МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

РАДІОТЕХНІКА

Всеукраїнський
міжвідомчий науково-технічний збірник

Засновано в 1965 р.

ВИПУСК 205

Харків
Харківський національний
університет радіоелектроніки
2021

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.

Сборник включен в Перечень научных профессиональных изданий Украины, категория «Б», технические и физико-математические науки (утверждено приказами МОНУ от 17.03.2020 № 409, от 02.07.2020 № 886, от 24.09.2020 № 1188) по специальностям: 171 – Электроника; 172 – Телекоммуникации и радиотехника; 173 – Авионика; 125 – Кибербезопасность; 151 – Автоматизация и компьютерно-интегрированные технологии; 152 – Метрология и информационно-измерительная техника; 153 – Микро- и наносистемная техника; 163 – Биомедицинская инженерия; 105 – Прикладная физика и наноматериалы.

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.

Сайт: rt.nure.ua

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

За зміст статті відповідальні автори.

Відповідальні випускові: *І.Д. Горбенко, д-р техн. наук, проф.*, *І.В. Свид, канд. техн. наук, доц.*
Технічний секретар *О.С. Полякова.*

Рекомендовано Вченою радою Харківського національного університету радіоелектроніки, протокол №6/15-1 від 02.07.2021.

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.***

ЗМІСТ

CONTENT

МЕТОДИ ТА АЛГОРИТМИ КРИПТОГРАФІЧНОГО ЗАХИСТУ ІНФОРМАЦІЇ

УДК 004.056.55 **DOI:10.30837/rt.2021.2.205.01**

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

ОСНОВНІ ПОЛОЖЕННЯ ТА РЕЗУЛЬТАТИ ПОРІВНЯННЯ ВЛАСТИВОСТЕЙ ЕЛЕКТРОННИХ ПІДПИСІВ ПОСТКВАНТОВОГО ПЕРІОДУ НА АЛГЕБРАЇЧНИХ РЕШІТКАХ

Вступ

Постквантові проєкти стандартів електронних підписів (ЕП) Falcon [1] та Dilithium [2] є фіналістами, тобто двома із трьох переможців другого раунду конкурсу NIST США [3]. Наразі вони досліджуються на третьому раунді і при позитивних результатах дослідження можуть бути прийняті в якості міжнародних постквантових стандартів ЕП. При їх побудуванні використано математичний апарат алгебраїчних решіток та відповідні методи. При подальшому дослідженні та порівнянні вказаних постквантових проєктів стандартів ЕП, як з теоретичних, так і практичних позицій, основоположним є обґрунтування вимог до параметрів та ключів та у цілому обчислення основних показників згідно прийнятих умовних та безумовних критеріїв [4, 5]. Важливим при таких дослідженнях є визначення достатності забезпечення гарантованості їх захищеності від класичних, квантових, спеціальних та атак на основі помилок [1 – 5]. Вказане може бути забезпеченим, у тому числі, засобом обґрунтованого вибору розмірів загальних параметрів та ключів [6, 7] та практичного їх побудування згідно з прийнятою моделлю безпеки [4, 8, 9]. Але при виборі розмірів загальних параметрів та ключів виникає суттєве протиріччя між властивостями проєктів стандартів ЕП Falcon та Dilithium щодо стійкості та складності перетворень. Так, збільшення розмірів загальних параметрів та ключів приводить до збільшення складності перетворень, і навпаки [6, 7]. Що стосується теоретичних методів щодо побудування загальних параметрів та ключів для ЕП Falcon та Dilithium, то вони у цілому зрозумілі, а якщо використовувати [1, 2, 6, 7], то існує можливість їх побудувати і для більш високих рівнів безпеки, наприклад запропонованих та реалізованих у [6, 7] 6 та 7 рівнів безпеки, коли може бути забезпечено 384 та 512 біт захисту від класичних атак і 192 та 256 біт захисту від квантових атак, а також захищеність від спеціальних атак та атак на основі помилок [1 – 3, 10 – 17]. У цілому, в певному плані існує необхідність уточнення теоретичних питань та безумовна необхідність практичного побудування загальних параметрів та ключів для забезпечення 6 та 7 рівнів безпеки та оцінки і порівняння вказаних проєктів ЕП згідно з запропонованими безумовними, умовними та прагматичними критеріями [6, 7]. Причому, вирішення цих проблемних питань зводиться до реалізації відповідних моделей безпеки, побудованих на основі реальних моделей порушника та загроз.

Мета статі:

1) Аналіз проблемних питань вибору розмірів параметрів та ключів для постквантових проєктів ЕП, побудованих на основі математичних методів Falcon [1] та Dilithium [2], та особливості їх реалізації, в тому числі і реалізації відповідно до прийнятої моделі безпеки.

2) Порівняльний аналіз стійкості та складності проєктів стандартів ЕП Falcon та Dilithium у залежності від розмірів параметрів та ключів, в тому числі для 6 та 7 рівнів безпеки.

3) Розробка пропозицій стосовно рішень щодо прийняття в якості національних постквантових стандартів ЕП на основі математичних методів Falcon та Dilithium.

4) Визначення впливу безумовних, умовних та прагматичних критеріїв на переваги при прийнятті рішення щодо стандартизації ЕП на основі математичних методів Falcon та

Dilithium, в тому числі з урахуванням наявності патентів та необхідності отримання ліцензій тощо.

1. Основні параметри та ключі ЕП Falcon та Dilithium

1.1. Основні параметри та ключі алгоритму Falcon та дані щодо їх розмірів

Нехай розмір відкритого ключа є (PK_SIZE), розмір ЕП (DS_SIZE), а час (складність) перевірки підпису є (CHECK_TIME). В якості загального параметру виберемо степінь поліному n. Основний розгляд проведемо для п'ятого рівня безпеки щодо захищеності від квантових атак з використанням квантових комп'ютерів [1, 2, 18 – 20]. Вище прийняті позначення вибрані з урахуванням узгодженості з програмним забезпеченням.

При дослідженнях використано наступні формули для алгоритму Falcon [1].

Довжина відкритого ключа:

PK_SIZE=$1+n*\log_2 q$, де:

n – степінь полінома, модуль, що визначає основне поле x^n+1, n=1024 (5 рівень безпеки);

q – просте число, модуль для коефіцієнтів,

q=12289, $\log_2 q$=14; DS_SIZE – довжина ЕП.

ЕП обчислюється з використанням довжини випадкового компонента nonce (40 октетів), упакованого формату поліному, коефіцієнти якого задовольняють розподілу Гауса. Довжина коефіцієнту в упакованому форматі залежить від його значення і може займати від 8 до 23 бітів. Тому щодо довжини ЕП автори використовують середнє значення.

1.2. Основні параметри та ключі алгоритмів Dilithium та Falcon, та дані щодо їх розмірів

При дослідженнях використаємо такі основні формули для алгоритму Dilithium [2].

Довжина відкритого ключа:

PK_SIZE=SEED_SIZE+$n*k*(\log_2 q–d)$, де:

SEED_SIZE – розмір випадкового компоненту для відновлення матриці A (32 октету);

n – степінь полінома, модуль, визначає основне поле x^n+1, n=256;

q – просте число, модуль для коефіцієнтів, q=8380417;

d – кількість бітів в молодшій частині коефіцієнтів поліному (d=13);

k – довжина вектору, який складається з поліномів. Кожний поліном містить n коефіцієнтів розміром $w=\log_2 q–d$ бітів; k=8, w=10.

Довжина ЕП Dilithium:

DS_SIZE=SEED_SIZE+$l*n*\log_2 2\gamma_1+(\omega+k)$, де:

SEED_SIZE – розмір компоненту для відновлення поліному по модулю 3, який складається з заданої кількості ненульових елементів;

γ_1 – задає розмір коефіцієнту поліномів $\gamma_1=2^{19}$;

l – довжина вектору, який складається з поліномів (l=7). Кожний поліном містить n коефіцієнтів розміром $w=\log_2 2\gamma_1$ (w=20);

ω – кількість ненульових елементів в бітовій матриці переносів h (ω=75);

k – кількість рядків в матриці переносів (k=8).

Результати розрахунків розмірів відкритих ключів є (PK_SIZE), ЕП (DS_SIZE) та складностей (часових) перевірки підписів (CHECK_TIME) для алгоритмів Falcon та Dilithium для 5-го рівня безпеки наведені і в табл. 1.

Таблиця 1

Порівняння основних показників для алгоритмів Falcon та Dilithium

Алгоритм	PK_SIZE (октетів)	DS_SIZE (октетів)	CHECK_TIME (тактів)
Dilithium, n=256	2592	4595	279936
Falcon, n=1024	1793	1233.29[1]	168498

Як видно з табл. 1, алгоритм Dilithium згідно з наведеними параметрами програє алгоритму Falcon.

[1] Розмір ЕП змінюється в зв'язку з особливістю кодування поліному, коефіцієнти якого задовольняють розподілу Гауса. Кількість бітів для кодування одного коефіцієнту може змінюватись в інтервалі 8 – 24 біта, 8 бітів – найбільш імовірна довжина, 24 біта – найменш імовірна.

1.3. Порівняння основних параметрів та ключів ЕП Falcon та ЕП Dilithium для алгоритмів 7 рівня безпеки

На основі математичних методів Falcon [1] та Dilithium [2] розроблено проєкти стандартів ЕП «Вершина 1» та «Вершина 2» 7 рівня криптостійкості 512 біт захисту від класичних атак та 256 біт захисту від квантових атак [6, 7]. В табл. 2 наведені основні показники для цих алгоритмів в разі збільшення степені n та застосування решти необхідних параметрів.

Таблиця 2

Порівняння основних показників для алгоритмів 7 рівня безпеки, побудованих на базі методів ЕП Falcon та Dilithium

Алгоритм	PK_SIZE (октетів)	DS_SIZE (октетів)
Прототип Dilithium, n=512	5824	10708
Прототип Falcon, n=2048	3585	4884

Також, аналогічно як із табл. 1, із табл. 2 випливає, що алгоритм Dilithium відповідно до розглянутих параметрів програє алгоритму Falcon, але генерація ключів та формування ЕП для алгоритму Falcon та його прототипів потребує від розробника застосування багатьох різних алгоритмів та форматів, що може привести к значному звуженню практичного застосування алгоритму ЕП Falcon.

