99,99 %. Attack повертає правильні припущення для (*a*) знака, (*b*) експоненти, (*c*), множення мантиси (*c*) і додавання (*d*), підтверджуючи наш підхід. Для цього коефіцієнта правильні значення можна отримати з достовірністю понад 99,99 % менш ніж за 10 тисяч вимірювань

На рис. 6 показано результати запропонованих атак, і це підтверджує, що атаки працюють на практиці та можуть зламати Falcon. Атака виконується на коефіцієнт 0xC06017BC8036b580, що означає, що правильний знаковий біт – 0x1, експонента – 0x406, а біти мантиси – 0x017bc8036b580 (старші біти – 0x00BDE40, а нижчі – 0X36b580). Часові графіки кореляції на рис. 6, *a* – *d* відповідно кількісно визначають, що кореляційні сліди ($r_{i,j}$) й для правильних і хибно-позитивних припущень для знака, експоненти та мантиси перетинають 99,99 % довірчий інтервал, тоді як ті, що знаходяться в межах є істинно негативними, тобто ми не маємо помилково негативного випадку. Графік показує лише до найкращих

21 припущення для частини мантиси для візуальної ясності – атака оцінює до 2^{27} припущень для бітів мантиси вищого порядку та 2^{25} припущень для бітів нижчого порядку.

Рис. 6, *c*, *d* підтверджують нашу гіпотезу про запропоновану атаку. Дійсно, коли атака виконується на множення бітів мантиси, результати на рисунку 6, *c* показують, що значні помилкові спрацьовування відбуваються на першому кроці атаки, коли вона спрямована на множення мантиси. Результат кореляції п'яти найкращих припущень, які включають правильне припущення та чотири помилкові позитивні результати, насправді є абсолютно однаковими (дещо відрізняються на рисунку для наочності). Однак результати на рис.6, *d* підтверджують, що за допомогою подібної стратегії розширення та скорочення всі хибні спрацьовування усуваються, коли атака потім відстежує припущення на попередньому кроці, зосереджуючись на проміжних доповненнях [15].

На рис. 6, *e* – *h* зображено еволюцію кореляції, взяту за вибірку часу з найменшим витоком на рис. 6, *a* – *d* відповідно. Ці графіки вимірюють кількість слідів, необхідних для досягнення статистично значущої (99,99 %) кореляції. Витік правильних припущень стає статистично значущим лише з тисячею вимірювань під час атаки на експоненту та додавання мантиси, тоді як інші припущення зникають [15].

Найскладнішою частиною атаки, з точки зору кількості необхідних вимірювань, є вилучення знакового біта, і потрібно близько кількох сотень вимірювань, щоб отримати правильне значення, і приблизно 9 тисяч вимірювань, щоб зробити витік статистично значущим для цього прикладу. Загалом, вимірювання для всіх коефіцієнтів можна впевнено отримати менш ніж за 10 тисяч вимірювань. Зауважте, що витік знакових бітів є симетричним для припущень позитивного та негативного знаку. Це справді очікувано і не створює проблем для атаки, оскільки правильне припущення постійно має позитивну кореляцію в максимальній точці витоку (тобто червона лінія на рис. 6, *e* виглядає однаково як для негативних, так і для позитивних знаків) [15].

Було здійснено запропоновану атаку на Falcon-512, але така сама атака може бути застосована до іншого набору параметрів, тобто Falcon-1024, оскільки Falcon використовує однакові арифметичні реалізації з плаваючою крапкою для обох наборів параметрів.

2.4. Особливості

А. Обмеження подібної атаки

В роботі запропоновано загальну атаку, яка працює без профілювання цільового пристрою шляхом (пере)конфігурації секретного ключа, і це все ще практично. Можна розширити атаку за допомогою шаблонів або методів профілювання на основі машинного навчання і використання кращого вимірювального обладнання.

Б. Можливі контрзаходи

Найпопулярнішими методами захисту побічних каналів є приховування та маскування. У той час як приховування має на меті зробити енергоспоживання постійним, маскування має на меті рандомізацію проміжних значень, оброблених реалізацією. Хоча SABRE, ще один фіналіст постквантової сигнатури NIST, нещодавно запропонував замасковану реалізацію, вона ще не існує для Falcon.

В. NTT проти FFT – перспектива щодо побічних каналів

Ґрунтуючись на аналізі, можна стверджувати, що FFT, ймовірно, має нижчий рівень витоку потужності/ЕМ випромінювання за побічними каналами порівняно з NTT. Хоча для подібної атаки на FFT потрібно близько 10 тисяч трас, NTT виявилася вразливою навіть з одним слідом. Ймовірно, це пов'язано з великою кількістю нелінійності. У той час як NTT застосовує модульне скорочення простих p, цього не відбувається у FFT. Таким чином, атака розрізнить і усуне неправильні припущення в NTT значно простіше і швидше.

У цьому розділі продемонстровано першу атаку побічним каналом на постквантового фіналіста NIST Falcon. Результати підтвердили, що подібні атаки дійсно працюють на практиці з кількома тисячами вимірювань і без квантового комп'ютера. Таким чином, робота

обґрунтовує необхідність ефективних заходів протидії таким атакам і врахування відповідних накладних витрат у показниках продуктивності апаратного/програмного забезпечення [15].

3. Атака помилками на Falcon – BEARZ

BEARZ [6] представляє собою атаку помилками на схему електронного підпису Falcon. Дана атака була вперше запропонована Sarah McCarthy та ін. у 2019 році. Заснована вона на вилученні базису через переривання рекурсії або обнулення. Викликаючи помилки в роботі схеми, атака добуває закриту ключову інформацію. Моделлю зловмисника передбачено можливості пропуску команд та обнулення змінних.

3.1. Рекурсія алгоритму Falcon

Оскільки алгоритм вибірки для Falcon є рекурсивною формою GPV відбірника, атака буде спиратися на переривання рекурсивного виклику на початку. Як було сказано вище, Falcon [7, 12] має два рекурсивні виклики на верхньому рівні алгоритму ffSampling. Наприклад, задано вектор (t_0, t_1), алгоритм застосовується спочатку для t_1, а потім для t_0. Кожен елемент (t_0, t_1) безперервно ділиться на два вектори довжини $n = n/2$ поки n не буде дорівнювати 2. Після цього відбираються коефіцієнти гаусового розподілу для отримання вектору вибірки (z_0, z_1). Це спричиняє присутність на верхньому рівні алгоритму двох рекурсивних гілок. Для вектору довжини n, перші $n/2$ значень будуть представлені реальними коефіцієнтами, а другі $n/2$ – уявними [6].

Успішна атака потребуватиме переривання рекурсивного виклику у необхідній точці таким чином, щоб тільки $m-n$ місць були б заповнені. Але через характер функцій FFT така операція є неможливою.

3.2. FFT: злиття та розділення

Функції злиття та розділення для схеми Falcon виконуються у домені FFT. Це означає, що нульові вхідні дані не представлені нулем, що в свою чергу є проблемою для атаки помилками [6]. Тому після отримання вектору решітки із гаусового зразка, до нього застосовується зворотня FFT функція FFT^{-1}. Це робиться з метою переконатися, що підпис не знаходиться у FFT домені. Коли всі перераховані кроки виконані, може бути використана та сама післяобробка, що і для DLP-атаки.

3.3. Атаки переривання рекурсії

Атаки переривання рекурсії класифікуються в залежності від місця, в якому відбувається переривання. Усі вони призводять до однакового вихідного формату вектору z; перші $(2n-m)$ коефіцієнтів $z = (z_0, z_1)$ примусово задаються нулями.

Переривання другої рекурсії (для $m = n$)

Атака здійснюється шляхом переривання в кінці виклику алгоритму вибірки, після першого рекурсивного виклику [6]. Якщо атака виявиться успішною, то z_1 заповниться вибраними коефіцієнтами, а z_0 залишиться повністю нульовим. Такий тип атаки може бути виконаний також для $m \le n$, з метою обнулення перших n коефіцієнтів.

Обнулення або атака пропуску (для $m \le n$)

Існує два варіанти подібної атаки. Перший: при передостанньому злитті встановити необхідні з вихідних коефіцієнтів в нуль шляхом пропуску операцій або подальшого обнулення необхідних коефіцієнтів [6]. Операції, які необхідно пропустити: $f[(u << 1) + 0] = t\_re$ та $f[(u << 1) + 1] = t\_re$ з коду (Prest та ін., 2017)($merge\_fft()$ функція у $falcon\_fft.c$) [4, 9].

Другий: встановити необхідну кількість початкових коефіцієнтів z_1 в нуль перед обчисленням відповідного вектору решітки, тобто перезаписати вихід z_1 відбірника [6].

Переривання посеред рекурсії (для $m \le n$)

Даний тип потребує попереднього обчислення (одноразового), зате дає змогу застосовувати помилку на стадії одновимірного гаусового відбірника [9, 10, 16]. Це є сприятливою умовою, оскільки передбачає простоту фізичного вбудовування.

Якщо необхідно прирівняти до нуля ліву половину z_1 (тобто $m = n/2$), то вектор лівої сторони (LHS) при останньому виклику $merge\_fft()$ має містити у своїй першій половині нульові значення, а у векторі правої сторони (RHS) перша половина коефіцієнтів має дорівнювати його другій половині. Кожен дійсний коефіцієнт вектору z_1 генерується як $z_1[2u] = f_0[u] + (f_1[u] - f_1[u + n/4])$ та $z_1[2u+1] = f_0[u] - (f_1[u] - f_1[u + n/4])$, де n – розмірність вектору вищого рівня, а $u \in \{0, ..., n/4 - 1\}$. Для обнулення $z_1[2u]$ та $z_1[2u+1]$, можна встановити $f_0[u] = 0$ та $f_1[u] = f_1[u + n/4]$ для кожного $u \in \{0, ..., n/4 - 1\}$. Наприклад, для $n = 512$, перші $n/4 = 128$ коефіцієнтів 256-мірного вектору LHS встановлюються рівними нулю, а перші 128 коефіцієнтів 256-мірного вектору RHS встановлюються рівними другим 128 коефіцієнтам цього ж вектору [6].

256-розмірній RHS вектор повинен містити в першій половині значення, рівні значенням другої половини. Для забезпечення цього 128-мірний вектор LHS має містити в собі дійсні значення, що дорівнювали б уявним значенням, а 128-розмірний RHS вектор має бути рівним нулю.

Це обумовлюється тим, що нам необхідно, щоб $f[2u] = f[2u + n/2]$ та $f[2u+1] = f[2u+1+n/2]$ були рівними для кожної ітерації u. З погляду функції $merge\_fft()$ це означає: $f_0[u] + (f_1[u] - f_1[u + n/4]) = f_0[u + n/4] + (f_1[u] + f_1[u + n/4])$ та $f_0[u] - (f_1[u] - f_1[u + n/4]) = f_0[u + n/4] - (f_1[u] + f_1[u + n/4])$. Щоб відповідати умовам рівнянь, можна встановити $f_0[u] = f_0[u + n/4]$ – так, що перша половина коефіцієнтів дорівнюватиме другій половині, та $f_1[u] = f_1[u + n/4] = 0$. Таким чином, рівняння будуть просто залежати від вектору подачі LHS [6].

Кожну нижню гілку для RHS можна встановити рівною нулю. Якщо взяти LHS 128-розмірний вектор: 64-мірний вектор LHS повинен мати рівні дійсні та уявні значення, а RHS має бути нульовим. Будь-який вектор, рівний нулю, повинен мати обидва вектори подачі, рівними нулю, тому гілки нижче цієї можуть бути обнулені [16]. Даний метод може бути застосований для будь-якого m, що відповідає умові $m = 2^k$, де $k \in \mathbb{Z}$.

3.4. Обробка після нападу

Після проведення попередніх кроків та отримання недійсного підпису, останнім кроком атаки є відновлення таємного базису з даного підпису [6, 16]. Припустимо, мається $2n - m$ перших коефіцієнтів вектору (z_0, z_1) рівних нулю. Тоді підпис Falcon s_2 буде обчислюватись як

$$s = t - z\hat{B}, \tag{2}$$

де $\hat{B}$ матрицею базису у FFT домені.

Але в нашому випадку важливою є лише друга половина s, тобто s_2:

$$s_2 = t_1 - z\left(\frac{-A(f)}{-A(F)}\right). \tag{3}$$

Оскільки t_1 встановлено в 0:

$$s_2 = -z\left(\frac{-A(f)}{-A(F)}\right). \quad (4)$$

До того ж перші $m-n$ коефіцієнтів z нульові, отже:

$$s_2 = -z_1(-A(F)). \quad (5)$$

А оскільки відомо, що деякі з коефіцієнтів z_1 нульові, отримуємо:

$$s_2 = (z_1[m-1]x^{n-m}F + ... + z_1[0]x^{n-1}F). \quad (6)$$

Таким чином, вдається отримати підрешітку решітки, яка була згенерована з використанням F. А отже, маючи декілька недійсних підписів, є можливим знаходження решітки, породженої F (F – короткий вектор у даній решітці). Далі за допомогою алгоритму BKZ можна знайти цей короткий вектор [6]. Знаючи F, можна отримати G, f і g з відкритого ключа h, і тим самим знайти таємний базис NTRU решітки.

3.5. Модель помилки та контрзаходи

У моделі помилки передбачено проведення аналізу побічних каналів [6]. Шляхом аналізу може бути виявлено вікно між першим і другим рекурсивним викликом. В межах цього вікна алгоритм може бути перервано. Атака обнулення може бути застосована у момент зберігання вектору в оперативній пам'яті, шляхом обнулення під час цього необхідних бітів [16]. Альтернативним варіантом може бути пропуск рядків коду. Даний тип атаки можна організувати за допомогою перепадів тактової частоти процесора [16].

Для протистояння BEARZ [6] атаці існують певні контрзаходи. Наприклад, подвійне обчислення підпису – один з найпростіших методів виявлення атак помилками. Таким чином, відразу після підписання повідомлення підписувач має змогу переконатися в тому, що на обладнання не здійснювались атаки помилками в момент підписання. Ще одним ефективним методом виявлення BEARZ атаки є перевірка того, що вибраний вектор не йде до нуля в певній точці вздовж своєї довжини в кінці ffSampler алгоритму.

3.6. Модель помилки та контрзаходи

Даний пункт демонструє порівняльні характеристики алгоритму Falcon в чистій реалізації та реалізації з застосованими контрзаходами.

Для порівняння були обрані два набори параметрів, які застосовуються для схеми ЕП Falcon [7, 12]. Дані параметри показані в табл. 1.

Таблиця 1

Набори параметрів для тестування

Набір параметрів	Розмірність (N)	Модуль (q)
Набір 1	512	12289
Набір 2	1024	

Порівняння реалізацій з різними контрзаходами наводиться в табл. 2. Тестування проводилося з використанням CPU AMD Ryzen 7 3750H @ 2,3 – 4,0 ГГц.

Таблиця 2

Результати аналізу продуктивності

Реалізація	Набір параметрів	% погіршення
Чиста реалізація (Thomas Prest та ін.)	Набір 1	-
	Набір 2	-
+ постійне за часом NTT	Набір 1	-
	Набір 2	-
+ перевірка після підпису	Набір 1	5,5
	Набір 2	5,5
+ нульова перевірка	Набір 1	34
	Набір 2	16,5
+ відкидання зразка (6,25%)	Набір 1	14
	Набір 2	8,5
+ перемішування Фішера-Йейтса	Набір 1	21
	Набір 2	3
+ BlindVector	Набір 1	12
	Набір 2	52
Рекомендований набір контрзаходів: + перевірка після підпису + відкидання зразка (6,25%) + перемішування Фішера-Йейтса	Набір 1	4,6
	Набір 2	4,4

Для більшої наочності результати представлено у вигляді діаграми:

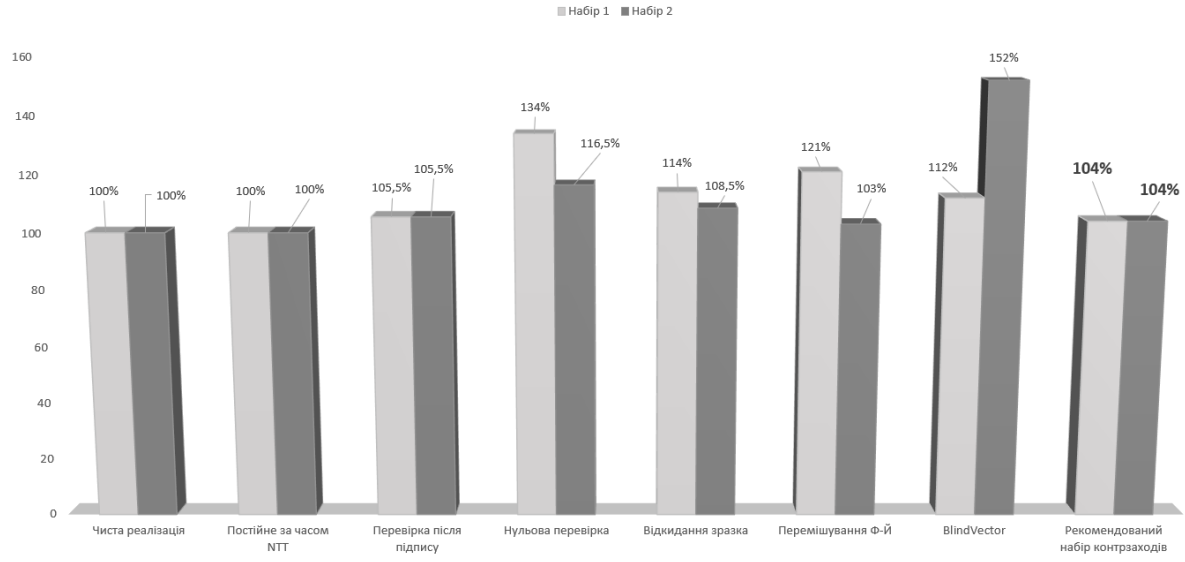


Рис. 7. Вплив контрзаходів на швидкодію

Як можна спостерігати в табл. 1 та на рис. 7, контрзаходи впливають на швидкодію Falcon незначним чином. Рекомендований набір контрзаходів погіршує продуктивність менш ніж на 5 %. Такий хороший результат зумовлений ефективністю процесів вибірки та перевірки.

Для перевірки контрзаходу проти атак помилками було 100 разів проведено процедуру підписання. Для набору 1 швидкість була в діапазоні 50 операцій на секунду, а для набору 2 – 20 операцій на секунду. Отримані дані свідчать про його ефективність проти часових атак. Також, можна додати, що контрзахід нульової перевірки виявляє атаку зі 100 % успіхом. Даний захід можна рекомендувати як мінімальний і достатній контрзахід.

Висновки

1. Після проведення аналізу вразливостей та стійкості алгоритму Falcon проти спеціальних атак можна сказати, що недоліком даного алгоритму є мала кількість досліджень стійкості проти спеціальних атак. Було показано, що певні математичні компоненти, які використовуються в алгоритмі ЕП Falcon, є вразливими до атак, заснованих на аналізі часових показників.

2. Також, що відбірник секретної інформації заснований на решітках, який застосовується у алгоритмі Falcon, настільки ж вразливий до атак помилками, як і відбірник, що використовується в альтернативних схемах підпису. Через це при стандартизації чи впровадженні слід розглядати можливість фізичних атак.

3. Вразливими до спеціальних атак є вибірка Гауса, теоретичне перетворення числа (NTT) та функція гешування HashtoPoint.

4. Аналіз можливих контрзаходів для протидії спеціальним атакам, показав, що вони впливають на швидкодію Falcon незначним чином. Рекомендований набір контрзаходів погіршує продуктивність менш ніж на 5 %. Такий перспективний результат зумовлений ефективністю процесів вибірки та перевірки.

5. Отже, незважаючи на те, що відбірник схеми Falcon все ж є вразливим до атак з помилками, ефективність компонентів та математики даного алгоритму електронного підпису сприяє тому, що він здатен конкурувати з іншими схемами, навіть з контрзаходами проти цих атак.

6. Розглянуто атаку побічними каналами на Falcon. Така атака є атакою з відомим відкритим текстом, яка використовує електромагнітні вимірювання пристрою для отримання секретних ключів підпису, які потім можна використовувати для підробки підписів у довільних повідомленнях. Запропонована атака націлена на унікальне множення з плаваючою комою в рамках швидкого перетворення Фур'є алгоритму Falcon за допомогою нової стратегії розширення та скорочення, яка отримує змінні знака, мантиси та експоненти без помилкових спрацьовувань.

7. Отримані значення з плаваючою комою відображаються назад у коефіцієнти секретного ключа. Подібна атака, зокрема, не вимагає попередньої характеристики профілю потужності цільового пристрою або створення спеціальних вхідних даних. Натомість статистичні відмінності на отриманих даних достатні для успішного виконання запропонованого диференційного електромагнітного аналізу. Результати на ARM-Cortex-M4, що працює з довідковим програмним забезпеченням FALCON NIST, показують, що приблизно 10 тисяч вимірювань достатньо, щоб отримати весь ключ.

8. Запропоновано загальну атаку, яка працює без профілювання цільового пристрою шляхом (пере)конфігурації секретного ключа, і це все ще практично. Є можливість розширити атаку за допомогою шаблонів або методів профілювання на основі машинного навчання і використання кращого вимірювального обладнання.

9. Основними методами протидії подібним атакам можна вважати приховування та маскування. Які, наприклад, реалізовані у SABRE [4], ще одному фіналісті конкурсу NIST.

Список літератури:

1. P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer // SIAM Review, vol. 41, no. 2, pp. 303– 332, 1999.

2. J. Proos et al. Shor's discrete logarithm quantum algorithm for elliptic curves // Quantum Info. Comput., vol. 3, no. 4, pp. 317–344, Jul. 2003.

3. P. Kocher et al. Differential power analysis // Advances in Cryptology – CRYPTO' 99, 1999, pp. 388–397.
4. Post-Quantum Cryptography. Round 3 Submissions. 2020. URL: https://csrc.nist.gov/Projects/post-quantum-cryptography/round-3-submissions.
5. L´eo Ducas, Vadim Lyubashevsky and Thomas Prest. Efficient Identity-Based Encryption over NTRU Lattices. 2014. URL: https://eprint.iacr.org/2014/794.pdf.
6. Sarah McCarthy, James Howea, Neil Smythb, Séamus Brannigan, and Máire O'Neill. BEARZ Attack FALCON: Implementation Attacks with Countermeasures on the FALCON signature scheme. 2019. URL: https://eprint.iacr.org/2019/478.pdf.
7. Pierre-Alain Fouque, Jeffrey Hoffstein, Paul Kirchner, Vadim Lyubashevsky, Thomas Pornin, Thomas Prest, Thomas Ricosset, Gregor Seiler, William Whyte, Zhenfei Zhang. Falcon: Fast-Fourier Lattice-based Compact Signatures over NTRU Specifications v1.2. 2020. URL: https://falcon-sign.info/falcon.pdf.
8. Verbauwhede I., Karaklajic D., and Schmidt J.-M. The Fault Attack Jungle – A Classification Model to Guide You. 2011. URL: https://www.esat.kuleuven.be/cosic/publications/article-2046.pdf.
9. J. Ahrens and U. Dieter. Extension of forsythe's method for random sampling from the normal distribution. 1973. URL: https://www.ams.org/journals/mcom/1973-27-124/S0025-5718-1973-0329190-8/S0025-5718-1973-0329190-8.pdf.
10. Thomas Prest. Gaussian Sampling in Lattice-Based Cryptography. 2015. URL: https://tel.archives-ouvertes.fr/tel-01245066v2/document.
11. Patrick Longa and Michael Naehrig. Speeding up the Number Theoretic Transform for Faster Ideal Lattice-Based Cryptography. 2016. URL: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/05/RLWE-1.pdf.
12. Офіційний сайт ЕП Falcon. URL: https://falcon-sign.info.
13. Hodgers P., Regazzoni F., Gilmore R., Moore C., and Oder T. State-of-the-art in physical side-channel attacks and resistant technologies. 2016. URL: https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5a63fd691&appId=PPGMS.
14. Robert Primas. Side-Channel Attacks on Efficient Lattice-Based Encryption. 2017. URL: https://diglib.tugraz.at/download.php?id=5a1def5f2e7fa&location=browse.
15. E. Karabulut and A. Aysu. Falcon down: Breaking falcon post-quantum signature scheme through side-channel attacks // 2021 58th ACM/IEEE Design Automation Conference (DAC), pp. 691-696, 2021.
16. Дерев'янко Я.А., Горбенко І.Д. Вимоги та результати оцінки захищеності перспективного електронного підпису від спеціальних атак. 2020. URL: http://www.viti.edu.ua/files/zbk/2020/c_2020.pdf.

Надійшла до редколегії 10.09.2022

Відомості про авторів:

Дерев'янко Ярослав Андрійович – студент кафедри безпеки інформаційних систем і технологій, факультету комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна; Україна; e-mail: yarik0009258@gmail.com; ORCID: https://orcid.org/0000-0002-3290-3373

Горбенко Іван Дмитрович – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ "Інститут Інформаційних Технологій", головний конструктор, Україна; e-mail: gorbenkoi@iit.kharkov.ua; ORCID: https://orcid.org/0000-0003-4616-3449

УДК 004.056: 004.056.5 **DOI:10.30837/rt.2022.3.210.04**

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

ДОСЛІДЖЕННЯ ОСНОВНИХ СХЕМ ШИФРУВАННЯ З МОЖЛИВІСТЮ ПОШУКУ У БАЗАХ ДАНИХ, ЯКІ ПІДТРИМУЮТЬ SQL

Вступ

Аутсорсинг зберігання та обробки даних на сторонніх серверах, таких як сервери хмар, широко використовується і демонструє вибухове зростання [1]. Однак у міру збільшення масштабу, цінності та централізації даних зростають проблеми безпеки та приватності. Існує виявлений ризик того, що дані, які зберігаються в базах даних, будуть скомпрометовані [2]. А цього, відповідно до різних міжнародних законів і стандартів таких, як: Загальний регламент захисту персональних даних Європейського Союзу (General Data Protection Regulation – GDPR) [3], Стандарт безпеки даних індустрії платіжних карток (Payment Card Industry Data Security Standard – PCI DSS) [4], Закон про переносимість та підзвітність медичного страхування (Health Insurance Portability and Accountability Act – HIPAA) [5, 6] та деякими іншими, не можна допустити. Це стимулювало дослідження у сфері безпечного управління даними та підвищило їх актуальність.

Шифрування – це стандартний підхід до забезпечення конфіденційності даних, що передаються на аутсорсинг так званим чесним, але допитливим серверам хмар. Шифрування унеможливлює доступ до даних без ключів як для інсайдерів (insider), так і для сторонніх осіб (outsider). Однак традиційні схеми шифрування позбавляють користувачів певних функціональних можливостей над даними, таких, наприклад, як пошук [7]. У цьому випадку для пошуку слів у зашифрованих зовнішніх даних користувачам необхідно отримати весь набір даних із відповідного сховища, розшифрувати його та здійснити пошук локально. Такий підхід створює серйозні проблеми з продуктивністю, які зводять нанівець переваги аутсорсингу, внаслідок чого для більшості застосунків він стає неприйнятним.

Проблема пошуку за зашифрованими даними викликала великий інтерес як у наукових колах, так і в індустрії. Однак, як можна помітити [8 – 10], дослідження з шифрування з можливістю пошуку (searchable encryption – SE) більшою мірою зосереджені на сценарії користувача, який передає на аутсорсинг зашифрований набір документів (таких як електронна пошта, медичні записи тощо) і хотів би продовжити пошук за ключовими словами у цьому зашифрованому наборі даних. Хоча на практиці багато компаній, організацій, установ різних форм власності зберігають дані в базах, що використовують реляційну модель даних. Широко поширена мова SQL дозволяє користувачам зберігати, запитувати та оновлювати свої дані у зручній для них формі. Бази даних SQL (до того ж NewSQL та деякі NoSQL бази даних також дозволяють працювати в парадигмі SQL-запитів) забезпечують швидкий пошук та вилучення записів за умови, що сервер SQL може зчитувати вміст даних. Однак шифрування зазвичай заважає серверу зчитувати необхідні дані, а отже, ускладнює пошук у зашифрованих базах даних. Зокрема, механізм криптографічного захисту даних для пошуку за зашифрованими даними, що зберігаються в базі даних, повинен дозволяти серверу ефективно обробляти пошукові запити, не маючи доступу до відкритих даних. Крім того, запити до зашифрованої бази даних мають бути зручними для користувача. Запити зазвичай виражаються стандартною мовою, такою, наприклад, як SQL. Запити повинні імітувати функції пошуку, так само, якби дані не були зашифровані. Безпосереднє застосування рішень для пошуку необхідної інформації в зашифрованих БД даних, що підтримують SQL, не є простим завданням.

Відомі дві основні проблеми конфіденційності [2, 11]. По-перше, власник даних повинен бути впевнений, що дані, які зберігаються на сайті постачальника послуг, захищені від крадіжки даних сторонніми особами. По-друге, дані мають бути захищені навіть від постачаль-

ників послуг (допустимого (valid) користувача, відомого як інсайдер), якщо самим постачальникам не можна довіряти. При цьому зазвичай розрізняють:

– противників / зловмисників, які є напівчесними (semi-hones; або чесними, але допитливими – honest-but-curious), тобто вони наслідують запропоновані протоколи, але можуть пасивно намагатися отримати додаткову інформацію з повідомлень, які вони спостерігають;

– противників, які є шкідливими (malicious), що означає, що вони активно бажають виконувати будь-які дії, необхідні для отримання додаткової інформації чи впливу на роботу системи.

Крім того, розрізняють зловмисників, які зберігаються протягом усього терміну служби бази даних, та тих, які отримують моментальний знімок в один момент часу. Більшість активних досліджень у галузі технології захищеного пошуку розглядає напівчесний захист від постійного внутрішнього противника [2]. У цьому роботі ми також зосередимося на цьому аспекті.

Проєктування захищеної пошукової системи – це баланс між безпекою, функціональністю, продуктивністю та зручністю використання. Описи безпеки зосереджені на інформації, що розкривається або просочується зловмиснику, який має доступ до сервера бази даних. Функціональність насамперед характеризується типами запитів, на які може відповісти захищена база даних. На продуктивність та зручність використання впливають структури даних бази даних та механізми індексування, а також необхідні обчислювальні та мережеві витрати.

Основне завдання в аналізованому аспекті захищеної системи пошуку на віддаленому сервері, якому не довіряють, спрямована на те, щоб сервер нічого не дізнався про дані, що зберігаються в захищеній базі даних, або про запити, а запитувач нічого не дізнався, крім результатів запиту, при цьому залишити можливість використовувати запити SQL типу на зашифрованих даних. Для цього застосовуються різні підходи, системи та методи. Однак, незважаючи на велику різноманітність запропонованих варіантів, немає домінуючого рішення для всіх випадків використання. Не існує найбільш захищеної пошукової системи або набору методів. Користувачі повинні розуміти характеристики системи та компроміси для свого варіанта використання. Тому розглянемо далі деякі з основних існуючих рішень.

1. CryptDB

CryptDB – це система, яка забезпечує безпечне зберігання конфіденційних даних у віддалених БД, у тому числі, що обслуговуються у хмарних сервісах. Вона працює, виконуючи SQL-запити до зашифрованих даних, використовуючи набір ефективних схем шифрування, які підтримують SQL. CryptDB може пов'язувати ключі шифрування з паролями користувачів, так що елемент даних може бути розшифрований лише за допомогою пароля одного з користувачів, які мають доступ до цих даних. В результаті адміністратор бази даних ніколи не отримає доступу до розшифрованих даних, і навіть якщо всі сервери будуть скомпрометовані, зловмисник не зможе розшифрувати дані жодного користувача, що не увійшов до системи [12].

В основі CryptDB лежить використання можливостей СУБД MySQL або PostgreSQL. Основна перевага CryptDB полягає у виконанні запитів над зашифрованими даними, що уможливлює застосування CryptDB на практиці – використання чітко визначеного набору SQL-операторів, кожен з яких може ефективно обробляти зашифровані дані.

У CryptDB розглядаються дві основні загрози [12]:

– Загроза 1 – допитливий адміністратор бази даних – пасивний противник, який намагається дізнатися про конфіденційні дані (шляхом відстеження на сервері СУБД), але CryptDB перешкоджає цьому;

– Загроза 2 – зловмисник, який може отримати повний контроль над застосунками та серверами СУБД. У цьому випадку CryptDB не може надати будь-яких гарантій користува-

чам, які вже працювали з застосунком під час атаки, але все ж таки може забезпечити конфіденційність даних інших користувачів, що вийшли з системи.

Архітектура CryptDB складається з двох частин: проксі-сервера бази даних та немодифікованої СУБД (рис. 1).

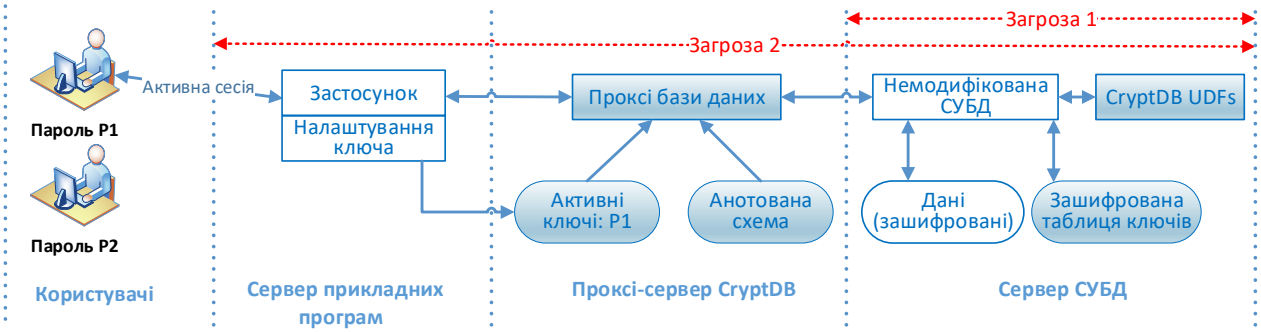


Рис. 1. Архітектура CryptDB

CryptDB використовує функції, що визначаються користувачем (User Defined Functions – UDFs) для виконання криптографічних операцій в СУБД. Прямокутники та закруглені прямокутники представляють процеси та дані відповідно. Затінення вказує на компоненти, додані CryptDB. Пунктирні лінії вказують на поділ між комп'ютерами користувачів, сервером прикладних програм, сервером, на якому працює проксі-сервер бази даних CryptDB (який зазвичай збігається з сервером прикладних програм), і сервером СУБД. CryptDB усуває два типи загроз, показаних пунктирними лініями. При загрозі один допитливий адміністратор бази даних з повним доступом до сервера СУБД відстежує особисті (private) дані, і в цьому випадку CryptDB запобігає доступу адміністратора баз даних до будь-якої приватної інформації. При загрозі 2 зловмисник отримує повний контроль як над програмним, так і над апаратним забезпеченням програми, проксі та серверами СУБД, і в цьому випадку CryptDB гарантує, що зловмисник не зможе отримати дані, що належать користувачам, які не увійшли до системи (наприклад, користувач 2).

Проксі-сервер використовує секретні ключі для шифрування всіх даних, що вставлені. Основна ідея обчислень над зашифрованими даними полягає у тому, щоб дозволити серверу СУБД виконувати обробку запитів до зашифрованих даних, як це було б із незашифрованою базою даних, тобто дозволити йому обчислювати певні функції над елементами даних на основі зашифрованих даних. Наприклад, якщо СУБД необхідно виконати команду угруповання GROUP BY за деяким стовпцем, то сервер СУБД повинен вміти визначати, які елементи в цьому стовпці рівні один одному, але не фактичне значення кожного.

Противник може отримати доступ до ключів, які використовуються для шифрування всієї бази даних. Рішення полягає в тому, щоб зашифрувати різні елементи даних (наприклад, дані, що належать різним користувачам) із різними ключами. Противник, який атакує сервер прикладних програм або проксі-сервер, тепер може розшифрувати лише дані користувачів, що увійшли до системи в даний момент (дані, що зберігаються на проксі-сервері). Дані неактивних користувачів зашифровані ключами, які недоступні зловмиснику і залишаться конфіденційними. Для успішної роботи CryptDB використовує три основні рішення.

1. До зашифрованих баз даних звернення здійснюється за допомогою SQL-запитів, що дозволяють виконувати такі операції, як перевірка на рівність, порівняння порядку, агрегація / підсумовування та з'єднання (це уможливлює здійснення на практиці обробку зашифрованих даних). CryptDB шляхом адаптації відомих схем шифрування (для перевірки на рівність, підсумовування, порівняння порядку) та використання нового криптографічного методу із збереженням конфіденційності для з'єднань, CryptDB шифрує кожен елемент даних таким чином, щоб СУБД могла працювати з перетвореними даними. В основному використовується шифрування із симетричним ключем.

2. Шифрування, що настроюється, на основі запитів. Деякі схеми шифрування пропускають на сервер СУБД більше інформації про дані, ніж інші, але вони необхідні для обробки певних запитів. Щоб уникнути цього, CryptDB ретельно (залежно від запитів) налаштовує схему шифрування для будь-якого заданого елемента даних, використовуючи так звані «цибулини» (onions) шифрування. Onions – це новий спосіб компактного зберігання кількох зашифрованих текстів один в одному в базі даних та запобігання дорогому повторному шифруванню.

CryptDB має можливість перемикатися «на льоту» між різними криптографічними схемами залежно від типу операції, що виконується. Це реалізовано за рахунок «цибулинного» багатоступінчастого шифрування, коли дані зашифровані в кілька шарів різними алгоритмами. У кожного шару свій ключ і свій список операцій, що підтримуються. На нижньому шарі використовуються найнадійніші алгоритми, а операції у верхніх шарах можливі без розшифровування нижніх шарів.

3. Зв'язування ключів шифрування з паролями користувачів, щоб кожен елемент даних у БД можна було розшифрувати тільки за допомогою ланцюжка ключів, що базується на паролі одного з користувачів, що мають доступ до цих даних. В результаті, якщо користувач не авторизований у застосунку, і якщо зловмисник не знає пароль користувача, то зловмисник не може розшифрувати дані користувача, навіть якщо СУБД та сервер прикладних програм повністю скомпрометовані.

CryptDB використовує для різних запитів до бази даних наступні типи шифрування.

Випадковий (Random – RND). RND забезпечує максимальну безпеку в CryptDB: IND-CPA (indistinguishability under chosen plaintext attack – нерозрізненість при атаці за вибраним відкритим текстом), схема є ймовірнісною, що означає, що два рівні значення відображаються в різні шифртексти з ймовірністю близькою до одиниці. З іншого боку, RND не дозволяє ефективно виконувати будь-які обчислення із зашифрованим текстом. Ефективна конструкція RND полягає у використанні блокового шифру, такого як AES або Blowfish, у режимі зчеплення блоків шифртексту (Cipher Block Chaining – CBC) разом із випадковим вектором ініціалізації (IV).

Детермінований (Deterministic – DET). DET має трохи більш слабку гарантію, але, як і раніше, забезпечує надійний захист: можливий витік тільки тих зашифрованих значень, які відповідають одному й тому ж значенню даних, детерміновано генеруючи один і той же зашифрований текст для того самого відкритого тексту. Цей рівень шифрування дозволяє серверу виконувати перевірки на рівність, що означає, що він може виконувати вибірку з предикатами рівності, з'єднання за еквівалентністю, GROUP BY, COUNT, DISTINCT і т. д. З погляду криптографії DET має бути псевдовипадковою перестановкою (pseudo-random permutation – PRP).

Шифрування зі збереженням порядку (Order-preserving encryption – OPE) дозволяє встановлювати відносини порядку між елементами даних на основі їх зашифрованих значень, не розкриваючи самих даних. Якщо $x < y$, то $\mathrm{OPE}_K(x) < \mathrm{OPE}_K(y)$ для будь-якого секретного ключа K. Отже, якщо стовпець зашифрований за допомогою OPE, сервер може виконувати запити діапазону за наявності зашифрованих констант $\mathrm{OPE}_K(c_1)$ і $\mathrm{OPE}_K(c_2)$, відповідних діапазону $[c_1, c_2]$. Сервер також може виконувати SQL запити з конструкціями, функціями, що агрегують, ORDER BY, MIN, MAX, SORT тощо. OPE є слабкішою схемою шифрування, ніж DET, оскільки вона розкриває порядок. Таким чином, проксі-сервер CryptDB показуватиме серверу стовпці, зашифровані за допомогою OPE, тільки якщо користувачі запитують запити порядку цих стовпців.

Гомоморфне шифрування (Homomorphic encryption – HE) – це безпечна ймовірна схема шифрування (безпека рівня IND-CPA), що дозволяє серверу виконувати обчислення із зашифрованими даними, при цьому остаточний результат розшифровується на проксі-сервері. Хоча повністю гомоморфне шифрування (fully homomorphic encryption – FHE) дуже повіль-

не, гомоморфне шифрування для певних операцій є ефективним. Зокрема, для підтримки додавання був реалізований метод Пайе (Paillier) [13]. У Пайе множення двох зашифрованих значень призводить до шифрування суми значень, тобто $\mathrm{HE}_K(x)\cdot\mathrm{HE}_K(y)=\mathrm{HE}_K(x+y)$, де множення виконується за модулем деякого значення відкритого ключа. Для обчислення агрегатів SUM проксі-сервер замінює SUM викликами функції користувача (UDF), яка виконує множення Пайе в стовпці, зашифрованому за допомогою HE. HE також можна використовувати для обчислення середніх значень, якщо сервер СУБД повертає суму та кількість окремо, а також для збільшення значень (наприклад, SET *id=id+1*).

З'єднання (JOIN та OPE-JOIN). Для забезпечення рівності між двома стовпцями потрібна окрема схема шифрування, оскільки використовуються різні ключі для DET, щоб запобігти кореляції між стовпцями. JOIN також підтримує всі операції, дозволені DET, а також дозволяє серверу визначати значення, що повторюються між двома стовпцями. OPE-JOIN дає змогу виконувати з'єднання за відносинами порядку.

Пошук слова (SEARCH). SEARCH використовується для пошуку в зашифрованому тексті для підтримки таких операцій, як оператор LIKE в MySQL. Для кожного стовпця, який вимагає пошук, текст розбивається на ключові слова, використовуючи стандартні роздільники (або використовуючи спеціальну функцію вилучення ключових слів). Потім у цих словах видаляються повтори, випадково міняються місцями слова, а потім шифрується кожне зі слів, використовуючи схему [14], доповнюючи кожне слово до однакового розміру. SEARCH майже так само безпечний, як RND: шифрування не повідомляє серверу СУБД, чи повторюється певне слово в декількох рядках, але воно дає витік кількості ключових слів, зашифрованих за допомогою SEARCH. Зловмисник може оцінити кількість окремих або повторюваних слів (наприклад, шляхом порівняння розміру шифртекстів SEARCH і RND для тих самих даних).

Коли користувач виконує такий запит, як:

```
SELECT * FROM messages WHERE msg LIKE "%alice%",
```

проксі передає серверу СУБД токен (лазівку, Trapdoor), що є шифруванням "alice". Сервер не може розшифрувати токен, щоб визначити слово, що лежить в його основі. За допомогою функції, що визначається користувачем, сервер СУБД перевіряє, чи відповідає яке-небудь зашифроване слово в будь-якому повідомленні токену. У цьому підході все, що сервер дізнається в результаті пошуку, це те, чи відповідає токен повідомленню чи ні, і це відбувається тільки для токенів, запрошених користувачем. При цьому слід звернути увагу, що SEARCH дозволяє CryptDB шукати за ключовими словами лише за повним словом. Він не може підтримувати довільні звичайні вирази.

Таким чином, автори CryptDB [12] у своєму рішенні запропонували платформу, яка підтримує SQL-запити до зашифрованих даних. Це рішення ґрунтується на різних рівнях шифрування із збереженням властивостей, таких як детерміноване (DET) та шифрування із збереженням порядку (OPE), що застосовуються до стовпця таблиці SQL. Щоб запросити зашифровану базу даних, CryptDB перетворює незашифрований SQL-запит на його зашифрований еквівалент і розшифровує цільові шари. Основний недолік CryptDB полягає в тому, що щоразу, коли видаляється один шар, схема шифрування стає слабкою [9].

2. MONOMI

MONOMI – система безпечного виконання операцій над конфіденційними даними на ненадійному сервері бази даних [15]. MONOMI працює шляхом шифрування всієї бази даних та виконання запитів до зашифрованих даних. MONOMI була реалізована з використанням PostgreSQL та призначена для виконання аналітичних SQL-запитів.

Існуючі проблеми, які змушена вирішувати MONOMI:

– по-перше, запити до великих наборів даних часто обмежені можливостями системи введення-виводу, наприклад, читання даних із диска або потокове передавання через пам'ять.

В результаті схеми шифрування, які значно збільшують розмір даних можуть уповільнити обробку запитів;

– по-друге, аналітичні запити вимагають складних обчислень, які можуть бути неефективними для виконання над зашифрованими даними. Натомість на практиці повинні використовуватися ефективні схеми шифрування, які можуть виконувати лише певні обчислення (наприклад, використання шифрування із збереженням порядку для сортування та порівняння тощо). Завдання полягає в тому, щоб розділити запит на частини, які можуть бути виконані з використанням доступних схем шифрування на ненадійному сервері та частини, які повинні бути виконані на довіреному клієнті;

– по-третє, деякі методи обробки запитів за зашифрованими даними можуть прискорити одні запити, але уповільнити інші, що вимагає ретельного планування виконання кожного запиту для бази даних, що розглядається, і комбінації запитів.

Способи вирішення зазначених проблем:

а) вводиться роздільне клієнт-серверне виконання складних запитів, при якому виконується максимально можлива частина запиту за зашифрованими даними на сервері, а компоненти запитів, що залишилися, виконуються шляхом надсилання зашифрованих даних довіреному клієнту, який розшифровує дані та обробляє запити у звичайному режимі;

б) вводиться ряд методів, що підвищують продуктивність для певних типів запитів (але не обов'язково для всіх), включаючи попереднє обчислення для кожного рядка, ефективне за простором / економічне шифрування (space-efficient encryption), групове гомоморфне додавання (grouped homomorphic addition) та попередню фільтрацію (pre-filtering);

в) до архітектури додаються проєктувальник для оптимізації фізичного розміщення даних на сервері та планувальник для прийняття рішення про те, як розділити виконання запиту між клієнтом та сервером. Проєктувальник і планувальник необхідні, тому що «жадібне» застосування всіх методів або «жадібне» виконання всіх обчислень на сервері може призвести до надмірних накладних витрат пам'яті та / або неефективного виконання запитів.

Оскільки розробка MONOMI базувалася на CryptDB, вона має аналогічні властивості, що стосуються безпеки. Хоча ненадійний сервер зберігає лише зашифровані дані, він все ж таки може отримати інформацію про вихідні дані у вигляді відкритого тексту трьома способами:

1) деякі схеми шифрування розкривають інформацію, необхідну для обробки запитів (наприклад, детерміноване шифрування виявляє дублікати для перевірки рівності);

2) деякі схеми шифрування можуть пропускати більше інформації, ніж потрібно. Наприклад, схема із збереженням порядку дає частковий витік інформації про відкритий текст;

3) сервер дізнається, які рядки відповідають кожному предикату, обчисленому на сервері. Наприклад, рядки, які відповідають стовпцю LIKE '%keyword%'.

Комбінуючи ці джерела інформації, сервер, який є противником, може отримати додаткову інформацію про рядки або запити за допомогою статистичних методів.

MONOMI ніколи не зберігає текстові дані на сервері і використовує лише схеми шифрування, необхідні для роботи програми. MONOMI дозволяє адміністратору додатково обмежувати схеми шифрування, які використовуються для особливо чутливих стовпців. Наприклад, вимога шифрування із збереженням порядку (найслабша схема MONOMI) не повинна використовуватися для стовпців, у яких зберігаються номери кредитних карток або номерів соціального страхування.

Архітектура MONOMI представлена на рис. 2.

Під час налаштування системи проєктувальник MONOMI запускається на довіреній клієнтській машині та визначає її ефективну фізичну конфігурацію для ненадійного сервера. Щоб визначити основні характеристики робочого навантаження (workload – робочим навантаженням, як правило, є будь-яка програма, яка працює на комп'ютері; сьогодні терміни «робоче навантаження», «застосунок», «програмне забезпечення» та «програма» використовуються як взаємозамінні) для досягнення хорошої продуктивності проєктувальник приймає в

якості вхідних даних репрезентативну підмножину запитів і статистику за даними, наданими користувачем. Користувачі не зобов'язані застосовувати проєктувальник, і замість цього можуть вручну вводити стратегію шифрування або змінювати створену ними стратегію.

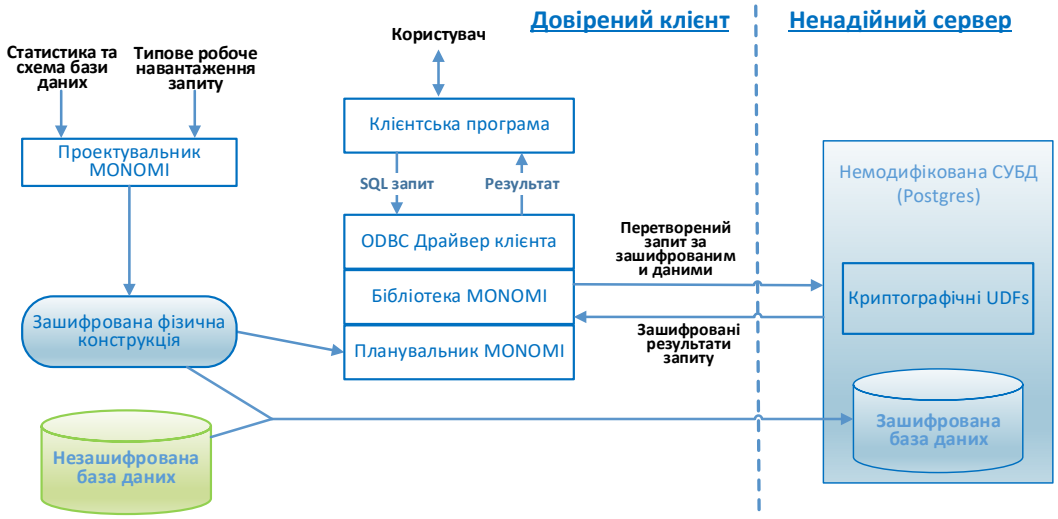


Рис. 2. Загальна архітектура MONOMI

При нормальній роботі програми видають немодифіковані запити SQL за допомогою бібліотеки MONOMI ODBC (Open Database Connectivity), яка є єдиним компонентом, що має доступ до ключів розшифрування. Бібліотека ODBC використовує планувальник, щоб визначити найкращий план виконання запиту з поділом для застосунків клієнт-сервер.

Враховуючи план виконання, бібліотека видає один або кілька запитів до зашифрованої бази даних, яка не має доступу до ключів розшифрування та може виконувати операції лише із зашифрованими даними. База даних запускає немодифіковане програмне забезпечення СУБД з декількома визначеними функціями користувача (UDFs), що надаються MONOMI, які реалізують операції із зашифрованими даними.

MONOMI шифрує всі дані, що зберігаються в базі даних, хоча на практиці неконфіденційні дані можна зберігати як відкритий текст для підвищення ефективності. Після того як клієнтська бібліотека отримає проміжні результати з бази даних, розшифрує їх і виконає всі операції, що залишилися, які не можуть бути ефективно виконані на сервері, результати надсилаються до застосунку, як якщо б вони виконувались у стандартній базі даних SQL.

Схеми шифрування, що використовуються MONOMI, приклади операцій SQL, які вони дозволяють над зашифрованими даними на сервері, та інформація, що розкривається зашифрованими текстами кожної схеми за відсутності будь-яких запитів, наведено у табл. 1.

Таблиця 1

Схеми шифрування, що використовуються у MONOMI

Схема шифрування	**SQL-операції**	**Витік (Leakage)**
Randomized AES + CBC	Жодної	Ні
Deterministic AES + режими: CMC [16] або FFX [17]	a = const, IN, GROUP BY, equi-join	Дублікати
OPE [18]	a > const, MAX, ORDER BY	Order + partial Незашифрований текст [19]
Paillier [13]	a + b, SUM(a)	Ні
SEARCH [12, 14]	шаблон LIKE	Ні

Основні характеристики MONOMI.

– реалізована поверх PostgreSQL;

– працює шляхом шифрування всієї бази даних та виконання запитів до зашифрованих даних;

– бібліотека MONOMI ODBC є єдиним компонентом, що має доступ до ключів розшифрування;

– представляє новий підхід, що базується на роздільному виконанні запиту на клієнт-сервері, що дозволяє виконати частину запиту на ненадійному сервері поверх зашифрованих даних. Для інших елементів запиту MONOMI завантажує проміжні результати на клієнт;

– бібліотека MONOMI ODBC отримує проміжні результати з бази даних, розшифровує їх і виконує всі операції, що залишилися, які не можуть бути ефективно виконані на сервері;

– реалізовано кілька методів, що покращують продуктивність: попереднє обчислення рядка, просторово-ефективне шифрування, групове гомоморфне додавання та попередня фільтрація;

– оскільки ці оптимізації добре працюють для одних запитів та неефективні для інших, MONOMI має планувальник, щоб визначити найкращий спосіб виконання запиту.

3. Seabed

Seabed – система, що забезпечує ефективну аналітику великих зашифрованих наборів даних [20]. Seabed використовує нову адитивно-симетричну схему гомоморфного шифрування (ASHE – additively symmetric homomorphic encryption) для ефективного виконання великомасштабних агрегацій. Крім того, Seabed представляє нову схему рандомізованого шифрування під назвою Splayed ASHE або SPLASHE, яка в деяких випадках може запобігти частотним атакам на основі допоміжних даних. ASHE і базовий варіант SPLASHE задовольняють стандартному поняттю семантичної безпеки (IND-CPA), у той час як розширений варіант SPLASHE доказово не пропускає більше інформації, ніж кількість значень вимірювань, які часто й рідко зустрічаються в базі даних. Прототип Seabed реалізований на базі Apache Spark (уніфікований аналітичний механізм (фреймворк) з відкритим кодом для великомасштабної обробки даних). На відміну від CryptDB та Monomi, система заснована не на реляційній базі даних, а на файловій системі, що призначена для зберігання неструктурованої інформації. Дані зберігаються в HDFS (Hadoop Distributed File System – файлова система, призначена для зберігання файлів великих розмірів, поблоково розподілених між вузлами обчислювального кластера) з використанням серіалізації Google Protobuf (Protobuf – це схема серіалізації, запропонована Google; ця схема не залежить від мови та платформи. Вона підтримує такі мови, як java, c, go, python та деякі інші, а також забезпечує підтримку файлів мультиплатформних бібліотек).

Схема Seabed представлена на рис. 3.

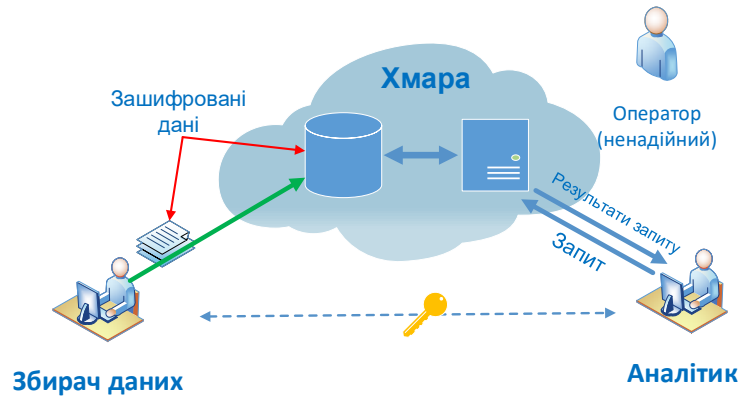


Рис. 3. Схема Seabed

Збирач даних (Data collector) збирає велику кількість даних, шифрує їх та завантажує у хмару, якій не довіряє. Аналітик може генерувати запити для обробника запитів у хмарі. Відповіді будуть зашифровані, але аналітик може розшифрувати їх за допомогою секретного ключа, який є спільним із збирачем даних. Робоче навантаження, яке передбачається підтримувати, складається із запитів у стилі OLAP (OnLine Analytical Processing – інтерактивна аналітична обробка) для великих наборів даних.

Основні компоненти Seabed наведено рис. 4.

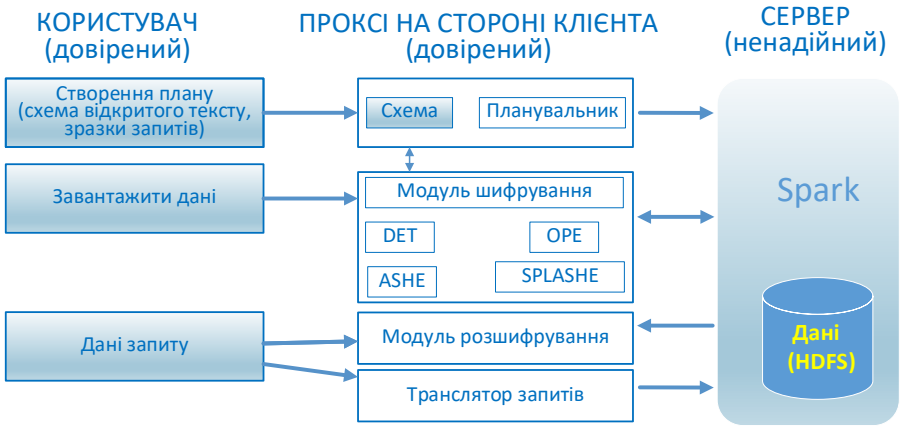


Рис. 4. Основні компоненти Seabed

Користувач взаємодіє з проксі-сервером клієнта Seabed, який працює у довіреному середовищі. Проксі, своєю чергою, взаємодіє з недовіреним сервером Seabed. Як і попередні системи, Seabed приховує від користувачів всі криптографічні операції, тому вони взаємодіють із системою так само, як зі стандартною системою Spark. Користувач може видавати три види запитів:

1. Створення плану (Create Plan). Спочатку користувач надає схему у вигляді відкритого тексту та зразок запиту, заданого для планувальника Seabed. Планувальник використовує їх і спеціальну процедуру визначення схем шифрування стовпців.
2. Завантаження даних (Upload Data). Потім користувач надсилає дані у вигляді відкритого тексту в модуль шифрування Seabed. Дані шифруються за допомогою необхідної схеми шифрування, а записи додаються до таблиці, що зберігається у хмарі. Це безперервний процес.
3. Запит даних (Query Data). Під час аналізу користувач надсилає сценарій запиту транслятору запитів Seabed, який модифікує запити обробки зашифрованих даних перед їх відправкою на сервер. Завдання транслятора запитів – перехоплювати немодифіковані запити клієнта та переписувати їх відповідно до схеми зашифрованого набору даних. При цьому в системі підтримуються принципи, введені в CryptDB та MONOMI. Єдина технічна відмінність від попередніх систем полягає в тому, що вони працюють з реляційними базами даних, тому і вихідною, і цільовою мовою транслятора є SQL. Однак Seabed працює на Spark, тому цільова мова – Scala та Spark API. Сервер виконує запити та відповідає модулю розшифрування проксі-сервера. Після розшифрування та подальшої обробки (якщо така потрібна) результати відправляються назад користувачеві.

Планувальник даних визначає, як шифрувати кожен стовпець у схемі з огляду на список конфіденційних стовпців, складений користувачем. Користувач також надає зразок набору запитів, який використовується планувальником для вибору алгоритму шифрування. Крім того, щоб використовувати розширений SPLASHE, користувач вказує кількість різних значень, які може набувати кожен стовпець, і частотний розподіл цих значень.

Модуль шифрування шифрує відкритий текст.

У разі використання адитивної симетричної схеми гомоморфного шифрування (ASHE) передбачається, що відкриті тексти беруться з адитивної групи $\mathbb{Z}_n = \{0,1,...,n-1\}$. Також

передбачається, що відправник і одержувач, що шифрують і розшифровують зашифрований текст, спільно використовують таємний ключ k, а також псевдовипадкову функцію (PRF) $F_k : I \to \mathbb{Z}_n$, що перетворює ідентифікатор з множини I у випадковий номер із $\mathbb{Z}_n$. Один із можливих варіантів PRF: $F_k = H(i \| k) \bmod n$, для $i \in I$, де H – криптографічна геш-функція (моделюється як випадкова функція), $\|$ позначає конкатенацію, розмір діапазону H кратний n. Іншим варіантом може бути використання AES, коли він використовується як псевдовипадкова перестановка.

На рис. 5 представлений загальний огляд ASHE у контексті Seabed. Наведена на рис. 5 схема відповідає стандартному поняттю семантичної безпеки (CPA).

Транслятор запитів перехоплює немодифіковані запити клієнта і переписує їх відповідно до обраного способу шифрування даних.

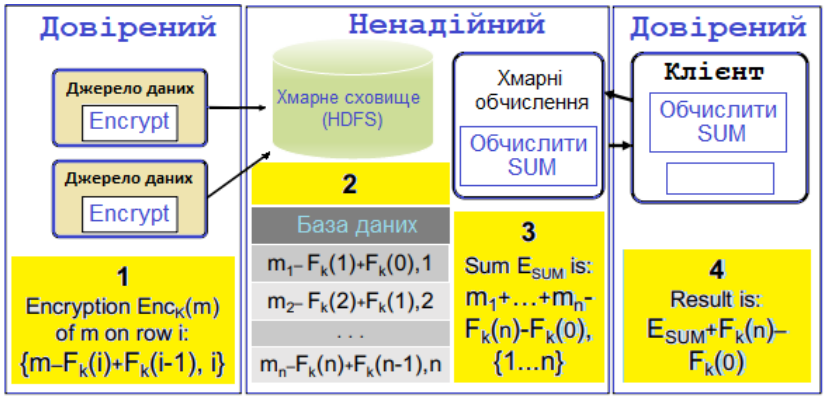


Рис. 5. Основні компоненти Seabed та схема ASHE

Передбачається, що противник є пасивним (чесним, але допитливим), тобто противник намагатиметься дізнатися про конфіденційні дані, але не буде їх активно спотворювати або іншим чином втручатися в роботу системи. Також передбачається, що зловмисник може спробувати провести частотну атаку. Особливо це актуально в тих випадках, коли стовпець може мати невелику кількість значень, а хмара знає, що якесь значення буде найбільш поширеним.

4. Arx

Arx – практично реалізована та багатофункціональна система баз даних, яка шифрує дані за допомогою семантично безпечних схем шифрування [20]. Система Arx реалізована на базі NoSQL СУБД MongoDB.

Передбачається, що сервер БД може бути розміщений у приватній або публічній хмарі.

Модель загроз Arx передбачає, що зловмисник не контролює і не спостерігає за даними або виконанням на стороні клієнта і може отримати доступ лише до сторони сервера, що складається з проксі-сервера Arx та серверів бази даних.

Arx вважає атакуючих сервер зловмисників пасивними (чесними, але допитливими): зловмисники вивчають дані на стороні сервера, щоб зібрати конфіденційну інформацію, але дотримуються протоколу і не змінюють бази даних або результатів запитів. Крім того, у моделі Arx зловмисник не може впровадити будь-які нові запити, оскільки він не має доступу до клієнтської програми або секретних ключів на клієнтському проксі, а лише до сервера. При цьому розглядаються два типи пасивних зловмисників, так званих офлайн та онлайн зловмисників із різними гарантіями для кожного з них.

Так офлайн зловмисник може викрасти одну копію бази даних, що складається із зашифрованих колекцій та індексів. Ця копія не містить даних у пам'яті, пов'язаних із виконанням поточних запитів (які підпадають під дію онлайн-зловмисника). Онлайн-зловмисник – це звичайний пасивний зловмисник: він може реєструвати та відстежувати будь-яку інформацію, доступну на сервері (тобто всі зміни в базі даних, весь стан у пам'яті та всі запити) у

будь-який момент часу протягом будь-якого періоду часу. При цьому Arx приховує параметри в запитах, але не операції, що виконуються.

Arx не покладається на якесь довірене обладнання на сервері. Основне завдання запропонованої архітектури (рис. 6) – не змінювати наявний сервер БД та застосунки, що з ним працюють. Застосунок та система бази даних залишаються незмінними. Натомість Arx вводить два компоненти між застосунком та сервером БД: довірений клієнтський проксі (client proxy) та недовірений проксі-сервер (server proxy).

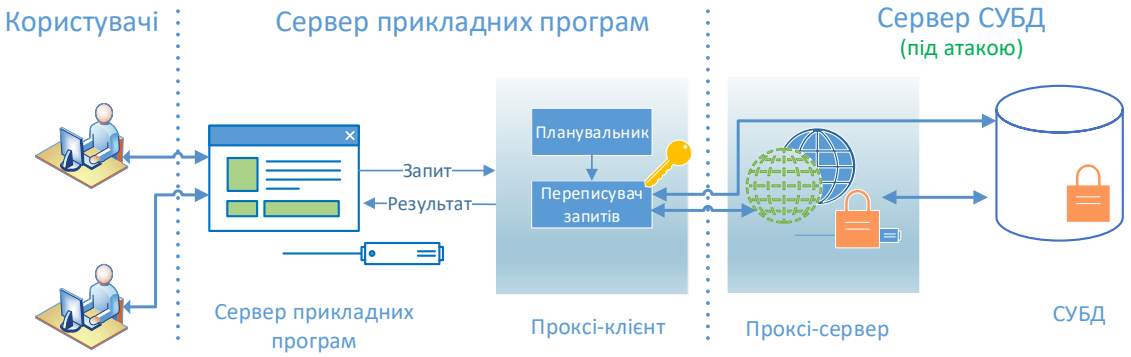


Рис. 6. Архітектура Arx

Довірений клієнтський проксі розгорнутий на сервері застосунків, ненадійний проксі-сервер розгорнутий на сервері СУБД, сервер СУБД розміщений у приватній або публічній хмарі. Клієнтський проксі перехоплює запити та шифрує конфіденційну інформацію. Проксі-сервер підтримує індекси зашифрованих даних та виконує вхідні запити. Сірі прямокутники зображують компоненти, представлені Arx, інші елементи представляють існуючі компоненти. Ключ вказує, що конфіденційні дані у компоненті завжди залишаються суворо зашифрованими.

Слід зазначити, що клієнтський проксі експортує до застосунку той же API, що і сервер БД, тому застосунок не потрібно змінювати. Проксі-сервер взаємодіє із сервером БД, викликаючи його незмінений API (наприклад, надаючи запити). Іншими словами, проксі-сервер веде себе як незмінний клієнт сервера БД. На відміну від CryptDB, Arx не може використовувати функції, що визначаються користувачем замість проксі-сервера, тому що проксі-сервер повинен взаємодіяти з сервером БД кілька разів для кожного запиту клієнта.

Клієнтський проксі зберігає головний (master) ключ. Він переписує запити, шифрує дані та перенаправляє переписані запити на проксі-сервер для виконання разом із допоміжними криптографічними токенами. Останній перенаправляє всі запити без будь-яких конфіденційних полів безпосередньо на сервер БД. Клієнтський проксі легковажний. Він не зберігає БД і робить набагато менше роботи, ніж сервер. Клієнтський проксі зберігає метадані (інформацію про схему), невелику кількість стану та, можливо, кеш. Сервер виконує дорогу частину запитів до БД, фільтруючи та поєднуючи безліч документів у невеликий набір результатів.

Щоб використовувати Arx, адміністратор вказує наступну інформацію під час налаштування системи:

1. Анотовану схему (необов'язково): унікальні поля, поля, які є конфіденційними (якщо є) і розміри полів.
2. Операції, які виконуються з конфіденційними полями.
3. Поля, які мають бути проіндексовані.

По-перше, адміністратор використовує наступний API:

$$collection = \{field_1: info_1, ..., field_n: info_n\},$$

щоб анотувати поля в колекції. Ця анотація необов'язкова, але, якщо вона надана, вона покращує продуктивність Arx. Для *info* слід вказати "*unique*", якщо значення в полі є унікальними, наприклад, SSN (Social Security Number). Arx автоматично визначає первинні ключі як

унікальні. У *info* також може вказуватись максимальна довжина поля, що допомагає Arx вибрати більш ефективну схему шифрування.

За замовчуванням Arx шифрує всі поля в БД. Однак адміністратор може явно перевизначити цю поведінку, вказавши інформацію як неконфіденційну для певного поля. Цю опцію слід використовувати тільки в тому випадку, якщо адміністратор вважає, що це поле не є конфіденційним, і хоче зменшити накладні витрати на шифрування, або якщо Arx не підтримує обчислення в цьому полі, але адміністратор все ще хоче використовувати Arx для інших полів. Однак при цьому слід враховувати, що хоча деякі поля самі по собі не можуть бути конфіденційними, вони можуть призвести до витоку допоміжної інформації про інші поля в базі даних. Отже, адміністратор має обирати такі поля з обережністю.

По-друге, Arx необхідно знати шаблони запитів, які будуть виконуватись у базі даних. Наприклад, для запиту `select * from T where age=10`, Arx необхідно знати, що проводитиметься перевірка на рівність за віком. Тобто Arx необхідно знати, які операції над якими полями виконуються, але не константи, які будуть запитуватись. Адміністратор може вказати ці операції безпосередньо, або надати трасування запуску застосунку, і Arx автоматично їх ідентифікує.

По-третє, Arx повинен знати перелік звичайних індексів, створених застосунком. Arx потрібна ця інформація, щоб забезпечити ті ж самі асимптотичні гарантії продуктивності, що й для незашифрованої бази даних.

Для пошуку даних Arx представляє два нові індекси бази даних:

– Arx-RANGE для запитів із діапазону.

– Arx-EQ для запитів на рівність.

Підтримувані Arx операції: INSERT, DELETE, UPDATE, SELECT (в тому числі з можливістю використання агрегуючих функцій: суми, суму квадратів, мінімуму і максимуму, середнього, агрегування, з мінімальною постобробкою на клієнтському проксі та інших).

Arx збільшує загальний обсяг даних, що зберігаються в базі даних, тому що: зашифровані тексти більше, ніж відкриті тексти для певних схем шифрування, і до документів додаються додаткові поля для виконання певних операцій, таких, наприклад, як перевірки на рівність з використанням EQ або токени для індексації ArxEq. Крім того, індекси ArxRange більші, ніж звичайні дерева B+, тому що кожен вузол у дереві індексів зберігає спотворені схеми. При цьому Arx підтримує менше запитів, ніж CryptDB.

5. CipherSweet

CipherSweet – серверна бібліотека, розроблена Paragon Initiative Enterprises для реалізації шифрування на рівні полів з можливістю пошуку [22]. CipherSweet використовує так зване сліпе індексування зі стратегіями нечіткої фільтрації та фільтрації Блума [23], щоб забезпечити швидкий пошук зашифрованих даних з мінімальним їх витоком. Фільтр Блума – це компактна імовірнісна структура даних, що підтримує запити на членство у множині. Фільтр Блума допускає хибні спрацьовування (твердження, що елемент є частиною множини, коли він таким не є), але ніколи не призводить до хибно негативних результатів (повідомлення про те, що існуючий елемент відсутній у множині). Сліпі індекси не слід плутати з традиційними індексами, введеними для підвищення продуктивності в системах управління базами даних.

Кожен сліпий індекс (аналог розглянутого вище зашифрованого індексу I) у кожному стовпці використовує ключ, відмінний від ключа шифрування та кожного іншого ключа сліпого індексу. Це не дозволяє використовувати оператори LIKE або пошук за звичайними виразами, але дозволяє індексувати перетворення (наприклад, підрядки) відкритого тексту, гешованого за допомогою окремого ключа.

Ідея техніки сліпого індексування, що використовується в CipherSweet, є досить простою. Розглянемо деяку таблицю A (табл. 2) бази даних з атрибутами: *id, field_1, field_2,*

де містяться дані (*data_1, data_2, data_3, data_4, data_5*). Для пошуку відповідних даних по атрибуту *field_1* можна використати, наприклад, наступний простий запит:

```
SELECT * FROM table WHERE field_1 = data_1,
```

результатом якого буде два картежі з *id* рівним 1 та 2.

Таблиця 2

Таблиця *A*

id	***field_1***	***field_2***
1	data_1	data_2
2	data_1	data_3
3	data_4	data_5

Однак при зашифруванні даних сервер не має змоги порівняти *field_1=data_1*. У зв'язку з чим виникає необхідність модифікації вихідної таблиці *A*. А саме потрібно введення додаткового атрибуту – індексу *field_1_index*, в якому будуть міститися так звані метадані, пов'язані з даними атрибуту *field_1* (табл. 3).

Таблиця 3

Таблиця A_1

id	***field_1_index***	***field_1***	***field_2***
1	index_1	data_1	data_2
2	index_1	data_1	data_3
3	index_4	data_4	data_5

Тепер можна зашифровувати відповідні дані стовпця *field_1*, використовуючи певний криптостійкий шифр. CipherSweet використовує автентифіковане шифрування із симетричним ключем та випадковим вектором ініціалізації (IV).

Наприклад, компонент CipherSweet FIPSCrypto надає безпечний інтерфейс AEAD (Authenticated Encryption with Associated Data – автентифіковане шифрування зі зв'язаними даними – клас блокових режимів шифрування, при якому частина повідомлення шифрується, частина залишається відкритою, і все повідомлення повністю автентифіковано), використовуючи FIPS 140-2 (та інші додаткові документи FIPS). Алгоритми FIPSCrypto забезпечують полегшене шифрування даних за допомогою алгоритму AES-256 у режимі лічильника (CTR). Зашифровані тексти автентифікуються за допомогою HMAC-SHA384. IV (вектори ініціалізації) / нонси (nonce – одноразовий номер) генеруються за допомогою стійкого криптографічно генератора псевдовипадкових чисел. Для розділення ключів використовується HKDF-HMAC-SHA384.

При цьому, щоб можна було використовувати операцію порівняння, необхідно виконати ще певні дії. А саме, у відповідні поля атрибуту *field_1_index* необхідно записати результат застосування до даних криптографічно стійкої псевдовипадкової функції (pseudorandom function family – PRF; наприклад, PBKDF2 або Argon2).

Для визначеності позначимо операцію шифрування за допомогою криптостійкого шифру, як Enc/Dec (зашифрування та розшифрування відповідно) та операцію обчислення псевдовипадкової функції PRF. Тоді результат перетворених даних на недовіреному сервері можна подати так (табл. 4).

Таблиця 4

Вміст таблиці A_1 після перетворень

id	***field_1_index***	***field_1***	***field_2***
1	index_1=PRF(data_1)	Enc(data_1)	data_2
2	index_1=PRF(data_1)	Enc(data_1)	data_3
3	index_4=PRF(data_4)	Enc(data_4)	data_5

Для реалізації описаного принципу потрібно: модифікувати запити для роботи з даними та інтегрувати три криптографічні функції Enc, Dec та PRF у логіку програми. При цьому слід зазначити, що алгоритми шифрування та псевдовипадкової функції використовують різні ключі. CipherSweet використовує ряд методів розширення ключів, щоб перетворити один ключ, який обробляється об'єктом Постачальник ключів і може бути отриманий зі сторонніх служб керування ключами, в:

– один окремий ключ шифрування для кожного зашифрованого поля у кожній таблиці;

– множина різних ключів для обчислення сліпих індексів, по одному для кожного індексу у кожному зашифрованому полі кожної таблиці.

На рис. 7 схематично представлено процес розширення ключів.

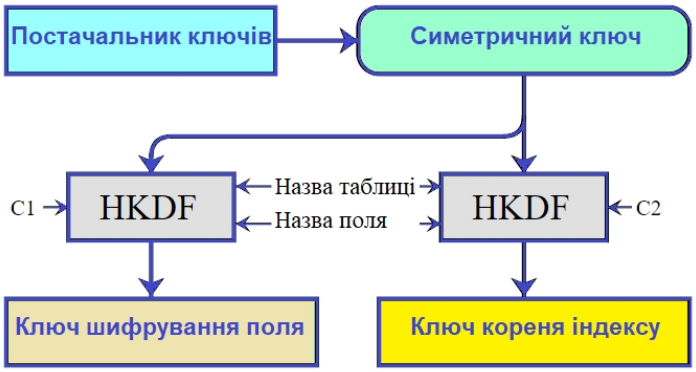


Рис. 7. Процес розширення ключів у CipherSweet

HKDF (HMAC (Hash-based Message Authentication Code) Key Derivation Function) – це проста функція формування ключів (KDF), заснована на коді автентифікації повідомлень HMAC; C1 і C2 – константи, які були обрані таким чином, щоб між ними була відстань Хеммінга 32*4=128b (C1 – це байт `0xB4` повторений 32 рази; C2 – байт `0x7E` повторений 32 рази); Ключ шифрування поля – секретний ключ шифрування даних відповідного атрибуту таблиці; Ключ кореня індексу – кореневий ключ для кожного сліпого індексу. Відповідний ключ кожного індексу обчислюється шляхом використання HMAC-SHA256 імені таблиці, імені атрибута / поля та імені індексу як вихідні дані.

Таким чином, наприклад, запит INSERT може виглядати так:

```
INSERT INTO A1 (id, field_1_index, field_1, field_2)
VALUES (1, PRF(data_1), Enc(data_1), data_2).
```

Запит SELECT для вибірки зашифрованих у стовпці *field_1* даних у цьому випадку перетворюється на такий вид:

```
SELECT * FROM table WHERE field_1_index=PRF(data_1).
```

Після отримання результату запиту програма може скористатися функцією Dec і розшифрувати дані атрибута *field_1* із записів з *id* 1 і 2.

При цьому у контексті безпеки слід відзначити важливу особливість. Результат застосування PRF буде завжди тим самим (значення індексів *field_1_index*) для двох однакових значень даних (табл. 5). Це потенційний витік, тому що сервер на основі аналізу зашифрованих даних та їх індексів, до яких він потенційно має доступ, може визначити, які записи в таблиці містять однакові значення зашифрованих даних за атрибутами, що підтримують пошук. Щоб зменшити негативні наслідки цього витоку, CipherSweet пропонує зберігати на сервері усічене значення індексів (табл. 6).

Таблиця 5

Вміст таблиці A_1 до усічення значення індексів

id	***field_1_index***	***field_1***	***field_2***
1	index_1[64]	Enc(data_1)	data_2
2	index_1[64]	Enc(data_1)	data_3
3	index_4[64]	Enc(data_4)	data_5

Де index_*i*=PRF(data_*i*).

Таблиця 6

Вміст таблиці A_1 після усічення значення індексів

id	***field_1_index***	***field_1***	***field_2***
1	index_1[<64]	Enc(data_1)	data_2
2	index_1[<64]	Enc(data_1)	data_3
3	index_4[<64]	Enc(data_4)	data_5

Припустимо, що псевдовипадкова функція генерує індекси розміром 64 байти. Тоді шляхом запису в таблицю A_1 БД усіченого значення індексу (наприклад, розмірністю менше 64 байт), можна знизити наслідки потенційного витоку схеми, тим самим підвищить безпеку. Однак сервер не зможе однозначно визначати записи, які необхідно включати до результату запиту. Тобто зростає можливість включення до результату запиту зайвих записів. У цьому випадку на застосунок покладається відповідальність фільтрації всіх нерелевантних записів після отримання результату запиту від сервера та розшифрування даних.

Параметр усічення (один із параметрів безпеки) може конфігуруватися користувачем. CipherSweet надає спеціальний планувальник для його розрахунку. Для використання планувальника необхідно оцінити очікувану кількість записів, що включають атрибути із зашифрованими даними, та розподілити дані, що підлягають захисту [24].

Підхід сліпого усічення індексу є одним з найбільш ефективних методів пом'якшення наслідків атаки витоку, оскільки він приховує шаблон пошуку регульованим чином за рахунок точності запиту. Коли сліпий індекс усікається до певної кількості біт, його можна розглядати як фільтр Блуму для пошуку в базі даних.

Основним параметром безпеки сліпих індексів є верхня межа витоку відкритого тексту (середня кількість рядків, що повертаються), яку можна виразити наступним чином [10]:

$$C = R / 2^{-S} \tag{1}$$

де $S = \sum_{i=0}^{n} \min(L_i, K_i)$; n – кількість сліпих індексів; L_i – довжина сліпого індексу (у бітах); K_i – довжина ключа (у бітах); R – низка зашифрованих записів, які використовують ці сліпі індекси. Зазвичай рекомендується підбирати параметри таким чином, щоб $2 \le C < \sqrt{R}$. Якщо $C < 2$, то зловмисник зможе зробити висновки, що деякі відкриті тексти ідентичні, що порушує стандартне поняття безпечної схеми. В іншому випадку, якщо $C > \sqrt{R}$, то буде мати місце занадто багато колізій, що сильно вплине на продуктивність (необхідна продуктивність не буде досягнута).

Насправді безпечною верхньою межею витоку відкритого тексту є максимальна кількість очікуваних збігів. У загальному випадку число n та довжину кожного індексу L_i слід мінімізувати. Чим більше індексів створено, тим більше впевненості набуває зловмисник. У той же час більші індекси корисніші, ніж більш короткі індекси.

Незважаючи на певні переваги та практичну реалізацію можливості здійснення пошуку за зашифрованими даними, CipherSweet підтримує мінімальну функціональність запитів (тільки на рівність – equality) і має відносно низьку продуктивність, хоча й надає високий рівень безпеки.

6. Acra

Acra – це набір інструментів забезпечення безпеки даних протягом усього їх життєвого циклу в сучасних розподілених системах [25]. Acra забезпечує шифрування на рівні застосунків, маскування, токенізацію, контроль доступу, запобігання витоку з бази даних та можливості виявлення вторгнень у зручному, дружньому для розробників пакеті.

Основні функції безпеки, що надаються Acra:

– Криптографічна безпека: дані шифруються під час зберігання та передачі за допомогою шифрування на рівні застосунків та на транспортному рівні з надійною взаємною автентифікацією.

– Шифрування із можливістю пошуку.

– Вибіркове шифрування: є можливість вибору, що, де і як потрібно зашифрувати.

– Маскування даних та токенізація: анонімізація або псевдонімізація даних із збереженням вихідного формату.

– Інструменти управління ключами: гнучке управління переносом / ротацією / відкликанням відповідно до наявних потреб у навантаженні та архітектурою даних.

– Брандмауер запитів SQL: запобігає SQL ін'єкції, блокує несанкціоновані та підозрілі запити.

– Система виявлення вторгнень: виявляє виток даних за допомогою шкідливих записів та інтеграції з SIEM (Security Information and Event Management).

– Моніторинг та реєстрація операцій: є можливість відстеження виконуваних операцій та подій, що контролюють дії з даними.

– Захищене від несанкціонованого доступу ведення журналу аудиту.

– Політики: виразна мова політик, що дозволяє настроювати поведінку Acra у великих інфраструктурах.

Компоненти забезпечення безпеки

– AcraServer також відомий як SQL-проксі, який працює як прозорий проксі-сервер зашифрування / розшифрування з базами даних SQL. Програма не знає, що дані зашифровані до того, як вони потраплять до бази даних, база даних також не знає, що хтось зашифрував дані. AcraServer підтримує шифрування, шифрування з можливістю пошуку, маскування, токенізацію, брандмауер SQL, ведення журналу та ведення журналу аудиту.

– AcraTranslator, також відомий як служба API. Це сервер API, який надає більшість функцій Acra у вигляді HTTP/gRPC API з клієнтськими SDK (software development kit) та захистом трафіку. AcraTranslator не залежить від бази даних і покладає на застосунок відповідальність за фактичне розміщення даних у сховищі. Він підтримує шифрування, створення гешів з можливістю пошуку, маскування, токенізацію, ведення журналу.

– AnyProxy. Це сервер API, який працює між кількома мікросервісами / застосунками, керованими API.

– Клієнтські SDK. Acra надає додаткові SDK для шифрування даних (AcraWriter), для розшифрування даних (AcraReader) або для роботи з AcraTranslator.

Компоненти Acra сумісні з численними реляційними СУБД (MySQL v5.7+, PosgtreSQL v9.4-v11, MariaDB v10.3; Google Cloud SQL, Amazon RDS), сховищами об'єктів та ключ-значення (key-value – KV), хмарними платформами, зовнішніми системами управління ключами (key management systems – KMS), системами балансування навантаження.

Далі детальніше розглянемо можливості Acra у контексті шифрування з можливістю пошуку, заснованого на підході «сліпого індексування», розробленого в рамках проєкту CipherSweet.

На відміну від існуючих рішень, дана система забезпечує суворий поділ обов'язків, що гарантує відсутність витоку криптографічних ключів із застосунку, безпечне зберігання та управління криптографічними ключами, а також набір додаткових функцій безпеки, які відповідають реальним загрозам. На рис. 8 представлений принцип проксі-опосередкованого пошуку зашифрованих даних зі сліпим індексуванням.