2. Рекомендації щодо практичної реалізації алгоритму ЕП Falcon

Нижче розглянуто практичні рекомендації із генерації ключів для алгоритму Falcon [1]. На наш погляд, цей опис може значно полегшити реалізацію алгоритму для розробників. Ці рекомендації особливо актуальні в умовах збільшення значення степені полінома n до 2048 для забезпечення підвищення криптостійкості включно до 7 рівня [4, 5]. Для цього в якості решітки в алгоритмі застосовується NTRU решітка.

В якості секретних ключів для забезпечення 5 рівня безпеки [1, 3] застосовуються:

- короткі вектори f, g, коефіцієнти яких задовольняють розподілу Гауса та за модулем не перевищують 32;
- короткі вектори F, G, коефіцієнти яких за модулем не перевищують 128 та задовольняють NTRU рівнянню:

$$f^*G - Fg = q. \tag{1}$$

В якості відкритого ключа застосовується поліном, обчислений за формулою

$$h = f^{-1} * g \text{ в полі } Z[x]/(\phi, q), \tag{2}$$

де ϕ – поліном $x^{1024}+1$.

Для обчислення відкритого ключа h застосовується NTT формат [1], що допустимо завдяки спеціальному вибору q. Але випадковий поліном f може не мати інверсію в полі $Z[x]/(\phi, q)$, тоді необхідно сформувати нову пару поліномів f, g. Обчислення ключів f, g, коефіцієнти яких задовольняють розподілу Гауса, не тривіальна задача, особливо в умовах забезпечення захисту від спеціальних атак на основі константного часу обчислень. Обмеження на значення квадратичної норми для поліномів f, g також може привести до необхідності повтору операцій, що впливає на час генерації ключів. Перевірка квадратичної норми виконується не тільки для самих поліномів f, g, а також для поліномів, обчислених за формулою

$$f' = \frac{qf^*}{ff^*+gg^*},\ g' = \frac{qg^*}{ff^*+gg^*}. \tag{3}$$

Для обчислень (3) використовується комплексне подання чисел. Позначення f^*, g^* означають комплексно поєднані значення. Для поліномів f, g виконання операцій множення та ділення поліномів застосовує FFT формат. Таким чином, обчислення f, g, h потребує застосування двох спеціальних форматів NTT та FFT, а для множення за модулем застосовують арифметику Монтгомері [1, 23].

Для рішення NTRU рівняння в [1] запропонована така методика.

2.1. Поступовий перехід від поліномів для поля x^n+1 до поля $x+1$.

Поліном в полі x^n+1 містить n коефіцієнтів. При поступовому переході з поля x^r+1 до поля $x^{r/2}+1$ кількість коефіцієнтів зменшується в 2 рази, а значення цих коефіцієнтів збільшується. Для перетворення поліному p, заданого в полі x^r+1 в поліном $p\_$, заданого в полі $x^{r/2}+1$, виконуються наступні операції:

2.1.1. Для полінома p формуються два поліноми $p1$, $p2$ розміром $r/2$. Для цього в поліном $p1$ записуються коефіцієнти полінома p з парними, а в поліном $p2$ – з непарними номерами, тобто $p1[k]=p[2k]$; $p2[k] = p[2k+1]$; ($k = 0, \ldots r/2–1$);

2.1.2. Обчислюються поліноми $e_2=p_1^2$ та $o_2=p_2^2$ за модулем $x^{r/2}+1$;

2.1.3. Обчислюються коефіцієнти поліному – результату $p\_$ за формулами:

- $p\_[i+1]=e_2[i+1]–o_2[i]$ ($i=0\ldots r/2–2$);
- $p\_[0]=e_2[0]+o_2[i])$; ($i=t/2–1$).

Однакові операції виконуються для поліномів f, g. Позначимо відповідні результати $pf\_$ та $pg\_$ для поліномів f, g відповідно.

Для виконання операцій над коефіцієнтами необхідно виконувати опрацювання довгих чисел, ці операції виконуються без врахування модуля q. Так, при переході від поля $x^{1024}+1$ до поля x^1+1 отримаємо поліноми f', g', які містять по одному коефіцієнту значення якого для поля x^1+1 займає більше ніж 6000 бітів, тобто операції необхідно виконувати над даними, довжина яких значно перевищує довжину даних для інших несиметричних алгоритмів, наприклад, RSA.

Також в [1] не зберігають результатів перетворень, а для кожного поточного значення r для виконання зворотного перетворення обчислюють їх повторно, тому рекомендується їх зберігати, що збільшить навантаження на пам'ять, але зменшить час обчислення для кроку 3.

В результаті виконання цієї операції отримаємо два масиви поліномів, один для поліному f, інший для поліному g. Позначимо ці масиви відповідно pf, pg, $pf[0]$, що дає співпадання поліномому f, gp [0] з поліномом g. $pf[1]$ – результат перетворення полінома f в поліном в полі $x^{n/2}+1$, а $pg[1]$ – результат перетворення поліному g з поліному в полі $x^{n/2}+1$, … у поліном $pf[\log_2 n]$ – результат перетворення полінома f в поліном в полі $x+1$, $pg[\log_2 n]$ – результат перетворення полінома g в поліном в полі $x+1$. Поліноми $pf[\log_2 n]$, $pg[\log_2 n]$ містять по одному коефіцієнту. Позначимо їх lf, lg відповідно.

2.2. Для отриманих значень виконується розширений алгоритм Евкліда для вирішення діафантового порівняння і знаходження значень s, t, та найбільшого спільного дільника для lf, lg:

2.2.1. $lf*s+lg*t=gcd\ (lf, lg)$;

2.2.2. Якщо $gcd\ (lf, lg)\neq 1$, то знову переобчислюються значення f, g, h. Якщо $gcd\ (lf, lg)=1$, то після обчислення $lG=q*s$ та $lF=-q*t$ отримаємо тотожність: $lf*lG–lg*lF=q$ для поля $x+1$.

Значення lG, lF розглядаються як поліноми G, F для поля x^1+1, вони записуються замість $pf[\log_2 n]$ та $pg[\log_2 n]$ і далі необхідно для них виконати зворотне перетворення поліномів для обчислення F та G для поля x^n+1.

2.3. Виконується поступовий зворотний перехід від поля $x+1$ до полів x^2+1 … x^n+1. Вхідними даними для цього етапу є масиви поліномів pf, pg, отримані для полів $x+1$, x^2+1 … x^n+1. Результатом є поліноми F, G.

Для кожного $r = 1, 2, 4, \ldots n/2$ виконуються кроки 2.3.1–2.3.3.

2.3.1. Для зворотного переходу від поліному, заданому в полі x^r+1 до поліному p, заданому в полі $x^{2r}+1$ застосовуються поліноми pfr та pfr_2 з масиву pf для полів x^r+1, $x^{2r}+1$, поліноми pgr та pgr_2 з масиву $\underline{pg}$ для полів x^r+1, $x^{2r}+1$ відповідно. Далі, в процесі перетворення, виконуються такі операції:

- створюється 2 поліноми p_1, p_2 для поля $x^{2r}+1$. Поліном $p1$ формується з поліному *pfr* за рахунок виконання операцій:

$p_1[2k]=pfr\,[k]$; $p_1[2k]=0$; (k=0, 1, …r).

Поліном p_2 формується з поліному pg_2r за рахунок виконання таких операцій:

$p_2[2k]=pg_2r[2k]$; $p_2[2k+1]=-pg_2r[2k+1]$; (k=0, 1, …r);

- обчислюється поліном *pF* для поля $x^{2r}+1$: $pF_2r=p_1*p_2$ для поля $x^{2r}+1$;

- створюється 2 поліноми p_1, p_2 для поля $x^{2r}+1$. Поліном p_1 формується з поліному *pgr* за рахунок виконання операцій:

$p_1[2k]=pgr[k]$; $p_1[2k]=0$; (k=0, 1, …r);

Поліном p_2 формується з поліному pf_2r за рахунок виконання операцій:

$p_2[2k]=pf_2r\,[2k]$; $p_2[2k+1]=-pf_2r[2k+1]$; (k=0, 1,…r);

- обчислюється компонент масиву *pG* для поля $x^{2r}+1$: $pG_2r=p_1*p_2$ для поля $x^{2r}+1$.

2.3.2. Після переходу від поля $x^{r}+1$ до поля $x^{2r}+1$ виконується операція зменшення коефіцієнтів поліномів (reduce).

Вхідні дані для цієї операції:

значення *r*;

поліноми для поля $x^{2r}+1$: pf_2r, pg_2r;

поліноми *pF*, *pG*.

Результатом виконання таких операцій є поліноми *pF*, *pG*.

В результаті виконання цієї операції довжини коефіцієнтів поліномів зменшуються таким чином, щоб вони могли записатися в число з плаваючою точкою без втрати значущих цифр (для поточної реалізації передбачається 53 біта). Зменшення довжини виконується за рахунок виділення старших цифр поліномів.

Операція спочатку виконується для пари поліномів (pf_2r, pg_2r). Для пари визначається максимальна довжина коефіцієнту поліному (max, бітів), обчислюється параметр зсуву (scale=max–53) і виконується зсув усіх коефіцієнтів поліномів пари на задане значення в сторону молодших бітів (операція ділення коефіцієнтів на 2^{scale}). В результаті отримаємо поліноми, в яких довжина кожного коефіцієнту не перевищує 53 біта.

Обчислюється загальна квадратична норма для поліномів pf_2r, pg_2r. Позначимо її *pfg* (операція виконується для FFT формату).

2.3.3. Далі виконується цикл поступового зменшення довжини коефіцієнтів для пари (*pF*, *pG*):

- для пари визначається максимальна довжина коефіцієнту поліному (Max, бітів);

- якщо Max<max, то подальше зменшення неможливо, виконується вихід з циклу;

- обчислюється параметр зсуву (scale=Max–53) і виконується обчислення $lF_1=lF/2^{\text{scale}}$, $lG_1=lG/2^{\text{scaleт}}$;

- обчислюється загальна квадратична норма для поліномів lF_1, lG_1. Позначимо її *pFG* (формат FFT);

- обчислюється значення *d=pFG*/*pfg* (формат FFT);

- перехід від FFT формату до звичайного формату і округлення коефіцієнтів поліному *d* до цілих коефіцієнтів, якщо усі коефіцієнти дорівнюють 0, то вихід з циклу;

- обчислення $lF_1=d*lf \bmod x^{2r}+1$; $lG_1=d*lg \bmod x^{2r}+1$ (*lf*, *lg* – поліноми в полі $x^{2r}+1$);

- обчислення $lF=lF–lF_1*2^{\text{Max}-\text{max}}$; $lG=lG–lG_1*2^{\text{Max}–\text{max}}$.

Таким чином, генерація ключів передбачає застосування наступних методів і форматів:

- формування поліномів з коефіцієнтами, розміри яких обмежені і задовольняють розподілу Гауса;
- застосування NTT формату для фіксованого простого *q* для формування відкритого ключа і множення поліномів в разі застосування CRT формату представлення великих чисел;
- застосування FFT формату для виконання операції ділення для поліномів.

Перехід на початок формування ключів виконується, якщо:

- поліном *f* не має зворотного значення (1);

- квадратична норма коефіцієнтів поліномів f і g перевищує задане значення;
- квадратична норма коефіцієнтів поліномів f' і g' (2) перевищує задане значення;
- найбільший загальний дільник поліномів lf, lg не дорівнює 1.

Як показує аналіз алгоритмів формування ключів, найбільш ресурсно важкими є операції множення поліномів за модулем з коефіцієнтами – довгими числами. Зазвичай ці операції виконуються для декількох поліномів, що спрощує задачу їх паралельного виконання. Зберігання проміжних даних в загальній пам'яті, що практикують автори [1], напроти, ускладнюють задачу паралельного виконання. Тому рекомендується для тимчасового зберігання даних в функціях застосовувати локальну пам'ять.

3. Методи та алгоритми побудови загальних параметрів для ЕП типу Falcon у залежності від їх розмірів

3.1. Постановка проблеми побудування загальносистемних параметрів для Falcon N для 256 та 512 біт безпеки

Одним із основних завдань конкурсу є розробка та прийняття постквантового чи постквантових стандартів ЕП. Фіналістами другого етапу конкурсу NIST стали три механізми ЕП – Dilithium, Falcon та Rainbow [10].

Попередні дослідження показали, що серед схем ЕП на решітках дещо відрізняється від інших кандидатів Falcon [1, 21], він також має перспективи щодо прийняття в якості міжнародного стандарту ЕП. Домінуючим основним концептуальним підходом до проєктування механізму ЕП Falcon є використання перетворення типу «геш-і-підпис» [1, 21]. Перевагою такого підходу є доведена стійкість в межах моделі квантового випадкового оракула. В процесі досліджень виявлено, що навіть його аналізу присвячено значно менше робіт, ніж, щодо інших, наприклад проєкту ЕП Dilithium [2]. Крім того, як і щодо інших, при проєктуванні ЕП Falcon були прийняті обмеження щодо максимальних рівнів безпеки у вигляді максимально 256 біт проти класичного та 128 біт проти квантового криптоаналізу. Тому, як з точки зору теорії, так і практики, генерація загальносистемних параметрів для використання 384 і 512 біт безпеки проти класичного криптоаналізу та 192 і 256 біт безпеки проти квантового криптоаналізу є важливою проблемною задачею.

В якості основних властивостей розглянемо результати первинного аналізу щодо відомих атак на ЕП Falcon. При цьому врахуємо обмеження та практичні алгоритми обчислення загальносистемних параметрів та їх оптимізації для 256 та 512 біт безпеки проти класичного та не менше 128 та 256 біт проти квантового криптоаналізу. Важливим є проведення додаткових досліджень та прийняття рішення щодо реалізації рівнів 384 біт безпеки проти класичного та не менше 192 біт проти квантового криптоаналізу щодо ЕП Falcon. Сутність механізму Falcon наведено в [1, 21] і нижче не розглядається.

3.2. Аналіз атак на ЕП Falcon

Перетворення GPV [1], що використане в ЕП Falcon, вимагає, щоб геш-функція H була захищена від колізій. Це означає, що розмір солі в бітах повинен бути не меншим за 2λ, де λ – рівень безпеки, що вимагається. Проте, за умовами конкурсу NIST [4, 3] кількість запитів на вироблення ЕП (signature queries) є не більшою за $qs=2^{64}$, що вимагає цього розміру у вигляді $\lambda+\log_2(qs)$. В табл. 3 наведено вимоги щодо розмірів до для 5 – 7 рівнів безпеки ЕП Falcon.

Таблиця 3

Розмір (ентропія) початкового значення r (солі) в бітах

Безпека	Розмір r	Розмір r з врахуванням вимог NIST
256	512	320
384	768	448
512	1024	576

Попередній аналіз показав, що основними атаками щодо ЕП Falcon є атаки на відновлення особистого ключа з відкритого ключа ЕП та атаки на підробку ЕП. Розглянемо ці атаки.

3.3. Атаки на відновлення особистого ключа ЕП Falcon

Атаки на відновлення особистого ключа з відкритого ключа щодо Falcon можуть зводиться до вирішення проблеми NTRU [1]. У ряді криптосистем, стійкість яких ґрунтується на проблемі NTRU, коли поліноми f та g мають коефіцієнти з множини значень $\{0, 1, -1\}$, виникає проблема забезпечення стійкості проти комбінованих атак. Це робить можливим реалізувати різні комбінаторні атаки. Наприклад, для [24] найефективнішою атакою є гібридна атака, яка знаходить частину вектора комбінаторними шляхами, якщо не збільшувати розмір модуля n. Тому [24] для захисту від цієї атаки потрібно збільшити рівень безпеки до 2^{512}. Щодо Falcon такі атаки неможливі, оскільки поліноми f, g змінюються, точніше, вибираються згідно з нормальним розподілом з заданими параметрами. В даному випадку простір можливих значень поліномів збільшується настільки, що застосування комбінаторних методів стає неефективним. Тому залишається прямий шлях відновлення особистого ключа з відкритого засобом редукції базису решітки [1, 15 – 17]. При цьому, чим менші значення має норма найменшого вектора (f, g), тим більша криптостійкість системи. В криптосистемі Falcon поліноми генеруються над полем

$$Z_q[X]/(\phi(x)), \deg(\phi) = n$$

з математичним очікуванням рівним 0. Перетворення спираються на [1, 21], у яких детально досліджувалися можливості застосування алгоритмів вибірки нормально розподілених величин.

А основною умовою захисту ЕП Falcon від атак на відновлення особистого ключа ЕП з відкритого шляхом редукції є така

$$\left(\frac{B}{2\pi e}\right)^{\frac{2B-2n+1}{2B-2}} < \sqrt{\frac{3eB}{8n}}. \tag{4}$$

3.4. Атаки на підробку ЕП

Атаки на безпосередньо підробку ЕП можуть бути найбільш загрозливими. Тому іншим методом є атака підробки ЕП. При реалізації такої атаки потрібно знайти достатньо короткий вектор s. Відповідно, це можливо зробити, редукувавши базис так, щоб виконувалася умова [6]

$$\| b_1^* \| < \beta. \tag{5}$$

Причому оцінити $\| b_1^* \|$ можливо таким же чином, що і у попередньому випадку, тобто у такій послідовності:

$$\| b_1^* \| \approx GH(B)^{\frac{2n+1-2}{2(B-1)}} \det(\Lambda)^{\frac{1}{2n}} = GH(B)^{\frac{2n-1}{2B-2}} \sqrt{q} = \left(\frac{B}{2\pi e}\right)^{\frac{2n-1}{2B-2}} \sqrt{q}. \tag{6}$$

Отримуємо, що умовою захисту від атак на підробку підпису є

$$\left(\frac{B}{2\pi e}\right)^{\frac{2n-1}{2B-2}} \sqrt{q} < \beta. \tag{7}$$

Для практичного прийняття рішення необхідно визначитись щодо того, як обирати параметр β. Розробники ЕП Falcon N пропонують використовувати значення $\sigma = 1.55\sqrt{q}$ для

полінома $\phi = x^n + 1$ і $\sigma = 1.32 * 2^{1/4} * \sqrt{q}$ для полінома $\phi = x^n - x^{n/2} - 1$. Такий вибір σ базується на результатах роботи [1, 21], і параметр β для полінома $\phi = x^n + 1$ обчислюється як:

$$\beta = 1.2 * \sigma \sqrt{2nq}. \tag{8}$$

Для полінома $x^n - x^{n/2} - 1$ β обчислюється таким чином:

$$\beta^2 = \frac{(1.2 * \sigma * 2n\sqrt{q})^2}{n}. \tag{9}$$

Якщо підставити значення β у рівняння для оцінки захищеності від атак на підробку підпису, то також обидві сторони будуть пропорційні значенню $\sqrt{q}$. Середньоквадратичні відхилення при вибірці поліномів з нормального розподілу підібрані таким чином, щоб від q складність атаки не залежала. Проте на параметр q існує безліч інших обмежень, які впливають на його вибір.

Параметр q обирається згідно наступних міркувань [1, 9, 21, 14]:

- для захисту від алгебраїчних атак q має бути простим числом;
- якщо q буде занадто малим (порядку $q \approx n$), то будуть можливі BKW атаки;
- якщо q буде занадто великим ($q \approx n^{2.83}$), то будуть можливі атаки на підполе;
- якщо використовується поле $x^n + 1$, то для реалізації ефективного множення повинне виконуватися рівняння $q \equiv 1 \bmod 2n$;
- якщо використовується поле $x^n - x^{n/2} - 1$, то для реалізації ефективного множення повинне виконуватися рівняння $q \equiv 1 \bmod 3n$.

Необхідно відмітити, що стійкість найкращого алгоритму пошуку найменшого вектору оцінюється як $2^{0.292B}$, де B – розмір блоку при редукції. Якщо при криптоаналізі застосовувати алгоритм Гровера, то нижня оцінка класичної стійкості в 256 біт складає $2^{0.265B}$ квантової стійкості (при класичній стійкості в 256 біт). Тому, для ЕП на решітках квантова стійкість при класичній стійкості 256 біт набагато більше, ніж 128 біт [21, 22].