Основним компонентом схеми Acra SE є так званий AcraServer, який працює як реверсивний (reverse) проксі (прозорий проксі-сервер шифрування / розшифрування). Він знаходиться між застосунком та базою даних. Застосунок не знає, що дані зашифровані до того, як вони потраплять до бази даних, база даних також не знає, що хтось зашифрував дані (тому цей режим часто називають режимом прозорого шифрування).

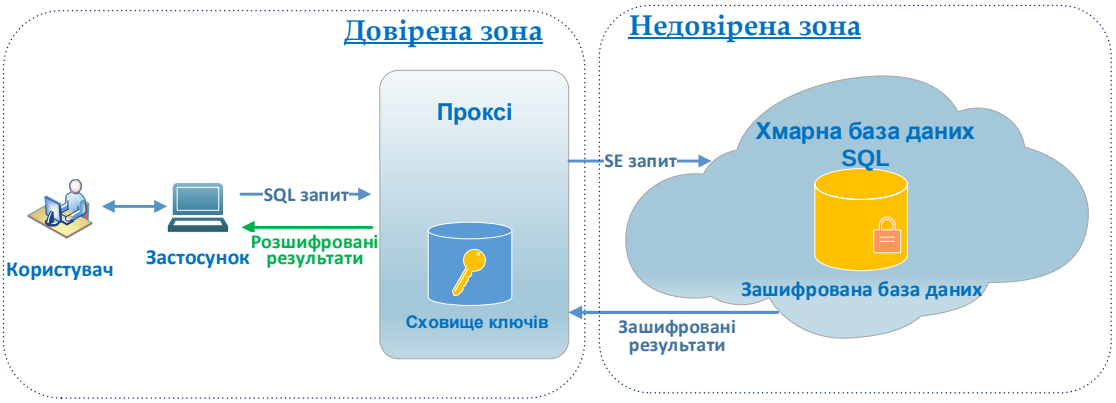


Рис. 8. Використання проксі-сервера між програмою
та ненадійним постачальником сховища

Криптографічна схема обчислення сліпого індексу є алгоритм T, на вхід якого надходять вихідні дані R і симетричний ключ K, на виході – рядок, що повертається: $UT_R = T(R, K)$. До кращих кандидатів для схеми розрахунку сліпого індексу належать алгоритми сімейства HMAC [26, 27].

На рис. 9 показані типові потоки даних (включаючи конфіденційні дані) з різних застосунків (Application 1, 2, …, N) до бази даних і навпаки. Застосунок видає запит, AcraServer виконує всі криптографічні операції (за потреби) і надсилає запит (який може бути змінено) далі до бази даних. База даних обробляє запит та надсилає результат назад. AcraServer отримує результат з бази даних, виконує (за потреби) криптографічні операції та відправляє результат (який також може бути змінений) далі у застосунок.

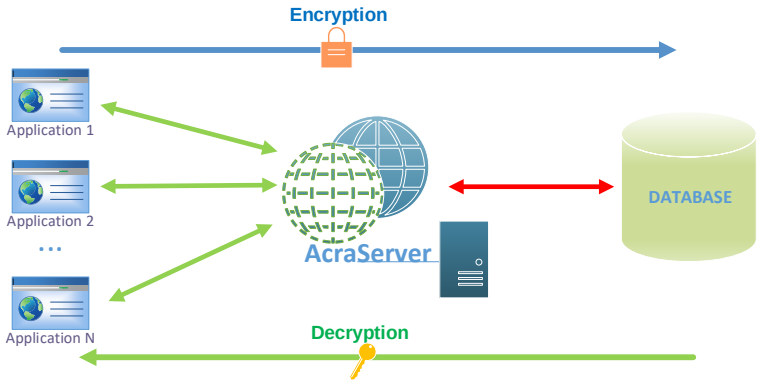


Рис. 9. Потоки даних

Слід зазначити, що один екземпляр AcraServer може працювати з кількома застосунками, але тільки з однією базою даних. AcraServer підтримує два режими роботи:

– стандартний;

– прозорий.

У стандартному режимі шифрування (створення AcraStructs) виконується на стороні застосунку. AcraStructs – це багатоцільовий криптографічний контейнер, у якому зберігаються зашифровані дані (і пов'язані з ними метадані) у певному форматі. AcraStruct може зберігатися як двійковий запис у таблиці бази даних або як великий двійковий об'єкт у файловій системі. Детальніше AcraStruct AS можна представити у вигляді наступної послідовності (масиву байтів) змінної довжини:

$$AS = Tag \,||\, PK_{rand} \,||\, K_{rand}^{PK_A} \,||\, len \,||\, Ciphertext \tag{2}$$

де Tag – спеціальний тег, що відзначає початок AcraStruct; PK_{rand} – відкритий ключ із випадково згенерованої пари ключів ECDH (Elliptic-Curve Diffie–Hellman): (PK_{rand}, SK_{rand}); $K_{rand}^{PK_A} = W(PK_A, SK_A)$ – випадковий симетричний ключ, PK_A – відкритий ключ із пари ключів ECDH (PK_{rand}, SK_{rand}), що належить клієнту A; len – представлення з прямим порядком слідування байтів (масив байтів довжиною 8) цілочисленної змінної, що визначає довжину зашифрованих даних; $Ciphertext = E_{K_{rand}}(R)$ – дані R, зашифровані за схемою AEAD.

У прозорому режимі шифрування виконується на AcraServer. Це дозволяє легко інтегрувати Acra SE в існуючу інфраструктуру без зміни вихідного коду застосунку.

Слід зазначити, що функції шифрування (і безпечного пошуку) AcraServer можна налаштувати кожного стовпця. Це означає, що кожна таблиця в базі даних може бути повністю зашифрована (кожний стовпець), частково (деякі стовпці зашифровані, деякі ні) або повністю не зашифрована.

Безпечний пошук вимагає запровадження двох процедур:

– безпечного завантаження (це модифікація типових запитів INSERT та UPDATE – застосування алгоритму SUpload) зашифрованих даних з індексами у ненадійне сховище;

– безпечного вилучення (це модифікація типового запиту SELECT – застосування алгоритму SSelect) даних з сховища.

Всі властивості безпеки шифрування Acra з можливістю пошуку дуже схожі на властивості безпеки CipherSweet, що створює ризик атак із частково відомим відкритим текстом. У зв'язку з чим автори роботи [10] дають практичні рекомендації для безпеки при використанні інструменту шифрування з можливістю пошуку Acra SE, а саме:

– не створювати сліпі індекси для надзвичайно конфіденційних даних (дані, які не повинні бути розкриті за будь-яку ціну);

– створювати тільки мінімальну кількість сліпих індексів (чим більше індексів – тим більше витоків метаданих);

– якщо дані, які необхідно проіндексувати, надзвичайно чутливі і мають дуже низьку ентропію, щоб їх можна було безпечно помістити в сліпий індекс, їх можна гешувати разом з деякими іншими даними (використання складового індексу);

– сліпий індекс доцільно перетворити до усіченого виду, щоб зменшити витік інформації за рахунок збільшення ймовірності колізій, що призведе до «неправильних» результатів виконання операції SELECT, які мають бути відфільтровані після розшифрування (рекомендується, щоб для будь-якого заданого значення вираз $2 \le C < \sqrt{R}$ завжди залишався істинним (див. вираз (1));

– повинен бути включений безпечний та автентифікований зв'язок між програмою та AcraServer

– та деякі інші.

Однак незважаючи на певні рішення, спрямовані на забезпечення безпеки зберігання та пошуку конфіденційних даних, Acra, як і CipherSweet, яка була взята як прототип схеми шифрування з можливістю пошуку, підтримує мінімальну функціональність запитів, а саме тільки на рівність.

Підбиваючи підсумки, наведемо в табл. 7 деякі узагальнені характеристики [2, 10] деяких систем SE, які мають реалізації та підтримують реляційні бази даних.

Доцільно також відзначити, що CryptDB дозволяє виконати більшість функцій СУБД із втратою продуктивності (накладними витратами) менше ніж 30 % [12]. Для Blind Seer [28] втрата продуктивності більшості запитів становить від 20 до 300 %. У SisoSPIR [33] повідомляється про уповільнення продуктивності на 500 % порівняно з базовою системою MySQL під час перевірки рівності ключових слів (keyword equality) та запитів діапазону (range). Acra SE має уповільнення на два порядки порівняно із звичайним виконанням операцій у PostgreSQL [10]. Найбільший вплив мають криптографічні операції (шифрування в SUpload і розшифрування в SSelect). Крім того, результати запитів SELECT можуть містити нерелевантні рядки при малих розмірах сліпого індексу (1-2 байти), що може призвести до подальшого зниження продуктивності.

Таблиця 7

Порівняльна характеристика деяких систем SE

Система	**Підтримувані операції**	**Додаткові можливості**	**Наявність відкритого вихідного коду**
CryptDB [12]	Equality, Boolean, Range, Sum, Join, Update	Автентифікація користувача, контроль доступу	Так
Blind Seer [28, 29]	Equality, Boolean, Keyword, Range, Update	Політика запитів	Ні
OSPIR-OXT [30, 31, 32]	Equality, Boolean, Keyword, Range, Substring, Wildcard, Update	Політика запитів	Ні
SisoSPIR [33]	Equality, Keyword, Range, Substring	Політика запитів	Ні
CipherSweet [22]	Equality	Управління ключами	Так
Acra [10]	Equality	Автентифікація, політика запитів, виявлення вторгнень, управління ключами, моніторинг та спостереження	Acra: Так Acra SE: Ні

Крім того, слід пам'ятати, що кожній з наведених вище схем властиві різні типи витоку даних. Наприклад, CryptDB є найшвидшим і найпростішим у розгортанні. Однак, як тільки якийсь стовпець БД використовується у запиті, CryptDB розкриває статистику по всьому набору даних для цього стовпця. Blind Seer та OSPIR-OXT також передають інформацію на сервер, але в основному про дані, що повертаються запитом. Таким чином, вони підходять для установок, де запитується невелика частина бази даних.

SisoSPIR підходить, якщо регулярно запитується велика частина даних. Однак SisoSPIR не підтримує логічні запити, що є обмеженням. CipherSweet підтримує лише запити на рівність, але пропонує просту та зрозумілу модель безпеки, яка, по суті, є компромісом між часом та пам'яттю, що створює ризик атак із частково відомим відкритим текстом. Acra заснована на шифруванні CipherSweet з можливістю пошуку, адаптованим до схеми з використанням проксі. Поряд з безпечним пошуком, в Acra забезпечується суворий поділ обов'язків, що гарантує відсутність витоку криптографічного ключа із застосунку, належне управління ключами та додаткові функції безпеки, що відповідають реальним загрозам для конфіденційних даних, що зберігаються у зовнішньому сховищі.

Висновки

1. Технологія захищених баз даних відкриває нові можливості у використанні хмарних сховищ, оскільки вселяє впевненість у власника даних у безпеці його збережених даних, у тому числі за рахунок використання можливостей захищеного пошуку. Захищені системи пошуку криптографічно ізолюють ролі читання, запису та адміністрування бази даних. Цей поділ обмежує непотрібний доступ адміністратора та захищає дані у разі злому системи.

2. Сьогодні пошук у захищених базах даних досяг переломного моменту у своїй зрілості. Однак, незважаючи на велику кількість існуючих академічних досліджень і практичних схем, в даний час немає домінуючого рішення для всіх випадків використання, не існує найбільш захищеної пошукової системи або набору методів. Користувачі повинні розуміти характеристики системи та компроміси для свого варіанта використання. Проєктування таких систем є балансом між безпекою, функціональністю, продуктивністю і зручністю використання. Коли схема покращується в одному аспекті, зазвичай доводиться жертвувати іншими. Ця задача ускладнюється постійною спеціалізацією баз даних, оскільки деяким користувачам потрібні функціональні можливості баз даних SQL, NoSQL або NewSQL. Ця еволюція баз даних продовжуватиметься, і спільнота захищеного пошуку повинна мати можливість швидко надавати функціональні можливості, сумісні з усіма типами баз, сховищ даних.

3. У всіх досить ефективних систем шифрування з можливістю пошуку є загальна проблема – вони пропускають шаблон пошуку, який показує, чи було виконано два пошукові запити по тому самому ключовому слову чи ні. Отже, шаблон пошуку надає інформацію щодо частоти появи кожного запиту. І ця інформація може бути додатково використана за допомогою статистичного аналізу, що дозволяє зловмиснику отримати повну інформацію про ключові слова відкритого тексту, що значно знижує переваги безпеки шифрування даних.

4. Результат аналізу розглянутих систем та практичних рішень показав, що не завжди запропоновані методи шифрування однаково придатні. Більше того, незважаючи на гнучкий вибір шифрування, кількість типів SQL запитів залишається обмеженою.

5. Для здійснення процедур шифрування та розшифрування даних, планування порядку виконання запитів, попередньої обробки даних в архітектуру системи обробки вводяться додаткові компоненти (довірені проксі-сервери, планувальники тощо), що призводить до ускладнення системи, збільшення обсягів необхідної пам'яті та збільшення часу виконання запитів. Усе це змушує проводити подальші дослідження альтернативних підходів задля забезпечення безпечної роботи з віддаленими базами, сховищами даних.

Список літератури;

1. Abadi D., Ailamaki A., Andersen D., Bailis P., Balazinska M., Bernstein P., Boncz P., Chaudhuri S., et al. The Seattle Report on Database Research // ACM SIGMOD Record. 2019. 48. P. 44–53.

2. Fuller B., Varia M., Yerukhimovich A., Shen E., Hamlin A., Gadepally V., Shay R., Mitchell J. D., Cunningham R. K. Sok: Cryptographically protected database search // 2017 IEEE Symposium on Security and Privacy (SP), 2017. P. 172–191. https://doi.org/10.1109/SP.2017.10.

3. General Data Protection Regulation GDPR. URL: https://gdpr-info.eu/ (дата звернення: 12.06.2022).

4. Payment Card Industry (PCI) Data Security Standard. Requirements and Testing Procedures Version 4.0. 2022. URL: https://www.pcisecuritystandards.org/documents/PCI-DSS-v4_0.pdf (дата звернення: 12.06.2022).

5. Atchinson B. K., Fox D. M. From the field: the politics of the health insurance portability and accountability act. Health affairs. 1997. 16(3). P. 146-150.

6. Scholl M., Stine K., Hash J., Bowen P., Johnson A., et al. NIST Special Publication 800-66 Revision 1. An Introductory Resource Guide for Implementing the Health Insurance Portability and Accountability Act (HIPAA) Securi-

ty Rule. 2008. URL: https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-66r1.pdf (дата звернення: 12.06.2022).

7. Bösch, C., Hartel, P., Jonker, W., Peter, A. A survey of provably secure searchable encryption. ACM Computing Surveys (CSUR). 2014. 47(2). P. 1–51.

8. Єсін В. І., Вілігура В. В. Дослідження основних методів і схем шифрування з можливістю пошуку // Радіотехніка. 2022. № 209. С. 138–155.

9. Azraoui M., Önen M., Molva R. Framework for Searchable Encryption with SQL Databases. CLOSER. 2018. P. 57–67.

10.Pilyankevich E., Kornieiev D., Storozhuk A. Proxy-Mediated Searchable Encryption in SQL Databases Using Blind Indexes. Cryptology ePrint Archive. 2019.

11.Hacigümüş H., Iyer B., Li C., Mehrotra S. Executing SQL over encrypted data in the database-service-provider model // Proceedings of the 2002 ACM SIGMOD international conference on Management of data. 2002. P. 216–227. https://doi.org/10.1145/564691.564717.

12.Popa R. A., Redfield C. M., Zeldovich N., Balakrishnan H. CryptDB: protecting confidentiality with encrypted query processing // Proceedings of the Twenty-Third ACM Symposium on Operating Systems Principles. SOSP '11. 2011. P. 85–100. https://doi.org/10.1145/2043556.2043566.

13.Paillier P. Public-Key Cryptosystems Based on Composite Degree Residuosity Classes // Stern, J. (eds) Advances in Cryptology - EUROCRYPT ’99. EUROCRYPT 1999. Lecture Notes in Computer Science, 1999. Vol 1592. P. 223–238. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-48910-X_16.

14.Song D. X., Wagner D., Perrig A. Practical techniques for searches on encrypted data // Proceeding 2000 IEEE symposium on security and privacy. S&P 2000. IEEE, 2000. P. 44–55. https://doi.org/10.1109/SECPRI.2000.848445.

15.Tu S. L., Kaashoek M. F., Madden S. R., Zeldovich N. Processing analytical queries over encrypted data // Proceedings of the VLDB Endowment. 2013. 6(5). P. 289–300. https://doi.org/10.14778/2535573.2488336.

16.Halevi S., Rogaway P. A Tweakable Enciphering Mode // Boneh, D. (eds) Advances in Cryptology - CRYPTO 2003. CRYPTO 2003. Lecture Notes in Computer Science, 2003. Vol 2729. P. 482–499. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-45146-4_28.

17.Bellare M., Rogaway P., Spies T. Addendum . The FFX mode of operation for format-preserving encryption // A parameter collection for enciphering strings of arbitrary radix and length, Draft 1.0, NIST. 2010. URL: https://csrc.nist.gov/CSRC/media/Projects/Block-Cipher-Techniques/documents/BCM/proposed-modes/ffx/ffx-spec2.pdf.

18.Boldyreva A., Chenette N., Lee Y., O’Neill A. Order-Preserving Symmetric Encryption // Joux, A. (eds) Advances in Cryptology - EUROCRYPT 2009. EUROCRYPT 2009. Lecture Notes in Computer Science, 2009. Vol 5479. P. 224–241. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-01001-9_13.

19.Boldyreva, A., Chenette, N., O’Neill, A. Order-Preserving Encryption Revisited: Improved Security Analysis and Alternative Solutions. In: Rogaway, P. (eds) // Advances in Cryptology – CRYPTO 2011. CRYPTO 2011. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 2011. Vol. 6841. P. 578–595. https://doi.org/10.1007/978-3-642-22792-9_33.

20.Papadimitriou A., Bhagwan R., Chandran N., Ramjee R., Haeberlen A., Singh H., Modi A., Badrinarayanan S. Big data analytics over encrypted datasets with seabed // 12th USENIX symposium on operating systems design and implementation (OSDI 16). 2016. P. 587–602.

21.Poddar R., Boelter T., Popa R. A. Arx: an encrypted database using semantically secure encryption // Proceedings of the VLDB Endowment. 12(11). 2019. P. 1664–1678. https://doi.org/10.14778/3342263.3342641.

22.CipherSweet. URL: https://ciphersweet.paragonie.com/ (дата звернення: 12.06.2022).

23.Tarkoma S., Rothenberg C. E., Lagerspetz E. Theory and practice of bloom filters for distributed systems // IEEE Communications Surveys & Tutorials. 2011. 14(1). P. 131–155.

24.Blind Index Planning. URL: https://ciphersweet.paragonie.com/node.js/blind-index-planning. (дата звернення: 12.06.2022).

25.Cossack Labs Knowledge Base. Acra in a nutshell. URL: https://docs.cossacklabs.com/acra/ (дата звернення: 12.06.2022).

26.Bellare M., Canetti R., Krawczyk H. Keying Hash Functions for Message Authentication. In: Koblitz, N. (eds) // Advances in Cryptology - CRYPTO ’96. CRYPTO 1996. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 1996. Vol 1109. P. 1–15. https://doi.org/10.1007/3-540-68697-5_1.

27.Turner J. M. The keyed-hash message authentication code (HMAC) // Federal Information Processing Standards Publication. 2008. 198(1). P. 1–13.

28.Pappas V., Krell F., Vo B., Kolesnikov V., Malkin T., Choi S. G., Bellovin S. Blind seer: A scalable private DBMS // 2014 IEEE Symposium on Security and Privacy. IEEE, 2014. P. 359-374.

29.Fisch B. A., Vo B., Krell F., Kumarasubramanian A., Kolesnikov V., Malkin T., Bellovin S. M. Malicious-client security in blind seer: a scalable private DBMS // 2015 IEEE Symposium on Security and Privacy. 2015. P. 395–410.

30.Cash D., Jarecki S., Jutla C., Krawczyk H., Roşu MC., Steiner M. Highly-Scalable Searchable Symmetric Encryption with Support for Boolean Queries. In: Canetti, R., Garay, J.A. (eds) // Advances in Cryptology – CRYPTO

2013. CRYPTO 2013. Lecture Notes in Computer Science. 2013. Vol. 8042. P. 353–373. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-40041-4_20.

31.Jarecki S., Jutla C., Krawczyk H., Rosu M., Steiner M. Outsourced symmetric private information retrieval // Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security. 2013. P. 875–888.

32.Faber S., Jarecki S., Krawczyk H., Nguyen Q., Rosu M., Steiner M. Rich Queries on Encrypted Data: Beyond Exact Matches. In: Pernul, G., Y A Ryan, P., Weippl, E. (eds) Computer Security – ESORICS 2015. ESORICS 2015. Lecture Notes in Computer Science. 2015. Vol 9327. P. 123–145. Springer, Cham. https://doi.org/10.1007/978-3-319-24177-7_7.

33.Ishai Y., Kushilevitz E., Lu S., Ostrovsky R. Private Large-Scale Databases with Distributed Searchable Symmetric Encryption // Sako, K. (eds) Topics in Cryptology - CT-RSA 2016. CT-RSA 2016. Lecture Notes in Computer Science. 2016. Vol. 9610. P. 90–107. Springer, Cham. https://doi.org/10.1007/978-3-319-29485-8_6.

Надійшла до редколегії 15.09.2022

Відомості про авторів:

Есін Віталій Іванович – д-р техн. наук, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, Україна; e-mail: v.i.yesin@karazin.ua; ORCID: https://orcid.org/0000-0003-1977-7269

Вілігура Владислав Вікторович – аспірант кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Харківський національний університет імені В.Н. Каразіна, Україна; e-mail: viligura93@gmail.com; ORCID: https://orcid.org/0000-0002-1137-2382

УДК 004.056.5 **DOI:10.30837/rt.2022.3.210.05**

М.В. ЄСІНА, канд. техн. наук, Є.В. ОСТРЯНСЬКА, І.Д. ГОРБЕНКО, д-р техн. наук

СТАН ТРЕТЬОГО РАУНДУ ПРОЦЕСУ СТАНДАРТИЗАЦІЇ ПОСТКВАНТОВОЇ КРИПТОГРАФІЇ NIST

Вступ

В останні роки спостерігається стійкий прогрес у створенні квантових комп'ютерів. У разі реалізації великомасштабних квантових комп'ютерів вони будуть загрожувати безпеці багатьох широко використовуваних криптосистем з відкритим ключем. Схеми встановлення ключів і цифрові підписи, що засновані на факторизації, дискретних логарифмах і криптографії на еліптичних кривих, найбільш сильно постраждають. Симетричні криптографічні примітиви, такі як блокові шифри і геш-функції, будуть порушені незначно. Внаслідок цього було активізовано дослідження щодо пошуку криптосистем на відкритих ключах, які були б захищені від криптоаналітиків як з квантовими, так і з класичними комп'ютерами. Цю область часто називають постквантовою криптографією (PQC), або іноді квантово-стійкою криптографією. Її мета полягає в розробці схем, які можна розгорнути в існуючих комунікаційних мережах та протоколах без суттєвих змін.

Національний інститут стандартів і технологій знаходиться в процесі вибору одного або декількох криптографічних алгоритмів з відкритим ключем за допомогою відкритого конкурсу. Нові стандарти криптографії з відкритим ключем визначатимуть одну або кілька додаткових цифрових підписів, шифрування з відкритим ключем і алгоритми встановлення ключів. Передбачається, що ці алгоритми будуть здатні добре захищати конфіденційну інформацію в недалекому майбутньому, в тому числі після появи квантових комп'ютерів.

Після багаторічного огляду кандидатів NIST вибрав 26 алгоритмів для переходу до 2-го раунду оцінки у січні 2019 року [1]. Ці алгоритми розглядалися як найбільш перспективні кандидати для можливої стандартизації і були обрані на основі як внутрішнього аналізу, так і відгуків спільноти. Під час 2-го раунду ці кандидати були піддані більш детальному аналізу з боку NIST і більш широкого криптографічного співтовариства. Після ретельного обговорення NIST вибрав сім фіналістів та вісім альтернативних варіантів, щоб перейти до 3-го раунду в липні 2020 року [2]. Намір NIST полягав у стандартизації невеликої кількості фіналістів наприкінці 3-го раунду, а також невеликої кількості альтернативних кандидатів після 4-го раунду.

3-й раунд розпочався в липні 2020 року і тривав приблизно 18 місяців. Під час 3-го раунду відбувся більш ретельний аналіз теоретичних та емпіричних доказів, що використовуються для обґрунтування безпеки кандидатів. Також проводилось ретельне оцінювання їх продуктивності, використовуючи оптимізовані реалізації на різних програмних та апаратних платформах.

Після трьох раундів оцінки та аналізу NIST вибрав перші алгоритми, які він стандартизує в результаті процесу стандартизації PQC.

Метою цієї статті є огляд та аналіз стану оцінювання та відбору процесу стандартизації постквантової криптографії NIST. У звіті узагальнено кожен із 15 алгоритмів-кандидатів 3-го раунду та визначено обрані для стандартизації, а також ті, які продовжуватимуть оцінюватись у 4-му раунді аналізу. Алгоритм шифрування на відкритому ключі та встановлення ключа, який буде стандартизовано, – Crystals-Kyber. Цифрові підписи, що будуть стандартизовані, – Crystals-Dilithium, Falcon та Sphincs+. Незважаючи на те, що обирається декілька алгоритмів підпису, NIST рекомендує Crystals-Dilithium як основний алгоритм, що повинен бути реалізований. Крім того, чотири альтернативні алгоритми-кандидати встановлення ключа проходять до 4-го раунду оцінювання: BIKE, Classic McEliece, HQC та SIKE. Ці кандидати все ще розглядаються для майбутньої стандартизації [3].

1. Критерії оцінювання та процес відбору кандидатів

NIST обрав 15 алгоритмів-кандидатів для 3-го раунду. Сім з п'ятнадцяти алгоритмів були обрані у якості алгоритмів-фіналістів, в той час як інші вісім були позначені як «альтернативні варіанти» [2]. Набір фіналістів включав алгоритми, які NIST вважав найбільш перспективними, такими, що відповідають більшості випадків використання, і найімовірніше, що будуть готові до стандартизації незабаром після закінчення 3-го раунду. Альтернативні кандидати вважалися потенційними кандидатами для майбутньої стандартизації, швидше за все, після чергового раунду оцінки. Деякі з альтернативних кандидатів мають гірші характеристики ефективності, ніж фіналісти, але можуть бути вибрані для стандартизації на основі високої впевненості NIST у їх безпеці. Інші мають прийнятну ефективність, але потребують додаткового аналізу чи іншої роботи, щоб забезпечити достатню гарантію їх безпеки для стандартизації NIST. Крім того, деякі альтернативні кандидати були обрані на основі прагнення NIST до різноманітності в майбутніх постквантових стандартах безпеки, або на їх потенціалі для подальшого вдосконалення.

Сім фіналістів включали у себе чотири механізми інкапсуляції ключів (KEM) та три механізми цифрового підпису. З восьми альтернативних варіантів п'ять – KEM та три – цифрові підписи. Командам подання було дозволено внести незначні модифікації та повторно подати свої пакети, які повинні були відповідати тим же вимогам, що і оригінальні подання. Повні оновлені технічні характеристики були розміщені на веб-сайті PQC NIST [3] 23 жовтня 2020 року для публічного огляду.

Таблиця 1

Фіналісти третього раунду

Шифрування на відкритому ключі/KEM	Цифрові підписи
Classic McEliece	Crystals-Dilithium
Crystals-Kyber	Falcon
NTRU	Rainbow
Saber	

Таблиця 2

Альтернативні кандидати третього раунду

Шифрування на відкритому ключі/KEM	Цифрові підписи
BIKE	GeMSS
FrodoKEM	Picnic
HQC	SPHINCS+
NTRU Prime	
SIKE	

1.1. Критерії оцінювання

У Call for Proposals NIST [4] визначено три широкі аспекти критеріїв оцінки, які будуть використовуватися для порівняння відповідних алгоритмів в процесі стандартизації PQC NIST: 1) безпека, 2) вартість і продуктивність та 3) характеристики алгоритму і реалізації. Ці критерії описані нижче разом з обговоренням того, як вони вплинули на оцінювання кандидатів у 2-му раунді.

Як і у випадку з минулими конкурсами Розширений стандарт шифрування (AES) і Безпечний алгоритм гешування 3 (SHA-3), безпека є найбільш важливим фактором, що NIST використовує при оцінці кандидатів на постквантові алгоритми. Нинішні стандарти з відкритим ключем NIST використовуються у самих різних додатках, включаючи Інтернет-протоколи, такі як TLS, SSH, IKE, IPsec і DNSSEC, а також для сертифікатів, підпису програмного коду і безпечних завантажувачів. Нові стандарти NIST на відкритому ключів забезпечать постквантову безпеку для кожного з цих додатків.

Для кількісної оцінки безпеки можливих алгоритмів NIST дав три можливі визначення безпеки – два для шифрування і одне для підпису. NIST також визначив п'ять категорій безпеки для класифікації обчислювальної складності атак, які порушують визначення безпеки (див. [5]).

NIST також згадував інші бажані властивості безпеки, такі як пряма безпечність, стійкість до атак бічними каналами та багатоключових атак, а також стійкість до неправильного використання. У деяких випадках NIST закликає представників внести незначні зміни, щоб забезпечити або вдосконалити ці додаткові бажані властивості безпеки (наприклад, додавання відкритої солі до шифртекстів, щоб уникнути багатоцільових атак на KEM).

Що стосується схем шифрування загального призначення і встановлення ключів, то у Call for Proposals [5] запрошувалися «семантично безпечні» схеми щодо атаки на основі адаптивно вибраного шифртексту (еквівалентно безпеці IND-CCA2). Для одноразових випадків використання NIST також приймав алгоритми, що забезпечують семантичну безпеку щодо атаки на основі вибраного відкритого тексту (безпека IND-CPA). IND-CCA2 безпека не потрібна в строго одноразових випадках використання, і спроба задовольнити більш суворі вимоги IND-CCA2 безпеки може спричинити за собою значні втрати продуктивності для деяких схем. Схеми цифрового підпису повинні були забезпечити екзистенційно непідробні підписи стосовно атаки на основі адаптивно вибраного повідомлення (EUF-CMA безпека). Автори заохочувалися, але не були зобов'язані надавати докази безпеки у відповідних моделях.

П'ять категорій безпеки, визначені у [5], були засновані на обчислювальних ресурсах, необхідних для виконання певних атак методом перебору проти існуючих стандартів NIST для AES і SHA в різних моделях вартості обчислень, як класичних, так і квантових.

1.2. Порівняльний аналіз кандидатів 3-го раунду

Що стосується вартості та продуктивності, то початковий запит пропозицій [5] визначав вартість як другий за важливістю критерій при оцінці алгоритмів-кандидатів. Вартість включає в себе обчислювальну ефективність генерації ключа і операцій з відкритим і особистим ключем, витрати на передачу відкритих ключів і підписів або шифртекстів, а також витрати на реалізацію в термінах RAM (оперативної пам'яті) або підрахунку гейтів.

При порівнянні загальної ефективності алгоритмів були розглянуті як обчислювальні витрати, так і витрати на передачу даних. Для використання загального призначення оцінка загальної ефективності розглядалася як витрати на передачу відкритого ключа на додаток до підпису або шифртексту під час кожної транзакції. Для KEM також враховували вартість генерації ключа, оскільки багато додатків використовують нову ключову пару KEM для кожної транзакції для забезпечення прямої секретності. Для алгоритмів підпису вартість генерації ключів вважалася менш важливою.

На рис. 1 показані числа обчислювальної продуктивності з [6] для процесора x86-64 з розширеннями AVX2 для Kyber, NTRU та Saber для категорій безпеки 1 та 3. На рис. 2 показані загальні витрати для Kyber, NTRU та Saber, коли додається вартість передачі даних. Рис. 2 було створено за допомогою орієнтовної вартості 2000 циклів/байт.

Інкапсуляція та декапсуляція є дуже швидкою з усіма трьома схемами. Незважаючи на те, що Saber має найнижчу загальну вартість завдяки меншим відкритим ключам та шифротекстам, різниця у вартості між Kyber та Saber не була достатньо великою, щоб вважатися значною.

Вартість генерації ключа для ntruhps2048677 або ntruhrss701 є приблизно в 11 разів більшою, ніж для KYBER512. Однак, як показано на рис. 2, загальна вартість використання цих схем, як правило, домінує у витратах передачі даних, і тому більша частина різниці в загальній вартості наборів параметрів NTRU порівняно з Kyber та Saber – через дещо більші відкриті ключі та шифротексти NTRU. Як результат, загальна вартість ntruhps2048677 менше, більш ніж на 30 %, ніж для KYBER512. Крім того, оскільки відкриті ключі та шифро-

тексти для наборів параметрів категорії 1 та 3 для всіх трьох схем, ймовірно, вписуються в один Інтернет-пакет, їхні числа продуктивності можуть вважатися порівнянними. Можна також зазначити, що, згідно з [6], вартість генерації ключів для ntruhps2048677 або ntruhrss701 порівнянна з витратами на генерацію ключа для криптографії на еліптичних кривих при кривій P-256, яка широко використовується для обміну тимчасовими ключами.

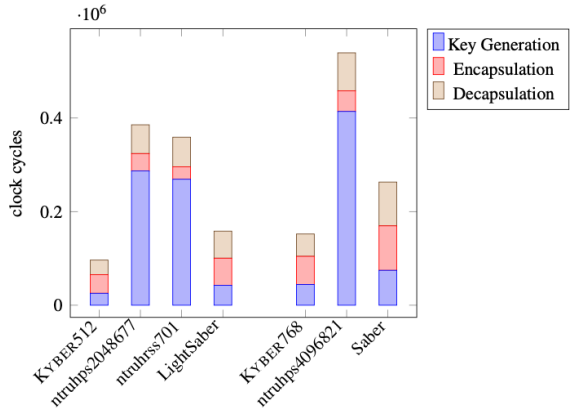


Рис. 1. Порівняльний аналіз KEM на процесорах x86-64 з розширеннями AVX2

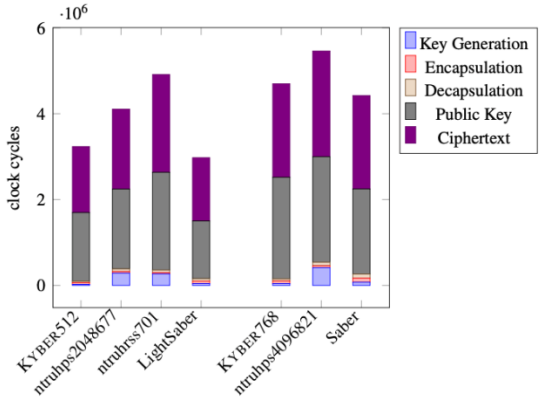


Рис. 2. Порівняльний аналіз KEM на процесорах x86-64
з розширеннями AVX2 з 2000 циклів/байт витратами на передачу

Звіт про стан 2-го раунду вибрав Dilithium і Falcon як фіналістів схеми підпису загального призначення та вказав на намір обрати щонайбільше одного з них [2]. Третій фіналіст, Rainbow, хоча й має привабливий профіль продуктивності для додатків, що вимагають малих підписів або швидкої перевірки, зазнав втрат безпеки, таким чином, показники ефективності Rainbow будуть далі опущені.

На рис. 3 показано показники обчислювальної продуктивності з [6] для процесора x86-64 з розширеннями AVX2 для Dilithium і Falcon. На відміну від рис.1, цей рисунок не включає вартість генерації ключів, оскільки ключі підпису не генеруються на основі кожної транзакції. На рис. 4 показані «загальні витрати» для Dilithium та Falcon з урахуванням вартості передачі відкритого ключа та підпису. При використанні процесора x86-64 генерація підпису за допомогою Dilithium відбувається трохи швидше, ніж за допомогою Falcon. Однак у загальних витратах на використання цих схем переважає передача даних, тому загальна вартість Falcon нижча через менший розмір відкритого ключа та підпису. Для більшості додатків, які використовують процесор x86-64 або подібний, показники продуктивності для Dilithium або Falcon мають бути прийнятними. Однак, на відміну від підписів Falcon, підписи Dilithium не можуть поміститися в один Інтернет-пакет, тому адаптація деяких додатків для використання Dilithium може виявитися складнішою, ніж їх адаптація для використання Falcon (наприклад, [7, 8]).

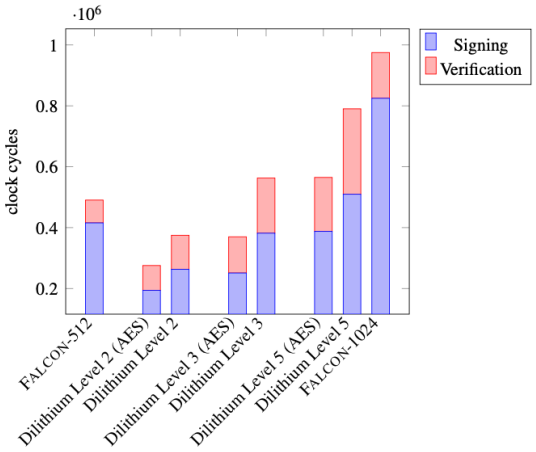


Рис. 3. Порівняльний аналіз підпису на процесорах x86-64
з розширеннями AVX2

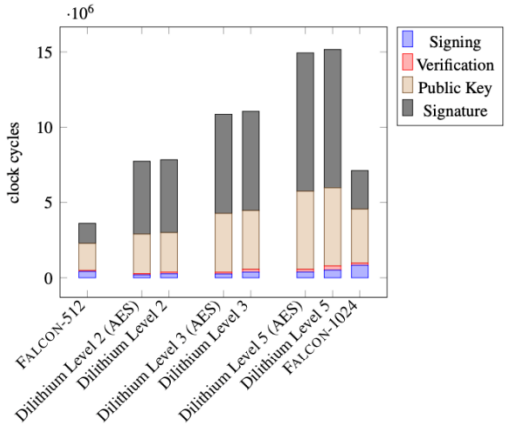


Рис. 4. Порівняльний аналіз підпису на процесорах x86-64
з розширеннями AVX2 з 2000 циклів/байт витратами на передачу

2. Попередня інформація щодо моделей та визначення безпеки

У цьому розділі представлені деякі складні обчислювальні проблеми, які є загальними для багатьох схем на основі кодів, багатовимірних схем або схем на решітках, досліджених у процесі стандартизації NIST PQC. Інші складні обчислювальні проблеми будуть згадані за потреби в описі окремих кандидатів у розд. 4.

2.1. На основі коду

Складність проблем загального та синдромного декодування (і деяких їх варіантів) є складовою аргументу безпеки для трьох KEM на основі коду, які проходять до 4-го раунду: BIKE, Classic McEliece і HQC. Усі три схеми забезпечують IND-CPA безпеку PKE з доказами, які залежать від однієї з цих двох обчислювальних проблем.

Нехай $C\_(n,k)$ двійковий лінійний код. Нехай $\mathbb{F}_2$ позначає скінченне поле двох елементів. Тоді набір із 2^k кодових слів C утворює k-вимірний підпростір $\mathbb{F}_2^n$. Для будь-якого вектора $v \in \mathbb{F}_2^m$, $m \in \mathbb{N}$ нехай $|v|$ позначатиме вагу Хеммінга v.

П р о б л е м а 1. (Проблема декодування синдрому (вирішення)). Дано $(n-k)\times n$ матрицю перевірки парності H для C, вектор $y \in \mathbb{F}_2^{n-k}$ і ціль $t \in \mathbb{N}$, визначити, чи існує $x \in \mathbb{F}_2^n$, що задовольняє $Hx^T = y$ і $|x| \le t$.

П р о б л е м а 2. (Проблема пошуку кодового слова (вирішення)). Дано $(n-k)\times n$ матрицю перевірки парності H для C і ціль $w\in\mathbb{N}$, визначити, чи існує $x\in\mathbb{F}_2^n$, який задовольняє $Hx^T=0$ і $|x|=w$.

Для загального двійкового лінійного коду C Berlekamp, McEliece та van Tilborg показали, що ці дві проблеми є NP-повними [9]. Це не гарантує, що будь-який даний криптографічний екземпляр проблеми є складним.

Найефективніші відомі атаки проти KEM на основі коду базуються на декодуванні набору інформації (ISD). Цей підхід ігнорує структуру двійкового коду та прагне відновити вектор помилки на основі його низької ваги Хеммінга. Ці методи виникли з алгоритму Прангe в 1962 році [10] і з тих пір зазнали низки вдосконалень. Також вивчалися квантові версії алгоритмів ISD [11 – 14]. Ці результати представляють загальне прискорення класичних алгоритмів ISD на основі Гровера та вказують на те, що ISD можна прискорити майже так само, як і пошук грубою силою.

2.2. На основі багатовимірних перетворень

Аргументи безпеки для двох багатовимірних схем підпису, GeMSS і Rainbow, залежать від складності проблеми *MQ* і проблеми MinRank.

П р о б л е м а 3. (Багатовимірна квадратична (*MQ*) поліноміальна проблема). Дано скінченне поле $\mathbb{F}$ і систему з m квадратичних поліномів з n змінними x_i:

$$f_k(x_1,\ldots,x_n)=\sum_{1\le i\le j\le n}a_{ij}^{(k)}x_ix_j+\sum_{1\le i\le n}b_i^{(k)}x_i+c^{(k)}=0, \tag{1}$$

для k від 1 до m, де $a_{ij}^{(k)}$, $b_i^{(k)}$, $c^{(k)}$ усі в $\mathbb{F}$, визначити, чи існує розв'язок у $\mathbb{F}^n$.

П р о б л е м а 4. (Проблема MinRank (вирішення)). Дано скінченне поле $\mathbb{F}$, k матриць M_i розміром $m\times n$ із записами в $\mathbb{F}$ та обмеження рангу r, визначити, чи існують значення $c_i\in\mathbb{F}$, які задовольняють наступне рівняння:

$$rank\left(\sum_{i=1}^{k}c_iM_i\right)\le r. \tag{2}$$

Проблема *MQ* була NP-складною в усіх полях [15]. У [16] показано, що проблема MinRank є NP-складною. Важливо відзначити, що коли цільовий ранг r фіксований, проблема MinRank має поліноміальну складність; таким чином, багатовимірні криптосистеми зазвичай вимагають великого значення r для будь-якого пов'язаного екземпляра MinRank. Відомо, що жодна проблема не є складною в середньому випадку, а також NP-складність не означає, що випадки, які виникають із криптографічних схем, нерозв'язні.

Найефективніші відомі загальні атаки на проблему *MQ* включають алгоритми базису Гробнера, такі як F4/F5, див. [17, 18], і алгоритми лінеаризації, такі як XL, див. [19]. Найефективніші атаки для MinRank відрізняються залежно від розміру та кількості матриць і цільового рангу. Основні методи включають комбінаторні методи пошуку, вперше введені в [20], і метод опорних міноpів, див. [21].

2.3. На основі алгебраїчних решіток

7 із 15 кандидатів 3-го раунду є криптосистемами на основі решітки. Ці криптосистеми пов'язані з великою кількістю академічних досліджень, які наголошують на (асимптотичній) доказовій безпеці, заснованій на найгіршому сценарії складності проблем решітки. Ранньою віхою в цьому напрямку досліджень стала стаття 1996 року Ajtai [22], яка визначила проблему короткого цілого розв'язку (SIS) і пов'язала її середню складність із найгіршою складніс-

тю пошуку коротких векторів у кожній цілочисельній решітці, даючи односторонні функції на основі решітки та односторонні функції з секретом на основі решітки.

Нижче коротко описано різні базові проблеми безпеки для кожної з цих систем:

П р о б л е м а 5. (Проблема короткого цілого розв'язку ($SIS_{n,m,q,\beta}$)). Нехай n, m, q – додатні цілі числа, і нехай β є позитивним дійсним числом. Дано матрицю $\mathbf{A} \in \mathbb{Z}_q^{n \times m}$, вибрану рівномірно навмання, необхідно знайти ненульовий цілий вектор $\mathbf{z} \in \mathbb{Z}^m$ з евклідовою нормою $\|\mathbf{z}\| \le \beta$, такий що $\mathbf{Az} = \mathbf{0} \in \mathbb{Z}_q^n$.

П р о б л е м а 6. (Проблема пошуку – $NTRU_{R,q,\mathcal{D},\gamma}$). Нехай q – додатне ціле число, γ – додатне дійсне число, а R – кільце форми $R = \mathbb{Z}_q[x]/\Phi$ (де Φ – монічний поліном). Дано елемент $h \in R$, взятий з деякого розподілу $\mathcal{D}$, такий, що існує ненульовий $(f, g) \in R^2$, який задовольняє $h \cdot f = g \bmod q$ і має малі евклідові норми $\|f\|$, $\|g\| \le \sqrt{q}/\gamma$, необхідно знайти таку пару (f, g).

Наступні кілька проблем – це всі типи проблем з навчанням з помилками (LWE). Для вектору $\mathbf{s} \in \mathbb{Z}_q^n$ і розподілу помилок χ необхідно визначити розподіл $A_{\mathbf{s},\chi}$ навчання з помилками (LWE) над $\mathbb{Z}_q^n \times \mathbb{Z}_q$, вибравши $\mathbf{a} \in \mathbb{Z}_q^n$ рівномірно навмання, вибравши $e \leftarrow \chi$ над $\mathbb{Z}$ і вивівши пару $(\mathbf{a}, b)$ де $b = \langle \mathbf{s}, \mathbf{a} \rangle + e \bmod q$.

П р о б л е м а 7. (Проблема пошуку – $LWE_{n,m,q,\mathcal{B},\chi}$). Нехай $\mathbf{s} \in \mathbb{Z}_q^n$ вибрано з деякого розподілу $\mathcal{B}$. Дано m вибірок $(\mathbf{a}_1, b_1), \ldots, (\mathbf{a}_m, b_m) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, що взяті незалежно випадковим чином із розподілу $A_{\mathbf{s},\chi}$, необхідно знайти $\mathbf{s}$.

П р о б л е м а 8. (Проблема прийняття рішень – $LWE_{n,m,q,\mathcal{B},\chi}$). Нехай $\mathbf{s} \in \mathbb{Z}_q^n$ вибрано з деякого розподілу $\mathcal{B}$. Не знаючи $\mathbf{s}$, дано m вибірок $(\mathbf{a}_1, b_1), \ldots, (\mathbf{a}_m, b_m) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, розрізнити наступні два випадки: (i) вибірки беруться незалежно від розподілу $A_{\mathbf{s},\chi}$, або (ii) вибірки беруться незалежно від рівномірного розподілу над $\mathbb{Z}_q^n \times \mathbb{Z}_q$.

Далі визначаємо деякі алгебраїчно структуровані проблеми SIS/LWE. Як правило, у цих алгебраїчно структурованих варіантах кільце R вважається поліноміальним кільцем ступеня n у формі $R = R_q = \mathbb{Z}_q[X]/(f(X))$ для деякого натурального числа q. Варіанти $f(X)$, які розглядаються в третьому раунді, мають вигляд $f(X) = X^{2d} + 1$, як у KYBER, Saber, Dilithium і FALCON. Окремо $f(X) = X^n - 1$ і $f(X) = X^{n-1} + X^{n-2} + \ldots + X + 1$ використовуються NTRU, а $f(X) = X^p - X - 1$ для простого числа p вибирається NTRU LPrime і sNTRU Prime. У третьому раунді використання алгебраїчних-SIS/LWE здебільшого набуло формулювання на основі модулів, як показано нижче.

П р о б л е м а 9. (Задача Module – $SIS_{R,m,k,q,\beta}$). Дано $\mathbf{a}_1, \ldots, \mathbf{a}_m \in R_q^k$ векторів поліномів $\mathbf{a}_m \in R_q^k$, вибраних рівномірно навмання, розглянемо їх як рядки матриці $\mathbf{A} \in R_q^{m \times k}$. Потім необхідно знайти ненульовий поліноміальний вектор $\mathbf{z} \in R_q^k$ з нормою $\|\mathbf{z}\| \le \beta$ так, що $\mathbf{Az} = 0$.

П р о б л е м а 10. (Проблема прийняття рішення – $LWE_{R,m,q,\mathcal{B},\chi}$). Нехай $\mathbf{s} \in R_q^k$ вибрано з деякого розподілу $\mathcal{B}$. При невідомому $\mathbf{s}$ дано m вибірок $(\mathbf{a}_1, b_1), \ldots, (\mathbf{a}_m, b_m) \in R_q^n \times R_q$, ро-

зрізнити наступні два випадки: (i) кожна вибірка складається незалежно від розподілу $A_{R,\mathbf{s},\chi}$ (аналог розподілу LWE $A_{\mathbf{s},\chi}$, але над R_q), або (ii) кожна вибірка складається незалежно від рівномірного розподілу над $R_q^k \times R_q$.

Нарешті, представляємо сімейство проблем навчання з округленням (LWR). Різниця між LWE і LWR полягає в тому, що вибірки формуються як округлені внутрішні продукти, а не незалежно від вибірки розподілу помилок χ. Тобто, зразки LWR приймають форму $(\mathbf{a}_i, b_i) \in \mathbb{Z}_q^n \times \mathbb{Z}_q$, де $b_i = \lfloor \langle \mathbf{s}, \mathbf{a}_i \rangle \rceil_p$, а $\lfloor \cdot \rceil_p : \mathbb{Z}_q \to \mathbb{Z}_p$ (для $p < q$) є модульною функцією округлення, визначеною як $\lfloor x + q\mathbb{Z} \rceil_p := \lfloor x \cdot (p/q) \rceil + p\mathbb{Z}$.

Оцінка вартості вирішення цих критичних проблем безпеки на примірниках решітки реального світу є дуже нетривіальною, оскільки передбачає вибір найкращого типу атаки та оптимізацію параметрів атаки, щоб знайти найкраще можливе рішення із заданою кількістю обчислювальних ресурсів. Теоретичні межі та комп'ютерне моделювання використовуються для того, щоб оцінити вартість вирішення надзвичайно великих випадків цих проблем. Останніми роками це було предметом інтенсивних досліджень, які призвели до надійних оцінок конкретної безпеки криптосистем на основі решітки.

2.4. Моделі безпеки IND-CPA, IND-CCA2 та EUF-CMA

У оригінальному CFP [5] NIST дав визначення безпеки, які слід було сприймати як твердження того, що NIST вважав відповідною моделлю атаки. NIST планував стандартизувати KEM, які дозволили б «семантично безпечне» шифрування або інкапсуляцію ключів для загального використання – зокрема, схему, яка забезпечує нерозрізнення зашифрованих текстів під адаптованою атакою зашифрованого тексту. Грубо кажучи, схема є безпечною в цьому визначенні, якщо жоден зловмисник не може розрізнити «шифрування виклику» двох повідомлень на свій вибір, незважаючи на те, що він має доступ оракула до шифрування та розшифрування (останнє не можна використовувати під час виклику). Ця властивість позначається як IND-CCA2 безпека у науковій літературі [23]. Також для позначення цієї властивості використовуються терміни IND-CCA або CCA безпека.

Майже всі кандидати KEM, представлені NIST, досягли цієї функції, спочатку вказавши схему шифрування з відкритим ключем IND-CPA. Схема шифрування IND-CPA забезпечує нерозрізнення зашифрованих текстів під час обраної атаки відкритого тексту; це те саме визначення, що й вище, за винятком того, що криптоаналітик не має доступу оракула до розшифрування. Потім були створені повні KEM IND-CCA2 шляхом поєднання схем шифрування IND-CPA з певним типом перетворення Фудзісакі – Окамото (FO).

Для схем підписів відповідним визначенням безпеки була екзистенційна непідроблюваність під адаптованою атакою на повідомлення. Грубо кажучи, у цьому визначенні криптоаналітик отримує доступ оракула до функції підпису та повинен створити дійсний підпис для повідомлення, яке раніше не було підписано оракулом. У науковій літературі ця властивість позначається EUF-CMA безпека [23].

На додаток до цих визначень безпеки існують додаткові властивості безпеки. Хоча такі властивості не є обов'язковими для подання, вони можуть бути бажаними.

3. Стандартизація постквантової криптографії

Під час 3-го раунду було отримано деякі криптоаналітичні результати, які мали значний вплив на вибір NIST. Атака на GeMSS [24] різко знизила його безпеку та підірвала впевненість NIST у його стійкості. Цей результат призвів до виключення GeMSS з розгляду для стандартизації NIST.

Алгоритм Rainbow також зазнав значних атак під час 3-го раунду [25, 26]. Перша атака на початку 3-го раунду спричинила втрату наборів параметрів від 20 до 55 біт безпеки в моделі RAM, причому набори параметрів з вищим рівнем безпеки втрачали більше бітів

безпеки. За цим слідувала більш серйозна атака наприкінці 3-го раунду, що призвело до відновлення особистого ключа для параметрів категорії безпеки 1 трохи більше, ніж за два дні обчислень на одному ноутбуці. Через брак впевненості в безпеці NIST не вибрав Rainbow для стандартизації.

NIST також вирішив вилучити FrodoKEM, NTRU Prime та Picnic з розгляду для стандартизації. FrodoKEM – це кандидат, заснований на решітці, якого було обрано як альтернативний варіант під час 2-го туру. FrodoKEM в основному вирізняється тим, що він не покладається на структуровані решітки (на відміну від фіналістів Kyber, NTRU та Saber). У той час як NIST має намір вибрати принаймні один додатковий KEM, не заснований на структурованих решітках, для стандартизації після 4-го раунду, три інші альтернативи KEM (BIKE, HQC і SIKE) краще підходять для цієї ролі, ніж FrodoKEM.

FrodoKEM загалом має гіршу продуктивність, ніж ці три, тому не розглядатиметься надалі для стандартизації. NTRU Prime також було висунуто як альтернативний варіант, оскільки він вважався менш перспективним порівняно з фіналістами. Під час 3-го раунду не було результатів, які б суттєво змінили цю точку зору. Оскільки NIST буде стандартизувати один із (на основі структурованих решіток) фіналістів KEM, NTRU Prime не було обрано для продовження процесу. Схожа ситуація була і з підписами. Picnic не було обрано, оскільки NIST вирішив стандартизувати Sphincs+. Picnic та Sphincs+ мають подібні профілі ефективності (невеликі відкриті ключі та великі підписи) і підходять для тих же випадків використання. Sphincs+ і Picnic мають кілька версій, що робить пряме порівняння витрат та ефективності більш складним. Однак у кожного з них є набагато більша вартість та набагато гірша продуктивність порівняно з Dilithium та Falcon, що робить ці критерії менш важливими. Безпека Picnic не краща, ніж у Sphincs+, і NIST вважає, що, хоча Sphincs+ є зрілою конструкцією, Picnic та пов'язані з ним схеми продовжуватимуть отримувати користь від майбутніх досліджень та вдосконалень.

Вибираючи між подібними алгоритмами KEM, вартість та ефективність були значними критеріями відбору. NIST розглядав результати тестування продуктивності, надані спільнотою [6, 27] на декількох платформах при визначенні ефективності обчислень.

Одним із важких виборів, з якими стикнувся NIST, було прийняття рішення між Kyber, NTRU та Saber. Усі троє були обрані фіналістами і були дуже порівнянні один з одним. NIST впевнений у безпеці, яку забезпечує кожен. Більшість додатків зможуть використовувати будь-яку з них без суттєвих штрафів на продуктивність. Як зазначається, на завершення 2-го раунду NIST мав намір стандартизувати лише один із цих фіналістів, оскільки всі троє базувалися на структурованих решітках. Проблеми, пов'язані з патентами, були фактором рішення NIST протягом 3-го раунду, оскільки NIST дізнався про різні сторонні патенти. Однією з відмінностей між Kyber, Saber та NTRU є конкретне припущення щодо безпеки, що кожен покладається на безпеку. NIST вважає проблему MLWE, від якої залежить Kyber, трохи переконливішою, ніж інші припущення, такі як MLWR або проблема NTRU. NIST також високо оцінив специфікацію команди Kyber, яка включала ретельний і детальний аналіз безпеки. Що стосується продуктивності, то Kyber був майже найкращим (якщо не найкращим) у більшості тестів.

Решту обраних кандидатів KEM (BIKE, Classic McEliece, HQC, SIKE) продовжуватимуть оцінювати у 4-му раунді. І BIKE, і HQC засновані на структурованих кодах і будуть придатними як KEM загального призначення, що не ґрунтується на решітках. NIST може вибрати максимум одного з цих двох кандидатів для стандартизації по завершенню 4-го раунду. SIKE залишається привабливим кандидатом для стандартизації через його невеликі розміри ключа та шифротексту. NIST сподівається, що подальше вивчення SIKE триватиме протягом 4-го раунду. Classic McEliece був фіналістом, але наразі не стандартизується NIST. Хоча він вважається захищеним, NIST ще не передбачає, що він буде широко використовуватись через великий розмір відкритого ключа. Таким чином, ще немає терміновості для стандартизації Classic McEliece.

У [2] NIST вказав на намір вибрати щонайменше одного з Dilithium та Falcon, оскільки обидва базуються на структурованих решітках і можуть використовуватися в більшості додатків. Зрештою, NIST вирішив вибрати обидві схеми для стандартизації. Генерація ключа та підпису для Falcon, схоже, потребує більшої кількості ресурсів (гейтів та RAM), ніж Dilithium, що може зробити Falcon непридатним для впровадження на обмежених пристроях, особливо у випадках, коли вимагається захист від атак бічними каналами. Крім того, NIST визнає, що простіша конструкція ключа та генерації підписів Dilithium допоможе забезпечити безпечні реалізації. З цих причин NIST вибрав Dilithium як основний алгоритм підпису, який він рекомендує для загального використання, і надасть пріоритет його стандартизації.

NIST розуміє, що деякі додатки не працюватимуть так, як вони були розроблені, якщо підпис та дані, що підписуються, не будуть вписуватися в один Інтернет-пакет. Для цих додатків складність реалізації генерації підпису Falcon може не викликати занепокоєння, але труднощі з модифікацією додатків для роботи з більшим розміром підпису Dilithium можуть створити бар'єр для переходу до постквантових схем підпису. З цієї причини NIST вирішив також стандартизувати Falcon. Враховуючи загальну кращу продуктивність Falcon, коли генерацію підписів не потрібно виконувати на обмежених пристроях, багато додатків можуть вважати за краще використовувати Falcon, ніж Dilithium, навіть у випадках, коли розмір підпису Dilithium не буде перешкодою для реалізації.

Для того щоб не покладатися повністю на безпеку решіток, NIST також стандартизує Sphincs+. Безпека алгоритму підпису Sphincs+ добре зрозуміла, хоча він набагато більший та повільніший, ніж алгоритми підпису на решітках.

Підводячи підсумок, NIST обрав чотири алгоритми з 3-го раунду для стандартизації та чотири алгоритми для просування до 4-го раунду для подальшої оцінки та вивчення (див. табл. 3 та 4 для списку цих алгоритмів).

Таблиця 3

Алгоритми, які слід стандартизувати

Шифрування на відкритому ключі/KEM	Цифрові підписи
Crystals–Kyber	Crystals-Dilithium
	Falcon
	Sphincs+

Таблиця 4

Кандидати, що переходять до четвертого раунду

Шифрування на відкритому ключі/KEM	Цифрові підписи
BIKE	
Classic McEliece	
HQC	

Висновки

1. Основними алгоритмами, рекомендованими NIST для більшості випадків використання, є Crystals-Kyber (встановлення ключа) і Crystals-Dilithium (цифрові підписи). Крім того, схеми підпису Falcon і Sphincs+ також будуть стандартизовані. Кандидати BIKE, Classic McEliece, HQC, SIKE перейшли для подальшого вивчення до 4-го раунда оцінювання.

2. NIST створить нові проєкти стандартів для цих алгоритмів, координуючи команди подання, щоб гарантувати, що стандарти узгоджуються зі специфікаціями. У рамках процесу розробки NIST шукатиме інформацію про те, які конкретні набори параметрів слід включити, зокрема для будь-якої категорії безпеки 1. Після завершення стандарти будуть опубліковані для громадського обговорення. Після завершення періоду коментарів NIST перегляне проєкти стандартів, якщо це необхідно, на основі отриманих відгуків. Потім відбудеться остаточний розгляд, затвердження та процес оприлюднення. NIST сподівається опублікувати готовий стандарт до 2024 року.

3. NIST продовжує розглядати різноманітність обчислювальних припущень щодо складності як важливу довгострокову мету безпеки для своїх стандартів. NIST стандартизує практично ефективні схеми різних сімей криптосистем, щоб зменшити ризик того, що єдиний прорив у криптоаналізі залишить світ без життєздатного стандарту як для встановлення ключів, так і для цифрових підписів. Тим не менш, NIST не відчуває необхідності встановлювати ці стандарти відразу, а скоріше надасть пріоритет тим схемам, які здаються найближчими до того, щоб бути готовими до стандартизації та широкого прийняття. NIST вважає, що ця стратегія врівноважує прагнення до різноманітності з необхідністю ретельно перевіряти всі стандарти, перш ніж вони будуть видані.

4. Четвертий раунд оцінювання та аналізу відбуватиметься подібно до попередніх раундів. Як і раніше, чотирьом алгоритмам-кандидатам буде дозволено вносити відносно незначні модифікації у свої матеріали, які повинні бути подані до NIST і повинні відповідати тим же вимогам, що визначені в [5]. Подальші відомості та інструкції будуть надані на форумі PQC. Після завершення 4-го раунду NIST може вирішити вибрати деяких із кандидатів 4-го раунду для стандартизації.

5. Незважаючи на те, що 3-й раунд завершується і NIST почне розробляти перші стандарти PQC, зусилля зі стандартизації в цій галузі триватимуть ще деякий час. Це не слід інтерпретувати як те, що користувачі повинні чекати, щоб прийняти постквантові алгоритми. NIST сподівається на швидке впровадження цих перших стандартизованих алгоритмів і видасть майбутні вказівки щодо переходу. Перехід, безсумнівно, матиме багато складнощів, і виникнуть проблеми для деяких випадків використання, таких як пристрої IoT або прозорість сертифікатів.

Список літератури:

1. Alagic G., Alperin-Sheriff J., et al. (2019) Status report on the first round of the NIST post-quantum cryptography standardization process (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8240. – Режим доступу: https://doi.org/10.6028/NIST.IR.8240.
2. Alagic G., Alperin-Sheriff J., et al. (2020) Status report on the second round of the NIST post-quantum cryptography standardization process (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8309. Режим доступу: https://doi.org/10.6028/NIST.IR.8309.
3. NIST PQC. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization.
4. Schank J. (2021) Category 5 NTRU parameters. [Електронний ресурс]. Режим доступу: https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/t1JCgzSS-uk/m/VXXQaJgFCQAJ.
5. National Institute of Standards and Technology (2016) Submission requirements and evaluation criteria for the post-quantum cryptography standardization process. Режим доступу: https://csrc.nist.gov/CSRC/media/Projects/Post-Quantum-Cryptography/documents/call-for-proposals-final-dec-2016.pdf.
6. Bernstein D., Lange T. (eds.), eBACS: ECRYPT Benchmarking of Cryptographic Systems – SUPERCOP (2020). [Електронний ресурс]. Режим доступу: https://bench.cr.yp.to/supercop.html.
7. Shulman H., Goodman J., et al. (2021) PANEL: PQC considerations for DNSSEC, Third PQC Standardization Conference. [Електронний ресурс]. Режим доступу: https://www.nist.gov/video/third-pqc-standardization-conference-session-v-applications.
8. Bindel N. (2021) Suitability of 3rd round signature candidates for vehicle-to-vehicle communication // Workshop Record of the Third PQC Standardization Conference. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/Presentations/2021/suitability-of-3rd-round-signature-candidates-for.
9. Berlekamp E., McEliece R,, van Tilborg H. (1978) On the inherent intractability of certain coding problems (corresp.) // IEEE Transactions on Information Theory 24(3): 384-386. Режим доступу: https://doi.org/10.1109/TIT.1978.1055873.
10. Prange E. (1962) The use of information sets in decoding cyclic codes // IRE Transactions on Information Theory 8(5): 5-9. Режим доступу: https://doi.org/10.1109/TIT.1962.1057777.
11. Bernstein D. J. (2010) Grovervs. McEliece. Post-Quantum Cryptography, ed Kachigar G., Tillich J. P. (2017) Quantum information set decoding algorithms. Post-Quantum Cryptography, eds Lange T., Takagi T. (Springer International Publishing, Cham), pp. 69-89.
12. Kachigar G., Tillich J. P. (2017) Quantum information set decoding algorithms. Post-Quantum Cryptography, eds Lange T., Takagi T. (Springer International Publishing, Cham), pp. 69-89.

13. Kirshanova E. (2018) Improved quantum information set decoding. Post-Quantum Cryptography, eds Lange T., Steinwandt R. (Springer International Publishing, Cham), pp. 507-527.
14. Esser A., Ramos-Calderer S., et al. (2021) An optimized quantum implementation of ISD on scalable quantum resources, Cryptology ePrint Archive, Report 2021/1608. Режим доступу: https://ia.cr/2021/1608.
15. Patarin J., Goubin L. (1997) Trapdoor one-way permutations and multivariate polynominals // Proceedings of the First International Conference on Information and Communication Security ICICS'97 (Springer-Verlag, Berlin, Heidelberg), p. 356-368.
16. Buss J. F., Frandsen G. S., Shallit J. O. (1996) The computational complexity of some problems of linear algebra. BRICS Report Series 3(33). Режим доступу: https://doi.org/10.7146/brics.v3i33.20013.
17. Faugere J. C. (1999) A new efficient algorithm for computing Grobner bases (F_4). Journal of Pure and Applied Algebra 139(1): 61-88. Режим доступу: https://doi.org/https://doi.org/10.1016/S0022-4049(99)00005-5.
18. Faugere J. C. (2002) A new efficient algorithm for computing Grobner bases without reduction to zero (F_5) // Proceedings of the 2002 International Symposium on Symbolic and Algebraic Computation ISSAC'02 (Association for Computing Machinery, New York, NY, USA), p. 75-83. Режим доступу: https://doi.org/10.1145/780506.780516.
19. Courtois N., Klimov A., Patarin J,, Shamir A. (2000) Efficient algorithms for solving overdefined systems of multivariate polynomial equations. Advances in Cryptology – EUROCRYPT 2000, ed Preneel B. (Springer Berlin Heidelberg, Berlin, Heidel- berg), pp. 392-407.
20. Goubin L., Courtois N. T. (2000) Cryptanalysis of the TTM cryptosystem. Advances in Cryptology – ASIACRYPT 2000, ed Okamoto T. (Springer Berlin Heidelberg, Berlin, Heidelberg), pp. 44-57.
21. Bardet M., Bros M., et al. (2020) Improvements of algebraic attacks for solving the rank decoding and MinRank problems. Advances in Cryptology – ASIACRYPT 2020, eds Moriai S., Wang H. (Springer International Publishing, Cham), pp. 507-536.
22. Ajtai M. (1996) Generating hard instances of lattice problems (extended abstract). Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing STOC'96 (Association for Computing Machinery, New York, NY, USA), p. 99-108. Режим доступу: https://doi.org/10.1145/237814.237838.
23. Katz J., Lindell Y. (2020) Introduction to Modern Cryptography (Chapman & Hall/CRC), 3rd Ed.
24. Tao C., Petzoldt A., Ding J. (2020) Improved key recovery of the HFEv- signature scheme // Cryptology ePrint Archive, Report 2020/1424. Режим доступу: https://ia.cr/2020/1424.
25. Beullens W. (2021) Improved cryptanalysis of UOV and Rainbow. Advances in Cryptology – EUROCRYPT 2021, eds Canteaut A., Standaert F. X. (Springer International Publishing, Cham), pp. 348-373.
26. Beullens W. (2022) Breaking Rainbow takes a weekend on a laptop, Cryptology ePrint Archive, Report 2022/214. Режим доступу: https://ia.cr/2022/214.
27. pqm4: Post-quantum crypto library for the ARM Cortex-M4 (2020). [Електронний ресурс]. Режим доступу: https://github.com/mupq/pqm4.

Надійшла до редколегії 03.09.2022

Відомості про авторів:

Єсіна Марина Віталіївна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, доцент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; науковий співробітник-консультант АТ «Інститут Інформаційних технологій»; Україна; e-mail: m.v.yesina@karazin.ua; ORCID: https://orcid.org/0000-0002-1252-7606

Острянська Єлизавета Вадимівна – аналітик з систем захисту інформації, АТ «Інститут Інформаційних технологій», Україна; e-mail: antelizza@gmail.com

Горбенко Іван Дмитрович – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ "Інститут Інформаційних Технологій", головний конструктор, Україна; e-mail: gorbenkoi@iit.kharkov.ua; ORCID: https://orcid.org/0000-0003-4616-3449

УДК 004.056.5 **DOI:10.30837/rt.2022.3.210.06**

Є.В. ОСТРЯНСЬКА, М.В. ЄСІНА, канд. техн. наук, І.Д. ГОРБЕНКО, д-р техн. наук

АНАЛІЗ ПОГЛЯДІВ ЄВРОПЕЙСЬКОГО СОЮЗУ НА КВАНТОВО-ПОСТКВАНТОВІ ОБМЕЖЕННЯ

Вступ

Практично всім асиметричним криптографічним схемам, які зараз використовуються, загрожує потенційна розробка потужних квантових комп'ютерів. Постквантова криптографія є одним із способів боротьби з цією загрозою. Її безпека базується на складності математичних проблем, які наразі вважаються нерозв'язними ефективно – навіть за допомогою квантових комп'ютерів.

Однак із розробкою потужного квантового комп'ютера, на якому можна використовувати алгоритми Шора, безпека криптографії з відкритим ключем, яка використовується сьогодні, у майбутньому опиниться під серйозною загрозою. Це також впливає на типові схеми узгодження ключів, що є важливим елементом для захисту конфіденційності даних. Особливо це стосується даних, які повинні бути конфіденційними протягом тривалого періоду часу. Наприклад, якщо сьогодні зловмисник зафіксує обмін ключами, цілком можливо, що в майбутньому, коли стануть доступними криптографічно релевантні квантові комп'ютери, криптоаналітик зможе обчислити спільний ключ, розшифрувати та прочитати дані, зашифровані ним. Цей сценарій також відомий як «зберегти зараз, розшифрувати пізніше».

Щоб протистояти загрозі сучасної асиметричної криптографії з боку квантових комп'ютерів, виникла нова галузь криптографічних досліджень: постквантова криптографія.

Постквантова криптографія займається розробкою та дослідженням асиметричних криптосистем, які, згідно з сучасними знаннями, не можуть бути зламані навіть потужними квантовими комп'ютерами. Ці методи базуються на математичних задачах, для розв'язання яких на сьогодні невідомі ані ефективні класичні алгоритми, ані ефективні квантові алгоритми. У сучасних дослідженнях застосовуються різні підходи до реалізації постквантової криптографії, серед них: криптографія на основі кодів, криптографія на основі решітки, криптографія на основі гешування, криптографія на основі ізогенії та багатовимірна криптографія.

Метою статті є огляд обчислювальної моделі квантових комп'ютерів; квантових алгоритмів, які найбільше впливають на сучасну криптографію; ризику створення криптографічно-релевантних квантових комп'ютерів (CRQC); безпеки симетричної криптографії та криптографії з відкритим ключем за наявності CRQC; зусилля зі стандартизації NIST PQC; перехід до квантово-стійкої криптографії з відкритим ключем; актуальність та поточний стан розвитку квантово-стійкої криптографії у Європейському Союзі. Також висвітлюється хід найважливіших зусиль у цій галузі: стандартизації постквантової криптографії NIST.

1. Попередні визначення квантово-безпечної криптографії

Квантово-стійка криптографія – це криптографія, яка спрямована на надання криптографічних функцій і протоколів, які залишаються безпечними, навіть, якщо створено великомасштабні відмовостійкі квантові комп'ютери [1]. В останні роки спостерігається стійкий прогрес у створенні квантових комп'ютерів. У разі реалізації великомасштабних квантових комп'ютерів вони будуть загрожувати безпеці багатьох широко використовуваних криптосистем з відкритим ключем. Схеми встановлення ключів і цифрові підписи, засновані на факторизації, дискретних логарифмах і криптографії на еліптичних кривих, найбільш сильно постраждають. Симетричні криптографічні примітиви, такі як блокові шифри і геш-функції, будуть порушені незначно. Як результат, було проведено активізацію досліджень щодо пошуку криптосистем на відкритих ключах, які були б захищені від зловмисників як з квантовими, так і з класичними комп'ютерами. Цю область часто називають постквантовою криптографі-

єю (PQC), або іноді квантово-стійкою криптографією. Мета полягає в розробці схем, які можна розгорнути в існуючих комунікаційних мережах та протоколах без суттєвих змін.

2. Перспектива розвитку та загроза квантових комп'ютерів

Наразі немає однозначної відповіді та дуже незрозуміло, коли і навіть, якщо CRQC колинебудь буде побудовано. Розрив між сучасними квантовими комп'ютерами та передбачуваними CRQC величезний, і галузь стикається з деякими найближчими проблемами, такими як, наприклад, відсутність відомих програм для квантових комп'ютерів Noisy Intermediate-Scale (NISQ), які, як очікується, будуть створені найближчими роками.

Найкраща поточна оцінка, яка є на даний час, полягає в тому, що комітет експертів у 2019 році дійшов висновку, що поява CRQC протягом наступного десятиліття буде дуже несподіваною.

Колективні зусилля, спрямовані на створення квантових комп'ютерів, які можуть виконувати великі алгоритми, є широкими та складними. Є кілька публічно відомих залучень як в наукових колах, так і в промисловості [2, 3]. Існує також безліч потенційних реалізацій квантового комп'ютера (наприклад, з точки зору того, як реалізувати фізичні кубіти та квантові вентилі), які вивчаються та пропонуються, з надпровідними кубітами та кубітами на основі захоплених іонів, які є популярними кандидатами. Однак існує величезна різниця між сучасними гучними маленькими квантовими комп'ютерами та передбачуваними CRQC [4]. Так, на рис. 1 зображено передбачувану структуру майбутніх квантових комп'ютерів з виправленням помилок.

Рис. 1. Передбачувана структура майбутніх квантових комп'ютерів з виправленням помилок

Остання та найкраща оцінка, яка є на даний момент часу, полягає в тому, що комітет експертів з наукових кіл та промисловості у звіті 2019 року дійшов висновку, що поява CRQC протягом наступного десятиліття буде дуже несподіваною [4]. У тому ж звіті стверджується, що немає жодних відомих практичних застосувань для квантових комп'ютерів із шумовим проміжним масштабом (NISQ), які можна побачити в найближчі роки. У звіті також стверджується, що в іншому випадку галузь створення квантових комп'ютерів може стати суворо залежною від державного фінансування, якщо в найближчому майбутньому не вдасться знайти корисні програми.

Крім технічних аргументів і аргументів щодо зручності використання, оцінка прогресу квантових комп'ютерів ще більше ускладнюється іншими аспектами. По суті, те, що роблять учасники спільноти безпеки (тобто промисловість, наукові кола, уряди та організації зі стандартизації), це спостерігають один за одним, щоб оцінити, як інші учасники оцінюють ризик. Тоді ми можемо зауважити, що спільнота безпеки в цілому, здається, спокійно очікує результатів стандартизації NIST PQC, яка обговорюється в розд. 3 [5].

Щоб оцінити, коли буде необхідний перехід до квантово-безпечної криптографії, дуже показовим є наступне міркування фізика-теоретика M. Mosca з [6], що схематично зображено на рис. 2.

Нехай:

- x – кількість років, протягом яких дані, що підлягають захисту, повинні залишатися в безпеці,
- y – кількість років, необхідних для перетворення відповідної системи на стійку до квантового комп'ютера криптографію,
- z – кількість років, які знадобляться для існування квантових комп'ютерів, які загрожують криптографії, яка зараз використовується.

Тоді, якщо $x+y>z$, у вас проблема!

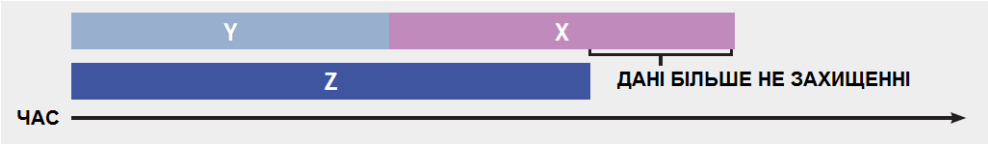


Рис. 2. Зображення «Теореми Mosca»

Це твердження стало відомим як «теорема Mosca», хоча це, звичайно, досить очевидне твердження.

Якщо перехід до квантово-безпечної криптографії розпочати сьогодні, він завершиться через y років. Наскільки великий y, залежить від різних факторів, таких як ступінь впливу на системи та доступність квантово-безпечних альтернатив. Отже, важливим першим кроком є оцінка та розробка плану міграції [7].

Таким чином, останні дані, які все ще були зашифровані за допомогою старих методів, будуть згенеровані через y років, а потім мають бути захищені ще на x років. У випадку спілкування в реальному часі цей період часу x може бути зникаюче малим. Натомість, наприклад, конфіденційна медична інформація має залишатися в безпеці протягом кількох десятиліть.

Припустимо, що виконується $x+y>z$. Потім можна перехопити останні дані, які ще не захищені квантово-безпечним способом, і розшифрувати їх протягом часу, протягом якого вони повинні бути захищені. Таким чином, перехід до квантово-безпечної криптографії має розпочатися досить рано, щоб $x+y<z$ все ще витримувався для всіх даних, які потрібно захистити. Але залишається під питанням наскільки великим має бути z.

Для національних систем безпеки BSI працює згідно з гіпотезою, що криптографічно відповідні квантові комп'ютери будуть доступні на початку 2030-х років [9, 10]. Слід підкреслити, що це твердження не слід розуміти як прогноз доступності квантових комп'ютерів, а скоріше представляє еталон для оцінки ризику. Тому BSI ініціював перехід до квантово-безпечної криптографії відповідно до програми федерального уряду «Квантові технології – від фундаментальних досліджень до ринку» [11].

3. Стандартизація постквантової криптографії

Постквантова криптографія займається розробкою та дослідженням асиметричних криптосистем, які, згідно з сучасними знаннями, не можуть бути зламані навіть потужними квантовими комп'ютерами. Ці методи базуються на математичних задачах, для розв'язання яких на сьогодні невідомі ані ефективні класичні алгоритми, ані ефективні квантові алгоритми. У сучасних дослідженнях застосовуються різні підходи до реалізації постквантової криптографії. До них належать, серед іншого:

- Криптографія на основі кодів: безпека схем на основі кодів ґрунтується на труднощах ефективного декодування загальних кодів з виправленням помилок.
- Криптографія на основі решітки: безпека схем на основі решітки базується на складності вирішення певних обчислювальних проблем на математичних решітках.

- Криптографія на основі гешування: безпека схем підпису на основі гешування базується на властивостях безпеки використаної геш-функції.
- Криптографія на основі ізогенії: схеми на основі ізогенії базують свою безпеку на тому факті, що важко знайти ізогенію між двома суперсингулярними еліптичними кривими, якщо така існує.
- Багатовимірна криптографія: безпека багатовимірної криптографії базується на припущенні, що багатовимірні поліноміальні системи рівнянь над скінченними полями важко вирішити.

Далі в цьому розділі будуть розглянуті лише перші три класи, оскільки постквантові схеми, рекомендовані BSI, належать лише до цих класів. Багатовимірні схеми мають довгу історію атак і виправлень. В даний час BSI не має наміру рекомендувати використання багатовимірних схем. Криптографія, заснована на ізогеніях (відображення між еліптичними кривими зі спеціальними властивостями), є цікавою темою дослідження, яку, на думку BSI, слід вивчити далі, перш ніж розглядати рекомендацію.

Як було вже наголошено, в останні роки постквантова криптографія набула значного значення: у серпні 2015 року Агентство національної безпеки США (NSA) попередило про вплив квантових комп'ютерів на безпеку криптографічних схем та ініціювало перехід на постквантові криптосистеми. Як виправдання NSA посилалося на досягнення у фізиці та технологіях, які могли б дозволити розробити криптографічно відповідний квантовий комп'ютер. NSA не назвало жодних конкретних постквантових алгоритмів, але послалося на майбутні стандарти Національного інституту стандартів і технологій (NIST).

Згідно з повідомленням NSA, NIST розпочав процес у листопаді 2016 року, наприкінці якого має бути доступний вибір постквантових схем. Цей процес проводиться в кілька раундів. До кінцевого терміну подання в листопаді 2017 року було подано 82 пропозиції, з яких 69 відповідали мінімальним критеріям і були прийняті NIST як кандидати в першому раунді. У січні 2019 року на основі публічних коментарів дослідницького співтовариства та внутрішнього аналізу NIST відібрав 26 із цих кандидатів для проходження до 2-го туру. Ці 26 кандидатів 2-го туру включають 17 схем асиметричного шифрування або узгодження ключів і 9 схем цифрових підписів. Потім, у липні 2020 року, NIST оголосив кандидатів, які пройдуть до 3-го туру. NIST розділив кандидатів 3-го туру на «фіналістів» і «альтернативних кандидатів». Причини, чому деякі схеми були названі альтернативними кандидатами, дуже різні.

Після трьох раундів оцінки та аналізу, NIST вибрав перші алгоритми, які він стандартизує в результаті процесу стандартизації PQC. Механізм інкапсуляції відкритого ключа (KEM), який буде стандартизований, – Crystals-Kyber. Цифрові підписи, які будуть стандартизовані, – Crystals-Dilithium, Falcon, та Sphincs+. Незважаючи на те, що вибираються кілька алгоритмів підпису, NIST рекомендує Crystals-Dilithium як основний алгоритм, який повинен бути реалізований. Крім того, чотири альтернативні алгоритми-кандидати KEM проходять до 4-го раунду оцінки: BIKE, Classic McEliece, HQC та SIKE. Ці кандидати будуть розглянуті для майбутньої стандартизації після завершення 4-го раунду.

Через терміновість переходу на процедури узгодження ключів, стійкі до квантового комп'ютера, BSI вперше рекомендував дві з цих схем у своїй технічній настанові TR-02102-1 [12] ще на початку 2020 року. У той же час німецька крипто індустрія має надію, що це має надати орієнтацію та дозволити їй розробляти продукти, готові до ринку, на ранній стадії, і це допоможе BSI зосередити дослідження безпеки на відповідних алгоритмах. Ці дві схеми – FrodoKEM [13] на основі решітки та Classic McEliece на основі коду [14], обидві з яких на той час перебували у 2-му раунді процесу NIST. У той час як Classic McEliece зараз є серед фіналістів 3-го раунду, FrodoKEM був включений до списку альтернативних кандидатів.

Паралельно з процесом NIST, а також у контексті постквантової криптографії, існують інші види стандартизації. Наприклад, Китайська асоціація криптологічних досліджень (CACR) проводила національний конкурс з 2018 по 2019 рік.

BSI вітає процес NIST як метод визначення стандартів у прозорому міжнародному процесі, який потім можна використовувати в усьому світі. Це особливо відноситься до окремого процесу стандартизації німецьких або європейських алгоритмів. «Поширення» міжнародних стандартів перешкоджатиме сумісності та зменшить ринкові можливості виробників криптовалюти. Крім того, розподіл персоналу та дослідницьких ресурсів призведе до нижчої якості оцінювання тих алгоритмів, які в кінцевому підсумку будуть обрані.

3.1. Транспортування ключів

Процес NIST спочатку шукав методи як для транспортування ключів, так і для шифрування. Однак стало очевидним, що подання, по суті, зосереджувалися на транспортуванні ключів і визначали асиметричне шифрування лише як попередній етап для такого механізму. В основному йдеться про методи FrodoKEM і Classic McEliece, рекомендовані BSI. FrodoKEM – це схема транспортування ключів на основі решітки, безпека якої базується на припущенні, що так звану проблему навчання з помилками (LWE) важко вирішити для класичних і квантових комп'ютерів. На відміну від багатьох інших схем на основі решіток у процесі NIST, базові решітки FrodoKEM не мають додаткової алгебраїчної структури. Хоча невідомо, чи така додаткова структура може бути використана криптоаналітиками, FrodoKEM таким чином усуває цей ризик. З іншого боку, FrodoKEM є дещо неефективним порівняно з деякими іншими схемами транспортування ключів на основі решітки. Додаткову інформацію про FrodoKEM також можна знайти в [15]. NIST виправдовує рішення включити FrodoKEM до списку альтернативних кандидатів тим, що хоча FrodoKEM має потенційні переваги в безпеці перед іншими схемами на основі решітки, він також пропонує нижчу продуктивність. Таким чином, стандартизація FrodoKEM, швидше за все, може почекати до закінчення 3-го раунду, і FrodoKEM також може служити «консервативною резервною копією», якщо криптоаналітичні досягнення будуть досягнуті щодо решіток з додатковою алгебраїчною структурою. Оскільки причина, по якій FrodoKEM не було включено до списку фіналістів 3-го раунду, не стосується безпеки схеми, BSI продовжує дотримуватися своєї рекомендації FrodoKEM.