3.5. Точність арифметики з плаваючою крапкою

Останнім невизначеним параметром, який необхідно вибрати для забезпечення рівнів стійкості 384 та 512 біт, є необхідна точність виконання операцій у арифметиці з плаваючою крапкою. Розробники Falcon для теоретичної оцінки використовували роботу [1], проте точність обиралася з практичних експериментів. З [1] видно, що рівень безпеки λ слабо впливає на потрібну точність. Основний вплив має кількість запитів на підпис $qs = 2^{64}$, тому є надія, що 53 бітів буде достатньо. Проте, це питання є предметом подальшого дослідження.

Також до недоліку ЕП Falcon необхідно віднести використання арифметики з плаваючою крапкою. Разом з використанням деревоподібних структур це ускладнює аналіз схеми до атак сторонніми каналами. Іншою проблемою є складність реалізації на малоресурсних пристроях.

3.6. Генерація параметрів для 256, 384, 512 біт стійкості ЕП Falcon

У цілому, якщо підсумувати наведене вище, знайти параметри n, q, β можна з системи нерівностей (10) [7]:

$$\begin{cases} \left(\frac{B}{2\pi e}\right)^{\frac{2n-1}{2B-2}} \sqrt{q} < \beta \\ \left(\frac{B}{2\pi e}\right)^{\frac{2B-2n+1}{2B-2}} < \sqrt{\frac{3eB}{8n}} \end{cases}, \tag{10}$$

де параметр β визначається як $\beta = 1.2 * \sigma\sqrt{2nq}$, якщо використовується поліном x^n+1, і $\beta^2 = \frac{(1.2 * \sigma * 2n\sqrt{q})^2}{n}$, якщо використовується поліном $x^n - x^{n/2} - 1$.

На основі (10) розроблено програмне забезпечення, з використанням якого були обчислені параметри n, q та β_2 для ЕП Falcon для відповідних поліномів для 256, 384, 512 біт безпеки, що наведені в табл. 4.

Отриману криптостійкість наведено в табл. 5. Криптостійкість наведено у форматі «стійкість» λ / розмір блоку.

Оцінки криптостійкості отримувалися з розміру блоку як $0{,}265B$ та $0{,}292B$. Такий підхід вважається класичним і використовувався як авторами Dilithium, так і авторами Falcon [1, 2]. Проте, оцінки є досить грубими. В табл. 5 для деяких квантових атак для 256 і 512 sn отримані значення є трохи меншими за необхідні, проте через грубість оцінки можна вважати, що вони досягають потрібного порога. Для 256 біт згенеровані параметри співпадають із параметрами, що згенеровані авторами Falcon.

Таблиця 4

Основні загальносистемні параметри Falcon *N* для 256, 384, 512 біт безпеки

Безпека	$\phi(x)$	n	q	β_2
256	x^n+1	1024	12289	87070769
384	$x^n - x^{n/2} - 1$	1536	18433	174141539
512	x^n+1	2048	12289	200928983

Обґрунтування та сутність оптимізації за параметром β наведено в [6].

У табл. 6 наведено значення для параметрів τ, n, які можливо використовувати на практиці та відповідні значення β, причому параметр q має значення 12289.

Таблиця 5

Криптостійкість Falcon *N* до атак на основі редукції решіток

Безпека	Стійкість до відновлення ключа (класична)	Стійкість до відновлення ключа (квантова)	Стійкість до підробки підпису (класична)	Стійкість до підробки підпису (квантова)
256	273/936	248/936	269/922	244/922
384	413/1417	375/1417	430/1474	390/1474
512	554/1899	503/1899	599/2053	544/2053

Таблиця 6

Альтернативні значення параметра β

n	τ	Ймовірність повтору	$\lfloor \beta \rfloor$
1024	1.1	10^{-9}	8382
	1.08	10^{-6}	8230
	1.07	0.0007	8153
	1.05	0.064	8001
2048	1.07	10^{-9}	11710
	1.055	10^{-6}	11545
	1.045	0.0002	11436
	1.03	0.025	11272

Оптимізовані альтернативні набори загальносистемних параметрів для 256 біт класичної стійкості зведені у табл. 7.

Таблиця 7

Оптимізовані альтернативні набори загальносистемних параметрів для 256 біт класичної стійкості

n	q	$\lfloor \beta^2 \rfloor$	Підробка ЕП клас. (біт)	Підробка ЕП квант. (біт)	Відновлення ключа клас. (біт)	Відновлення ключа квант. (біт)
1024	12289	70265242 (τ=1.1)	277	252	273	248
1024	12289	67733370 (τ=1.08)	279	253	273	248
1024	12289	66484856 (τ=1.07)	280	254	273	248
1024	12289	64022669 (τ=1.05)	282	255	273	248

Оптимізовані альтернативні набори загальносистемних параметрів для 512 біт класичної стійкості зведені у табл. 8.

Таблиця 8

Оптимізовані альтернативні набори загальносистемних параметрів для 512 біт класичної стійкості

n	q	$\lfloor \beta^2 \rfloor$	Підробка ЕП клас. (біт)	Підробка ЕП квант. (біт)	Відновлення ключа клас. (біт)	Відновлення ключа квант. (біт)
2048	12289	137125015 (τ=1.07)	618	561	554	503
2048	12289	133307337 (τ=1.055)	621	563	554	503
2048	12289	130792161 (τ=1.045)	622	565	554	503
2048	12289	127064310 (τ=1.03)	625	567	554	503

4. Метод та алгоритми побудови загальних параметрів ЕП типу Dilithium у залежності від їх розмірів

В цьому підрозділі наводяться окремі результати теоретичних та практичних досліджень щодо створення постквантового ЕП Dilithium на алгебраїчній решітці [2, 21, 22]:

- обґрунтування перспективного постквантового національного стандарту ЕП Dilithium на основі алгебраїчних решіток з відхиленням;
- методи обчислення системних параметрів для ЕП Dilithium 128, 256, 384 та 512 біт рівнів безпеки;
- генерація системних параметрів ЕП Dilithium та «Вершина 1» для 128, 256, 384, та 512 біт стійкості.

4.1. Обґрунтування перспективного постквантового національного стандарту електронного підпису на основі алгебраїчних решіток з відхиленням

Одним із видів криптографічних перетворень типу ЕП, що може бути включений в національний стандарт ЕП постквантового періоду, на наш погляд, може стати ЕП на алгебраїчних решітках типу Dilithium [1, 21].

4.2. Аналіз стійкості алгоритму Dilithium проти основних атак

Наразі в постквантовій криптології актуальними є завдання забезпечення криптографічної стійкості щодо квантових атак. Вона ґрунтується не вирішенні проблеми навчання з помилками.

На основі аналізу визначено [9 – 17, 24], що стосовно LWE можливо застосування таких атак:

1. BKW, коли LWE зводиться до SIS атаки.
2. Primal attack (Search-LWE зводиться до BDD атаки).
3. Dual attack (Decision-LWE зводиться до SIS).
4. Зведення до uSVP атаки пошуку короткого вектора.

Деталі щодо кожної з атак, а також їх теоретичне обґрунтування наведено у [2, 7].

4.3. Захищеність алгоритму ЕП від атак сторонніми каналами

В процесі проведення конкурсу на постквантовий стандарт ЕП особлива вимога висунута до захищеності кандидату на ЕП від атак сторонніми каналами. Тому така проблемна задача є актуальною, в першу чергу стосовно ЕП типу Dilithium та рішення для України «Вершина 1»[7].

Дослідження стосовно алгоритму ЕП Dilithium проведено за такими параметрами [5 – 7]:

- BKZ bock-size to break SIS=475;
- BKZ bock-size to break LWE=485;
- k=5; l=4; η=5; ς=4; β=275; ω=96.

Результати отримано методом програмного моделювання. Для проведення експерименту було згенеровано 10000 ключів та виконано 10000 підписів. Результат залежності часу підпису від номеру ключа наведено на рис. 1. Для 10000 ключів максимальне відхилення від нормалізованого середнього (дисперсія) усіх вимірів часу підпису повинно знаходитися в інтервалі –5.19676<=*d*<=6.62797(%), щоб вважати, що час підпису не залежить від ключа. Номери ключів, для яких було отримано мінімальне та максимальне значення при повтореннях вимірів не повинні співпадати.

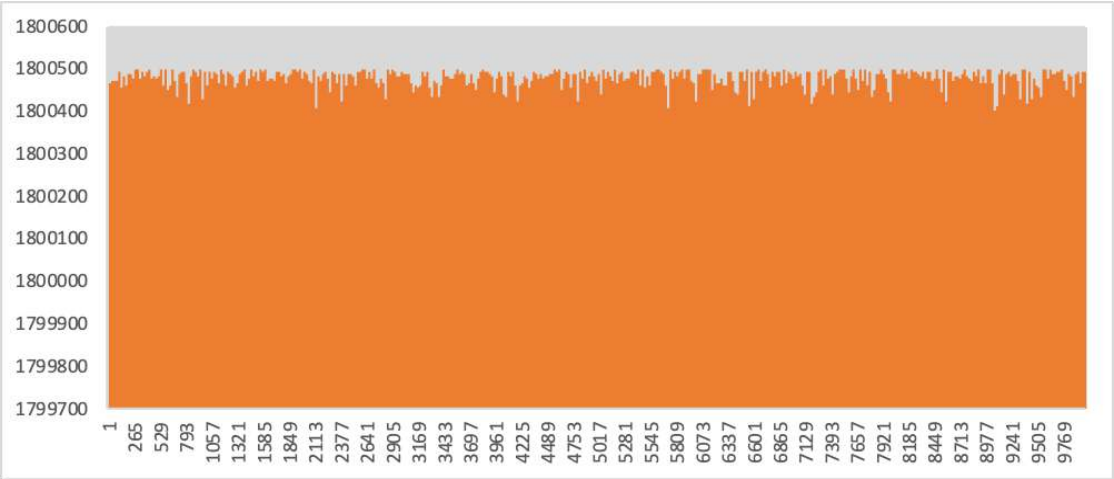


Рис. 1. Залежність часу підпису (у тактах процесору) від номеру ключа

Значення дисперсії $d\approx 2\%$, що свідчить про практично статистичну незалежність часу підпису від ключа, що є важливим з точки зору захищеності від атак сторонніми каналами.

4.4. Вибір параметрів 5-7 рівнів стійкості

Попередній аналіз показав, що значення параметрів, наведені в [7], не забезпечують при застосуванні в механізмі Dilithium стійкість ЕП від класичних атак на рівні 256 бітів. Фактично нижні оцінки стійкості наведено в рядках табл. 9 [21] з назвами Best Known Classical bit-cost і Best Known Quantum bit-cost окремо для кожної з двох задач – класичної та квантової, на складності яких базується стійкість. Це задачі SIS та LWE. Для кожної з них зазначені параметри обчислюються за формулами

$$\text{Best Known Classical bit-cost (класична атака)}=0{,}292b, \quad (11)$$

$$\text{Best Known Quantum bit-cost (квантова атака)}=0{,}265b, \quad (12)$$

де b є довжиною блоку (BKZ bock-size b to break SIS або LWE [7]). В ролі кінцевої оцінки стійкості використовується найменше з двох значень, обчислених для b, що є довжиною блоку для задачі SIS та задачі LWE відповідно.

Загальні обмеження [5 – 7], які необхідно врахувати при виборі параметрів для забезпечення стійкості на рівні $\lambda \in \{256, 384, 512\}$ бітів.

1. Геш-функція CRH, що використовується у ЕП, повинна бути стійкою до колізій [2]. Отже, довжина її вектору значень повинна бути, як мінімум, 2λ бітів. (Зокрема, при λ=256 функція CRH повинна приймати значення довжини 512, а ні 384, як в оригінальному методі [7]).

2. Криптографічна функція H, яка використовується, повинна бути стійкою відносно знаходження другого прообразу і, отже, приймати, принаймні, 2λ різних значень, що є поліномами з кільця R_q, які мають коефіцієнти 0, 1, –1 та містять точно h ненульових коефіцієнтів (зауважимо, що кількість таких поліномів дорівнює $2^h\binom{n}{h}$).

При λ=256 в [2, 7] рекомендується використовувати параметри n=256, h=60, і умова

$$2^h \binom{n}{h} \geq 2^\lambda \tag{13}$$

виконується.

При $\lambda \in \{384, 512\}$ та n=256 забезпечити виконання умови неможливо, якщо $h \leq n/2$=128. Отже, треба збільшити n до 512; при цьому числа q, γ_1, γ_2 можна залишити такими самими як в [2, 7]:

$$q=2^{23}-2^{13}+1,$$
$$\gamma_1=(q-1)/16,\ \gamma_2=\gamma_1/2. \tag{14}$$

3. Довжини векторів ρ та K, що використовуються, повинні бути не менше, ніж λ.

Таким чином, для забезпечення стійкості схеми ЕП на рівні $\lambda \in \{256, 384, 512\}$ необхідно:

1) використовувати геш-функцію CRH, значеннями якої є двійкові вектори довжини 2λ [7];

2) використовувати двійкові вектори ρ та K, що мають довжину λ;

3) покласти n=256, якщо λ=256; n=512, якщо $\lambda \in \{384, 512\}$;

4) вибрати просте число $q \equiv 1 \bmod 2n$ та обчислити γ_1, γ_2 за формулою (14);

5) обчислити вагу c як найменше натуральне h.

Зауважимо, що при n=512, $q=2^{23}-2^{13}+1$ час формування підпису може виявитися занадто великим; у цьому випадку треба збільшити q (приблизно в два рази).

4.5. Генерація загальносистемних параметрів ЕП Dilithium для 128, 256, 384, 512 біт стійкості

Для генерації системних параметрів удосконаленого ЕП Dilithium використаємо результати аналізу відомих атак на криптосистему, що наведені вище, та встановимо умови, за яких забезпечується захист від них. Основними атаками є [7]:

- відновлення особистого ключа на основі відкритого ключа;
- засобом підробки ЕП.

Враховуючи наведені критерії, алгоритм генерації загальносистемних параметрів виглядає наступним чином [2, 7, 21]:

1. Визначити потрібний рівень безпеки $\lambda \in \{256, 384, 512\}$.
2. Обрати значення N. Якщо λ=256, то N=256, інакше N=512.
3. Обрати значення q: q=8389417.
4. Обчислити γ_1 та γ_2 за формулами $\gamma_1 = (q-1)/16$, $\gamma_2 = \gamma_1/2$.
5. Обчислити значення η. За замовченням встановити η=2. На наступних кроках значення буде уточнене.
6. Встановити значення (k, l)=(2, 1).
7. Обчислити λ_1-стійкість до Primal Attack. Якщо стійкість менша за λ, то оновити параметри (k, l)=(k+1, l+1) та повернутися до кроку 7 або збільшити η та повернутися до кроку 7.
8. Обчислити λ_2-стійкість до Dual Attack. Якщо стійкість менша за λ, то оновити параметри (k, l)=(k+1, l+1) та повернутися до кроку 7 або збільшити η та повернутися до кроку 7.
9. Обчислити λ_3-стійкість до SIS з ζ_1. Якщо стійкість менша за λ, то оновити параметри (k, l)=(k+1, l+1) та повернутися до кроку 7 або збільшити k=k+1 та повернутися до кроку 7.
10. Обчислити λ_4-стійкість до SIS з ζ_2. Якщо стійкість менша за λ, то оновити параметри (k, l)=(k+1, l+1) та повернутися до кроку 7 або збільшити k=k+1 та повернутися до кроку 7.
11. Обчислити h як найбільше ціле, для якого виконується нерівність $2^h \binom{n}{h} \geq 2^\lambda$.
12. Обчислити d як найбільше ціле, для якого виконується нерівність $2^{d-1}h+1 \leq 2\gamma_2$.

13. Встановити $\beta=\eta h$ та зменшувати β, щоб ймовірність повтору циклу була достатньо малою.

14. Обчислити w=0,08nk (цей крок не впливає на криптостійкість та його можливо оптимізувати).

У табл. 9 наведено значення параметрів для удосконаленого ЕП Dilithium.

Таблиця 9

Значення параметрів для 256, 384, 512 біт стійкості

Набір	N, q	γ_1	γ_2	k, l	η	β	d	h	ω
256	(256,8380417)	523776	261888	(9,8)	2	144	14	60	184
384	(512, 8380417)	523776	261888	(7,5)	5	100	13	77	286
512	(512, 8380417)	523776	261888	(9,8)	2	74	13	118	368

Ймовірність повтору циклу при цьому складає для 256 біт – 0,15442678312246608, для 384 біт – 0,15609624568669475 і для 512 біт – 0,15247678668181552.

Результати оцінки криптостійкості для удосконаленого ЕП Dilithium (в бітах) з використанням параметрів з табл. 9 наведено в табл. 10.

Таблиця 10

Оцінки криптостійкості для удосконаленого ЕП Dilithium

Набір	Primal атака (клас.)	Primal атака (квант.)	Dual атака (клас.)	Dual атака (квант.)	SIS (класич.)	SIS (квант.)
256	298	270	296	269	293	266
384	440	399	438	397	503	456
512	582	*527*	579	525	590	535

5. Особливості застосування методики оцінки та результати порівняння постквантових алгоритмів ЕП на алгебраїчних решітках

В даному підрозділі наводяться пропозиції щодо застосування методики з оцінки та порівняння перспективних криптографічних перетворень типу ЕП, в першу чергу щодо криптографічної стійкості.

В табл. 11 наведені характеристики обраних для порівняння алгоритмів (значення швидкості криптоперетворень та генерації ключів наведено в тактах). В порівнянні приймали участь проєкти стандартів «Вершина 1» та «Вершина 2», а також алгоритм Dilithium, який за попередніми дослідженнями мав кращі результати серед постквантових алгоритмів підпису, що засновані на перетвореннях на алгебраїчних решітках. Стійкість алгоритмів «Вершина» 128 біт відповідає 3-му рівню стійкості NIST, 256 – 5-му, тому пропорційно для виконання порівняння згідно шкали оцінок попарного порівняння параметрам 384 був наданий 7-й рівень, а 512 – 9-й.

Таблиця 11

Характеристики алгоритмів ЕП, що засновані на перетвореннях на алгебраїчних решітках

Алгоритми	$I_{ст.}$	$l_{в.к.}$	$l_{о.к.}$	$l_{рез.}$	$T_{пр.}$	$T_{зв.}$	$T_{гк.}$
Dilithium_round3_sec2	2	1 312	3 504	2 420	259 172	118 412	124 031
Dilithium_round3_sec3	3	1 952	3 856	3 293	428 587	179 424	256 403
Dilithium_round3_sec5	5	2 592	5 792	4595	538 986	279 936	298 050
Вершина_1_128	3	1 472	3 488	2 693	133 340	109 818	90 328
Вершина_1_256	5	2 624	5 792	5 345	259 103	233 712	229 669
Вершина_1_384	7	4 528	9 088	6762	411 040	398 029	317 324
Вершина_1_512	9	5 824	11 008	10708	643 744	620 989	485 471
Вершина_2_128	3	897	4097	666	655 672	139 620	33 696 000
Вершина_2_256	5	1 793	8193	1 280	1 338 825	285 714	107 055 000
Вершина_2_512	9	3 585	5121	2 515	2 600 053	265 416	28493603229

В табл. 12 наведені результати досліджень – відносна перевага алгоритмів ЕП, що отримана методом попарних порівнянь за кожною з характеристик.

Таблиця 12

Відносна перевага алгоритмів ЕП за кожною з характеристик

Алгоритми	$I_{ст.}$	$l_{в.к.}$	$l_{о.к.}$	$l_{рез.}$	$T_{пр.}$	$T_{зв.}$	$T_{гк.}$
Dilithium_round3_sec2	0,0198	0,1770	0,1816	0,1131	0,1583	0,1857	0,2090
Dilithium_round3_sec3	0,0299	0,0965	0,1475	0,0606	0,0849	0,0984	0,1082
Dilithium_round3_sec5	0,0697	0,0655	0,0768	0,0507	0,0666	0,0506	0,0915
Вершина_1_128	0,0299	0,1395	0,1816	0,0800	0,3006	0,2195	0,2696
Вершина_1_256	0,0697	0,0655	0,0768	0,0339	0,1583	0,0716	0,1388
Вершина_1_384	0,1453	0,0327	0,0407	0,0261	0,0975	0,0348	0,0797
Вершина_1_512	0,2681	0,0233	0,0296	0,0173	0,0479	0,0218	0,0608
Вершина_2_128	0,0299	0,2487	0,1212	0,3211	0,0479	0,1466	0,0192
Вершина_2_256	0,0697	0,1108	0,0467	0,1989	0,0238	0,1130	0,0146
Вершина_2_512	0,2681	0,0406	0,0973	0,0984	0,0143	0,0581	0,0086

На рис. 2 відображено гістограму загальної відносної переваги алгоритмів ЕП з урахуванням вагових коефіцієнтів характеристик.