Класичний McEliece – це схема транспортування ключів на основі коду, заснована на варіанті Niederreiter [16] схеми шифрування McEliece [17], створеному за допомогою двійкових кодів Гоппи. Оригінальна криптосистема McEliece була представлена ще в 1978 році, тому вона має довгу історію незламування порівняно з іншими постквантовими криптосистемами. Одним із недоліків схеми є те, що вона вимагає дуже великих відкритих ключів порівняно з іншими кандидатами, що може зробити її використання проблематичним у деяких сценаріях.

Таблиця 1

Середня продуктивність одноядерного процесора на Intel Xeon E-2124 3,3 ГГц для деяких кандидатів на алгоритм NIST PQC KEM (і деяких поточних альтернатив, не пов'язаних з PQC) на рівні безпеки 1 NIST PQC

Алгоритм KEM	Генерація ключа	Інкапсуляція	Декапсуляція	Розмір відкритого ключа	Розмір інкапсуляції
NTRU (ntruhps2048509)	0.048 ms	0.0073 ms	0.012 ms	699 B	699 B
Kyber (kyber512)	0.0070 ms	0.011 ms	0.0084 ms	800 B	768 B
SABER (lightsaber2)	0.012 ms	0.016 ms	0.016 ms	672 B	736 B
Classic McEliece (mceliece348864)	14 ms	0.011 ms	0.036 ms	261120 B	128 B
SIKE (SIKEp434_compressed)	3.0 ms	4.4 ms	3.3 ms	197 B	236 B
ECDH (X25519) (non-PQC)	0.038 ms	0.044 ms	0.044 ms	32 B	32 B
ECDH (P-256) (non-PQC)	0.074 ms	0.18 ms	0.18 ms	32-64 B	32-64 B
RSA-3072 (non-PQC)	400 ms	0.027 ms	2.6 ms	384 B	384 B

Іншими фіналістами в процесі NIST серед схем узгодження ключів є Crystals-Kyber на основі структурованої решітки, NTRU та SABER. Іншими альтернативними кандидатами є методи BIKE та HQC на основі кодів, NTRU Prime на основі структурованої решітки та

схема SIKE на основі ізогеній. Як видно з табл. 1, SIKE має відносно невеликі відкриті ключі та зашифровані тексти (приблизно 400 байт, і їх можна стиснути приблизно до 200 байт), але час роботи на порядки повільніший, ніж у багатьох інших кандидатів.

3.2. Схеми підписів

Кандидатами-фіналістами для схем підписів у 3-му раунді процесу NIST є схеми на базі решітки Crystals-Dilithium і Falcon і багатовимірна схема Rainbow.

Безпека Crystals-Dilithium базується на задачах решітки module-LWE та module-SIS, які є структурованими варіантами задач LWE та SIS (Short Integer Solution (коротке ціле рішення)) відповідно. Загалом Dilithium має хорошу продуктивність, помірні розміри ключа та підпису, і його легше реалізувати, ніж Falcon згідно з NIST [18].

Безпека Falcon базується на проблемі SIS, створеній за допомогою так званих решіток NTRU, які також мають додаткову структуру. Однією з цілей розробки Falcon є компактність, тобто мінімізація суми розмірів відкритого ключа та підпису. Підписання та перевірка за допомогою Falcon також ефективні, але генерація ключів повільніша порівняно з Dilithium. Як видно з табл. 2 Dilithium має найменший рекомендований набір параметрів у цій заявці – на рівні 2.

Через нові атаки на багатовимірні методи в 3-му раунді процесу NIST наразі видно, що Rainbow не буде стандартизовано [19]. Крім того, NIST планує прийняти нові пропозиції щодо схем підпису протягом 6 – 12 місяців після завершення 3-го раунду процесу стандартизації. У цьому випадку особливо будуть розглядатися ті схеми, які не базуються на структурованих решітках, див. також [20].

Таблиця 2

Середні значення продуктивності на одноядерному Intel Xeon E-2124 3,3 ГГц для деяких кандидатів на алгоритм підпису NIST PQC (і деяких поточних альтернатив, не пов'язаних з PQC) на рівні безпеки 1 NIST PQC

Алгоритм підпису	Генерація ключа	Підпис	Верифікація	Розмір відкритого ключа	Розмір підпису
Falcon (falcon512dyn)	5.9 ms	0.23 ms	0.029 ms	897 B	666 B
Dilithium (dilithium2aes)	0.015 ms	0.041 ms	0.019 ms	1312 B	2420 B
Rainbow (rainbow1aclassic363232)	2.7 ms	0.017 ms	0.0087 ms	161600 B	64 B
SPHINCS+ (SPHINCS+- SHA-256-128s-simple)	27 ms	210 ms	0.28 ms	32 B	7856 B
LMS (using SHA-256, limited to 220 messages)	-	-	-	56 B	2828 B
Ed25519 (non-PQC)	0.014 ms	0.015 ms	0.050 ms	32 B	64 B
ECDSA (P-256) (non-PQC)	0.029 ms	0.041 ms	0.086 ms	64 B	64 B
RSA-3072 (non-PQC)	400 ms	2.6 ms	0.027 ms	384 B	384 B

Серед кандидатів у процесі NIST Sphincs+ [21] є консервативним вибором як метод на основі гешу без стану. NIST розглядав Sphincs+ на 3-му раунді конкурсу як безпосередньо доступну альтернативу, якщо криптоаналітичні досягнення обмежать впевненість у безпеці фіналістів. Інші альтернативні кандидати включають Picnic і GeMSS, де Picnic базується на симетричних примітивах і техніках з нульовим знанням, а GeMSS є багатовимірною схемою підпису.

І як було сказано вище, після трьох раундів оцінки та аналізу, NIST вибрав перші алгоритми, які він стандартизує в результаті процесу стандартизації PQC і серед цифрових підписів це будуть – Crystals-Dilithium, Falcon, та Sphincs+.

4. Подальший розвиток криптографічних протоколів, оцінки та рекомендації

Французький ANSSI вже прокоментував використання QKD у документі з позицією [22]. У ньому згадуються обмеження, які вже обговорювалися тут. Серед іншого, проблематичними вважаються складне та дорогоцінне придбання, велика кількість продемонстрованих атак бічними каналами на пристрої QKD, обмежений радіус дії та відсутність наскрізної безпеки на великих відстанях. ANSSI робить висновок, що постквантова криптографія надає

альтернативу, яка є простішою та дешевшою у реалізації та не підлягає багатьом обмеженням QKD. Тому слід зосередитися на просуванні постквантової криптографії як квантово-безпечної криптографії.

NSA також вказує на технічні обмеження QKD. До них належать необхідність розповсюдження ключів для автентифікації, дороге придбання спеціалізованого обладнання та висока вразливість до атак на фізичну реалізацію та атак типу «відмова в обслуговуванні». З цих причин NSA виступає проти використання QKD у системах національної безпеки, доки не будуть усунені вищезазначені обмеження.

NCSC Великобританії також виступає проти використання QKD в урядових і військових програмах [23].

Як обговорювалося раніше, QKD має багато практичних обмежень. Деякі з них можуть бути подолані в майбутньому. Особливо бажаною була б розробка квантових повторювачів для підтримки наскрізної безпеки. Однак цього не варто очікувати найближчими роками. Крім того, європейські продукти QKD наразі недоступні на ринку. Навіть, якщо європейські продукти розробляються, їх спочатку потрібно оцінити за критеріями, які ще належить розробити. Це правда, що BSI робить перші кроки в цьому напрямку з розробкою профілю захисту. Однак поки що цей профіль захисту обмежений протоколами підготовки та вимірювання та з'єднаннями «точка-точка» та вимагає подальшого створення обширної супровідної документації.

Беручи до уваги робочу гіпотезу про те, що криптографічно відповідний квантовий комп'ютер буде доступний на початку 2030-х років, BSI вважає, що вже зараз необхідно терміново вжити відповідних заходів для переходу на квантово-безпечні схеми. Сама по собі ця нагальність робить перехід до постквантової криптографії, стандартизація якої вже добре просунута в процесі NIST, явним пріоритетом з точки зору BSI. Крім того, постквантові алгоритми набагато гнучкіші, оскільки їх можна реалізувати в існуючій інфраструктурі, вони економічніші, не вимагають секретних попередньо розподілених ключів і пропонують наскрізну безпеку.

На відміну від класичних і постквантових схем, QKD обіцяє інформаційно-теоретичну безпеку. Однак для цього потрібні відповідні докази безпеки для практично використовуваних протоколів і найзагальнішої моделі атаки. З точки зору BSI, теоретичні основи QKD ще не були задовільно розроблені в цьому відношенні. З огляду на це та сприйнятливість реалізацій до атак із стороннього каналу, «надзахищені» оцінки QKD, які іноді робляться, здаються недоречними.

Отже, з точки зору BSI, до того як QKD можна буде рекомендувати як критично важливу для безпеки технологію для практичних застосувань, ще потрібно роз'яснити багато питань і вирішити обмеження. Однак QKD і постквантова криптографія мають потенціал доповнювати одна одну, особливо тому, що вони базуються на різних принципах. Використання QKD в даний час можливо в основному в контексті експериментів для випадків обмеженого використання, де практичні обмеження менш значні, в гібридному режимі як доповнення в поєднанні з класичними і постквантовими методами узгодження ключів. Крім того, це також може забезпечити наскрізну безпеку на великих відстанях. Подальші дослідження квантової комунікації вітаються також тому, що можуть бути багатообіцяючі програми поза криптографією.

Співтовариство з безпеки та відповідні учасники чекають на завершення стандартизації NIST PQC. Наприклад, NSA все ще рекомендує набір неквантово-стійких алгоритмів відкритого ключа для захисту «цілком-секретних» [24]. Як правило, такий матеріал потребує захисту протягом десятиліть. Те, що галузь може зробити сьогодні, – це забезпечити, щоб продукти були достатньо підготовлені для переходу на алгоритми з відкритим ключем із властивостями (розміри ключа та підпису/зашифрованого тексту), подібними до пропозицій 3-го раунду стандартизації NIST PQC, коли прийде час. Схеми в стандартизації NIST PQC призначені для заміни функціональних інтерфейсів для поточного встановлення відкритого ключа та

алгоритмів підпису. Це означає, що такі протоколи, як IKEv2 і TLS, можуть продовжувати працювати, як і сьогодні, з дещо іншими характеристиками продуктивності в асиметричній частині протоколів (з відкритим ключем), коли протоколи та сертифікати оновлено для підтримки нових алгоритмів. Комунікаційні накладні витрати PQC також можуть сприяти заповненню вікон перевантаження в таких протоколах, як TCP [25].

Важко сказати, що саме станеться, коли стандартизація NIST PQC закінчиться в найближчі кілька років. Ймовірним сценарієм є те, що інші організації, які розробляють стандарти, підуть за цим і, що важливі протоколи, такі як TLS і IKEv2, будуть оновлені для підтримки нових стандартизованих алгоритмів PQC. Цілком імовірно також, що нові алгоритми мають хорошу підтримку бібліотек у важливих бібліотеках програмного забезпечення, таких як OpenSSL, на той час, оскільки робота над реалізаціями продуктивного рівня (наприклад, ефективного та безпечного від атак із синхронізованим боковим каналом) уже триває. На відміну від цього, дещо більш невизначеним є те, наскільки швидко буде розгортання в додатках, підтримка апаратного забезпечення та підтримка PKI. Ці речі мають певну ціну (наприклад, накладні витрати, проблеми із застарілою взаємодією та витрати на розгортання/розробку) для відповідних учасників, і одним із важливих факторів тут може бути реальний прогрес, досягнутий у напрямку створення великомасштабних квантових комп'ютерів у наступні декілька років. Як грубу аналогію можна розглянути, наскільки повільно додатки та суб'єкти історично поступово відмовлялися від алгоритмів (наприклад, MD5, RC4, SHA-1), безпека яких поступово, але публічно погіршувалася через криптоаналіз [26]. Однак слід враховувати не лише вартість оновлення криптографічних алгоритмів і ризики безпеки від використання погіршених алгоритмів, а й чисту вартість репутації використання погіршених алгоритмів. Подія, яка може сприяти прийняттю PQC після завершення стандартизації PQC NIST, полягає в тому, що SDO та інші важливі суб'єкти, такі як NSA, не лише оновлять стандарти та вказівки для підтримки нових алгоритмів PQC, але й припинять використання алгоритмів відкритих ключів, які зараз використовуються.

Для майбутнього використання постквантової криптографії недостатньо стандартизувати криптографічні алгоритми. Швидше, необхідно також адаптувати криптографічні протоколи до нових алгоритмів. Це пов'язано, наприклад, з тим, що в багатьох протоколах дозволена довжина відкритих ключів обмежена і більше не достатня для нових алгоритмів. Однак суттевим моментом є те, що постквантові алгоритми, як правило, не слід використовувати окремо, а лише в гібридному режимі, тобто в поєднанні з класичною процедурою. Зміни в протоколах і стандартах повинні бути ініційовані та спільно розроблені галуззю. Ця робота вже триває для багатьох протоколів.

В останні роки були запущені великі міжнародні програми з просування квантових технологій. Наприклад, Федеральне міністерство освіти та досліджень (BMBF) оголосило про свій намір сприяти розробці довгострокової безпечної криптографії та її ефективному застосуванню в додатках у рамках науково-дослідницької програми федерального уряду з IT-безпеки «Самовизначення і безпека у цифровому світі 2015 – 2020» [27]. З цією метою в серпні 2018 року було опубліковано рекомендації щодо фінансування дослідницьких проєктів на тему «постквантової криптографії». У рамках цього плану протягом 2019 – 2022 років фінансується сім проєктів для інтеграції постквантової криптографії в програми (Aquorypt), інфраструктури відкритих ключів (FLOQI), криптобібліотека Botan (KBLS), обробка медичних даних (PQC4MED), вбудовані системи (QuantumRISC), мережі (QuaSiModO) та критичні інфраструктури (SIKRIN-KRYPTOV). Загальний обсяг усіх цих проектів становить 24,2 млн. євро, при цьому частка фінансування BMBF становить приблизно 16,1 млн. євро.

Під керівництвом BMBF дослідницька програма федерального уряду «Квантові технології – від фундаментальних досліджень до ринку» [11] забезпечить федеральне фінансування у розмірі 650 млн. євро на розвиток квантових технологій у Німеччині в період з 2018 по 2022 рік.

Стимулюючий і майбутній пакет федерального уряду передбачає загалом 2 млрд. євро на розвиток квантових технологій і, зокрема, квантових обчислень, з яких приблизно 1,1 млрд. євро виділено BMBF і приблизно 900 млн. євро Федеральному міністерству економіки та енергетики (BMWi).

Зокрема, у зв'язку з просуванням квантових обчислень консультативний комітет, призначений у жовтні 2020 року, склав «Roadmap Quantencomputing» від імені Федерального уряду [28]. Мотивуючись цим, BMBF ініціював конкретні заходи фінансування для «Демонстраційних збірок квантових комп'ютерів» та «Мережі прикладних програм для квантових обчислень» у рамках фінансування поточної програми «Квантові технології – від фундаментальних досліджень до ринку». Більшість фінансування, яким керує BMWi, зосереджено на Німецькому аерокосмічному центрі (DLR) з метою розробки німецького квантового комп'ютера та відповідного програмного забезпечення та додатків.

Ще однією важливою програмою є флагманська програма ЄС з квантових технологій, що розпочалася 1 жовтня 2018 року із загалом 24 дослідницькими проєктами. Програма розрахована на 10 років і має загальний обсяг 1 млрд євро. На першому етапі з жовтня 2018 року по вересень 2021 року вона мала забезпечити загальну суму 152 млн. євро для 24 проєктів.

Проєкти охоплюють аспекти «Фундаментальна наука», «Квантові симуляції», «Квантові датчики та метрологія», «Квантові комунікації» та «Квантові обчислення». Ці та їхні плани описані в «Програмі стратегічних досліджень» [29, 30]. Зокрема, програма з квантових обчислень містить два проєкти побудови європейського квантового комп'ютера. Цей проєкт OpenSuperQ зосереджений на надпровідних кубітах, подібних до IBM, Google і Rigetti Computing та проєкту AQTION, що використовує захоплені іони і метою якого є реалізація портативного та в принципі комерційного обладнання для квантових комп'ютерів на рівні понад 50 кубітів.

У той же час Європейське спільне підприємство високопродуктивних обчислень (EuroHPC JU) переслідує цілі створення європейської інфраструктури квантового комп'ютера та сприяння дослідженням та інноваціям у цій галузі. Після переорієнтації програми у вересні 2020 року бюджет на період 2021 – 2033 років тепер становить 8 млрд. євро і включає розробку інфраструктури квантових обчислень та квантового моделювання для інтеграції в інфраструктуру високопродуктивних обчислень (HPC). Є намір побудувати такий сучасний проєкт до 2023 року.

QuNET (див. також [31]) – це національний дослідницький проєкт квантового розподілу ключів з використанням різних технологій з обсягом проекту 165 млн. євро до 2026 року, з яких BMBF виділяє 125 млн. євро для фінансування. Основними інститутами, які беруть участь у QuNET, є Інститут прикладної оптики та точного машинобудування Фраунгофера (IOF), Інститут імені Фраунгофера Генріха Герца (HHI), Інститут зв'язку та навігації Німецького аерокосмічного центру (DLR-IKN) та Інститут Макса Планка для науки про світло (MPL). У рамках проєкту будуть розроблені концепції загальної мережі та необхідної архітектури системи, а також нові ключові технології для квантового зв'язку. Також будуть враховані вимоги до стандартизації та сертифікації загальних систем QKD. Частиною підпроєкту QuNET-alpha було встановлення зашифрованого з'єднання між BMBF і BSI у Бонні в серпні 2021 року. Його було розроблено як гібридну схему шляхом поєднання постквантової схеми та QKD для узгодження ключів.

З огляду на кількість відкритих проєктів та явну зацікавленість ЄС у квантових технологіях, можна зробити висновки, що квантові технології все ще знаходяться в зародковому стані, але вже беззаперечно, що вони мають величезний економічний потенціал і значною мірою впливатимуть на інформаційну безпеку. Технологія квантових сенсорів, квантовий зв'язок і квантові комп'ютери все більше стають центром успішного довгострокового економічного розвитку Німеччини та Європи.

Висновки

1. Криптографія з відкритим ключем, що використовується в даний час, така як RSA, Діффі – Геллман, Ель Гамаль або ECC, знаходиться під загрозою квантових обчислень. Сучасні криптографічно-релевантні квантові алгоритми по суті вимагають успішної квантової корекції помилок (QEC). Криптоаналітичні досягнення на основі вже наявних пристроїв Noisy Intermediate Scale Quantum (NISQ) не можуть бути виключені. Комерціалізація квантових обчислень уже почалася, наприклад, завдяки широкому поширенню Quantum as a Service (QaaS).

2. На даний момент постквантовим криптографічним схемам, як правило, ще не довіряють такою ж мірою, як усталеним криптосистемам, оскільки вони не були добре вивчені, наприклад, з точки зору стійкості до атак бічними каналами і безпеки реалізації. Водночас, однак, необхідно своєчасно переходити на квантово-безпечні схеми. З цієї причини ідея не використовувати постквантову криптографію окремо, а лише в поєднанні з усталеними алгоритмами, загалом отримала визнання.

3. Наразі немає однозначної відповіді та дуже незрозуміло, коли і навіть, якщо CRQC коли-небудь буде побудовано. Розрив між сучасними квантовими комп'ютерами та передбачуваними CRQC величезний, і галузь стикається з деякими найближчими проблемами, такими як, наприклад, відсутність відомих програм для квантових комп'ютерів Noisy Intermediate-Scale (NISQ), які, як очікується, будуть створені найближчими роками. Найкраща поточна оцінка, яка є на даний час, полягає в тому, що комітет експертів у 2019 році дійшов висновку, що поява CRQC протягом наступного десятиліття буде дуже несподіваною.

4. Однак ризик створення CRQC означає, що наразі розгорнуту криптографію з відкритим ключем необхідно замінити квантово-стійкими альтернативами. Наприклад, інформація, зашифрована за допомогою сучасної криптографії з відкритим ключем, може бути записана криптоаналітиками, а потім піддана атаці, якщо можна створити QRQC. Потенційна шкода, яку може завдати CRQC, є основою мотивації шукати контрзаходи, навіть, якщо у нас є невизначеність щодо того, коли та чи можна створити ці комп'ютери. Оновлення розгорнутих систем, які використовують криптографію з відкритим ключем, також може тривати багато років.

5. На безпеку симетричної криптографії (включаючи криптографічні геш-функції) CRQC (включаючи розміри ключів) практично не впливають. У той час як очікується, що алгоритм Шора зможе зламати сучасну криптографію з відкритим ключем за лічені години на одному CRQC, очікується, що алгоритм, який застосовується до симетричної криптографії, алгоритм Гровера, матиме гіпотетичний час роботи багато мільярдів років на аналогічному розмірі CRQC.

6. Стандарти, наприклад, щодо протоколів, і сертифіковані продукти все ще відсутні. QKD слід використовувати лише в гібридному режимі з класичними та постквантовими схемами узгодження ключів.

7. Наскільки добре можна атакувати криптографічні алгоритми за допомогою квантових комп'ютерів, залежить не лише від прогресу, досягнутого в створенні квантових комп'ютерів, але й значною мірою від алгоритмічних інновацій. Наприклад, чи існують криптографічно відповідні квантові алгоритми, які вимагають менше кубітів? Або що обійдеться з меншим чи без квантового виправлення помилок? Або мають меншу глибину контуру? Чи можна прискорити криптографічні атаки за допомогою квантових комп'ютерів спеціального призначення? Ці запитання показують, що важливо поєднувати дослідження квантових комп'ютерів і квантових алгоритмів.

8. Також досі залишається багато відкритих питань щодо постквантової криптографії. З одного боку, стійкість до атак бічними каналами і безпека впровадження цих криптосистем ще недостатньо досліджені. З іншого боку, необхідні подальші дослідження можливих криптоаналітичних досягнень, як з класичними, так і з квантовими комп'ютерами. Зокрема, питання про те, чи структуровані та неструктуровані решітки забезпечують однакову безпе-

ку, є важливим дослідницьким питанням, яке слід шукати. З огляду на ці питання у Європейському Союзі було створено багато проєктів та ініціатив щодо дослідження багатьох питань, що стосуються постквантової криптографії та зокрема побудови великомасштабного квантового комп'ютера.

Список літератури:

1. John Preuß Mattsson, Ben Smeets and Erik Thormarker Quantum-Resistant Cryptography. Ericsson Security Research. Режим доступу: https://arxiv.org/ftp/arxiv/papers/2112/2112.00399.pdf.

2. Diane Peters. The quest to build a reliable quantum computer, 14 October 2020. [Електронний ресурс]. Режим доступу: https://www.universityaffairs.ca/features/feature-article/the-quest-to-build-a-reliable-quantum-computer/.

3. The GSMA Internet Group Quantum Computing, Networking and Security, Version 1.0, March 2021. Режим доступу: https://www.gsma.com/newsroom/wp-content/uploads/IG-11-Quantum-Computing-Networking-and-Security.pdf.

4. National Academies of Sciences, Engineering, and Medicine; Division on Engineering and Physical Sciences; Computer Science and Telecommunications Board; Intelligence Community Studies Board; Committee on Technical Assessment of the Feasibility and Implications of Quantum Computing; Emily Grumbling and Mark Horowitz: "Quantum Computing Progress and Prospects", 2019. [Електронний ресурс]. Режим доступу: https://www.nap.edu/catalog/25196/quantum-computing-progress-and-prospects#toc.

5. Post-Quantum Cryptography PQC. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/post-quantum-cryptography-standardization/call-for-proposals.

6. M. Mosca Cybersecurity in an era with quantum computers: will we be ready? (2015). Режим доступу: https://eprint.iacr.org/2015/1075.pdf.

7. Federal Office for Information Security Quantum-safe cryptography – fundamentals, current developments and recommendations, 2022.05.18. [Електронний ресурс]. Режим доступу: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Brochure/quantum-safe-cryptography.pdf?__blob=publicationFile&v=4.

8. Internationale Fernmeldeunion ITU-T Recommendation X.509 10/2019, October 2019. [Електронний ресурс]. Режим доступу: https://www.itu.int/ITU-T/recommendations/rec.aspx?rec=X.509.

9. Deutscher Bundestag Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Dr. Anna Christmann, Kai Gehring, Margit Stumpp, weiterer Abgeordneter und der Fraktion BÜNDNIS 90 / DIE GRÜNEN – Drucksache 19/24762. Режим доступу: https://dserver.bundestag.de/btd/19/252/1925208.pdf.

10. Deutscher Bundestag Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Dr. Konstantin von Notz, Tabea Rößner, Dr. Irene Mihalic, weiterer Abgeordneter und der Fraktion BÜNDNIS 90 / DIE GRÜNEN – Drucksache 19/25549. Режим доступу: https://dserver.bundestag.de/btd/19/263/1926340.pdf.

11. Federal Ministry of Education and Research Quantum technologies – from basic research to market, A Federal Government Framework Programme, September 2018. Режим доступу: https://www.quantentechnologien.de/fileadmin/public/Redaktion/Dokumente/PDF/Publikationen/Federal-Government-Framework-Programme-Quantum-technologies-2018-bf-C1.pdf.

12. Federal Office for Information Security BSI TR-02102-1: Cryptographic Mechanisms: Recommendations and Key Lengths. [Електронний ресурс]. – Режим доступу: https://www.bsi.bund.de/EN/Service-Navi/Publications/TechnicalGuidelines/tr02102/BSITR02102.html.

13. M. Naehrig, E. Alkim, et al. FrodoKEM, National Institute of Standards and Technology, 2020. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

14. M. R. Albrecht, D. J. Bernstein, et al. Classic McEliece, National Institute of Standards and Technology, 2020. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

15. H. Hagemeier: Frodo is the "New Hope", BSI-Magazine 2020/01, S. 12-14. [Електронний ресурс]. Режим доступу: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/Magazin/BSI-Magazin_2020-01.pdf?__blob=publicationFile&v=1.

16. H. Niederreiter Knapsack-type cryptosystems and algebraic coding theory // Problems of Control and Information Theory, 15(2), S. 159-166, 1986.

17. R. J. McEliece A public-key cryptosystem based on algebraic coding theory // Technical report, NASA, 1978. Режим доступу: https://tmo.jpl.nasa.gov/progress_report2/42-44/44N.PDF.

18. D. Moody, G. Alagic, et al. Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process, 2020, NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD. Режим доступу: https://doi.org/10.6028/NIST.IR.8309.

19. W. Buellens Breaking Rainbow Takes a Weekend on a Laptop, February 2022. Режим доступу: https://eprint.iacr.org/2022/214.

20. National Institute of Standards and Technology NIST Status Update on the 3rd Round, July 2020. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/Presentations/2021/status-update-on-the-3rd-round.

21. A. Hülsing, D. J. Bernstein, et al. SPHINCS+, National Institute of Standards and Technology, 2020. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.

22. Agence nationale de la sécurité des systèmes d'information (ANSSI) Should Quantum Key Distribution be Used for Secure Communications?, Technical Position Paper, May 2020. Режим доступу: https://www.ssi.gouv.fr/uploads/2020/05/anssi-technical_position_papers-qkd.pdf.

23. National Cyber Security Center Quantum security technologies, Whitepaper, 24. March 2020. Режим доступу: https://www.ncsc.gov.uk/pdfs/whitepaper/quantum-security-technologies.pdf.

24. NSA/CSS Commercial National Security Algorithm Suite. [Електронний ресурс]. – Режим доступу: https://apps.nsa.gov/iaarchive/programs/iad-initiatives/cnsa-suite.cfm.

25. Bas Westerbaan Sizing Up Post-Quantum Signatures for the Web, 31 October, 2021. [Електронний ресурс]. Режим доступу: https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/anE3sBUWZS0.

26. Daniel J. Bernstein Boring crypto, University of Illinois at Chicago & Technische Universiteit Eindhoven. Режим доступу: http://cr.yp.to/talks/2015.10.05/slides-djb-20151005-a4.pdf.

27. Federal Ministry of Education and Research Self-determined and secure in the digital world 2015-2020, The German Government's research framework programme on IT security, March 2015. [Електронний ресурс]. Режим доступу: https://www.forschung-it-sicherheit-kommunikationssysteme.de/service/publikationen/self-determined-and-secure-in-the-digital-world-2015-2020.

28. VDI Technologiezentrum GmbH Roadmap Quantencomputing, October 2020. [Електронний ресурс]. Режим доступу: https://www.quantentechnologien.de/fileadmin/public/Redaktion/Dokumente/PDF/Publikationen/Roadmap-Quantencomputing-bf-C1.pdf.

29. EU Quantum Technologies Flagship Strategic Research Agenda, March 2020. [Електронний ресурс]. Режим доступу: https://qt.eu/about-quantum-flagship/resources/.

30. European Commission New Strategic Research Agenda on Quantum Technologies, February 2020. [Електронний ресурс]. Режим доступу: https://digital-strategy.ec.europa.eu/en/news/new-strategic-research-agenda-quantum-technologies.

31. Deutscher Bundestag: Antwort der Bundesregierung auf die Kleine Anfrage der Abgeordneten Manuel Höferlin, Frank Sitta, Grigorios Aggelidis, weiterer Abgeordneter und der Fraktion der FDP – Drucksache 19/17500 – Hochsicheres Quantennetzwerk QuNET. Режим доступу: https://dserver.bundestag.de/btd/19/183/1918355.pdf.

Надійшла до редколегії 05.09.2022

Відомості про авторів:

Острянська Єлизавета Вадимівна – аналітик з систем захисту інформації, АТ «Інститут Інформаційних технологій», Україна; e-mail: antelizza@gmail.com

Єсіна Марина Віталіївна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, доцент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; науковий співробітник-консультант АТ «Інститут Інформаційних технологій»; Україна; e-mail: m.v.yesina@karazin.ua; ORCID: https://orcid.org/0000-0002-1252-7606

Горбенко Іван Дмитрович – д-р техн. наук, професор, Харківський національний університет імені В. Н. Каразіна, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, АТ "Інститут Інформаційних Технологій", головний конструктор, Україна; e-mail: gorbenkoi@iit.kharkov.ua; ORCID: https://orcid.org/0000-0003-4616-3449

UDC 004.725.4+517.9 **DOI:10.30837/rt.2022.3.210.07**

Y. KOTUKH, V. LUBCHAK, O. STRAKH

NEW CONTINUOUS-DISCRETE MODEL FOR WIRELESS SENSOR NETWORKS SECURITY

1. Introduction

A wireless sensor network (WSN) is a group of «smart» sensors with a wireless infrastructure designed to monitor the environment. This technology is the basic concept of the Internet of Things (IoT). The WSN can transmit confidential information while working in an insecure environment. Therefore, appropriate safety measures must be considered in the network design. However, computational node constraints, limited storage space, an unstable power supply, and unreliable communication channels, and unattended operations are significant barriers to the application of cybersecurity techniques in these networks.

There are mathematical models for studying the prevalence of malicious software, which can be global (by the topology of communication between WSN nodes, but not by their characteristics) [1 – 8], or individual (by individual features of nodes, but not by the global nature of their interaction) [9 – 15] models. In addition, existing models can be classified by types of interaction (continuous [3 – 5] and discrete [1, 2, 6 – 8, 9 – 15], deterministic [1, 3–5, 7, 10–14] and stochastic [2, 6, 8, 9, 15], etc.) and the use of mathematical apparatus (system of partial differential equations) [3, 4, 8], systems of ordinary differential equations [1, 5, 7], cellular automata [9, 10, 12], Markov chains [2, 6, 11], agent modelling [13 – 15], etc.). All existing models have certain specifics and possibilities for their application to build a strategy to protect WSN from malware. But they also have certain drawbacks. Given the peculiarities of obtaining data on the state of a group of nodes WSN, this process cannot be considered in a purely continuous or purely discrete mode. These two factors must be combined.

This article considers a new continuous-discrete model of malware propagation through wireless sensor network nodes, which is based on a system of so-called dynamic equations with impulsive effect on time scales.

2. Our approach

Consider some wireless sensor networks. Its continuous operation can be observed only at certain time intervals; at other intervals, the possibilities of observation are limited to individual point transmissions of relevant information. Therefore, to build a model, it is necessary to use mathematical objects at continuously discrete intervals. One of the theories that allow this is the theory of dynamic equations on time scales [16]. The key concepts of this theory that we need in the future are the time scale ($\mathbb{T}$) – an arbitrary closed non-empty subset of the real numbers, the forward jump operator ($\sigma(t) \coloneqq \inf\{\forall s \in \mathbb{T} : s > t\}$), delta derivative ($x^{\Delta}$), which is a generalization of the concepts of ordinary derivative and difference operator, as well as a matrix exponential function $e_A(t,s)$ [16].

Let the studied WSN have certain topological characteristics and each of its nodes is in one of the classes:

1) Susceptible (**S)**, where the sensors are not infected by malware but have susceptible to such software individual computational characteristics.

2) Exposed (**E)**, through the sensors of which the malware has passed, but they cannot transmit it to adjacent sensors due to the individual characteristics of the latter and the features of the received software, as well as their characteristics;

3) Infected (**I)**, whose sensors are infected by malware and can attempt to infect others;

4) Recovered (**R)**, where the sensors of which acquire temporary immunity, after the successful removal of malware, or the establishment of security fixes;

5) Dead (**D)**, in which the sensors are not recoverable (for example, their power was quickly depleted when they were infected with malware; or due to physical damage not related to the software cannot work, etc.).

The individual characteristics due to which each node of WSN is in a particular class are influenced by various factors, including such factors that are not related to the characteristics of malicious software: type of sensor node, its computing power, power consumption, transmission, and information reception, data collection method, routing protocols, etc. To build a model of network operation, we define some vector $x(t) = col(x_1, x_2, x_3, x_4, x_5)$ – vector of quantitative values of network nodes of each of the above five classes (*S*, *E*, *I*, *R*, *D*) at every moment of time observation (t). So, if we consider the operation of network nodes without possible intrusions, the network model will be some system of dynamic equations on time scales in the form:

$$x^{\Delta} = A(t)x + f(t), \tag{1}$$

where $x(t) \in C^1_{rd}(\mathbb{T}_{(t_0)}; \mathbb{R}^5)$ – 5-dimensional vector column of rd-continuous, Δ-differentiable [16] functions, $\mathbb{T}_{(t_0)} := [t_0; \infty)_{\mathbb{T}} = [t_0; \infty) \cap \mathbb{T}$, $A(t)$ – (5×5) matrix, the components of which are rd-continuous functions, $f(t) \in C_{rd}(\mathbb{T}_{(t_0)}; \mathbb{R}^5)$ – rd-continuous vector-valued function. In this model, the value determines the initial time of observation, and the components and $f(t)$ – characteristics of deterministic communication between five classes of nodes of the whole wireless sensor network. In addition, certain individual features of nodes (in particular, their duty cycle, human service factor, etc.) at some point in time make it possible to determine the quantitative parameters of the network itself, which can be mathematically described as some boundary conditions for the system (1). These conditions will include the initial condition regarding the number of nodes available at the initial time t_0:

$$x_1(t_0) + x_2(t_0) + x_3(t_0) + x_4(t_0) + x_5(t_0) = n.$$

All such conditions, in general, can be represented by a linear vector functional $\ell : \mathbb{R}^5 \to \mathbb{R}^m$, where m – total number of conditions. Therefore, taking into account system (1), we will have a boundary value problem:

$$x^{\Delta} = A(t)x + f(t), \ \ell x = \alpha, \tag{2}$$

where $\alpha \in \mathbb{R}^m$ – m-dimension vector constant. Because the condition $m = 5$ is not assumed, then the boundary value problem (1), (2) is a Fredholm. Necessary and sufficient conditions of solvability of such problems using the method of pseudo-inverse matrices [17] were obtained in [18].

Note now that under the influence of malware at certain points in time t_k $(k = 1, 2, \ldots)$ there is a change in the parameters of WSN, which is not related to its natural functioning. Factors in these changes may be related, for example, to the type of malware itself, the mechanism by which it is distributed, or the purpose for which the malicious code is distributed. Then such moments in the proposed model will determine the presence of the corresponding impulsive action:

$$x(t_k + 0) = B_k x(t_k) + a_k, \quad k = 1, 2, \ldots, p. \tag{3}$$

The conditions for the existence of solutions of the Fredholm boundary value problem, which consists of a linear inhomogeneous dynamic system (1), boundary condition in (2), and impulsive action (3), were obtained in [19] as such a result.

T h e o r e m 1. *If* $A(t) \in C_{rd}(\mathbb{T}_{(t_0)}; \mathbb{R}^{5\times5})$, $B_k \in M_5(\mathbb{R})$, $k = \overline{1,p}$, *then inhomogeneous boundary value problem (2), (3) is solvable if and only if the inhomogeneities* $f(t) \in C_{\mathrm{rd}}\left([a;b]_{\mathbb{T}_+} / \{t_k\}; \mathbb{R}^5\right)$, $a_k \in \mathbb{R}^5$, $\forall k = \overline{1,p}$ *and* $\alpha \in \mathbb{R}^m$ *satisfy the following conditions*

$$P_{Q_d^*}(\alpha - \ell F(\cdot)) = \theta_d, \tag{4}$$

where $P_{Q_d^*}$ – $(d \times m)$ *matrix, which consists of* d ($d := m - \operatorname{rank} Q$) *linearly independent rows of the* $(m \times m)$ *matrix* (*orthoprojector*) $P_{Q^*} : \mathbb{R}^m \to N(Q^*)$, $P_{Q^*} := I_m - QQ^+$, Q^+ – $(5 \times m)$ *matrix, which is the unique matrix pseudo-inverse according to Moore–Penrose [17] to the matrix* $Q = \ell S_A(\cdot, t_0)$ – $(m \times 5)$ *constant matrix,* $S_A(t,s)$ – *the impulsive transition matrix, associated with the sequence* $\{B_k, t_k\}_{k=1}^p$ *and normalized at the point* t_0, *which has the form*:

$$S_A(t,s) = \begin{cases} e_A(t,s), & t_{k-1} \le s \le t \le t_k; \\ e_A(t,t_k+0)(I+B_k)e_A(t_k,s), & t_{k-1} \le s \le t_k < t < t_{k+1}; \\ e_A(t,t_k+0)\prod\limits_{s<t_j\le t}\left[(I+B_j)\times\right. & \\ \left.\times e_A(t_j,t_{j-1}+0)\right]\left(I+B_i\right)e_A(t_i,s), & t_{i-1} \le s < t_i < \\ & < \ldots < t_k < t < t_{k+1}, \end{cases}$$

$F(t) = \int\limits_{t_0}^{t} S_A(t,\sigma(s))f(s)\Delta s + \sum\limits_{a<t_j<t} S_A(t,t_j+0)a_j$. *Only for those and only those inhomogeneities,* a_k, α, *for which the condition* (4) *holds, the problem* (2), (3) *possesses an r-parameter* ($r := 5 - \operatorname{rank} Q$) *family of linearly independent solutions*:

$$x(t;c_r) = S_A(t,t_0)P_{Q_r}c_r + G\left(\begin{bmatrix} f \\ a_k \\ \alpha \end{bmatrix}\right)(t), \quad c_r \in \mathbb{R}^r, \tag{5}$$

where P_{Q_r} – $(5 \times r)$ *matrix, which consists of r linearly independent columns of* (5×5) *matrix* (*orthoprojector*) $P_Q : \mathbb{R}^5 \to N(Q)$, $P_Q := I_5 - Q^+Q$ *and*

$G\left(\begin{bmatrix} f \\ a_k \\ \alpha \end{bmatrix}\right)(t) := F(t) + S_A(t,t_0)Q^+\left\{\alpha - \ell\int\limits_{t_0}^{\cdot} S_A(\cdot,\sigma(s))f(s)\Delta s - \ell\sum\limits_{a<t_j<\cdot} S_A(\cdot,t_j+0)a_j\right\}$ – *generalized Green operator of inhomogeneous boundary value problem* (2), (3).

So, having the corresponding numerical values of inhomogeneities, which are obtained from the corresponding conditions of connectivity of node classes, their characteristics and features of malware, it is possible to simulate the operation of the entire wireless sensor network as a boundary value problem for an impulsive dynamic system on time scales of the form:

$$\begin{aligned} &x^{\Delta} = A(t)x + f(t), \ t \in \mathbb{T}_{(t_0)} \\ &x(t_k+0) = B_k x(t_k) + a_k, \quad k = 1,2,\ldots,p, \\ &\ell x = \alpha, \end{aligned}$$

which, under certain conditions (4), gives the predicted consequences in the form of solutions (5).

To prevent unwanted consequences due to the spread of malware, using the proposed model, we have various options, including adjusting the conditions that affect the parameters of inhomogeneities f, a_k and α.

3. Conclusions

The current level of development of equipment and technologies is characterized by the constant expansion of the variety and complexity of mechanical and controllable objects, the functioning of which takes place in a continuously discrete mode over time. One such object is the process of spreading malicious software in wireless sensor networks, the constant growth of which is due to their use as the only type of self-organized data network with the least complexity and low cost.

It should be noted that despite the long history of sensor networks, the concept of their construction has not been fully formed. Therefore, the study of certain properties of such networks is very important for both domestic and world science. Moreover, for strategically important industries of the country, in particular national cybersecurity, the protection of wireless sensor networks is a very important component. This paper proposes a new model of malware distribution, which is described by some boundary value problem for an impulsive dynamic system on time scales.

References:

1. Liu B. Malware propagations in wireless ad hoc networks / B. Liu, W. Zhou, L. Gao, H. Zhou, T. H. Luan, S. Wen // IEEE Trans. Dependable Secure. Comput. 2018. Vol. 15. P. 1016–1026.

2. Wu X. Nodes availability analysis of NB-IoT based heterogeneous wireless sensor networks under malware infection / X. Wu, Q. Cao, J. Jin, Y. Li, H. Zhang // Wirel. Commun. Mob. Comput. 2019. Vol. 2019.

3. Queiruga-Dios A., Encinas A. H., Martín-Vaquero J., Encinas L. H. Malware propagation models in wireless sensor networks: a review, 2016 // International Joint Conference «SOCO'16-CISIS'16-ICEUTE'16». 2017. Vol. 527. P. 648–657.

4. Zhu L., Zhao H., Wang X. Stability and bifurcation analysis in a delayed reaction-diffusion malware propagation model // Comput. Math. Appl. 2015. Vol. 69. P. 852–875.

5. Feng L. Modeling and stability analysis of worm propagation in wireless sensor network / L. Feng, L. Song, Q. Zhao, H. Wang // Math. Probl. Eng. 2015. Vol. 2015. P. 1–8.

6. Shen S. A non-cooperative non-zero-sum game-based dependability assessment of heterogeneous WSNs with malware diffusion / S. Shen, H. Ma, E. Fan, K. Hu, S. Yu, J. Liu, Q. Cao // J. Netw. Comput. Appl. 2017. Vol. 91. P. 26–35.

7. Acarali D. Modelling the spread of botnet malware in IoT-based wireless sensor networks / D. Acarali, M. Rajarajan, N. Komninos, B. B. Zarpelão // Secur. Commun. Netw. 2019. Vol. 2019. https://doi.org/10.1155/2019/3745619.

8. Shen S. SNIRD: disclosing rules of malware spread in heterogeneous wireless sensor networks / S. Shen, H. Zhou, S. Feng, J. Liu, Q. Cao // IEEE Access. 2019. Vol. 7. P. 92881–92892.

9. Wang Y., Li D., Dong N. Cellular automata malware propagation model for WSN based on multi-player evolutionary game // IET Netw. 2018. Vol. 7. P. 129–135.

10. A. M. del Rey, J. H. Guillén, G. R. Sánchez. Modeling malware propagation in wireless sensor networks with individual-based models // Conference of the Spanish Association for Artificial Intelligence. Springer. Cham. Switzerland. 2016. P. 194–203.

11. Wang T. Propagation modeling and defending of a mobile sensor worm in wireless sensor and actuator networks / T. Wang, Q. Wu, S. Wen, Y. Cai, H. Tian, Y. Chen, B. Wang // Sensors. 2017. Vol. 17(1). P. 139.

12. F. K. Batista, Á. M. del Rey, S. Quintero-Bonilla, A. Queiruga-Dios. A SEIR model for computer virus spreading based on cellular automata, 2017 // International Joint Conference «SOCO'17-CISIS'17-ICEUTE'17». 2018. Vol. 649. P. 641–650.

13. Bose A., Shin K. G. Agent-based modeling of malware dynamics in heterogeneous environments // Secur. Commun. Netw. 2013. Vol. 6. P. 1576–1589.

14. Hosseini S., Azgomi M. A., Rahmani A. Agent-based simulation of the dynamics of malware propagation in scale-free networks // Simulation. 2016. Vol. 92. P. 709–722. https://doi.org/10.1177/0037549716656060

15. Batista F. K., del Rey A. M., Queiruga-Dios A. A new individual-based model to simulate malware propagation in wireless sensor networks // Sensors. 2020. Vol 8 (3). P. 410. https://doi.org/10.3390/math8030410.

16. Bohner M., Peterson A. Dynamic equations on time scales. An introduction with applications. MA. Boston: Birkhauser Boston Inc. 2001.

17. Boichuk A. A., Samoilenko A. M. Generalized inverse operators and fredholm boundary-value problems. Netherlands. Utrecht: Koninklijke Brill NV. 2004.

18. Agarwal R. Fredholm boundary value problems for perturbed systems of dynamic equations on time scales /

R. Agarwal, M. Bohner, A. Boˇichuk, O. Strakh // Mathematical Methods in the Applied Sciences. 2014. https://doi.org/10.1002/mma.3356.

19. Strakh O. P. Linear noetherian boundary-value problems for impulsive dynamic systems on a time scale // Journal of Mathematical Sciences. 2014. Vol. 201 (3). P. 400–406. https://doi.org/10.1007/s10958-014-1999-4Lee, S.hyun. & Kim Mi Na (2008) This is my paper // ABC Transactions on ECE, Vol. 10, No. 5, pp.120–122.

Received 20.09.2022

Information about authors:

Yevgen Kotukh – Associate Professor of Cybersecurity in Sumy State University, Ukraine; e-mail: yevgen-kotukh@gmail.com; ORCID: https://orcid.org/0000-0003-4997-620X

Volodymyr Lubchak – a Head of Cybersecurity Department in Sumy State University, Ukraine; e-mail: v.liubchak@dcs.sumdu.edu.ua; ORCID: https://orcid.org/0000-0002-7335-6716

Oleksandr Strakh – Assistant Professor of the Department of Cybersecurity in Sumy State University, Ukraine; e-mail: o.strakh@dcs.sumdu.edu.ua; ORCID: https://orcid.org/0000-0002-7680-5716

RADIOLOCATION AND RADIONAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

УДК 629.7.022 **DOI:10.30837/rt.2022.3.210.08**

В.М. КАРТАШОВ, д-р техн. наук, М.В. РИБНИКОВ, О.В. КАРТАШОВ,
В.О. ПОСОШЕНКО, канд. техн. наук

АНАЛІЗ МЕТОДІВ АКУСТИЧНОЇ ПЕЛЕНГАЦІЇ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

Вступ

Галузь, пов'язана з виробництвом та застосуванням безпілотних літальних апаратів (БПЛА), в даний час інтенсивно розвивається: розширюється номенклатура літальних апаратів, з'являються та реалізуються нові можливості їх застосування у різних галузях діяльності людини. Однак стрімке поширення БПЛА призвело до збільшення потенційної можливості порушення чинного законодавства, а також збільшення кількості різних інцидентів, пов'язаних з їх використанням.

Малі БПЛА є найбільш складними об'єктами виявлення для сучасних засобів розвідки та контролю повітряного простору [1], тому щодо малих БПЛА потрібен зовсім інший підхід до виявлення, ніж пропонують традиційні методи. Специфіка малих БПЛА полягає в тому, що через малі масогабаритні параметри вони можуть бути доставлені і запущені в місцях, заборонених для польотів, а оскільки вони працюють на малих висотах, то для забезпечення виявлення на великих площах знадобиться значна кількість розподілених систем виявлення. Це значно збільшує вартість комплексу виявлення. Серед різних методів виявлення БПЛА акустична локація є найбільш економічно доцільним рішенням. Акустичний метод пеленгації ґрунтується на визначенні напрямку приходу акустичного сигналу, що генерується електродвигунами БПЛА, після чого координати та дистанцію об'єкта можна визначити за допомогою пасивних методів локації з використанням як мінімум двох содарів [2].

Серед різних методів пеленгації джерел акустичних сигналів [3] найбільш популярні методи, засновані на визначенні різниці в часі приходу сигналу в різні точки простору, а саме метод узагальненої крос-кореляції з фазовим перетворенням (GCC-PHAT), і метод, заснований на формуванні променю, а саме алгоритм керованої потужності відгуку з фазовим перетворенням (SRP-PHAT).

В [4, 5] описано практичне застосування алгоритму SRP-PHAT для знаходження кутів приходу сигналу БПЛА. Автори використовують кілька пар мікрофонів, визначають пік функції взаємної кореляції для кожної пари датчиків, а результуючий напрямок приходу сигналу знаходять методом найменших квадратів. Незважаючи на те, що алгоритм є досить простим і мало витратним у обчислювальному плані, його роздільна здатність обмежена шириною діаграми спрямованості, тому для збільшення роздільної здатності потрібно збільшувати кількість елементів мікрофонної решітки (МР). Як відомо [6, 7], основна потужність акустичного сигналу БПЛА зосереджена на низьких частотах, тому для забезпечення прийнятної роздільної здатності знадобиться досить велика апертура МР.

Методи високої роздільної здатності, відомі в літературі [8], дозволяють дещо зменшити розмір апертури МР, проте вони також мають свої недоліки, зумовлені, перш за все, використанням при обробці досить вузького спектру частот корисного сигналу, а отже, малої частини його енергії.

В статті проаналізовано переваги і недоліки відомих методів пеленгації акустичних сигналів з метою вибору найбільш підходящих алгоритмів для пеленгації БПЛА у тих чи інших практичних ситуаціях. Отримання якісних показників аналізованих алгоритмів здійснювалося методом статистичного комп'ютерного моделювання в середовище Matlab.

Класичні методи пеленгації акустичних сигналів

Алгоритм GCC-PHAT ґрунтується на знаходженні функції взаємної кореляції сигналів двох мікрофонів, розміщених у різних точках простору. При цьому функція кореляції визначається як [9]

$$\tilde{R}_{ij}(\tau) = \sum_{k=1}^{L-1} \Phi \cdot X_i(k) \cdot X_j(k)^* \cdot \exp^{j\omega\tau}, \quad (1)$$

де $X(k)$ – комплексна амплітуда вхідних сигналів i-го і j-го мікрофонів; $k = (0...L)$ – відлік дискретного перетворення Фур'є; символ $()^*$ – позначає комплексне спряжене; τ – часова затримка сигналів двох рознесених у просторі каналів. Φ – вагова функція PHAT для відбілювання сигналу, яка позитивно позначається на стійкості даного алгоритму до реверберацій та шумів, і визначається як

$$\Phi = \frac{1}{\left| X_i(k) \cdot X_j(k)^* \right|}. \quad (2)$$

Часова затримка сигналів визначається за найбільшим піком функції кореляції

$$\tilde{\tau} = \arg\max R(\tau). \quad (3)$$

Тоді напрям приходу корисного сигналу при використанні алгоритму GCC-PHAT можна знайти за допомогою виразу

$$\hat{\theta} = \sin^{-1}\left(\frac{v \cdot \tilde{\tau}}{d} \right), \quad (4)$$

де d – відстань між мікрофонами у приймальній акустичній решітці; v – швидкість поширення звуку у атмосфері.

Алгоритм SRP-PHAT можна розглядати як розширення алгоритму GCC-PHAT [10], він заснований на обчисленні та попарному підсумовуванні вихідних даних GCC-PHAT у N елементній МР у сітці кутів сканування θ. Потужність керованого відгуку можна знайти за формулою [11]

$$P(\theta) = \sum_{i}^{N} \sum_{j=i+1}^{N} \tilde{R}_{ij}(\tau(\theta)). \quad (5)$$

Оціночний напрямок приходу сигналу буде відповідати максимуму потужності керованого відгуку

$$\hat{\theta} = \arg\max P(\theta). \quad (6)$$

Алгоритм SRP-PHAT набув своєї популярності завдяки стабільній роботі в реверберіційних середовищах, яка досягається завдяки використанню функції PHAT, що дозволяє не враховувати амплітуду сигналу. В [12] було запропоновано підхід, заснований на середньоарифметичному нормуванні некогерентної комбінації частот (NAM); у ньому на відміну від SRP-PHAT нормування застосовується до амплітуд частот просторового спектру (ПС).

Так, нормований ПС розраховується за формулою [12]

$$P(f,\theta) = \sum_{f=0}^{L} \frac{P(f,\theta)}{\max(P(f))}. \quad (7)$$

На рис. 1 представлено порівняння ПС при впливі на вхід двох сигналів БПЛА, що діють з напрямів 25 та 32 град. ПС отримані за допомогою алгоритмів SRP, SRP-PHAT та SRP-NAM при відношенні сигнал шум (ВСШ) 20 дБ.

Порівняння точності визначення кутових координат приходу акустичних сигналів методами SRP, SRP-PHAT і SRP-NAM проводилося шляхом знаходження середньоквадратичного відхилення (СКВ) між знайденим значенням кута $\hat{\theta}_k$ і істинними значеннями φ_k як функції від ВСШ, за формулою [13]

$$RMSE = \sqrt{E\left[\frac{1}{K}\sum_{j=1}^{J}\sum_{l=1}^{K}[(\hat{\theta}_j - \varphi_j)^2]\right]}, \quad (8)$$

де $j = (1...J)$ – кількість джерел випромінювання, $l = (1...K)$ – кількість випробувань, E – модуль числа.

На рис. 2 зображено графік СКВ, отриманий за кількості випробувань $K = 100$ у процесі моделювання, у діапазоні ВСШ від -20 до 20 дБ для двох сигналів БПЛА, діючих з напрямів 25 і 35 град. Кількість елементів МР – 30, смуга пропускання 0-4000Гц.

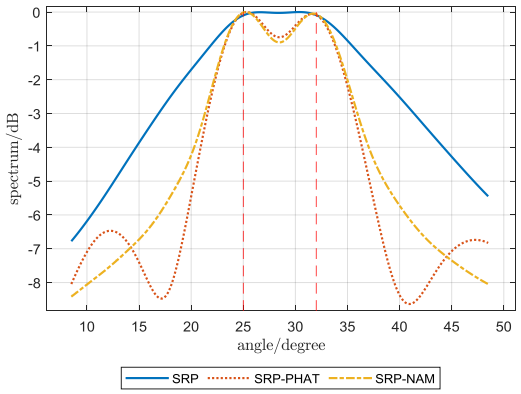


Рис. 1. Просторові спектри, отримані за допомогою алгоритмів SRP, SRP-PHAT та SRP-NAM, при впливі на вхід системи двох сигналів БПЛА, що діють із напрямків 25 та 32 град

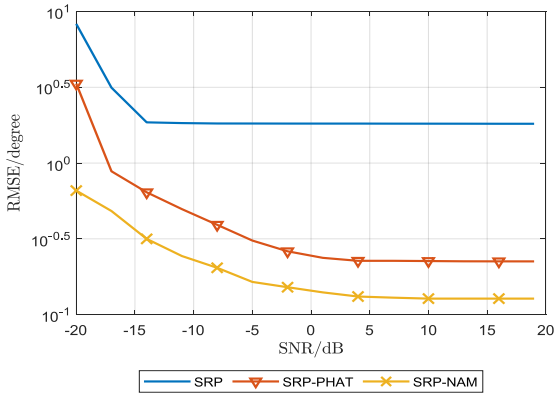


Рис. 2. Графік СКВ визначення кутових координат об'єктів в залежності від ВСШ для адгоритмів SRP, SRP-PHAT та SRP-NAM

Як видно з рис. 1, алгоритм SRP-PHAT розрізняє джерела сигналів на відміну від класичного алгоритму SRP, але при цьому на його виході є хибні піки. Алгоритм SRP-NAM усуває проблему з хибними піками, забезпечуючи при цьому невелике збільшення роздільної здатності.

Як випливає з представленого графіка, алгоритм SRP-NAM забезпечує найкращу точність оцінки та більшу стійкість за умов низьких значень ВСШ. Алгоритм SRP-PHAT дещо програє у зазначених показниках, а алгоритм SRP, який широко обговорюється в літературі та рекомендується до використання на практиці для пеленгування акустичних сигналів БПЛА, має суттєво гірші якісні характеристики порівняно з першими двома алгоритмами.

Методи високої роздільної здатності

Як відомо, методи формування променя схильні до обмеження Релея [14]. У разі, коли є обмеження по апертурі, доцільно застосовувати методи високої роздільної здатності, які дозволяють долати зазначене обмеження. Найбільш відомі такі методи як MUSIC або ROOT-MUSIC, які засновані на розкладанні сигнального та шумового підпросторів [15].

Однак ці алгоритми спочатку розроблялися для вузькосмугових сигналів, і, незважаючи на те, що в сигналі БПЛА присутні вузькосмугові складові, до яких можна застосувати

вузькосмугові алгоритми MUSIC та ROOT-MUSIC, частоти піків спектра не є стаціонарними в часі та змінюються залежно від частоти обертання електродвигуна [16].

Широкосмугова адаптація вузькосмугових алгоритмів полягає у розкладанні сигналів на окремі частотні складові за допомогою швидкого перетворення Фур'є або банків фільтрів [17], та застосування вузькосмугового алгоритму до кожної частотної складової f з подальшим підсумовуванням отриманих просторових спектрів. Такий метод прийнято називати некогерентною широкосмуговою обробкою [18].

Основна ідея алгоритму MUSIC полягає у тому, щоб відокремити сигнал від шуму, використовуючи властивість ортогональності їх просторів через власне розкладання кореляційної матриці прийнятого сигналу.

Враховуючи, що сигнал і шум не корельовані, кореляційну матрицю вхідного сигналу можна представити як [19]

$$R = U_s \Lambda_s U_s^T + U_n \Lambda_n U_n^T, \tag{9}$$

де U_s – матриця власних векторів, що охоплюють сигнальний підпростір, U_n – матриця власних векторів, що охоплюють шумовий підпростір, Λ_s, Λ_n – діагональні матриці власних значень сигналів та шумів.

При сортуванні власних значень кореляційної матриці в порядку зменшення, матриці власних векторів сигналу та шумів можна представити такими виразами

$$U_s = (q_1, ..., q_p), \tag{10}$$

$$U_n = (q_{p+1}, ..., q_N), \tag{11}$$

де q – власні вектори кореляційної матриці, p – кількість джерел випромінювання, N – кількість елементів МР.

З (10), (11) можемо бачити, що максимальна кількість джерел випромінювання, яку можна визначити за допомогою алгоритму MUSIC, дорівнює $p\max = N - 1$, кількість джерел випромінювання до того ж має бути апріорі відома. На практиці кількість джерел випромінювання можна визначити за методами [20, 21].

Просторово-частотний спектр некогерентної MUSIC (IMUSIC) можна знайти за формулою [22]

$$P(f, \theta) = \frac{1}{a(f, \theta) U_n(f) U_n^T(f) a^T(f, \theta)}, \tag{12}$$

де T – операція транспонування; $a(f, \theta)$ – вектор положення, який визначається як

$$a(f, \theta) = \exp(-j 2\pi f (n-1) d \sin\theta / v), \tag{13}$$

де n – порядковий номер мікрофона, d – відстань між мікрофонами.

Напрямок приходу широкосмугового сигналу $\hat{\theta}$ можна знайти за формулою

$$\hat{\theta} = \arg\max\left(\sum_{i-1}^{N} P_i(\theta)\right), \tag{14}$$

де i – індекс частоти.

На рис 3. зображено просторові спектри, отримані з використанням алгоритмів ненормованого IMUSIC та нормованого IMUSIC методом моделювання при дії двох сигналів БПЛА з напрямків 25 та 28 град. Смуга пропускання 0 – 4000 Гц, ВСШ 20 дБ.

Для того щоб зрозуміти який частотний діапазон, слід враховувати при формуванні загальної оцінки кутового положення об'єктів, що спостерігаються, представимо нормований

просторово-частотний спектр, отриманий за допомогою алгоритму IMUSIC при впливі на вхід мікрофонної решітки сигналів двох БПЛА з напрямків 25 і 28 град. (рис. 4).

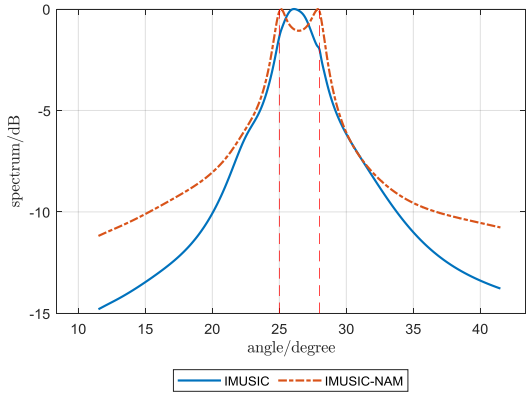


Рис. 3. Просторові спектри, отримані з використанням алгоритмів ненормованого IMUSIC та нормованого IMUSIC при дії двох сигналів БПЛА з напрямків 25 та 28 град

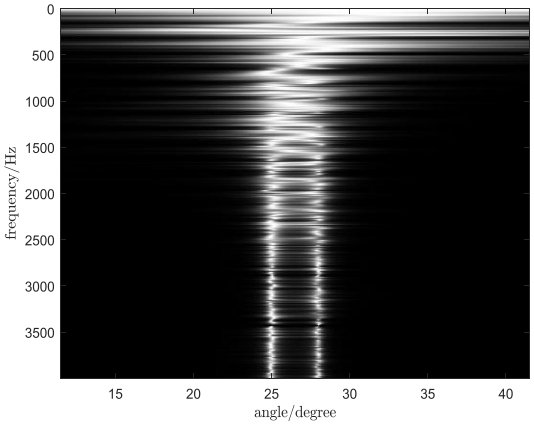


Рис. 4. Нормований просторово-частотний спектр, отриманий з використанням алгоритму MUSIC при впливі на вхід мікрофонної решітки сигналів двох БПЛА з напрямків 25 і 28 град

На рис. 3 у першому випадку, коли просторовий спектр отриманий без нормування по частоті, сигнали не розрізняються по кутах, у другому випадку, при використанні нормування за частотою забезпечується розрізнення сигналів. Як і у випадку з методом формування променю, нормування оцінок дозволяє знизити вплив великих потужностей на низьких частотах.

Як видно з рис. 4, в області низьких звукових частот роздільна здатність не є достатньою, оскільки для більш довгих хвиль потрібна більша відстань між елементами мікрофонної решітки. Також роздільна здатність методу IMUSIC сильно залежить від ВСШ. На рис. 4 можемо бачити, що в області високих звукових частот не всі просторові частоти розрізняються за кутовими координатами. Це пов'язано з низьким ВСШ на певних частотах, оскільки основна потужність сигналу зосереджена в низькочастоної області спектру, а високі частоти схильні до більшого загасання в атмосфері [23].

На рис. 5 представлено просторові спектри, отримані при використанні алгоритмів IMUSIC та IMUSIC-NAM, для діапазону часових частот від 3000 до 4000 Гц при впливі на вхід мікрофоної решітки сигналів двох БПЛА з напрямків 25 і 28 град, відношенні сигналу шум 20 дБ. На рис. 6. представлено просторові спектри, отримані для діапазону частот від 3000 до 4000 Гц з використанням алгоритму IMUSIC при відношенні сигналу шум 0 дБ.

Як видно з рис. 5, на високих частотах нормований некогерентний MUSIC забезпечує більш глибокий провал у просторовому спектрі між джерелами сигналів, ніж NAM-MUSIC, а отже забезпечує кращу роздільну здатність.

У першому випадку при формуванні загальної оцінки враховуються всі оцінки ПС у вказаному діапазоні частот, у другому випадку при формуванні загальної оцінки відсіюються ті з них, у яких не спостерігається роздільної здатності двох сигналів.

Як видно із рис. 6, відсіювання оцінок ПС з недостатньою роздільною здатністю при формуванні загальної оцінки в зазначеному частотному діапазоні, дозволяє істотно збільшити глибину провалу в просторовому спектрі в діапазоні кутів між об'єктами, що спостерігаються, до 1,3 Дб.

Таким чином, метод некогерентної MUSIC здатний забезпечити кращу роздільну здатність порівняно з класичними методами, але його ефективність забезпечується лише при високих ВСШ, що ускладнює його використання в умовах низьких значень ВСШ. Ще один із недоліків некогерентної MUSIC полягає у зниженні ефективності оцінки для когерентних сигналів, проте цю проблему можна вирішити, застосувавши просторове згладжування до кореляційної матриці [22].

Розвитком алгоритму некогерентної MUSIC стосовно задачі виявлення БПЛА є використання відсіювання частотних компонентів з низьким ВСШ при формуванні оцінки, і нормування амлітуд окремих частот з метою зменшення впливу на оцінку потужніших гармонік на низьких частотах.

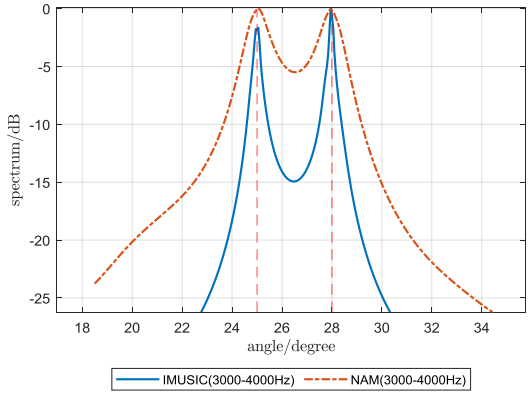


Рис. 5. Просторові спектри, отримані при використанні нормованого та ненормованого IMUSIC для діапазону частот 3000 – 4000 Гц при впливі на вхід антеної решітки сигналів двох БПЛА з напрямків 25 і 28 град

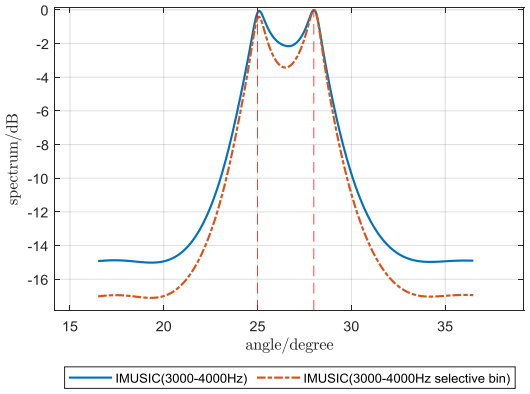


Рис. 6. Просторові спектри, отримані з використанням алгоритму IMUSIC для діапазону частот від 3000 до 4000 Гц щодо сигнал-шум 0 дБ з відсіювання поганих оцінок і без відсіювання

У [23] запропонований некогерентний алгоритм перевірки ортогональності проєктованих підпросторів (TOPS), в якому на відміну від MUSIC не потрібно апріорі знати кількість джерел випромінювання. Основна ідея алгоритму TOPS полягає в тому, щоб визначити, чи ортогональні підпростір сигналу опорної частоти і підпростір шуму проміжної частоти на кожному з можливих напрямків. Якщо ортогональність існує, то напрямок є істинним. У цьому алгоритмі використовується діагональна матриця як функція частоти та кутової координати, яка використовується для перетворення підпростору сигналу опорної частоти на кожну частотну точку, після чого будується нова матриця з шумовим підпростором. Діагональна матриця перетворення має вигляд [21]

$$\left|\Phi(f,\theta)\right|_{n,n} = \exp(-j2\pi f(n-1)d\sin\theta/c). \quad (15)$$

Використовуючи діагональну матрицю перетворення на проміжній частоті $\Delta f = f_i - f_0$ та матрицю власних векторів, що охоплює сигнальний підпростір U_0 на опорній частоті f_0, яку на практиці вибирають як найпотужнішу серед інших частот, перетворену матрицю власних векторів, що охоплюють підпростір сигналу, можна знайти

$$U_i(\theta) = \Phi(\Delta f,\theta)\cdot U_0. \quad (16)$$

Роздільна здатність оцінки просторових кутів приходу сигналу залежить від точності знаходження розрахункової кореляційної матриці, яка, у свою чергу, залежить від кількості накопичених вхідних відділків та ВСШ. При обробці сигналів TOPS застосовується метод проєціювання підпростору, щоб зменшити витік компонентів підпростору сигналу в шумовому підпросторі, що оцінюється. Тоді матриця проєкції $P(\theta)$

$$P(\theta)_i = I - (a^T(\Delta f,\theta)\cdot a(\Delta f,\theta))^{-1}\cdot a^T(\Delta f,\theta)\cdot a(\Delta f,\theta), \quad (17)$$

де I – одинична матриця елементів МР.

Проєційована матриця власних векторів визначається наступним чином:

$$U_i^{'}(\theta) = P_i(\theta) \cdot U_i(\theta). \tag{18}$$

Отримавши матрицю векторів, що охоплюють підпростір сигналу опорної частоти, і матрицю векторів, що охоплюють підпростір шуму на кожній частоті, можна побудувати оціночну матрицю

$$D(\theta)_i = [U_1^{'T}(\theta)G_1 \,|\, U_2^{'T}(\theta)G_2 \,|\cdots U_i^{'T}(\theta)G_i], \tag{19}$$

де G_i – матриця власних векторів, що охоплюють шумовий підпростір проміжної частоти.

Визначити справжній напрямок на об'єкт можна, обчисливши ранг оціночної матриці, оскільки, коли імовірний кут приходу сигналу дорівнюватиме істинному куту приходу, ранг матриці $D(\theta)$ буде відсутнім. У випадку, якщо кут θ приходу сигналу не дорівнює справжньому куту падіння, матриця $D(\theta)$ матиме повний ранг. На практиці через наявність шуму підпростір сигналу матимемо певні помилки, оскільки матриця $D(\theta)$, як правило, не матиме рангу. Ми можемо визначити, наскільки матриця близька до неповноцінного рангу за мінімальним сингулярним значенням матриці. Тоді оцінка напрямку приходу сигналу за допомогою алгоритму TOPS може бути отримана так:

$$\hat{\theta} = \arg\max \frac{1}{\sigma(\theta)_{\min}}, \tag{20}$$

де $\sigma(\theta)_{\min}$ – мінімальне сингулярне значення оцінної матриці $D(\theta)$.

На рис. 7 зображено два просторові спектри, отримані методом TOPS, у разі приходу на вхід МР двох сигналів, з напрямків 25 та 28 град. Кількість елементів МР – 30, відстань між елементами МР кратна половині довжині хвилі на частоті 4000 Гц, опорна частота становить $f_0 = 2000$ Гц, у першому випадку верхня межа смуги пропускання становить 4000 Гц, у другому випадку – 8000 Гц.

Залежності СКВ від ВСШ для алгоритмів IMUSIC, IMUSIC-NAM та TOPS, наведені на рис. 8, також отримані з використанням формули (8), при діапазоні ВСШ від -10 до 20 дБ, при дії БПЛА з напрямків 25 та 32 град. Кількість елементів МР – 30, смуга пропускання 1000 – 4000 Гц.

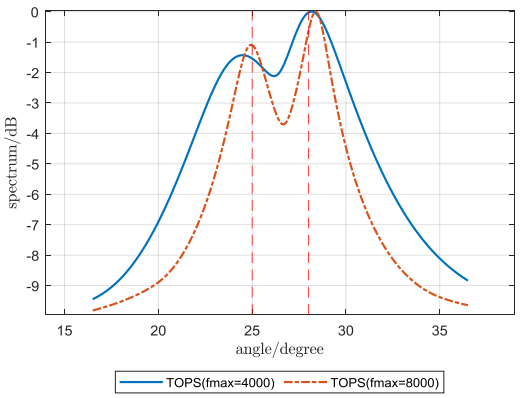


Рис. 7. Просторові спектри, отримані за допомогою алгоритму TOPS, у разі приходу на вхід МР двох сигналів з напрямків 25 та 28 град

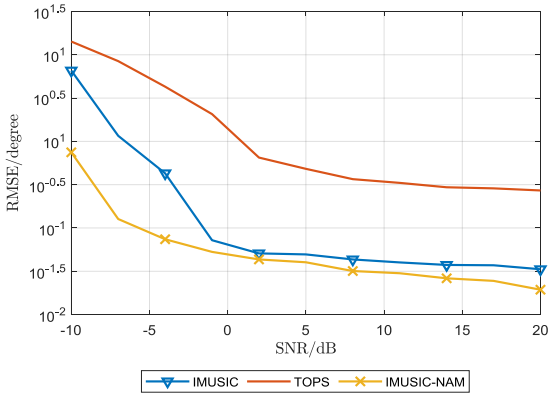


Рис. 8. Графіки СКО залежно від ВСШ для алгоритмів IMUSIC, TOPS та IMUSIC-NAM

На рис. 7 можемо спостерігати деяке відхилення оцінки від справжнього напряму, що є недоліком цього методу. Для збільшення роздільної здатності та зменшення відхилення в оцінці опорну частоту слід вибирати, ґрунтуючись на найбільшій амплітуді частоти у верхній частині смуги пропускання.

З графіків на рис. 8 можемо бачити, що алгоритм TOPS поступається IMUSIC і IMUSIC-NAM в точності оцінок, у той час як IMUSIC-NAM перевершує обидва методи і показує кращу стійкість у області низьких ВСШ.

Варто відзначити стійкість алгоритму до накладання ПС, що дозволяє збільшити роздільну здатність, встановивши відстань між елементами МР кратну середній довжині хвилі в смузі пропускання, а не довжині хвилі, що відповідає найвищій частоті (як у випадку з некогерентним MUSIC, який може сильно «страждати» від накладання діапазону, особливо за високі значення ВСШ).

Алгоритм TOPS є обчислювально більш витратним, ніж IMUSIC. Зниження обчислювальної складності даного алгоритму можна досягти, обмеживши кут сканування, для чого необхідно провести початкову оцінку класичним методом, і далі в межах ширини діаграми спрямованості зробити обробку алгоритмом TOPS, при цьому збільшивши верхню межу смуги пропускання вдвічі.

Висновки

У літературі відомо досить багато методів пеленгації різних джерел випромінювання. Проте завдання пеленгації БЛЛА щодо їх акустичного випромінювання має низку істотних особливостей. Вони пов'язані, перш за все, з особливостями структури акустичного сигналу, що випромінюється БПЛА в процесі польоту, а також з особливостями поширення звукових хвиль в атмосфері, які схильні до значної вітрової рефракції, мають різний рівень згасання в атмосфері різних частотних складових спектра сигналу та ін. Алгоритм пеленгування GCC-PHAT, заснований на визначенні різниці моментів часу приходу сигналу в дві або більше рознесені в просторі точки, є обчислювально економічним і досить простим для визначення напрямку на БПЛА, однак він не здатний розрізняти більше одного джерела випромінювання в межах діаграми спрямованості. Методи, засновані на формуванні променю, також є порівняно простими у реалізації та обчислювально ефективними, вони також більш стійкі при низьких значеннях ВСШ. Алгоритм SRP-NAM має більшу точність визначення кутів, ніж SRP-PHAT, тому він може бути адекватною заміною алгоритму SRP-PHAT.

Методи високої роздільної здатності забезпечують кращу роздільну здатність у напрямку, ніж класичні методи, що у разі обмеження по апертурі МР є позитивним фактором при виборі алгоритму при проєктуванні станції пеленгування БПЛА. Алгоритм IMUSIC-NAM може бути використаний у задачах виявлення та пеленгування БПЛА, оскільки дозволяє вирішити проблему впливу більшої амплітуди сигналу на низьких частотах.

Подальша робота з удосконалення методу некогерентної MUSIC буде спрямована на отримання ефективного алгоритму, здатного адаптивно виділяти основні піки частотного спектру сигналів БПЛА для збільшення працездатності методу в умовах слабкого ВСШ. Метод TOPS – відносно новий некогерентний метод високої роздільної здатності, який не потребує апріорного знання кількості джерел випромінювання. Він має велику стійкість до проникання високих частот ПС сигналу в область низьких частот ПС, що дає перевагу в роздільній здатності, коли довжини хвиль, які враховуються в оцінці, менші за подвійну відстань між елементами МР. У той самий час метод TOPS є більш витратним в обчислювальному відношенні, до того ж забезпечує невелике зміщення оцінки кутових координат. Метод TOPS, хоч і меншою мірою, ніж MUSIC, але також втрачає ефективність для формування оцінок для когерентних сигналів. Подальші зусилля щодо вдосконалення алгоритму TOPS слід спрямувати на зменшення обчислювальних витрат, а також збільшення точності одержуваних оцінок.

Список літератури:

1. Макаренко С. И., Тимошенко А. В., Васильченко А. С. Анализ средств и способов противодействия беспилотным летательным аппаратам. Ч. 1. Беспилотный летательный аппарат как объект обнаружения и поражения // Системы управления, связи и безопасности. 2020. № 1. С. 109-146.

2. Daobilige Su / Robotic Sound Source Mapping using Microphone Arrays / thesis / University of Technology Sydney / 2017.

3. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Бабкин С.И., Шейко С.А., Селезнев И.С. Алгоритмы пеленгации беспилотных летательных аппаратов по их акустическому излучению // Радиотехника. 2019. Вып. 196. С. 22-31.

4. Iqbal Muhammad Amjad., Zhao Zhao., Xu ZhiYong., Saad Ur Rehman,. 3-D Localization of UAV and Detection based on Harmonics Index and Spectral Entropy Criteria To cite this article // OP Conference Series Materials Science and Engineering 853 012037.

5. A. Sedunov, H. Salloum, A. Sutin, N. Sedunov and S. Tsyuryupa. UAV Passive Acoustic Detection // IEEE International Symposium on Technologies for Homeland Security (HST), 2018, P. 1-6..

6. Карташов В. М., Олейников В. Н., Шейко С. А., Бабкин С. И., Корытцев И. В., Зубков О. В. Особенности обнаружения и распознавания малых беспилотных летательных аппаратов // Радиотехника. 2018. Вып. 195. С. 235-243.

7. Cabell, Randolph, Robert G. McSwain and Ferdinand W. Grosveld. Measured Noise from Small Unmanned Aerial Vehicles // Noise-Con 2016.

8. H. Krim and M. Viberg, Two decades of array signal processing research: the parametric approach // IEEE Signal Processing Magazine, vol. 13, P. 67-94, July 1996.

9. C. Knapp., G. Carter. The generalized correlation method for estimation of time delay // IEEE Transactions on Acoustics, Speech, and Signal Processing. 1976.Vol. 24. P. 320–327.

10. A. Ramamurthy, H. Unnikrishnan and K. D. Donohue. Experimental performance analysis of sound source detection with SRP PHAT-β // IEEE Southeastcon. 2009. P. 422-427.

11. Tribikram Kundu. Acoustic source localization // ISSN 0041-624X Vol. 54, Issue 1, 2014. P. 25-38.

12. D. Salvati., C. Drioli., G. L. Foresti. Incoherent Frequency Fusion for Broadband Steered Response Power Algorithms in Noisy Environments // IEEE Signal Process. 2014. Lett, Vol. 21, Num. 5.P. 581-585.

13. Wang Ben., Wang Wei., Gu Yujie., Lei Shujie. Underdetermined DOA Estimation of Quasi-Stationary Signals Using a Partly-Calibrated Array // Sensors. 2017. 17. 702.

14. I.A. McCowan. Robust Speech Recognition using Microphone Arrays, PhD Thesis / Queensland University of Technology, Australia, 2001.

15. S. Visuri, H., V. Koivunen, Subspace-based direction-of-arrival estimation using nonparametric statistics // IEEE Transactions on Signal Processing. 2001. Vol. 49, no. 9. P. 2060-2073.

16. Олейников В.Н., Зубков О.В., Карташов В.М., Корытцев И.В., Бабкин С.И., Шейко С.А., Селезнев И.С. Экспериментальная оценка эффективности аглоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению // Радиотехника. 2019. Вып. 199. С. 29-37.

17. Kartashov V.M., Oleynikov V.N., Zubkov O.V., Korytsev I.V., Babkin S. I., Sheiko S.A., Kolendovskaya M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles // Telecommunications and Radio Engineering. New York. 2020. Vol. 79, №9. P.769-780. DOI: 10.1615/TelecomRadEng.v79.i9.40

18. T. Engin Tuncer., Benjamin Friedlander / Classical and Modern Direction-of-Arrival Estimation, 2009.

19. Feng-Gang Yan., Zhi-Kun Chen., Ming-Jian Sun., Yi Shen., Ming Jin, Two-Dimensional Direction-of-Arrivals Estimation Based on One-Dimensional Search Using Rank Deficiency Principle // International Journal of Antennas and Propagation. 2015, Article ID 127621, 8 pages, 2015.

20. H. Akaike. A new look at the statistical model identification / Automatic Control // IEEE Transactions on. Dec 1974. Vol. 19, no. 6. P. 716–723.

21. M. Wax., T. Kailath. Detection of signals by information theoretic criteria / Acoustics, Speech and Signal Processing // IEEE Transactions on. Apr 1985. Vol. 33, no. 2. P. 387–392.

22. M. Wax, T.-J. Shan and T. Kailath. Spatio-temporal spectral analysis by eigenstructure methods // IEEE Trans. Acoust. Speech Signal Process, Aug. 1984. Vol. ASSP-32, no. 4. P. 817-827.

23. 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 06.09.2022

Відомості про авторів:

Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373

Рибников Микола Володимирович – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: mykola.rybnykov@nure.ua; ORCID: https://orcid.org/0000-0003-1340-8788

Карташов Олександр Володимирович – Харківський національний університет радіоелектроніки, здобувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: mservicekh1@gmail.com

Посошенко Віталій Олександрович – канд. техн. наук, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vitalii.pososhenko@nure.ua; ORCID: https://orcid.org/0000-0003-0867-9161

УДК 629.7.022 **DOI:10.30837/rt.2022.3.210.09**

В.М. ОЛЕЙНІКОВ, канд. техн. наук, В.М. КАРТАШОВ, д-р техн. наук,
С.О. ШЕЙКО, канд. техн. наук, О.В. ЗУБКОВ, канд. техн. наук, О.І. ОЛЕЙНІКОВА

ВИЗНАЧЕННЯ МІСЦЯ ПОЛОЖЕННЯ МАЛОРОЗМІРНИХ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ ЗА АКУСТИЧНИМ ВИПРОМІНЮВАННЯМ

Вступ

Області застосування малорозмірних безпілотних літальних апаратів (БПЛА) постійно розширюються. Використання малорозмірних БПЛА стало серйозним проривом у сфері технічних досягнень останнього десятиліття. На сьогоднішній день вони широко використовуються в різних галузях діяльності людини: моніторинг, інспекція, спостереження, логістика, транспортування, розвідка та картографування, телекомунікації тощо. Універсальність застосування та простота використання призводять до можливості несанкціонованого застосування БПЛА у районах розташування об'єктів критичної інфраструктури. Висока маневреність, невеликі габарити, мала помітність БПЛА породжують проблему їхнього оперативного виявлення. Особливе місце займає повсюдне застосування БПЛА у військовій сфері: це розвідка, спостереження, цілевказівка. Захист критичних інфраструктур та військових об'єктів від таких загроз обумовила необхідність розробки систем виявлення, розпізнавання та локалізації малорозмірних БПЛА. Для вирішення зазначених проблем широкого розповсюдження набули радіолокаційні, оптикоелектронні та акустичні методи.

Радіолокаційний метод [1 – 3] забезпечує необхідну дальність дії, дозволяє визначати координати та здійснювати траєкторні вимірювання, проводити ідентифікацію цілі по сигнатурі радіолокаційного сигналу. Стан навколишнього середовища не істотно впливає на дальність виявлення цілі та інші характеристики станцій. Малі геометричні розміри і відповідно низька радіолокаційна помітность БПЛА обмежують ефективність цього методу. Основний недолік методу – відсутність скритності спостереження.

Пасивні оптичні та інфрачервоні методи [4 – 6] забезпечують скритність вимірювань, можливість розпізнавання об'єктів за аналізом їх зображень, можливість вимірювання відстаней до віддалених об'єктів або об'єктів з малим коефіцієнтом віддзеркалення. Оптичні та інфрачервоні камери спостереження не працюють у складних метеорологічних умовах. Застосування інфрачервоного методу обмежено низьким тепловим випромінюванням БПЛА.