Як видно, найбільшу перевагу має алгоритм «Вершина 1» з параметрами стійкості 128 біт, для більш стійких параметрів перевага вже у алгоритму «Вершина 2». «Вершина 2» на даному етапі програє через свою низьку швидкодію, якщо її реалізація буде більш оптимізована, то вже «Вершина 2» буде на першому місці.

Таким чином, зроблено порівняння алгоритмів, що пройшли до третього етапу NIST, а також алгоритмів проєктів стандарту «Вершина 1» та «Вершина 2». При порівнянні використовувалися два методи порівняння для отримання більш точної оцінки в залежності від вимог до алгоритмів ЕП.

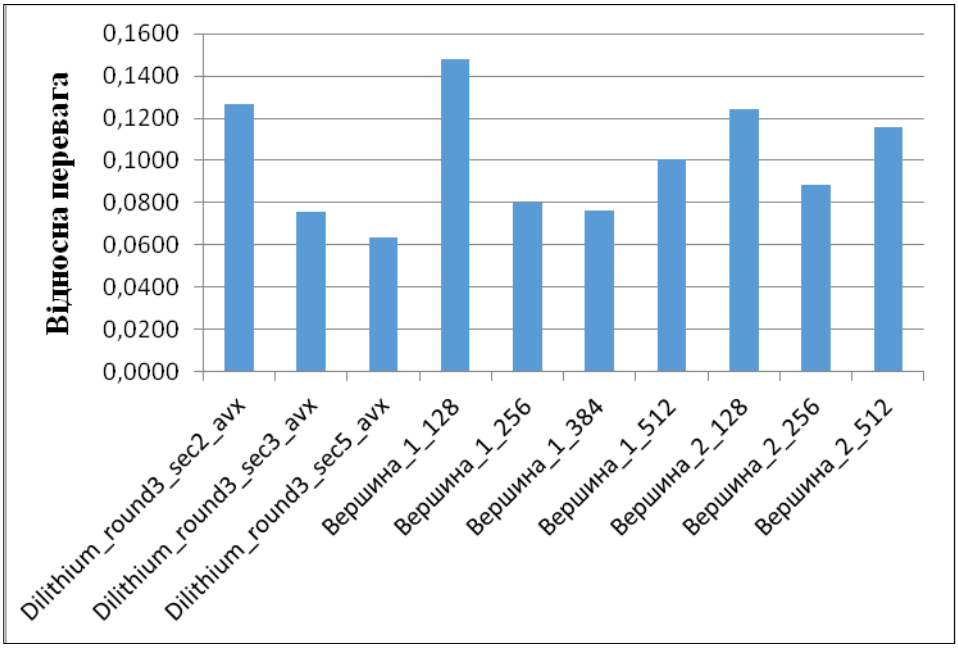


Рис. 2. Переваги алгоритмів ЕП

Висновки

1. Математичні методи Falcon та Dilithium є теоретичними основами для створення постквантових проєктів стандартів ЕП Falcon та Dilithium. Ці ЕП є фіналістами міжнародного конкурсу NIST США. Наразі вони досліджуються на третьому раунді і при позитивних результатах дослідження можуть бути прийняті в якості міжнародних постквантових стандартів ЕП. При їх побудування використано математичний апарат алгебраїчних решіток та відповідні методи.

2. Як з теоретичних, так і практичних позицій, основоположним щодо цих ЕП є обґрунтування вимог до параметрів та ключів та їх побудування за умови достатності забезпечення гарантованості їх захищеності від класичних, квантових, спеціальних та атак на основі помилок

3. При виборі розмірів загальних параметрів та ключів виникає суттєве протиріччя між властивостями проєктів стандартів ЕП Falcon та Dilithium щодо стійкості та складності перетворень. Збільшення розмірів загальних параметрів та ключів приводить до збільшення складності перетворень, і навпаки.

4. Що стосується теоретичних методів побудування загальних параметрів та ключів для ЕП Falcon та Dilithium, то вони у цілому зрозумілі, а якщо використовувати [1, 2, 6, 7], то існує можливість їх побудувати і для більш високих, але обґрунтованих рівнів безпеки – 6-го та 7-го, коли може бути забезпечено 384 і 512 біт захисту від класичних атак та 192 і 256 біт захисту від квантових атак, а також захищеність від спеціальних атак та атак на основі помилок.

5. Результати розрахунків розмірів відкритих ключів (PK_SIZE), електронних підписів (DS_SIZE) та складностей (часових) перевірки підписів (CHECK_TIME) для алгоритмів ЕП Falcon та Dilithium для 5-го рівня безпеки наведені в табл. 1, які свідчать про те, що ЕП Dilithium для 5-го рівня безпеки програє ЕП Falcon.

6. На основі математичних методів Falcon та Dilithium розроблено проєкти стандартів ЕП «Вершина 1» та «Вершина 2» 7-го рівня криптостійкості 512 біт захисту від класичних атак та 256 біт захисту від квантових атак. В табл. 2 наведені основні показники для цих алгоритмів в разі збільшення степені n та застосування решти необхідних параметрів.

7. Генерація ключів та формування електронного підпису для алгоритму Falcon та його прототипів потребує від розробника застосування багатьох різних алгоритмів та форматів, що може привести к значному звуженню практичного застосування алгоритму ЕП Falcon.

8. У [1] не зберігають результатів перетворень, а для кожного поточного значення r для виконання зворотного перетворення обчислюють їх повторно, тому рекомендується їх зберігати, що збільшить навантаження на пам’ять, але зменшить час обчислення. Для тимчасового зберігання даних в функціях рекомендується застосовувати локальну пам'ять.

9. Попередній аналіз дозволяє зробити висновки, що Falcon все таки має перспективи щодо прийняття в якості міжнародного стандарту ЕП. Домінуючим основним концептуальним підходом до проєктування механізму ЕП Falcon є використання перетворення типу «геш-і-підпис». Перевагою такого підходу є доведена стійкість в межах моделі квантового випадкового оракула.

10. Безпосередньо ЕП Falcon та його аналізу присвячено значно менше робіт, ніж щодо інших, наприклад проєкту ЕП Dilithium. Крім того, як і щодо інших, при проєктуванні ЕП Falcon були прийняті обмеження щодо максимальних рівнів безпеки у вигляді максимально 256 біт проти класичного та 128 біт проти квантового криптоаналізу.

11. Попередній аналіз показав, що основними атаками щодо ЕП Falcon є атаки на відновлення особистого ключа з відкритого ключа ЕП та атаки на підробку ЕП.

12. Атаки на відновлення особистого ключа з відкритого ключа щодо Falcon можуть зводиться до вирішення проблеми NTRU. У ряді криптосистем, стійкість яких ґрунтуються на проблемі NTRU, коли поліноми f та g мають коефіцієнти з множини значень {0, 1, –1}, виникає проблема забезпечення стійкості проти комбінованих атак.

13. Атаки безпосередньої підробки ЕП Falcon можуть бути найбільш загрозливими. Тому іншим впливом є атака підробки ЕП. При реалізації такої атаки потрібно знайти достатньо короткий вектор s. Відповідно, це можливо зробити, редукувавши базис так, щоб виконувалася необхідні умови.

14. До недоліку ЕП Falcon необхідно віднести використання арифметики з плаваючою крапкою. Разом з використанням деревоподібних структур це ускладнює аналіз методу до

атак сторонніми каналами. Іншою проблемою є складність реалізації перетворень на малоресурсних пристроях.

15. Одним із видів криптографічних перетворень типу ЕП, що може бути включений в національний стандарт ЕП постквантового періоду, на наш погляд, може стати ЕП на алгебраїчних решітках типу Dilithium.

16. Наведені у табл. 11 та 12 результати порівняння проєктів стандартів ЕП «Вершина 1» та «Вершина 2», а також алгоритму Dilithium, показали, що кращі результати серед постквантових алгоритмів ЕП, що засновані на перетвореннях на алгебраїчних решітках, мають проєкти ЕП «Вершина 1» та «Вершина 2».

17. Кращі показники у алгоритмів «Вершина 1» та «Вершина 2» («Вершина 2» на даному етапі програє через свою низьку швидкодію, якщо її реалізація буде більш оптимізована, то вже «Вершина 2» буде на першому місці).