Акустичний метод [7 – 9], незважаючи на відносно невисоку дальність дії, дозволяє визначати просторові координати БПЛА за його акустичним випромінюванням, забезпечувати ідентифікацію БПЛА шляхом формування акустичних сигнатур на основі частотно-часового аналізу звукових сигналів. Відносно невелика дальність виявлення БПЛА для цього методу компенсується пасивним режимом роботи, можливістю локалізації БПЛА, що знаходиться на гранично малих висотах і дальностях. Максимальна дальність виявлення та локаліації БПЛА залежить від швидкості польоту, ракурсу, характеристик діаграми спрямованості акустичного випромінювання, рівня фонового шуму, погодних умов, параметрів системи прийому та обробки акустичних сигналів. Для вирішення проблеми місцезнаходження малорозмірних та низькошвидкісних БПЛА на відносно невеликих дистанціях найбільш ефективним вважається акустичний метод.

Особливості акустичного випромінювання БПЛА

Для розробки алгоритмів визначення місцезнаходження БПЛА необхідно знати особливості формування його акустичного поля та характеристики акустичного випромінювання. Аналіз спектральної щільності потужності (СЩП) акустичного випромінювання різних моделей малорозмірних БПЛА, рис.1, показав наявність вузькосмугових тональних та широкосмугових шумоподібних складових з переважним випромінюванням повітряного гвинта [10 – 13, 14].

При обертанні гвинта виникають коливання тиску повітря, що відбуваються за рахунок витіснення повітря, об’єм якого дорівнює об’єму лопаті гвинта. Це призводить до появи шуму витіснення. Частота проходження лопатей повітряного гвинта дорівнює частоті обертання ротора, помноженої на число лопатей. Спектр шуму гвинта має гармонійні складові частоти обертання ротора і гармоніки лопатевої частоти. Дискретні складові спектру акустичного випромінювання (АВ), пов'язані з шумом обертання та взаємодії, як правило, мають на 15 – 20 дБ вищі рівні, ніж широкосмуговий шум обтікання лопаті. Частота проходження лопатей АВ БПЛА знаходиться в межах від 80 до 250 Гц. Кількість гармонік лопатевої частоти – від 10 до 40. Зі збільшенням відстані до БПЛА, внаслідок поглинання звуку атмосфері, високочастотні гармоніки істотно послаблюються рівня фонового шуму. У процесі польоту БПЛА спектральні лінії акустичного випромінювання розширюються, оскільки контролер БПЛА постійно регулює швидкість кожного двигуна для балансування БПЛА та підтримки стабільного польоту, рис. 1. Більш детально процес регулювання швидкості двигунів, що супроводжується зміною спектра АВ, видно на спектрограмі рис. 2.

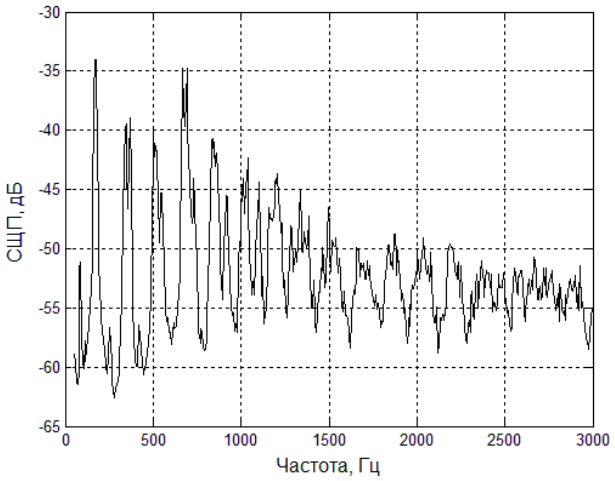


Рис. 1. Спектральні складові АВ квадрокоптера DJI Phantom 3 у процесі польоту

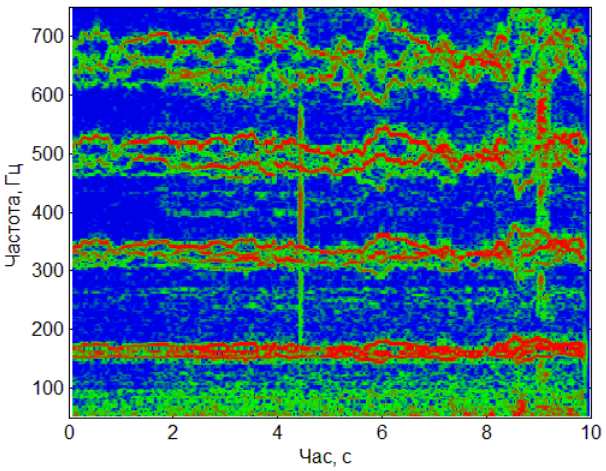


Рис. 2. Спектрограма складових АВ квадрокоптера DJI Phantom 3 у процесі польоту

Характеристика спрямованості АВ БПЛА визначає розподіл випромінюваної акустичної енергії у просторі. На рис. 3 представлено переріз 3D характеристики спрямованості сумарного АВ гвинтомоторної групи квадрокоптера DJI Phantom 3 та характеристики спрямованості АВ його електродвигунів. АВ гвинтомоторної групи має виражену просторову спрямованість, основне випромінювання відбувається у верхню півсферу. З підвищенням номера гармоніки АВ спостерігається ускладнення форми характеристики спрямованості – вона стає більш порізаною, з великою глибиною провалів, ширина пелюсток зменшується, відбувається зміна напрямку основного випромінювання [14].

Характеристика спрямованості АВ електродвигунів БПЛА суттєво відрізняється від характеристики спрямованості випромінювання гвинтомоторної групи, оскільки має іншу природу формування акустичного сигналу. При зміні ракурсу спостереження БПЛА відбувається зміна рівнів спектральних складових акустичного випромінювання у відповідності з характеристикою спрямованості відповідної гармоніки. Це призводить до зміни рівня акустичного сигналу та модифікації спектру АВ, а також впливає на максимальну дальність виявлення та локалізації БПЛА.

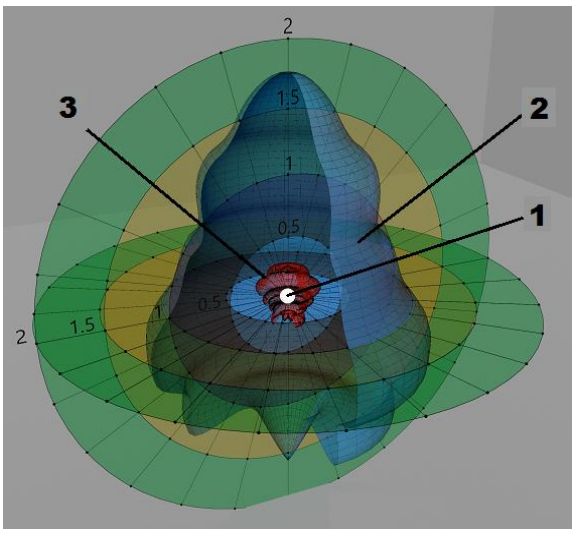


Рис. 3. Переріз 3D характеристики спрямованості АІ БПЛА DJI Phantom 3: 1 – місцезнаходження БПЛА;2 – сумарна характеристика спрямованості АВ гвинтомоторної групи квадрокоптера DJI Phantom 2; 3 – характеристики спрямованості АВ електродвигунів

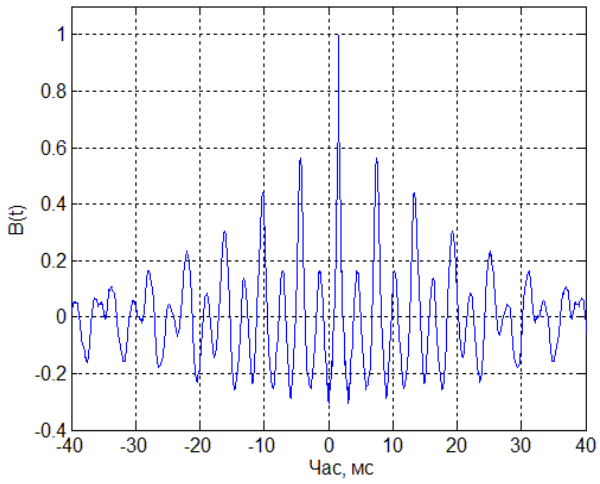


Рис. 4. ВКФ акустичних сигналів БПЛА

Акустичний сигнал гвинтомоторної групи БПЛА відноситься до класу широкосмугових сигналів, що не дозволяє використовувати фазові методи вимірювання кутових координат. З теорії сигналів відомо, що ширина інтервалу кореляції $\Delta\tau$ обернено пропорційна ширині спектра сигналу Δfc:

$$\Delta\tau = \frac{1}{\Delta fc}.$$

Застосування кореляційної обробки широкосмугових сигналів АІ БПЛА забезпечує отримання вузької основної пелюстки взаємної кореляційної функції (ВКФ), рис. 4, що дозволяє реалізувати високу точність просторового дозволу акустичної системи визначення місцезнаходження.

Пасивні методи визначення місцезнаходження БПЛА з акустичного випромінювання

При визначенні кутових координат джерел акустичного випромінювання (ДАВ) знаходять напрямок, що відповідає напрямку нормалі до фронту хвилі, випромінюваної акустичним джерелом. Застосовуючи метод тріангуляції [15], за наявності кількох незалежних вимірів, можна визначити місце розташування ДАВ.

Для визначення координат джерел акустичного випромінювання широке застосування знаходять методи з використанням мікрофонних решіток (МР) [16], які поділяються на класичні методи, методи надроздільності та метод різниці часу приходу (TDOA – time difference of arrival).

Класичний метод обробки полягає у введенні в оброблюваний сигнал відносних часових затримок та підсумовуванні (метод Бартлетта). При скануванні необхідного кутового сектора знаходиться напрямок з найбільшою потужністю, що відповідає оцінці напрямку приходу корисного сигналу. У цьому випадку просторовий спектр формується з використанням наступного виразу:

$$P(\theta) = \frac{S(\theta)R_{xx}S^T(\theta)}{S(\theta)S^T(\theta)} \ ,$$

де $S(\theta)$ – вектор відгуку антеної решітки; R_{xx} – просторова кореляційна матриця розміром N на N елементів.

Можливості використання класичних методів для пеленгування БПЛА щодо їх акустичного випромінювання значною мірою обмежені тим, що пеленгування можливе за наявності в діаграмі спрямованості системи лише одного ДАВ. Присутність у діаграмі спрямованості декількох ДАВ призводить до появи аномальних помилок при оцінці пеленгу, оскільки амплітудно-фазовий розподіл акустичного поля буде суперпозицією декількох хвиль. У силу малих розмірів апертури антеної системи розділення ДАВ в цьому випадку виявляється скрутним.

Методи надроздільності задач пеленгування включають велику групу алгоритмів 16]: алгоритми лінійного передбачення; алгоритми Кейпону, засновані на розкладанні просторової кореляційної матриці за власними векторами; метод MUSIC, EV; алгоритми, засновані на поділі сигнального та шумового просторів – ROOT-MUSIC та ESPRIT та ін.

Для алгоритму Кейпона просторовий спектр розраховується відповідно до виразу

$$P(\theta) = \frac{1}{S(\theta)R_{xx}^{-1}S^T(\theta)} \ .$$

Оцінкою істинного напрямку приходу сигналу є кут, який відповідає піковому значенню спектра. У порівнянні з класичним, метод Кейпона, що характеризується в більшості випадків більш високою роздільною здатностю, вимагає лише додаткового обернення матриці.

Метод різниці часу приходу або різницево-дальномірний метод (РДМ) забезпечує вимірювання кутових координат та дальності до джерела випромінювання. Для реалізації РДМ визначення просторових координат використовуються не менше чотирьох мікрофонів, що дозволяють отримати три незалежні різниці часу приходу до джерела випромінювання.

За певної різниці часу приходу до окремих мікрофонів акустичного сигналу можна визначити кутове положення джерела випромінювання. Вимірювання різниці часу приходу до окремих мікрофонів МР проводиться шляхом обчислення положення максимуму ВКФ сигналів на часовій осі, що приймаються відповідними мікрофонами. Обчислення ВКФ акустичних сигналів БПЛА, отриманих з мікрофонів МР, виконується за формулою:

$$r_{ij}(\tau) = \frac{1}{T}\int_0^T S_i(t)S_j(t+\tau)dt,$$

де $S_{i(}t)$, $S_{j(}t)$ – смугові акустичні сигнали на вході i-го та j-го мікрофонів; T – час аналізу.

ВКФ головної пелюстки акустичних сигналів БПЛА, отриманих з мікрофонів МР рис. 4, має вигляд близький до δ-функції. Це дозволяє при великих співвідношеннях сигнал/шум визначати величину зсуву часу приходу акустичного широкосмугового сигналу між мікрофонами МР з точністю до кроку дискретизації сигналів.

При відомій відстані d між i-м та j-ми мікрофонами та затримки часу приходу τ, можна визначити кутове положення джерела випромінювання:

$$\alpha = \arcsin(\frac{\pi c_{зв}}{d}),$$

де $c_{зв}$ – швидкість поширення звуку.

Для кожної пари мікрофонів вимірювальної бази існує поверхня положення, що визначається як геометричне місце точок, різниця відстаней яких до фокусів (точки в яких розта-

шовані мікрофони) є постійна величина. Поверхня положення для РДМ – гіперболоїд обертання. За наявності трьох пар мікрофонів оцінка розташування джерела сигналу визначається як точка перетину трьох відповідних поверхонь положення.

Різницево-дальномірний метод через можливість застосування завадостійких кореляційних алгоритмів вимірювання часових затримок сигналів має суттєві переваги перед іншими методами.

У розглянутих методах використовується різна кількість мікрофонів та різні конфігурації МР. Кількість мікрофонів, що використовуються для реалізації методів формування променя, значно більша, ніж у методі TDOA, оскільки збільшення числа мікрофонів має домінуючий вплив на точність локалізації ДАВ.

Експериментальна оцінка ефективності розглянутих алгоритмів пеленгування БПЛА [17] у відкритому просторі показують хорошу відповідність заданим значенням пеленгу. Слід зазначити, що з використанням методів Бартлетта і Кейпона істотно зростає дисперсія оцінок пеленгу при азимутах понад 60^{o}. Для методу TDOA дисперсія оцінок пеленгу значно менше, ніж для методів Бартлетта і Кейпона. Результати визначення пеленгу джерела акустичного випромінювання у широкій смузі частот у відкритому просторі показують хорошу відповідність заданим значенням пеленгу при застосуванні методу TDOA.

Розглянемо приклади реалізації пасивних систем для визначення об'єктів у повітряному середовищі з акустичного випромінювання гвинтомоторної групи БПЛА.

У роботі [18] розглядається мікрофонна решітка із чотирьох мікрофонів, розташованих у вершинах тетраедра. Таке розташування мікрофонів дозволило отримати просторову вибірковість при пеленгації БПЛА. Характеристика спрямованості формується шляхом затримки та підсумовування сигналів, прийнятих мікрофонами решітки. Застосування фільтра Вінера на виході формувача характеристики спрямованості дозволило підвищити співвідношення сигнал/завада. Цикл обробки сигналу у смузі 80 – 2000 Гц становив 4 с. За цими даними, з використанням адаптивного фільтра Калмана, проводилася побудова траєкторії руху БПЛА. Максимальна відстань виявлення для БПЛА становила 600 м з ймовірністю 99 % та вірогідністю помилкової тривоги 3 %.

В акустичну систему для відстеження розташування БПЛА університету Чжецзян (Ханчжоу, Китай) [19] входить дві мікрофонні решітки (МР) у формі тетраедра та пристрій обробки сигналів та підготовки даних для реалізації алгоритмів локалізації та відстеження розташування БПЛА. МР складається з чотирьох мікрофонів, розміщених у вершинах тетраедра. Відстань між кожним мікрофоном решітки та центром нижньої грані тетраедра – 1 м, відстань між центрами двох МР – 14 м.

Всі мікрофони мають вітрозахист, захист від впливу гідрометеорів і для ослаблення ефекту багатопроменевості розташовуються над звукопоглинаючим матеріалом. Метод отримання координат БПЛА – різницево-дальномірний. Під час обробки траєкторних даних використовується фільтр Калмана.

Для перевірки ефективності алгоритму оцінки розташування БПЛА використовуються дані GPS трекінгу. На рис. 5 представлено порівняльні оцінки розташування БПЛА за даними GPS та акустичних вимірювань. На рис. 6 представлені оцінки помилки місцезнаходження БПЛА у тестовому польоті, з яких видно, що більш ніж у 95 % помилки оцінки розташування на дистанції до 100 м нижче 6 м та у 80 % оцінки помилки розташування знаходяться в межах 2 м.

У Технологічному інституті Стівенса розроблено систему акустичного спостереження DADS [20], яка призначена для відстеження розташування та ідентифікації БПЛА з акустичного випромінювання. Система спостереження складається з трьох мікрофонних решіток, розташованих по периметру контрольованого об'єкта. Мікрофонні датчики в мікрофонних решітках розташовані у вершинах тетраедра. Дані з кожної решітки передаються каналом Wi-Fi на систему обробки.

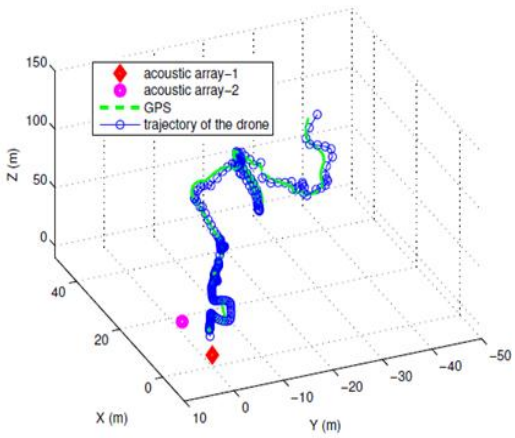


Рис. 5. Порівняльні оцінки розташування БПЛА за даними GPS та за результатами акустичних вимірювань [19]

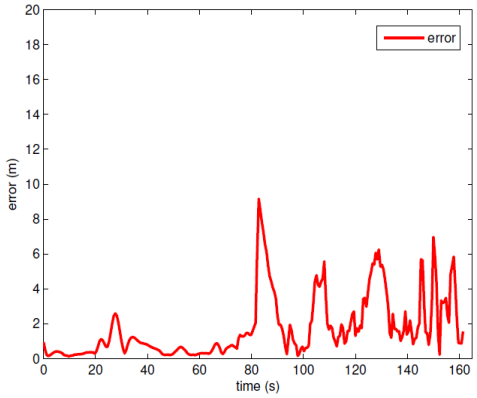


Рис. 6. Оцінки помилки визначення БПЛА в тестовому польоті [19]

Огляд простору ведеться на основі алгоритму, що базується на розрахунку значення взаємної кореляційної функції акустичних сигналів, прийнятих окремими мікрофонними датчиками мікрофонних решіток та їх подальшою обробкою.

Підвищення точності вимірювання різниці моментів часу приходу сигналу досягається збільшенням відношення сигнал/шум сигналу, що приймається за рахунок застосування мікрофонних датчиків. Мікрофонний датчик [20, 21] складається з М електретних мікрофонних капсулів (М=16). Масив електретних капсулів забезпечує більш високу чутливість, ніж використання одиночного мікрофона. Використання таких мікрофонних датчиків покращує відношення сигнал/шум порівняно з одиночним мікрофоном з коефіцієнтом М.

Калібрування орієнтації системи DADS виконується шляхом випромінювання білого шуму з динаміка з відомою позицією GPS, а потім проводиться корекція орієнтації системи DADS на основі різниці між виявленим напрямком та напрямком, розрахованим за даними GPS розташування БПЛА.

Експериментальні дослідження проводилися з різними конфігураціями системи DADS та кількома типами БПЛА у різних умовах довкілля. Так, визначення розташування БПЛА Inspire 2 проводилося тріангуляційним методом із застосуванням двох мікрофонних решіток, рознесених на відстань 60 м, рис. 7.

Результат виміру відображається у вигляді 3D треку спільно з даними GPS, розміщеного на БПЛА. Максимальна дальність місцезнаходження БПЛА становила 250 м при рівні акустичних завад 45 – 50 дБА.

Максимальна відстань виявлення була досягнута в момент знаходження БПЛА на нормалях до акустичних баз (сторін трикутника) і становила 200 м з достовірністю 50 % і 100 м з достовірністю 100 % [20].

Методи локалізації акустичних сигналів засновані на використанні обмеженої кількості мікрофонів та алгоритмі оцінки напрямку приходу акустичних сигналів TDOA, не дозволяють однозначно локалізувати декілька одночасно присутніх БПЛА. Ця проблема вирішується за рахунок використання великих мікрофонних решіток, що збільшує можливості виділення акустичного сигналу з певного напрямку. Це досягається використанням алгоритмів надроздільності сигналів, що дозволяє однозначно визначати місце розташування одного або декількох БПЛА як у двох, так і в трьох вимірах.

Перспективною технологією локалізації акустичного випромінювання БПЛА є поєднання акустичних та оптичних методів, що реалізована в акустичних камерах. Акустичні камери – технічні пристрої, що дозволяють визначити положення джерел акустичних коливань із накладенням оптичного зображення колірної карти інтенсивності акустичного сигналу.

Рис. 7. Схема розгортання 3 вузлів супроводу, прогнозована ймовірність супроводу маневруючого БПЛА з використанням декількох вузлів, біла лінія-GPS трек, зелена лінія – акустичний трек, жовті трикутники – розташування вузлів [20]

Використовуваний метод зветься методом бімформінгу. Метод бімформінгу є способом обробки сигналів мікрофонної решітки з метою визначення просторового розподілу джерел акустичного сигналу. Основна ідея методу [22] полягає в когерентному підсумовуванні акустичних сигналів мікрофонів мікрофонної решітки для збільшення рівня сигналу, що випромінюється з фокусної точки, та мінімізації вкладів сигналів з усіх інших точок, розташованих на площині відображення. У найпростішому випадку методу бімформінгу сигнали мікрофонів складаються із затримкою часу, величина затримки залежить від положення точки фокусування та положення мікрофонів. Для цього заздалегідь розраховуються затримки сигналів від кожного вузла сітки площини відображення до кожного мікрофона МР. Сформований акустичний промінь послідовно проходить всі вузли сітки, при цьому для фокусування на відповідний вузол використовуються розраховані затримки сигналів.

У роботі [23] для виявлення та локалізації БПЛА використовується масив акустичних датчиків акустичної камери Distran Omni 360. Камера складається з мікрофонних решіток з 120 елементів, розташованих на акустично прозорій сферичній поверхні. Для отримання оптичного зображення використовується ширококутна камера, яка розміщена в центрі мікрофонної решітки. Це дозволяє після обробки акустичних сигналів отримати акустичне зображення поверхні відображення, накладене на оптичне зображення досліджуваного об'єкта.

Особливість цієї акустичної камери і те, що на відміну планарних акустичних камер, вона здатна створювати повністю сферичні акустичні зображення, тобто можна відстежувати джерела звуку у будь-якому напрямі. Акустична камера створює сферичні зображення рівня звукового тиску, де кожен елемент роздільної здатності відноситься до рівня звукового тиску звуків, що випромінюються з відповідного напрямку.

Рис. 8. Образ акустичних сигналів, накладений на оптичне зображення квадрокоптера Parrot AR Drone 2.0 [23]

Використання мікрофонної решітки з 120 елементів дозволило, залежно від типу дрона, виявити і відобразити на дистанції від 150 м (Parrot AR Drone 2.0) до 290 м (DJI Phantom2) у присутності заважаючих джерел звуку навколишнього міського середовища. Образ акустичних сигналів БПЛА добре видно на оптичному зображенні рис. 8.

Структура пасивного содару з мінімально можливою кількістю мікрофонів

Розглянемо структуру пасивного содару (ПС) із мінімально можливою кількістю мікрофонів. Для визначення розташування БПЛА використовується пасивний содар, реалізований з використанням чотирьох каналів прийому акустичних сигналів.

Прийом АВ БПЛА здійснюється з використанням конденсаторних вимірювальних мікрофонів Superlux ECM-999. Виходи мікрофонів підключаються через симетричний

аудіоінтерфейс XLR до входів чотириканальної зовнішньої звукової карти Behringer U-Phoria UM2. Акустичний сигнал оцифровувався з частотою дискретизації F_s= 48 кГц та розрядністю 24 біта.

У пасивному содарі використано різницево-дальномірний метод вимірювання координат. Для його реалізації використовується мікрофонні решітки з мінімальною кількістю мікрофонів, що дозволяє реалізувати вимірювальні бази в азимутальній та кутомісній площині. На рис. 9 представлена фотографія мікрофонних решіток М1-М4 пасивного содару, що знята в умовах проведення польових вимірювань. Мікрофонні решітки містять дві азимутальні бази в горизонтальній площині М1-М2 – пеленгатор 1, М2-М3 – пеленгатор 2, (*d1*=2 м, *d2*=2 м) і одну кутомісну базу у вертикальній площині М2-М4- пеленгатор 3, (*d3* = 1 м). Висота установки мікрофонів азимутальних баз 1,3м.

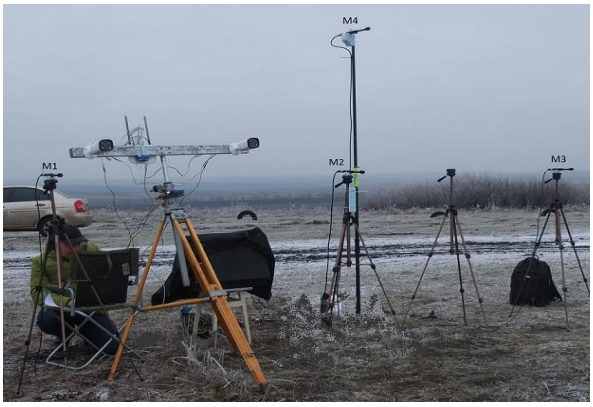


Рис. 9. Мікрофонна решітка М1-М4 пасивного содару

На виході кожного з пеленгаторів формується оцінка кутового положення джерела звуку щодо бази відповідного пеленгатора. Для визначення часу запізнення сигналів між двома мікрофонами знаходиться положення максимуму взаємокореляційної функції. Різниця в часі прибуття перераховується в різницю відстаней до точки розташування джерела звуку. Поверхня положення є поверхнею двопорожнинного гіперболоїда обертання у фокуси якого вміщені мікрофони. Якщо відстані від фокусів до цілі великі порівняно з розмірами бази, то гіперболоїд обертання навколо цілі практично збігається зі своєю асимптотою – конусом, вершина якого збігається з серединою бази. Розташування джерела звуку визначається як точка перетину трьох гіперболоїдів з мікрофонами, які розташовані в їх фокусах.

Обробка результатів вимірювання пасивного содару

Обробка даних пеленгів, отриманих з мікрофонів відповідної бази пеленгатора, включає видалення низькочастотних складових спектра акустичного сигналу БПЛА, що забезпечує виключення впливу атмосферних та техногенних шумів на результати подальшої обробки. Далі проводяться статистично забезпечені вимірювання ВКФ для кожної вимірювальної бази, знаходження положення максимуму головної пелюстки ВКФ, розрахунок азимутів та кута місця за даними відповідних вимірювальних баз, видалення некоректних значень часового ряду кутових даних, згладжування даних медіанним фільтром. Потім проводиться обробка траєкторних вимірювань фільтром Калмана та згладжування даних фільтром ковзного середнього. Додатково проводяться допоміжні розрахунки потужності сигналу, оцінка співвідношення с/ш оцінки ефективної смуги спектра акустичного випромінювання БПЛА.

Чинники, що впливають на величину похибки визначення координат БПЛА

Похибка визначення координат БПЛА різницево-дальномірним методом залежить від точності вимірювання різниці моментів часу приходу сигналу між мікрофонами масиву та розміру азимутальних та кутомісних вимірювальних баз.

Для кожного пеленгатора, що входить до складу пасивного содару, існує поверхня положення для ДАВ, що визначається як геометричне місце точок, різниця відстаней яких до фокусів (*r1, r2*) в яких розташовані мікрофони, є постійна величина. Поверхні положення можна характеризувати товщиною, яка визначається роздільною здатністю по різниці дальностей *r=r1-r2* у пасивному содарі:

$$\Delta r = c/\Delta f_{0,}$$

де f_o – ефективна ширина спектра акустичного сигналу випромінювання БПЛА; c – швидкість звуку.

Як випливає з результатів експериментальних досліджень, ефективна ширина спектра акустичного сигналу випромінювання БПЛА [14] лежить в межах 300 – 700 Гц. Відповідно, очікується величина Δr в межах 0,5 – 1,1 м.

Розглянемо похибки визначення положення БПЛА щодо акустичного випромінювання. На рис. 10 представлено взаємне розташування елементів пеленгатора содару, що працює на основі різницево-дальномірного методу. Джерело акустичного випромінювання розташовано у точці B, d – база між мікрофонами М1, М2, φ – кут, під яким видно базу d. інформаційний параметр

$$W = r1\text{-}r2,$$

де $r1$ і $r2$ – відстань від ДАВ до двох мікрофонів М1, М2.

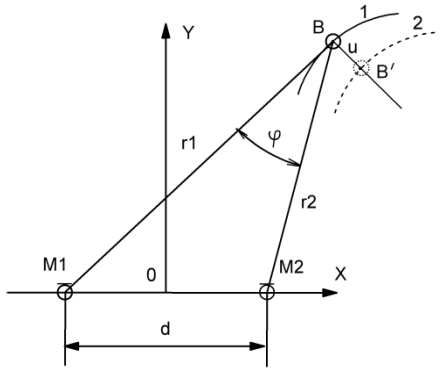


Рис. 10. Зсув лінії положення в різницево-дальномірному содарі

Зсув лінії положення з позиції 1 в позицію 2 (рис. 10), викликаний похибкою вимірювання різниці відстаней ΔW, дорівнює

$$u = \frac{\Delta W}{2\sin\dfrac{\varphi}{2}}.$$

Середньоквадратичний розмір похибки визначення лінії положення σ_u визначається як

$$\sigma_u = \frac{\sigma_W}{2\sin\dfrac{\varphi}{2}},$$

де σ_W – середньоквадратичне значення похибки вимірювання ΔW.

Оскільки в содарі похибка вимірювання параметра W складається з похибок складових її елементів $\sigma_w = \sigma_{r1} + \sigma_{r2} = \sigma_\tau$, отримуємо:

$$\sigma_u = \frac{c\sigma_\tau}{2\sin\dfrac{\varphi}{2}}.$$

Середньоквадратична похибка виміру часу взаємного запізнення визначається як:

$$\sigma_\tau = \frac{\tau_k}{2}\sqrt{\frac{1+\rho_{uv}^2}{n\rho_{uv}^2}},$$

де ρ_{uv} – коефіцієнт кореляції вхідних сигналів $u(t)$ і $v(t)$; τ_k – ширина огинаючої автокореляційної функції корисного сигналу на рівні 0,5; $n=\Delta f_e T_e$ – коефіцієнт кореляційного накопичення; T_e – еквівалентний час інтегрування; Δf_e – еквівалентна ширина спектру.

Міра роздільної здатності за часом взаємного запізнення визначається протяжністю кореляційної функції акустичних сигналів, що оброблюються, і визначається як

$$\Delta r = \frac{c\tau_k}{2}.$$

Таким чином, у різницево-дальномірному содарі середньоквадратичне значення похибки лінійного зміщення σ_u залежить від похибки вимірювання інтервалу різниці часу прибуття σ_τ та кута φ, під яким видно базу пеленгатора.

Мінімальні похибки вимірювання розташування БПЛА виникають за φ=π, тобто. коли БПЛА перебуває безпосередньо над базою пеленгатора. При віддаленні БПЛА від бази пеленгатора похибка зростає обернено пропорційно sin*(φ/2)*. Збільшення розміру бази пеленгатора *d* призводить до збільшення кута φ і відповідно до зменшення похибки вимірювання розташування БПЛА, але при цьому з'являються динамічні похибки вимірювання розташування, пов'язані з ефектом Доплера.

Вплив ефекту Доплера виявляється в тому, що радіальні швидкості джерела звуку щодо кожного з двох мікрофонів вимірювальної бази содару відрізняються через рознесення мікрофонів у просторі. Відповідно, існує різниця радіальних швидкостей джерела звуку між мікрофонами кожної пари мікрофонів вимірювальної бази содару. Це, у свою чергу, призводить до додаткового зсуву положення ВКФ та модифікації координат, що визначаються содаром.

Результати імітаційного моделювання з визначення місцезнаходження БПЛА з використанням пасивного содару

Для дослідження ефективності алгоритмів визначення місцезнаходження в середовищі MATLAB створена імітаційна модель процесу обробки сигналів БПЛА, в якій задаються координати мікрофонів мікрофонної решітки пасивного содару і траєкторія руху БПЛА. Параметри мікрофонних решіток пасивного содару в модельному експерименті відповідають параметрам решітки в натурному експерименті. У процесі моделювання безперервно для кожної дискрети часу розраховуються затримки акустичного сигналу, що з'являються під час прийому сигналу у кожному мікрофоні. Також задається необхідне відношення сигнал/шум сигналів, що обробляються, проводиться облік відзеркалень від земної поверхні відповідно до заданого коефіцієнта відзеркалення. З цих даних формується математична модель сигналів на вході содара. Зіставляючи траєкторні параметри руху БПЛА, задані в імітаційній моделі та результати розрахунку координат БПЛА, отримані шляхом обробки сформованих акустичних сигналів, можна отримати значення абсолютних величин похибки використовуваного алгоритму визначення місцезнаходження та умов проведення модельного експерименту.

Як видно з наведеного аналізу, ключовим параметром, що впливає на точність визначення координат БПЛА різницево-дальномірним методом є точність визначення часу взаємного запізнення. За допомогою імітаційної моделі проаналізуємо величину інстументальних похибок місця визначення БПЛА в залежності від частоти дискретизації акустичного сигналу і дальності. Для проведення експерименту задамо положення центру мікрофонної решітки координатою (0,0,1). Як уже зазначалося, параметри мікрофонних решіток пасивного содару в модельному експерименті відповідають параметрам решітки в натурному експерименті (*d1=2 м, d2=2 м, d3=1 м*). Рух БПЛА задамо горизонтальною лінійною траєкторією з положення (R, -30, 20) в положення (R, 30, 20) зі швидкістю 2 м/с. Задамо співвідношення сигнал/шум 40 дБ. Для кількох значень дальності R (50, 75, 100, 125, 150 м) і частоти дискретизації акустичного сигналу (48, 96, 192 кГц) проведемо вимірювання положення БПЛА на трасі польоту. Залежність середнього значення інструментальної похибки визначення БПЛА від дальності при різних значеннях частоти дискретизації акустичного сигналу представлені на рис. 11.

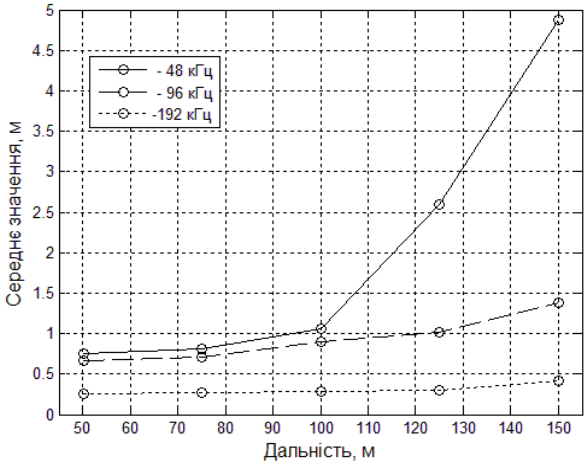
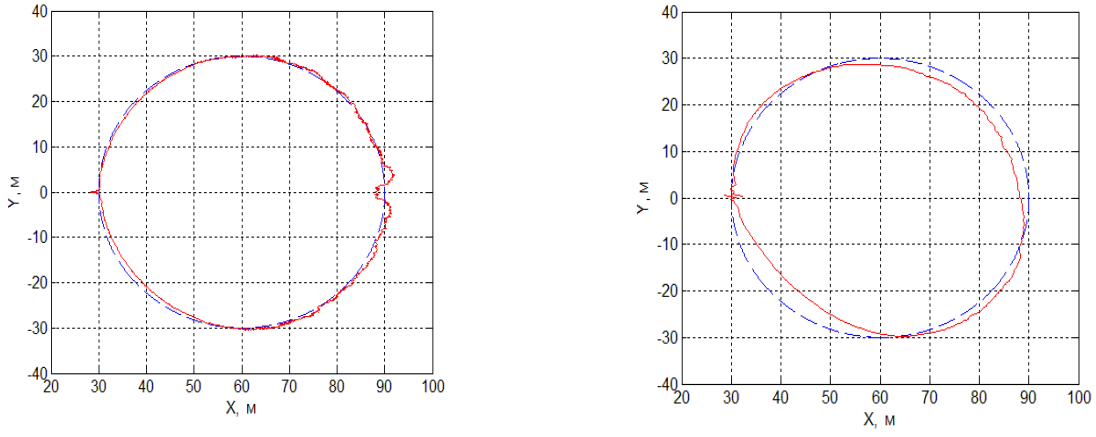


Рис. 11. Залежність середнього значення інструментальної похибки визначення БПЛА від дальності при різних значеннях частоти дискретизації акустичного сигналу

При заданих розмірах баз содару та частоті дискретизації 48 кГц інструментальна похибка місцезнаходження БПЛА істотно зростає на дальності більше 100 м. Очевидні шляхи зниження інструментальної похибки місцезнаходження БПЛА полягають у збільшенні розміру баз і частоти дискретизації акустичного сигналу.

Розглянемо як виявляється ефект Доплера при моделюванні руху БПЛА у просторі. Задамо в модельному експерименті рух БПЛА по колу діаметром 60 м, видалення центру кола від содару 60 м, висота польоту 20 м, кругові швидкості руху 2 та 20 м/с. Швидкість 20 м/с відповідає максимальній швидкості руху квадрокоптера DJI Phantom 3 PRO. Зіставлення заданих траєкторій руху БПЛА та результатів, отриманих у процесі імітаційного моделювання, представлені на рис.12.



а *б*

Рис. 12. Зіставлення заданої траєкторії руху БПЛА (пунктирна лінія) та траєкторії, побудованої за результатами обробки сформованих акустичних сигналів БПЛА при імітаційному моделюванні (суцільна лінія); *а* – кругова швидкість 2 м/с, *б* – кругова швидкість 20 м/с.

Для кругової швидкості 2 м/с спостерігається практично повна відповідність заданих параметрів траєкторії та параметрів траєкторії, отриманих під час імітаційного моделювання. Для кругової швидкості 20 м/с є помітна відмінність заданої та отриманої у модельному експерименті траєкторій. Кругова траєкторія руху БПЛА трансформується на еліптичну. Зі збільшенням кругової швидкості відмінності зростають. Так при круговій швидкості 2 м/с величина середнього значення та середньоквадратичного відхилення оцінок відмінності від заданих параметрів траєкторії становлять 0,82 м та 0,57 м, для 20 м/с відповідно 3,56 м та 1,84 м. Різниця компонент радіальної швидкості стає помітною у разі швидкорухаючого БПЛА і при рознесенні мікрофонів вимірювальних баз на значну відстань.

Результати натурних експериментів щодо визначення місця розташування БПЛА з використанням пасивного содару

Визначення розташування БПЛА за допомогою пасивного содару проводили в польових умовах. Як БПЛА використовувався квадрокоптер DJI Phantom 3 PRO. Вимірювання пара-

метрів атмосфери проводилося мобільною метеостанцією. Температура повітря – 2,4 °С, вологість – 55 %, тиск – 1038 гПа. Акустичний шум довкілля 35 – 37 дБА. Координати БПЛА, отримані стандартним приймачем GPS DJI Phantom 3 PRO, використовувалися для верифікації координат БПЛА, отриманих за допомогою ПС. Географічні координати БПЛА перетворювалися на декартові та поєднувалися з системою координат ПС.

Для верифікації результатів вимірювань розташування БПЛА використовувалися дві тестові траси. На тестовій трасі 1 здійснювався проліт по прямій лінії із зависанням БПЛА над контрольними точками (50, 100, 150, 200 м) на 5 – 10 с на висоті 5 м. Вимірювання довжини ділянок тестової траси для встановлення маркерів контрольних точок проводилося за допомогою геодезичної рулетки. На тестовій трасі 2 здійснювався проліт по спіралі, що піднімається, діаметром 50 м, видалення центру спіралі від содару 50 м, число витків спіралі 3, максимальна висота підйому 60 м.

Обробка результатів натурних вимірювань показала деяку невідповідність результатів вимірювань координат у контрольних точках з даними, отриманими в натурному експерименті. Для коректного використання результатів вимірювань необхідно провести юстування кутового положення азимутальних баз. Для юстування вимірювальної системи ПС використовувався аудіозапис акустичних сигналів БПЛА при зависанні над контрольними точками з видаленням 50 і 100 м. Похибка позиціонування квадрокоптера для цих точок не перевищує 0,5 м.

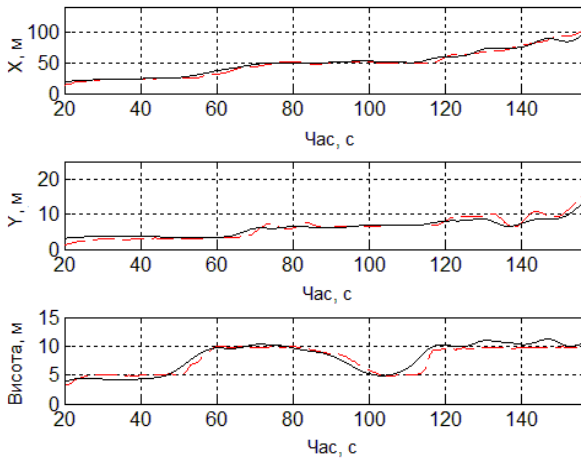


Рис. 13. Траса 1. Зіставлення елементів траєкторій польоту БПЛА, побудованих за координатами, отриманими за допомогою приймача GPS (пунктирна лінія) та пасивного содару (суцільна лінія)

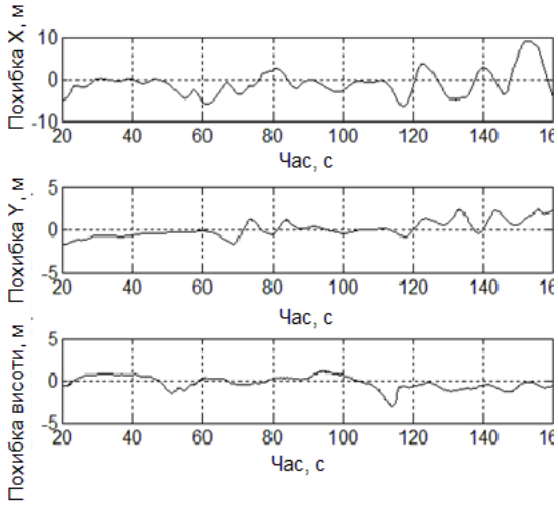


Рис. 14. Траса 1. Похибки визначення розташування БПЛА за результатами акустичних вимірювань

Юстування положення мікрофонів реалізується шляхом введення поправки в одну з азимутальних баз, при якій забезпечується збіг виміряного результату розташування джерела звуку з відомим. Використовуючи дані про висоту у контрольних точках, проводиться юстування ПС по кутомісній базі. Введення поправки до кутового положення азимутальних баз дозволило скоригувати результати вимірювань, насамперед це стосується результатів вимірювання дальності.

Доцільним є використання калібрування содару по акустичним реперним джерелам звуку, встановленим на відомій дальності при проведенні вимірювань.

На рис. 13 наведено зіставлення елементів траєкторій польоту БПЛА, побудованих за координатами, отриманими за допомогою приймача GPS та пасивного содару на трасі 1, представлене у декартових координатах. На рис. 14 показані абсолютні значення похибок позиціонування щодо координат, отриманих за допомогою приймача GPS. Так на дальності до 100 м (до 160 с польоту) абсолютні значення похибок визначення координат у 95 % випадків не перевищують 3 м, визначення висоти не більше 1 м.

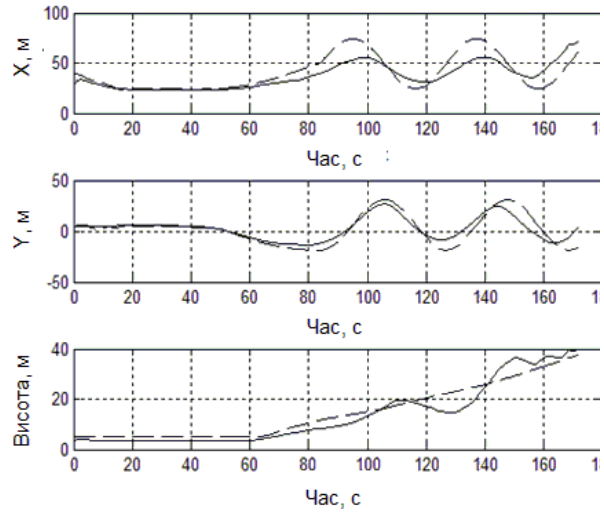


Рис. 15. Траса 2. Зіставлення елементів траєкторій польоту БПЛА, побудованих за координатами, отриманими за допомогою приймача GPS (пунктирна лінія) та пасивного содару (суцільна лінія)

Можна відзначити задовільну відповідність зіставних результатів вимірювань для цієї траси.

На рис. 15 наведено зіставлення елементів траєкторій польоту БПЛА, координат, отриманих за допомогою приймача GPS і пасивного содару на трасі 2, представлене в декартових координатах. Кругова швидкість руху БПЛА – 4 м/с. Абсолютні значення помилок визначення координат для траси 2 істотно вищі ніж на трасі 1, що обумовлено змінами швидкості та напрямки руху БПЛА та проявом ефекту Доплера.

Висновки

1. У роботі розглянуто низку методів визначення місцезнаходження БПЛА за акустичним випромінюванням. Встановлено переваги та недоліки окремих методів. Для визначення координат джерел акустичного випромінювання широке застосування знаходять методи з використанням мікрофонних решіток, які поділяються на класичні методи, методи надроздільності та метод різниці часу приходу.

2. Перспективною технологією визначення місця розташування акустичного випромінювання БПЛА є технологія бімформінгу, в якій здійснюється поєднання акустичних та оптичних методів обробки сигналів.

3. Проаналізовано фактори, що впливають на величину похибки визначення координат БПЛА. Збільшення розмірів баз содару призводить до зменшення помилки виміру місцезнаходження БПЛА, але при цьому збільшуються динамічні помилки виміру розташування БПЛА, пов'язані з ефектом Доплера.

4. Інструментальна похибка вимірювання дальності визначається похибками вимірювання кутів приходу акустичних сигналів по азимутальних баз, яка пов'язана з розміром вимірювальних баз содару і частотою дискретизації акустичних сигналів.

5. Для заданої конфігурації мікрофонної решітки содару методом імітаційного моделювання отримані оцінки інструментальної похибки визначення місцезнаходження БПЛА та похибки, викликані ефектом Доплера. Обробка результатів натурних вимірювань, проведених за допомогою содару, показує, що абсолютні значення похибок визначення координат БПЛА на дальності до 100 м у 95 % випадків не перевищують 3м, визначення висоти не більше 1 м. Вимірювання азимуту та кута місця при цьому досить точні при високій роздільній здатності.

6. Мікрофонні решітки содару є прецизійним пристроєм, зовнішні механічні впливи можуть призвести до порушення її функціонування. Доцільним є регулярне використання калібрування содару по акустичним реперним джерелам звуку, встановленим на відомій дальності.

Список літератури:

1. Oleynikov V., Zubkov O., Kartashov V., Koryttsev I., Sheiko S., Babkin S. Experimental estimation of direction finding to unmanned air vehicles algorithms efficiency by their acoustic emission // 2019 International Scientific-Practical Conference on Problems of Infocommunications – Science and Technology. 2019. P.175 – 178.

2. Макаренко С. И., Тимошенко А. В., Васильченко А. С. Анализ средств и способов противодействия беспилотным летательным аппаратам. Ч. 1. Беспилотный летательный аппарат как объект обнаружения и поражения // Системы управления, связи и безопасности. 2020. № 1. С. 109 –146.

3. Карташов В.М. и др. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : ХНУРЭ, 2014. 312 с.

4. Kartashov V. Oleynikov O. Zubkov, S. Sheiko. Optical detection of unmanned air vehicles on a video stream in a real-time // The Fourth International Conferenceon Information and Telecommunication Technologies and Radio Electronics (UkrMiCo'2019). Odessa, Ukraine. 2019. 4 p.

5. Карташов В.М., Олейников В.Н., Колендовская М.М., Тимошенко Л.П., Капуста А.И., Рыбников Н.В. Комплексирование изображений при обнаружении беспилотных летательных аппаратов // Радіотехніка. 2020. Вип. 201. С. 120 – 129.

6. Sergiyenko O., Rodriguez-Quiñonez J.C. Developing and Applying Optoelectronics in Machine Vision. IGI Global, 2016. 341 p.

7. Oleynikov V.N., Zubkov O.V., Kartashov V.M., Korytsev I.V., Babkin S.I., Sheiko S.A. Investigation of detection and recognition efficiency of small unmanned aerial vehicles on their acoustic radiation // Telecommunications and Radio Engineering. 2019. Vol. 78, Iss. 9. P.759 – 770.

8. Kartashov V., Oleynikov V., Koryttsev I., Sheyko S., Zubkov O., Babkin S., Selieznov I. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles // 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). 2020. 4 p.

9. Kartashov V.M., Oleynikov V.N., Sheyko S.A., Babkin S.I., Koryttsev I.V., Zubkov O.V., Anokhin M.A. Information characteristics of sound radiation of small unmanned aerial vehicles // Telecommunications and Radio Engineering. 2018. Vol.77, Iss. 10. pp. 915 – 924.

10. Олейников В.Н., Шейко С.А., Бабкин С.И. Исследование характеристик акустического излучения малых беспилотных летательных аппаратов// Сб. науч. трудов VI Междунар. радиоэлектронного форума "Прикладная радиоэлектроника. Состояние и перспективы развития (МРФ-2017)" Междунар. науч. конф. "Радиолокация. Спутниковая навигация. Радиомониторинг". 24-26 октября 2017 г. Харьков, Украина. Харьков : Точка. С.107 – 111.

11. Kartashov,V.,Oleynikov,V.,Koryttsev,I.,Zubkov,O.,Babkin,S.,Sheiko,S. Processing and recognition of small unmanned vehicles' sound signals. // 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, 2018, P. 1 – 5.

12. Карташов В.М., Олейников В.Н., Шейко С.А., Бабкин С.И., Корытцев И.В., Зубков О.В., Анохин М.А. Информационные характеристики звуковых сигналов малых беспилотных летательных аппаратов // Радиотехника. 2017. Вып 191. С. 181 – 187.

13. Kartashov V.M., Oleynikov V.N, Sheyko S.A., Babkin S.I., Koryttsev I.V., Zubkov O.V., Anokhin M.A. Information characteristics of sound radiation of small unmanned aerial vehicles // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhika). 2018. Vol. 77(10). P. 915 – 924.

14. 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.

15. Алмазов В.Б. Методы пассивной радиолокации. Харьков : Изд-во Военной инженерной радиотехнической академии противовоздушной обороны имени Л.А. Говорова, 1974. 85 с.

16. В.М. Карташов, И.В. Корытцев, В.Н. Олейников, О.В. Зубков, С.И. Бабкин, С.А. Шейко, Н.А. Левский, И.С. Селезнев. Алгоритмы пеленгации беспилотных летательных аппаратов по их акустическому излучению // Радиотехника. Вып.196. 2019. С. 22 – 31.

17. Олейников В.Н., Зубков О.В., Карташов В.М., Корытцев И.В., Бабкин С.И., Шейко С.А., Селезнев И.С. Экспериментальная оценка эффективности аглоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению // Радиотехника. 2019. Вып. 199. С. 29 – 37.

18. M. Benyamin and G. H. Goldman. Acoustic detection and tracking of a Class I UAS with a small tetrahedral microphone array. Adelphi, MD, 2014 [Online]. Available: http://www.arl.army.mil/arlreports/2014/ARL-TR-7086.pdf.

19. Xianyu Chang A. Surveillance System for Drone Localization and Tracking Using Acoustic Arrays / Xianyu Chang, Chaoqun Yang, Zhiguo Shi // Published Computer Science. 2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM) 2018. URL: https://ieeexplore.ieee.org/document/8448409 (дата обращения: 22.10.2020) .

20. Sedunov A. et al. Stevens drone detection acoustic system and experiments in acoustics UAV tracking // 2019 IEEE International Symposium on Technologies for Homeland Security (HST). IEEE, 2019. 1 – 7.

21. Alexander Sedunov. Passive acoustic localization of small aircraft / Alexander Sedunov, Alexander Sutin, Hady Salloum, Nikolay Sedunov // 166th Meeting of the Acoustical Society of America, San Francisco, CA, 02-06 December 2013, Signal Processing in Acoustics: Paper 2pSP2. URL: https://www.researchgate.net/publication/258249414_Passive_acoustic_localization_of_small_aircraft .

22. Johnson D.H., Dudgeon D.E. Array signal processing: concepts and techniques. 1st ed. Signal processing series. Prentice – Hall, Upper Saddle River, NJ, 1993.

23. Detection and tracking of drones using advanced acoustic cameras Joël Busset, Florian Perrodin, Peter Wellig, Beat Ott, Kurt Heutschi, Torben Rühl, Thomas Nussbaumer Proceedings Volume 9647, Unmanned/Unattended Sensors and Sensor Networks XI; and Advanced Free-Space Optical Communication Techniques and Applications; 96470F (2015)

Надійшла до редколегії 30.08.2022

Відомості про авторів:

Олейніков Володимир Миколайович – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vladimir.oleinikov@nure.ua; ORCID: https://orcid.org/0000-0002-3358-5987 .

Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; e-mail: volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373

Шейко Сергій Олександрович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, e-mail: sergiy.sheiko@nure.ua, ORCID: https://orcid.org/0000-0003-1638-4478.

Зубков Олег Вікторович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, e-mail: Oleh.zubkov@nure.ua, ORCID: https://orcid.org/0000-0002-8528-6540.

Олейнікова Олена Іванівна – Харківський національний університет радіоелектроніки, старший викладач кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна, e-mail: elena.oleynikova@nure.ua, ORCID: https://orcid.org/0000-0002-5601-7018

УДК 621.372(075) **DOI:10.30837/rt.2022.3.210.10**

В.М. КАРТАШОВ, д-р техн. наук, В.М. ОЛЕЙНІКОВ, канд. техн. наук,
І.С. СЕЛЄЗНЬОВ, О.В. КАРТАШОВ

ДІАГРАМИ СПРЯМОВАНОСТІ АКУСТИЧНОГО ВИПРОМІНЮВАННЯ БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

Вступ

Безпілотні літальні апарати (БПЛА) набули широкого застосування в багатьох областях людської діяльності . Вони здатні виконувати великий набір корисних функцій [1 – 3]: доставка вантажів, здійснення аерофотозйомки поверхні, виявлення порушень технічного стану об'єктів, оцінка ступеня вирубки лісових масивів тощо. Військові використовують БПЛА для виконання таких задач, як розвідка місцевості, ретрансляція радіосигналів, постановка перешкод радіозасобам супротивника та ін. В той же час БПЛА можуть нести також значну фізичну або інформаційну загрозу у господарській діяльності, приватному житті людей [2, 4]. Тому задача виявлення безпілотних літальних апаратів і протидії їх несанкціонованим діям є дуже актуальною.

Найчастіше для виявлення безпілотних літальних апаратів використовуються радіолокаційні, акустичні та оптичні методи і засоби [1, 2, 5]. Функціонування пасивних акустичних локаторів – содарів засновано на прийомі звукових коливань пружного середовища, випромінюваних БПЛА у процесі польоту. Основними джерелами акустичного випромінювання безпілотних літальних апаратів є двигуни БПЛА та їх гвинти.

Дослідженням шумових показників БПЛА присвячені роботи [5 – 15]. Як випливає з результатів досліджень, сумарний спектр акустичного випромінювання (АВ) малого БПЛА обумовлений гармонійними та широкосмуговими складовими. У шумі силової установки БПЛА, що має поршневий двигун повітряного охолодження, особливо за відсутності в його вихлопному тракті глушника, визначальним джерелом акустичного випромінювання є саме двигун. Експериментальні дослідження шумів БПЛА проводилися в різних умовах: у камері, що немає луни [11], та в реальних умовах при польотах в атмосфері [7, 10, 12, 13]. Значний інтерес мають для практики АВ БПЛА з електричними гвинтомоторними системами, актуальним є також питання вивчення особливостей формування сумарного звукового поля у процесі їхнього польоту.

Рівень акустичного шуму електричних рухових систем значно нижчий, ніж двигунів внутрішнього згоряння. У разі застосування малих БПЛА, що мають малі габаритні розміри та малопотужні електродвигуни і не мають радіоканалу керування, основним напрямом їх виявлення стають акустичні спостереження [16, 17].

Таким чином, в області дослідження АВ БПЛА отримано деякі результати, але вони досить суперечливі. В деяких літературних джерелах стверджується, що АВ безпілотних літальних апаратів є практично ізотропним, а результати експериментів інших авторів показують, що воно є досить спрямованим. Тому потрібне уточнення наявних характеристик АВ БПЛА. Переважна кількість експериментальних результатів у літературі [3 – 16], присвячених цьому питанню, представлені у вигляді двомірних залежностей. У даному дослідженні ставиться завдання з отримання тривимірних залежностей, що дозволить більш наочно представити отримані в експерименті результати вимірювань.

1. Опис експериментальної установки

Спрямованість звукового поля безпілотних літальних апаратів характеризується деякими загальними закономірностями, а, з іншого боку, кожен тип БПЛА має свої певні особливості, які також позначаються в процесі його виявлення та пеленгації. Шум повітряного гвинта утворюється, в основному, внаслідок взаємодії лопатей з навколишнім середовищем у процесі створення тяги та при витісненні повітря з фіксованого обсягу середовища. Генерація

акустичного випромінювання може відбуватися також і при аеродинамічній взаємодії лопатей з турбулентними утвореннями в потоці, що набігає. Відповідно до цього шум малозавантаженого гвинта зазвичай поділяють на шум обертання та широкосмуговий «вихревий» шум.

Об’єктом дослідження в даній роботі виступав БПЛА громадянського типу «Dji Phantom 3», зовнішній вигляд якого представлено на рис. 1 із такими характеристиками:

- тип планера: мультикоптер;
- кількість гвинтів: 4;
- кількість лопатей на гвинтах: 2;
- тип двигунів: електричні;
- маса: 1280 грам;
- максимальний час польоту: 25 хв;
- максимальна швидкість польоту: 58 км/год;
- максимальна висота польоту: 500 м.

Рис. 1. Зовнішній вигляд БПЛА «Dji Phantom 3»

Експеримент з запису АВ БПЛА проводився у «заглушеній» камері з розмірами 3м*3м*2.5 м (ширина*довжина*висота), стіни якої покриті звукопоглинальними панелями із поверхнею спеціальної геометричної форми, що має вигляд пірамід.

Мікрофон розташовувався на металевій жорстко закріпленій штанзі, яка має можливість обертання навколо об’єкту дослідження та фіксації у необхідному положенні (рис. 2). Експериментальна установка включає БПЛА, закріплений на поворотній штанзі, мікрофон для запису звуку, та поворотну штангу для мікрофону.

Задача з отримання характеристик спрямованості літального апарату виконується наступним чином: БПЛА обертається на поворотному пристрої за азимутом, займаючи деякі фіксовані положення через певні проміжки. В певному секторі та на певній відстані від БПЛА обертається за кутом місця мікрофон, який фіксує рівень акустичного тиску шуму БПЛА. Так, при фіксованому кутовому положенні БПЛА мікрофон обертається від 0 до 180°, фіксуючи в кожному кутовому положенні рівень шуму у просторі навколо БПЛА. За отриманими результатами далі будуються характеристики спрямованості акустичного шумового поля літального апарату.

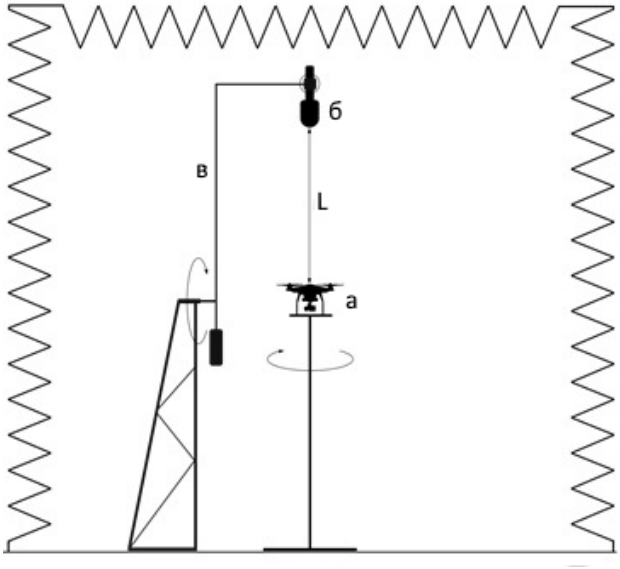


Рис. 2. Схема розташування засобів експериментальної установки для вимірювання характеристик спрямованості акустичного випромінювання БПЛА у вертикальній площині, вид спереду:
а – об’єкт дослідження; б – мікрофон; в – штанга мікрофону;
L – відстань від об’єкта вимірювання до мікрофону (1 м)

Акустичне випромінювання об’єкту дослідження записувалось з використанням мікрофону «Superlux ECM999», який має так характеристики:

- тип: конденсаторний;
- характеристика спрямованості: неспрямований;
- частотний діапазон: 20 – 20 кГц;
- динамічний діапазон: 106 дБ;
- сигнал/шум: 70 дБ;
- імпеданс: 200 Ом;
- максимальний звуковий тиск: 132 дБ.

Мікрофон підключався до зовнішньої звукової карти «Behringer U-Phoria UMC404HD» з використанням аудіо-інтерфейсу XLR, якій забезпечує можливість підключення необхідної апаратури при записі акустичного випромінювання (рис. 3).

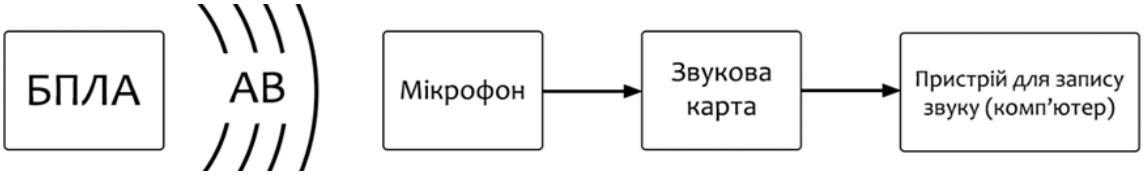
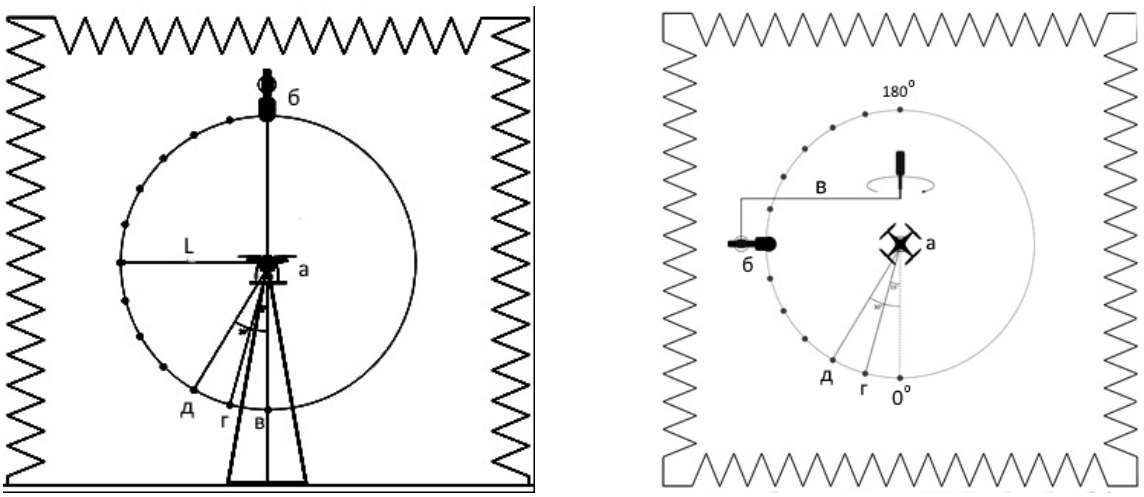


Рис. 3. Функціональна схема з запису акустичного випромінювання БПЛА

Під час тестових замірів було встановлено, що відстані у 1 м достатньо для мінімізації впливу повітряного потоку від лопатей на запис та отримання результатів вимірювання, придатних для подальшої обробки програмними засобами.

Запис АВ БПЛА у вертикальній площині проводився із кроком у 15°, починаючи від 0° до 180°. Точки із вертикальним зміщенням у просторі, у яких проводився запис, позначено на рис. 4, *а* позначками – в, г, д.

На рис. 2 та 4 літера L позначає відстань від БПЛА до мікрофону, яка дорівнює 1 м.



а *б*

Рис. 4. Схема розташування компонентів експериментальної установки для вимірювання характеристик спрямованості звукового поля БПЛА: *а* – у вертикальній площині, вид збоку, *б* – у горизонтальній площині; а – об’єкт дослідження; б – мікрофон; в – штанга мікрофону; г, д – точки, у яких проводився запис сигналу

Для того щоб отримати характеристики спрямованості АВ БПЛА у об’ємному вигляді, потрібно мати дані у вертикальній і у горизонтальній площині. Запис АВ БПЛА у горизонтальній площині також проводився із кроком у 15° в діапазоні кутів від 0° до 180°. Точки у просторі, у яких виконувався запис із горизонтальним зміщенням, позначено на рис. 4, *б* позначками – г, д.

Запис АВ проводився в першому експерименті спочатку при відсутності, а далі при наявності гвинтів, на повній швидкості їх обертання, а у другому експерименті – у режимі статичного польоту БПЛА (планування на місці).

Після того, як акустичне випромінювання об’єкта досліджень було записано єдиною звуковою доріжкою, звуковий ряд було «нарізано» у програмному пакеті Sony Sound Forge на окремі фрагменти, що відповідають точкам у просторі, у яких проводився запис.

2. 2Д подання характеристик звукового поля БПЛА

За допомогою інструментів програмного пакету MathLab було виконано спектральний аналіз акустичного сигналу БПЛА, якій подано на рис. 5, *а*. З рис. 5, *а* видно, що АВ БПЛА представляє собою за структурою широкосмуговий сигнал. Найбільшу потужність мають спектральні складові у частотному діапазоні до 500 Гц, де найбільшою за амплітудою є перша гармоніка, а далі має місце зменшення складових спектру до рівня шуму навколишнього середовища.

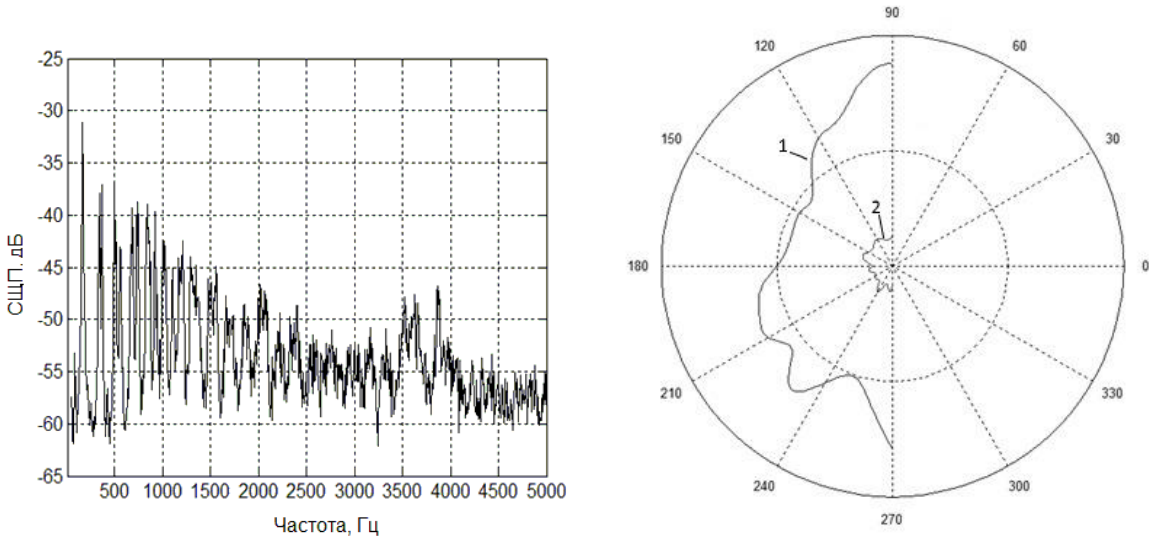
Звукове поле повітряного гвинта формується при періодичній дії на повітряне середовище лопатей БПЛА [10, 12, 14, 15]. Коливання тиску повітря за рахунок витіснення із середовища обсягу, рівного обсягу лопаті гвинта, при його обертанні, призводять до появи шуму витіснення. Основна частота проходження лопатей повітряного гвинта дорівнює частоті обертання ротора, помноженої на число лопатей. Спектр шуму гвинта має гармонійні складові частоти обертання ротора і гармоніки лопатевої частоти (рис. 5, *а*). Частоти гармонійних складових у спектрі шуму повітряного гвинта визначаються відповідно до виразу $f_в = knN$, де k – номер гармоніки, n – частота обертання ротора (про/с), N – число лопатей.

Крім цього, внаслідок вихору біля лопатей повітряного гвинта виникає так званий «вихровий звук» – широкосмуговий шум обтікання лопаті. При обтіканні лопаті повітрям утворюється прикордонний шар, в якому зосереджено дію в'язкості, а стікання з лопаті прикордонного шару призводить до вихроутворення. Інтенсивність вихорів залежить від форми

лопаті і швидкості потоку, що набігає. Спектр шуму вихроутворення є безперервним за частотою (рис. 5, *а*).

Дискретні складові спектру АВ, пов'язані з шумом обертання та взаємодії, і, як правило, мають на 15 – 20 дБ вищі рівні, ніж широкосмуговий шум обтікання лопаті.

Діаграми спрямованості АВ БПЛА у горизонтальній площині, для випадку обертання усіх гвинтів на максимальній швидкості (позначена цифрою 1), та АВ двигунів без гвинтів (цифра 2) представлені на рис. 5, *б*.



а *б*

Рис. 5. *а* – Спектр акустичного випромінювання квадрокоптера «Dji Phantom 3»,
б – Діаграми спрямованості АВ гвинтів (1) та двигунів БПЛА (2) у вертикальній площині

З рис. 5, *б* бачимо, що при відсутності гвинтів акустичне випромінювання досліджуваного об’єкта значне слабкіше за рівнем.

На рис. 6 подано діаграми спрямованості перших чотирьох гармонійних складових спектру АВ у горизонтальній площині для випадку присутності гвинтів БПЛА та максимальної швидкості їх обертання. Також слід зазначити, що характеристики, зображені на рис. 5 та 6, були отримані у вільному просторі.

Аналіз спрямованості АВ БПЛА у вертикальній площині у певної смуги частот показує, що з підвищенням номеру гармоніки спостерігаються зміна форми характеристики спрямованості.

Аналіз спрямованості АВ БПЛА у вертикальній площині у певної смуги частот показує, що з підвищенням номеру гармоніки спостерігаються зміна форми характеристики спрямованості.

З отриманих результатів випливає, що залежно від ракурсу спостереження БПЛА, спектральні складові акустичного випромінювання на частотах гармонік, що визначаються характеристиками спрямованості, мають різні рівні.

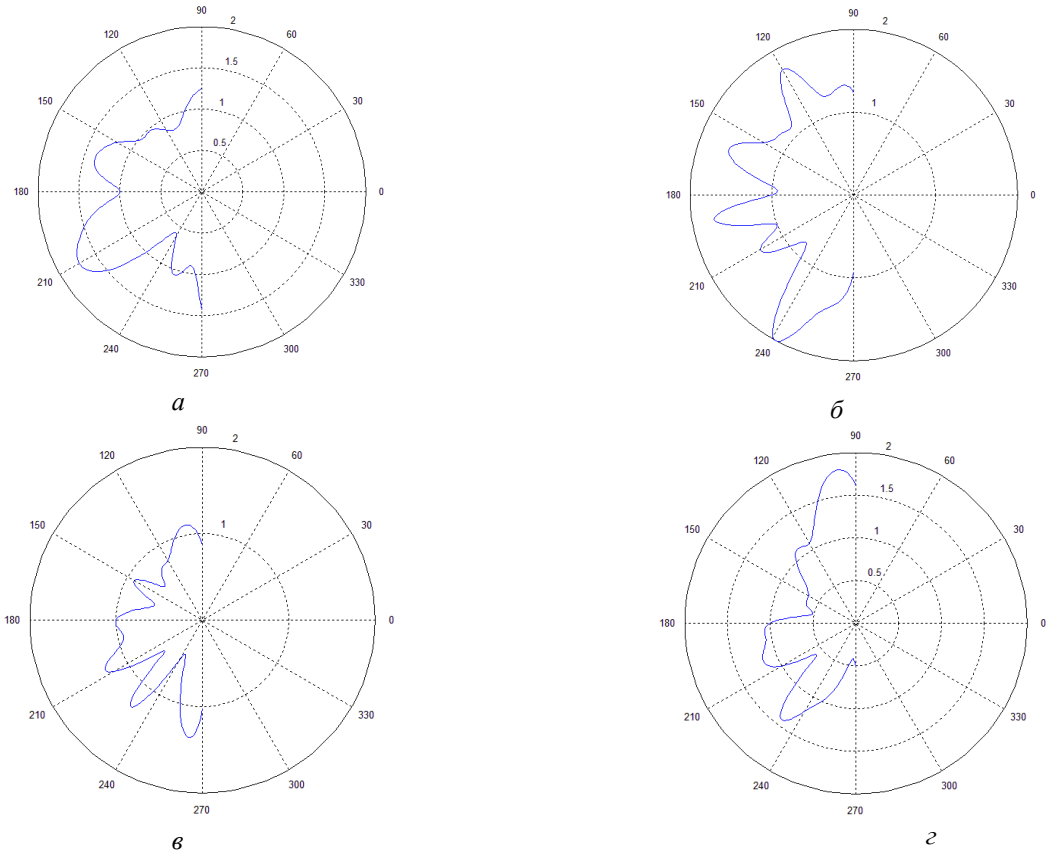


Рис. 6. Діаграми спрямованості гармонік спектру АВ БПЛА у вертикальній площині для випадку максимальної швидкості обертання гвинтів:
а – першая гармоніка; *б* – друга гармоніка; *в* – третя гармоніка; *г* – четверта гармоніка

На рис. 7 – 11 зображено діаграми спрямованості акустичного випромінювання БПЛА у вертикальній площині, отримані при зміні значень азимуту від 0° до 180°, з кроком у 15°. Експеримент проводився у заглушеній камері при невеликій швидкості обертання гвинтів БПЛА, оскільки спостерігався ефект підвищеного шуму при потраплянні набігу повітря від гвинтів на мікрофон. Тривалість запису у кожній точці простору становила 10 с. Також слід зазначити, що при обертанні БПЛА по азимуту під час вимірювання бортова навігаційна система постійно намагається компенсувати положення безпілотника у просторі. Оскільки навігаційна система БПЛА орієнтована на підтримання корпусу у заданому положенні (а також за тангажем), спостерігається змінення обертів гвинтів. Як видно з рисунків, кожному значенню азимуту відповідає своя за формою характеристика спрямованості, але слід відзначити, що деякі характеристики частково повторюють одна одну.

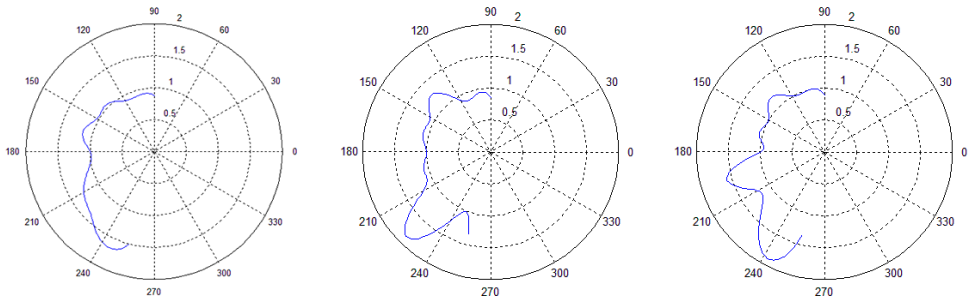


Рис. 7. Діаграми спрямованості АВ БПЛА у вертикальній площині при мінімальній швидкості обертання гвинтів для значень азимуту: *а* – 0°; *б* – 15°; *в* – 30°

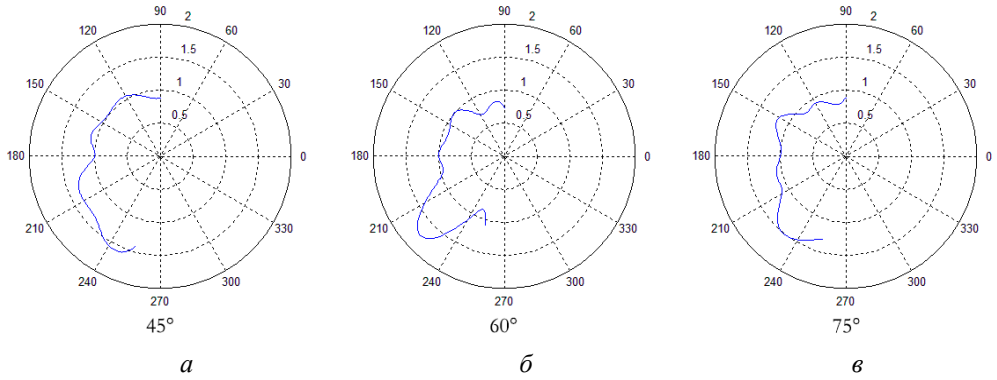


Рис. 8. Діаграми спрямованості АВ БПЛА у вертикальній площині при мінімальній швидкості обертання гвинтів для значень азимуту: *а* – 45°; *б* – 60°; *в* – 75°

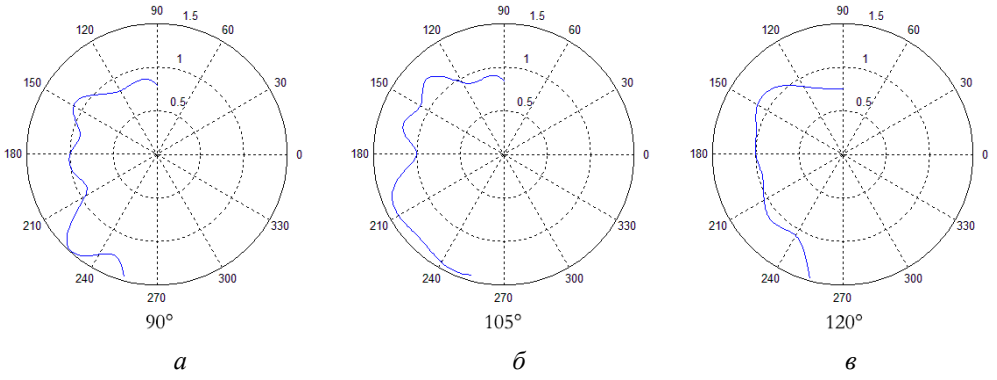


Рис. 9. Діаграми спрямованості АВ БПЛА у вертикальній площині при мінімальній швидкості обертання гвинтів для значень азимуту: *а* – 90°; *б* – 105°; *в* – 120°

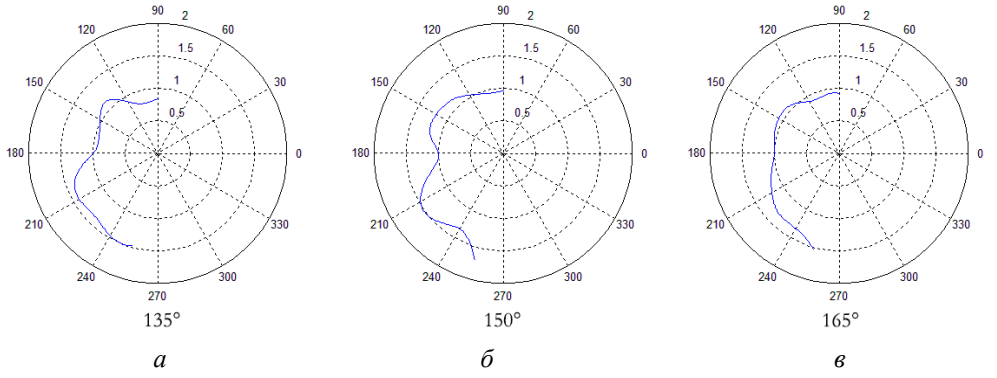


Рис. 10. Діаграми спрямованості АВ БПЛА у вертикальній площині при мінімальній швидкості обертання гвинтів для значень азимуту: *а* – 135°; *б* – 150°; *в* – 165°

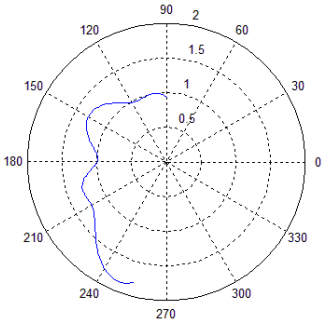


Рис. 11. Діаграма спрямованості АВ БПЛА у вертикальній площині при мінімальній швидкості обертання гвинтів для значення азимуту 180°

3. 3Д подання характеристик звукового поля БПЛА

Для побудови 3D характеристик акустичного випромінювання БПЛА використовувались дані, отримані при вимірюванні двомірних діаграм спрямованості у вертикальній площині в тринадцяти перерізах при значеннях азимуту: 0°, 15°, 30°, 45°, 60°, 75°, 90°, 105°, 120°, 135°, 150°, 165° та 180° (рис. 7 – 11).

Програмним пакетом для реалізації тривимірних візуалізацій було обрано пакет Maxon Cinema 4D. Широкий інструментарій по роботі з камерою та наявність великої кількості модифікаторів у цій програмі дозволяють створення анімацій. Є значний набір інструментів для підтримки моделювання, малювання, скульптингу, композитингу, трекінгу, анімації та рендерингу. Cinema 4D є універсальною комплексною програмою для створення та редагування дво- та тривимірних ефектів та об'єктів. Програмування здійснюється чотирма мовами: Python, C++, C.O.F.F.E.E., Xpresso. Текстурування виконується за допомогою середовищ 3ds Max та Maya.

Підготовка з подання результатів експериментальних вимірювань у тривимірному просторі проводилось комбінованим методом. На першому етапі було створено контури моделі за допомогою сплайнів (рис. 12), а на другому – побудована за цими сплайнами полігональна геометрія об’єкту моделювання .

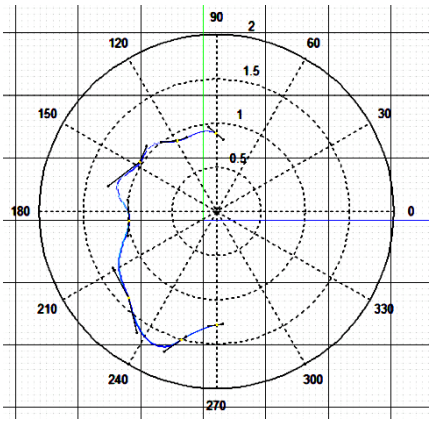


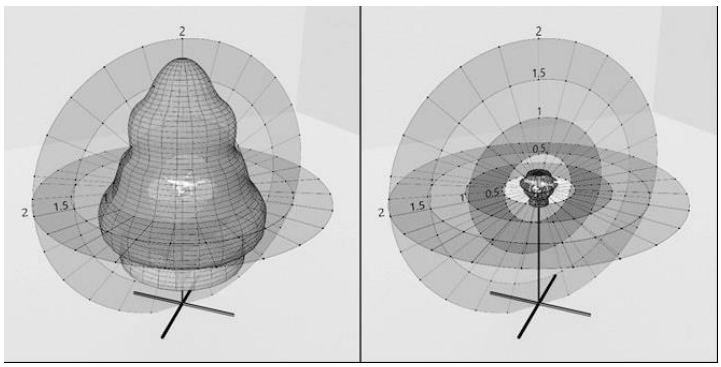
Рис. 12. Контур моделі діаграми спрямованості об’єкта, створений за допомогою сплайну, у вертикальній площині для значення азимуту 0°

На рис. 13 зображено 3D діаграми спрямованості АВ гвинтів та двигунів БПЛА, побудовані за результатами експериментальних досліджень.