Список літератури:

1. Falcon: Fast-Fourier Lattice-based Compact Signatures over NTRU. Round 3 Submissions. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/projects/post-quantum-cryptography/round-3-submissions.
2. Léo Ducas Crystals-Dilithium: Algorithm Specifications and Supporting Documentation. Round 3 Submissions. [Електронний ресурс]. Режим доступу: https://csrc.nist.gov/Projects/post-quantum-cryptography/round-3-submissions.
3. Gorjan Alagic NISTIR 8309 Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process / Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, John Kelsey, Yi-Kai Liu, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone. Режим доступу: https://doi.org/10.6028/NIST.IR.8309.
4. Chen L Report on Post-Quantum Cryptography / Chen L, Jordan S, Liu Y-K, Moody D, Peralta R, Perlner RA, Smith-Tone D // (National Institute of Standards and Technology, Gaithersburg, MD), NIST Interagency or Internal Report (IR) 8105 (2016). Режим доступу: https://doi.org/10.6028/NIST.IR.8105.
5. Горбенко І. Д. Методи, методика та результати порівняльного аналізу кандидатів на постквантовий стандарт електронного підпису / І. Д. Горбенко, О. Г. Качко, М. В. Єсіна, В. А. Пономар // XX Ювілейна Міжнар. наук.-практ. конференція "Безпека інформації в інформаційно-телекомунікаційних системах", 22-24 травня, 2018, м. Буча. С. 96-97.
6. Горбенко І. Д. Генерація загальносистемних параметрів для криптосистеми Falcon для 256, 384, 512 біт безпеки / І. Д. Горбенко, С.О. Кандій, М.В. Єсіна, Є.В. Острянська // Радіотехніка. 2020. Вип. 202. С. 57-63.
7. Горбенко І. Д. Методи обчислення системних параметрів для електронного підпису «Crystals-Dilithium» 128, 256, 384 та 512 біт рівнів безпеки / І. Д. Горбенко, А. М. Олексійчук, О. Г. Качко, Ю. І. Горбенко, М. В. Єсіна, С. О. Кандій // Радіотехніка. 2020. Вип. 202. С. 5-28.
8. Yesina Maryna Comparative Analysis of Key Encapsulation Mechanisms / Maryna Yesina, Mikolaj Karpinski, Volodymyr Ponomar, Yurij Gorbenko, Tomasz Gancarczyk, Uliana Iatsykovska // Proceedings of the 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS). September 18-21, 2019, Metz, France. Volume 1. – P. 7-12.
9. Горбенко Ю. І. Моделі загроз щодо асиметричних криптоперетворень перспективного електронного підпису / Ю.І. Горбенко, М.В. Єсіна, В.В. Онопрієнко, Г.А. Малєєва // Радіотехніка. 2020. Вип. 202. С. 72-78.
10. Горбенко Ю. І. Аналіз стійкості постквантового електронного підпису Dilithium до атак на помилки / Ю.І. Горбенко,О.С. Дроздова // Радіотехника. 2020. Вип. 202. С. 49 – 56.
11. Daniel J. Bernstein, Tanja Lange, Christiane Peters Attacking and defending the McEliece cryptosystem. [Електронний ресурс]. Режим доступу: https://link.springer.com/chapter/10.1007/978-3-540-88403-3_3.
12. Martin Albrecht, Shi Bai, Leo Ducas A Subfield Lattice Attack on Overstretched NTRU Assumptions. In Matthew Robshaw and Jonathan Katz, editors, Advances inCryptology – CRYPTO 2016, pages 153–178, Berlin, Heidelberg, 2016. Springer BerlinHeidelberg.
13. Paul Kirchner, Pierre-Alain Fouque Revisiting Lattice Attacks on Overstretched NTRU Parameters. In Jean-Sebastien Coron and Jesper Buus Nielsen, editors, Advancesin Cryptology – EUROCRYPT 2017, pages 3-26, Cham, 2017. Springer InternationalPublishing.
14. Vadim Lyubashevsky, Daniel Wichs Simple lattice trapdoor sampling from a broad class of distributions. In Jonathan Katz, editor, PKC 2015, volume 9020 of LNCS, pages 716 – 730, Gaithersburg, MD, USA, March 30 – April 1, 2015. Springer, Heidelberg, Germany.
15. Martin R. Albrecht On the complexity of the BKW algorithm on LWE / Martin R. Albrecht, Carlos Cid, Jean-Charles Faugère, Robert Fitzpatrick, Ludovic Perret //. Designs, Codes and Cryptography, 74: 325-354, 2015.
16. Ronald Cramer, Léo Ducas, Benjamin Wesolowski Short stickelberger class relations and application to ideal-SVP. In Coron and Nielsen, pages 324-348.
17. Avrim Blum, Adam Kalai, Hal Wasserman Noise-tolerant learning, the parity problem, and the statistical query model. Journal of the ACM, 50(4): 506-519, July 2003.
18. Квантовый компьютер. [Електронний ресурс]. Режим доступу: http://www.tadviser.ru/index.php/.
19. Каптьол Є. Ю. Аналіз можливостей та особливості програмування задач криптології на квантовому комп'ютері /Є. Ю. Каптьол, І.Д. Горбенко // Радіотехніка. 2020. Вип. 202. С. 37-48.

20. Горбенко Ю. І. Аналіз можливостей квантових комп'ютерів та квантових обчислень для криптоаналізу сучасних криптосистем / Ю. І. Горбенко, Р. С. Ганзя // Східно-європейський журнал передових технологій. 2014. № 1/9 (67). С. 8-15.

21. Vadim Lyubashevsky CRYSTALS-Dilithium. Submission to the NIST Post-Quantum Cryptography Standardization [NIS] / Vadim Lyubashevsky, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Peter Schwabe, Gregor Seiler, Damien Stehlé //, 2017. Режим доступу: https://pq-crystals.org/dilithium.

22. Олексійчук А. М. Обґрунтування перспективного постквантового національного стандарту електронного підпису на основі решіток / А. М. Олексійчук, В. А. Кулібаба, М. В. Єсіна, С. О. Кандій, Є. В. Острянська, І. Д. Горбенко // Радіотехніка. 2020. Вип. 200. С. 5-14.

23. Качко О. Г. Оптимізація алгоритму множення поліномів для NTRU – побітних алгоритмів / О. Г. Качко, Ю. І Горбенко, В. А. Пономар, М. В. Єсіна, С. О. Кандій // Радіотехніка. 2020. Вип. 200. С. 15-24.

24. Nick Howgrave-Graham A hybrid lattice-reduction and meet-in-the-middle attack against NTRU. In Alfred Menezes, editor, CRYPTO 2007, volume 4622 of LNCS, pages 150–169, Santa Barbara, CA, USA, August 19-23, 2007. Springer, Heidelberg, Germany.

Надійшла до редколегії 05.04.2021

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, головний конструктор АТ «Інститут інформаційних технологій»; 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, начальник відділу програмування АТ «Інститут інформаційних технологій», Україна, e-mail: iit@iit.kharkov.ua, ORCID: https://orcid.org/0000-0001-9249-0497

Потій Олександр Володимирович – д-р техн. наук., професор, заступник Голови Державної служби спеціального зв'язку та захисту інформації; Україна; e-mail: potav@ua.fm; ORCID: https://orcid.org/0000-0002-2366-0541

Олексійчук Антон Миколайович – д-р техн. наук, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "КПІ", професор спеціальної кафедри №1; Україна; e-mail: alex-dtn@ukr.net; ORCID: https://orcid.org/0000-0003-4385-4631

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

Єсіна Марина Віталіївна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, старший викладач кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: rinayes20@gmail.com; ORCID: https://orcid.org/0000-0002-1252-7606

Стельник Ігор Валерійович – Адміністрація Державної служби спеціального зв'язку та захисту інформації України, заступник директора Департаменту.

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 621.391:519.2 **DOI:10.30837/rt.2021.2.205.02**

А.М. ОЛЕКСІЙЧУК, д-р техн. наук, О.С. ШЕВЧУК

ОЦІНКИ ЕФЕКТИВНОСТІ АТАК НА ОСНОВІ ПІДІБРАНИХ ВІДКРИТИХ ТЕКСТІВ НА КРИПТОСИСТЕМУ РАО-НАМА НАД СКІНЧЕННОЮ АБЕЛЕВОЮ ГРУПОЮ

Вступ

Криптосистема Рао – Нама [1] являє собою симетричну версію кодової криптосистеми Мак-Еліса [2], запропоновану з метою позбутися слабкостей, притаманних найпершим симетричним кодовим схемам шифрування [3, 4]. Майже одразу після опублікування цієї криптосистеми з'явилися атаки на неї на основі підібраних відкритих текстів [5, 6], що привело до появи різноманітних удосконалень та модифікацій оригінальної криптосистеми (див. [7], де можна знайти огляд публікацій, присвячених новітнім версіям криптосистеми Рао – Нама).

В [6] запропоновано так звану еквіваленту криптосистему Рао – Нама, яка описується рівнянням шифрування (над полем з двох елементів) вигляду $c = mA + e$, де відкритий текст m і шифрований текст c є двійковими векторами довжини k і n відповідно, секретний ключ A є двійковою матрицею розміру $k \times n$ і рангу k, а e являє собою випадковий двійковий вектор, який вибирається з певної множини M потужності N, яка зберігається в секреті. При цьому для забезпечення однозначності розшифрування різні елементи множини M повинні мати різні синдроми (тобто бути різними за модулем коду з твірною матрицею G). Така версія криптосистеми Рао – Нама не відрізняється за стійкістю від оригінальної криптосистеми з [1], проте характеризується помітно меншою довжиною ключа [6, с. 126]. В [5, 6] описано низку атак на основі підібраних відкритих текстів як на оригінальну, так і на еквівалентну криптосистеми Рао – Нама, але аналіз ефективності цих атак потребує подальших досліджень.

Метою статті є отримання оцінок ефективності (трудомісткості при заданій верхній межі ймовірності помилки) атак на криптосистему, яка узагальнює еквівалентну схему шифрування Рао – Нама на випадок скінченної абелевої групи (зауважимо, що необхідність дослідження подібних версій криптосистеми Рао – Нама обумовлена їх розглядом у нещодавніх публікаціях; див. [7]). Представлено дві атаки, які будуються на основі підібраних відкритих текстів. Перша з них не згадується у відомих авторам цієї статті працях і за певних (визначених нижче) умов дозволяє відновлювати секретний ключ еквівалентної криптосистеми Рао – Нама за $O(kN^2)$ операцій в середньому.

Друга атака являє собою узагальнено-спрощений варіант відомої атаки Стройка-ван Тілбурга [5, 6]. Показано, що складність цієї атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи, над якою розглядається криптосистема Рао – Нама. Отримано оцінку ймовірності тривіальності стабілізатора за умови випадкового вибору множини M за рівноймовірною схемою. З цієї оцінки випливає, що атака Стройка-ван Тілбурга є в середньому помітно більш ефективною в порівнянні із найгіршим випадком, розглянутим в [5, 6].