Як бачимо на рис. 13, сумарне випромінювання всіх спектральних складових АВ не є ізотропним, воно має виражену просторову спрямованість. Характеристика спрямованості АВ електродвигунів БПЛА суттєво відрізняється від характеристики спрямованості випромінювання гвинтомоторної групи, оскільки має іншу фізичну природу формування сигналу. Для АВ електродвигунів спостерігається ефект екранування у нижній півсфері, обумовлений наявністю пластикового кожуха двигуна.

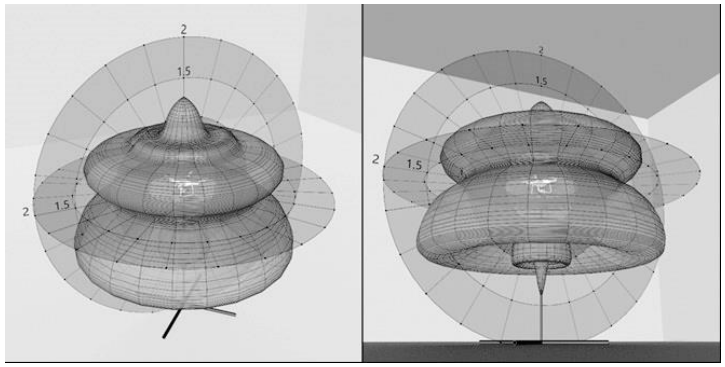
На рис. 14 зображено 3D діаграми спрямованості АВ БПЛА для чотирьох гармонічних складових з акустичного сигналу. Поверхні отримані за допомогою модифікатора, який за контурами обертання сформованих сплайнів дозволяє отримати тривимірні фігури кожної з діаграм спрямованості.

Тривимірні характеристики відображають, що інтенсивність АВ БПЛА приймає найбільші значення у напрямках витиснення повітря гвинтами, а форма просторової діаграми спрямованості відповідає сферичному закону.



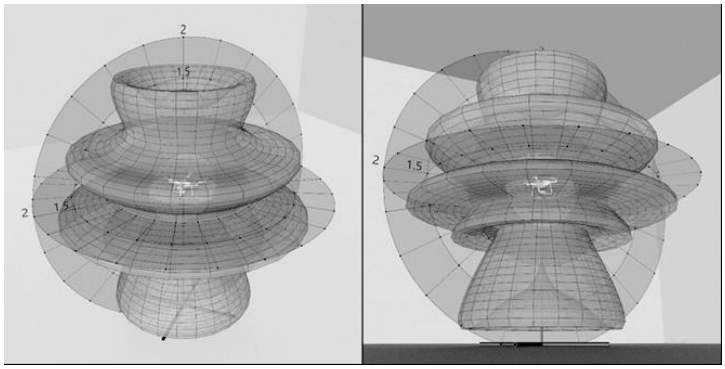
а *б*

Рис. 13. Тривимірні діаграми спрямованості АВ гвинтів (*а*) та двигунів (*б*) БПЛА



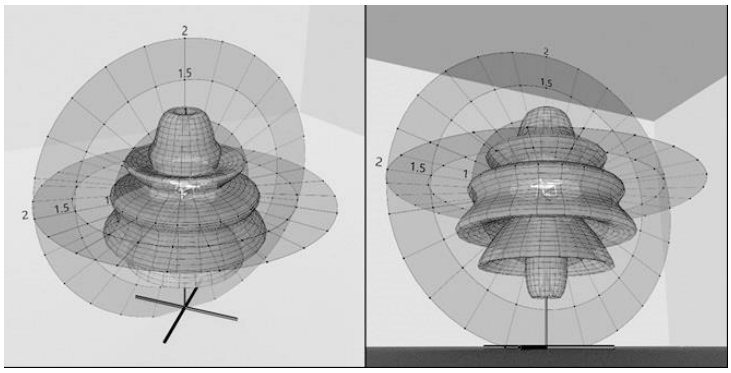
а *б*

14.1



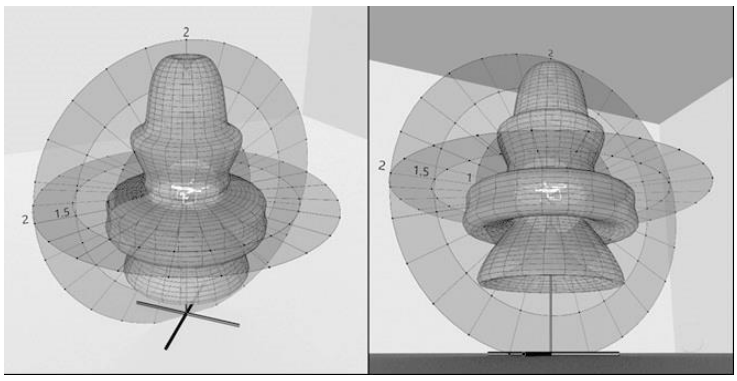
а *б*

14.2



а *б*

14.3



а *б*

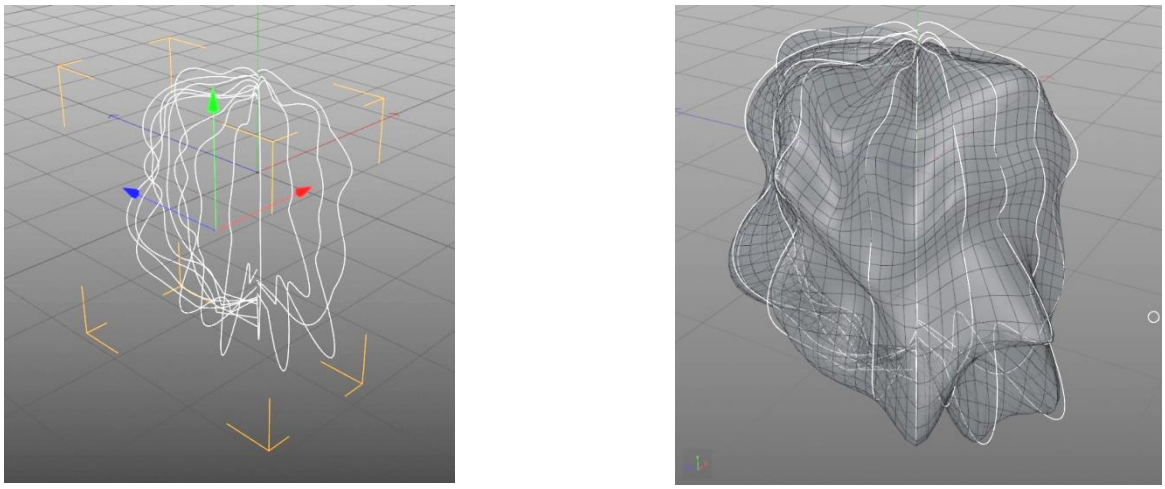
14.4

Рис. 14. 3D діаграми спрямованості АВ БПЛА для чотирьох гармонічних складових акустичного сигналу – вид зверху (*а*) та вид знизу (*б*): 14.1 – перша гармоніка; 14.2 – друга гармоніка; 14.3 – третя гармоніка; 14.4 – четверта гармоніка

Як видно з результатів експериментів, АВ поширюється за сферичним законом, але в залежності від гармоніки тривимірна фігура буде різною.

Якщо використовувати усі сплайни, які відповідають перерізам у вертикальної площині, розташованим згідно з замірами у експерименті через крок у 15° по азимуту, то можна створити подобу каркасу для тривимірного об’єкту (рис. 15, *а*).

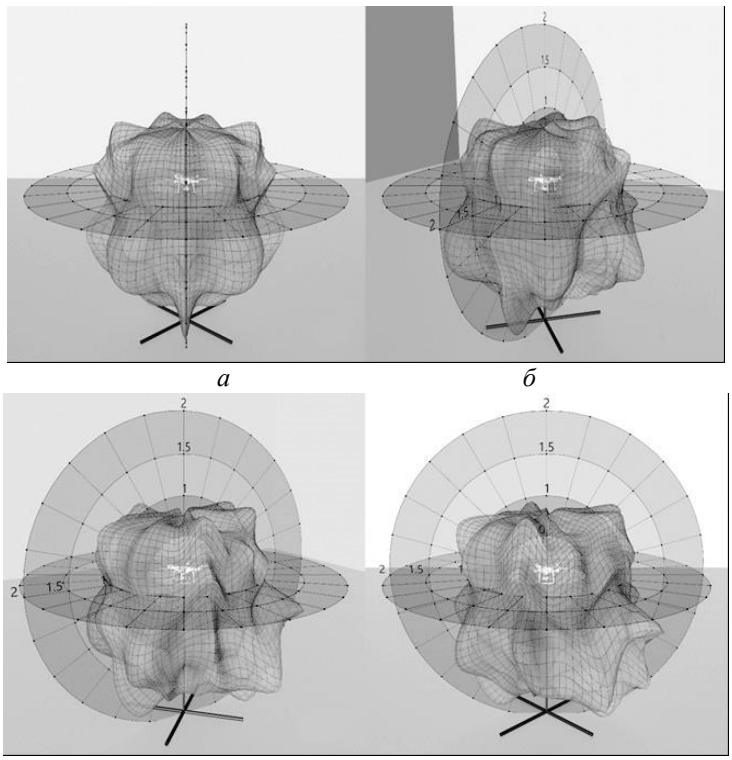
Наступним кроком є з’єднання отриманої геометричної конструкції полігонами, у результаті чого отримаємо тривимірну діаграму спрямованості у такому вигляді (рис. 15, *б*).



а *б*

Рис. 15. Побудова тривимірної діаграми спрямованості БПЛА:
а – набір сплайнів для створення 3D моделі діаграми спрямованості;
б – зображення тривимірної діаграми спрямованості у робочому середовищі програми

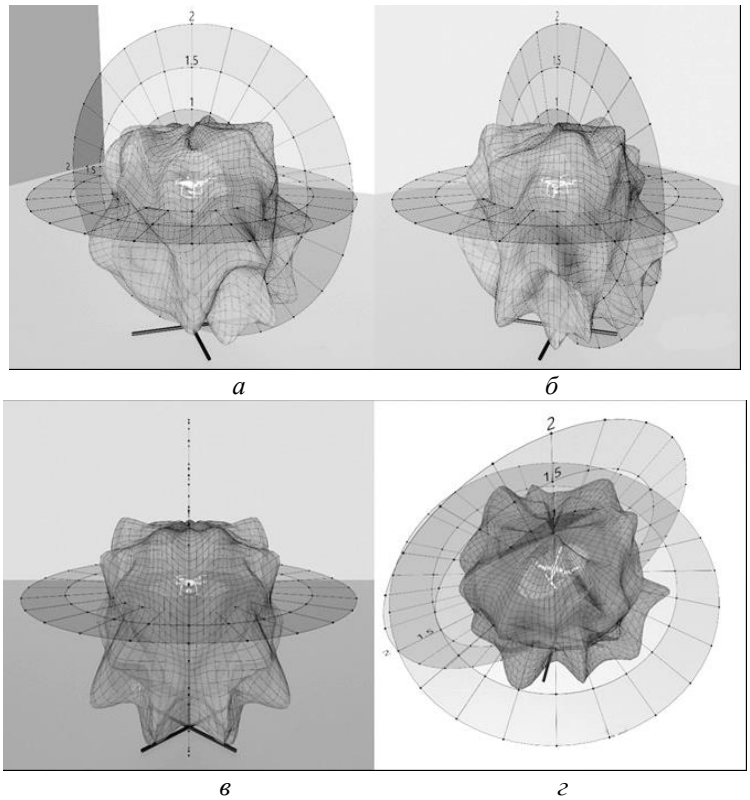
Задля отримання більш наочного зображення потрібно зробити рендер змодельованих тривимірних діаграм спрямованості АВ БПЛА з накладанням вимірювальних лімбів, після чого характеристики спрямованості будуть мати такий вигляд (рис. 16, 17).



а *б*

в *г*

Рис. 16. Тривимірна візуалізація оброблених результатів експерименту по вимірюванню діаграми спрямованості АВ при спостереженні з точок, що відповідають різним значенням азимута: *а* – 0°; *б* – 30°; *в* – 60°; *г* – 90°



а *б*

в *г*

Рис. 17. Тривимірна візуалізація оброблених результатів експерименту по вимірюванню діаграми спрямованості АВ при спостереженні з точок, що відповідають різним значенням азимута: *а* – 120°; *б* – 150°; *в* – 180°; *г* – вид зверху

Отримані фігури, що характеризують просторовий розподіл акустичної енергії випромінювання БПЛА за кутовими координатами, можна охарактеризувати як псевдовипадкові. З цього висновку випливає, що дальність виявлення і спостереження БПЛА в реальних умовах є величиною статистичної, яка залежить від ракурсу його спостереження.

Висновки

Експериментальні дослідження структури та параметрів звукового поля квадрокоптера показали, що спектри сигналу мають яскраво виражені гармонійні складові із частотами, кратними частоті обертання гвинта. У роботі показано, що найбільшу потужність мають спектральні складові у частотному діапазоні до 500 Гц, де перша гармоніка має найбільшу амплітуду, а далі має місце зменшення складових спектру до рівня шуму навколишнього середовища.

Важливе значення мають діаграми спрямованості випромінювання БПЛА, що характеризують тиск АВ у різних напрямках. У роботі отримано двовимірні та тривимірні діаграми АВ БПЛА з гвинтами та без гвинтів, коли працює тільки двигун літального апарату. Показано, що при відсутності гвинтів акустичне випромінювання значне слабкіше за рівнем. Також було представлено експериментальні дані у вигляді тривимірних діаграм АВ для чотирьох гармонік акустичного сигналу і проаналізовано, які зміни в просторової спрямованості АВ БПЛА спостерігаються за змінами тривимірної фігури зі зміною частоти гармоніки.

Показано, що просторовий розподіл як повної енергії (у всьому діапазоні частот) акустичного сигналу, так і енергії окремих його спектральних (гармонічних) складових є суттєво анізотропним. З цього висновку випливає, що дальність виявлення і спостереження БПЛА в реальних умовах є величиною статистичною, яка залежить від ракурсу його спостереження. Відповідно, акустична спостережуваність (величина, аналогічна ефективної площі розсіювання в радіолокації) повинна описуватися статистично. Дослідження законів розподілу ймовірностей акустичної спостережуваності БПЛА в залежності від кутів може стати завданням подальших досліджень у цій галузі.

Список літератури:

1. Кошкин Р.П. Беспилотные авиационные системы. Москва : Стратегические приоритеты, 2016. 676 с.
2. Ситнік О.В., Карташов В.М. Радіотехнічні системи : навч. посіб. Харків : Компанія СМІТ, 2009. 430 с.
3. Kartashov V., Oleynikov V., Koryttsev I., Zubkov O., Babkin S., Sheiko, S. Processing and Recognition of Small Unmanned Vehicles Sound Signals. International Scientific-Practical Conference on Problems of Infocommunications Science and Technology, PIC S and T 2018 – Proceedings 31 January 2019. P. 392 – 396.
4. Oleynikov V., Zubkov O., Kartashov V., Koryttsev I., Sheiko S., Babkin S. Experimental estimation of direction finding to unmanned air vehicles algorithms efficiency by their acoustic emission // 2019 International Scientific-Practical Conference «Problems of Infocommunications – Science and Technology, PIC S and T. 2019 – Proceeding». 2019. P.175 – 178.
5. Sadasivan S. Acoustis siqnature of an unmanned air vehicle – exploitation for aircraft localisation and parameter estimation / S. Sadasivan, M. Gurubasavaraj, S.R. Sekar // Eronautical DEF SCI J. 2001. Vol. 51, № 3. P. 279 – 283.
6. Massey K. Noise Measurements of Tactical UAVs / K. Massey, R. Gaeta // Georgia Inst. of Technology / GTRI / ATAS, Atlanta. 16th AIAA / CEAS Aeroacoustics Conference. American Institute of Aeronautics and Astronautics, 2010. P. 1 – 16.
7. Marino L. Experimental analysis of UAV-propellers noise // 16th AIAA/CEAS Aeroacoustics Conference. University "La Sapienza", Rome, Italy. American Institute of Aeronautics and Astronautics, 2010. P. 1 – 14.
8. Pham T. TTCP AG-6: Acousting detection and tracking of UAVs / T.Pham, N.Srour // U.S. Army Research Laboratory. Proc. of SPIE. 2004. Vol. 54. P. 24 – 29.
9. Zelnio A.M. Detection of small aircraft using an acoustic array. Thesis. B.S. // Electrical Engineering, Wright State University. 2007. 55 p.
10. G. Sinibaldi, L. Marino. Experimental analysis on the noise of the propellers for small UAV // Applied Acoustics. 74 (2013). P.79 – 88.
11. Nanyaporn Intaratep, W. Nathan Alexander, William J. Devenport, Sheryl M. Grace, Amanda Dropkin. Experimental Study of Quadcopter Acoustics and Performance at Static Thrust Conditions // Aeroacoustics Conferences 30 May – 1 June, 2016, Lyon, France 22nd AIAA/CEAS Aeroacoustics Conference. P. 1 – 6.

12. Kartashov V.M., Oleynikov V.N, Sheyko S.A., Babkin S.I., Koryttsev I.V., Zubkov O.V., Anokhin M.A. Information characteristics of sound radiation of small unmanned aerial vehicles // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika). 2018. Vol. 77 (10). P. 915 – 924.

13. Козерук С.О., Коржик О.В. Виявлення малих літальних апаратів за акустичним випромінюванням // Visnyk NTUU KPI Series Radiotekhnika Radiobuduvannia. 2019. Iss. 76. P. 15 – 20.

14. Мошков П.М., Самохин В.Ф. Оценка влияния числа лопастей и диаметра на шум воздушного винта // Вестник Самарского ун-та. Аэрокосмическая техника, технологии и машиностроение. 2016. Т. 15, No 3. С. 25 – 34.

15. Заславский Ю.М., Заславский В.Ю. Акустический шум низколетящего квадрокоптера // NOUSE Theory and Practice. 2019. Т.5, № 3. С.21 – 27.

16. Oleynikov V. M., Kartashov V. M., Babkin S. I., Zubkov O. V., Koryttsev I. V., Sheyko S. A., Seleznov I. S. Structure and parameters of propeller unmanned aerial vehicles' sound fields // Telecommunications and Radio Engineering. 2020. Vol. 79, Iss. 17. P. 1539 – 1550.

17. Карташов В. М., Олейников В. Н., Шейко С. А., Бабкин С. И., Корытцев И. В., Зубков О. В. Особенности обнаружения и распознавания малых беспилотных летательных аппаратов // Радиотехника. 2018. Вып. 195. С. 235 – 243.

Надійшла до редколегії 11.09.2022

Відомості про авторів:

Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373

Олейніков Володимир Миколайович – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vladimir.oleinikov@nure.ua; ORCID: https://orcid.org/0000-0002-3358-5987

Селєзньов Іван Сергійович – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, email: ivan.seleznov@nure.ua, ORCID: https://orcid.org/0000-0002-0731-7540

Карташов Олександр Володимирович – Харківський національний університет радіоелектроніки, здобувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, email: mservicekh1@gmail.com

УДК 621.396.967.2 **DOI: 10.30837/rt.2022.3.210.11**

І.В. СВИД, канд. техн. наук, М.Г. ТКАЧ, А.О. СЄРІКОВ, О.В. КОРОТІЧ,
С.В. ДАЦЬКО, Д.О. СУХОРУКОВ, Т.С. МАЧОНІС

ОБРОБКА ІНФОРМАЦІЇ МЕРЕЖ РАДІОЛОКАЦІЙНИХ СИСТЕМ СПОСТЕРЕЖЕННЯ ПОВІТРЯНОГО ПРОСТОРУ

Вступ

До основних джерел інформації про повітряну обстановку у системі контролю повітряного простору відносяться оглядові первинні радіолокатори [1, 2], вторинні радіолокаційні системи [3 – 7] та системи ідентифікації за ознакою «свій-чужий» (Identification Friend or Foe (IFF)) [8 – 12]. В свою чергу первинні радіолокаційні системи в залежності від територіального розташування передавача та приймача діляться на однопозиційні та багатопозиційні [13 – 18]. При цьому слід зазначити, що аналіз інформаційної безпеки однопозиційних радіолокаторів [15, 16] показує їх вразливість як широкому спектрі ненавмисних та навмисних завад, так і визначенні їх місця розташування. Це обумовлено простотою як виявлення випромінюючого передавача зондувального сигналу в однопозиційних радіолокаторах, так і оцінки його координат. Природно це зумовило основний недолік однопозиційних радіолокаторів – низька стійкість до завад та живучість. Перехід до мережі радіолокаційних систем дозволяє значно послабити вплив навмисно спрямованих завад [17 – 19] на пункти прийому багатопозиційної мережі, у зв'язку з неможливістю оцінити їх координати. Це дозволяє стверджувати, що мережі радіолокаційних систем мають більш високу стійкість до навмисних та внутрісистемних завад у порівнянні з однопозиційними радіолокаційними системами [20 – 22].

Для підвищення якості інформаційного забезпечення споживачів інформація мережі радіолокаційних систем обробляються [23]. При цьому обробка радіолокаційної інформації може здійснюватися як на сигнальному рівні [24 – 26], так і на рівнях як первинної [27], так і вторинної [28] обробки інформації. Це визначає складність здійснення міжетапної оптимізації обробки інформації як в однопозиційних радіолокаторах, так і у мережах радіолокаційних систем [29] спостереження повітряного простору.

Можна стверджувати, що у відомих роботах [26 – 29], зокрема, проведено систематичне введення в теорію, розробку та подано результати досліджень технології обробки інформації в радіолокаційних мережах систем спостереження повітряного простору. Розглянуто як класичну теорію, так і методи обробки радіолокаційної інформації на наведених вище етапах обробки інформації радіолокаційних систем. Представлена технологія обробки радіолокаційної інформації доцільна як в управлінні повітряним рухом, так і в системі протиповітряної оборони. Названі системи описуються схожими алгоритмами обробки радіолокаційної інформації, і, в цілому, мають загальні математичні основи.

Також у [29] розглянуто різні аспекти оптимальної обробки як сигнальних даних, так й інформації. При цьому показано, що розглянутий підхід оптимізації обробки як сигналів, так і радіолокаційної інформації дозволяє суттєво покращити характеристики, порівняно з існуючим підходом до обробки радіолокаційних даних. При цьому слід зазначити, що деякі алгоритми оптимізації обробки радіолокаційних сигналів дозволяють прогнозувати продуктивність вже на етапі проєктування, а також служать для об'єднання інформації для багатоцільового відстеження з використанням розподіленої архітектури відстеження.

У роботах [24, 28] показано, що при виконуваних процедурах на етапах обробки інформації систем радіолокаційного спостереження оптимізація, як виявлення, так і вимірювання координат повітряних об'єктів, можлива тільки при розподіленій обробці інформації у мережах спостереження. При цьому значення аналогового порогу для виявлення сигналу використовується і у якості параметру при спільній оптимізації обробки радіолокаційних інформації.

Метою запропонованої роботи є синтез та аналіз оптимальної структури міжетапної обробки сигнальних даних та радіолокаційної інформації первинної обробки в мережі радіолокаційних систем спостереження повітряного простору.

Загальна характеристика структури міжетапної обробки інформації в мережах радіолокаційних систем спостереження повітряного простору

Обробка інформації мережі радіолокаційних систем – це приведення інформації, отриманої на пунктах прийому ехосигналів, у вигляд, придатний для використання споживачами. Слід зазначити, що обробка інформації, зокрема інформації мережі радіолокаційних систем, та поєднання інформації мережі радіолокаційних систем спостереження виконується за допомогою сучасних інформаційних технологій. Саме інформаційні технології дозволяють реалізувати автоматичний збір, обробку, зберігання, передачу та видачу радіолокаційної інформації користувачам, підвищуючи при цьому практично всі показники якості інформації мережі радіолокаційних систем.

Складність структури системи обробки інформації мережі радіолокаційних систем не дозволяє проводити формалізацію та аналіз її роботи загалом. У зв'язку з цим всю структуру обробки інформації мережі радіолокаційних систем необхідно розділили на кілька частин, що є функціонально цілісними [14, 28]. Однак слід зазначити, що послідовна процедура обробки інформації мережі радіолокаційних систем ускладнює проведення міжетапної оптимізації обробки інформації, що призводить до зниження якості виявлення повітряного об'єкта [29, 30] та оцінки його координат [31, 32].

В мережах радіолокаційних систем обробка радіолокаційної інформації поділяється, як правило, на чотири етапи:

- обробка сигнальної інформації, яка полягає у виявленні сигналів, що приймаються, і оцінці їх параметрів;
- первинна обробка даних, що полягає у виявленні повітряного об’єкта та вимірі його координат з відповідними якістю та похибками. Цей процес складається з кількох етапів, до яких можна віднести: виявлення пакета сигналів, що приймаються; формування повного пакету сигналів, що приймаються; визначення дальності до повітряного об’єкта та оцінка його азимуту. Всі ці інформаційні задачі реалізуються з використанням оптимальних алгоритмів, що ґрунтуються на критеріях мінімуму помилок прийняття рішення та результатів вимірювання;
- вторинна обробка інформації передбачає визначення параметрів траєкторії кожного виявленого повітряного об’єкта за сигналами однієї чи кількох позицій мережі радіолокаційних систем, включаючи операції ототожнення позначок повітряного об’єкта;
- третинна обробка інформації полягає у поєднанні параметрів траєкторій повітряного об’єкта, отриманих різними приймальними пунктами мережі радіолокаційних систем з ототожненням траєкторій.

Зміст кожного із зазначених вище етапів свідчить про їхню важливість та необхідність для створення повної картини повітряної обстановки в зоні відповідальності.

При цьому слід зазначити, що поточний вектор стану повітряного об’єкта з відповідною матрицею точності вимірювання координат повітряного об’єкта складається після закінчення первинної обробки інформації.

Для виконання задач первинної обробки радіолокаційної інформації з приймальних пунктів мережі радіолокаційних систем поступають сигнальні дані, які несуть інформацію про виявлення сигналів, що приймаються, тобто $x_i = 1$, коли в елементі часового розділення відбулося перевищення порога виявлення; у випадку коли не відбулося перевищення порога виявлення, то – $x_i = 0$.

Таким чином, для першого етапу сигнальної обробки інформації частковими показниками якості обробки інформації буде імовірність правильного виявлення приймаємих сигналів D_S яка може бути визначена виходячи з співвідношення

$$D_S = f\left[q_S, F_S = f(z_S) = const\right], \quad (1)$$

де q_s – відношення сигнал/шум сигналу, що приймається, F_S – імовірність хибної тривоги виявлення сигналу, що приймається, z_S – поріг виявлення сигналу.

Оптимальність рішення задачі виявлення сигналів приймається, як правило, на основі критерію Неймана – Пірсона [33], який зводиться до максимізації імовірності правильного виявлення сигналів при обмеженні на імовірність хибного виявлення. Ці дві зазначені імовірності є показниками якості виявлення сигналів. Операції оцінки параметрів сигналів в загальному випадку оптимізуються, як правило, за критерієм мінімуму середнього ризику.

Для первинної обробки інформації мережі радіолокаційних систем показником якості обробки інформації є імовірність правильного виявлення повітряного об'єкта $P = D_1$ при фіксованій імовірності хибної тривоги, котру можливо оцінити з співвідношення

$$D_1 = f\left[N, C, F_1 = f(N, C, z_s) = const\right], \quad (2)$$

де N – пачка прийнятих сигнальних даних, C – цифровий поріг виявлення повітряного об'єкта.

Рішення про виявлення повітряного об'єкта з показниками якості F_1 та D_1 поступає на вимірювач координат повітряного об'єкта. Оцінка координат миттєвого положення повітряного об'єкта виконується одночасно з виявленням повітряного об'єкта. Основне завдання вимірювача координат повітряного об'єкта полягає у тому, щоб на основі аналізу отриманої послідовності нулів та одиниць оцінити оптимальним чином дальність та азимут повітряного об'єкта.

Наведений опис щодо виявлення сигналів та повітряного об'єкта в мережі радіолокаційних систем наочно показує, що етапна реалізація обробки інформації з одного боку спростила проведення оптимізації обробки всередині кожного етапу обробки, проте з іншого боку ускладнила проведення сумісної оптимізації як виявлення повітряного об'єкта, так і вимірювання координат повітряного об'єкта. Дійсно, стабілізація імовірності хибного виявлення повітряного об'єкта повинна здійснюватися аналоговим порогом виявлення сигналу, що складно забезпечити у прийнятій на практиці системі обробки інформації.

Синтез оптимальної структури обробки інформації в мережі радіолокаційних систем спостереження повітряного простору

Для підвищення якості інформаційного забезпечення споживачів мережа радіолокаційних систем потребує проведення обробки інформації на всіх етапах. Будемо враховувати, що мережа радіолокаційних систем складається з R приймальних пунктів, кожний з яких має M елементів дозволу щодо дальності. За такої постановки питання спільна оптимальна обробка інформації мережі радіолокаційних систем може здійснюватися у двох варіантах:

- при поєднанні рішень на рівні виявлення сигналів мережі радіолокаційних систем;
- при поєднанні рішень на рівні виявлення повітряного об'єкта.

Синтезуємо та здійснимо аналіз синтезованої структури оптимальної обробки інформації мережі радіолокаційних систем при вказаних вище варіантах поєднання інформації. При такій постановці питання у нас є сигнальні дані, які потупають від мережі радіолокаційних систем за N послідовних періодів повторення зондуючого сигналу.

Завдання виявлювача повітряного об'єкта полягає в тому, щоб на основі аналізу вхідної послідовності нулів та одиниць прийняти рішення (оптимальним чином) про наявність або відсутність повітряного об'єкта у прийнятій послідовності.

На приймальних пунктах мережі радіолокаційних систем прийняті сигнали після оптимальної лінійної обробки та детектування порівнюються у пороговому пристрої з величиною порогу, який визначає імовірність виявлення приймаємих сигналів. Після порогового пристрою на подальшу обробку поступають рішення, тобто сигнальні дані. Це дозволяє стверджувати, що з приймальних пунктів мережі радіолокаційних систем споживачу надається сукупність сигнальної інформації x_i, з показниками якості виявлення, які визначаються пороговою напругою, а також відношенням сигнал/шум приймаємої реалізації (1).

В цьому разі спостерігач має R, матрицю сигнальної інформації $\vec{X} = \left\| x_{ij} \right\|$, де $\left\| x_{ij} \right\| = 1$, коли в елементі часового розділення $i = \left(\overline{1, M}\right)$, $j = \left(\overline{1, N}\right)$, який відповідає просторовому дозволу, що розглядається, відбулося перевищення порога; коли ж не відбулося то $x_{ij} = 0$.

Таким чином, для рішення задачі виявлення необхідно отримати відношення правдоподібності та порівняти його з аналоговим порогом, обраним у відповідності до допустимої імовірності хибної тривоги виявлення повітряного об'єкта. Функції правдоподібності для гіпотез H_1 (наявності сигналу) та H_0 (відсутності сигналу) при цьому можливо записати в наступному вигляді:

$$L\left(x_i \left| H_1\right.\right) = \prod_{i=1}^{N} P_{sp}^{x_i}\left(x_i\right)\left[1 - P_{sp}\left(x_i\right)\right]^{1-x_i}, \tag{3}$$

$$L\left(x_i \left| H_0\right.\right) = \prod_{i=1}^{N} P_{p}^{x_i}\left(x_i\right)\left[1 - P_{p}\left(x_i\right)\right]^{1-x_i}, \tag{4}$$

де x_i – об'єднана послідовність нулів та одиниць з виходів приймальних пунктів мережі радіолокаційних систем.

Використовуючи вирази (3) та (4), відношення правдоподібності можна записати у вигляді

$$l\left(x_i\right) = \frac{L\left(x_i \left| H_1\right.\right)}{L\left(x_i \left| H_0\right.\right)} = \prod_{i=1}^{N}\left(\frac{P_{sp}(x_i)}{P_p(x_i)}\right)^{x_i}\left[\frac{1 - P_{sp}\left(x_i\right)}{1 - P_p\left(x_i\right)}\right]^{1-x_i} \geq l_0. \tag{5}$$

Здійснивши логарифмування виразу (5) та перетворивши даний вираз, отримуємо:

$$\sum_{i=1}^{N} x_i \eta_i \geq C, \tag{6}$$

де $\eta_i = \ln\frac{p_{sp}(x_i)[1 - p_{sp}(x_i)]}{p_p(x_i)[1 - p_p(x_i)]}$, $C = \ln l_0 - \sum_{i=1}^{N} \ln\frac{1 - p_{sp}(x_i)}{1 - p_p(x_i)}$.

Таким чином, алгоритм оптимального виявлення повітряного об'єкта (6) в мережі радіолокаційних систем зводиться до сумування вагових коефіцієнтів η_i, які визначаються формами діаграм направленості антен приймальних пунктів мережі радіолокаційних систем, що відповідають позиціям пачки, де $x_i = 1$:

Виходячи з викладеного, можна стверджувати, що для прийняття рішення про виявлення повітряного об'єкта, при сумісній обробці на рівні сигнальної інформації, обробці належить сукупність нулів та одиниць x_{ij}. У цьому випадку можна зробити висновок, що x_{ij} – це випадкова величина, яка відповідає розподілу Бернуллі:

$$P\left(x_{ij}\right) = P_{ijr}^{x_{ij}}\left(1 - P_{ijr}\right)^{1-x_{ij}},$$

де P_{ij} – імовірність перевищення порога в i-м часовому каналі обробки інформації.

При відсутності сигналу $P_{ij} = F_{rj}$ це імовірність хибної тривоги, а при дії сигналу $P_{ij} = D_{ij}$ – імовірність виявлення сигналу.

Будемо враховувати, що на вхід пристрою сумісної обробки усього масиву попередніх рішень поступає сукупність всіх можливих комбінацій x_{ij} як при відсутності, так и при наявності сигналу (гіпотези приведених вище випадкових величин x_{ij}).

Сумісні розподіли імовірності (H_0 та H_1), тобто $P(x_{ij}|H_0)$ та $P(x_{ij}|H_1)$ довільні, однак відомі. Для будь-якої конкретної сукупності x_{ij} можна сформувати наступне відношення правдоподібності:

$$\Lambda = P(x_{ij}|H_1)/P(x_{ij}|H_0). \tag{7}$$

Таким чином, порівняння Λ з порогом, який визначається допустимою імовірністю хибної тривоги, забезпечує оптимальне, за критерієм Неймана – Пірсона, рішення про наявність чи відсутність сигналу приймається x_{ij}.

У зв'язку з тим, що шуми в каналах часової обробки незалежні, то можна записати:

$$P(x_{ij}|H_0) = \prod_{j=1,i=1}^{N,M} P(x_{ij}|H_0) = \prod_{j=1,i=1}^{N,M} F_{ijr}^{x_{ij}} (1-F)^{1-x_{ij}}. \tag{8}$$

При дії сигналу щодо перевищення порогів у всіх каналах обробки інформації можливо враховувати незалежними подіями. В цьому випадку вираз (8) можемо записати в вигляді

$$P(x_{ij}|H_1) = \prod_{i=1,j=1}^{M,N} P(x_{ij}|H_1) = \prod_{i=1,j=1}^{M,N} D_{ijr}^{x_{ij}} \left(1-D_{ij}\right)^{1-x_{ij}}. \tag{9}$$

Враховуючи вирази (8) и (9), вираз (7) можна записати так:

$$\Lambda = \frac{\prod_{i=1,j=1}^{M,N} D_{ij}^{x_{ij}} \left(1-D_{ij}\right)^{1-x_{ij}}}{\prod_{i=1,j=1,}^{M,N} F_{ij}^{x_{ij}} \left(1-F_{ij}\right)^{1-x_{ij}}}. \tag{10}$$

Після логарифмування виразу (10) отримуємо:

$$L = \ln\Lambda = \sum_{i=1,j=1}^{M,N} x_{ij}\left(\ln D_{ij} - \ln F_{ij}\right) + (1-x_{ij})\left[\left(1-\ln D_{ij}\right) - \left(1-\ln F_{ij}\right)\right].$$

Позначимо множники при x_{ij}:

$$Q_{ij} = \ln D_{ij} - \ln F_{ij} - \ln(1-D_{ij}) + \ln(1-F_{ij}) = \ln\left(\frac{D_{ij}(1-F_{ij})}{(1-D_{ij})F_{rj}}\right).$$

В цьому випадку, коли відкинуті складові, які не залежать від x_{ij}, то отримуємо оптимальний, за критерієм Неймана – Пірсона, алгоритм виявлення повітряного об'єкта при поєднанні попередніх рішень всіх часових та просторових каналів обробки інформації:

$$L = \sum_{i=1,j=1}^{M,N} Q_{ij} x_{ij} \underset{>}{\overset{<}{-}} k, \tag{11}$$

де k – цифровій поріг, який визначається імовірністю хибного виявлення повітряного об'єкта.

Алгоритм (11) отриманий для випадку, коли поєднання інформації здійснюється на рівні виявлення сигналів.

Для випадку поєднання інформації на рівні виявлення повітряних об'єктів алгоритм можна записати так:

$$L_1 = \sum_{j=1,}^{M} Q_{ir} x_{ir} + \sum_{i=1}^{N} Q_i x_i \underset{>}{<} C. \qquad (12)$$

В розглядаємих варіантах обробки інформації імовірність виявлення повітряного об'єкта оптимізується за рахунок сумісної оптимізації виявлення сигналів та виявлення повітряного об'єкта. Таким чином, для реалізації запропонованого алгоритму повинна бути створена інформаційна база зберігання радіолокаційної інформації на потрібну кількість зондування передатчика мережі радіолокаційних систем, в кожному елементі якої повинні зберігатися сигнальні дані з показниками якості їх отримання.

На основі виразів (11) та (12) можливо розрахувати імовірності виявлення повітряного об'єкта для різних значень відповідних величин та відповідного вирішального правила, тобто оцінити два варіанта оптимізації обробки інформації в мережі радіолокаційних систем.

Аналіз якості обробки та поєднання даних в мережах радіолокаційних системах спостереження повітряного простору

На рис. 1 представлені залежності імовірності виявлення повітряного об'єкта $D = f(q, k/m, N = 25)$ для способу попереднього поєднання результатів виявлення сигналів, а на рис. 2 – для способу попереднього поєднання результатів виявлення повітряного об'єкта при фіксованій імовірності хибної тривоги виявлення повітряного об'єкта.

Представлені залежності отримані при пачці сигналів, що приймаються $N = 25$, цифровому порозі прийняття рішення об виявленні повітряного об'єкта $C = 12$ при загальному числі приймальних пунктів мережі радіолокаційних систем що дорівнює $m = 4$ и для різних логік прийняття рішення про виявлення повітряного об'єкта k/m. Порівняльний аналіз рис. 1 показує, що найбільш оптимальна логіка прийняття рішення для метода попереднього поєднання результатів виявлення сигналів реалізується при $k/m = 2/4$, а найгірша – при $k/m = 4/4$. Порівняльний аналіз імовірності виявлення повітряного об'єкта при $q = 1{,}52$ показує, що для логіки прийняття рішення $k/m = 4/4;\ 3/4;\ 2/4;\ 1/4$ імовірність виявлення повітряного об'єкта складає 0,4; 0,81; 0,91; 0,78 відповідно.

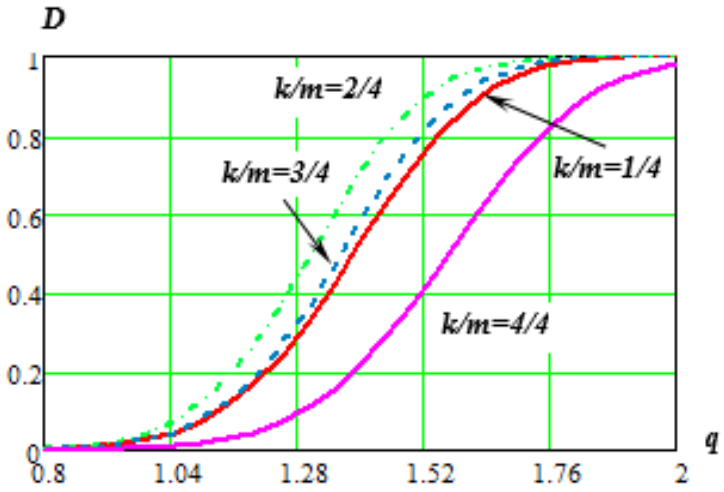


Рис. 1. Імовірність виявлення повітряного об'єкта для способу попереднього поєднання результатів виявлення сигналів

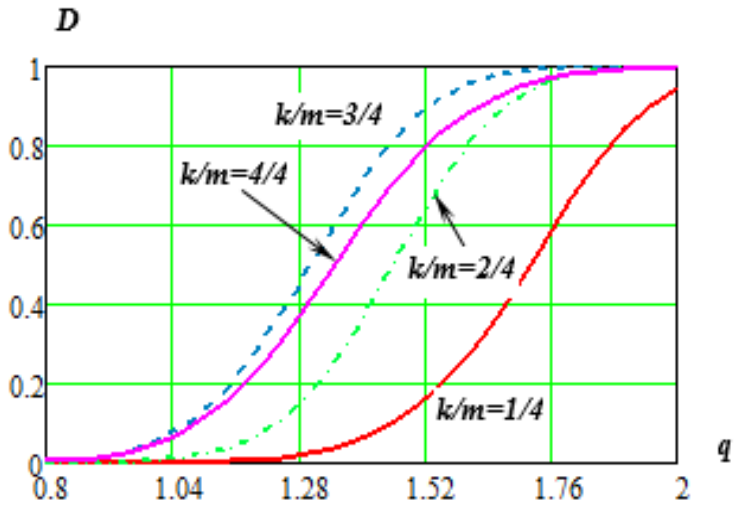


Рис. 2. Імовірність виявлення повітряного об’єкта для способу попереднього поєднання результатів виявлення повітряного об’єкта

Порівняльний аналіз рис. 2 показує, що найбільш оптимальна логіка прийняття рішення при попередньому поєднанні результатів виявлення повітряного об’єкта реалізується при $k/m=3/4$, а найгірша – при $k/m=1/4$. Порівняльний аналіз імовірності виявлення повітряного об’єкта при $q=1,52$ показує, що для логік прийняття рішення $k/m=4/4;\ 3/4;\ 2/4;\ 1/4$ імовірність виявлення повітряного об’єкта складає 0,8; 0,91; 0,62; 0,18 відповідно.

Аналіз рис. 1 та 2 дозволяє провести порівняльний аналіз двох методів поєднання інформації в мережі радіолокаційних систем.

Висновки

Синтезована структура обробки радіолокаційної інформації мережі радіолокаційних систем спостереження повітряного простору, яка дозволила здійснити міжетапну оптимізацію обробки як сигнальних даних, так й інформації первинної обробки.

Слід зазначити, що синтезована структура сумісної оптимальної обробки як сигнальних даних мережі радіолокаційних систем, так й інформації первинної обробки дозволила реалізувати два способи обробки інформації. Наведені розрахунки показали, що для спосіб обробки інформації, при якому поєднання інформації здійснюється на рівні прийняття рішень про виявлення повітряних об’єктів в кожному каналі обробки сигнальних даних, має деякі переваги в якості обробки інформації мережі радіолокаційних систем у порівнянні з варіантом поєднання інформації на етапі обробки сигналів, який використовується в теперішній час. Однак для способу поєднання інформації на рівні прийняття рішень про виявлення повітряних об’єктів потік передаваємої інформації на пункт сумісної обробки значно зменшується. Все це дозволяє підвищити якість обробки інформації в системі контролю повітряного простору.

Список літератури:

1. M. Skolnik. Improvements for air-surveillance radar // Proceedings of the 1999 IEEE Radar Conference. Radar into the Next Millennium (Cat. No.99CH36249), 1999, pp. 18-21, doi: 10.1109/NRC.1999.767195.

2. I. Svyd, I. Obod, O. Maltsev, V. Andrusevich, B. Bakumenko and O. Vorgul. Optimal Measurement of Signal Data Parameters of Requesting Radar Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2021, pp. 138-141. doi: 10.1109/UKRCON53503.2021.9575235.

3. F. L. Neindre, G. Ferre, D. Dallet, F. Letellier and K. Pitois. A Successive Interference Cancellation-based Receiver for Secondary Surveillance Radar // IEEE Transactions on Aerospace and Electronic Systems, 2022. doi: 10.1109/TAES.2022.3193649.

4. I. Obod, I. Svyd, O. Maltsev and S. Starokozhev. The Effect of Masking Interference on the Quality of Request Signal Detection in Aircraft Responders of the Identification Friend or Foe Systems // 2020 IEEE International Confer-

ence on Problems of Infocommunications. Science and Technology (PIC S&T), 2020, pp. 721-726. doi: 10.1109/PICST51311.2020.9467955.

5. M. Barbary, A. S. Hafez and T. Crew. An Industrial Design and Implementation Approach of Secondary Surveillance Radar System // 2021 International Telecommunications Conference (ITC-Egypt), 2021, pp. 1-9. doi: 10.1109/ITC-Egypt52936.2021.9513961.

6. I. Svyd, I. Obod, O. Maltsev and A. Hlushchenko. Secondary Surveillance Radar Response Channel Information Security Improvement Method // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), 2020, pp. 341-345. doi: 10.1109/DESSERT50317.2020.9125018.

7. M. Leonardi and D. D. Fausto. Secondary Surveillance Radar Transponders classification by RF fingerprinting // 2018 19th International Radar Symposium (IRS), 2018, pp. 1-10. doi: 10.23919/IRS.2018.8448244.

8. І. Свид, І. Обод. Завадостійкість радіолокаційних систем ідентифікації за ознакою «свій-чужий». Харків : Мадрид, 2021. С. 253. doi: 10/30837/978-617-7988-76-1.

9. Y. Jiang, Z. Yang, C. Bo, and D. Zhang. Continuous IFF response signal recognition technology based on capsule network // Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2021, pp. 455-468, doi: 10.1007/978-3-030-90196-7_39.

10. I. Svyd, I. Obod and O. Maltsev. Interference Immunity Assessment Identification Friend or Foe Systems // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 287-306, 2021. doi: 10.1007/978-3-030-71892-3_12.

11. T. M. Schuck, B. Shoemaker and J. Willey. Identification friend-or-foe (IFF) sensor uncertainties, ambiguities, deception and their application to the multi-source fusion process // Proceedings of the IEEE 2000 National Aerospace and Electronics Conference. NAECON 2000. Engineering Tomorrow (Cat. No.00CH37093), 2000, pp. 85-94. doi: 10.1109/NAECON.2000.894896.

12. V. Semenets, I. Svyd, I. Obod, O. Maltsev and M. Tkach. Quality Assessment of Measuring the Coordinates of Airborne Objects with a Secondary Surveillance Radar // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 105-125, 2021. doi: 10.1007/978-3-030-71892-3_5.

13. I. Ivashko, O. Krasnov and A. Yarovoy. Performance analysis of multisite radar systems // 2013 European Microwave Conference, 2013, pp. 1771-1774, doi: 10.23919/EuMC.2013.6687021.

14. Толюпа С.В., Дружинін В.А., Гордієвський О.Т. Розпізнавання групових об'єктів у багатопозиційних системах оперативного супроводження // Сучасний захист інформації. 2012. № 1. С. 66-70.

15. Обод І.І., Стрельницький О.О. Інформаційна безпека інформаційної мережі систем спостереження повітряного простору // Системи обробки інформації. 2015. № 9(134). С. 96-98.

16. Обод І.І., Стрельницький О.О. Захист інформації в мережі систем спостереження повітряного простору // Системи обробки інформації. 2016. № 2(139). С. 47-49.

17. J. Xu, X. -Z. Dai, X. -G. Xia, L. -B. Wang, J. Yu and Y. -N. Peng. Optimizations of Multisite Radar System with MIMO Radars for Target Detection // IEEE Transactions on Aerospace and Electronic Systems, vol. 47, no. 4, pp. 2329-2343, OCTOBER 2011. doi: 10.1109/TAES.2011.6034636.

18. I. Svyd, I. Obod, O. Maltsev, O. Vorgul, V. Chumak and B. Bakumenko. Estimation of the Spatial Coordinates of Air Objects in Synchronous Radar Networks for Airspace Observation // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), 2021, pp. 425-428. doi: 10.1109/PICST54195.2021.9772227.

19. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях радиолокации // Системи обробки інформації. 2006. № 9(58). С. 69-75.

20. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях // Системи обробки інформації. 2006. № 9(58). С. 69-71.

21. H. You, X. Jianjuan, G. Xin. Radar Data Processing with Applications // Publishing House of Electronics Industry. 2016. doi: 10.1002/9781118956878.

22. Chen Su, Chuanyun Zou, Liangyu Jiao, Qianglin Zhang // A MIMO Radar Signal Processing Algorithm for Identifying Chipless RFID Tags. Sensors (Basel). 2021 Dec 12;21(24):8314. doi: 10.3390/s21248314

23. Обод І.І., Стрельницький О.О., Андрусевич В.А. Методи підвищення якості інформаційного забезпечення системами спостереження повітряного простору // Системи обробки інформації. 2014. № 4(120). С. 53-55.

24. Обод І.І., Шевцова В.В. Порівняльний аналіз запитальних систем передачі інформації системи контролю повітряного простору // Зб. наук. праць Харків. нац. ун-ту Повітряних Сил. 2013. № 1(34). С. 123-125.

25. І. Обод, І. Свид, О. Мальцев. Обробка даних радіолокаційних систем спостереження повітряного простору : навч. посібник. Харків : Мадрид, 2021. 255 с.

26. J. Li, P. Stoica. MIMO Radar Signal Processing. Wiley-IEEE Press, 2008. 448 p.

27. S. M. Wu, G. A. Ybarra and W. E. Alexander. A complex optimal signal-processing algorithm for frequency-stepped CW data // IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 45, no. 6, pp. 754-757, June 1998. doi: 10.1109/82.686697.

28. Толюпа С.В., Дружинін В. А., Наконечний В.С., Цьопа Н.В., Батрак Є.О. Методи та алгоритми обробки радіолокаційної інформації у багатопозиційних системах зі змінною просторовою конфігурацією. Київ : Логос, 2014. 230 с.

29. Обод И.И. Обнаружение воздушных целей системой вторичной радиолокации // Радіоелектронні і комп'ютерні системи. 2005. № 3. С.25-28.

30. G. Lee, S. Lee, K. Kim and N. Kwak. Probabilistic Track Initiation Algorithm Using Radar Velocity Information in Heavy Clutter Environments // 2018 15th European Radar Conference (EuRAD), 2018, pp. 277-280. doi: 10.23919/EuRAD.2018.8546666.

31. Conte E., Daddio E., Farina A., and Longo M. Multistatic radar detection – Synthesis and comparison of optimum and suboptimum receivers // IEEE Proceedings F: Communications Radar and Signal Processing. 1983. vol. 130, no. 6, pp. 484-494.

32. I. Prokopenko, V. Vovk and K. Prokopenko. Fast resource management algorithm for multi-position radar systems // 2015 16th International Radar Symposium (IRS), 2015, pp. 1045-1051. doi: 10.1109/IRS.2015.7226339.

33. V. Andrusevich and I. Obod. Assessment of the Quality of Information Support by Air Radar Surveillance Systems // Advanced Information Systems, vol. 5, no. 2, pp. 78-82, 2021. doi: 10.20998/2522-9052.2021.2.10.

34. I. Obod. Integrated Coordinate-and-Time Support for the Address Inquiry in the Secondary Radar Systems // Telecommunications and Radio Engineering, vol. 53, no. 3, pp. 54-56, 1999. doi: 10.1615/telecomradeng.v53.i3.100.

35. I. Prokopenko, V. Vovk, S. Stavitsky and V. Medvedev. Optimization of use of resource in multi-position radar systems // 2014 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS), 2014, pp. 92-97. doi: 10.1109/MRRS.2014.6956673.

36. I. Obod, I. Svyd, O. Vorgul, O. Maltsev, O. Datsenko and N. Boiko. Optimization of Data Processing Structure for Multi-Position Radar Surveillance Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2021, pp. 133-137. doi: 10.1109/UKRCON53503.2021.9575286.

37. І.І. Обод, В.В. Шевцова. Пропускна спроможність відповідачів запитальних систем передачі польотної інформації // Системи обробки інформації. 2013. № 1(108). С. 105-108.

Надійшла до редколегії 30.08.2022

Відомості про авторів:

Свид Ірина Вікторівна – кандидат технічних наук, доцент, завідувач кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: iryna.svyd@nure.ua; ORCID: http://orcid.org/0000-0002-4635-6542

Ткач Марія Геннадіївна – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: mariia.zavorotna@nure.ua; ORCID: http://orcid.org/0000-0002-4248-7633

Сєріков Антон Олександрович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: anton.sierikov1@nure.ua; ORCID: https://orcid.org/0000-0002-3917-2008

Коротіч Олексій Віталійович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: oleksii.korotich@nure.ua; ORCID: https://orcid.org/0000-0002-7213-6666

Дацько Сергій Валерійович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: serhii.datsko@nure.ua; ORCID: https://orcid.org/0000-0002-2524-8702

Сухоруков Дмитро Олексійович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: dmytrosukhorukov10@gmail.com; ORCID: https://orcid.org/0000-0002-5772-286X

Мачоніс Тадас Сігітасович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: machonis7@gmail.com; ORCID: https://orcid.org/0000-0001-7656-2948

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS
ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

УДК 541.64:542.06:678 **DOI:10.30837/rt.2022.3.210.12**

В.М. БОРЩОВ, д-р техн. наук, О.М. ЛІСТРАТЕНКО, канд. техн. наук,
М.А. ПРОЦЕНКО, канд. техн. наук, І.Т. ТИМЧУК, канд. техн. наук, О.В. КРАВЧЕНКО,
О.В. СУДДЯ, М.І. СЛІПЧЕНКО, д-р фіз.-мат. наук, Б.М. ЧІЧКОВ, д-р техн. наук

ВИСОКОТЕПЛОПРОВІДНІ КОМПОЗИТНІ ПОЛІІМІДНІ МАТЕРІАЛИ

Вступ

Проблема відведення тепла, що виділяється під час роботи напівпровідникових приладів та пристроїв (наприклад, мікросхем, компонентів комп'ютерної техніки, світлодіодних джерел світла тощо), в останні роки стає домінуючою [1, 2]. Тепло, що виділяється, необхідно відводити в навколишній простір, в іншому випадку напівпровідникові прилади перегріваються, що різко знижує надійність їх роботи. Традиційно ця проблема вирішувалася за рахунок застосування матеріалів з високим коефіцієнтом теплопровідності (головним чином металів, їх сплавів, оксидів та нітридів).

Альтернативою застосування металів та їх сплавів як тепловідвідних елементів стали теплопровідні полімерні матеріали (ТПМ). Однак основною перешкодою при цьому є низька теплопровідність високомолекулярних матеріалів. Для більшості полімерних матеріалів, що випускаються промисловістю, коефіцієнт теплопровідності становить (0,1–0,3) Вт/(м·К). Тобто традиційні полімерні матеріали є теплоізолюючими матеріалами, які не здатні проводити тепло. Підвищення коефіцієнтів теплопровідності та температуропровідності полімерних матеріалів можливе за рахунок модифікації властивостей базових полімерів шляхом введення мікро-, субмікро- або нанорозмірних наповнювачів з високою теплопровідністю. У якості таких наповнювачів в теперішній час використовуються порошки металів (Ag, Al, Cu, Fe та інш.), графіт, нітриди і оксиди металів [3, 4]. Застосування ТПМ дозволяє відмовитися від використання металів під час виготовлення тепловідвідних елементів конструкцій. Заміна дорогих металів теплопровідними полімерними композиціями дозволяє значно знизити собівартість електронних приладів та пристроїв, а також суттєво зменшити їхню вагу [5]. Залежно від сфери застосування ТПМ з високою теплопровідністю повинні мати або хороші електроізоляційні властивості, або високу електропровідність. Наприклад, електроізоляційні ТПМ застосовуються як елементи теплопровідних підкладок і друкованих плат, що знаходяться під напругою, а електропровідні ТПМ – як елементи приладів та пристроїв, що підлягають заземленню в процесі експлуатації [6]. Застосування наповнювачів у вигляді різних вуглецевмісних матеріалів (графіт, технічний вуглець, вуглецеві нанотрубки) дозволяє отримати ТПМ з високою тепло- та електропровідністю [7]. З іншого боку, наповнення полімерних матриць порошками нітридів, карбідів або оксидів металів з високим коефіцієнтом теплопровідності, таких як, наприклад, AlN, SiC, ZnO і SiO_2 забезпечує хороші діелектричні властивості ТПМ разом з високою теплопровідністю [8, 9].

Розробка перспективних технологій створення полімерно-неорганічних нанокомпозитних матеріалів – одна з областей сучасного матеріалознавства, що найбільш динамічно розвиваються. Заміна традиційних полімерних матеріалів композитами на основі тих же полімерів, що наповнені різними наночастинками (НЧ), дозволяє досягти суттєвого підвищення рівня властивостей матеріалів – підвищити механічну міцність та жорсткість матеріалу, термостійкість, покращити низку інших властивостей. Теплопровідні полімерні композиційні матеріали, що містять неорганічні наповнювачі, широко застосовуються в електро- та теплотехніці, електроніці. Композиційні матеріали, як правило, отримують шляхом механічного або ультразвукового змішування полімеру та наповнювача, що пройшов попередню обробку органомодифікуючими сполуками для надання йому органофільних властивостей.

Серед полімерних композиційних матеріалів поліімідні матеріали, незважаючи на відносну високу вартість, займають одне з лідируючих місць завдяки термостійкості та високим показникам міцності, тому розробка нових композиційних матеріалів на основі поліімідних зв'язуючих завжди буде викликати відчутний науковий і промисловий інтерес. На сьогодні поліімідні матеріали мають широкий спектр можливих застосувань як матеріали, що тривало працюють у вузлах машин та приладів, а також як ізоляційні покриття. Поліімідні матеріали можуть бути використані при температурах (+250 – +500)°С (залежно від часу та навколишнього середовища), при кріогенних температурах, при впливі проникаючої радіації з поглиненими дозами до 10^4 Мрад і більше, при високих механічних навантаженнях та при поєднанні цих умов. Вони ефективно застосовуються в електротехніці та радіоелектроніці, авіаційній, ракетно-космічній та інших галузях промисловості, оскільки можуть суттєво знизити вагу та габаритні розміри виробів, підвищити їх надійність, питому потужність та робочу температуру. Аналіз науково-технічних даних та практичних робіт з розробок термостійких полімерних матеріалів, зокрема, поліімідів, показав перспективність використання термопластичних поліамідокислот (ПАК) та лаків, на основі яких одержують вільні поліімідні (ПІ) плівки з використанням мономерів різної хімічної будови.

Однак безперервно зростаючі вимоги до матеріалів призводять до необхідності розширення типового асортименту плівок, що випускаються, і створення нових видів поліімідних систем зі спеціальними властивостями, у тому числі з високими електроізоляційними і, в той же час, теплопровідними характеристиками [10].

Таким чином, метою виконаної роботи (огляду) було проведення пошуку та аналізу даних та результатів теоретичних і експериментальних досліджень, матеріалів дисертацій, літературних джерел та патентів у галузі поліімідних композиційних матеріалів. Узагальнення отриманих даних та рекомендацій щодо створення нових перспективних теплопровідних електроізоляційних поліімідних матеріалів, у тому числі в гнучких поліімідних лакофольгових шаруватих матеріалах. Особливо пошуку рекомендацій, що перевірені на практиці щодо створення композитних поліімідних плівок із суттєво збільшеною теплопровідністю від типових значень 0,12 Вт/(м•К) до 5 – 10 Вт/(м•К), що дозволяють використовувати їх також як термоінтерфейси для комутуючих плат та кабелів у різних радіоелектронних пристроях. У тому числі у приймачах детекторів електромагнітних випромінювань, у світлодіодних джерелах світла та сонячних модулях як космічного, так і наземного застосування, для забезпечення оптимальних теплових режимів.

1. Предмет та методи досліджень

1.1. Композити на основі полімерних матриць та методи їх отримання

Ненаповнені полімери, як було зазначено раніше, у їх природному стані є утеплювачами, теплопровідність яких становить 0,1 – 0,3 Вт/(м•К). Низько та середньо наповнені полімери мають теплопровідність 0,3 – 2 Вт/(м•К), що є недостатнім значенням ефективного розсіювання тепла, необхідного для багатьох технічних застосувань. Основна відмінність між низько-, середньо- та високонаповненими системами полягає в механізмі перенесення тепла. У низько- та середньонаповнених системах наповнювачі мікронного розміру не створюють безперервних шляхів для теплового потоку, так що передача тепла в таких композитах визначається, головним чином, полімерною матрицею. У високо наповнених системах частинки пов'язані одна з одною, створюючи безперервну сітку, через структуру якої здійснюється передача тепла.

Основна ідея ефективного управління теплофізичними характеристиками композитного матеріалу при високих ступенях наповнення полягає у максимізації теплопровідних шляхів поряд з мінімізацією граничного теплового опору наповнювач-наповнювач і наповнювач-матриця. Високонаповнені (> 50 об. %) полімерні композити можуть мати теплопровідність до 32 Вт/(м•К), і, отже, можуть бути ефективними, з практичної точки зору, теплопровідними матеріалами [2, 11, 12].

Схематичне зображення залежності теплопровідності полімерних композитів від вмісту наповнювача наводиться на рис. 1 [12].

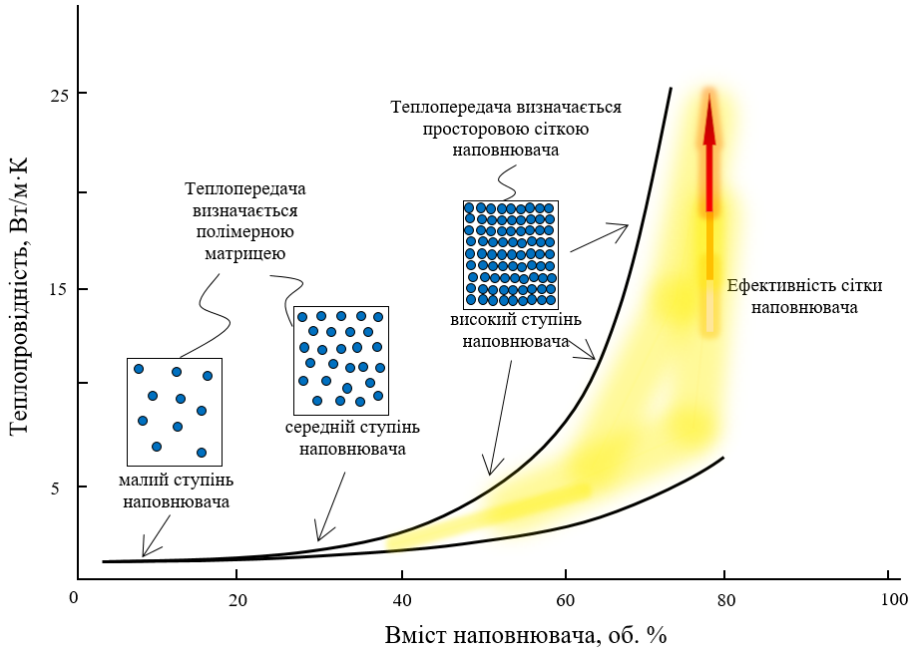


Рис. 1. Схематична залежність теплопровідності полімерних композитів від вмісту наповнювача

Якщо здатність твердих тіл проводити тепло пов'язана з їхньою структурою, складом та в основному характеризується коефіцієнтом теплопровідності, то розсіювання тепла відбувається на межі розділу фаз повітря – тверде тіло. Закони теплопередачі в режимі так званої природної конвекції такі, що є деяка конкретна межа кількості тепла, яка може бути поглинена з одиниці поверхні, що тепло віддає навколишньому повітрю. Ця кількість тепла не залежить від теплопровідності матеріалу, що віддає тепло (чи то деревина, метал, пластмаса або папір). Для охолодження в цілому це означає, що підвищення коефіцієнта теплопровідності має сенс лише до того моменту, поки кількість тепла, що транспортується через тіло, не досягне значення, яке може бути максимально прийнято (розсіяно) повітрям на останньому етапі, що лімітує. Відповідно до розрахунків, «ефективно» працююча величина коефіцієнта теплопровідності коливається в районі 5 – 10 Вт/(м•К). Подальше збільшення вже надмірно і не призводить до збільшення теплознімання загалом (рис. 2) [2].

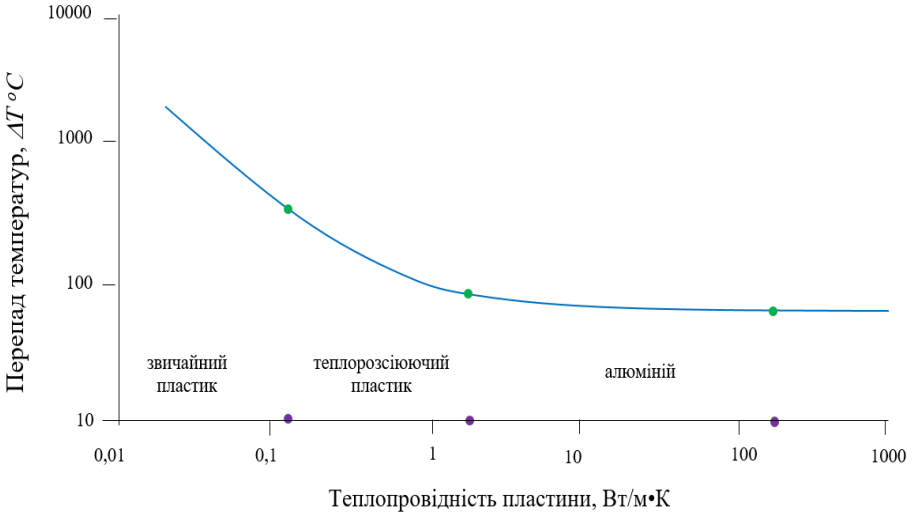


Рис. 2. Вплив теплопровідності матеріалів пластин із різною теплопровідністю на нерівномірність її температурного поля

Тим самим підтверджено, що теплопровідний потенціал міді, алюмінію та інших високотеплопровідних дорогих матеріалів використовується в режимі природного охолодження у кращому випадку лише на одну десяту своїх можливостей, а їх застосування є технічно надмірним. Саме корпуси, монтажні плати та інші численні деталі з полімерів є по суті об'єднуючим, інтегруючим середовищем для взаємодії тепловиділяючих радіоелектронних та інших функціональних елементів. Тому зрозуміло прагнення розробників використовувати більш дешеві полімери, у тому числі полііміди, для охолодження електронних пристроїв.

Глибина впливу наповнювача на властивості матеріалу зростає по мірі збільшення концентрації частинок в матриці. Однак на практиці такі концентрації, особливо наночастинок, вдається варіювати в досить вузьких межах: зазвичай концентрації наночастинок, що вводяться в полімер, складають одиниці відсотків. В силу високої поверхневої активності наночастинок вони характеризуються вираженою тенденцією до агрегації та агломерації з утворенням кластерів, стопок (наношари) та пачок-джгутів (частки циліндричної форми) окремих наночастинок. Утворення таких агрегатів призводить до гетерогенізації структури матеріалу та різкого зниження ефективності впливу нанонаповнювача на його властивості.

Відповідно, найважливішим етапом будь-яких технологій виготовлення нанокомпозитних матеріалів є процес введення наночастинок наповнювача в об'єм матричного полімеру та їх змішування. Завданням, яке необхідно вирішити при організації цього процесу, є забезпечення однорідного розподілу наночастинок, які вносяться в об’єм матриці. Від того, наскільки успішно вирішено це завдання, залежить гранична концентрація нанонаповнювача, яку вдається створити у нанокомпозитному матеріалі без втрати позитивного ефекту.

Розрізняють два основні механізми змішування: просте та диспергуюче. Під простим змішуванням розуміють процес, в результаті якого відбувається статистично випадковий розподіл частинок вихідних компонентів в об’ємі суміші без зміни їх початкових розмірів. Диспергуюча суміш – це процес змішування, який супроводжується зміною (зменшенням) початкових розмірів частинок компонентів, який пов'язаний з їх дробленням, руйнуванням агрегатів, деформуванням і розпадом дисперсної фази і т.д.

Основне завдання диспергуючого змішування зруйнувати агрегати твердих частинок і розподілити їх в об’ємі рідкого полімеру.

При створенні полімерних нанокомпозитів із вже готовим нанонаповнювачем використовуються три основні методи:

− змішання в розчині (для розчинних в органічних розчинниках полімерів);

− змішання в розплаві (для термопластичних полімерів);

− in-situ полімеризація.

Добре відомо, що дуже складно ефективно диспергувати наночастинки в розчиннику простим перемішуванням. Обробка високопотужним ультразвуком є досить ефективною для формування дисперсії НЧ. Ультразвукова обробка широко використовується при диспергуванні, емульгуванні, дробленні та активації частинок. За допомогою ультразвуку можна ефективно зруйнувати агрегати та клубки НЧ. Дослідження на різних дисперсіях агломератів наночастинок з різним вмістом твердої частини продемонстрували значні переваги ультразвуку в порівнянні з іншими технологіями, такими як роторні мішалки, поршневі гомогенізатори, кульові млини та колоїдні млини. Для змішування в розчині полімерна матриця повинна бути розчинна, щонайменше, в одному розчиннику. Це проблематично для багатьох полімерів.

Змішування в розплаві є методом, який має достатньо поширене застосування і є досить простим методом, особливо часто він використовується у випадку термопластичних полімерів. У процесі формування з розплаву НЧ механічно диспергують в полімерній матриці за допомогою змішувача з великою силою зсуву при високій температурі. Цей підхід є простим та сумісним з існуючими промисловими технологіями. Сили зсуву руйнують агрегати НЧ і запобігають їх утворенню. Недоліком цього способу є те, що цей спосіб дає дисперсію НЧ у полімерній матриці, яка значно гірша порівняно з дисперсією, яка може бути досягнута за

допомогою змішування в розчині. Крім того, НЧ повинні бути меншими через високу в'язкість композитів при більш високому вмісті НЧ.

При використанні методу In-situ полімеризації НЧ диспергуються у мономері з подальшою полімеризацією. Причому вищий відсоток наповнювачів може бути легко диспергований, і вони утворюють сильну взаємодію з полімером матриці. Цей метод застосовується для приготування композитів з полімерами, які не можуть бути оброблені змішуванням у розчині або змішування в розплаві, наприклад, нерозчинні та термічно нестійкі полімери [13].

1.2. Матеріали порошків мікро- та наночастинок для створення електроізолюючих теплопровідних полімерних композиційних матеріалів

Для підвищення теплопровідності полімерних матеріалів при збереженні необхідних їх властивостей, у тому числі високих електроізоляційних характеристик, найчастіше при створенні полімерних композитів та нанокомпозитів використовуються порошки мікро- та наночастинок наступних широко відомих та застосовуваних у промисловому виробництві діелектричних теплопровідних матеріалів, таких як BN, AlN, CN, TiO_2 та Al_2O_3 та ін. (табл. 1).

Таблиця 1

Теплопровідність наповнювачів

Тип наповнювача	Коефіцієнт теплопровідності, Вт/м•К
Нітрид бору	180
Нітрид алюмінію	285
Нітрид вуглецю	50 – 110
Оксид алюмінію	20
Оксид титану	3 – 8

Практична можливість суттєвого збільшення теплопровідності електроізоляційних полімерних композитних матеріалів, у тому числі поліімідів, при введенні їх в об’єм як наповнювачів діелектричних теплопровідних матеріалів підтверджується наведеними нижче прикладами:

- введення частинок нітриду бору (НБ). У роботі [14] досліджували теоретичними та експериментальними методами теплопровідність полімерних нанокомпозитів, що містять гексагональний НБ (ГНБ). При невпорядкованому розташуванні 6,3 об. % ГНБ підвищення теплопровідності становило 300 %. При цьому орієнтування, функціоналізація чи покриття даних частинок поліімідом (ПГНБ) сприяє ще більшому підвищенню теплопровідності. Полімерні композиційні матеріали (ПКМ), що містять 60 об. % частинок ПГНБ, мають теплопровідність 3,3 Вт/(м•К), у той час як без покриття – 2,6 Вт/(м•К);

- введення частинок нітриду алюмінію. У роботі [15] повідомляється, що теплопровідність композиційного епоксидного матеріалу, що містить 50 об. % необробленого нітриду алюмінію (AlN), становила 1,25 Вт/(м•К);

- введення частинок Al_2O_3 і НБ. У роботі [16] представлено результати досліджень впливу розмірів та геометрії частинок змішаних наповнювачів різних складів на термічні та механічні властивості гібридних композитних систем. Вивчено вплив системи, що складається з полігональних частинок Al_2O_3 та пластин НБ, на теплопровідність ПКМ. ПКМ, що містить 30 об. % гібридного наповнювача Al_2O_3+НБ (1 мкм), мав значно більшу теплопровідність (0,57 Вт/(м•К)), ніж наповнений;

- введення частинок TiO_2 та НБ. У роботі [17] вивчено теплофізичні властивості поліімід на основі діангідриду трициклодецентетракарбонової кислоти та 4,4′-діамінодифенілового ефіру та його плівкових композицій, у тому числі з частинками нітриду бору та діоксиду титану. Визначено значення теплопровідності для поліімідних плівок із наповнювачем частинок нітриду бору та частинок TiO_2 з розмірами частинок до 10 мкм. Теплопровідність

плівок ПІ + нітрид бору (40 мас. %) склала 1,24 Вт/(м•К), а теплопровідність плівок ПІ+TiO_2 склала близько 1,12 Вт/(м•К);

- введення частинок нітриду вуглецю. У роботі [18] дослідники розглядають варіант наповнювача – нітрид вуглецю. Як і графен – це речовина з двовимірною структурою, проте на відміну від нього нітрид вуглецю має істотно меншу електропровідність. Завдяки унікальній шаруватій структурі, багаточисельним функціональним групам та відповідній забороненій зоні нанолисти нітриду вуглецю (CNNS) використовуються для покращення теплопровідності поліімідної плівки до 2,04 Вт/(м•К).

1.3. Поліімідні композиційні матеріали та методи введення мікро- та наночастинок у поліімідну матрицю

На даний час на основі ароматичних поліімідів різними дослідницькими групами розроблено ряд ізолюючих теплопровідних поліімідних композиційних плівок з теплофізичними характеристиками, що в кілька разів перевершують вихідні полііміди, у тому числі розроблені ефективні технології їх створення із застосуванням наповнювачів мікро і наночастинок різних діелектриків з високою теплопровідністю.

У роботі [17] виготовляли поліімідні композитні плівки на основі діангідриду трициклодецентетракарбонової кислоти та 4,4′-діамінодифенілового ефіру, у тому числі з нітридом бору та діоксидом титану. Сумішові композиції з нітридом бору та діоксидом титану (розмір частинок до 10 мкм) отримували гомогенізацією добавок, що модифікують, в кількості 1 – 40 мас. % при кімнатній температурі протягом години в розчинах поліміду N,N′-диметилацетаміді концентрацією 20 мас. %. Потім із розчинів відливали плівки товщиною 55 – 65 мкм на скляних підкладках.

У роботі [18] показаний зручний спосіб введення наночастинок нітриду вуглецю в полімерну матрицю. Розчин наночастинок та поліамінової кислоти в диметилформаміді наносять на скляну підкладку, яку потім нагрівають до 60°С протягом чотирьох годин. При нагріванні підкладки розчинник випаровується, залишаючи на поверхні тонку плівку. Цю плівку відпалюють у печі при температурах від 60 до 250°С. При цьому відбувається процес імідізації – формування остаточного полімеру поліміду. Дослідники з'ясували, що в процесі випаровування розчину та імідізації наночастинки нітриду вуглецю взаємодіють з полімерною матрицею та самоорієнтуються у горизонтальній площині. Орієнтація наночастинок має велике значення – саме правильно збудовані частинки нітриду вуглецю дозволяють створити шлях для проходження тепла плівкою. Отримані дослідниками плівки показали відчутне зростання коефіцієнта теплопровідності в горизонтальній площині поліімідного матеріалу – 2,04 Вт/(м·К) для зразка з 20 мас. % вмістом наночастинок, проти 0,18 Вт/(м·К) у чистого поліміду. Показано, що нанокомпозитні плівки ПІ/CNNS зберігають високі електроізоляційні властивості та термічну стабільність. Ця робота розширює сферу застосування CNNS та забезпечує простий та ефективний підхід до проєктування поліімідних матеріалів з високою теплопровідністю.

У [19] для створення плівкових композитів на основі лаку АД 9103 (ТУ 6-19-283-85) використовувалися неорганічні наповнювачі TiO_2 (рутил, розмір частинок 30 – 40 нм) та Al_2O_3 (корунд, розмір частинок 40 – 80 нм). Поліімідні плівкові композити виготовляли у три етапи:

1. Приготування композицій з різними масовими співвідношеннями полімер/наповнювач змішуванням розчину поліпіромелітамідокислоти в диметилформаміді.

2. Ультразвукова обробка сумішевих композицій на генераторі УЗД-3,5 протягом 20 хвилин.

3. Нанесення сумішевих композицій на скляні підкладки та затвердіння за заданим температурним режимом.

Режими затвердіння сумішевих композицій для отримання плівкових композитів були обрані відповідно до даних для формування поліімідних плівок на основі АД 9103:

- 80 °С – 1 год, 100 °С – 1 год, 150 °С – 1 год, 200 °С – 1 год, 250 °С – 1 год, нагрівання до 320 °С;

- швидкість нагрівання до заданих температур, 5 °С/ хв;

- вміст неорганічних наповнювачів становило 3 та 10 мас. %.

У роботі [20] об'єктами дослідження були плівки на основі жорстколанцюгових поліімідів. Поліімід на основі лаку марки АД-9103 ІС, що промислово випускається (ТУ 6-19-247-84). Як наповнювач використовували алмазну шихту (АШ), яка отримана детонаційним синтезом із суміші вибухових речовин.

Для отримання композитних поліімідних плівок лак АД-9103 ІС змішували з АШ в кількості 5 і 10 мас. %. Диспергування сумішевих композицій проводили на ультразвуковому лабораторному генераторі «ЛУЗД-1.5/1П» протягом 25 хвилин. Раніше проведені експерименти показали, що мінімальний оптимальний час диспергування сумішевих композицій з АШ становить 25 хвилин, оскільки менший час обробки призводить до формування плівки з дефектами, які утворені агрегованими частинками високодисперсного наповнювача. Такі дефекти ведуть до зниження властивостей міцності плівкових зразків, а подальше збільшення часу диспергування не впливає значимо на якість одержуваних плівок. Для отримання поліімідних композитних плівок суміші, які приготовлені, аплікатором наносили на скляну підкладку, потім сушили при поступовому нагріванні до 300 °С зі швидкістю 5 °С/хв. Після цього термостатували ПІ плівки при 200 ^{o}С протягом однієї години.

У роботі [21] описані способи одержання нанокомпозитних матеріалів на основі поліімідних матриць та нанорозмірних наповнювачів. Як було зазначено раніше, всі способи отримання нанокомпозитів спрямовані в кінцевому підсумку на підвищення однорідності розподілу наночастинок в об’ємі полімеру і за рахунок цього на отримання якісних нанокомпозитів з розширеним діапазоном концентрацій наночастинок, які можуть бути введені в полімер до початку інтенсивних процесів агрегації.

Однак реально для кожного виду нанонаповнювача існує певна межа цієї концентрації, перевищення якого не призводить до подальшого підвищення рівня властивостей матеріалу. У цій роботі запропоновано рішення, спрямоване на подальше розширення діапазону концентрацій наночастинок, введення яких у полімерні матриці не призводить до розвитку агрегаційних процесів. Зазначений технічний результат досягається в запропонованому способі отримання нанокомпозитних матеріалів на основі поліімідних матриць і нанорозмірних наповнювачів – наночастинок як мінімум двох типів: наночастинок шарової геометрії, нановолокон, нанотрубок і наноконусів/дисків, за рахунок спільного або послідовного введення певних кількостей цих наночастинок в об’єм матричного полімеру. При цьому концентрація наночастинок кожного типу залишається досить низькою для того, щоб вони залишалися однорідно розподіленими в об’ємі полімеру та не утворювали агрегатів. А сумарна концентрація наночастинок є досить високою для того, щоб забезпечити значний виграш у властивостях отриманого нанокомпозиту. Для приготування однорідних дисперсій наночастинок у розчиннику перед їх введенням у розчин полімеру використовують ультразвукові методи диспергування. Суміші розчину полімеру та дисперсії наночастинок гомогенізуються за допомогою механічних пристроїв, які перемішують, лопастного типу протягом 24 годин при швидкості 1000 об/хв.

Плівки, які відлиті з нанокомпозитних розчинів на плоскі підкладки за допомогою фільтр-шаблонів з регульованим зазором, піддають сушінню протягом двох годин при 80 або 90 °С з наступною термообробкою в режимі нагрівання до 250 °С зі швидкістю 3 °С/хв або до 360 °С зі швидкістю 5 °С /хв з наступною витримкою при цій температурі протягом 30 або 20 хвилин відповідно. Для виготовлення нанокомпозитного матеріалу можуть бути використані різні види наночастинок. В якості полімерних матриць при приготуванні нанокомпозитів за методом, що описується, можуть бути використані різні поліімiди, зокрема термопластичні ароматичні поліімiди, що містять чотири і більше ароматичних циклів в елементарній ланці. В результаті реалізації перерахованих технічних рішень вдається ввести у полімер

наночастинки в сумарних концентраціях, що перевищують максимальні концентрації наночастинок одного типу, які можуть бути диспергованні в полімері до початку їх інтенсивної агрегації. За рахунок цього досягається більш суттєве зростання величин необхідних характеристик матеріалу, ніж реалізованих при введенні наночастинок одного типу в концентрації, що дорівнює сумі концентрацій різних наночастинок, що вводяться відповідно до запропонованого способу.

У роботі [22] повідомляється, що зменшення розміру частинок наповнювачів у композиційному матеріалі до нанорозмірів не призводить до істотної зміни теплопровідності порівняно з матеріалами-прототипами, що використовують мікронні та субмікронні частинки наповнювача (при близьких значеннях наповнювача). Зменшення розміру наночастинок менш ніж 20 нм призводить до значного зниження значення коефіцієнта теплопровідності. Зокрема, повідомляється, що при дослідженні суміші мікронних і нанорозмірних частинок наповнювача теплопровідності композитів збільшуються. При цьому розмір мікро- і наночастинок наповнювачів у суміші, що використовується, впливає на властивості композиту не так помітно, як їх масове співвідношення в суміші. Було встановлено, що відношення розміру мікрочастинок до розміру наночастинок не повинно перевищувати 1000, але і не повинно бути менше 100, так як за цих умов досягається оптимальний розподіл наночастинок між мікрочастинками і утворюється достатньої кількості теплопровідних шляхів. Варіювання кількості наповнювача в матеріалі дозволило встановити, що високі теплопровідні та діелектричні властивості досягаються при утриманні частинок наповнювача в композиті не менше 55 мас. % та зростають при подальшому збільшенні вмісту наповнювача. Збільшення вмісту наповнювача понад 90 мас. % призводить до погіршення деформаційно-міцнісних властивостей (погіршуються характеристики міцності та пластичності). Максимальна теплопровідність спостерігалася для варіанта композиційного матеріалу, в якому як наповнювач використовується суміш мікро- і наночастинок порошків наповнювачів з переважним вмістом мікрочастинок. Оптимальний варіант загального наповнення сумішами порошків полімерної матриці становив щонайменше 80 %. При складі сумішей порошків 80 % із середніми розмірами частинок до 10 мкм та 20 % із середніми розмірами частинок до 100 нм матеріали мали теплопровідність до 5,3 Вт/(м·К), а при складі сумішей порошків 20 % із середніми розмірами частинок до 10 мкм та 80 % із середніми розмірами частинок до 100 нм, композиційні матеріали (КМ) мали теплопровідність до 4,3 Вт/(м·К). При складі сумішей порошків 70 % із середніми розмірами частинок до 10 мкм та 30 % із середніми розмірами частинок до 20 нм КМ мали теплопровідність до 8,5 Вт/(м·К), а при складі сумішей порошків 30 % із середніми розмірами частинок до 10 мкм та 70 % із середніми розмірами частинок до 20 нм КМ мали теплопровідність до 4,1 Вт/(м·К).

Теплопровідні нанокомпозити виготовлялися шляхом змішування полімеру та порошків наповнювача із застосуванням попередньої обробки наповнювачів органомодифікуючими сполуками для надання їм органофільних властивостей. Такий підхід дозволяє максимально розширити діапазон концентрацій наночастинок, введення яких у полімерні матриці не призводить до розвитку процесів агрегації.

Висновки

У процесі проведення пошуку та аналізу науково-технічної інформації було розглянуто різні способи підвищення теплопровідності полімерних матеріалів, у тому числі, композиційних поліімідних плівок. Однак, незважаючи на те, що основна ідея ефективного управління теплофізичними характеристиками полімерних матеріалів при високих ступенях наповнення мікро- та наночастинок полягає у максимізації теплопровідних шляхів поряд з мінімізацією граничного теплового опору «наповнювач-наповнювач» та «наповнювач-матриця», вирішувати це завдання на практиці досить не просто. Пов'язано це, перш за все, з тим, що матеріалів наповнювачів, що володіють високою теплопровідністю і високими електроізолюючими властивостями, не так вже й багато. Крім того, вартість субдисперсних і нанодис-

персних теплопровідних порошків також досить висока, а способи їх введення в матрицю полімерів, а особливо поліміідну матрицю, досить складні і трудомісткі, що суттєво підвищує вартість композитних теплопровідних поліімідних матеріалів.

Вплив наповнювача на властивості матеріалу зростає зі збільшенням концентрації частинок в матриці. Однак на практиці ці концентрації вдається варіювати в досить вузьких межах через різке зниження якості композиційних плівок, що виготовляються. Особливо це стосується наночасток. Відповідно, найважливішим етапом будь-яких технологій виготовлення композитних теплопровідних матеріалів є процес введення мікро-, субмікро- або нанорозмірних наповнювачів з високою теплопровідністю наповнювача в об'єм матричного полімеру та їх змішання. Головним завданням, яке необхідно вирішити при організації цього процесу, є забезпечення однорідного розподілу частинок, що вводяться, в обсязі матриці. Від того, наскільки успішно буде вирішено це завдання, залежить гранична концентрація наповнювача, яку вдається створити в композитному матеріалі без втрати позитивного ефекту.

Проведений аналіз експериментальних даних із різних джерел підтверджує, що в даний час вже досягнуто суттєвих практичних результатів у підвищенні теплопровідних властивостей експериментальних лабораторних зразків композиційних поліімідних плівок. Коефіцієнти теплопровідності таких плівок можуть бути в межах від 1,12 до 8,5 Вт/(м•К) і більше.

Однак при цьому сучасні теплопровідні електроізолюючі поліімідні плівки, що промислово випускаються, наприклад, теплопровідна поліімідна плівка DuPont™ Kapton® MT та її китайський аналог теплопровідна плівка типу KYPI – MT компанії Suzhou (Сучжоу) Kying industrial materials Co.ltd., мають теплопровідність лише від 0,36 до 0,46 Вт/(м•К). А високоефективна теплопровідна поліімідна плівка DuPont™ Kapton® MT+ має найвищу теплопровідність серед усіх поліімідних плівок, представлених на ринку, яка не перевищує 0,75 – 0,8 Вт/(м•К). Але при цьому вони зберігають високу напругу електричного пробою, механічну стійкість та гнучкість [23 – 25].

Отже, завдання створення недорогих, але високоякісних, теплопровідних поліімідних композиційних матеріалів, що промислово випускаються, з досить високими показниками теплопровідності (5 – 10 Вт/(м•К)) і без погіршення їх характеристик міцності і пластичності в даний час є актуальною і технічно затребуваною. Такі поліімідні композити дозволять ще більше покращити електроізолюючі та теплові характеристики електродвигунів та трансформаторів, а також замінити керамічні плати та створити гнучкі теплопровідні ланцюги у різних радіоелектронних пристроях. У тому числі у приймачах детекторів електромагнітних випромінювань та у світлодіодних джерелах світла, у сонячних модулях як космічного, так і наземного застосування та багатьох інших галузях науки та техніки.

Список літератури:

1. Yoo Y. Thermal conductive carbon filled polymer composites / Y. Yoo [et al.] // Proc. 18th Intern. Conf. Comp. Mat. Korea, August 21-26. 2011. P. 3 20.

2. A. Krivatkin, Yu. Saunenko. Heat-dissipating plastics a challenge to aluminum // Semicond. light. technol. 2010. № 1. P. 54–56 [in Russian].

3. Yung K.C. Effect of AlN content on the performance of brominated epoxy resin for printed circuit board substrate / K.C. Yung [et al.] // Polym. Sci. B: Polym. Phys. 2007. V. 43, № 13. P. 1662–1674.

4. Khumalo V.M. Polyethylene/synthetic boehmite alumina nanocomposites: Structure, thermal and rheological properties / V.M. Khumalo, J. Karger-Kocsis, R. Thomann // Express Polymer Letter. 2010. V. 4, № 5. P. 264–274.

5. Han Z. Thermal conductivity of carbon nanotubes and their polymer nanocomposites : a review / Z. Han, A. Fina // Progress in Polymer Science. 2011. V. 36, № 7. P. 914–944.

6. V. Kosnyrev. Thermal Conductive Materials of the Bergquist Company // Power Electron. 2008. № 2. P. 118–122 [in Russian].

7. Piao M. Preparation and characterization of expanded graphite polymer composite films for thermoelectric applications / M. Piao [et al.] // Physica Status Solidi (b) 2013. V. 250, № 12. P. 2529–2534.

8. Lebedev S.M. Novel polymeric composites with nonlinear current-voltage characteristic / S.M. Lebedev, O.S. Gefle, A.E. Strizhkov // IEEE Transactions on Dielectrics and Electrical Insulation 2013. V. 20, № 1. P. 289–295.

9. Wang X. Large-surface-area BN nanosheets and their utilization in polymeric composites with improved thermal and dielectric properties / X. Wang [et al.] // Nanoscale Research Letters 2012. V. 7. P. 662–669.

10. A.S. Egorov, et al. Investigation of the processes of modification of polyimide systems intended for the creation of composite multilayer materials // Plastic masses. 2019. № 5-6. P. 6-8 [in Russian].
11. Yanfei Xu, Xiaoxue Wang, Jiawei Zhou, Bai Song, Zhang Jiang, Elizabeth M. Y. Lee, Samuel Huberman, Karen K. Gleason, Gang Chen. Molecular engineered conjugated polymer with high thermal conductivity // Science Advances. 2018.
12. L.K. Oliferov. Mechanochemical synthesis of functional nanostructured polymer-based composites : Dis. Ph.D. (2016). P. 154 [in Russian].
13. A.S. Egorov. Development of technology for new composite materials modified with silicon carbide and carbon nanotubes : Dis. Ph.D. (2018). P.301 [in Russian].
14. Thermal conductivity of carbon nanotube and hexagonal boron nitride polymer composites / Tabkh Paz Majid, Shajari Shaghayegh, Mahmoodi Mehdi, Park Dong-Yeob, Suresh Hamsini, Park Simon S. // Composites. B. 2016. 100. P. 19-30.
15. Synergetic effect of thermal conductive properties of epoxy composites containing functionalized multi-walled carbon nanotubes and aluminum nitride / Teng Chih-Chun, Ma Chen-Chi M., Chiou Kuo-Chan, Lee Tzong-Ming // Composites. B. 2012. 43. № 2. P. 265-271.
16. E.A. Nikolaeva, A.N. Timofeev, K.V. Mikhailovsky, Methods for increasing the thermal conductivity coefficients of polymers and polymer composite materials // Information technology bulletin. 2018; 15 (1): 156-168 [in Russian].
17. B.A. Zhubanov, et al., Thermophysical properties of alicyclic polyimide // Chem. J. of Kazakhstan. №3. 2014. p. 15-24 [in Russian].
18. Wang Y., Zhang X., Ding X., Zhang P., Shu M., Zhang Q., Gong Y., Zheng K., Tian X. Imidization-induced Carbon Nitride Nanosheets Orientation towards Highly Thermally Conductive Polyimide Film with Superior Flexibility and Electrical Insulation // Composites Part B, https:// doi.org/10.1016/j.compositesb.2020.108267.
19. K.B. Vernigorov. Influence of structural features of dispersed-filled polyimides on their resistance to high-energy oxygen plasma : Abstract. Dis. Ph.D. (2012). P. 24 [in Russian].
20. S.V. Kryuchkova et al. Influence of the diamond charge on the stability of polyimide films to the effect of atomic oxygen // Vestn. Moscow un-that. ser. 2. chem. 2017. V. 58. №5. P. 223-229 [in Russian].
21. R.U. Patent 2,636,084 (2017) [in Russian].
22. R.U. Patent 2,600,110 (2016) [in Russian].
23. Теплопровідна поліімідна плівка DuPont™ Kapton® MT, https://www.dupont.com/products/kapton-mt.html // Офіційний сайт (дата звернення 03.08.2022).
24. Теплопровідна електроізолююча поліімідна плівка типу KYPI- MT, http://ru.kying.com // Офіційний сайт (дата звернення 03.08.2022).
25. Теплопровідна поліімідна плівка DuPont™ Kapton® MT +, https://www.dupont.com/products/kapton-mt-plus.html. // Офіційний сайт (дата звернення 03.08.2022).

Надійшла до редколегії 30.05.2022

Відомості про авторів:

Борщов Вячеслав Миколайович – д-р техн. наук, професор, ТОВ «Науково-виробниче підприємство «ЛТУ», перший заступник директора – головний конструктор; Україна; e-mail: viatcheslav.borshchov@cern.ch; ORCID: https://orcid.org/0000-0002-5579-8932

Лістратенко Олександр Михайлович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», провідний науковий співробітник; Україна; e-mail: sasha.listratenko.12@gmail.com; ORCID: https://orcid.org/0000-0001-7643-5295

Проценко Максим Анатолійович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», начальник відділення – заступник головного конструктора; Україна; e-mail:max.protsenko.1978@gmail.com; ORCID: https://orcid.org/0000-0001-9313-1701

Тимчук Ігор Трохимович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», головний технолог; Україна; e-mail: ihortymchuk78@gmail.com; ORCID: https://orcid.org/0000-0002-6436-7253

Кравченко Олександр Вікторович – ТОВ «Науково-виробниче підприємство «ЛТУ», заступник начальника відділу; Україна; e-mail: kravcenkoaleksandr671@gmail.com; ORCID: https://orcid.org/0000-0002-7145-4304

Суддя Олександр Валерійович – ТОВ «Науково-виробниче підприємство «ЛТУ», науковий співробітник; Україна; e-mail: 4el1195@gmail.com; ORCID: https://orcid.org/0000-0002-2403-979X

Сліпченко Микола Іванович – д-р фіз.-мат. наук, професор, Інститут сцинтиляційних матеріалів НАНУ, провідний науковий співробітник; Україна; e-mail: naukovets.big@gmail.com; ORCID: https://orcid.org/0000-0002-4242-4800

Чічков Борис Миколайович – д-р техн. наук, професор, Інститут квантової оптики, Ганноверський університет імені Лейбниця, завідувач лабораторії нановиробництва; Германія; e-mail: chichkov@iqo.uni-hannover.de; ORCID: https://orcid.org/0000-0002-8129-7373

УДК 621.362 **DOI:10.30837/rt.2022.3.210.13**

М.А. ЯСНОГОРОДСЬКИЙ

ВИКОРИСТАННЯ РІЗНИХ МАТЕРІАЛІВ В ЯКОСТІ МЕТАЛЕВОГО КОМПОНЕНТА В МЕТАМАТЕРІАЛЬНОМУ ТЕРМОФОТОВОЛЬТАЇЧНОМУ ВИПРОМІНЮВАЧІ

Вступ

Термофотовольтаїчні (ТФВ) пристрої – це технологія, яка перетворює теплову енергію або тепло в корисну електроенергію. Термофотовольтаїчні системи є перспективними для використання сонячної енергії, відпрацьованого тепла та тепла від розпаду радіоізотопів або спалювання палива. Системи ТФВ працюють шляхом нагрівання випромінювача, який випромінює світло, яке перетворюється на електрику. Однією з ключових проблем є розробка випромінювача, який не тільки переважно випромінює світло в певних діапазонах довжин хвиль, але й одночасно задовольняє інші інженерні обмеження. Пристрої ТФВ зазвичай складаються з чотирьох компонентів: емітер, фільтр, фотодіод і зовнішнє джерело тепла. Є багато джерел тепла, з яких ТФВ може виробляти електрику, включаючи сонячне світло, горіння, радіоізотопи, та інші теплі тіла (наприклад, промислові печі). Як правило, пристрої ТФВ працюють в інфрачервоному діапазоні електромагнітного спектру із загальною температурою джерела понад 1200 – 1500°C. Однак, щоб отримати найбільш ефективне перетворення тепла в енергію, температура джерела ТФВ повинна мати пік випромінювання чорного тіла, відповідний енергії забороненої зони відповідного діода ТФВ.

Для випромінювача у термфотовольтаїчному пристрої може бути використаний будь-який матеріал, який нагрівається до високої температури. Особливо корисним випромінювачем є селективний випромінювач, який, переважно, випромінює в певній області довжин хвиль. Немає одного способу зробити випромінювач; існує багато типів випромінювачів, кожен з яких може мати різну геометричну конфігурацію та окремий набір матеріалів. Одним з перспективних нових класів випромінювачів є метаматеріали, які демонструють високі оптичні характеристики. Метаматеріал (ММ) можна налаштувати так, щоб він краще відповідав сонячному спектру, що дозволить розробити широкосмугові ширококутні метаматеріали, які могли б покращити збір світла в сонячних елементах. А метаматеріали з ширококутним відгуком можуть приймати світло під різними кутами. У випадку з сонячними елементами це означає більше збору світла та менше відбитого або «витраченого» світла[1].

Задача

У цій статті розглянуто використання ТФВ діоду з антимоніду галію (GaSb), чия енергія забороненої зони відповідає температурі чорного тіла 1600°C. А саме, метою даної роботи було визначити умови, потрібні для якісного та доцільного використання різних матеріалів задля функціонування ММ-випромінювача, а також, змоделювати та визначити доцільність застосування таких матеріалів, як платина, золото та ніхром в ММ-випромінювачі та надати рекомендації щодо подальших досліджень за цим напрямком. При температурах джерела близько 1600°C ТФВ прилад можна інтегрувати з сучасним виробництвом електроенергії об'єктів з великою користю. Відповідно до закону Кірхгофа, оптичне поглинання дорівнює оптичному випромінюванню, і використовуючи цей закон, було показано, що ідеальний поглинач ММ, продемонстрований у літературі, також працюватиме як випромінювач ММ для застосувань у термофотовольтаїчних пристроях. Результати цього дослідження спрямовані на інтеграцію ТФВ пристроїв в існуючі технології.

Матеріали та методи

Високотемпературні випромінювачі відіграють вирішальну роль у термофотовольтаїчному перетворенні енергії. Відповідно до закону Стефана – Больцмана, потужність випромі-

нювання чорного тіла пропорційна температурі. Отже, висока робоча температура є сприятливою для досягнення високої потужності випромінювання. У той же час пік спектральної густини потужності чорного тіла переміщується в бік коротших хвиль з більш високою температурою. Таким чином, для даного спектрального коефіцієнта випромінювання та для фіксованого положення забороненої зони фотоелектричної комірки ефективність перетворення також зростає з температурою. Теплове випромінювання від чорного тіла охоплює широкий діапазон довжин хвиль, де більша частина енергії випромінюється в області довгих хвиль. Оскільки енергія фотонів нижче ширини забороненої зони, фотоелектричний елемент не перетворює довгохвильові фотони в електрику.

Крім того, оскільки ці фотони зрештою все одно поглинаються, наприклад, у корпусі ця поглинена потужність фотонів низької енергії призведе до значного підвищення температури фотоелектричної комірки, та, таким чином, до зниження її зовнішньої квантової ефективності. Передні поверхневі фільтри можна використовувати для переробки довгохвильових фотонів, тобто для гальмування передачі фотонів низької енергії та повернення їх назад до випромінювача, що зменшить радіаційні втрати. Крім того, ефективність перетворення ТФВ пристрїв можна збільшити, якщо зменшити випромінювання на більших довжинах хвилі, оскільки, енергія, що випромінюється в іншому випадку, наприклад, чорним тілом тепер залишається всередині випромінювача. У цьому контексті спектрально селективні випромінювачі особливо важливі для генерації короткохвильового теплового випромінювання.

Ідеальний тепловий випромінювач ТФВ повинен забезпечувати вузькосмугове випромінювання з енергією трохи вище за ширину забороненої зони фотоелемента, оскільки енергії фотонів, що набагато перевищують ширину забороненої зони, створюють проблему термалізації фононних носіїв з можливим і непотрібним нагріванням фотоелектричного елемента. У практичних системах ТФВ досить селективного випромінювача, який забезпечує ступінчасту функцію у своїй спектральній характеристиці зі сходинкою, розташованою на енергії забороненої зони (Eg) фотоелектричної комірки, оскільки популяція квантових станів фотонів слідує розподілу Бозе – Ейнштейна, які вже стрімко наближаються до високих енергій.

Технологія випромінювача ТФВ працює через нагрівання (кондуктивне або оптичне). Коли випромінювач ТФВ нагрівається, він випромінює оптичний спектр, визначений як кривою випромінювання чорного тіла, так і будь-якими специфічними наноструктурами на поверхні випромінювача. Щоб випромінювати фотони з більшою енергією, випромінювач ТФВ необхідно нагріти до більш високих температур. Таким чином, створюється компроміс із випромінювачами ТФВ спектрально чутливих структур, які можуть витримувати високі температури для випромінювання фотонів високої енергії. З цією метою проведена значна робота зі створення дизайну ідеальних випромінювачів/поглиначів у видимому/ІЧ-випромінюванні[5].

Більшість випромінювачів ММ демонструють вузькі смуги поглинання з повною шириною напів-максимума приблизно 1μm під час роботи від ІЧ до видимого діапазону [3 – 11]. Проте симуляція передбачає можливість створення ідеального вузькосмугового випромінювача з шириною 500 нм на оптичних частотах [11, 12]. Зменшуючи повну ширину напівмаксимума спектрального випромінювача, можна підвищити ефективність діода шляхом зменшення генерації фононів у діоді ТФВ через невідповідність енергії забороненої зони/енергії фотона. Це забезпечить перевагу над випромінювачами на фотонних кристалах (PhC), які зазвичай мають ширший спектр випромінювання [13 – 26].

На відміну від випромінювачів PhC, шаблони ММ різних розмірів можна розбивати на мозаїку, щоб створити дизайнерський спектр оптичного випромінювання, як показано в [10], що відповідає безпосередньо спектру поглинання діода ТФВ. Однак існує серйозна перешкода, яку необхідно подолати при виготовленні ММ випромінювача ТФВ. Усі пристрої ММ, описані в цьому полі для програм ГГц і ТГц, виготовляються з Al, Au або Cu. Щоб використовувати їх для ТФВ пристроїв, ці пристрої повинні працювати при високих температурах. Наприклад, ширина забороненої зони GaSb, загального матеріалу ТФВ, ідеально підходить

для поглинання фотонів 1,7μm, що відповідає спектру температури чорного тіла з центром у 1600°C. Чорне тіло з центром 1600°C містить фотони, які важко перетворити в діоді ТФВ (тобто фотони з довжиною хвилі більше 1,7μm). Таким чином, має бути реалізований спектральний випромінювач, який випромінює лише фотони трохи коротші за 1,7μm. Вплив перегріву ідеального поглинача ММ можна побачити в роботі Ванга та ін., що підтверджує, що зміна форми ММ має різкий вплив на вибірковість довжини хвилі [27].

Для вирішення цієї проблеми пропонується ідеальний поглинач із ММ платини. Цей вибір матеріалу створює труднощі, оскільки рухливість вільних електронів і частота зіткнень платини нижчі, ніж традиційні металізації у ММ, які створюють нижчу електропровідність, ніж золото, вольфрам, алюміній, мідь та інші [28]. Однак платина пропонує перевагу різко підвищеної точки плавлення (1768°C), а також знижене окислення при високій температурі порівняно з іншими поширеними матеріалами ММ.

Ефективну діелектричну та проникну здатність метаматеріалу, а отже, і оптичну реакцію, можна розрахувати шляхом визначення оптичного імпедансу Z, показаного в рівнянні [29]:

$$Z = \sqrt{\frac{(1+r)^2-t^2}{(1-r)^2-t^2}}, \tag{1}$$

де r – коефіцієнт відбиття і t – коефіцієнт пропускання. Крім того, дійсний і уявний показники заломлення n необхідно визначити за допомогою рівнянь [30]:

$$Im(n) = \pm Im\left(\frac{cos^{-1}(\frac{1}{2t}[1-r^2+t^2])}{kd}\right), \tag{2}$$

та

$$Re(n) = \pm Re\left(\frac{cos^{-1}(\frac{1}{2t}[1-r^2+t^2])}{kd}\right) + \frac{2m\pi}{kd}, \tag{3}$$

де d – товщина ММ, а k – падаючий хвильовий вектор, та m – це ціле число, яке використовується для виправлення помилок розгалуження.

Ефективну проникність μ_{eff}, і діелектричну проникність ε_{eff}, можна знайти за допомогою рівнянь:

$$\mu = nZ \tag{4}$$

та

$$\varepsilon = n/Z \tag{5}$$

Загальноприйнятим методом характеристики властивостей електромагнітного розсіювання однорідного матеріалу є ідентифікація його імпедансу z і показника заломлення n. Хоча можна повністю визначити розсіювання в термінах z і n, часто зручніше вибрати інший набір аналітичних змінних, які мають пряму інтерпретацію матеріалу. Такими змінними є електрична проникність і магнітна проникність. Як n, так і z, а отже, ε і μ, є частотно-залежними складними функціями, які задовольняють певні вимоги, засновані на причинно-наслідковому зв'язку. Для пасивних матеріалів $Re(n)$ та $Im(n)$ повинні бути більше нуля [29, 30].

Як було зазначено на початку, ТФВ системи повинні працювати у вузькому діапазоні параметрів. Тепловий випромінювач має бути достатньо гарячим (>800 °C), щоб значна кількість випромінюваної потужності була вище забороненої зони фотоелектричного елемента. Однак внутрішня квантова ефективність фотоелектричних елементів значно погіршується, оскільки їхня заборонена зона знижується нижче 0,6 еВ. Навіть у найбільш сприятливих екстремумах цих обмежень закон зміщення Віна говорить нам, що пік чорного тіла при 1000°C

становить 2,27 мкм. Коли температури підвищуються, спектр чорного тіла стає краще узгодженим з фотоелементами, але лише певні матеріали з бажаними оптичними властивостями мають достатньо високі точки плавлення, щоб вижити, і навіть тоді їх оптичні властивості в гарячому стані відрізняються від оптичних властивостей при кімнатній температурі. Таким чином, високоефективна ТФВ система повинна містити високоякісний фотоелектричний матеріал з низькою забороненою зоною, гарячий, але не надто гарячий випромінювач і здатність зменшувати або рециркулювати енергію, яка випромінюється нижче ширини забороненої зони.

Отримані характеристики досліджуваних матеріалів

Для подальшого дослідження була використана програма CST Studio Suite та за основу взятий вже змодельована структура [31], а також матеріали, які відповідають вищезазначеним вимогам, а саме платина, золото та ніхром. Даний пристрій змодельований з поляризацією падаючого E-поля в X-площині і даними про проникність, отриманими з [28]. На рис. 1 зображена елементарна комірка структури досліджуваного випромінювача. Також, як видно з рис. 1, при резонансі виникає поверхневий струм, який утворює діелектричну проникність. Створюючи індуковану діелектричну та індуковану проникність, виникає новий, індукований показник заломлення. Коли новий показник заломлення збігається з показником навколишнього середовища, на поверхні ММ не виникає відбиття. Результатом є збільшення оптичного пропускання до задньої площини заземлення ММ та, в ідеалі, близько 100 % поглинання (що відповідає коефіцієнту випромінювання 100 %).

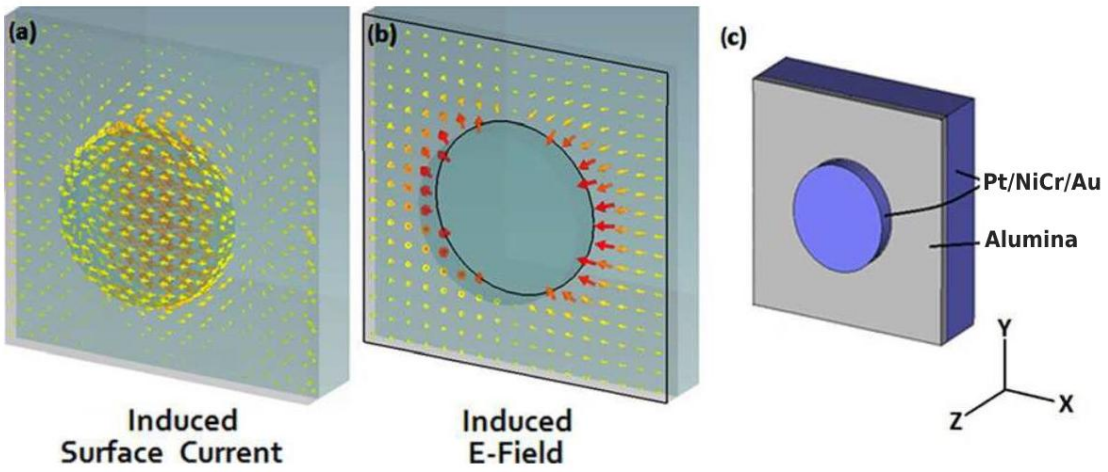


Рис.1. (a) – індукований поверхневий струм через оптичну взаємодію з візерунком ММ; (b) – індуковане електричне поле, що демонструє величину E-поля в Z-площині; (c) – структура ММ, використана в цій роботі, що складається з платинової/ніхромової/золотої задньої площини заземлення, діелектричної прокладки Al_2O_3 та візерунку ММ платини

Ця структура була виготовлена за допомогою фізичного осадження з парової фази на хімічно чисту сапфірову підкладку. Потім було використано осадження атомного шару для створення діелектричної прокладки та електрона променева літографія для створення візерунка ММ.

Для порівняння металевого компонента в даному пристрої було взято декілька матерілів – NiCr, AU та вищезгадана платина (Pt). Для початку приведемо порівняння характеристик, таких, як температура плавлення, питома електропровідність та питомий електричний опір, зазначені в табл. 1.

Таблиця 1

Характеристики досліджуваних матеріалів

Параметр	NiCr	Au	Pt
Температура плавлення	1400°C	1050°C	1750°C
Питома електропровідність	0,1	4.1	1,02
Питомий електричний опір	1,1	0,024	0,1

За допомогою формул (1) – (5) та характеристик матеріалів, які потенційно могли би використовуватись в якості металевого компонента, в даному пристрої було змодельовано залежність поглинання від довжини хвилі для пристрою з використанням кожного з запропонованих матеріалів. Результати моделювання можна побачити на рис. 2.

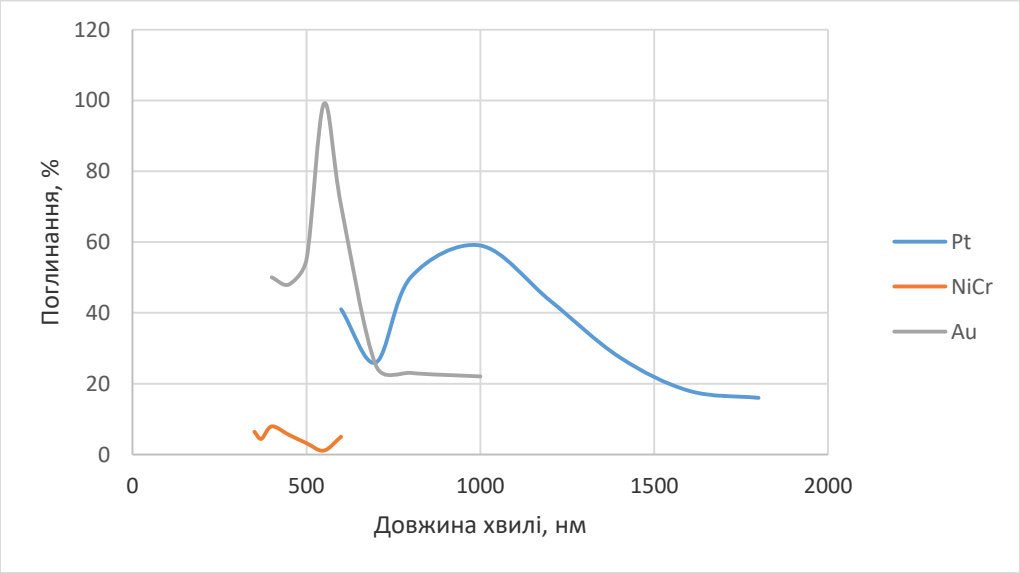


Рис. 2. Залежність поглинання від довжини хвилі для пристрою з використанням платини, ніхрому та золота

Відповідно до рис. 2 маємо наступні результати (табл. 2).

Таблиця 2

Показники поглинання та довжина хвилі в максимальній точці для досліджуваних матеріалів

Параметр	NiCr	Au	Pt
Показник поглинання в максимальній точці, %	8	99	59
Довжина хвилі, нм	400	550	1000

Для даного змодельованого пристрою було запропоновано декілька матеріалів в якості металевого компонента випромінювача, такі, як платина, ніхром та золото. Оскільки усі запропоновані матеріали мають високу температуру плавлення (більше 1400°C), тому випромінювач, в якому вони можуть використовуватись, може витримувати суворі умови, що спостерігаються в емітерних каскадах ТФВ систем. Як показано в табл. 2 та рис. 2, показники поглинання золота та платини виглядають найбільш ефективними для подальшого використання в ТФВ системах. Але, повна ширина на рівні напівмаксимуму експериментальної структури в разі використання платини має становити приблизно 500 нм, в той час, як для золота та ніхрому близько 200 – 300 нм, що мало для ММ-випромінювачів. Цей резонан-

сний пік можливо посилити за допомогою електронно-променевої літографії з вищою роздільною здатністю, яка створить набагато більш вузьку смугу піку випромінювання.

Висновки

В роботі розглядаються такі матеріали, як платина, ніхром та золото, які можуть бути використані в якості металевого компонента у метаматеріальному електронному випромінювачі, який демонструє здатність витримувати кілька циклів нагрівання до температур, що перевищують 650°C.

В результаті моделювання та порівняння характеристик та графіку поглинання зроблено висновок, що, незважаючи на те, що платина є ідеальним кандидатом для металевого компонента в даному приладі, але поглинання золота є кращим при менших довжинах хвиль. В той же час, оскільки золото має набагато нижчу температуру плавлення, ніж платина (1050°C проти 1750°C відповідно), це робить її використання в даному приладі більш проблематичним через потребу використання компонента при температурах більше ніж 800 – 900°C, що є граничними значеннями до температури плавлення компонента з золота. Стосовно ніхрому, то можна побачити, що слабкою стороною даного варіанту є значення поглинання, що дорівнює приблизно 8 % в максимальній точці. Але температура плавлення, яка дорівнює 1400°C, може дозволити використання даного сплаву в приладах, що не потребують великих значень поглинання, але мусять бути надійними та стійкими до високих температур.

Дане моделювання може бути використано для подальших досліджень та покращення технології виготовлення ММ-випромінювачів за допомогою застосування електронно-променевої літографії з високою роздільною здатністю задля звуження смуги піку випромінювання. Необхідним буде практичне дослідження доцільності використання даних матеріалів, та, можливо, декількох інших (як, наприклад, вольфрам) у випромінювачі. Також планується подальше дослідження подібних структур для визначення їх ККД за різних умов та взаємодії з різними матеріалами.

Список літератури:

1. Explaining metamaterials and metasurfaces – properties and applications [Інтернет-посилання]. nanowerk.com (2021). https://www.nanowerk.com/what-are-metamaterials.php
2. Yongzheng Wen, Ji Zhou. Metamaterial route to direct photoelectric conversion. State Key Laboratory of New Ceramics and Fine Processing, School of Materials Science and Engineering, Tsinghua University, Beijing 100084, People's Republic of China, 2019
3. C. M. Watts, X. Liu, and W. J. Padilla. Metamaterial Electromagnetic Wave Absorbers. Adv. Mater. 24, OP98–OP120 (2012).
4. N. Liu, M. Mesch, T. Weiss, M. Hentschel, and H. Giessen // Nano Lett. 10, 2342–2348 (2010).
5. M. Yan // Opt. 15, 025006 (2013).
6. J. Hao, J. Wang, X. Liu, W. J. Padilla, L. Zhou, and M. Qiu // Appl. Phys.Lett. 96, 251104 (2010).
7. W.-C. Chen, M. Koirala, X. Liu, T. Tyler, K. G. West, C. M. Bingham, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla, e-print arXiv:1212.2868v1.
8. X. Liu, T. Starr, A. F. Starr, and W. J. Padilla // Phys. Rev. Lett. 104, 207403 (2010).
9. W. J. Padilla. FiO/LS Technical Digest, OSA, 2012.
10. X. Liu, T. Tyler, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla // Phys. Rev. Lett. 107, 045901 (2011).
11. T. Maier and H. Brueckl // Opt. Lett. 35(22), 3766 (2010).
12. J. Hao, L. Zhou, and M. Qiu // Phys. Rev. B 83, 165107 (2011).
13. S. Lin, J. G. Fleming, and J. Moreno, Sand Report No. SAND2003-0845, Sandia National Laboratories, March 2003. Appl. Phys. Lett. 104, 201113 (2014)
14. J. G. Fleming, S. Y. Lin, I. El-Kady, R. Biswas, and K. M. Ho // Lett. Nat. 417, 52 (2002).
15. J. M. Gee, J. B. Moreno, S. Y. Lin, and J. G. Fleming // Conference Record of the 29th IEEE Photovoltaic Specialists Conference 2002 (IEEE, Piscataway, NJ), pp. 896–899.
16. H. Sai and H. Yugami // Appl. Phys. Lett. 85, 3399 (2004).
17. S. Y. Lin, J. G. Fleming, and J. B. Moreno // Appl. Phys. Lett. 83(2), 380 (2003).
18. G. B. Farfan, M. F. Su, M. M. Reda Taha, and I. El-Kady // Proc. SPIE 7609, 76090V (2010).
19. M. U. Pralle, N. Moelders, M. P. McNeal, I. Puscasu, A. C. Greenwald, J. T. Daly, E. A. Johnson, T. George, D. S. Choi, I. El-Kady, and R. Biswas // Appl. Phys. Lett. 81(25), 4685 (2002).

20. I. Celanovic, F. O'Sullivan, N. Jovanovic, M. Qi, and J. Kassakian // Proc. SPIE 5450, 416 (2004).
21. S.-Y. Lin, J. G. Fleming, E. Chow, J. Bur, K. K. Choi, and A. Goldberg // Phys. Rev. B 62(4), R2243 (2000).
22. H.-K. Fu, Y.-W. Jiang, M.-W. Tsai, S.-C. Lee, and Y.-F. Chen, J. Appl.Phys. 105, 033505 (2009).
23. N. Jovanovic, I. Celanovic, and J. Kassakian // AIP Conf. Proc. 890, 47 (2007).
24. I. Celanovic, N. Jovanovic, and J. Kassakian // Appl. Phys. Lett. 92, 193101 (2008).
25. D. L. C. Chan, M. Soljac˘ic´, and J. D. Joannopoulos // Opt. Express 14(19), 8785 (2006).
26. V. Rinnerbauer, S. Ndao, Y. X. Yeng, W. R. Chan, J. J. Senkevich, J. D. Joannopoulos, M. Soljacic, and I. Celanovic // Energy Environ. Sci. 5, 8815 (2012).
27. Jing Wang, Yiting Chen, Xi Chen, Jiaming Hao, Min Yan, and Min Qiu. Photothermal reshaping of gold nanoparticles in a plasmonic absorber // Opt. Express 19, 14726-14734 (2011)
28. Robert J. Bell, Mark A. Ordal, and Ralph W. Alexander. Equations linking different sets of optical properties for nonmagnetic materials // Appl. Opt. 24, 3680-3682 (1985)
29. X. Chen, T. M. Grzegorczyk, B.-I. Wu, J. Pacheco Jr., and J. A. Kong. Robust method to retrieve the constitutive effective parameters of metamaterials // Phys. Rev. E 70, 016608 (2004).
30. D. R. Smith, S. Schultz, P. Markos, and C. M. Soukoulis. Determination of effective permittivity and permeability of metamaterials from reflection and transmission coefficients // Phys. Rev. B 65, 195104 (2002).
31. Corey Shemelya, Dante DeMeo, Nicole Pfiester Latham, Xueyuan Wu, Chris Bingham, Willie Padilla, Thomas E. Vandervelde. Stable high temperature metamaterial emitters for thermophotovoltaic applications // Appl. Phys. Lett. 104, 201113 (2014).

Надійшла до редколегії 03.09.2022

Відомості про автора:
Ясногородський Максим Андрійович – НТУУ "Київський політехнічний інститут імені Ігоря Сікорського", аспірант кафедри мікроелектроніки, факультету електроніки; Україна; email: m.yasnogorodskyi-me24@lll.kpi.ua; ORCID: https://orcid.org/0000-0001-9829-9802

INFORMATION METHODS OF RADIO ENGINEERING
ІНФОРМАЦІЙНІ МЕТОДИ РАДІОТЕХНІКИ

УДК 621.391:519.246.8 **DOI:10.30837/rt.2022.3.210.14**

В.А. ТИХОНОВ, д-р фіз.-мат. наук, В.М. КАРТАШОВ, д-р техн. наук, О.В. КАРТАШОВ, В.О. ПОСОШЕНКО, канд. техн. наук

МАТЕМАТИЧНІ МОДЕЛІ НЕСТАЦІОНАРНИХ ВИПАДКОВИХ ПРОЦЕСІВ У СВВП ПОДАННІ

Вступ

Значний інтерес при вирішенні прикладних завдань, пов'язаних з аналізом випадкових процесів, викликають методи та математичні моделі, що надають можливість дослідити статистичні характеристики складених, а також нестаціонарних випадкових процесів, що описують широкий клас фізичних явищ [1 – 3].

Актуальним є завдання виділення довготривалих корельованих складових акустичного сигналу безпілотного літального апарату (БПЛА), що формують спектральні піки в низькочастотної області спектра [4 – 6]. Виділення спектру сигналу БПЛА в області низьких частот з використанням адекватних математичних моделей дозволяє ефективно виділяти і розпізнавати БПЛА на фоні завад, що формуються іншими джерелами звуку [7 – 9]. Використання моделі авторегресії дозволяє також оцінювати параметри та виділяти сигнали [10, 11] на фоні завад, розпізнавати людей за їх голосами [12]. Актуальною є, зокрема, проблема оцінки довготривалої зміни клімату Землі під дією антропогенних факторів, яка цікавить фахівців багатьох галузей науки [1].

Досить складними для аналізу є нестаціонарні випадкові процеси з трендом та сезонною складовою. Оскільки процеси в статистичній радіотехніці часто представляють у вигляді вектора, координатами якого є його відліки, цей вектор корисно представити як послідовність підвекторів меншої довжини, ніж сам вектор. Наприклад, для тимчасових рядів із середньомісячних температур із сезонною складовою, довжина підвектора становитиме 12 відліків.

Реальні та імітаційні нестаціонарні випадкові процеси, що розглядаються в роботі, з трендом і сезонною складовою представляються моделлю складеного векторного випадкового процесу (СВВП) [13]. При цьому довжина підвектора дорівнює періоду сезонної складової. Фактично у такому поданні відліки часового ряду замінюються їх сукупністю, тобто підвекторами. Аналізуються статистичні зв'язки для підвекторів, а не як завжди, для відліків процесу. Якщо довжина підвектора дорівнює одиниці, всі операції у поданні СВВП еквівалентні звичайним операціям для часових рядів [13, 14].

Метою дослідження є удосконалення методу та моделі для оцінювання статистичних характеристик складених та нестаціонарних випадкових процесів, у тому числі різних складових часових рядів параметрів атмосфери. Отримані результати можуть бути використані для аналізу середньострокових і довгострокових змін атмосферних умов, уточнення результатів, отриманих традиційними методами математичної статистики, а також в інших галузях людської діяльності.

Основні властивості випадкового процесу в СВВП поданні

Під корельованим СВВП будемо розуміти процес $\vec{x}^n[t]$, у якому існують статистичні зв'язки між підвекторами $\vec{x}_i$. Уявимо процес $\vec{x}^n[t]$ у вигляді послідовності $m = N/n$ підвекторів

$$\vec{x}^n[t] = \{\vec{x}_1, \vec{x}_2, ..., \vec{x}_{N/n}\}.$$

Кожен підвектор визначається n координатами вектора $\vec{x}^n[t]$:

$$\vec{x}_1 = \{x[1], x[2],..., x[n]\},\ \vec{x}_2 = \{x[n+1], x[n+2],..., x[2n]\},\ldots$$

$$\ldots, \vec{x}_i = \{x[(i-1)n+1],..., x[in]\},\ldots, \vec{x}_{N/n} = \{x[N-n+1],..., x[N]\},$$

де i – номер підвектора, N – номер останнього відліку вектора. Якщо кількість відліків вектору некратна довжині підвектора n, то береться ціла частина цього числа, тобто $N/n \sim \lfloor N/n \rfloor$. При такому поданні можна повніше досліджувати закономірності зміни векторів.

СВВП подання зручно записати в матричному вигляді, в якому рядками матриці є координати підвекторів завдовжки n

$$\vec{x}^n[t] = \begin{bmatrix} \vec{x}_1^n \\ \vec{x}_2^n \\ \vdots \\ \vec{x}_m^n \end{bmatrix} = \begin{bmatrix} x[1] & x[2] & ... & x[n] \\ x[n+1] & x[n+2] & ... & x[2n] \\ \vdots & & & \\ x[N-n+1] & x[N-n+2] & ... & x[N] \end{bmatrix}.$$

Таким чином, процес представляється послідовністю підвекторів меншої довжини n.

Середні значення СВВП $\vec{x}^n[t]$, що складаються з підвекторів довжиною n, визначаються виразом

$$\bar{\vec{x}}^n = \frac{1}{m}\sum_{i=1}^{m} \bar{\vec{x}}_i, \tag{1а}$$

де середні значення підвекторів дорівнюють

$$\bar{\vec{x}}_i = \frac{1}{n}\sum_{v=1}^{n} x[in+v]. \tag{1б}$$

Формули (1) дають можливість визначити середні вектори та підвектори. Тоді середнє значення вектора часового ряду має координати, що дорівнюють середнім значенням відповідних підвекторів процесу. Наприклад, середнє від часового ряду, що складається із середніх значень підвекторів, можна інтерпретувати як послідовність

$$\overline{\vec{x}^n[t]} = [\overline{\vec{x}_1}, \overline{\vec{x}_2},..., \overline{\vec{x}_m}]^T, \tag{2}$$

де середнє $\bar{\vec{x}}_i$ визначається (1б).

Очевидно, центрування СВВП полягає у відніманні з кожного підвектора середнього часового ряду

$$\overline{\vec{x}^n[t]} = \overline{[\bar{\vec{x}}_1, \bar{\vec{x}}_2,..., \bar{\vec{x}}_m]^T},$$

тобто

$$\vec{x}_c^n[t] = [\vec{x}_1 - \overline{\vec{x}^n[t]}, \vec{x}_2 - \overline{\vec{x}^n[t]},...\vec{x}_m - \overline{\vec{x}^n[t]}]^T.$$

Далі вважатимемо, що аналізована вибірка центрована.

Слабкий тренд температур на фоні сильних сезонних коливань можна виявити, усереднюючи процес у СВВП поданні. Знайдемо тренд процесу, представленого на рис. 1, шляхом усереднення. Для цього вихідний часовий ряд середньомісячних температур представимо у вигляді ряду середньорічних даних

$$\overline{x^n[t]} = \overline{x_1[t]}, \overline{x_2[t]},..., \overline{x_m[t]}, \tag{3а}$$

де

$$\overline{x_i[t]} = \frac{1}{n}\sum_{k=1}^{n} x_k[t],\ \ n = 12. \tag{3б}$$

У (3а) та (3б) час визначено як рік вимірювань температур.

Для отримання трендів додатково згладимо середньорічні значення даних (3а) низькочастотним ковзним фільтром. Щоб накласти на тренд вихідні дані, тобто щоб довжина вихідних даних та даних, отриманих усередненням, збігалася, продовжимо кожне значення тренду на 12 відліків, рівних відповідному відліку тренду. Тоді отримаємо часові ряди трендів, рівні за тривалістю початковій довжині даних. Отриманий таким способом тренд та вихідні дані представлені на рис. 1. Зауважимо, що процедура використовуваного ковзного усереднення, коли при усередненні враховуються $l/2$ попередніх значень часового ряду та $l/2$ наступних, не дає змоги отримати на початку часового ряду та наприкінці усереднення з тією ж точністю, як і всередині часового ряду.

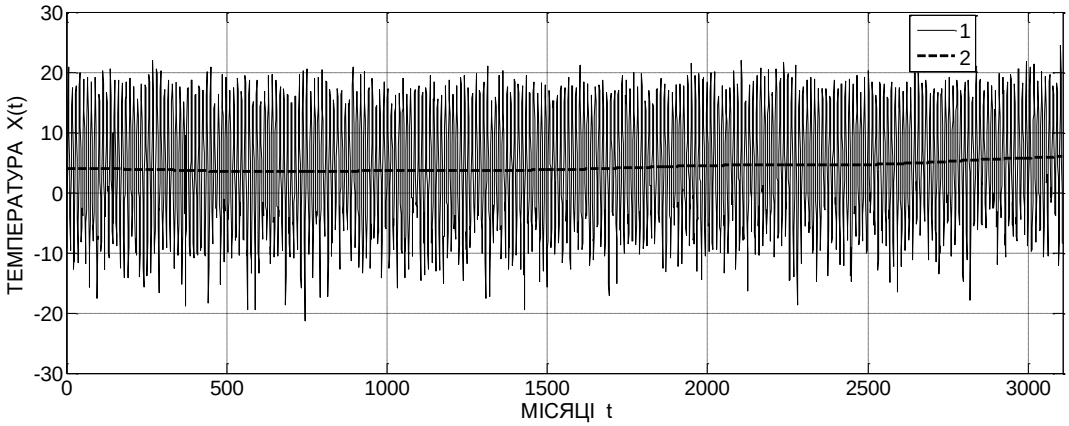


Рис. 1. Дані середньомісячних температур та оцінка їх довготривалого тренду

Нестаціонарний процес із трендом, показаний на рис. 1, являє собою дані про середньомісячні температури. Повна довжина ряду становить 3108 значень, отриманих обсерваторією за 259 років – з 1752 по 2010 рік. Дані містять слабкий тренд, яскраво виражені сезонні коливання температур з періодом 12 місяців та стаціонарний процес.

Аналіз тренду на рис. 1 показує, що з 1752 по 1803 рік температура падала з 4,05 до 3,45°С. Спад температури склав 0,6°С. Після 1804 року спостерігалося зростання температури з 3,45 до 5,96°С у 2010 році. Приріст температури за цей термін становив 2,51°С.

Модель лінійного передбачення нестаціонарних процесів із сезонною складовою та трендом

Зупинимося на принципах побудови моделі авторегресії проінтегрованого ковзного середнього (АРПКС) [14]. Нехай нестаціонарний процес $x[t]$ містить тренд, сезонну складову та стаціонарний процес, який описуватимемо у загальному випадку моделлю авторегресії – ковзного середнього (АРКС). У найпростішому випадку тренд і сезонну складову можна врахувати за допомогою операторів зсуву. Мультиплікативний процес $\omega[t]$ без сезонної складової та тренду можна записати у вигляді [15]

$$\omega[t] = \nabla^d \omega_1[t] = \nabla^d \nabla_s x[t] = \left(1 - z^{-1}\right)^d \left(1 - z^{-s}\right) x[t]. \quad (4)$$

Для виключення сезонної складової з процесу $x[t]$ застосовується оператор взяття різниці $\nabla_s = 1 - z^{-s}$, де z^{-s} – оператор зсуву, дія якого визначається виразом $z^{-s} x[t] = x[t-s]$, а s – дорівнює періоду сезонної складової. Тоді процес без сезонної складової, але з трендом і стаціонарною складовою, записується у вигляді

$$\omega_1[t] = \nabla_s x[t] = \left(1 - z^{-s}\right) x[t].$$

Для виключення сезонної складової з процесу $x[t]$ застосовується оператор взяття різниці $\nabla_s = 1 - z^{-s}$, де z^{-s} – оператор зсуву, дія якого визначається виразом $z^{-s} x[t] = x[t-s]$, а

s – дорівнює періоду сезонної складової. Тоді процес без сезонної складової, але з трендом і стаціонарною складовою, записується у вигляді

$$\omega_1[t] = \nabla_s x[t] = (1 - z^{-s})x[t].$$

Щоб виключити також тренд з часового ряду $\omega_1[t]$, необхідно вплинути на нього оператором $\nabla^d = (1 - z^{-1})^d$, тобто

$$\omega[t] = (1 - z^{-1})^d \omega_1[t]. \quad (5)$$

Для лінійного тренду вважають $d = 1$, а для квадратичного тренду необхідно брати $d = 2$ і т. д. Для лінійного тренду, з (5) отримуємо першу дискретну похідну процесу $\omega_1[t] = \nabla x[t] = x[t] - x[t-1]$. Для квадратичного тренду використовується друга дискретна похідна процесу $\omega_1[t] = \nabla^2 x[t] = \nabla(x[t] - x[t-1]) = (x[t] - 2x[t-1] + x[t-2])$.

Видалення сезонної складової та тренду операторами зсуву дозволяє врахувати ці складові і потім відновити.

У більш загальної мультиплікативної моделі АРПКС нестаціонарного часового ряду з трендом і сезонної складової, наприклад, що складається з даних усередині місяця, враховуються кореляції спостережень у послідовні місяці цього року. Модель також визначає кореляції спостережень місяця у послідовні роки. Модель АРПКС такого процесу описується загальною мультиплікативною моделлю виду [16]

$$\Phi(z^{-1})\varphi_P(z^{-s})\nabla^d \nabla_s^D x[t] = Q_q(z^{-1})\theta_\vartheta(z^{-s})a[t], \quad (6)$$

де оператори зсуву авторегресії (АР) і ковзного середнього (КС) дорівнюють:

$$\nabla_s^D = (1 - z^{-s})^D,\ \Phi(z^{-1}) = 1 - \sum_{i=1}^{p} \Phi[i] z^{-i},\ Q(z^{-1}) = 1 - \sum_{i=1}^{q} Q[i] z^{-i},\quad \varphi_P(z^{-s}) = 1 - \sum_{i=1}^{P} \varphi[i] z^{-is},$$

$$\theta_\vartheta(z^{-s}) = 1 - \sum_{i=1}^{\vartheta} \theta[i] z^{-is}.$$

Для моделі АР, з рівняння (6) випливає

$$x[t] = \sum_{i=1}^{p} \Phi[i] x[t-i] + a[t]. \quad (7)$$

Зауважимо, що оператори, які усувають сезонні коливання (6), діють не тільки на ці складові процесу, але і на інші складові нестаціонарного процесу. Видалення сезонної складової оператором $\nabla_s^D = (1 - z^{-s})^D$ спотворює тренд і навіть його видаляє. Цей оператор слабко впливає на стаціонарну складову процесу АРПКС. Видалення тренду оператором $\nabla^d = (1 - z^{-1})^d$ впливає на властивості сезонної складової та на стаціонарну складову процесу.

Модель авторегресії нестаціонарного процесу у СВВП поданні

Застосування СВВП подання сезонної складової враховує статистичні зв'язки підвекторів. При використанні СВВП уявлення, довжина підвектора дорівнює періоду сезонної складової. Модель АРПКС у СВВП поданні описується виразом

$$\Phi^n(z^{-1})\nabla^d \nabla_s^D \vec{x}^n[t] = Q^n(z^{-1})\vec{a}^n[t], \quad (8)$$

де оператори АР та КС набувають вигляду

$$\Phi^n(z^{-1}) = 1 - \sum_{i=1}^{p} \Phi^n[i] z^{-i},\ Q^n(z^{-1}) = 1 - \sum_{i=1}^{q} Q^n[i] z^{-i}.$$

У СВВП поданні модель АРПКС може бути записана і в більш складній формі, аналогічній (6):

$$\Phi_p^n(z^{-1})\varphi_P^n(z^{-s})\nabla^d \nabla_s^D \vec{x}^n[t] = Q_q^n(z^{-1})\theta_\vartheta^n(z^{-s})\vec{a}^n[t].$$

Зупинимося докладніше на процесі АР в СВВП представленні. Його можна отримати з (8), якщо покласти, що оператори, які усувають тренд і сезонність, рівні 1, а коефіцієнти КС рівні 0. Нижче розглядаються властивості СВВП на імітаційних процесах із заданими статистичними характеристиками. В якості статистичних характеристик використовується частота піку і його ширина смуги параметричної спектральної щільності потужності. Такі процеси можна отримати методом формуючого фільтра за параметрами АР моделей лінійного передбачення [13,17], використовуючи зв'язок коефіцієнтів АР, частот піків та їх ширин смуги. Різницеве рівняння АР СВВП має вигляд

$$\vec{x}^n[t] = \sum_{i=1}^{p} \Phi^n[i]\vec{x}^n[t-i] + \vec{a}^n[t]. \tag{9}$$

Умова оптимальності моделі АР СВВП полягає у статистичній незалежності підвекторів $\vec{a}^n[t]$. Для моделі АР СВВП помилки $\vec{a}^n[t]$ мають бути некорельованими, тобто $E\{\vec{a}^n[t]\vec{a}^n[t-i]\} = 0,$ при $i \neq 0$.

Роботу формуючого фільтра, що генерує процес АР у СВВП поданні, можна описати рівнянням (9). Затримки мають довжину n відліків. Згенерований процес АР СВВП складається з корельованих підвекторів, сформованих фільтром. У якості породжувального процесу використовуються підвектори типу білого шуму $\vec{a}^n[t]$ довжиною n. Природно, що процеси АР та АР у СВВП уявленні не збігаються. При $n = 1$ модель АР СВВП вироджується у звичайну модель АР випадкового процесу.

Для центрованого часового ряду формула оцінки кореляційної функції у СВВП поданні має вигляд

$$R^n[k] = (\vec{x}_i^n, \vec{x}_{i+k}^n) = \frac{1}{m-k} \sum_{i=1}^{m-k} \sum_{l=1}^{n} (x[in+l]x[in+l+kn]), \tag{10}$$

де $k-$ зсув часу кореляційної функції, $m-$ кількість підвекторів завдовжки n у часовому ряді, змінний індекс l приймає значення $l = 1,...,n.$ Скалярний добуток підвекторів $\vec{x}_i^n$ визначається виразом

$$c_{i,v}^n = (\vec{x}_i^n, \vec{x}_v^n) = \sum_{l=1}^{n} c_{i,v}^l.$$

Щоб отримати рівняння для розрахунку коефіцієнтів підвекторів АР процесу, помножимо (9) на $\vec{x}[t-k]$ та усереднимо. Після нескладних змін знайдемо рівняння типу Юла – Уокера для розрахунку параметрів моделі АР СВВП:

$$R^n[k] = \sum_{i=1}^{p} \Phi^n[i]R^n[i-k], \quad k = 1,2,...p. \tag{11}$$

При $k = 0$ отримаємо вираз, що пов'язує дисперсію підвекторів процесу, векторів процесу та векторів помилки передбачення

$$R^n[0] = \sum_{i=1}^{p} \Phi^n[i]R^n[i] + D_a^n.$$

Параметричні спектри СВВП процесів описуються параметрами АР їх моделей. З огляду на збіги форм (9) і (10), всі властивості моделі АР випадкових процесів справедливі й у моделі АР СВВП. Так, вираз для параметричної оцінки спектральної щільності потужності (СЩП) за моделлю АР у СВВП поданні має вигляд

$$P^n(f) = \frac{D_a^n}{\left|1 - \sum_{i=1}^{p} \Phi^n[i] e^{-j2\pi fiT}\right|^2}. \qquad (12)$$

Параметричну СЩП процесу в СВВП поданні, розраховану за (12) при $n = 12$, показано на графіку 1 (рис. 2). На графіку СЩП видно характер зміни підвекторів (середньомісячних температур) з часом. З аналізу графіка випливає, що сезонна складова часового ряду температур має довгостроковий тренд. Цей висновок можна зробити з графіка СЩП, який має найвищий максимум на нульовій частоті.

Графік 2 на рис. 2, що є класичним варіантом параметричної СЩП часового ряду, має суттєві відмінності від графіка 1 на рис. 2. Параметрична СЩП процесу, розрахована згідно з (12) при $n = 1$, має гострий пік на частоті, що відповідає періоду 12 (рис. 2).

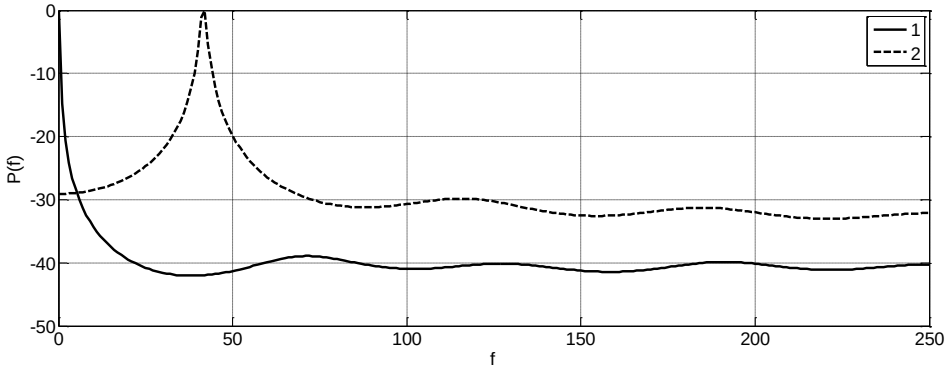


Рис. 2. Спектри АР(8) середньомісячних температур: 1 – параметрична СЩП за моделлю АР(8) у СВВП поданні, 2 – СЩП за моделлю АР(8) середньомісячних температур

Моделювання випадкового процесу АР у СВВП поданні

Промоделюємо імітаційний процес у СВВП поданні з трендом, у якого зміна підвекторів у часі описуються вузькосмуговим процесом другого порядку. СВВП АР процес утворювався підвекторами білого шуму, які подавалися на фільтр, що формує АР, з лініями затримки, рівними довжині підвектора. Параметри формуючого АР фільтра становили: центральна частота – 100, ширина смуги – 5, частота дискретизації 500. Коефіцієнти АР(2) для цих параметрів спектра дорівнювали: $\Phi[1] = 0{,}5989$, $\Phi[2] = -0{,}9391$. Потім генерувався корельований випадковий СВВП процес з довжиною підвектора 12 відліків. Зміна підвекторів у СВВП поданні моделювала зміну сезонної складової з періодом 12 у часі. На рис. 3 показаний цей процес довжиною 1608 відліків з трендом, що лінійно змінюється від 0 до 4, у якого сезонна складова описувалася СВВП моделлю із зазначеними частотними параметрами СЩП. Тренд адитивно поєднувався з АР(2) процесом у СВВП поданні.

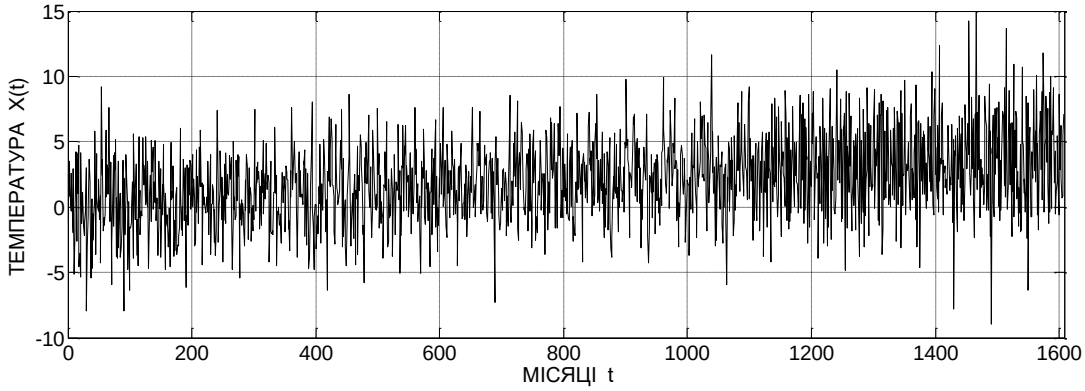


Рис. 3. Модель випадкового СВВП процесу АР(2) з адитивним довготривалим трендом

Помітний на рис. 4 тренд процесу оцінювався за описаним вище алгоритмом за допомогою формул (1) і (3). Згладжування проводилося ковзним фільтром з періодом згладжування 20. Порівняння істинного тренду та оцінного (рис. 4) показує їх хороший збіг.

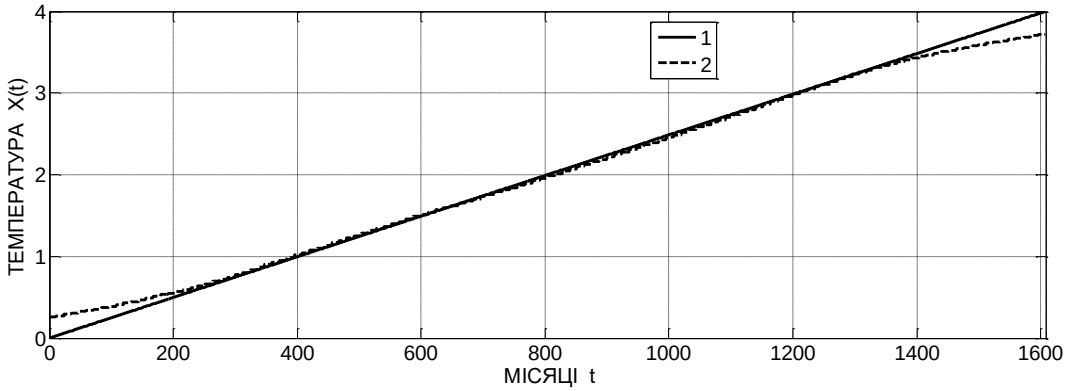


Рис. 4. Точність оцінки тренду: 1 – теоретичний тренд, 2 – оцінка тренда

Після видалення процесу тренду шляхом віднімання з процесу його оцінки, за параметричною СЩП оцінювався спектр зміни підвекторів у часі. Для цього за вибіркою оцінювали значення кореляційної функції (10) і підставлялися в систему рівнянь Юла – Уокера (11). При вирішенні цієї системи рівнянь знаходилися коефіцієнти АР(2). На рис. 5 представлена параметрична оцінка СЩП, знайдена за (12), з використанням моделі СВВП з довжиною підвектора 12. Теоретичне значення СЩП знаходилося за формулою (11), при $n=1$, безпосередньо за коефіцієнтами АР(2), що задаються. Порівняння оцінки СЩП з теоретичним значенням СЩП моделі показує, що вони добре збігаються.

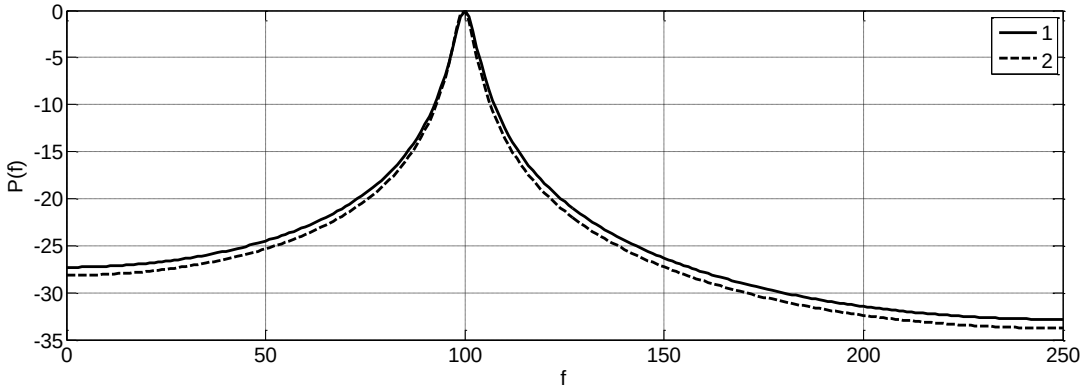


Рис. 5. Параметричні спектри за моделями АР(2): 1 – теоретична СЩП,
2 – оцінка СЩП сезонної складової моделі в СВВП поданні

Зауважимо, що оцінка СЩП без урахування моделі СВВП знайдена за (10), (11) та (12) при довжині підвектора $n=1$, схожа на спектр білого шуму (рис. 6). Однак значення функції автокореляції можуть бути ненульовими при більших зсувах часу, ніж ті, які використовувалися для оцінки. В даному випадку зсув часу дорівнює 8, так як з метою оцінки параметричної СЩП застосовувалася модель АР(8). Параметрична оцінка СЩП білого шуму квазірівномірна у всій області частот. Цьому відповідають коефіцієнти АР, близькі до нуля. Модель АР(8) використовуваної вибірки білого шуму має коефіцієнти АР: $\Phi[1]=0{,}019$; $\Phi[2]=0{,}002$; $\Phi[3]=0{,}012$; $\Phi[4]=-0{,}010$; $\Phi[5]=-0{,}030$; $\Phi[6]=0{,}006$; $\Phi[7]=-0{,}008$; $\Phi[8]=-0{,}040$. Таким чином, процес з властивостями білого шуму у СВВП поданні може мати ненульову кореляцію і вузькосмугову СЩП.

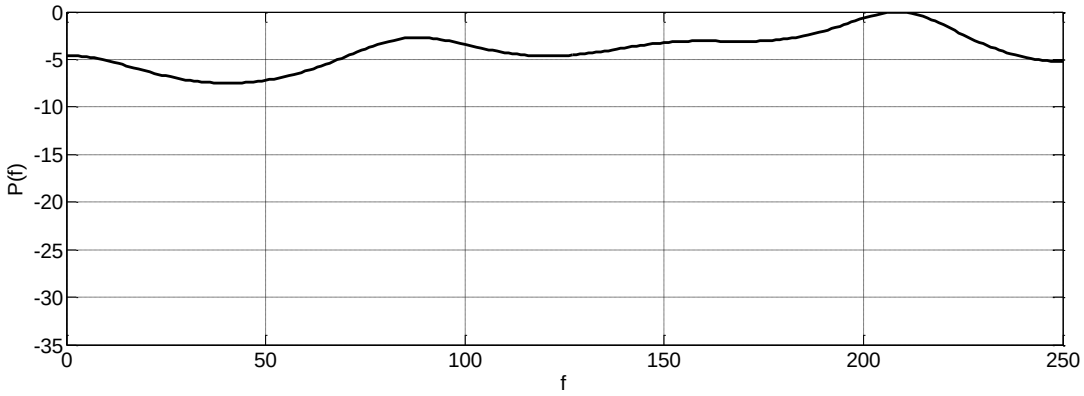


Рис. 6. Параметрична оцінка СЩП за моделлю АР(8)

Більш складна закономірність зміни сезонної складової нестаціонарного процесу в СВВП поданні моделювалася в наступному прикладі. Нестаціонарний процес включав лінійний тренд, що змінювався від 0 до 2. Підвектор був довжиною також 12 відліків. Зміна підвекторів описувалося моделлю АР(4) з частотами піків 50 і 120 і відповідними ширинами смуг 5 і 15. Вибір адитивної суміші тренду і АР процесу, представлена на рис. 7.

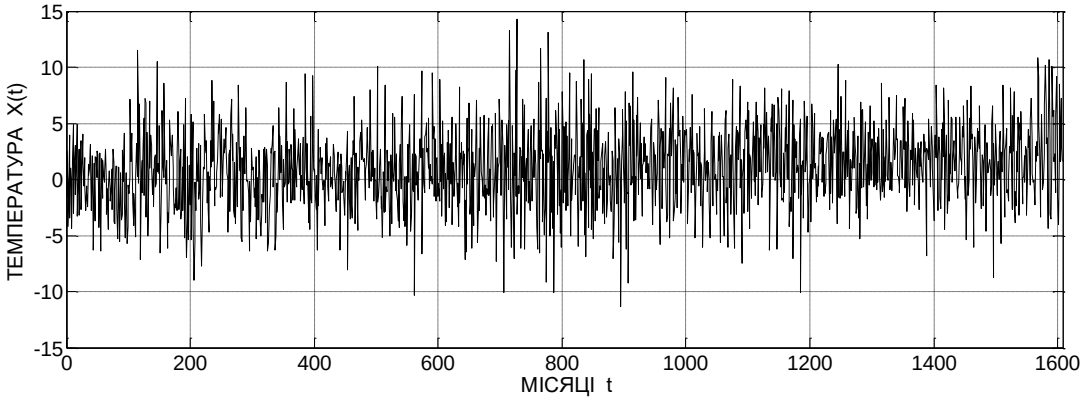


Рис. 7. Модель адитивного випадкового СВВП процесу АР(4) із довготривалим трендом

Тренд процесу, представлений на рис. 8, оцінювався описаним вище способом за допомогою формул (1) і (3). Згладжування проводилося ковзним фільтром з періодом згладжування 40. Порівняння істинного тренду та оцінного (рис. 8) показує їх хороший збіг.

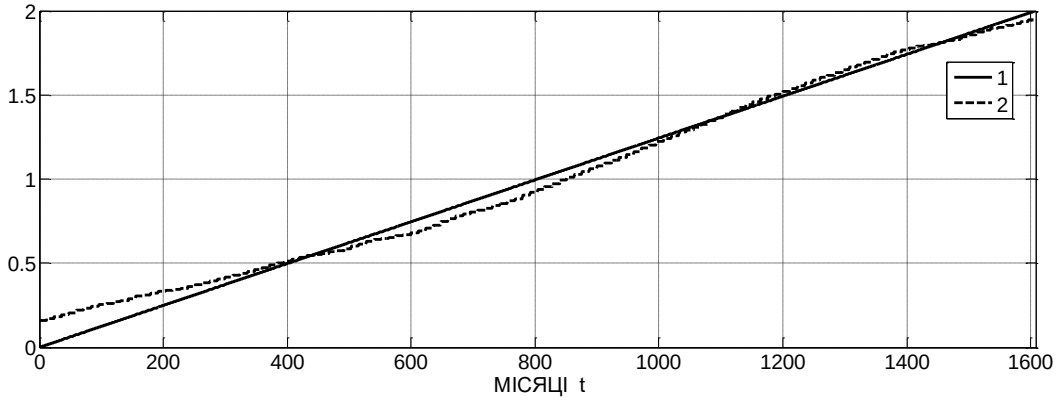


Рис. 8. Точність оцінки тренду: 1 – теоретичний тренд, 2 – оцінка тренда

Після видалення із суміші оцінки тренду за СЩП оцінювалася зміна підвекторів у часі. На рис. 9 представлена двомодова параметрична АР(4) СВВП оцінка СЩП, знайдена по (10). Порівняння оцінки СЩП із її теоретичним значенням показує, що вони добре збігаються.

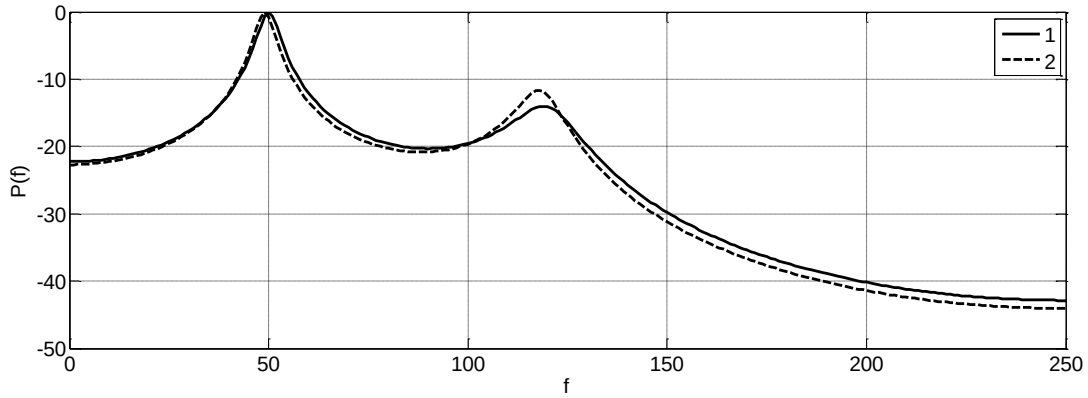


Рис. 9. Параметричні спектри за моделями АР(4):
1 – теоретична СЩП, 2 – оцінка СЩП сезонної складової моделі у СВВП поданні

Оцінка СЩП, знайдена за моделями АР з використанням (10), (11) та (12), при довжині підвектора не виявляє кореляції для зсувів часу 8 і схожа на спектр білого шуму (рис. 10). Зазначимо, що параметрична оцінка СЩП білого шуму приблизно рівномірна для всіх частот. Цьому відповідають коефіцієнти АР(8), близькі до нуля: $\Phi[1] = -0{,}040$; $\Phi[2] = 0{,}048$; $\Phi[3] = 0{,}045$; $\Phi[4] = -0{,}194$; $\Phi[5] = -0{,}067$; $\Phi[6] = 0{,}031$; $\Phi[7] = -0{,}037$; $\Phi[8] = 0{,}055$. Таким чином, процес із властивостями білого шуму, у СВВП поданні, може мати ненульову кореляцію.

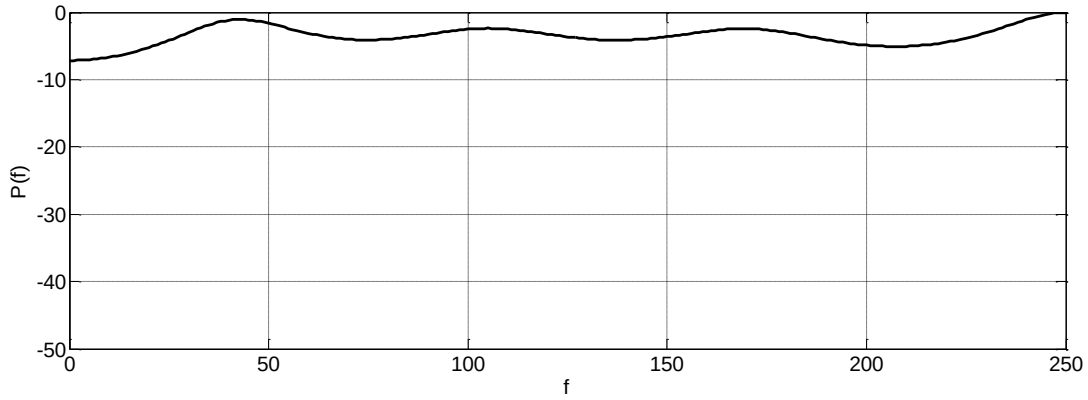


Рис. 10. Параметрична оцінка СЩП за моделлю АР(8)

Висновки

Показано деякі аспекти використання моделі СВВП для аналізу та моделювання нестаціонарних даних, що містять трендову та сезонну складові. Використання моделі СВВП дозволяє аналізувати також вид нестаціонарності, що враховує довготривалу зміну сезонної складової. Показано, що в рамках моделі СВВП шляхом знаходження середніх значень за період сезонної складової можна виділити тренд. При виділенні тренду слід використовувати усереднені сезонні дані за період, що дорівнює довжині підвектора. Подальше згладжування випадкової трендової складової здійснюється ковзним усередненням за допомогою низькочастотного фільтра.

Досліджувалися також оцінки зміни сезонної складової процесу за допомогою моделі АР з використанням СВВП уявлення. Достовірність оцінювання сезонної складової визначалася за СПМ АР моделі в СВВП поданні, у якої довжина підвекторів. Показано, що застосування моделі СВВП дозволяє знаходити закономірності процесів з властивостями білого шуму (у роботі використовувався зсув кореляційної функції, якій дорівнює 8).

Список літератури:

1. Карташов В.М., Тихонов В.А., Олейников В.Н. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : ХНУРЭ, 2014. 312 с.
2. Калистратова М.А., Кон А.И. Радиоакустическое зондирование атмосферы. Москва : Наука, 1985. 200 с.
3. Тихонов В.А., Чеботарёва Д.В. Прогнозирование потока данных в сетях мобильной связи // Матеріали Ш Міжнар. наук.-практ. конференції «Наукоємні технології в інфокомунікаціях» (23 - 25 травня 2019 р., Харків ; Кам′янець-Подільський, Україна). Харків : Мадрід, 2019. С. 126-127.
4. Ситнік О.В., Карташов В.М. Радіотехнічні системи : навч. посібник. Харків : Сміт, 2009. 448 с.
5. Кошкин Р.П. Беспилотные авиационные системы. Москва : Стратегические приоритеты, 2016. 676 с.
6. Макаренко С. И., Тимошенко А. В., Васильченко А. С. Анализ средств и способов противодействия беспилотным летательным аппаратам. Ч. 1. Беспилотный летательный аппарат как объект обнаружения и поражения // Системы управления, связи и безопасности. 2020. № 1. С. 109-146.
7. 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. 2019. Vol. 78, Is. 9. P. 771-781.
8. Oleynikov V. N., Zubkov O. V., Kartashov V. M., Korytsev I. V., Babkin S. I., Sheiko S. A. Investigation of detection and recognition efficiency of small unmanned aerial vehicles on their acoustic emission // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 9. P. 759-770.
9. В.А. Тихонов, В.М. Карташов, В.М. Олейников, В.И. Леонидов, Л.П. Тимошенко, И.С. Селезнев, Н.В. Рыбников. Обнаружение-распознавание беспилотных летательных аппаратов с использованием составной модели авторегрессии их акустического излучения // Вісник НТУУ «КПІ». Радіотехніка. Радіоапаратобудування. 2020. Вип. №81. С. 38–46. DOI: https://doi.org/10.20535/RADAP.2020.81.38-46.
10. Омельченко В.А., Безрук В.М., Коваленко Н.П. Распознавание заданных радиосигналов при наличии неизвестных сигналов на авторегрессионной основе // Радиотехника. 2001. № 123. С. 195–199.
11. Дробахин О.О. Автоматизация процесса распознавания сигналов дефектоскопа на основе модели линейного предсказания // Дефектоскопия. 1985. № 10. С. 64–67.
12. Рамишвили Г.С. Автоматическое распознавание говорящего по голосу. Москва : Радио и связь, 1981. 224 с.
13. Тихонов В.А., Безрук В.М. Модели линейного предсказания в статистической радиотехнике. Харків : ХНУРЕ, 2020. 468 с.
14. Марпл.-мл. С. Л. Цифровой спектральный анализ и его приложения. Москва : Мир, 1990. 584 с.
15. Бокс Дж., Дженкинс Г. Анализ временных рядов : пер. с. англ. Москва : Мир, 1974. Вып.1. 406 с.
16. Brockwell P.J., Davis R.A. Introduction to Time Series and Forecasting. Springer, 2002. P. 434.
17. Кармалита В.А. Цифровая обработка случайных колебаний. Москва : Машиностроение, 1986. 180 с.

Надійшла до редколегії 29.08.2022

Відомості про авторів:

Тихонов В'ячеслав Анатолійович – д-р ф.-м. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри інформаційно-мережної інженерії; Україна; e-mail: vyacheslav.tykhonov@nure.ua; ORCID: https://orcid.org/0000-0002-4618-4787

Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; e-mail: volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373

Карташов Олександр Володимирович – Харківський національний університет радіоелектроніки, здобувач кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; e-mail: mservicekh1@gmail.com

Посошенко Віталій Олександрович – Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vitalii.pososhenko@nure.ua; ORCID: https://orcid.org/0000-0003-0867-9161

УДК 621.372(075.8) **DOI:10.30837/rt.2022.3.210.15**

О.В. ЛАЗОРЕНКО, д-р фіз.-мат. наук, А.А. ОНИЩЕНКО,
Л. Ф. ЧОРНОГОР, д-р фіз.-мат. наук

МЕТОД КОРИГУЮЧОЇ ФУНКЦІЇ ДЛЯ ФРАКТАЛЬНОГО АНАЛІЗУ

Вступ

На початку XXI сторіччя революційні ідеї фракталізації, створені видатним американським вченим Б. Мандельбротом [1], спричинили появу принципово нового напрямку у сучасній радіофізиці – фрактальної радіофізики [2]. Фрактальний аналіз є одним із ефективних інструментів для дослідження фрактальних властивостей сигналів і процесів різного походження. Одними із основних числових характеристик, що використовуються у численних методах фрактального аналізу, є відповідні фрактальні розмірності [2]. Між тим, відомо, що точність оцінки цих розмірностей у переважній більшості випадків є досить невеликою, що не може задовольняти, у першу чергу, дослідників-практиків. Такими міркуваннями пояснюється актуальність даної роботи.

Метою роботи є створення простого й ефективного методу підвищення точності оцінювання фрактальних розмірностей у методах фрактального аналізу сигналів і процесів.

Модельні фрактальні та мультифрактальні сигнали

Для проведення досліджень, про які йдеться нижче, нами було створено великий набір з двох десятків модельних фрактальних (точніше кажучи, монофрактальних) і мультифрактальних сигналів. Але через обмеженість об’єму статті наведемо тільки три з них: дві моделі для монофрактальніх сигналів та одну – для мультифрактальних.

Модель 1. Ця модель є моделлю детермінованого фрактального НШС (ФНШС) сигналу, що базується на узагальненій функції Вейєрштраса [3], в якій всі випадкові фази дорівнюють нулю:

$$s_1(t) = \left[1 - b^{2D-4}\right] \frac{\sum_{n=0}^{M} b^{(D-2)n} \cos\left(2\pi s b^n t\right)}{1 - b^{(2D-4)(M+1)}},$$

де t – часова змінна, b – параметр масштабування за часом, D – фрактальна розмірність сигналу, $1 < D < 2$, M – кількість гармонік, які використовуються для побудови фізичного фракталу (якщо $M \to \infty$, то ми отримуємо математичний фрактал).

Модель 2. Дана модель є моделлю косинусної функції Вейєрштраса – Мандельброта [4]:

$$s_2(t) = \sum_{n=-\infty}^{+\infty} \frac{1 - \cos(\lambda^n t)}{\lambda^{(2-D)n}},$$

з фрактальною розмірністю $D = 1.5$, $\lambda > 1$. Вона є строго однорідною, а її графік є самоафінним [5].

Модель 3. Складна модель мультифрактального сигналу, що утворена адитивною сумою двох ФС, які базуються на модельному ФНШС сигналі $s_1(t)$. Перший з них має фрактальну розмірність $D = 1.8$, а його амплітуда зменшується за лінійним законом. Другий має $D = 1.2$, натомість його амплітуда зростає також за лінійним законом. Дана модель є типовою моделлю мультифрактального сигналу, оскільки є мультифрактальною як у глобальному, так і у локальному сенсі.

Метод коригуючої функції

Нехай є конкретний окремий метод монофрактального аналізу, який дозволяє отримати оцінку D^* невідомої фрактальної розмірності D досліджуваного сигналу, що представлено дискретним вектором даних, що містить N відліків. Оцінка D^* є невідомою нелінійною функцією величин D та N, тобто $D^* = f(D,N)$.

Отримаємо значення цієї функції на дискретній сітці: $D_{ij}^* = f(D_i, N_j)$, $i = \overline{1,n}$, $j = \overline{1,m}$. Для цього необхідно обрати модельний фрактальний сигнал (ФС), для якого величина його фрактальної розмірності D є заздалегідь відомою та може змінюватись у діапазоні $1 \le D < 2$. Чим щільнішою є дискретна сітка, тим краще, але збільшення величин n і m обмежується розумним об'ємом вектору даних, що зберігає величини D_{ij}^*, та часом їх обчислення для обраного методу монофрактального аналізу. Кожна величина D_{ij}^* представляється власною інтервальної оцінкою:

$$D_{ij}^* = \overline{D_{ij}^*} \pm \Delta D_{ij}^*, \tag{1}$$

де точкове значення $\overline{D_{ij}^*}$ та похибка ΔD_{ij}^* визначаються конкретним методом монофрактального аналізу.

Отже, функцію на дискретній сітці $D_{ij}^* = f(D_i, N_j)$, $i = \overline{1,n}$, $j = \overline{1,m}$ отримано.

Між тим, основною задачею є отримання оцінки невідомого значення фрактальної розмірності D аналізованого сигналу за відомим значенням D^* при відомому фіксованому значенні кількості відліків $N = N_{sig}$. Іншими словами, за фіксованого значення N_{sig} функція $D^* = f(D, N_{sig})$ як функція однієї змінної на проміжку $1 \le D < 2$ повинна мати обернену функцію $D = f^{-1}(D^*, N_{sig})$. Відомо, що це відбуватиметься тільки тоді, коли функція $D^* = f(D, N_{sig})$ є монотонною там. У даному випадку функція $D^* = f(D, N_{sig})$ повинна бути зростаючою функцією D на вказаному проміжку.

За порівняно великих значень N_{sig} ніяких проблем з цим не виникає. Але по мірі зменшення N_{sig} для функції $D_{ij}^* = f(D_i, N_j)$ знаходиться таке значення $N_{\min}$, нижче якого ($j = \overline{1,(\min-1)}$) указана вище монотонність функції $D^* = f(D, N_{sig})$ вже порушується. Це пояснюється наступними міркуваннями.

Наприклад, модельний ФНШС сигнал $s_1(t)$ з $D = 1.50$ (рис. 1) задано великою кількістю відліків (наприклад, $N = 131072 = 2^{17}$). Зменшувати кількість відліків N можна з використанням двох принципово різних стратегій.

Стратегія 1 («стратегія децимації»). Почнемо зменшувати N на кожному кроці вдвічі, використовуючи операцію децимації початкового сигналу. За використання прийнятної для друкованої продукції роздільної здатності (600 точок на дюйм) до $N = 4096 = 2^{12}$ неозброєним оком ми не побачимо жодних змін.

Але з теорії сигналів відомо, що така операція є еквівалентною пропусканню сигналу через фільтр високих частот. У результаті цього його спектр поступово звужуватиметься, а його фрактальні властивості погіршуватимуться.

Отже, зміни у отримуваній фрактальній розмірності D в такому випадку пояснюються змінами саме спектрального складу початкового сигналу.

Стратегія 2 («стратегія звуження вікна»). Друга стратегія полягає не в децимації початкового сигналу, а у поступовому звуженні (наприклад, вдвічі) прямокутного вікна у часовій області, крізь яке ми дивимось на сигнал. Якщо для першої стратегії після кожного кроку маємо «погіршену» копію початкового сигналу, то для другої стратегії бачимо тільки його окрему частину (рис. 2). Видно, що із зменшенням величини N характер змін досліджуваного сигналу є зовсім іншим. А тому й оцінки фрактальної розмірності такого сигналу будуть мати дещо інші значення, про що йтиметься у цій роботі нижче. До того ж указана вище монотонність функції $D^* = f(D, N_{sig})$ буде порушуватись за інших фіксованих значень N_{sig}, а тому величини $N_{\min}$, отримані для обох стратегій, можуть відрізнятися.

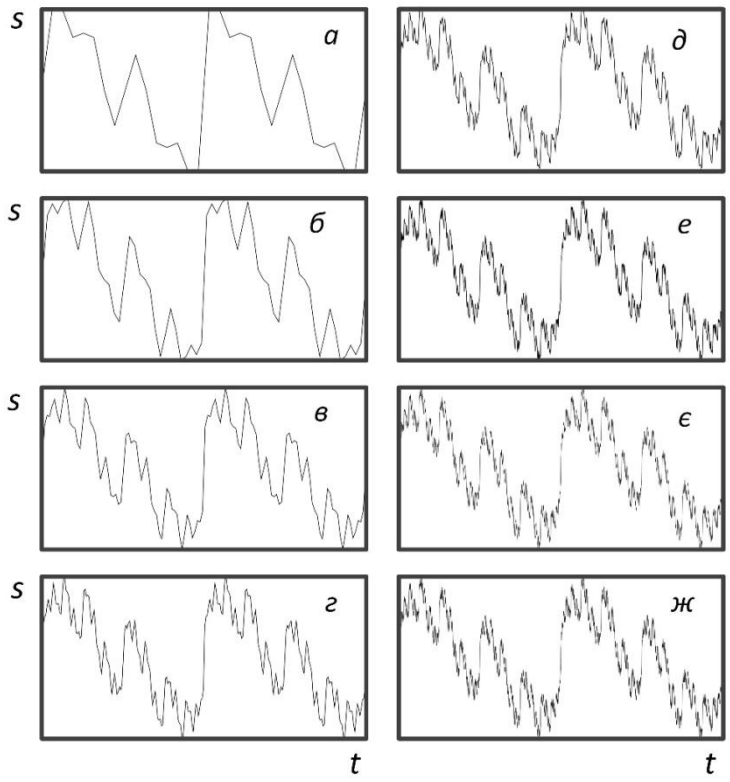


Рис. 1. Модель детермінованого ФНШС сигналу s_1 у часовій області для різних з $D = 1.50$ для різних значень кількості відліків N (стратегія 1): *а* – 32, *б* – 64, в – 128, *г* – 256, *д* – 512, *е* –1024, *є* – 2048, *ж* – 4096

Тепер декілька слів про застосування двох даних стратегій на практиці. Коли нам треба аналізувати довільний сигнал із заданою кількістю його відліків N, то цілком слушним виглядає використання першої стратегії, оскільки ми вважаємо, що перед нами є цілий сигнал, що має певні монофрактальні властивості.

Але коли потрібно будувати фрактограму (так часто називають часову залежність будь-якої фрактальної розмірності $D = D(t)$ [6]), наприклад при застосуванні методів узагальненого фрактального аналізу (УФА) [7] та динамічного фрактального аналізу (ДинФА) [8]), то на досліджуваний сигнал обов'язково накладається ковзаюче у часовій області вікно. Тоді стає у пригоді саме друга стратегія.

Повернемось до величини $N_{\min}$. Її сенс полягає в тому, що вона є мінімальним значенням довжини сигналу, для якого ще можна використовувати даний метод монофрактального аналізу. Цей результат є важливим для практиків, оскільки досі не існувало загального підходу до відповіді на дане запитання. Зауважимо також, що чим щільнішою є дискретна сітка, на якій пораховано функцію $D_{ij}^* = f(D_i, N_j)$, $i = \overline{1,n}$, $j = \overline{1,m}$, тим точнішою є оцінка величини $N_{\min}$.

Далі будуємо власне коригуючу функцію:

$$Cf_{ij} = \frac{D_{ij}^*}{D_i} = \frac{f(D_i, N_j)}{D_i},\ i = \overline{1,n},\ j = \overline{\min, m}. \qquad (2)$$

Зауважимо, що для $j = \overline{1,(\min-1)}$ робити цього немає сенсу, оскільки там не виконується умова монотонності функції $D^* = f(D, N_{sig})$. Отже, коригуюча функція є функцією двох дискретних змінних Cf_{ij}, $i = \overline{1,n}$, $j = \overline{\min, m}$.

Далі аналізуємо кількість відліків аналізованого сигналу N_{sig}. Вважаємо, що умова $N_{sig} \geq N_{\min}$ вже виконана (інакше всі подальші дії втрачають сенс через принципову неможливість коректного застосування обраного методу фрактального аналізу).

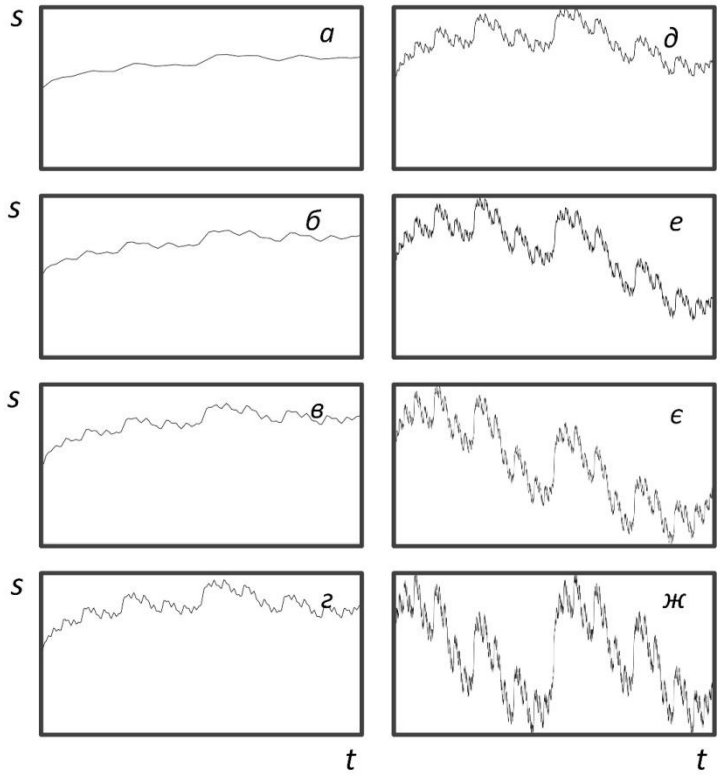


Рис. 2. Модель детермінованого ФНШС сигналу s_1 у часовій області для різних з $D = 1.50$ для різних значень кількості відліків N (стратегія 2): *а* – 32, *б* – 64, *в* – 128, *г* – 256, *д* – 512, *е* –1024, *є* – 2048, *ж* – 4096

Визначаємо величину l з умови $N_l \leq N_{sig} < N_{l+1}$. Після цього будуємо дискретний вектор Cf_i на основі коригуючої функції Cf_{ij}. Для цього, використовуючи лінійну апроксимацію коригуючої функції за змінною N між N_l та N_{l+1} для кожного фіксованого значення $i = \overline{1,n}$, отримуємо:

$$Cf_i = Cf_{il} + \\ +\left(Cf_{i(l+1)} - Cf_{il} \right) \frac{N_{sig} - N_l}{N_{(l+1)} - N_l}. \quad (3)$$

Тепер, маючи дискретний вектор Cf_i, можна обчислити відповідні дискретні значення $D_i^* = Cf_i \cdot D_i$, $i = \overline{1,n}$. Далі, вважаючи залежність між D і D^* лінійною на кожному відрізку $D_p \leq D \leq D_{p+1}$, $1 \leq p \leq (n-1)$, можна сконструювати наступну шматково-безперервну функцію $D^* = Cf(D)$, $1 \leq D < 2$:

$$\frac{D^* - D_p^*}{D_{p+1}^* - D_p^*} = \frac{D - D_p}{D_{p+1} - D_p}, \; 1 \leq p \leq (n-1).$$

Саме ця функція й дозволяє за відомою величиною оцінки фрактальної розмірності D^*, яка обчислюється в обраному методі монофрактального аналізу, отримати величину фрактальної розмірності:

$$D = D_p + \frac{D^* - Cf_p \cdot D_p}{Cf_{p+1} \cdot D_{p+1} - Cf_p \cdot D_p} \left(D_{p+1} - D_p \right), \quad (4)$$

де величина p, яка приймає натуральні значення та задовольняє умові $1 \leq p \leq (n-1)$, визначається зі співвідношення

$$Cf_p \cdot D_p \leq D^* \leq Cf_{p+1} \cdot D_{p+1}.$$

Зважаючи на те, що згідно з співвідношенням (1) як обчислювані під час моделювання оцінки величин D_{ij}^*, $i = \overline{1,n}$, $j = \overline{1,m}$, так і отримувані для аналізованого сигналу значення фрактальних розмірностей D^* задано в інтервальній формі

$$D^* = \overline{D^*} \pm \Delta D^*,$$

оцінку фрактальної розмірності D також слід записати в аналогічному вигляді:

$$D = \bar{D} \pm \Delta D, \tag{5}$$

де відповідно до формул (2) – (4) маємо:

$$\overline{D} = D_p + \frac{\overline{D^*} - \overline{Cf_p} \cdot D_p}{\overline{Cf_{p+1}} \cdot D_{p+1} - \overline{Cf_p} \cdot D_p}\left(D_{p+1} - D_p \right); \tag{6}$$

$$\overline{Cf_i} = \overline{Cf_{il}} + \left(\overline{Cf_{i(l+1)}} - \overline{Cf_{il}} \right)\frac{N_{sig} - N_l}{N_{(l+1)} - N_l}; \tag{7}$$

$$\overline{Cf_{ij}} = \frac{\overline{D_{ij}^*}}{D_i}; \tag{8}$$

$$\Delta D = \left(\overline{D} - D_p \right)\left[\frac{\Delta D^* + \Delta Cf_p \cdot D_p}{\overline{D^*} - \overline{Cf_p} \cdot D_p} + \frac{D_{p+1} \cdot \Delta Cf_{p+1} + D_p \cdot \Delta Cf_p}{\overline{Cf_{p+1}} \cdot D_{p+1} - \overline{Cf_p} \cdot D_p} \right]; \tag{9}$$

$$\Delta Cf_i = \Delta Cf_{il} + \frac{N_{sig} - N_l}{N_{(l+1)} - N_l}\left(\Delta Cf_{i(l+1)} + \Delta Cf_{il} \right); \tag{10}$$

$$\Delta Cf_{ij} = \frac{\Delta D_{ij}^*}{D_i}. \tag{11}$$

Зауважимо, що оскільки відповідно до формули (3) величини Cf_i залежать від кількості відліків аналізованого сигналу N_{sig}, тобто $Cf_i = Cf_i(N_{sig})$, то й оцінювана фрактальна розмірність D також залежить від N_{sig}. Її ж залежність від конкретного методу монофрактального аналізу міститься в величинах D_{ij}^*, $i = \overline{1,n}$, $j = \overline{1,m}$, що обчислюються з використанням модельних сигналів.

Якщо на практиці для побудови коригуючої функції використовувати тільки детерміновані сигнали, то у найпростішому випадку досить застосування одного модельного ФС. Але у разі використання модельних стохастичних ФС, де фрактальні властивості мають статистичний характер, необхідно обирати досить велику кількість модельних сигналів, а отримані результати усереднювати по ансамблю реалізацій, тобто

$$D_{ij}^* = \left\langle \overline{D_{ij}^*} \right\rangle \pm \left\langle \Delta D_{ij}^* \right\rangle,$$

де $\langle\ \rangle$ – операція усереднення по ансамблю модельних стохастичних ФС. Те саме можна зробити й у випадку, коли застосовується цілий набір моделей детермінованих ФС.

Таким чином, метод коригуючої функції (КФ) рекомендовано використовувати для покращення характеристик будь-якого методу монофрактального аналізу, в якому здійснюється оцінка фрактальної розмірності досліджуваного сигналу або процесу.

Практичне застосування методу КФ

Оскільки під час аналізу експериментальних даний у якості методів монофрактального аналізу автори роботи зазвичай використовують методи УФА та ДинФА, а також класичний метод нормованого розмаху [9], то розповімо про застосування методу КФ, у першу чергу, для регуляризаційної D_R [10], клітковой D_B [11], варіаційної D_V [12] та херстової D_H [13] фрактальних розмірностей.

Як було вказано, для кожної з цих розмірностей слід побудувати КФ, використовуючи набір модельних ФС із заздалегідь точно відомими значеннями фрактальної розмірності.

Таблиця 1

Кількість відліків, N	Теоретичне значення фрактальної розмірності D									
	1.10	1.20	1.30	1.40	1.50	1.60	1.70	1.80	1.90	1.99
32	1.45 ±0.05	1.47 ±0.05	1.50 ±0.05	1.52 ±0.05	1.54 ±0.05	1.57 ±0.06	1.57 ±0.06	1.58 ±0.06	1.60 ±0.06	1.64 ±0.06
64	1.40 ±0.03	1.43 ±0.03	1.46 ±0.03	1.50 ±0.03	1.53 ±0.03	1.56 ±0.03	1.59 ±0.03	1.62 ±0.03	1.64 ±0.03	1.72 ±0.02
128	1.35 ±0.02	1.38 ±0.02	1.42 ±0.02	1.46 ±0.02	1.49 ±0.02	1.53 ±0.02	1.56 ±0.02	1.60 ±0.02	1.65 ±0.02	1.78 ±0.01
256	1.34 ±0.01	1.38 ±0.01	1.42 ±0.01	1.46 ±0.01	1.50 ±0.01	1.55 ±0.01	1.59 ±0.01	1.64 ±0.01	1.66 ±0.01	1.81 ±0.01
512	1.33 ±0.01	1.37 ±0.01	1.42 ±0.01	1.47 ±0.01	1.52 ±0.01	1.58 ±0.01	1.63 ±0.01	1.68 ±0.01	1.71 ±0.01	1.82 ±0.01
1024	1.32 ±0.01	1.37 ±0.01	1.42 ±0.01	1.47 ±0.01	1.53 ±0.01	1.59 ±0.01	1.64 ±0.01	1.70 ±0.01	1.74 ±0.01	1.82 ±0.01
2048	1.315 ±0.005	1.362 ±0.005	1.416 ±0.005	1.473 ±0.005	1.533 ±0.005	1.595 ±0.004	1.655 ±0.004	1.713 ±0.004	1.769 ±0.003	1.861 ±0.002
4096	1.312 ±0.004	1.363 ±0.003	1.419 ±0.003	1.479 ±0.003	1.543 ±0.003	1.607 ±0.003	1.670 ±0.003	1.730 ±0.003	1.782 ±0.002	1.882 ±0.002
8192	1.309 ±0.003	1.362 ±0.002	1.421 ±0.002	1.484 ±0.002	1.551 ±0.002	1.620 ±0.002	1.688 ±0.002	1.753 ±0.002	1.807 ±0.002	1.842 ±0.001
16384	1.307 ±0.002	1.361 ±0.002	1.421 ±0.002	1.486 ±0.002	1.554 ±0.002	1.625 ±0.002	1.694 ±0.001	1.761 ±0.001	1.824 ±0.001	1.857 ±0.06
32768	1.305 ±0.001	1.361 ±0.001	1.422 ±0.001	1.488 ±0.001	1.558 ±0.001	1.629 ±0.001	1.700 ±0.001	1.767 ±0.001	1.831 ±0.001	1.882 ±0.000
65536	1.304 ±0.001	1.360 ±0.001	1.423 ±0.001	1.489 ±0.001	1.561 ±0.001	1.633 ±0.001	1.706 ±0.001	1.775 ±0.001	1.838 ±0.001	1.859 ±0.000
131072	1.303 ±0.001	1.360 ±0.001	1.423 ±0.001	1.490 ±0.001	1.563 ±0.001	1.636 ±0.001	1.710 ±0.001	1.785 ±0.001	1.848 ±0.001	1.881 ±0.000

У нашому випадку такою зручною моделлю є модель s_1, що описує ФНШС сигнал. Всі необхідні розрахунки було проведено з використанням FracLab Toolbox (версія 2.2) [14] для системи комп'ютерної математики (СКМ) MATLAB/SciLab, а також оригінального програмного забезпечення, створеного авторами роботи.

Побудова коригуючих функцій для різних фрактальних розмірностей проводилася окремо для кожної з обох стратегій. У якості прикладу тут наводимо оцінки величин херстової розмірності D_H у межах стратегії 1 (див. табл. 1). Сірим кольором у таблиці виділено рядки, де умови монотонності функції $D^* = f(D, N_{sig})$ не виконуються. Варто зазначити, що для стратегії 1 величина $N_{\min}$ складає 64 для D_H і D_R, 512 для D_B та 256 для D_V, що вкрай важливо для практиків.

Крім того, слід відзначити загальну тенденцію зменшення величини оцінюваної фрактальної розмірності D^* зі збільшенням N за фіксованого значення D. За тих самих умов також існує тенденція зменшення величини відповідної відносної похибки $\Delta D^* / D^*$.

Слід зазначити, що для стратегії 2 величина $N_{\min}$ складає 32 для D_H і D_R, 256 для D_B та 64 для D_V. На відміну від результатів застосування стратегії 1 тут спостерігається тенденція збільшення величини оцінюваної розмірності D^* зі збільшенням N за фіксованого значення D. За тих самих умов також існує тенденція зменшення величини відповідної відносної похибки $\Delta D^* / D^*$.

Ґрунтуючись на отриманих результатах, для кожної з двох стратегій і кожної з чотирьох фрактальних розмірностей (регуляризаційної D_R, кліткової D_B, варіаційної D_V та херстової D_H), на основі співвідношень (5) – (11) сконструйовано коригуючі функції та створено відповідне програмне забезпечення в середовищі СКМ MATLAB/SciLab. Для програмної реалізації коригуючих функцій використовувалися результати моделювання з кроком 0.01 по величині фрактальної розмірності, що, на нашу думку, є цілком прийнятним для практичних цілей.

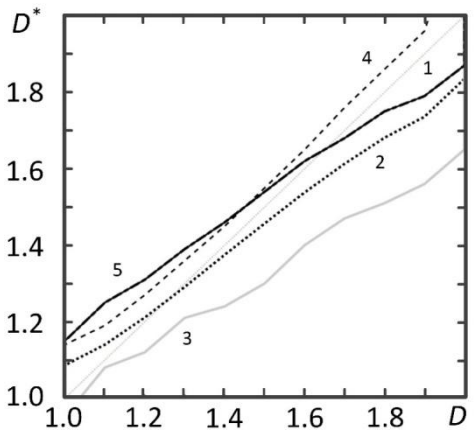


Рис. 3. Залежності $D^* = f(D, N_{sig})$ між оцінюваною D^* та скоригованою методом КФ D фрактальними розмірностями за фіксованого значення кількості точок досліджуваного ФС ($N_{sig} = 1024$) для: 1 – ідеального випадку, 2 – херстової розмірності $D_H^* = f(D_H)$, 3 – кліткової розмірності $D_B^* = f(D_B)$, 4 – регуляризаційної розмірності $D_R^* = f(D_R)$, 5 – варіаційної розмірності $D_V^* = f(D_V)$

Зазначимо, що найгірший результат у сенсі величин $N_{\min}$ для обох стратегій показала кліткова розмірність D_B. Це було досить очікувано, оскільки при її обчисленні використовувався метод звичайних кліток – один з найстаріших методів фрактального аналізу. Він має певну ваду, що полягає у так званому ефекті виснаження (див., наприклад, [15]), коли для дуже малих розмірів кліток залежність, що апроксимується лінійною регресією, яка дозволяє отримати величину D_B, починає істотно відрізнятися від лінійної. Останнє призводить до зменшення точності оцінювання D_B. Тому у роботі ми не апроксимуємо останню третину точок, де саме й проявляється цій ефект.

Корисним і цікавим здається також оцінка поведінки залежності $D^* = f(D, N_{sig})$ з використанням методу КФ за фіксованого значення N_{sig}. Результати такої оцінки, отриманої для регуляризаційної D_R, кліткової D_B, варіаційної D_V та херстової D_H фрактальних розмірностей при $N_{sig} = 1024$, наведено на рис. 3. Видно, що на відміну від ідеальної залежності ($D^* = D$ – пряма 1) реальні залежності $D_R^* = f(D_R)$ (крива 4), $D_B^* = f(D_B)$ (крива 3), $D_V^* = f(D_V)$ (крива 5) та $D_H^* = f(D_H)$ (крива 2) є, як і зазначалося вище, нелінійними.

Саме тому на практиці враховувати такі залежності зручно з використанням засобів комп'ютерної техніки, що саме й здійснювалося в даній роботі.

Аналіз модельних сигналів

Гарною ілюстрацією ефективності метода КФ є його застосування у методі ДинФА, коли у якості обчислюваної фрактальної розмірності використовується херстова розмірність D_H. Тут використовуємо ті самі моделі, що й в попередньому пункті. Результати аналізу зручно представляти в наступному форматі, який розглянемо на прикладі моделі 2 (рис. 4, *а*, *є*).

У лівому стовпчику під графіком самого сигналу у часовій області (рис. 4, *а*) розташовано традиційну для метода ДинФА функцію $D_H(t,T)$ (рис. 4, *б*), обчислювану з використанням методу нормованого розмаху у прямокутному спектральному вікні шириною T, центр якого розташовано на часовій осі в точці з координатою t. Геометричним образом функції $D_H(t,T)$ є поверхня, яку зображено з використанням кольорової схеми, наведеної на рис. 4, *б* праворуч від графіка цієї функції.

Нижче містяться чотири графіки залежностей $D_H(t,T_1)$ (рис. 4, *е*), $D_H(t,T_2)$ (рис. 4, *д*), $D_H(t,T_3)$ (рис. 4, *г*) та $D_H(t,T_4)$ (рис. 4, *в*) для фіксованих значень ширини вікна $T_1 = 0.004$, $T_2 = 0.833$, $T_3 = 1.667$ і $T_4 = 2.500$, які також позначено стрілками на рис. 4, *б* (ліворуч від графіка двовимірної функції $D_H(t,T)$). Для порівняння на рис. 4, *в – е, з – ї* також штрихованою лінією продемонстровано ідеальну часову залежність $D(t)$ фрактальної розмірності модельного ФС.

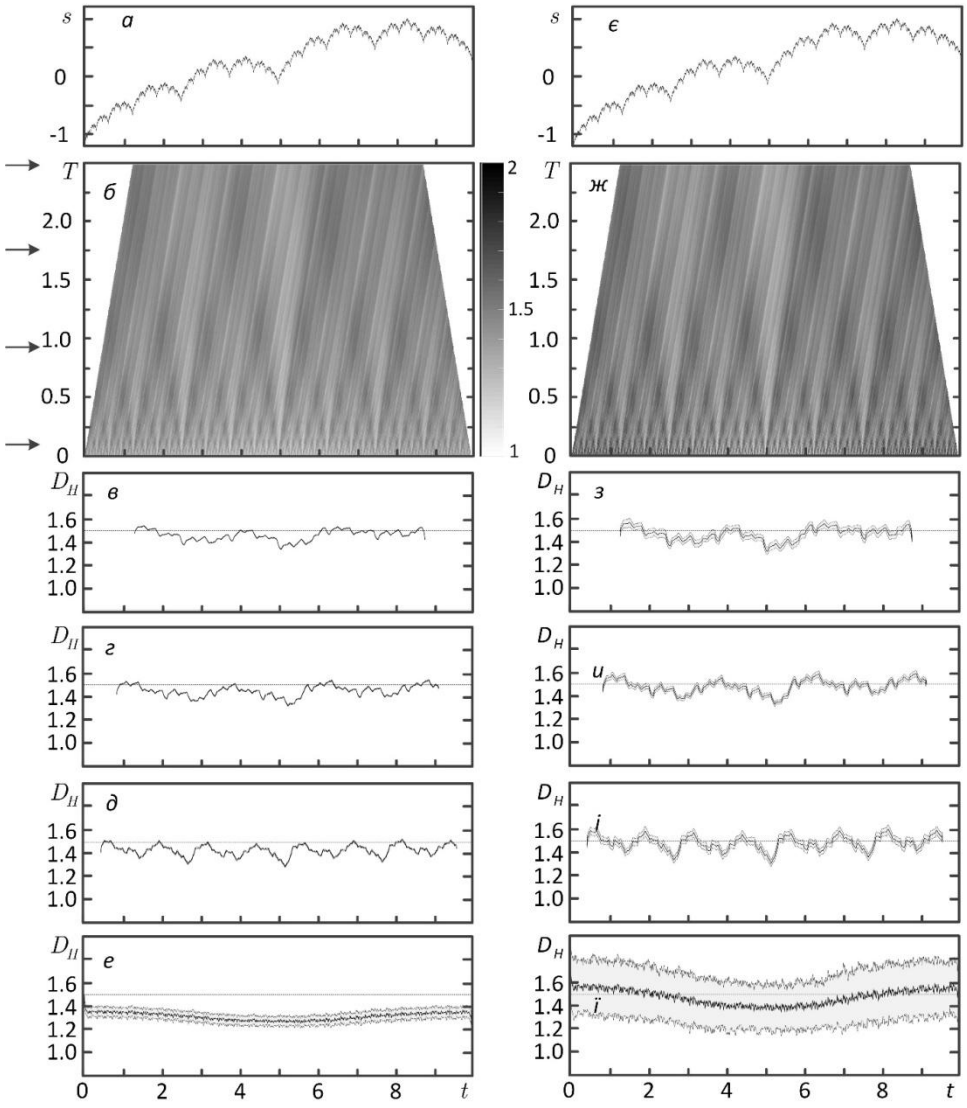


Рис. 4. Результати застосування методу КФ під час аналізу модельного ФС (модель 1) з використанням метода ДинФА: *а*, *є* – ФС у часовій області, *б* – херстова розмірність $D_H(t,T)$, *в, г, д, е* – херстова розмірність $D_{Hi}(t) \equiv D_H(t,T_i)$ за фіксованих значень T_i (позначено стрілками ліворуч на панелі б); *ж, з, і, ї* – те ж саме, але з використанням методу КФ

У правому стовпчику під графіком модельного сигналу у часовій області (рис. 4, *є*) перебуває скоригована з використанням методу КФ функція $D_H(t,T)$ (рис. 4, *ж*). Нижче неї розташовано відповідні скориговані функції $D_H(t,T_1)$ (рис. 4, *ї*), $D_H(t,T_2)$ (рис. 4, *і*), $D_H(t,T_3)$ (рис. 4, *и*) та $D_H(t,T_4)$ (рис. 4, *з*), що відповідають тим самим значенням ширини вікна, що й у лівому стовпчику.

Сірою заливкою на рис. 4, *в – е, з – ї* показано довірчі інтервали оцінюваних залежностей (рівень надійності 0.9). Зазначимо, що як й очікувалося, похибка оцінювання херстової розмірності ΔD_H зростає із зменшенням ширини вікна T. Розрахунки $D_H(t,T)$ проводилися для величин ширини вікна N від 32 до 2048 відліків, чому відповідає змінення безрозмірної змінної T від $T_1 = 0.004$ до $T_4 = 2.500$.

Проводячи порівняння лівого (нескориговані значення) та правого (скориговані значення) стовпчиків, можна зробити наступні висновки.

Результати застосування методу ДинФА для аналізу моделі 2 (рис. 4, *б – е*) мали певну ваду, яка полягає у тому, що оцінки херстової розмірності виявлялися суттєво заниженими, особливо для найбільш вузьких вікон, чому відповідають малі значення змінної T. Тут же

маємо можливість пересвідчитись, що використання методу КФ істотно покращує ситуацію. Натомість, певною платою за це є збільшення ширини довірчого інтервалу, що добре помітно на рис. 4, *з – ї*.

Дослідження моделі 3, що є складною моделлю мультифрактального сигналу, здавалось би, виходить за межі тематики даної статті. Натомість це не зовсім так, оскільки на практиці дослідникові заздалегідь взагалі невідомо, з яким саме (монофрактальним, мультифрактальним або взагалі нефрактальним) сигналом він матиме справу.

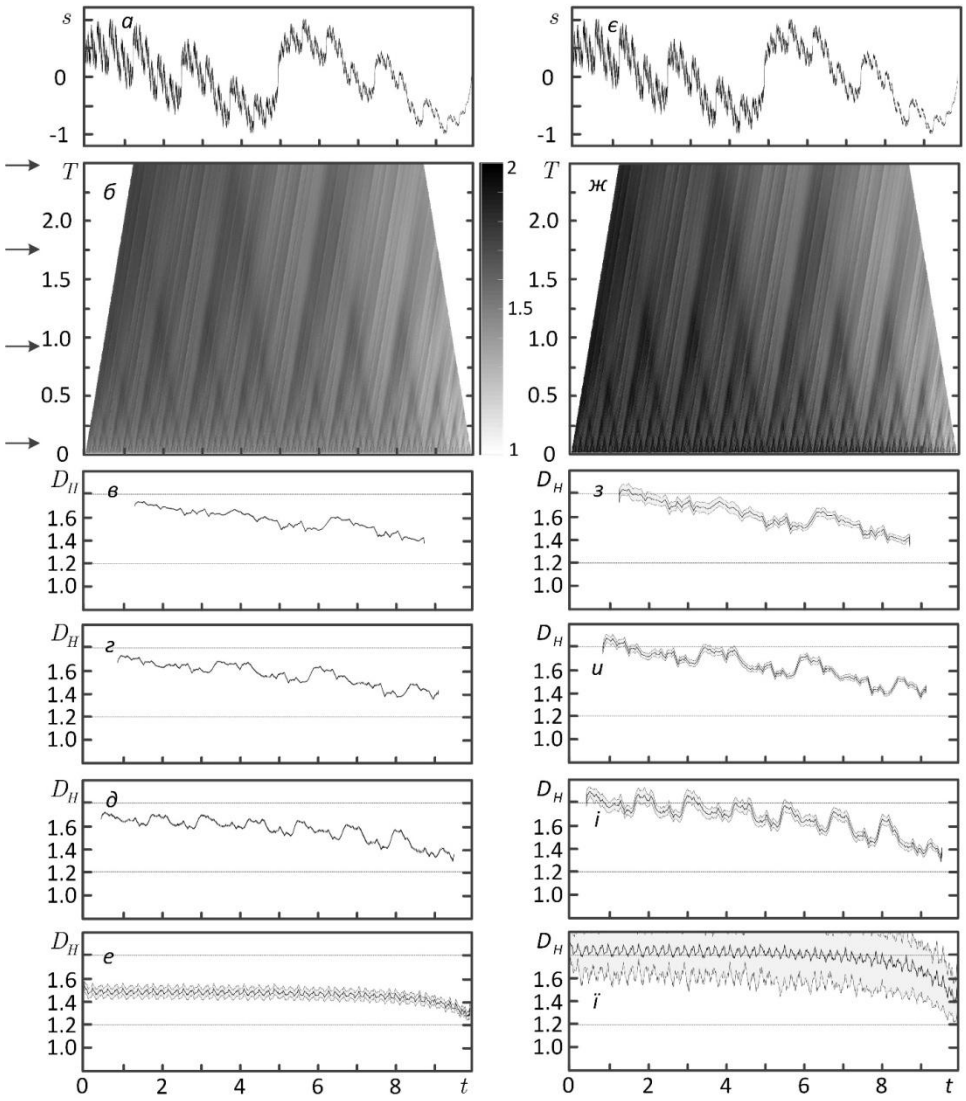


Рис. 5. Результати застосування методу КФ під час аналізу модельного ФС (модель 3) з використанням метода ДинФА: а, є – ФС у часовій області, б – херстова розмірність $D_H(t,T)$, в, г, д, е – херстова розмірність $D_{Hi}(t) \equiv D_H(t,T_i)$ за фіксованих значень T_i (позначено стрілками ліворуч на панелі б); ж, з, і, ї – те ж саме з використанням методу КФ. Сірим кольором позначено довірчі інтервали

Як було вже сказано, модель 3 є принципово мультифрактальною як у глобальному, так і у локальному сенсі, оскільки у кожен момент часу в ній присутні одразу дві монофрактальні компоненти з різними фрактальними розмірностями та різними відношеннями амплітуд цих компонент. Оскільки амплітуда компоненти з більшою фрактальною розмірністю ($D = 1.8$) лінійно зменшується, а компоненти з меншою фрактальною розмірністю ($D = 1.2$) лінійно зростає, оцінка їх адитивної суми у часі постійно змінюється. Для звичайного методу ДинФА (рис. 5, *в – е*) можна побачити, що із зменшенням ширини вікна ця оцінка стає все більш неточною. Особливо це помітно на рис. 5, *е*. Натомість застосування методу КФ (рис. 5, *з – ї*) істотно покращує ситуацію.

Із аналізу результатів моделі 3 також можна побачити, що найгірші результати спостерігаються при малих значеннях фрактальної розмірності D. Але останній результат не є вадою методу КФ. Слід зазначити, що коли фрактальна розмірність D наближається до 1, сам ФС потроху втрачає фрактальні властивості та при $D = 1$ вироджується у гладку криву. Саме в цих умовах метод нормованого розмаху (див., наприклад, [16 – 20]), на якому ґрунтується обчислення херстової розмірності, дійсно показує найгірші результати. Для інших фрактальн-

них розмірностей, які ми оцінювали у наших дослідженнях (регуляризаційної D_R, кліткової D_B, варіаційної D_V) при малих D, результати методу КФ виглядали значно кращими.

Висновки

1. Створено метод КФ, який дозволяє компенсувати завжди існуючу нелінійність залежності між істинним значенням фрактальної розмірності та її оцінкою, здійсненою з використанням обраного методу монофрактального аналізу сигналів і процесів за відомої кількості відліків дискретного вектору даних досліджуваного сигналу. Ідея методу полягає у побудові та застосуванні спеціальної КФ з використанням набору модельних ФС із заздалегідь відомими значеннями фрактальної розмірності.

2. У рамках двох різних стратегій зменшення кількості відліків дискретного вектору даних досліджуваного модельного сигналу із використанням КФ для кліткової D_B, варіаційної D_V, регуляризаційної D_R та херстової D_H фрактальних розмірностей знайдено мінімальну кількість відліків $N_{\min}$, за якої ще можна оцінювати відповідні розмірності. Встановлено, що $N_{\min}$ дорівнює 64 для D_H і D_R, 512 – для D_B та 256 – для D_V у рамках «стратегії децимації» та 32 – для D_H і D_R, 256 – для D_B та 64 – для D_V у рамках «стратегії звуження вікна».

3. За умов використання мінімально дозволеної кількості відліків $N_{\min}$ завдяки застосуванню метода КФ максимальне відхилення оцінюваної фрактальної розмірності від істинного відомого значення, яке дорівнювало: 1) у рамках «стратегії децимації» для D_B – 23 %, для D_R – 32 %, для D_V – 56 %, для D_H – 27 %; 2) у рамках «стратегії звуження вікна» для D_B – 23 %, для D_R – 21 %, для D_V – 12 %, для D_H -31 %, тепер не перевищує $5-7$ %, а самі істинні значення даних фрактальних розмірностей впевнено потрапляють до довірчого інтервалу отриманої оцінки, побудованого для рівня надійності 0.9.

Список літератури:

1. Mandelbrot B. B. The Fractal Geometry of Nature. New York : W. H. Freeman and Company, 1982. 468 p.
2. Лазоренко О. В., Черногор Л. Ф. Фрактальная радиофизика. 1. Теоретические основы // Радиофизика и радиоастрономия. 2020. Т. 25, № 1. С. 3 – 77.
3. West B. J., Bologna M., Grigolini P. Physics of Fractal Operators. New York : Springer-Verlag, 2003. 349 p.,
4. Feder J. Fractals. New York and London : Springer, 1988. 284 p.
5. Bandt C., Barnsley M., Devaney R., Falconer K. J., Kannan V., and Vinod Kumar P. B., eds. Fractals, Wavelets, and their Applications // Contributions from the International Conference and Workshop on Fractals and Wavelets (Springer Proceedings in Mathematics & Statistics). Switzerland: Springer Int. Publ., 2014. 508 p.
6. Raghavendra B. S, Narayana Dutt D. Computing fractal dimension of signals using multiresolution box-counting method // J. Inf. Math. Sci. 2010. Vol. 6, No. 1. P. 50 – 65.
7. Chemogor L. F., Lazorenko O. V. and Onishchenko A. A. Fractal Analysis of the Gravitational Waves as a Unique Ultra-Wideband Process // Proc. 9th International Conference on Ultrawideband and Ultrashort Impulse Signals, September 4-7, 2018, Odessa, Ukraine. Odessa, 2018. P. 34 – 39.
8 Onishchenko A., Chemogor L., Lazorenko O. Dynamical Fractal Analysis of the Acoustic Ultra-Wideband Signal Caused by the Chelyabinsk Meteoroid. Eskiehir Technical Univ // J. of Sei. and Tech. A -Appl. Sei. and Eng. 2019. Vol. 20. P. 188 – 192.
9. Hardy H. H., Beier R. A. Fractals in Reservoir Engineering. Singapore, New Jersey, London, Hong Kong : World Scientific, 1994. 359 p.
10. Roueff F., Levy-Vehel J. A Regularization Approach to Fractional Dimension Estimation. M. M. Novak // Fractals 98, Oct 1998, Valleta, Malta. World Scientific, 1998.
11. Feldman D. P. Chaos and Fractals. An Elementary Introduction. Oxford : Oxford University Press, 2012. 408 p.
12. Prigarin S. M., Hahn K., Winkler G. Variational dimension of random sequences and its application // Numerical Analysis and Applications. 2009. Vol. 2, No. 4. P. 352 – 363.
13. Mandelbrot B. B. Fractals: Form, Chance and Dimension. San Francisco : W. H. Freeman and Company, 1977. 468 p.
14. Legrand P. & Levy-Vehel, J. Signal and image processing with FRACLAB. FRACTAL04 // Complexity and Fractals in Nature, 8th International Multidisciplinary Conference, 2004.
15. Tong H. Dimension Estimation and Models. New Jersey : World Scientific, 1993. 223 p.

16. Levy-Vehel J., Lutton E. Fractals in Engineering. New Trends in Theory and Applications. New York : Springer-Verlag, 2005. 289 p.
17. Turcotte D. L. Fractals and Chaos in Geology and Geophysics. Cambridge : Cambridge University Press, 1997. 398 p.
18. Seuront L. Fractals and Multifractals in Ecology and Aquatic Science. Boca Raton : CRC Press, 2010. 344 p.
19. Alessio E., Carbone A., Castelli G., Frappietro V. Second-order moving average and scaling of stochastic time series // The European Physical Journal B – Condensed Matter. 2002. Vol. 27, No. (2). P. 197 – 200.
20. Taqqu M. S., Teverovsky V., Willinger W. Estimators for Long-Range Dependence: An Empirical Study // Fractals. 1995. Vol. 03, No. 04. P. 785 – 798.

Надійшла до редколегії 07.09.2022

Відомості про авторів:
Лазоренко Олег Валерійович – д-р фіз.-мат. наук, доцент, Харківський національний університет імені В. Н. Каразіна, завідувач кафедри загальної фізики, Україна; e-mail: Oleg.V.Lazorenko@karazin.ua; ORCID: https://orcid.org/0000-0002-0250-8671
Онищенко Андрій Анатолійович – Харківський національний університет радіоелектроніки, старший викладач кафедри фізики, Україна; e-mail: andrey.onishchenko@nure.ua, ORCID: https://orcid.org/0000-0002-2118-9119
Чорногор Леонід Феоктистович – д-р фізико-математичних наук, професор, Харківський національний університет імені В. Н. Каразіна, завідувач кафедри космічної радіофізики, Україна; e-mail: Leonid.F.Chernogor@karazin.ua; ORCID: https://orcid.org/0000-0001-5777-2392

УДК 621.396.967.2 **DOI:10.30837/rt.2022.3.210.16**

О.І. РОМАНОВ, д-р техн. наук, І.В. СВИД, канд. техн. наук,
Н.І. КОРНІЄНКО, А.О. РОМАНОВ

УПРАВЛІННЯ ОПТИЧНОЮ МЕРЕЖЕЮ КОНТРОЛЕРОМ SDN НА БАЗІ ONOS

Постановка проблеми

Одним з найбільш перспективних напрямків розвитку сучасних телекомунікацій є технологія SDN. Питання побудови мереж на базі цієї технології знаходяться в центрі уваги представників науково-дослідних організацій, університетів і операторів мобільного зв'язку.

Архітектура програмно-конфігурованої мережі складається з трьох рівнів: додатків, управління та інфраструктури, пов'язаних один з одним через відкриті API-інтерфейси (рис. 1).

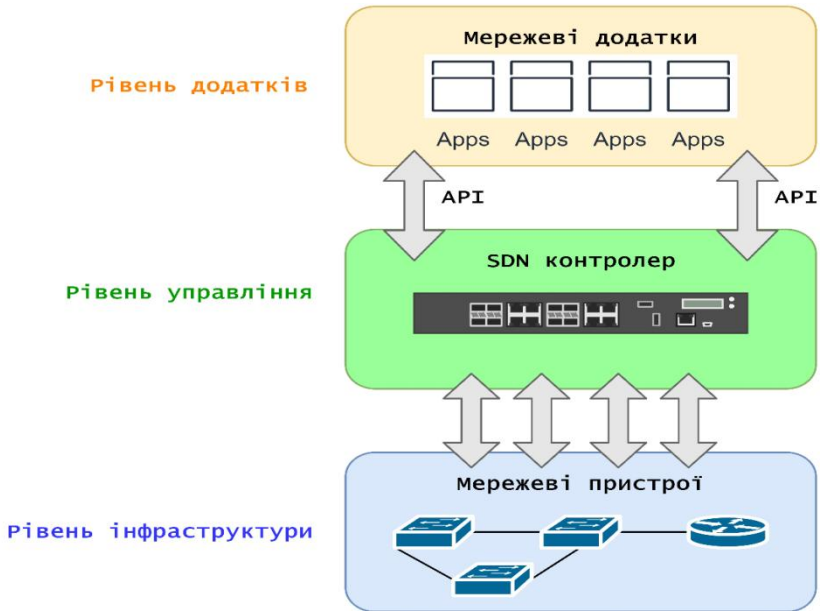


Рис. 1. Загальна архітектура SDN

Рівень додатків містить функціональні блоки з набором програмного забезпечення, яке може вирішувати окремі завдання управління, надавати сучасні сервіси користувачам, обробляти статистичні дані про стан мережевих елементів, забезпечити автоматизовану зміну протоколів при налаштуванні мережі, віртуалізовувати мережеві функції, балансувати навантаження та інше.

Рівень управління побудований на базі контролера, який є централізованим органом управління мережею. На цьому рівні забезпечується управління політиками і трафіком в мережі. Рівень інфраструктури містить як фізичні, так і віртуальні елементи мережі, у яких є тільки виконавчі функції. Усі задачі управління вирішує контролер, який доводить ці рішення до рівня інфраструктури.

Аналіз останніх досліджень і публікацій

Сьогодні ведуться дослідження і розробляються пропозиції щодо практичної реалізації цієї технології в рамках ряду проєктів. Ці проєкти охоплюють створення інфраструктури широкосмугового доступу наступного покоління, побудову сервісної платформи (services

platform) надання послуг, платформи для побудови SD-RAN і Core RAN мереж операторів мобільного зв'язку 5G/6G. Найбільший внесок в розвиток даного напрямку вносять представники консорціуму Open Networking Foundation (ONF). ONF стоїть на чолі всіх проєктів з побудови SDN і займається прискоренням процесів їх практичної реалізації [1].

Опубліковано ряд документів, що описують принципи побудови і функціонування мереж SDN. У роботах [2 – 6] розглянуто загальні вимоги, системні підходи і узагальнено архітектуру мереж SDN. У роботах [2, 7, 8, 11] розглянуто завдання і особливості протоколів, які використовуються при вирішенні різних завдань в мережах SDN. У роботах [9, 10, 12] описано особливості побудови елементів мережі SDN та порядок їх взаємодії в процесі обслуговування потоків інформації. У роботах [13 – 17] розглянуто принципи побудови оптичної транспортної мережі та надано рекомендації щодо забезпечення безпеки їх функціонування. Найбільш повно і систематизовано матеріал викладено в роботах [18 – 21]. У статтях [22 – 26] досліджуються різні аспекти обслуговування інформаційних потоків і акцентується увага на необхідності виконання вимог щодо забезпечення безпеки мереж.

Постановка завдання

Найбільш складні завдання в мережах SDN стоять перед рівнем управління. Вирішення цих завдань покладається на контролер, основним елементом якого є мережева операційна система. Розглянемо особливості побудови відкритої мережевої операційної системи (ONOS), функціональний склад її елементів, протоколи та інтерфейси, що дозволяють представити мережу SDN у вигляді моделі.

Виклад основного матеріалу

Функціональна структура контролера SDN. Контролер забезпечує обслуговування трафіку відповідно до політики, яка встановлена оператором мережі. Прибравши площину управління з мережевого обладнання, контролер реалізує централізовану систему керування, спрощує автоматичне керування мережею і забезпечує інтеграцію і адміністрування бізнес-додатків [3, 27 – 29].

Архітектура SDN (рис. 2) не визначає внутрішній дизайн або реалізацію контролера SDN. Це може бути:

- один монолітний процес;
- конфедерація ідентичних процесів, організованих для розподілу навантаження або захисту один одного від збоїв;
- набір окремих функціональних компонентів, які працюють за певним набором правил;
- контролер SDN може використовувати зовнішні сервіси для виконання деяких функцій, наприклад для розрахунку шляху.

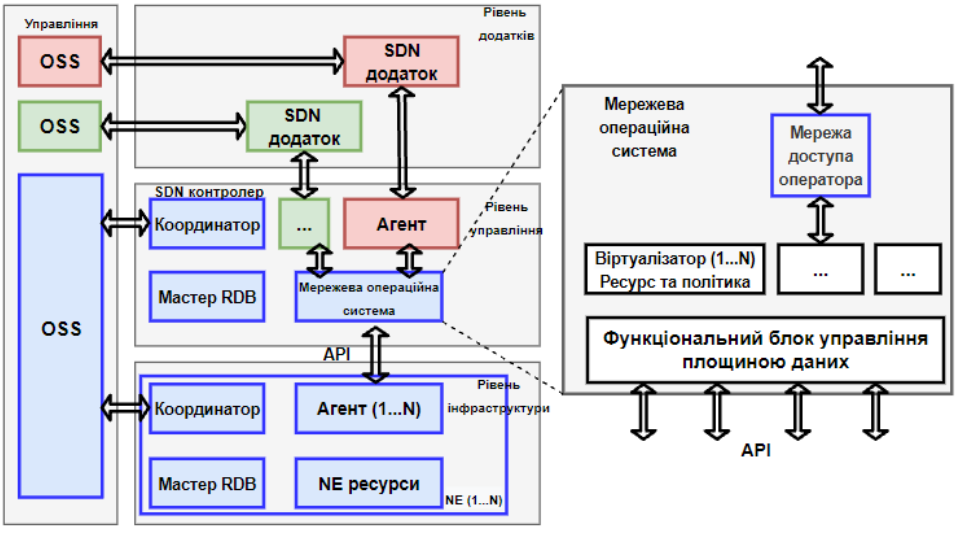


Рис. 2. Функціональні ресурси контролера

Допускається будь-яка комбінація цих альтернатив: контролер SDN розглядається як чорний ящик, який виконує певні функції. Компоненти контролера можуть розміщуватися на довільних обчислювальних платформах, включаючи обчислювальні ресурси в центрах обробки даних.

Так само як OSS управляє ресурсами і станами, контролер підпорядковується тією ж вимогою координації з будь-якими контролерами SDN, які залучені для спільного використання. Кілька компонентів менеджера або контролера можуть мати спільний доступ до мережевих ресурсів, але для виконання вимог SDN вони повинні бути:

- налаштовані на управління непересічними наборами ресурсів;
- синхронізовані один з одним, щоб ніколи не виникали суперечності або суперечливі команди.

Функціональний блок управління площиною даних ефективно володіє доступними йому підлеглими ресурсами і управляється відповідно до інструкцій OSS/координатора або віртуалізатора. Ресурси записуються у вигляді примірника інформаційної моделі. Доступ до моделі здійснюється через агента на підпорядкованому рівні. Оскільки область дій контролера SDN може охоплювати кілька мережевих елементів або навіть кілька віртуальних мереж, функціональний блок управління площиною даних повинен забезпечити скоординовану роботу. Ця функція зазвичай називається оркестрацією.

База даних ресурсів (RDB) моделює поточний екземпляр інформаційної моделі і необхідні можливості.

Координатор є функціональним компонентом, який діє від імені людини оператора. Він забезпечує надання інформації про політику обслуговування в мережі. Він передає керуючу інформацію для елементів мережі: моделі даних, площини управління і площині додатків. Тому функціональні блоки координатора розміщені повсюдно.

В архітектурі SDN віртуалізація – це виділення віртуальних абстрактних ресурсів для конкретних додатків. Контролер SDN пропонує сервіси додатків у вигляді примірника інформаційної моделі, яка абстрактно описує доступні ресурси і політику їх використання.

Віртуалізатор – це функціональний об'єкт, в якому зберігається примірник інформаційної моделі ресурсів і політики їх використання.

Віртуалізатор створюється OSS/координатором для кожного клієнтського додатку. OSS/координатор розподіляє ресурси і визначає політику, які повинен використовувати віртуалізатор в процесі надання послуг з додатками через інтерфейс API. Далі віртуалізатор створює агента, в який записується свій екземпляр інформаційної моделі доступних ресурсів і політики його використання для надання сервісу певної програми.

Будь-який протокол повинен закінчуватися функціональним об'єктом. Модель контролер-агент підходить для відносин між контрольованим і контролюючим об'єктом. Контрольований об'єкт призначається агентом, функціональним компонентом, який представляє ресурси і можливості клієнта в середовищі сервера. Агент в контролері SDN на рівні N представляє ресурси і дії, доступні з додатком на рівні N+1.

Архітектура відкритої мережевої операційної системи (ONOS). Розглянемо принципи побудови архітектури відкритої мережевої операційної системи ONOS та її особливості (рис. 3). Слід зазначити, що ONOS є основним функціональним блоком контролера SDN. Тому в науково-технічній літературі ONOS часто ототожнюють з контролером мережі SDN. Слід мати на увазі, що це не зовсім так. Контролер складається з ряду функціональних блоків, одним з яких є ONOS. Однак у зв'язку з тим, що ONOS відповідає за вирішення всіх інтелектуальних завдань управління оптичною мережею, в її архітектурі використовуються підходи, описи та рівні, які використовувалися як при описі архітектури SDN, так і при описі контролера цієї мережі.

Слід зазначити, що розробка ONOS є окремим проєктом ONF з відкритим кодом, випущеним під ліцензією Apache 2.0 [30]. Крім того, опис системи ONOS є загальнодоступним [31].

ONF вважає, що платформа ONOS розроблена для задоволення потреб операторів у створенні рішень операторського рівня, які забезпечують гнучкість створення та розгортання нових динамічних мережевих служб із спрощеними API. ONOS має забезпечувати підтримку як конфігурації, так і керування мережею в реальному часі. Такий підхід позбавить від необхідності запускати протоколи керування маршрутизацією та комутацією в мережевій інфраструктурі оператора.

Одним із можливих способів подальшого підвищення ефективності управління є перенесення ONOS у хмару. Це дозволить кінцевим користувачам створювати нові мережеві програми без необхідності змінювати площину даних.

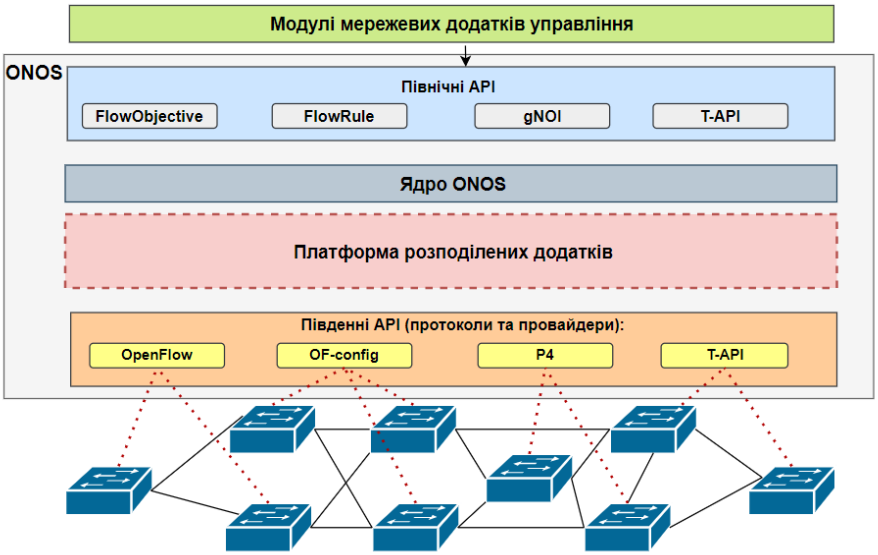


Рис. 3. Архітектура ONOS

Можливий варіант архітектури ONOS показаний на рис. 3. Видно, що архітектура ONOS є модульною за дизайном. Такий підхід дозволяє забезпечити оптимальне використання ресурсів, оскільки для кожного конкретного розгортання буде визначено оптимальну кількість необхідної підмножини модулів. Архітектура ONOS включає:

- ядро відкритої операційної системи, що відповідає за вирішення всіх інтелектуальних завдань управління мережею;

- платформу розподілених додатків, яка являє собою набір додатків для вирішення завдань управління оптичною мережею. Ця платформа побудована як розширювана модульна система. Крім того, кожен модульний функціональний блок вирішує певну задачу управління. Зі збільшенням переліку завдань управління, які необхідно вирішити, буде збільшуватися кількість модулів у платформі розподілених додатків;

- набір відкритих мережевих інтерфейсів. При цьому, як і в архітектурі SDN, пропонується використання двох типів інтерфейсів: відкритий південний інтерфейс (SBI), який є набором модулів для взаємодії з рівнем інфраструктури; відкритий північний інтерфейс (NBI), який являє собою набір модулів для взаємодії з функціональними блоками прикладного рівня.

Слід зазначити, що південний інтерфейс включає набір протоколів та інтерфейсів, таких як OpenFlow, OF-CONFIG, P4, T-API та інші. Завдання, які вирішують ці елементи системи, та їх важливість, визначаються умовами функціонування мережі. Наприклад, розглянемо роботу протоколів OF-CONFIG і OpenFlow.

OF-CONFIG (протокол керування та конфігурації OpenFlow) – це протокол для налаштування та керування робочим контекстом комутаторів OpenFlow.

Протокол OpenFlow – це протокол керування обробкою даних, що передаються через мережу передачі даних маршрутизаторами та комутаторами, який реалізує програмно-

конфігуровану мережеву технологію. Протокол OpenFlow передбачає, що комутатор OpenFlow (наприклад, комутатор Ethernet, який підтримує OpenFlow) налаштовано з різними артефактами, такими як IP-адреси контролерів OpenFlow та інших елементів мережі. Тобто комутатор попередньо налаштований.

Протокол OF-CONFIG призначений для віддаленого налаштування комутаторів OpenFlow. Він працює в повільнішому режимі порівняно з протоколом OpenFlow. Наприклад, протокол OF-CONFIG вирішує проблему створення матриць маршрутизації, яка згодом буде вирішуватися протоколом OpenFlow в реальному часі при обробці вхідних пакетів. Іншим прикладом протоколу OF-CONFIG може бути ввімкнення/вимкнення порту, що також не має відношення до обробки пакетів у реальному часі. OF-CONFIG представляє перемикач OpenFlow як абстракцію під назвою логічний перемикач OpenFlow (віртуальний перемикач). Протокол OF-CONFIG дозволяє конфігурувати параметри логічного комутатора OpenFlow так, щоб контролер OpenFlow міг спілкуватися та контролювати логічний комутатор OpenFlow через протокол OpenFlow. OF-CONFIG дозволяє динамічно пов'язувати ресурси з певними логічними перемикачами OpenFlow. Цей протокол може створити кілька віртуальних комутаторів з одного фізичного комутатора та призначити кожному певний ресурс. ONOS використовує ці протоколи для взаємодії з площиною даних, щоб реалізувати завдання керування мережею.

Останнім часом починають використовуватися P4, T-API. Вони все частіше використовуються для побудови мережевих моделей для вирішення задач керування. При цьому P4 демонструє більшу ефективність, ніж OpenFlow. Необхідно, щоб набір відкритих мережевих інтерфейсів і протоколів ONOS був універсальним для всіх сегментів і елементів мережі. Це дало б можливість будувати однорідні, масштабовані, універсальні системи з високим рівнем взаємодії всіх елементів.

Ця архітектура ONOS дозволить:

- спростити процес перевстановлення та оновлення програмного забезпечення на елементах мережі;
- знизити витрати на впровадження нових технологій та послуг всієї мережі оператора;
- поетапно збільшити кількість завдань управління, що вирішуються шляхом додавання нових функціональних блоків у складі платформи розподілених додатків;
- створити горизонтально-масштабовану систему, яка забезпечуватиме високий рівень відмовостійкості, що дуже важливо для забезпечення виконання заданих вимог до надійності функціонування централізованих систем управління.

Отже, можна сформулювати вимоги до архітектури ONOS:

По-перше, горизонтальна площина NBI повинна бути достатньо великою. Це пояснюється тим, що будь-який доступ до основного обладнання здійснюватиметься через ONOS. Таким чином, сукупність усіх північних API повинна бути достатньою для налаштування, підтримки та керування мережею.

Крім того, повинна бути багаторівнева система додатків і сервісів, що працюють на ONOS. Залежно від пріоритету завдання, яке вирішує додаток, і частоти його використання слід визначити рівень, що визначає ефективність завдання. Наприклад, додатки, які визначають головний функціональний блок керування у разі відмови елементів мережі, повинні мати пряме з'єднання з ONOS. А програми, які забезпечують встановлення нового програмного забезпечення, сертифікатів, параметрів конфігурації і в своїй роботі використовують протокол OF-CONFIG, можуть розташовуватися на рівні з високою затримкою.

Ще однією вимогою до ONOS є забезпечення одночасного двостороннього обміну інформацією через NBI та SBI. Як видно з рис. 3, програми використовують ONOS через NBI для керування мережею, а через SBI південні модулі передають інформацію про стан основної мережі до ядра ONOS.

Взаємодія між ядром ONOS і мережевими пристроями забезпечується набором протоколів і інтерфейсів, таких як OpenFlow, OF-CONFIG, P4, T-API, які забезпечують деталізацію

взаємодії з пристроями, тим самим ізолюючи ядро ONOS і додатки, що працюють на ньому, від деталей пристроїв мережі.

Ядро ONOS складається з ряду підсистем, кожна з яких відповідає за певний аспект функціонування мережі. Кожна підсистема підтримує власну абстракцію служби, яка відповідає за поширення параметрів стану мережі в кластері.

Сервіси ONOS побудовані за допомогою розподілених таблиць, які реалізовані за допомогою розподіленого сховища ключів/значень. ONOS використовує Atomix для зберігання. Це система на основі Java, яка включає: розподілену структуру даних; опис алгоритму прямого обміну повідомленнями між елементами; опис координації взаємодії, включаючи блокування небажаних функцій і обрання керівника; управління членством у групі.

Важливою особливістю Atomix є координація всіх екземплярів ONOS. Тут вирішуються дві задачі:

- кількість екземплярів ONOS, що працюють у будь-який момент часу, залежить від робочого навантаження та кількості реплікацій, необхідних для забезпечення доступності у разі збою. Примітив членства в групі Atomix використовується для визначення набору доступних екземплярів ONOS. Це дозволяє відстежувати справні та несправні екземпляри ONOS;

- основним завданням кожного екземпляра ONOS є моніторинг і підтримка підмножини фізичних комутаторів у мережі, де він обраний лідером. Визначення лідера здійснюється за допомогою функціонального блоку Atomix. Усі екземпляри ONOS можуть контролювати стан комутаторів. Однак управляти перемикачами може тільки лідер. Якщо примірник ONOS виходить з ладу, Atomix гарантує, що для перемикачів буде обрано нового лідера. Такий же підхід використовується при підключенні нового комутатора до мережі.

Сервіси та підсистеми ONOS. Хоча ONOS в значній мірі спирається на стандартні протоколи і моделі, наприклад OpenFlow, NETCONF, OpenConfig, його системна архітектура безпосередньо до них не прив'язана. ONOS модульна операційна система, вона складається із окремих підсистем (рис. 4).

Підсистема управління додатками ONOS бере на себе відповідальність за поширення артефактів додатків по кластеру. Вона забезпечує те, що всі вузли працюють з одним і тим же програмним забезпеченням. Базовий дистрибутив ONOS містить більше 175 додатків, які потрапляють в численні категорії, такі як програми для управління трафіком, драйвери пристроїв, утиліти, програми для моніторингу, готові моделі YANG.

Для того щоб взаємодіяти із зовнішнім світом, в ONOS є графічний інтерфейс та низка зовнішніх адаптерів, таких як REST API, CLI і розширюваний динамічний веб-інтерфейс. Інтерфейс gRPC з відкритим вихідним кодом використовує HTTP/2 для транспорту. Він надає такі функції, як аутентифікація, двонаправлена потокова передача, управління потоком, скасування і тайм-аути.

Додатки ONOS виступають як розширення ядра, вони можуть бути бібліотеками протоколів, драйверами або попередньо скомпільованими моделями.

Ядро ONOS складається із двох частин: мережевої та немережевої. Мережева частина підтримує виконання таких функцій: збір статистики, аналіз топології, конфігурація пристрою, створення віртуальних мереж, веб-груп, загалом підтримка служб для роботи в Інтернеті. Немережева частина виконує функції, необхідні для керування фізичною частиною SDN контролера, а саме обмін повідомленнями між контролерами, керування пам'яттю пристрою, налаштування операційної системи локально, підтримка графічного інтерфейсу для спрощеного керування системою.

Найбільш типовими є служби: хост, пристрій, посилання, топологія, майстер, кластер, конфігурація мережі, конфігурація компонентів, пакет. Дані служби використовуються багатьма додатками, оскільки вони надають інформацію про мережеві пристрої і їх топології. Однак сервісів існує набагато більше. У тому числі ті, які дозволяють додаткам програмувати поведінку мережі з використанням різних конструкцій і різних рівнів абстракції. Такі сервіси як: маршрут, Mcast, група, лічильник, правило потоку, ціль потоку, намір. Вони забезпе-

чують підтримку з'єднання по відповідним шляхам. А потім безперервно відстежують мережу, змінюючи шляхи з плином часу з метою забезпечення відповідності цілям, запропонованим наміром, в умовах мінливої мережевої ситуації.

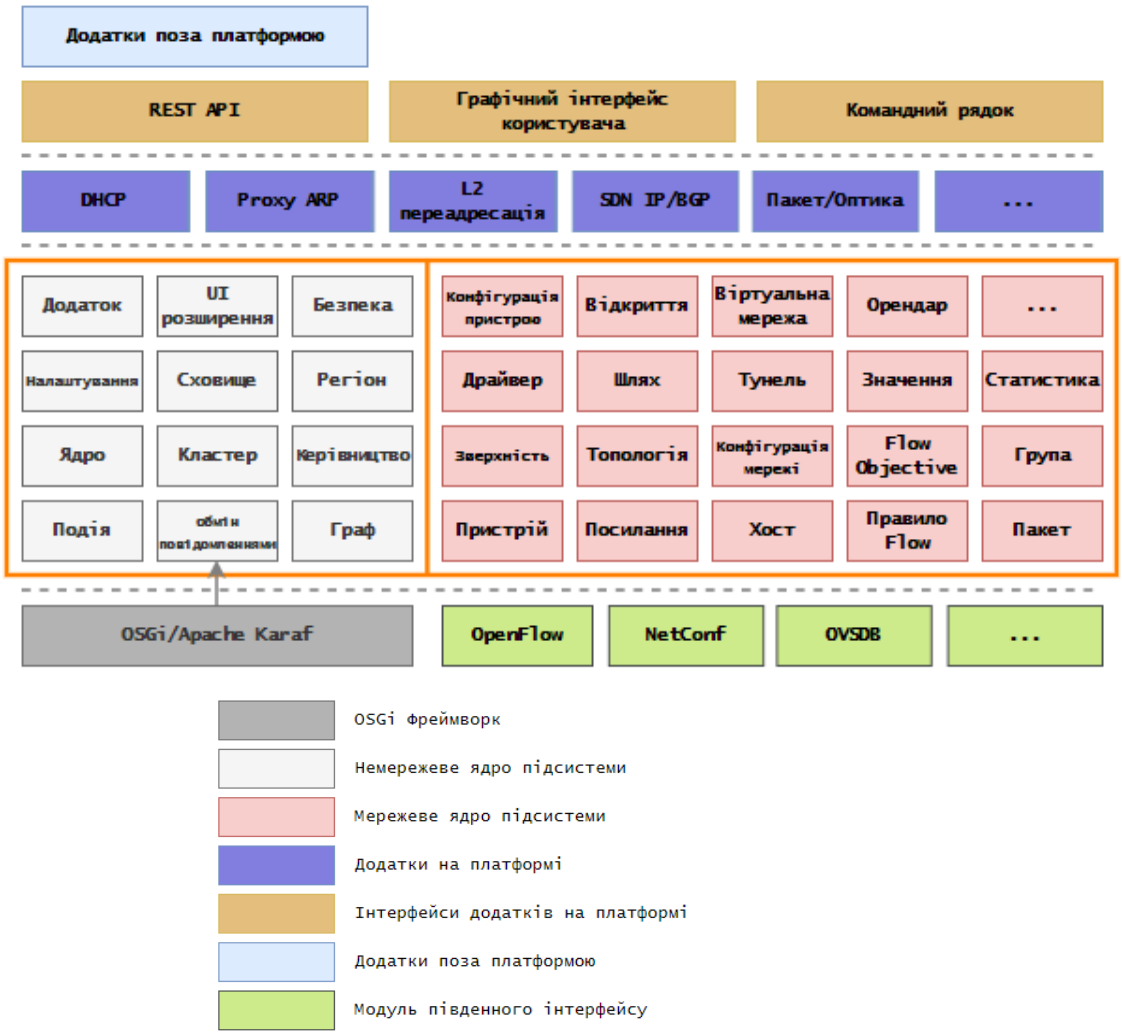


Рис. 4. Підсистеми ONOS

Кожна служба має власну розподілену базу даних і можливості обміну повідомленнями. Деякі програми можуть безкоштовно розширювати цей набір власними службами [32, 33].

Висновки

ONOS – це зручна платформа з відкритим вихідним кодом, яка дозволяє різним командам розробників спільно брати участь у проєктах з оновлення та вдосконалення системи управління. Використання ONOS дозволяє побудувати логічну централізовану площину управління в мережах SDN. Існуючий набір функціональних модулів, сервісів та інтерфейсів в ONOS дозволяє виконувати завдання управління оптичною мережею. Для подальшого розвитку ONOS необхідна розробка математичних моделей і методів оптимального вирішення задач керування в різних умовах експлуатації, які в майбутньому стануть програмними модулями прикладного рівня.

Список літератури:

1. Open Network Foundation. Accelerating the Adoption of SDN & NFV, 2021.

2. Open Network Operating System (ONOS) SDN Controller for SDN/NFV Solutions // Open Networking Foundation, 2021. [Online]. Available: https://opennetworking.org/onos.

3. What is SDN controller (software-defined networking controller)? // Definition from WhatIs.com, SearchNetworking, 2021. [Online]. Available: https://searchnetworking.techtarget.com/definition/SDN-controller-software-defined-networking-controller.

4. ONOS – Wikipedia , En.wikipedia.org, 2021. [Online]. Available: https://en.wikipedia.org/wiki/ONOS.

5. Apache License, Version 2.0 | Open Source Initiative, Opensource.org, 2021. [Online]. Available: https://opensource.org/licenses/Apache-2.0.

6. ONF TR-525 SDN Interoperability Event Technical Issues Report AppFest 2015.

7. K. Pentikousis. IETF RFC 7426. Request for Comments: 7426. ISSN: 2070-1721 EICT.

8. O. I. Romanov, M. V. Oryschuk and Y. S. Hordashnyk. Computing of influence of stimulated Raman scattering in DWDM telecommunication systems // 2016 International Conference Radio Electronics & Info Communications (UkrMiCo), 2016, pp. 1-4, doi: 10.1109/UkrMiCo.2016.7739622.

9. K. Pentikousis. ONOS. Security and Performance. Analysis // Report No. 1. September 19, 2017.

10. O. Romanov and V. Mankivskyi, Optimal Traffic Distribution Based on the Sectoral Model of Loading Network Elements // 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S&T), 2019, pp. 683-688, doi: 10.1109/PICST47496.2019.9061296.

11. O. Lemeshko and O. Yeremenko Linear optimization model of MPLS Traffic Engineering Fast ReRoute for link, node, and bandwidth protection // 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), 2018, pp. 1009-1013, doi: 10.1109/TCSET.2018.8336365.

12. O. Romanov, M. Nesterenko, L. Veres, R. Kamarali and I. Saychenko. Methods for Calculating the Performance Indicators of IP Multimedia Subsystem (IMS) // Advances in Information and Communication Technology and Systems, pp. 229-256, 2020. doi: 10.1007/978-3-030-58359-0_13.

13. O. Lemeshko, J. Papan, O. Yeremenko, M. Yevdokymenko and P. Segec. Research and Development of Delay-Sensitive Routing Tensor Model // IoT Core Networks Sensors, vol. 21, no. 11, p. 3934, 2021. doi: 10.3390/s21113934.

14. C. C. O'Connor, T. Vachuska, and B. Davie. Software-Defined Networks // A Systems Approach, 2021, p. 152.

15. K. Phemius, M. Bouet and J. Leguay. DISCO: Distributed multi-domain SDN controllers // 2014 IEEE Network Operations and Management Symposium (NOMS), 2014, pp. 1-4, doi: 10.1109/NOMS.2014.6838330.

16. J. Lam, S. Lee, H. Lee and Y. Oktian. Securing SDN Southbound and Data Plane Communication with IBC // Mobile Information Systems, vol. 2016, pp. 1-12, 2016. doi: 10.1155/2016/1708970.

17. O. I. Romanov, D. M. Fediushyna and T. T. Dong. Model And Method Of Li-Fi Network Calculation With Multipath Light Signals // 2018 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2018, pp. 1-4, doi: 10.1109/UkrMiCo43733.2018.9047550.

18. I. Obod, I. Svyd, O. Maltsev, G. Zavolodko, D. Pavlova and G. Maistrenko. Fusion the Coordinate Data of Airborne Objects in the Networks of Surveillance Radar Observation Systems // Data-Centric Business and Applications, pp. 731-746, 2020. doi: 10.1007/978-3-030-43070-2_31.

19. O. Romanov, E. Siemens, M. Nesterenko and V. Mankivskyi. Mathematical description of control problems in SDN networks // International Conference on Applied Innovations in IT (ICAIIT), pp. 33-39, 2021. Available: 10.25673/36582.

20. I. Svyd, I. Obod, O. Maltsev, T. Tkachova and G. Zavolodko. Optimal Request Signals Detection in Cooperative Surveillance Systems // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2019, pp. 1-5, doi: 10.1109/UKRCON.2019.8879840.

21. I. Obod, I. Svyd, O. Maltsev, G. Maistrenko, O. Zubkov and G. Zavolodko. Bandwidth Assessment of Cooperative Surveillance Systems. // 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT), 2019, pp. 1-6, doi: 10.1109/AIACT.2019.8847742.

22. D. Sanvito, D. Moro, M. Gullì, I. Filippini, A. Capone and A. Campanella. ONOS Intent Monitor and Reroute service: enabling plug&play routing logic // 2018 4th IEEE Conference on Network Softwarization and Workshops (NetSoft), 2018, pp. 272-276, doi: 10.1109/NETSOFT.2018.8460064.

23. D. Comer and A. Rastegarnia. Externalization of Packet Processing in Software Defined Networking // IEEE Networking Letters, vol. 1, no. 3, pp. 124-127, Sept. 2019, doi: 10.1109/LNET.2019.2918155.

24. O. Romanov, M. Nesterenko and V. Mankivskyi. The Method of Redistributing Traffic in Mobile Network // Data-Centric Business and Applications, pp. 159-182, 2021. doi: 10.1007/978-3-030-71892-3_7.

25. GitHub – OpenNetworkingFoundation/TAPI: ONF Transport API Repository (TAPI), GitHub, 2021. [Online]. Available: https://github.com/OpenNetworkingFoundation/tapi.

26. I. Svyd, I. Obod, O. Maltsev, T. Tkachova and G. Zavolodko. Improving Noise Immunity in Identification Friend or Foe Systems // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2019, pp. 73-77, doi: 10.1109/UKRCON.2019.8879812.

27. TAPI Overview – Open Transport Configuration & Control – Confluence, Wiki.opennetworking.org, 2021. [Online]. Available: https://wiki.opennetworking.org/display/OTCC/TAPI+Overview.

28. P. Littlewood, F. Masood, E. Follis. Optical transport network. Hannover : Ciena, 2014.

29. TAPI v2.1.3 Reference Implementation Agreement. TR-547. Version 1.0. July 2020, Opennetworking.org, 2018. [Online]. Available: https://opennetworking.org/wp-content/uploads/2020/08/ TR-547-TAPI-v2.1.3-Reference-Implementation-Agreement-1.pdf.

30. Open Network Operating System (ONOS) SDN Controller for SDN/NFV Solutions, Open Networking Foundation, 2021. [Online]. Available: https://opennetworking.org/onos/.

31. Software-Defined Networks: A Systems Approach – Software-Defined Networks: A Systems Approach Version 2.1-dev documentation , Sdn.systemsapproach.org, 2021. [Online]. Available: https://sdn.systemsapproach.org/index.html.

32. Chapter 6: Network OS – Software-Defined Networks: A Systems Approach Version 2.1-dev documentation, Sdn.systemsapproach.org, 2021. [Online]. Available: https://sdn.systemsapproach.org/onos.html.

33. O. Romanov, N. Korniienko, I. Obod and I. Svyd. Construction of the SDN Control Level Based on ONOS // 2021 IEEE International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo), 2021, pp. 127-132, doi: 10.1109/UkrMiCo52950.2021.9716691.

Надійшла до редколегії 12.07.2022

Відомості про авторів:

Романов Олександр Іванович – доктор технічних наук, професор, професор кафедри телекомунікацій, Інститут телекомунікаційних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Україна; email: a_i_romanov@ukr.net; ORCID: https://orcid.org/0000-0002-8683-3286

Свид Ірина Вікторівна – кандидат технічних наук, доцент, завідувач кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: iryna.svyd@nure.ua; ORCID: http://orcid.org/0000-0002-4635-6542

Корнієнко Надія Ігорівна – студентка кафедри телекомунікацій, Інститут телекомунікаційних систем Національного технічного університету України «Київський політехнічний інститут імені Ігоря Сікорського», Україна; email: nkornienko2000@ukr.net; ORCID: https://orcid.org/0000-0002-8402-8603

Романов Антон Олександрович – аспірант кафедри телекомунікаційних та радіоелектронних систем, Національний Авіаційний Університет, Україна; email: anton3329@gmail.com; ORCID: https://orcid.org/0000-0002-3425-0441

ABSTRACTS РЕФЕРАТИ РЕФЕРАТЫ

METHODS, ALGORITHMS AND TOOLS FOR CRYPTOGRAPHIC PROTECTION OF INFORMATION
МЕТОДИ, АЛГОРИТМИ ТА ЗАСОБИ КРИПТОГРАФІЧНОГО ЗАХИСТУ ІНФОРМАЦІЇ
МЕТОДЫ, АЛГОРИТМЫ И СРЕДСТВА КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ

UDC 004.056.5

Comparative characteristics of Crystals-Kyber and Skelya key encapsulation algorithms (DSTU 8961-2019) / *I.D. Gorbenko, Ye.G. Kachko, M.V. Yesina, V.A. Ponomar* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. No 210. P. 7 – 21.

In recent years, there has been a significant amount of research related to the development of quantum computers. If such a computer were to be built, it would be able to break existing public-key cryptosystems that are currently used for many purposes. This will seriously affect the privacy and integrity of digital communications, etc. That is why special attention is currently being paid to post-quantum cryptography, the main goal of which is the development of cryptographic systems that are protected from both quantum and classical attacks, and will also be able to interact with existing communication protocols and networks. In view of the significant importance of the practical application of directional encryption algorithms, at the international and state level, special attention was paid to the implementation of the proposed requirements for key encapsulation protocols. Key-establishment algorithms (KEA) form a common secret – the key for a symmetric encryption algorithm. The paper considers two KEA algorithms that use algebraic lattices: one of the finalists of the 3rd round Crystals-Kyber and the Skelya algorithm (DSTU 8961-2019). The Kyber algorithm first performs asymmetric encryption of a 32-byte message, and then generates a shared secret. The Skelya algorithm performs the same actions, but for asymmetric encryption, it uses messages of any length that do not exceed the maximum possible. That is why the last algorithm can be used not only as a KEA algorithm, but also as an asymmetric encryption algorithm. According to the NIST Security level, the Kyber algorithm provides cryptographic 1, 3, and 5 security levels, and the Rock algorithm provides cryptographic 3, 5, and 7 security levels. The cryptographic stability that is ensured for both algorithms is determined by a set of parameters. Thus, the purpose of this work is to review the details of the implementation of each of the mentioned algorithms, to compare the speed of the key generation, encapsulation and decapsulation algorithms for the Kyber and Skelya algorithms in terms of key data lengths, and the encapsulation result and computational complexity of both algorithms.

Key words: key encapsulation mechanisms; post-quantum cryptography; algebraic lattices; symmetric encryption; asymmetric encryption; Kyber; Skelya.

6 tabl. Ref: 6 items.

УДК 004.056.5

Порівняльна характеристика алгоритмів інкапсуляції ключів Crystals-Kyber та Скеля (ДСТУ 8961-2019) / *І.Д. Горбенко, О.Г. Качко, М.В. Єсіна, В.А. Пономар* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 7 – 21.

Останнім часом спостерігається значна кількість досліджень, які стосуються розробки квантових комп'ютерів. Якщо такий комп'ютер буде створено, то він зможе зламати існуючі криптосистеми з відкритими ключами, які зараз використовуються для багатьох цілей. Це серйозно вплине на конфіденційність і цілісність цифрових комунікацій, тощо. Саме тому особливу увагу наразі приділяють постквантовій криптографії, метою якої є розробка криптографічних систем, що захищені як від квантових, так і від класичних атак, а також зможуть взаємодіяти з протоколами і мережами зв'язку, що вже існують. З огляду на суттєву важливість застосування на практиці алгоритмів направленого шифрування, на міжнародному та державному рівнях особливу увагу було приділено впровадженню висунутих вимог до протоколів інкапсуляції ключів. Алгоритми інкапсуляції ключів (Key-establishment Algorithms, KEA) формують загальний секрет – ключ для симетричного алгоритму шифрування. В роботі розглянуто два KEA алгоритми, які застосовують алгебраїчні решітки: один з фіналістів 3-го раунду Crystals-Kyber та алгоритм Скеля (ДСТУ 8961-2019). Алгоритм Kyber спочатку виконує несиметричне шифрування повідомлення завдовжки 32 байти, а потім виконується формування загального секрету. Алгоритм Скеля виконує ті ж дії, але для несиметричного шифрування застосовує повідомлення будь-якої довжини, яка не перевищує максимально можливої. Ось чому останній алгоритм можна застосовувати не тільки як KEA алгоритм, а і як алгоритм несиметричного шифрування. Згідно з NIST Security level алгоритм Kyber забезпечує криптографічну стійкість 1, 3 та 5 рівнів, а алгоритм Скеля забезпечує криптографічну стійкість 3, 5 та 7 рівнів. Криптографічна стійкість, яка забезпечується, для обох алгоритмів визначається набором параметрів. Таким чином, метою цієї роботи є огляд деталей реалізації кожного з алгоритмів, виконано порівняння швидкодії алгоритмів генерації ключів, інкапсуляції та декапсуляції для алгоритмів Kyber та Скеля з боку довжин ключових даних, і результату інкапсуляції та обчислювальної складності обох алгоритмів.

Ключові слова: алгоритми інкапсуляції ключів; постквантова криптографія; алгебраїчні решітки; симетричне шифрування; асиметричне шифрування; Kyber; Скеля.

Табл. 6. Бібліогр.: 6 назв.

УДК 004.056.5

Сравнительная характеристика алгоритмов инкапсуляции ключей Crystals-Kyber и Скеля (ДСТУ 8961-2019) / *И.Д. Горбенко, Е.Г. Качко, М.В. Есина, В.А. Пономарь* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 7 – 21.

В последнее время наблюдается значительное количество исследований, касающихся разработки квантовых компьютеров. Если компьютер будет создан, то он сможет сломать существующие криптосистемы с открытыми ключами, которые сейчас используются для многих целей. Это серьезно повлияет на конфиденциальность и целостность цифровых коммуникаций. Именно поэтому особое внимание уделяется постквантовой криптографии, целью которой является разработка криптографических систем, защищенных как от квантовых, так и от классических атак, которые смогут взаимодействовать с уже существующими протоколами и сетями связи. Учитывая важность применения на практике алгоритмов направленного шифрования, на международном и государственном уровне особое внимание было уделено внедрению предъявляемых требований к протоколам инкапсуляции ключей. Алгоритмы инкапсуляции ключей (Key-establishment Algorithms, KEA) формируют общий секрет – ключ для симметричного алгоритма шифрования. В работе рассмотрены два алгоритма KEA, которые применяют алгебраические решетки: один из финалистов 3-го раунда Crystals-Kyber и алгоритм Скеля (ДСТУ 8961-2019). Алгоритм Kyber сначала выполняет несимметричное шифрование сообщения длиной 32 байта. Далее выполняется формирование общего секрета. Алгоритм Скала выполняет те же действия, но для несимметричного шифрования применяет сообщения любой длины, не превышающей максимально возможной. Вот почему последний алгоритм можно применять не только как алгоритм KEA, но и как алгоритм несимметричного шифрования. Согласно NIST Security level, алгоритм Kyber обеспечивает криптографическую устойчивость 1, 3 и 5 уровней, а алгоритм Скеля обеспечивает криптографическую устойчивость 3, 5 и 7 уровней. Обеспечиваемая криптографическая стойкость для обоих алгоритмов определяется набором параметров. Таким образом, целью этой работы является обзор деталей реализации каждого из алгоритмов, выполнение сравнение быстродействия алгоритмов генерации ключей, инкапсуляции и декапсуляции для алгоритмов Kyber и Скеля со стороны длин ключевых данных, и результата инкапсуляции и вычислительной сложности обоих алгоритмов.

Ключевые слова: алгоритмы инкапсуляции ключей; постквантовая криптография; алгебраические решетки; симметрическое шифрование; асимметрическое шифрование; Kyber; Скеля.

Табл. 6. Библиогр.: 6 назв.

UDC 004.056.5

Comparison of security arguments of promising key encapsulation mechanisms / *Yu.I. Gorbenko, S.O. Kandii* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 22 – 36.

The study of key encapsulation mechanisms on algebraic lattices is one of the important directions in modern post-quantum cryptography, since many mechanisms are already either standardized (ANSI X.9.98, DSTU 8961:2019 "Skelya") or are promising candidates for standardization (CRYSTALS-Kyber, FrodoKEM). The purpose of this work is to compare the security arguments of DSTU 8961:2019 "Skelya", CRYSTALS-Kyber, FrodoKEM key encapsulation mechanisms. The paper provides a comparison of theoretical evidence in the idealized random oracle (ROM) and quantum random oracle (QROM) models, as well as a comparison of specific values of security parameters in the core-SVP model, which is, in fact, a standard for lattice cryptography. Since all three key encapsulation mechanisms are based on different complex problems (NTRU, Module-LWE, LWE), a comparison of complex lattice theory problems and a comparison of their security arguments are additionally given. The strengths and weaknesses of the considered key encapsulation mechanisms are shown, and areas of research that require more detailed attention are highlighted.

Key words: post-quantum cryptography; algebraic lattices; DSTU 8961:2019 "Skelya", CRYSTALS-Kyber; FrodoKEM; ROM, QROM; core-SVP.

4 tabl. 3 fig. Ref: 35 items.

УДК 004.056.5

Порівняння аргументів безпеки перспективних механізмів інкапсуляції ключів / *Ю.І. Горбенко, С.О. Кандій* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 22 – 36.

Дослідження механізмів інкапсуляції ключів на алгебраїчних решітках є одним з важливих напрямів у сучасній постквантовій криптографії, оскільки багато механізмів вже або стандартизовані (ANSI X.9.98, ДСТУ 8961:2019 “Скеля”), або є перспективними кандидатами на стандартизацію (CRYSTALS-Kyber, FrodoKEM). Метою цієї роботи є порівняння аргументів безпеки механізмів інкапсуляції ключів ДСТУ 8961:2019 “Скеля”, CRYSTALS-Kyber, FrodoKEM. В роботі наведено порівняння як теоретичних доказів у ідеалізованих моделях випадкового оракула (ROM) та квантового випадкового оракула (QROM), так і порівняння конкретних значень параметрів безпеки у моделі core-SVP, яка, фактично, є стандартом для криптографії на решітках. Оскільки всі три механізми інкапсуляції ключів грунтуються на різних складних проблемах (NTRU,Module-LWE, LWE), то додатково наведено порівняння складних проблем з теорії решіток та порівняння аргументів їх безпеки. Показані сильні та слабкі сторони розглянутих механізмів інкапсуляції ключів та виділені напрямки досліджень, які потребують детальнішої уваги.

Ключові слова: постквантова криптографія; алгебраїчні решітки; ДСТУ 8961:2019 “Скеля”; CRYSTALS-Kyber; FrodoKEM; ROM, QROM; core-SVP.

Табл. 4. Іл. 3. Бібліогр.: 35 назв.

УДК 004.056.5

Сравнение аргументов безопасности перспективных механизмов инкапсуляции ключей / *Ю.И. Горбенко, С.О. Кандий* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 22 – 36.

Исследование механизмов инкапсуляции ключей на алгебраических решетках является одним из важных направлений в современной постквантовой криптографии, поскольку многие механизмы уже либо стандартизированы (ANSI X.9.98, ДСТУ 8961:2019 “Скеля”), либо являются перспективными кандидатами на стандартизацию (CRYSTALS-Kyber, FrodoKEM). Целью этой работы является сравнение аргументов безопасности механизмов инкапсуляции ключей ДСТУ 8961:2019 “Скеля”, CRYSTALS-Kyber, FrodoKEM. В работе представлены сравнения как теоретических доказательств в идеализированных моделях случайного оракула (ROM) и квантового случайного оракула (QROM), так и сравнение конкретных значений параметров безопасности у модели core-SVP, которая фактически является стандартом для криптографии на решетках. Поскольку все три механизма инкапсуляции ключей основываются на разных сложных проблемах (NTRU, Module-LWE, LWE), то дополнительно приведено сравнение сложных проблем из теории решеток и сравнение аргументов их безопасности. Показаны сильные и слабые стороны рассматриваемых механизмов инкапсуляции ключей и выделены направления исследований, требующие более подробного внимания.

Ключевые слова: постквантовая криптография; алгебраические решетки; ДСТУ 8961:2019 “Скеля”; CRYSTALS-Kyber; FrodoKEM; ROM, QROM; core-SVP.

Табл. 4. Ил. 3. Библиогр.: 35 назв.

UDC 004.056.5

FALCON signature vulnerability to special attacks and its protection / *Ya.A. Derevianko, I.D. Gorbenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. No 210. P. 37 – 52.

It is well known that quantum algorithms offer exponential speedup in solving the integer factorization and discrete logarithm problems that existing public-key systems rely on. Thus, post-quantum cryptography seeks alternative classical algorithms that can withstand quantum cryptanalysis. Growing concern about the quantum threat has prompted the National Institute of Standards and Technology (NIST) to invite and evaluate applications for a post-quantum cryptography standard, an ongoing process scheduled to be completed by 2023.

Falcon is an electronic signature algorithm based on the mathematics of algebraic lattices. The disadvantage of this algorithm is the small number of studies of resistance against special attacks, as well as attacks through side channels.

This material examines existing attacks on the implementation, and also analyzes the speed with applying countermeasures that would prevent such attacks. Although the Falcon scheme sampler, as well as certain mathematical transformations, are still vulnerable to attacks (which in turn allow the private key to be obtained), the efficiency of the components and mathematics of this signature algorithm make it competitive with other schemes, even with countermeasures against these attacks.

The work will also consider the attack by side channels on the Falcon. Such an attack is a known-plaintext attack that uses the device's electromagnetic radiation to derive secret signature keys, which can then be used to forge signatures in arbitrary messages. The obtained results show that Falcon is quite vulnerable to side-channel attacks and does not yet have protection against such attacks in the proposed implementation. Because of this, standardization or implementation should consider the possibility of physical attacks, as well as options for countering such attacks.

Key words: post-quantum cryptography; electronic signature algorithm; Falcon algorithm; NIST.

2 tabl. 7 fig. Ref: 16 items.

УДК 004.056.5

Вразливість ЕП FALCON до спеціальних атак та його захищеність / *Я.А. Дерев’янко, І.Д. Горбенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 37 – 52.

Добре відомо, що квантові алгоритми пропонують експоненціальне прискорення при розв’язанні завдань цілочисельної факторизації та (еліптичної кривої) дискретного логарифму, на які покладаються існуючі системи з відкритим ключем. Таким чином, постквантова криптографія шукає альтернативні класичні алгоритми, які можуть протистояти квантовому криптоаналізу. Зростаюче занепокоєння квантовою загрозою спонукало Національний інститут стандартів і технологій (NIST) запросити та оцінити заявки на стандарт постквантової криптографії, який є постійним процесом, який планується завершити до 2023 року.

Falcon – алгоритм електронного підпису, що заснований на математиці алгебраїчних решіток. Мінусом даного алгоритму є мала кількість досліджень стійкості проти спеціальних атак, а також атак побічними каналами.

У даному матеріалі розглядаються існуючі атаки на реалізацію, а також відбувається аналіз швидкодії при застосуванні контрзаходів, які б таким атакам перешкоджали. Незважаючи на те, що відбірник схеми Falcon, а також певні математичні перетворення, все ж є вразливими до атак (що в свою чергу дозволяє отримати приватний ключ), ефективність компонентів та математики даного алгоритму електронного підпису сприяє тому, що він здатен конкурувати з іншими схемами, навіть з контрзаходами проти цих атак.

Розглянуто атаку побічними каналами на Falcon. Така атака є атакою з відомим відкритим текстом, яка використовує електромагнітні вимірювання пристрою для отримання секретних ключів підпису, які потім можна використовувати для підробки підписів у довільних повідомленнях. Результати показують, що Falcon є до-

волі вразливим до атак побічними каналами та поки не має захисту від таких атак у запропонованій реалізації. Через це при стандартизації чи впровадженні слід розглядати можливість фізичних атак, а також варіанти протидії таким атакам.

Ключові слова: постквантова криптографія; алгоритм електронного підпису; алгоритм Falcon; NIST.

Табл. 2. Іл. 7. Бібліогр.: 16 назв.

УДК 004.056.5

Уязвимость ЭП FALCON к специальным атакам и его защищенность / *Я.А. Деревянко, И.Д. Горбенко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 37 – 52.

Хорошо известно, что квантовые алгоритмы предлагают экспоненциальное ускорение при решении задач целочисленной факторизации и (эллиптической кривой) дискретного логарифма, на которых основаны существующие системы с открытым ключом. Таким образом, постквантовая криптография ищет альтернативные классические алгоритмы, которые могут противостоять квантовому криптоанализу. Растущая обеспокоенность квантовой угрозой побудила Национальный институт стандартов и технологий (NIST) пригласить и оценить заявки на стандарт постквантовой криптографии, являющийся постоянным процессом, который планируется завершить к 2023 году.

Falcon – алгоритм электронной подписи, основанный на математике алгебраических решеток. Минусом данного алгоритма является малое количество исследований устойчивости против специальных атак, а также атак побочными каналами.

В данном материале рассматриваются существующие атаки на реализацию, а также происходит анализ быстродействия при применении контрмер, которые бы таким атакам препятствовали. Несмотря на то, что отборник схемы Falcon, а также определенные математические преобразования, все же уязвимы к атакам (что в свою очередь позволяет получить приватный ключ), эффективность компонентов и математики данного алгоритма электронной подписи способствует тому, что он способен конкурировать с другими схемами, даже с контрмерами против этих атак.

Рассмотрена атака побочными каналами на Falcon. Такая атака представляет собой атаку с известным открытым текстом, которая использует электромагнитные излучения устройства для получения секретных ключей подписи, которые затем можно использовать для подделки подписей в произвольных сообщениях. Результаты показывают, что Falcon довольно уязвим к атакам побочными каналами и пока не имеет защиты от таких атак в предложенной реализации. Поэтому при стандартизации или внедрении следует рассматривать возможность физических атак, а также варианты противодействия таким атакам.

Ключевые слова: постквантовая криптография; алгоритм электронной подписи; алгоритм Falcon; NIST.

Табл. 2. Ил. 7. Библиогр.: 16 назв.

UDC 004.056: 004.056.5

Researching basic searchable encryption schemes in databases that support SQL / *V.I. Yesin, V.V. Vilihura* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. No 210. P. 53 – 74.

Currently, many users prefer to outsource data to third-party cloud servers in order to mitigate the load of local storage. However, storing sensitive data on remote servers creates security challenges and is a source of concern for data owners. With ever-growing security and privacy concerns, it is becoming increasingly important to encrypt data stored remotely. However, the use of traditional encryption prevents the search operation in the encrypted data. One approach to solving this problem is searchable encryption. Solutions for search in secure databases cover a wide range of cryptographic techniques, although there is still no dominant solution. Designing secure search systems is a balance between security, functionality, performance, and usability. Therefore, this paper provides an overview of some of the important current secure search solutions. The main searchable encryption systems of databases that support SQL are considered. The strengths and weaknesses of the analyzed systems and the techniques implemented in them are highlighted. A comparative analysis of some characteristics of the compared systems is given. Attention is drawn to the fact that the ability to perform search operations in encrypted data leads to a complication of systems, an increase in the amount of required memory and query execution time. All this indicates the openness of the protected search problem and the need for further research in this direction to ensure secure work with remote databases and data warehouses.

Key words: database; data warehouse; confidentiality; encryption; searchable encryption.

7 tab. 9 fig. Ref: 33 items.

УДК 004.056: 004.056.5

Дослідження основних схем шифрування з можливістю пошуку у базах даних, які підтримують SQL / *В.І. Єсін, В.В. Вілігура* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 53 – 74.

Багато користувачів вважають за краще передавати дані на сторонні хмарні сервери, щоб знизити навантаження на локальне сховище. Однак збереження конфіденційних даних на віддалених серверах створює проблеми з безпекою та є джерелом занепокоєння для власників даних. У зв'язку з проблемами безпеки і приватності, що постійно зростають, стає все більш актуальним шифрування даних, що зберігаються віддалено. Однак використання традиційного шифрування перешкоджає виконанню операції пошуку зашифрованих даних. Одним із підходів до вирішення цієї проблеми є шифрування з можливістю пошуку. Рішення для пошуку в захищених базах даних охоплюють широкий спектр криптографічних методів, хоча домінуючого рішення досі не

існує. Проєктування захищених пошукових систем – це баланс між безпекою, функціональністю, продуктивністю та зручністю використання. Тому в роботі представлено огляд деяких важливих поточних рішень захищеного пошуку. Розглядаються основні системи шифрування з можливістю пошуку в базах даних, які підтримують мову структурних запитів SQL. Виділяються слабкі та сильні сторони аналізованих систем та реалізованих у них методів. Наводиться порівняльний аналіз деяких характеристик систем, що порівнюються. Звертається увага на той факт, що можливість здійснювати операції пошуку у зашифрованих даних призводять до ускладнення систем, збільшення обсягів необхідної пам'яті та часу виконання запитів. Все це свідчить про відкритість проблеми захищеного пошуку та необхідність проведення подальших досліджень у даному напрямку для забезпечення безпечної роботи з віддаленими базами, сховищами даних.

Ключові слова: база даних; сховище даних; конфіденційність; шифрування; шифрування з можливістю пошуку.

Табл. 7. Іл. 9. Бібліогр.: 33 назв.

УДК 004.056: 004.056.5

Исследование основных схем шифрования с возможностью поиска в базах данных, поддерживающих SQL / *В.И. Есин, В.В. Вилигура* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 53 – 74.

Многие пользователи предпочитают передавать данные на сторонние облачные серверы, чтобы снизить нагрузку на локальное хранилище. Однако хранение конфиденциальных данных на удаленных серверах создает проблемы с безопасностью и является источником беспокойства для владельцев данных. В связи с постоянно растущими проблемами безопасности и приватности становится все более актуальным шифрование данных, хранящихся удаленно. Однако использование традиционного шифрования препятствует выполнению операции поиска в зашифрованных данных. Одним из подходов к решению данной проблемы является шифрование с возможностью поиска. Решения для поиска в защищенных базах данных охватывают широкий спектр криптографических методов, хотя доминирующего решения до сих пор не существует. Проектирование защищенных поисковых систем – это баланс между безопасностью, функциональностью, производительностью и удобством использования. Поэтому в работе представлен обзор некоторых важных текущих решений защищенного поиска. Рассматриваются основные системы шифрования с возможностью поиска в базах данных, поддерживающих язык структурных запросов SQL. Выделяются слабые и сильные стороны анализируемых систем и реализованных в них методов. Приводится сравнительный анализ некоторых характеристик сравниваемых систем. Обращается внимание на тот факт, что возможность осуществлять операции поиска в зашифрованных данных приводят к усложнению систем, увеличению объемов необходимой памяти и времени выполнения запросов. Все это свидетельствует об открытости проблемы защищенного поиска и необходимости проведения дальнейших исследований в данном направлении для обеспечения безопасной работы с удаленными базами, хранилищами данных.

Ключевые слова: база данных; хранилище данных; конфиденциальность; шифрование; шифрование с возможностью поиска.

Табл. 7. Ил. 9. Библиогр.: 33 назв.

UDC 004.056.5

Status report on the third round of the NIST post-quantum cryptography standardization process / *M.V. Yesina, Ye.V. Ostrianska, I.D. Gorbenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. No 210. P. 75 – 86.

In recent years, there has been steady progress in the creation of quantum computers. If large-scale quantum computers are implemented, they will threaten the security of many widely used public-key cryptosystems. Key-establishment schemes and digital signatures based on factorization, discrete logarithms, and elliptic curve cryptography will be most affected. Symmetric cryptographic primitives such as block ciphers and hash functions will be broken only slightly. As a result, there has been an intensification of research on finding public-key cryptosystems that would be secure against cryptanalysts with both quantum and classical computers. This area is often called post-quantum cryptography (PQC), or sometimes quantum-resistant cryptography. The goal is to design schemes that can be deployed in existing communication networks and protocols without significant changes. The National Institute of Standards and Technology is in the process of selecting one or more public-key cryptographic algorithms through an open competition. New public-key cryptography standards will define one or more additional digital signatures, public-key encryption, and key-establishment algorithms. It is assumed that these algorithms will be able to protect confidential information well in the near future, including after the advent of quantum computers. After three rounds of evaluation and analysis, NIST has selected the first algorithms that will be standardized as a result of the PQC standardization process. The purpose of this article is to review and analyze the state of NIST's post-quantum cryptography standardization evaluation and selection process. The article summarizes each of the 15 candidate algorithms from the third round and identifies the algorithms selected for standardization, as well as those that will continue to be evaluated in the fourth round of analysis. Although the third round is coming to an end and NIST will begin developing the first PQC standards, standardization efforts in this area will continue for some time. This should not be interpreted as meaning that users should wait to adopt post-quantum algorithms. NIST looks forward to the rapid implementation of these first standardized algorithms and will issue future guidance on the transition. The transition will undoubtedly have many complexities, and there will be challenges for some use cases such as IoT devices or certificate transparency.

Key words: post-quantum cryptography; standardization; NIST; electronic signature; key transport.
4 tabl. 4 fig. Ref: 27 items.
УДК 004.056.5

Стан третього раунду процесу стандартизації постквантової криптографії NIST / *М.В. Єсіна, Є.В. Острянська, І.Д. Горбенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 75 – 86.

В останні роки спостерігається стійкий прогрес у створенні квантових комп'ютерів. У разі реалізації великомасштабних квантових комп'ютерів вони будуть загрожувати безпеці багатьох широко використовуваних криптосистем з відкритим ключем. Схеми встановлення ключів і цифрові підписи, що засновані на факторизації, дискретних логарифмах і криптографії на еліптичних кривих, найбільш сильно постраждають. Симетричні криптографічні примітиви, такі як блокові шифри і геш-функції, будуть порушені лише незначно. Як результат, було проведено активізацію досліджень щодо пошуку криптосистем на відкритих ключах, які були б захищені від криптоаналітиків як з квантовими, так і з класичними комп'ютерами. Цю область часто називають постквантовою криптографією (PQC), або іноді квантово-стійкою криптографією. Мета полягає в розробці схем, які можна розгорнути в існуючих комунікаційних мережах та протоколах без суттєвих змін. Національний інститут стандартів і технологій знаходиться в процесі вибору одного або декількох криптографічних алгоритмів з відкритим ключем за допомогою відкритого конкурсу. Нові стандарти криптографії з відкритим ключем визначатимуть одну або кілька додаткових цифрових підписів, шифрування з відкритим ключем і алгоритми встановлення ключів. Передбачається, що ці алгоритми будуть здатні добре захищати конфіденційну інформацію в недалекому майбутньому, в тому числі після появи квантових комп'ютерів. Після трьох раундів оцінки та аналізу, NIST вибрав перші алгоритми, які він стандартизує в результаті процесу стандартизації PQC. Метою цієї статті є огляд та аналіз стану оцінювання та відбору процесу стандартизації постквантової криптографії NIST. У статті узагальнено кожен із 15 алгоритмів-кандидатів третього раунду та визначено алгоритми, обрані для стандартизації, а також ті, які продовжуватимуть оцінюватись у четвертому раунді аналізу. Незважаючи на те, що третій раунд завершується і NIST почне розробляти перші стандарти PQC, зусилля зі стандартизації в цій галузі триватимуть ще деякий час. NIST сподівається на швидке впровадження цих перших стандартизованих алгоритмів і видасть майбутні вказівки щодо переходу. Перехід, безсумнівно, матиме багато складнощів, і виникнуть проблеми для деяких випадків використання, таких як пристрої IoT або прозорість сертифікатів.

Ключові слова: постквантова криптографія; стандартизація; NIST; електронний підпис; транспортування ключів.
Табл. 4. Іл. 4. Бібліогр.: 27 назв.
УДК 004.056.5

Состояние третьего раунда процесса стандартизации постквантовой криптографии NIST / *М.В. Есина, Е.В. Острянская, И.Д. Горбенко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 75 – 86.

В последние годы наблюдается устойчивый прогресс в разработке квантовых компьютеров. В случае реализации крупномасштабных квантовых компьютеров, они будут угрожать безопасности многих широко используемых криптосистем с открытым ключом. Схемы установки ключей и цифровые подписи, основанные на факторизации, дискретных логарифмах и криптографии на эллиптических кривых, пострадают наиболее сильно. Симметричные криптографические примитивы, такие как блочные шифры и геш-функции, будут нарушены лишь незначительно. Как результат, была проведена активизация исследований поиска криптосистем на открытых ключах, которые были бы защищены от криптоаналитиков как с квантовыми, так и с классическими компьютерами. Эту область часто называют постквантовой криптографией (PQC) или иногда – квантово-стойкой криптографией. Цель состоит в разработке схем, которые можно развернуть в существующих коммуникационных сетях и протоколах, без существенных изменений. Национальный институт стандартов и технологий находится в процессе выбора одного или нескольких криптографических алгоритмов с открытым ключом посредством открытого конкурса. Новые стандарты криптографии с открытым ключом будут определять одну или несколько дополнительных цифровых подписей, шифрование с открытым ключом и алгоритмы установки ключей. Предполагается, что эти алгоритмы будут способны хорошо защищать конфиденциальную информацию в скором будущем, в том числе после появления квантовых компьютеров. После трех раундов оценки и анализа NIST выбрал первые алгоритмы, которые он стандартизирует в результате процесса стандартизации PQC. Цель работы – обзор и анализ состояния оценки и отбора стандартизации постквантовой криптографии NIST. В статье обобщены каждый из 15 алгоритмов-кандидатов третьего раунда и определены алгоритмы, выбранные для стандартизации, а также те, которые будут продолжать оцениваться в четвертом раунде анализа. Несмотря на то, что третий раунд завершается и NIST начнет разрабатывать первые стандарты PQC, усилия по стандартизации в этой области продлятся еще некоторое время. NIST надеется на скорейшее внедрение этих первых стандартизированных алгоритмов и выдаст будущие указания по переходу. Переход, несомненно, будет иметь много сложностей, и возникнут проблемы для некоторых случаев использования, таких как IoT устройства или прозрачность сертификатов.

Ключевые слова: постквантовая криптография; стандартизация; NIST; электронная подпись; транспортировка ключей.
Табл. 4. Ил. 4. Библиогр.: 27 назв.

UDC 004.056.5

Analysis of views of the European Union on quantum-post-quantum limitations / *Ye.V. Ostrianska, M.V. Yesina, I.D. Gorbenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 87 – 98.

Virtually all asymmetric cryptographic schemes currently in use are threatened by the potential development of powerful quantum computers. Although there is currently no definite answer and it is very unclear when or even if CRQC will ever be built and the gap between modern quantum computers and the envisioned CRQC is huge, the risk of creating CRQC means that currently deployed public key cryptography must be replaced by quantum-resistant ones alternatives. For example, information encrypted using modern public key cryptography can be recorded by cryptanalysts and then attacked if a QRQC can be created. The potential harm that CRQC could cause is the basis of the motivation to seek countermeasures, even though we have uncertainties about when and if these computers can be built. Deployed systems that use public key cryptography can also take years to update. Post-quantum cryptography is one way to combat quantum computer threats. Its security is based on the complexity of mathematical problems that are currently considered unsolvable efficiently – even with the help of quantum computers. Post-quantum cryptography deals with the development and research of asymmetric cryptosystems, which, according to current knowledge, cannot be broken even by powerful quantum computers. These methods are based on mathematical problems for the solution of which neither efficient classical algorithms nor efficient quantum algorithms are known today. Various approaches to the implementation of post-quantum cryptography are used in modern research, including: code-based cryptography, lattice-based cryptography, hashing-based cryptography, isogeny-based cryptography, and multidimensional cryptography. The purpose of this work is to review the computational model of quantum computers; quantum algorithms, which have the greatest impact on modern cryptography; the risk of creating cryptographically relevant quantum computers (CRQC); security of symmetric cryptography and public key cryptography in the presence of CRQC; NIST PQC standardization efforts; transition to quantum-resistant public-key cryptography; relevance, views and current state of development of quantum-resistant cryptography in the European Union. It also highlights the progress of the most important effort in the field: NIST's standardization of post-quantum cryptography.

Key words: post-quantum cryptography; quantum computer; standardization; electronic signature; key transport.

2 tabl. 2 fig. Ref: 31 items.

УДК 004.056.5

Аналіз поглядів Європейського союзу на квантово-постквантові обмеження / *Є.В. Острянська, М.В. Єсіна, І.Д. Горбенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 87 – 98.

Практично всім асиметричним криптографічним схемам, які зараз використовуються, загрожує потенційна розробка потужних квантових комп'ютерів. Хоча наразі немає однозначної відповіді та дуже незрозуміло, коли і навіть якщо CRQC коли-небудь буде побудовано та розрив між сучасними квантовими комп'ютерами та передбачуваними CRQC величезний, однак ризик створення CRQC означає, що наразі розгорнуту криптографію з відкритим ключем необхідно замінити квантово-стійкими альтернативами. Наприклад, інформація, зашифрована за допомогою сучасної криптографії з відкритим ключем, може бути записана криптоаналітиками, а потім піддана атаці, якщо можна створити QRQC. Потенційна шкода, яку може завдати CRQC, є основою мотивації шукати контрзаходи, навіть якщо у нас є невизначеності щодо того, коли та чи можна створити ці комп'ютери. Оновлення розгорнутих систем, які використовують криптографію з відкритим ключем, також може тривати багато років. Постквантова криптографія є одним із способів боротьби з загрозами квантового комп'ютера. Її безпека базується на складності математичних проблем, які наразі вважаються нерозв'язними ефективно – навіть за допомогою квантових комп'ютерів. Постквантова криптографія займається розробкою та дослідженням асиметричних криптосистем, які, згідно з сучасними знаннями, не можуть бути зламані навіть потужними квантовими комп'ютерами. Ці методи базуються на математичних задачах, для розв'язання яких на сьогодні невідомі ані ефективні класичні алгоритми, ані ефективні квантові алгоритми. У сучасних дослідженнях застосовуються різні підходи до реалізації постквантової криптографії, серед них: криптографія на основі кодів, криптографія на основі решітки, криптографія на основі гешування, криптографія на основі ізогеній та багатовимірна криптографія. Метою цієї роботи є огляд обчислювальної моделі квантових комп'ютерів; квантових алгоритмів, які найбільше впливають на сучасну криптографію; ризику створення криптографічно-релевантних квантових комп'ютерів (CRQC); безпеки симетричної криптографії та криптографії з відкритим ключем за наявності CRQC; зусилля зі стандартизації NIST PQC; перехід до квантово-стійкої криптографії з відкритим ключем; аналіз актуальності, поглядів та поточного стану розвитку квантово-стійкої криптографії у Європейському Союзі. Також висвітлюється хід найважливіших зусиль у цій галузі: стандартизації постквантової криптографії NIST.

Ключові слова: постквантова криптографія; квантовий комп'ютер; стандартизація; електронний підпис; транспортування ключів.

Табл. 2. Іл. 2. Бібліогр.: 31 назв.

УДК 004.056.5

Анализ взглядов Европейского союза на квантово-постквантовые ограничения / *Е.В. Острянская, М.В. Есина, И.Д. Горбенко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 87 – 98.

Практически всем используемым сейчас асимметричным криптографическим схемам грозит потенциальная разработка мощных квантовых компьютеров. Пока нет однозначного ответа и очень непонятно когда, и даже если CRQC когда-нибудь будет построен и разрыв между современными квантовыми компьютерами и

предполагаемыми CRQC огромен. Однако риск создания CRQC означает, что развернутую криптографию с открытым ключом необходимо заменить квантово-устойчивыми альтернативами. Например, информация, зашифрованная с помощью современной криптографии с открытым ключом, может быть записана криптоаналитиками, а затем подвергнута атаке, если можно создать QRQC. Потенциальный вред, который может нанести CRQC, является основой мотивации искать контрмеры, даже если есть неопределенность относительно того, когда и можно ли вообще создать эти компьютеры. Обновление развернутых систем, использующих криптографию с открытым ключом, также может занять много лет. Постквантовая криптография – один из способов борьбы с угрозами квантового компьютера. Ее безопасность базируется на сложности математических проблем, которые сейчас считаются неразрешимыми эффективно – даже с помощью квантовых компьютеров. Постквантовая криптография занимается разработкой и исследованием асимметричных криптосистем, которые, согласно современным знаниям, не могут быть сломаны даже мощными квантовыми компьютерами. Эти методы базируются на математических задачах, для решения которых неизвестны ни эффективные классические алгоритмы, ни эффективные квантовые алгоритмы. В современных исследованиях применяются различные подходы к реализации постквантовой криптографии, в том числе: криптография на основе кодов, криптография на основе решетки, криптография на основе хеширования, криптография на основе изогений и многомерная криптография. Цель работы – обзор вычислительной модели квантовых компьютеров; квантовых алгоритмов, наиболее влияющих на современную криптографию; анализ риска создания криптографически релевантных квантовых компьютеров (CRQC); безопасности симметричной криптографии и криптографии с открытым ключом при наличии CRQC; обзор усилий по стандартизации NIST PQC; анализ возможностей перехода к квантово-устойчивой криптографии с открытым ключом, а также актуальности, взглядов и состояния развития квантово-стойкой криптографии в Европейском Союзе. Также освещается стандартизация постквантовой криптографии NIST.

Ключевые слова: постквантовая криптография; квантовый компьютер; стандартизация; электронная подпись; транспортировка ключей.

Табл. 2. Ил. 2. Библиогр.: 31 назв.

UDC 004.725.4+517.9

New continuous-discrete model for wireless sensor networks security / *Y. Kotukh, V. Lubchak, O. Strakh* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. No 210. P. 99 –103.

A wireless sensor network (WSN) is a group of "smart" sensors with a wireless infrastructure designed to monitor the environment. This technology is the basic concept of the Internet of Things (IoT). The WSN can transmit confidential information while working in an insecure environment. Therefore, appropriate security measures must be considered in the network design. However, computational node constraints, limited storage space, an unstable power supply, and unreliable communication channels, and unattended operations are significant barriers to the application of cybersecurity techniques in these networks. This paper considers a new continuous-discrete model of malware propagation through wireless sensor network nodes, which is based on a system of so-called dynamic equations with impulsive effect on time scales.

Key words: IoT; wireless network; security model; national cybersecurity.

Ref.: 19 items.

УДК 004.725.4+517.9

Нова неперервно-дискретна модель захисту бездротових сенсорних мереж / *Є. В. Котух, В. О. Любчак, О. П. Страх* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 99 –103.

Бездротові сенсорні мережі (БСМ) – це групи «розумних» датчиків з інфраструктурою бездротового зв’язку, призначені для моніторингу умов навколишнього середовища. Ця технологія є базовим поняттям Інтернету речей (IoT) та одним з найважливіших елементів в сучасних телекомунікаційних системах. БСМ можуть передавати важливу конфіденційну інформацію, працюючи у незахищеному середовищі, а, відтак, у проєктуванні мережі потрібно враховувати відповідні заходи безпеки. Однак обчислювальні обмеження вузлів, обмежений простір для зберігання даних, нестійке джерело живлення, ненадійний канал зв’язку та операції без нагляду є значними перешкодами для застосування технічних методів кібербезпеки у цих мережах.

Існують різні математичні моделі щодо вивчення поширення шкідливого програмного забезпечення (ПЗ) у БСМ: глобальні чи індивідуальні, неперервні чи дискретні, детерміновані чи стохастичні. Зазвичай ці моделі базуються на використанні систем диференціальних рівнянь у частинних похідних, систем звичайних диференціальних рівнянь, клітинних автоматів, ланцюгів Маркова, агентного моделювання тощо. Але, враховуючи особливості отримання даних щодо стану тієї чи іншої групи вузлів БСМ, процес поширення шкідливого ПЗ не можна розглядати у суто неперервному або суто дискретному за часом режимі. Ці два фактори мають бути поєднані.

Незважаючи на тривалу історію сенсорних мереж, концепція їх побудови до кінця не сформована. Тому дослідження окремих властивостей таких мереж є дуже важливим як для вітчизняної, так і для світової науки. Крім того, для стратегічно важливих галузей країни, зокрема національної кібербезпеки, захист бездротових сенсорних мереж є важливою складовою.

У статті розглянуто нову неперервно-дискретну модель поширення шкідливого ПЗ через вузли бездротової сенсорної мережі, яка базується на системі так званих динамічних рівнянь з імпульсним впливом на часовій шкалі. У якості змінних величин у цій системі взято кількість вузлів БСМ, які перебувають в одному із п'яти класів: сприйнятливому *S* (датчики не заражені шкідливим ПЗ, але мають сприйнятливі до нього обчислювальні характеристики), виявленому *E* (через датчики пройшло шкідливе ПЗ, яке не може передатися на суміжні датчики мережі), зараженому *I* (датчики заражені шкідливим ПЗ та мають можливість робити спроби зараження інших), відновленому *R* (датчики набувають тимчасового імунітету від шкідливого ПЗ) та віджилому *D* (датчики не підлягають відновленню). Крім того, у запропонованій моделі враховано можливість раптової зміни параметрів БСМ під впливом шкідливого ПЗ, яка не споріднена з її природним функціонуванням.

Ключові слова: IoT; бездротова мережа; модель захисту; національна кібербезпека.

Бібліогр.: 19 назв.

RADIOLOCATION AND RADIONAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ
РАДИОЛОКАЦИЯ И НАВИГАЦИЯ

UDC 629.7.022

Analysis of acoustic direction finding methods for unmanned aerial vehicles / *V.M. Kartashov, M.V. Rybnykov, A.V. Kartashov, V.A. Pososhenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 104 – 112.

Currently, classical means of detecting objects do not provide the necessary efficiency for detecting small UAVs, and acoustic location among the known methods for their observation is the most cost-effective solution.

The article analyzes the well-known methods of direction finding of acoustic signals in order to select algorithms for processing UAV signals. Obtaining qualitative indicators of the analyzed algorithms was carried out by the method of statistical computer modeling in the Matlab environment.

Based on the simulation results, it is shown that classical methods are the most stable under conditions of low signal-to-noise ratios. The GCC-PHAT direction finding algorithm, based on determining the difference in the time of arrival of a signal at spaced points, is computationally economical and simple enough to determine the direction to the UAV, but it is not capable of distinguishing more than one radiation source within the diagram orientation. Beamforming methods are also relatively easy to implement and computationally efficient, and are more robust at low signal-to-noise ratios. The SRP-NAM algorithm has a greater accuracy in determining angles than SRP-PHAT, so it can be an adequate replacement for the SRP-PHAT algorithm.

High-resolution methods provide better directional resolution than classical methods, which, in the case of a limited microphone array aperture, is a positive factor in the design of an UAV direction finding station. High resolution methods were considered: non-coherent MUSIC, non-coherent normalized MUSIC and TOPS method. It is shown that incoherent MUSIC gives poor results in distinguishing close UAV signals, since unequal estimates of the entire frequency range are concentrated during bearing formation. The incoherent normalized MUSIC algorithm is able to efficiently use the entire frequency range of the UAV acoustic signal. The TOPS algorithm is inferior to the incoherent normalized MUSIC algorithm, and on the other hand, it does not require a priori estimates of the number of radiation sources.

Key words: unmanned aerial vehicle; detection complex; direction finding station; sodar; microphone array; aperture; signal processing.

8 fig. Ref: 23 items.

УДК 629.7.022

Аналіз методів акустичної пеленгації безпілотних літальних апаратів / *В.М. Карташов, М.В. Рибников, О.В. Карташов, В.О. Посошенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 104 – 112.

Класичні засоби виявлення об'єктів не забезпечують необхідної ефективності стосовно виявлення малих безпілотних літальних апаратів (БПЛА), а акустична локація серед відомих методів їх спостереження є найбільш економічно доцільним рішенням.

У статті аналізуються відомі методи пеленгації акустичних сигналів з метою вибору алгоритмів для обробки сигналів БПЛА. Отримання якісних показників аналізованих алгоритмів здійснювалося методом статистичного комп'ютерного моделювання в середовище Matlab.

За результатами моделювання показано, що класичні методи є найбільш стійкими в умовах низьких відносин сигнал-шум. Алгоритм пеленгування GCC-PHAT, заснований на визначенні різниці моментів часу приходу сигналу в рознесені за простором точки, є обчислювально економічним і досить простим для визначення напрямку на БПЛА, однак він не здатний розрізняти більше одного джерела випромінювання в межах діаграми спрямованості. Методи, засновані на формуванні променю, також є порівняно простими у реалізації та обчислювально ефективними, вони більш стійкі при низьких значеннях сигнал-шум. Алгоритм SRP-NAM має більшу точність визначення кутів, ніж SRP-PHAT, тому він може бути адекватною заміною алгоритму SRP-PHAT.

Методи високої роздільної здатності забезпечують кращу роздільну здатність з напрямком, ніж класичні методи, що у разі обмеження по апертурі мікрофонних решіток є позитивним фактором при проектуванні станції пеленгування БПЛА. Розглядались такі методи високої роздільної здатності: некогерентний MUSIC, некогерентний нормований MUSIC та метод TOPS. Показано, що некогерентний MUSIC дає погані результати з розрізнення близьких сигналів БПЛА, оскільки при формуванні пеленгу осереднюються нерівнозначні оцінки всього діапазону частот. Алгоритм некогерентної нормованої MUSIC здатний ефективно використовувати весь діапазон частот акустичного сигналу БПЛА. Алгоритм TOPS поступається алгоритму некогерентної нормованої MUSIC, а з іншого боку він не вимагає апріорних оцінок кількості джерел випромінювання.

Ключові слова: безпілотний літальний апарат; комплекс виявлення; станція пеленгування; содар; мікрофонна решітка; апертура; обробка сигналів.

Іл. 8. Бібліогр.: 23 назв.

УДК 629.7.022

Анализ методов акустической пеленгации беспилотных летальных аппаратов / *В.М. Карташов, Н.В. Рыбников, А.В. Карташов, В.А. Посошенко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 104 – 112.

Классические средства обнаружения объектов не обеспечивают необходимой эффективности обнаружения малых беспилотных летательных аппаратов (БПЛА), а акустическая локация среди известных методов их наблюдения является наиболее экономически целесообразным решением.

В статье анализируются известные методы пеленгации акустических сигналов с целью выбора алгоритмов обработки сигналов БПЛА. Получение качественных показателей анализируемых алгоритмов производилось методом статистического компьютерного моделирования в среде Matlab.

По результатам моделирования показано, что классические методы являются наиболее устойчивыми в условиях низких отношений сигнал-шум. Алгоритм пеленгования GCC-PHAT, основанный на определении разности моментов времени прихода сигнала в разнесенные по пространству точки, является вычислительно экономичным и достаточно простым для определения направления на БПЛА, однако он не способен различать более одного источника излучения в пределах диаграммы направленности. Методы, основанные на формировании луча, также являются сравнительно простыми в реализации и вычислительно эффективными, они более устойчивы при низких значениях сигнал-шум. Алгоритм SRP-NAM имеет большую точность определения углов, чем SRP-PHAT, поэтому он может являться адекватной заменой алгоритму SRP-PHAT.

Методы высокого разрешения обеспечивают лучшее разрешение по направлению, чем классические методы, что в случае ограничения по апертуре микрофонных решеток является положительным фактором при проектировании станции пеленгования БПЛА. Рассматривались методы высокого разрешения: некогерентный MUSIC, некогерентный нормированный MUSIC и метод TOPS. Показано, что некогерентный MUSIC дает плохие результаты по разрешению близких сигналов БПЛА, поскольку при формировании пеленга используются неравнозначные оценки всего диапазона частот. Алгоритм некогерентного нормированного MUSIC способен эффективно использовать весь диапазон частот акустического сигнала БПЛА. Алгоритм TOPS уступает алгоритму некогерентного нормированного MUSIC, а с другой стороны, он не требует априорных оценок количества источников излучения.

Ключевые слова: беспилотный летательный аппарат; комплекс обнаружения; станция пеленгования; содар; микрофонная решетка; апертура; обработка сигналов.

Ил. 8. Библиогр.: 23 назв.

UDC 629.7.022

Determining the location of small unmanned aerial vehicles by acoustic radiation / *V.N. Oleynikov, V.M. Kartashov, S.A. Sheiko, O.V. Zubkov, E.I. Oleynikova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 113 – 127.

The features of the acoustic method for determining the location of the UAV using microphone arrays are considered: classical methods, super-resolution methods and the method of the difference in the arrival time of acoustic oscillations. A description of a sodar with a minimum number of microphones is given. The factors influencing the magnitude of the error in determining the coordinates of the UAV are analyzed. Estimates of the instrumental location error and errors caused by the Doppler effect were obtained by simulation modeling for a given configuration of the sodar microphone array. Using the considered sodar, the coordinates of a moving UAV were measured in open areas. Processing of field measurement results shows that the absolute values of errors in determining the coordinates of the UAV obtained using sodar at a distance of up to 100 m in 95% of cases do not exceed 3 m and in determining the height of no more than 1 m. Measurements of the azimuth and elevation angle are quite accurate at the same time high resolution.

Key words: acoustic radiation; cross correlation; angular coordinates; triangulation; microphone array; localization; passive sodar; quadrocopter; Doppler effect.

15 fig. Ref: 23 items.

УДК 629.7.022

Визначення розташування малорозмірних безпілотних літальних апаратів з акустичного випромінювання / *В.М. Олейніков , В.М. Карташов, С.О. Шейко, О.В. Зубков, О.І. Олейнікова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 113 – 127.

Розглянуто особливості акустичного метода визначення місця розташування БПЛА з використанням мікрофонних решіток: класичні методи, методи наддозволеності та метод різниці часу приходу акустичних коливань. Наводиться опис содару з мінімальним числом мікрофонів. Проаналізовано фактори, що впливають на величину похибки визначення координат БПЛА. Для заданої конфігурації мікрофонної решітки содару методом імітаційного моделювання отримані оцінки інструментальної похибки визначення та похибки, викликані ефектом Доплера. За допомогою розглянутого содару проведено вимір координат БПЛА, що переміщається в умовах відкритої місцевості. Обробка результатів натурних вимірювань показує, що абсолютні значення помилок визначення координат БПЛА, отримані за допомогою содару, на дальності до 100 м у 95 % випадків не перевищують 3 м, визначення висоти – не більше 1 м. Вимірювання азимуту та кута місця при цьому досить точні при високої роздільної здатності.

Ключові слова: акустичне випромінювання; взаємна кореляція; кутові координати; тріангуляція; мікрофонна решітка; локалізація; пасивний содар; квадрокоптер; ефект Доплера.

Іл. 15. Бібліогр.: 23 назв.

УДК 629.7.022

Определение местоположения малоразмерных беспилотных летательных аппаратов по акустическому излучению / *В.Н. Олейников, В.М. Карташов, С.А. Шейко, О.В. Зубков, Е.И. Олейникова* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 113 – 127.

Рассмотрены особенности акустического метода определения местоположения БПЛА с использованием микрофонных решёток: классические методы, методы сверхразрешения и метод разности времени прихода акустических колебаний. Приводится описание содара с минимальным числом микрофонов. Проанализированы факторы, влияющие на величину погрешности определения координат БПЛА. Для заданной конфигурации микрофонной решетки содара методом имитационного моделирования получены оценки инструментальной погрешности местоопределения и погрешности, вызванные эффектом Доплера. С помощью рассмотренного содара измерены координат перемещающегося БПЛА в условиях открытой местности. Обработка результатов натурных измерений показывает, что абсолютные значения ошибок определения координат БПЛА, полученные с помощью содара, на дальности до 100 м в 95% случаев не превышают 3 м, определения высоты – не более 1 м. Измерения азимута и угла места при этом достаточно точны при высокой разрешающей способности.

Ключевые слова: акустическое излучение; взаимная корреляция; угловые координаты; триангуляция; микрофонная решётка; локализация; пассивный содар; квадрокоптер; эффект Доплера.

Ил. 15. Библиогр.: 23 назв.

UDC 621.372(075)

Directional diagrams of acoustic radiation from unmanned aerial vehicles / *V.M. Kartashov, V.M. Oleynikov, I.S. Seleznyov, O.V. Kartashov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 128 – 140.

When solving the urgent task of detecting small unmanned aerial vehicles (UAVs) by their acoustic radiation (AR), there is a need to study the AR characteristics of UAVs. Therefore, considerable attention is paid in the literature to theoretical and experimental studies of the structure and parameters of the sound field created by UAVs.

This article is devoted to the experimental study of the directional diagrams of the acoustic radiation from the DJI Phantom 3 unmanned aerial vehicle. The UAV AR recording experiment was conducted in a "silenced" chamber, the walls of which are covered with sound-absorbing panels with a surface of a special geometric shape. The experimental setup includes the UAV mounted on a boom, a microphone for sound recording, and a boom for the microphone.

Studies of the structure and parameters of the sound field of the quadcopter have shown that the spectra of the emitted signal have pronounced harmonic components with frequencies multiples of the propeller rotation frequency. The spectral components have the greatest power in the frequency range up to 500 Hz, where the first harmonic has the largest amplitude, and then the spectrum components decrease to the ambient noise level.

Obtained two-dimensional and three-dimensional directional diagrams of the UAV AR with and without propellers when only the aircraft engine is operating. It is shown that in the absence of screws, the acoustic radiation is much weaker in level. The experimental data was also presented in the form of three-dimensional AR diagrams for the four harmonics of the acoustic signal, and it was analyzed what changes in the spatial orientation of the UAV AR are observed based on changes in the three-dimensional figure for each of the radiation harmonics.

It is shown that the spatial distribution of both the total energy (over the entire frequency range) of the acoustic signal and the energy of its individual spectral (harmonic) components is significantly anisotropic. It follows from this conclusion that the range of UAV detection and observation in real conditions is a statistical value that depends on the angle of its observation.

Key words: unmanned aerial vehicle; experiment; detection; signal; spectrum; acoustic radiation; directional diagram.

17 fig. Ref: 17 items.

УДК 621.372(075)

Діаграми спрямованості акустичного випромінювання безпілотних літальних апаратів / *В.М. Карташов, В.М. Олейніков, І.С. Селєзньов, О.В. Карташов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 128 – 140.

При вирішенні актуальної задачі виявлення малих безпілотних літальних апаратів (БПЛА) по їх акустичному випромінюванню (АВ) виникає необхідність дослідження характеристик АВ БПЛА. Тому значна увага в літературі приділяється теоретичним та експериментальним дослідженням структури і параметрів звукового поля, створюваного БПЛА.

Стаття присвячена експериментальному дослідженню діаграм спрямованості акустичного випромінювання безпілотного літального апарату «Dji Phantom 3». Експеримент з запису АВ БПЛА проводився у «заглушеній» камері, стіни якої покриті звукопоглинальними панелями із поверхнею спеціальної геометричної форми. Експериментальна установка включає БПЛА, закріплений на поворотній штанзі, мікрофон для запису звуку, та поворотну штангу для мікрофону.

Дослідження структури та параметрів звукового поля квадрокоптера показали, що спектри сигналу, що випромінюється, мають яскраво виражені гармонійні складові із частотами, кратними частоті обертання гвинта. Найбільшу потужність мають спектральні складові у частотному діапазоні до 500 Гц, де перша гармоніка має найбільшу амплітуду, а далі має місце зменшення складових спектру до рівня шуму навколишнього середовища.

Отримані двовимірні та тривимірні діаграми спрямованості АВ БПЛА з гвинтами та без гвинтів, коли працює тільки двигун літального апарату. Показано, що при відсутності гвинтів акустичне випромінювання значне слабкіше за рівнем. Експериментальні дані також було представлено у вигляді тривимірних діаграм АВ для чотирьох гармонік акустичного сигналу і проаналізовано, які зміни в просторової спрямованості АВ БПЛА спостерігаються за змінами тривимірної фігури для кожної з гармонік випромінювання.

Показано, що просторовий розподіл як повної енергії (у всьому діапазоні частот) акустичного сигналу, так і енергії окремих його спектральних (гармонічних) складових є суттєво анізотропним. З цього випливає, що дальність виявлення і спостереження БПЛА в реальних умовах є величиною статистичної, яка залежить від ракурсу його спостереження.

Ключові слова: безпілотний літальний апарат; експеримент; виявлення; сигнал; спектр; акустичне випромінювання; діаграма спрямованості.

Іл. 17. Бібліогр.: 17 назв.

УДК 621.372(075)

Диаграммы направленности акустического излучения беспилотных летательных аппаратов / *В.М. Карташов, В.М. Олейников, И.С. Селезнёв, А.В. Карташов* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 128 – 140.

При решении актуальной задачи обнаружения малых беспилотных летательных аппаратов (БПЛА) по их акустическому излучению (АИ) возникает необходимость исследования характеристик АИ БПЛА. Поэтому значительное внимание уделяется теоретическим и экспериментальным исследованиям структуры и параметров звукового поля, создаваемого БПЛА.

Статья посвящена экспериментальному исследованию диаграмм направленности акустического излучения беспилотного летательного аппарата Dji Phantom 3. Эксперимент по записи АИ БПЛА проводился в заглушенной камере, стены которой покрыты звукопоглощающими панелями с поверхностью специальной геометрической формы. Экспериментальная установка включает БПЛА, закрепленный на поворотной штанге, микрофон для записи звука и поворотную штангу для микрофона.

Исследования структуры и параметров звукового поля квадрокоптера показали, что спектры излучаемого сигнала имеют ярко выраженные гармонические составляющие с частотами, кратными частоте вращения винта. Наибольшую мощность имеют спектральные составляющие частотного диапазона до 500 Гц, где первая гармоника имеет наибольшую амплитуду, а далее имеет место уменьшение составляющих спектра до уровня шума окружающей среды.

Получены двумерные и трехмерные диаграммы направленности АИ БПЛА с винтами и без винтов, когда работает только двигатель летательного аппарата. Показано, что при отсутствии винтов акустическое излучение значительно слабее по уровню. Экспериментальные данные также были представлены в виде трехмерных диаграмм АИ для четырех гармоник акустического сигнала и проанализировано, какие изменения в пространственной направленности АИ БПЛА наблюдаются по изменениям трехмерной фигуры для каждой из гармоник излучения.

Показано, что пространственное распределение как полной энергии (во всем диапазоне частот) акустического сигнала, так и энергии отдельных его спектральных (гармонических) составляющих существенно анизотропно. Из этого следует, что дальность обнаружения и наблюдения БПЛА в реальных условиях является статистической величиной, которая зависит от ракурса его наблюдения.

Ключевые слова: беспилотный летательный аппарат; эксперимент; обнаружение; сигнал; спектр; акустическое излучение; диаграмма направленности.

Ил. 17. Библиогр.: 17 назв.

UDC 621.396.967.2

Processing of information from networks of airspace surveillance radar systems / *I.V. Svyd, M.G. Tkach, A.O. Sierikov, O.V. Korotich, S.V. Datsko, D.O. Sukhorukov, T.S. Machonis* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. № 210. P. 141 – 149.

The presented paper examines the principles and methods for processing of information from airspace surveillance radar networks. Information technologies allow implementing automatic collection, processing, storage, transmission and issuance of radar information to users. In this work, the synthesis and analysis of the optimal structure of the interstage processing of signal data and radar information of primary processing in the network of airspace surveillance radar systems is carried out. The quality of information from radar networks of airspace surveillance systems affects almost all indicators of the quality of the radar systems network functioning. The stages of radar information processing in radar systems networks are analyzed. The importance of specifying the above stages for creating a complete picture of the air situation in the area of responsibility is shown. To improve the quality of information support for consumers, a network of radar systems requires information processing at all stages. At each stage of information processing, information processing quality indicators were analyzed. This made it possible to show that the staged implementation of information processing, on the one hand, simplified the optimization of processing within each processing stage, but on the other hand, it made it difficult to carry out compatible optimization of both the detection of an aerial object and the measurement of the coordinates of an aerial object. The synthesized structure of processing radar information of the network of radar systems of airspace surveillance, which in its turn made it possible to carry out interstage optimization of processing of both signal data and primary processing information. Calculations have shown that the method of information processing, in which the combination of information is carried out at the level of decision-making on the detection of airborne objects in each signal data processing channel, has some advantages in the quality of information processing of the network of radar systems compared to the one currently used. time option combination information at the stage of signal processing. At the same time, for the method of combining information at the level of decision-making about the detection of aerial objects, the flow of transmitted information to the joint processing point is significantly reduced. All this allows improving the quality of information processing in the airspace control system.

Key words: radar system; network; airspace; surveillance system; data processing; signal data; interstage processing; optimization; quality of information processing; airspace control; air object.

2 fig. Ref: 37 items.

УДК 621.396.967.2

Обробка інформації мереж радіолокаційних систем спостереження повітряного простору / *І.В. Свид, М.Г. Ткач, А.О. Сєріков, О.В. Коротіч, С.В. Дацько, Д.О. Сухоруков, Т.С. Мачоніс* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 141 – 149.

Розглянуто принципи і методи обробки інформації мереж радіолокаційних систем спостереження повітряного простору. Інформаційні технології дозволяють реалізувати автоматичний збір, обробку, зберігання, передачу та видачу радіолокаційної інформації користувачам. Проведено синтез та аналіз оптимальної структури міжетапної обробки сигнальних даних та радіолокаційної інформації первинної обробки в мережі радіолокаційних систем спостереження повітряного простору. Якість інформації мереж радіолокаційних систем спостереження повітряного простору впливає практично на всі показники якості функціонування мережі радіолокаційних систем. Проаналізовано етапи обробки радіолокаційної інформації в мережах радіолокаційних систем. Показано важливість зазначення наведених етапів для створення повної картини повітряної обстановки в зоні відповідальності. Для підвищення якості інформаційного забезпечення споживачів мережа радіолокаційних систем потребує проведення обробки інформації на всіх етапах. На кожному етапі обробки інформації проаналізовано показники якості обробки інформації. Це дозволило показати, що етапна реалізація обробки інформації з одного боку спростила проведення оптимізації обробки всередині кожного етапу обробки, проте, з іншого боку, ускладнила проведення сумісної оптимізації, як виявлення повітряного об’єкта, так і вимірювання координат повітряного об’єкта. Синтезована структура обробки радіолокаційної інформації мережі радіолокаційних систем спостереження повітряного простору, яка дозволила здійснити міжетапну оптимізацію обробки як сигнальних даних, так й інформації первинної обробки. Наведені розрахунки показали, що спосіб обробки інформації, при якому поєднання інформації здійснюється на рівні прийняття рішень про виявлення повітряних об’єктів в кожному каналі обробки сигнальних даних, має деякі переваги в якості обробки інформації мережі радіолокаційних систем у порівнянні з варіантом поєднання інформації на етапі обробки сигналів, що використовується в теперішній час. При цьому, для способу поєднання інформації на рівні прийняття рішень про виявлення повітряних об’єктів потік інформації, що передається на пункт сумісної обробки, значно зменшується. Все це дозволяє підвищити якість обробки інформації в системі контролю повітряного простору.

Ключові слова: радіолокаційна система; мережа; повітряний простір; система спостереження ; обробка інформації; сигнальні дані; міжетапна обробки; оптимізація; якість обробки інформації; контроль повітряного простору; повітряний об'єкт.

Іл. 2. Бібліогр.: 37 назв.

УДК 621.396.967.2

Обработка информации радиолокационных систем наблюдения воздушного пространства / *И.В. Свид, М.Г. Ткач, А.А. Сериков, А.В. Коротич, С.В. Дацько, Д.А. Сухоруков, Т.С. Мачоніс* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 141 – 149.

Рассмотрены принципы и методы обработки информации радиолокационных сетей систем наблюдения воздушного пространства. Информационные технологии позволяют реализовать автоматический сбор, обработку, хранение, передачу и выдачу радиолокационной информации пользователям. Проведены синтез и анализ оптимальной структуры межэтапной обработки сигнальных данных и радиолокационной информации первичной обработки в сети радиолокационных систем наблюдения воздушного пространства. Качество информации сетей радиолокационных систем наблюдения воздушного пространства влияет практически на все показатели качества функционирования сети радиолокационных систем. Проанализированы этапы обработки радиолокационной информации в сетях радиолокационных систем. Показана значимость приведенных этапов для представления полной картины воздушной обстановки в зоне ответственности. Для повышения качества информационного обеспечения потребителей сеть радиолокационных систем требует проведения обработки информации на всех этапах. На каждом этапе обработки информации проанализированы показатели качества обработки информации. Это позволило показать, что этапная реализация обработки информации, с одной стороны, упростила проведение оптимизации обработки внутри каждого этапа обработки, однако, с другой стороны, усложнила проведение совместной оптимизации как обнаружения воздушного объекта, так и измерения координат воздушного объекта. Синтезирована структура обработки радиолокационной информации сети радиолокационных систем наблюдения воздушного пространства, позволившая осуществить межэтапную оптимизацию обработки сигнальных данных и информации первичной обработки. Расчеты показали, что способ обработки информации, при котором сочетание информации осуществляется на уровне принятия решений об обнаружении воздушных объектов в каждом канале обработки сигнальных данных, имеет некоторые преимущества в качестве обработки информации сети радиолокационных систем по сравнению с используемым в настоящее время вариантом сочетания информации на этапе обработки сигналов. При этом для способа сочетания информации на уровне принятия решений по обнаружению воздушных объектов поток передаваемой информации на пункт совместной обработки значительно уменьшается. Все это позволяет повысить качество обработки информации в системе контроля воздушного пространства.

Ключевые слова: радиолокационная система; сеть; воздушное пространство; система наблюдения; обработка информации; сигнальные данные; межэтапная обработка; оптимизация; качество обработки информации; контроль воздушного пространства; воздушный объект.

Ил. 2. Библиогр.: 37 назв.

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS
ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ
ФИЗИКА ПРИБОРОВ, ЭЛЕМЕНТОВ И СИСТЕМ

UDC 541.64:542.06:678

High-thermally conductive composite polyimide materials / *V.M. Borshchov, O.M. Listratenko, M.A. Protsenko, I.T. Tymchuk, O.V. Kravchenko, O.V. Syddia, M.I. Slipchenko, B.M. Chichkov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 150 – 159.

This review is devoted to analysis of works in the field of creating electrically insulating heat-conducting polyimide composite films based on powders of micro-, submicro- or nano-sized fillers with high dielectric and heat-conducting properties for use as effective thermal interface materials in various electronic devices in instrument making. Particular attention is paid to studies on the influence of the size of nano- and microparticles of inorganic fillers on the heat-nducting, dielectric, and physical-mechanical properties of nanocomposite polyimide materials. The analysis of the results of work on the study of the dependence of thermal conductivity on the ratios of micron and nanosized particles in mixtures and their number in polyimides and on the conditions of their polymerization was carried out to confirm the possibility of increasing the thermal conductivity values of promising polyimide materials from 0.12 W/(m•K) up to 5÷10 W/ (m•K). It is noted that the highest thermal conductivity of industrially produced modern polyimide films on market does not exceed 0.75÷0.8 W/(m•K). The task of creating inexpensive, but high-quality heat-conductive polyimide composite materials with sufficiently high thermal conductivity without deteriorating their strength and ductility characteristics is currently relevant and technically in demand.

Key words: electrical insulating heat-conductive polyimide composite films; micro, submicro and nano particles; inorganic fillers..

1 tab. 2 fig. Ref: 25 items.

УДК 541.64:542.06:678

Високотеплопровідні композитні поліімідні матеріали / *В.М. Борщов, О.М. Лістратенко, М.А. Проценко, І.Т. Тимчук, О.В. Кравченко, О.В. Суддя, М.І. Сліпченко, Б.М. Чічков* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 150 – 159.

Огляд присвячено аналізу робіт в області створення електроізоляційних теплопровідних поліімідних композиційних плівок на основі порошків мікро-, субмікро- або нанорозмірних наповнювачів, що мають високі діелектричні і теплопровідні властивості для застосування в якості ефективних термоінтерфейсних матеріалів в різних електронних пристроях в приладобудуванні. Особливу увагу приділено роботам з дослідження впливу розмірів нано- і мікрочастинок неорганічних наповнювачів на теплопровідні, діелектричні і фізико-механічні

властивості в тому числі нанокомпозиційних поліімідних матеріалів. Проаналізовано результати робіт з дослідження залежності теплопровідності від співвідношень мікронних і нанорозмірних частинок в сумішах і їх кількості в поліімідах і від умов їх полімеризації для підтвердження можливості збільшення значень теплопровідності перспективних поліімідних матеріалів від 0,12 Вт/(м•К) до 5÷10 Вт/(м•К). Зазначено, що найвища теплопровідність сучасних поліімідних плівок, що промислово випускаються та які представлені на ринку, не перевищує 0,75÷0,8 Вт/(м•К). Завдання створення недорогих, але високоякісних теплопровідних композиційних поліімідних матеріалів з досить високими показниками теплопровідності без погіршення характеристик їх міцності і пластичності є актуальним і технічно затребуваним.

Ключові слова: електроізоляційні теплопровідні поліімідні композиційні плівки; мікро-, субмікро- і наночастинки; неорганічні наповнювачі.

Табл. 1. Іл. 2. Бібліогр.: 25 назв.

УДК 541.64:542.06:678

Высокотеплопроводные композитные полиимидные материалы / *В.Н. Борщев, А.М. Листратенко, М.А. Проценко, И.Т. Тымчук, А.В. Кравченко, А.В. Судья, Н.И. Слипченко, Б.Н. Чичков* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 150 – 159.

Обзор посвящен анализу работ в области создания электроизоляционных теплопроводящих полиимидных композиционных пленок на основе порошков микро-, субмикро- или наноразмерных наполнителей, обладающих высокими диэлектрическими и теплопроводящими свойствами для применения в качестве эффективных термоинтерфейсных материалов в различных электронных устройствах в приборостроении. Особое внимание уделено работам по исследованию влияния размеров нано- и микрочастиц неорганических наполнителей на теплопроводящие, диэлектрические и физико-механические свойства нанокомпозиционных полиимидных материалов. Проведен анализ результатов работ по исследованию зависимости теплопроводности от соотношений микронных и наноразмерных частиц в смесях и их количества в полиимидах и от условий их полимеризации для подтверждения возможности увеличения значений теплопроводности перспективных полиимидных материалов от 0,12 Вт/(м•К) до 5÷10 Вт/(м•К).

Отмечено, что самая высокая теплопроводность промышленно выпускаемых современных полиимидных пленок, представленных на рынке, не превышает 0,75÷0,8 Вт/(м•К). А задача создания недорогих, но высококачественных теплопроводящих полиимидных композиционных материалов с достаточно высокими показателями теплопроводности без ухудшения характеристик их прочности и пластичности, в настоящее время является актуальной и технически востребованной.

Ключевые слова: электроизоляционные теплопроводящие полиимидные композиционные пленки; микро-, субмикро- и наночастицы; неорганические наполнители.

Табл. 1. Ил. 2. Библиогр.: 25 назв.

UDC 621.362

The use of various materials as a metal component in a metamaterial thermophotovoltaic emitter / *M.A. Yasnohorodskyi* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 160 – 166.

Thermophotovoltaics (TPV) is a process by which photons emitted by a heat emitter are converted into electrical energy by a photovoltaic cell. Selective heat emitters that can survive temperatures at or above 1000°C have the potential to significantly improve the energy conversion efficiency of a PV cell by limiting the emission of photons with energies below the band gap energy of a photovoltaic cell.

Waste heat can be a valuable source of energy if we can find a way to harvest it efficiently. Deviations from ideal absorption and ideal blackbody behavior lead to light losses. For selective emitters, any light emitted at wavelengths outside the bandgap energy of the photovoltaic system may not be efficiently converted, reducing efficiency. In particular, it is difficult to avoid emission associated with phonon resonance for wavelengths in the deep infrared, which cannot be practically converted. An ideal emitter would not emit light at wavelengths other than the bandgap energy, and much TFP research is devoted to designing emitters that approximate better this narrow emission spectrum.

TPV systems usually consist of a heat source, a radiator and a waste heat removal system. TFV cells are placed between the emitter, often a metal or similar block, and the cooling system, often a passive radiator.

Efficiency, heat resistance and cost are the three main factors for choosing a TPF emitter. The efficiency is determined by the absorbed energy relative to the incoming radiation. High temperature operation is critical because efficiency increases with operating temperature. As the temperature of the emitter increases, the radiation of the black body shifts toward shorter waves, which allows for more efficient absorption by photocells. This paper demonstrates the feasibility of using materials such as platinum, gold, and nichrome as a metal component in a metamaterial emitter with respect to their absorption and thermal stability.

Key words: metamaterials; thermophotoelectric devices; emitter; melting point; emitter.

2 tab. 2 fig. Ref: 31 items.

УДК 621.362

Використання різних матеріалів в якості металевого компонента в метаматеріальному термофотовольтаїчному випромінювачі / *М.А. Ясногородський* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 160 – 166.

Термофотовольтаїка (ТФВ) - це процес, за допомогою якого фотони, випромінювані тепловим випромінювачем, перетворюються в електричну енергію за допомогою фотоелектричного елемента. Селективні випромінювачі тепла, які можуть виживати при температурах на рівні або вище 1000°C, мають потенціал для значного підвищення ефективності перетворення енергії ТФВ шляхом обмеження випромінювання фотонів з енергією, нижчою від енергії забороненої зони фотоелектричної комірки.

Відпрацьоване тепло може бути цінним джерелом енергії, якщо ми зможемо знайти спосіб його ефективного збирання. Відхилення від ідеального поглинання та ідеальної поведінки чорного тіла призводять до втрат світла. Для селективних випромінювачів будь-яке світло, що випромінюється на довжинах хвиль, що не відповідають енергії забороненої зони фотоелектричної системи, може перетворюватися неефективно. Зокрема, важко уникнути випромінювання, пов'язаного з фононним резонансом, для довжин хвиль у глибокому інфрачервоному діапазоні, які неможливо практично перетворити. Ідеальний випромінювач не випромінював би світла на інших довжинах хвиль, окрім енергії забороненої зони, і багато досліджень ТФВ присвячено розробці випромінювачів, які краще наближають цей вузький спектр випромінювання.

Системи ТФВ зазвичай складаються з джерела тепла, випромінювача та системи відведення відпрацьованого тепла. Комірки ТФВ розміщуються між емітером, часто металевим або подібним блоком, і системою охолодження, часто пасивним радіатором.

Ефективність, термостійкість і вартість є трьома основними факторами для вибору випромінювача ТФВ. Ефективність визначається поглиненою енергією відносно вхідного випромінювання. Експлуатація при високій температурі має вирішальне значення, оскільки ефективність зростає з робочою температурою. Зі збільшенням температури випромінювача випромінювання чорного тіла зміщується в бік коротших хвиль, що дозволяє більш ефективно поглинати фотоелементами. У роботі продемонстровано доцільність використання таких матеріалів, як платина, золото та ніхром в якості металевого компонента в метаматеріальному випромінювачі відносно їх показників поглинання та термостійкості.

Ключові слова: метаматеріали; термофотоелектричні пристрої; емітер; температура плавлення; випромінювач.

Табл. 2. Іл. 2. Бібліогр.: 31 назв.

УДК 621.362

Использование различных материалов в качестве металлического компонента в метаматериальном термофотовольтаическом излучателе / *М.А. Ясногородский* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 160 – 166.

Термофотовольтаика (ТФВ) – это процесс, с помощью которого фотоны, излучаемые тепловым излучателем, превращаются в электрическую энергию с помощью фотоэлектрического элемента. Селективные излучатели тепла, которые могут выживать при температурах на уровне или выше 1000°C, имеют потенциал для значительного повышения эффективности преобразования энергии ТФВ путем ограничения излучения фотонов с энергией ниже энергии запрещенной зоны фотоэлектрической ячейки.

Отработанное тепло может быть ценным источником энергии, если мы сможем найти способ его эффективной уборки. Отклонения от идеального поглощения и идеального поведения черного тела приводят к потере света. Для селективных излучателей любой излучаемый свет на длинах волн, не соответствующих энергии запрещенной зоны фотоэлектрической системы, может преобразовываться неэффективно. В частности, трудно избежать излучения, связанного с фононным резонансом, для длин волн в глубоком инфракрасном диапазоне, которые невозможно практически преобразовать. Идеальный излучатель не излучал бы свет на других длинах волн, кроме энергии запрещенной зоны. Многие исследования ТФВ посвящены разработке излучателей, которые лучше приближают этот узкий спектр излучения.

Системы ТФВ обычно состоят из источника тепла, излучателя и отвода отработанного тепла. Ячейки ТФВ размещаются между эмиттером, часто металлическим или подобным блоком, и системой охлаждения, часто пассивным радиатором.

Эффективность, термостойкость и стоимость являются тремя основными факторами выбора излучателя ТФВ. Эффективность определяется поглощенной энергией относительно входящего излучения. Эксплуатация при высокой температуре имеет решающее значение, поскольку эффективность возрастает с рабочей температурой. С увеличением температуры излучателя излучение черного тела смещается в сторону более коротких волн, что позволяет более эффективно поглощать фотоэлементами. В работе продемонстрирована целесообразность использования таких материалов, как платина, золото и нихром в качестве металлического компонента в метаматериальном излучателе по отношению к их показателям поглощения и термостойкости.

Ключевые слова: метаматериалы; термофотоэлектрические устройства; эмиттер; температура плавления; излучатель.

Табл. 2. Ил. 2. Библиогр.: 31 назв.

INFORMATION METHODS OF RADIO ENGINEERING
ІНФОРМАЦІЙНІ МЕТОДИ РАДІОТЕХНІКИ
ИНФОРМАЦИОННЫЕ МЕТОДЫ РАДИОТЕХНИКИ

UDC 621.391:519.246.8

Mathematical models of non-stationary random processes in the SVVP representation / *V.A. Tikhonov, V.M. Kartashov, O.V. Kartashov, V.A. Pososhenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 167 – 176.

The work examines methods and mathematical models that provide the possibility of researching the statistical characteristics of complex and non-stationary random processes describing a wide class of physical phenomena. The proposed models can be used to study the processes observed in various fields of human activity, namely, to analyze the trajectories of unmanned aerial vehicles, their acoustic signals, meteorological processes reflecting the state of the atmosphere.

Real and simulated non-stationary random processes considered in the work are represented by the complex vector random process (CVRP) model. In this case, the length of the subvector is equal to the period of the seasonal component. In fact, in such a representation, the time series readings are replaced by their aggregate, i.e. subvectors. Statistical relationships are analyzed for subvectors, and not, as usual, for process counts. If the length of the subvector is equal to one, all operations in the SVVP representation are equivalent to the usual operations for time series.

The mathematical apparatus developed in the article was used to analyze changes in time series of atmospheric temperature observed over a long period of time; average annual temperatures were estimated with subsequent smoothing with a low-pass filter. The results obtained can be used to analyze medium-term and long-term changes in atmospheric conditions, refine the results obtained by traditional methods of mathematical statistics, analyze and predict data flows in mobile communication networks, as well as in other areas of human activity.

Key words: non-stationary process; model; autoregression; vector; trend; spectrum; shaping filter; trajectory; temperature.

10 fig. Ref: 17 items.

УДК 621.391:519.246.8

Математичні моделі нестаціонарних випадкових процесів у СВВП поданні / *В.А. Тихонов, В.М. Карташов, О.В. Карташов, В.О. Посошенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 167 – 176.

Розглядаються методи та математичні моделі, що забезпечують можливість дослідження статистичних характеристик складених, а також нестаціонарних випадкових процесів, що описують широкий клас фізичних явищ. Запропоновані моделі можуть бути використані для дослідження процесів, що спостерігаються у різних галузях людської діяльності, – для аналізу траєкторій руху безпілотних літальних апаратів, їх акустичних сигналів, метеорологічних процесів, що відображають стан атмосфери.

Реальні та імітаційні нестаціонарні випадкові процеси, що розглядаються, подаються моделлю складеного векторного випадкового процесу (СВВП). При цьому довжина підвектора дорівнює періоду сезонної складової. Фактично у такому поданні відліки часового ряду замінюються їх сукупністю, тобто підвекторами. Аналізуються статистичні зв'язки для підвекторів, а не як завжди, для відліків процесу. Якщо довжина підвектора дорівнює одиниці, всі операції у поданні СВВП еквівалентні звичайним операціям для часових рядів.

Розроблений математичний апарат використовувався для аналізу змін у часових рядах температури атмосфери, що спостерігаються тривалий час; виконано оцінювання середньорічних температур з подальшим ковзним згладжуванням низькочастотним фільтром. Результати можуть бути використані для аналізу середньострокових та довгострокових змін атмосферних умов, уточнення результатів, отриманих традиційними методами математичної статистики, аналізу та прогнозування потоків даних у мережах мобільного зв'язку, а також в інших галузях людської діяльності.

Ключові слова: нестаціонарний процес; модель; авторегресія; вектор; тренд; спектр; формуючий фільтр; траєкторія; температура.

Іл. 10. Бібліогр.: 17 назв.

УДК 621.391:519.246.8

Математические модели нестационарных случайных процессов в СВВП представлении / *В.А. Тихонов, В.М. Карташов, А.В. Карташов, В.А. Посошенко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 167 – 176.

Рассматриваются методы и математические модели, обеспечивающие возможность исследования статистических характеристик составных, а также нестационарных случайных процессов, описывающих широкий класс физических явлений. Предлагаемые модели могут быть использованы для исследования процессов, наблюдаемых в различных областях человеческой деятельности – для анализа траекторий движения беспилотных летательных аппаратов, их акустических сигналов, метеорологических процессов, отражающих состояние атмосферы.

Рассматриваемые в работе реальные и имитационные нестационарные случайные процессы представляются моделью составного векторного случайного процесса (СВСП). При этом длина подвектора равна периоду

сезонной составляющей. Практически в таком представлении отсчеты временного ряда заменяются их совокупностью, другими словами подвекторами. Анализируются статистические связи для подвекторов, а не, как обычно, для отсчетов процесса. Если длина подвектора равна единице, все операции по представлению СВВП эквивалентны обычным операциям для временных рядов.

Разработанный математический аппарат использовался для анализа изменений во временных рядах температуры атмосферы, наблюдаемых длительное время; выполнена оценка среднегодовых температур с последующим скользящим сглаживанием низкочастотным фильтром. Результаты могут использоваться для анализа среднесрочных и долгосрочных изменений атмосферных условий, уточнения результатов, полученных традиционными методами математической статистики, анализа и прогнозирования потоков данных в сетях мобильной связи, а также в других областях человеческой деятельности.

Ключевые слова: нестационарный процесс; модель; авторегрессия; вектор; тренд; спектр; формуючий фільтр; траектория; температура.

Ил. 10. Библиогр.: 17 назв.

UDC 621.372(075.8)

Corrective Function Method for the Fractal Analysis / *O.V. Lazorenko, A.A. Onishchenko, L.F. Chernogor* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 177 – 187.

One of the main numerical characteristics used in numerous methods of fractal analysis is the corresponding fractal dimensions. The accuracy of estimating these dimensions in the vast majority of cases is quite small, which cannot satisfy, first of all, researchers-practitioners. The method of the corrective function is put forward, which makes it possible to compensate for the ever-existing nonlinearity of the dependence between the true value of the fractal dimension and its estimation, performed using the selected method of monofractal analysis of signals and processes for a known number of samples of the discrete data vector of the investigated signal. The main idea of the method is to build and apply a special correction function using a set of model fractal signals with previously known values of the fractal dimension. The mathematical bases of the new method are outlined. Features of the practical application of the corrective function method are considered on the example of the evaluation of regularization, boxing, variation and Hurst fractal dimensions. For them, the minimum values of the number of samples of the discrete data vector of the investigated signal, at which these dimensions can still be estimated, are defined. Using a set of model monofractal and multifractal signals on the example of the dynamical fractal analysis method, the effectiveness of the created method of the corrective function is shown. It is proven that due to the application of the correction function method, the maximum deviation of the estimated fractal dimension from the true known value for the specified dimensions is reduced from 25 – 55% to 5 – 7%.

Key words: fractal; signal; process; analysis; method; dimension; estimation; accuracy; correction.

1 tab. 5 fig. Ref: 20 items.

УДК 621.372(075.8)

Метод коригуючої функції для фрактального аналізу / *О.В. Лазоренко, А.А. Онищенко, Л.Ф. Чорногор* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 177 – 187.

Одними із основних числових характеристик, що використовуються у численних методах фрактального аналізу, є відповідні фрактальні розмірності. Точність оцінки цих розмірностей у переважній більшості випадків є досить невеликою, що не може задовольняти, у першу чергу, дослідників-практиків. Запропоновано метод коригуючої функції, який дозволяє компенсувати завжди існуючу нелінійність залежності між істинним значенням фрактальної розмірності та її оцінкою, здійсненою з використанням обраного методу монофрактального аналізу сигналів і процесів за відомої кількості відліків дискретного вектору даних досліджуваного сигналу. Основна ідея методу полягає у побудові та застосуванні спеціальної коригуючої функції з використанням набору модельних фрактальних сигналів із заздалегідь відомими значеннями фрактальної розмірності. Викладено математичні засади нового метода. Розглянуто особливості практичного застосування методу коригуючої функції на прикладі оцінювання регуляризаційної, клітковой, варіаційної та херстової фрактальних розмірностей. Для них встановлено мінімальні значення кількості відліків дискретного вектору даних досліджуваного сигналу, за якої ці розмірності ще можна оцінювати. Із використанням низки модельних монофрактальних і мультифрактальних сигналів на прикладі методу динамічного фрактального аналізу продемонстровано ефективність створеного метода коригуючої функції. Доведено, що завдяки застосуванню метода коригуючої функції максимальне відхилення оцінюваної фрактальної розмірності від істинного відомого значення для вказаних розмірностей зменшено з 25 – 55 % до 5 – 7 %.

Ключові слова: фрактал; сигнал; процес; аналіз; метод; розмірність; оцінювання; точність; коригування.

Табл. 1. Іл. 5. Бібліогр.: 20 назв.

УДК 621.372(075.8)

Метод корректирующей функции для фрактального анализа / *О.В. Лазоренко, А.А. Онищенко, Л.Ф. Черногор* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 177 – 187.

Одними из основных числовых характеристик, используемых в многочисленных методах фрактального анализа, являются соответствующие фрактальные размерности. Точность оценки этих размерностей в подавляющем большинстве случаев достаточно невелика, что не может удовлетворять, в первую очередь, исследова-

телей-практиков. Предложен метод корректирующей функции, который позволяет компенсировать всегда существующую нелинейность зависимости между истинным значением фрактальной размерности и ее оценкой, осуществленной с использованием выбранного метода монофрактального анализа сигналов и процессов при известном количестве отсчетов дискретного вектора данных исследуемого сигнала. Основная идея метода заключается в построении и применении специальной корректирующей функции с использованием набора модельных фрактальных сигналов с заранее известными значениями фрактальной размерности. Изложены математические основы нового метода. Рассмотрено особенности практического применения метода корректирующей функции на примере оценивания регуляризационной, клеточной, вариационной и херстовой фрактальных размерностей. Для них установлено минимальное значение количества отсчетов дискретного вектора данных исследуемого сигнала, при котором эти размерности еще можно оценивать. С использованием ряда моделей монофрактальных и мультифрактальных сигналов на примере метода динамического фрактального анализа продемонстрирована эффективность созданного нового метода. Доказано, что благодаря применению метода корректирующей функции максимальное отклонение оцениваемой фрактальной размерности от истинного известного значения для указанных размерностей уменьшено с 25 – 55 % до 5 – 7 %.

Ключевые слова: фрактал; сигнал; процесс; анализ; метод; размерность; оценивание; точность; корректировка.

Табл. 1. Ил. 5. Библиогр.: 20 назв.

MEANS OF TELECOMMUNICATIONS
ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ
СРЕДСТВА ТЕЛЕКОММУНИКАЦИЙ

UDC 621.396.967.2

Optical Network Management by ONOS-Based SDN Controller / *O.I. Romanov, I.V. Svyd, N.I. Korniienko, A.O. Romanov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №210. P. 188 – 196.

The possibilities to manage the optical network with a logically centralized SDN control plane based on the Open Network Operating System (ONOS) are investigated. The structure of the controller and its main functional blocks are considered ensuring the collection of information about the state of network elements, the solution of the main control tasks, interaction of control systems built on different technological bases, are considered. The role and place of the open network operating system in the controller structure are shown, the description of the ONOS multilevel architecture in the form of a set of functional modules is given, the purpose and functions of the ONOS subsystems are analyzed, protocols and interfaces making it possible to present the SDN network as a model are described. The peculiarity of the model is that the managed network can be represented as a set of virtual network functions. Therefore, the control process becomes independent of which vendor's equipment was used to build the network, as well as whether the network is built on real physical elements or virtual ones. Using the ONOS allows you to build a logical centralized control plane in the SDN networks. The existing set of functional modules, services and interfaces in the ONOS allows you to perform optical network management tasks. For the further development of the ONOS, it is necessary to develop mathematical models and methods for the optimal solution of control problems in various operating conditions, which will become application-level software modules in the future.

Key words: Open Network Operating System, ONOS, Controller, SDN, Control Plane SDN, Data Plane SDN.

4 fig. Ref: 33 items.

УДК 621.396.967.2

Управління оптичною мережею контролером SDN на базі ONOS / *О.І. Романов, І.В. Свид, Н.І. Корнієнко, А.О. Романов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 210. С. 188 – 196.

Досліджуються можливості управління оптичною мережею логічно-централізованою площиною керування SDN на базі відкритої мережевої операційної системи ONOS. Розглянуто структуру контролера і його основні функціональні блоки, які забезпечують збір інформації про стан елементів мережі, вирішення основних завдань управління, взаємодію систем управління, побудованих на різній технологічній базі. Показані роль і місце відкритої мережевої операційної системи в структурі контролера, наведено опис багаторівневої архітектури ONOS у вигляді набору функціональних модулів, проаналізовано призначення і функції підсистем ONOS, дано опис протоколів і інтерфейсів, які дозволяють представити мережу SDN у вигляді моделі. Особливістю наведеної моделі є те, що керована мережа може бути представлена у вигляді набору віртуальних мережевих функцій. Тому процес управління стає незалежним від того, обладнання якого вендора використовувалося при побудові мережі, а також від того, побудована мережа на реальних фізичних елементах або віртуальних.

Використання ONOS дозволяє побудувати логічну централізовану площину управління в мережах SDN. Існуючий набір функціональних модулів, сервісів та інтерфейсів в ONOS дозволяє виконувати завдання управління оптичною мережею. Для подальшого розвитку ONOS необхідна розробка математичних моделей і методів оптимального вирішення задач керування в різних умовах експлуатації, які в майбутньому стануть програмними модулями прикладного рівня.

Ключові слова: відкрита мережева операційна система; ONOS; контролер; SDN; площина управління SDN; площина даних SDN.

Іл. 4. Бібліогр.: 33 назв.

УДК 621.396.967.2

Управление оптической сетью контроллером SDN на базе ONOS / *А.И. Романов, И.В. Свид, Н.И. Корниенко, А.А. Романов* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2022. Вып. 210. С. 188 – 196.

Исследуются возможности управления оптической сетью логически-централизованной плоскостью управления SDN на базе открытой сетевой операционной системы ONOS. Рассмотрена структура контроллера и его основные функциональные блоки, обеспечивающие сбор информации о состоянии элементов сети, решение основных задач управления, взаимодействие систем управления, построенных на разной технологической базе. Показаны роль и место открытой сетевой операционной системы в структуре контроллера, приведено описание многоуровневой архитектуры ONOS в виде набора функциональных модулей, проанализированы назначения и функции подсистем ONOS, дано описание протоколов и интерфейсов, позволяющих представить сеть SDN в виде модели. Особенностью данной модели является то, что управляемая сеть может быть представлена в виде набора виртуальных сетевых функций. Поэтому процесс управления становится независимым от того, оборудование какого вендора использовалось при построении сети, а также от того, построена ли сеть на реальных физических элементах или виртуальных. Использование ONOS позволяет выстроить логическую централизованную плоскость управления в сетях SDN. Существующий набор функциональных модулей, сервисов и интерфейсов ONOS позволяет выполнять задачи управления оптической сетью. Для дальнейшего развития ONOS необходима разработка математических моделей и методов оптимального решения задач управления в различных условиях эксплуатации, которые в будущем станут программными модулями прикладного уровня.

Ключевые слова: открытая сетевая операционная система; ONOS; контроллер; SDN; плоскость управления SDN; плоскость данных SDN.

Ил. 4. Библиогр.: 33 назв.

COLLECTION OF SCIENTIFIC PAPERS
RADIOTEKHNIKA
Issue 210
In English, Ukrainian and Russian

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 210
Англійською, українською, та російською мовами

СБОРНИК НАУЧНЫХ ТРУДОВ
РАДИОТЕХНИКА
Выпуск 210
На английском, украинском и русском языках

Коректор Л.І. Сащенко

Підп. до друку 30.09.2022. Формат 60x90/8. Папір офсет. Гарнітура Таймс. Друк. ризограф.
Ум. друк. арк**.** 13,9. Обл.-вид. арк. 12,8. Тираж 300 прим. Зам. № 494. Ціна договір.

Харківський національний університет радіоелектроніки (ХНУРЕ)
Просп. Науки, 14, Харків, 61166.

Оригінал-макет підготовлено і збірник надруковано у ПФ „Колегіум”, тел. (057) 703-53-74.
Свідоцтво про внесення суб’єкта видавничої діяльності до Державного реєстру видавців.
Сер. ДК №1722 від 23.03.2004.