1. Означення криптосистеми

Нехай G – скінченна абелева група порядку $q > 1$. Секретними ключами криптосистеми, що розглядається, є впорядковані набори $((g_1,...,g_k), M, \sigma)$, де $(g_1,...,g_k) \in G^k$, $M = \{z_1, \dots, z_N\} \subseteq G$, $\sigma : G \to G'$ – епіморфізм груп, ядро якого співпадає з підгрупою H, породженою елементами $g_1,...,g_k$. При цьому вважається, що виконані такі умови:

а) група H є прямою сумою циклічних підгруп, породжених елементами $g_1,...,g_k$ відповідно, тобто кожен елемент $g \in H$ допускає однозначне представлення у вигляді $g = m_1 g_1 + \cdots + m_k g_k$, де $m_i \in \overline{0, q_i - 1}$, q_i – порядок елемента g_i групи G, $i \in \overline{1, k}$;

б) елементи $z_1, \ldots, z_N$ множини M належать різним суміжним класам по підгрупі H, відмінним від цієї підгрупи.

Елементи з M зберігаються у вигляді таблиці $((\sigma(z_i), z_i) : i \in \overline{1, N})$, де кожен елемент z_i записано за адресою $\sigma(z_i)$. Зауважимо, що на підставі умови *б*) усі такі адреси є ненульовими попарно різними елементами групи G'.

За означенням множина відкритих текстів криптосистеми складається з усіх наборів $(m_1, \ldots, m_k)$, де $m_i \in \overline{0, q_i - 1}$, $i \in \overline{1, k}$. Для зашифрування відкритого тексту на ключі $((g_1,...,g_k), M, \sigma)$ з множини M вибирається випадковий рівноймовірний елемент z і обчислюється шифрований текст

$$c = m_1 g_1 + \cdots + m_k g_k + z. \qquad (1)$$

Для розшифрування цього шифротексту законний отримувач обчислює значення $\sigma(c)$, яке співпадає з елементом $\sigma(z)$ внаслідок означення епіморфізму σ. Далі отримувач знаходить елемент z за елементом $\sigma(z)$, використовуючи таблицю для зберігання множини M, обчислює повідомлення $c - z = m_1 g_1 + \cdots + m_k g_k$, за яким відновлює відкритий текст $(m_1, \ldots, m_k)$, спираючись на умову (а).

Як приклад розглянемо окремий випадок описаної криптосистеми, що будується над групою $G = (\mathbf{GF}(2)^n, \oplus)$ і являє собою еквівалентну версію класичної криптосистеми Рао – Нама, визначену в [6]. В цьому випадку $g_1,...,g_k$ є лінійно незалежними (над полем з двох елементів) двійковими векторами довжини n, які утворюють $k \times n$ матрицю A, що є твірною матрицею деякого двійкового лінійного коду C. Епіморфізм σ групи G в групу $G' = (\mathbf{GF}(2)^{n-k}, \oplus)$ визначається за формулою $\sigma(z) = Bz^T$, де B є перевірочною матрицею коду C, а z^T позначає вектор, транспонований до $z \in G$ (таким чином, в даному випадку $\sigma(z)$ є просто синдромом двійкового слова z). Згідно з формулою (1) шифротекст c отримується в результаті кодування відкритого тексту $(m_1, \ldots, m_k)$ кодом C з подальшим "накладанням" вектора помилок $z \in M$, який можна "зняти" на приймальному кінці системи зв'язку, знаючи множину M та синдром $\sigma(z)$.

В роботі Рао і Нама [1] розглянуто два способи формування множини M, перший з яких полягає у використанні певних заздалегідь визначених векторів, що мають вагу (Геммінга) приблизно $n/2$, а другий – у випадковому виборі цих векторів з урахуванням умови *б*). В [5] показано, що перший варіант не забезпечує належну стійкість криптосистеми Рао – Нама (внаслідок невеликої кількості та простої будови зазначених векторів), проте другий варіант є більш змістовним та потребує додаткових досліджень.

Для наведеної вище криптосистеми над скінченною абелевою групою G зазначений другий спосіб формування множини M узагальнюється таким чином.

Позначимо $r = |G| \cdot |H|^{-1}$ число суміжних класів групи G по підгрупі H. Тоді для формування випадкової множини M спочатку з імовірністю $\binom{r-1}{N}^{-1}$ вибирається множина з N ненульових суміжних класів по підгрупі H, після чого в кожному суміжному класі з імовірністю $|H|^{-1}$ вибирається один випадковий елемент, причому всі ці елементи вибираються

незалежно один від одного. Іншими словами, вважається, що кожна множина M з N елементів, яка задовольняє умові *б*), має однакову ймовірність $p(M)=\binom{r-1}{N}^{-1}|H|^{-N}$. В подальшому така схема формування множини M називається *рівноймовірною*.

Нижче розглядаються атаки на описану криптосистему, при проведенні яких вважається, що супротивник має доступ до оракула зашифрування з невідомим ключем $((g_1,...,g_k), M, \sigma)$. Метою атак є відновлення окремих частин ключа – множини M та впорядкованого набору $(g_1,...,g_k)$. *Ефективність атаки* характеризується її трудомісткістю (середньою або у найгіршому випадку) при заданій верхній межі ймовірності помилки атаки.

2. Алгоритм відновлення множини M

В [5, 6] (для випадку класичної криптосистеми Рао – Нама) описано природний *алгоритм відновлення множини* M, який полягає в наступному:

– вибрати натуральне число t;

– отримати набір $c_1,...,c_t$ шифрованих текстів, подаючи t разів на вхід оракула зашифрування відкритий текст $(m_1,\dots,m_k)=(0,\dots,0)$;

– покласти $M=\{c_1,...,c_t\}$.

Позначимо $t_0=t_0(N)$ найменше натуральне t, для якого множина, що складається з усіх різних елементів $c_1,...,c_t$, дорівнює M. Тоді t_0 є випадковою величною, що дорівнює найменшому числу частинок, які потрібно кинути в N скриньок для того, щоби усі скриньки були заповнені. (Дійсно, елементи множини M можна розглядати як скриньки, а шифровані повідомлення, отримані шляхом зашифрування нульового відкритого тексту, як частинки, що кидаються у скриньки випадково рівноймовірно та незалежно одна від одної) [8].

Для математичного сподівання випадкової величини t_0 справедливі співвідношення $\mathbf{E}t_0=N\sum_{i=1}^{N}1/i<N(\ln N+C+N^{-1})$, де $C=0{,}5772...$ – константа Ойлера [8, с. 18 ; 9, с. 108]. Це означає, що при $N>2$ середнє число зашифровувань, потрібних для успішного відновлення множини M, не перевищує $N(\ln N+1)$.

Для оцінки трудомісткості наведеного алгоритму відновлення множини M у найгіршому випадку можна скористатися відомою формулою для числа сюр'єктивних відображень t-множини в N-множину: $D(N,t)=\sum_{l=0}^{N}(-1)^l\binom{N}{l}(N-l)^t$ [10]. Для будь-якого $\delta\in(0,1)$ визначимо $t_1=t_1(\delta)$ як найменше натуральне t, для якого $D(N,t)\ge N^t(1-\delta)$. Тоді, застосовуючи наведений алгоритм з параметром $t=t_1$, отримаємо випадкову множину, яка співпадає з M із ймовірністю не менше $1-\delta$.

3. Алгоритм відновлення набору $(g_1,...,g_k)$

Для будь-якого $i\in\overline{1,k}$ позначимо e_i цілочисельний вектор довжини k, який має єдину ненульову, а саме, i-ту, координату, що дорівнює 1. Подаватимемо вектор e_i на вхід оракула зашифрування, поки вперше не отримаємо три різних шифрованих тексти:

$$c_1=g_i+z_{j_1},\ c_2=g_i+z_{j_2}\ \text{та}\ c_3=g_i+z_{j_3}, \tag{2}$$

де $z_{j_1}, z_{j_2}, z_{j_3}$ є незалежними випадковими рівноймовірними елементами множини M. Зауважимо, що на підставі леми 8.2.2 у [6] середнє число зашифрувань, які потрібно зробити для отримання трьох різних шифрованих текстів, дорівнює $3+1/(N-1)+1/(N-2)$.

Віднімаючи з першого рівняння (2) друге та третє відповідно, отримаємо, що

$$c_1 - c_2 = z_{j_1} - z_{j_2},\ c_1 - c_3 = z_{j_1} - z_{j_3}, \tag{3}$$

де $z_{j_1} \neq z_{j_2}$, $z_{j_1} \neq z_{j_3}$, $z_{j_2} \neq z_{j_3}$.

Припустимо, що множина M задовольняє такій *умові однозначності*: для будь-якого ненульового елемента $y \in G$ існує не більше однієї множини $\{z, z'\} \subseteq M$ такої, що $y = z - z'$.

В цьому випадку можна запропонувати наступний *алгоритм знаходження впорядкованого набору* $(g_1,...,g_k)$:

1) відновити множину M за допомогою алгоритму, наведеного в п. 2;

2) занумерувати елементи цієї множини довільним чином та побудувати таблицю T, яка складається з елементів $z_u - z_v$, записаних за адресами $\{u,v\}$, де $1 \le u < v \le N$, а $z_1, \dots, z_N$ є усі попарно різні елементи з M;

3) для кожного $i \in \overline{1,k}$:

– подавати вектор e_i на вхід оракула зашифрування, поки вперше не буде отримано три різних шифрованих тексти c_1, c_2, c_3;

– використовуючи таблицю T, знайти множини $\{u_1,v_1\}$ та $\{u_2,v_2\}$ такі, що $c_1 - c_2 = z_{u_1} - z_{v_1}$ та $c_1 - c_3 = z_{u_2} - z_{v_2}$;

– покласти $g_i = c_1 - z_{j_1}$, де j_1 – єдиний спільний елемент множин $\{u_1,v_1\}$ та $\{u_2,v_2\}$ (такий елемент напевно існує).

Твердження 1. Нехай $r = |G| \cdot |H|^{-1} \ge 14$ і випадкова множина M формується за рівноймовірною схемою (див. п. 1). Тоді ця множина задовольняє умові однозначності з ймовірністю не менше ніж $1 - N^4 |G|^{-1}$. При цьому наведений вище алгоритм відновлює набір $(g_1,...,g_k)$ із середньою трудомісткістю $O(kN^2)$.

Доведення. Перш за все, помітимо, що за умови однозначності система рівнянь (3) має єдиний розв'язок $(z_{j_1}, z_{j_2}, z_{j_3})$, який можна знайти за допомогою наведеного алгоритму, використовуючи рівності $\{u_1,v_1\} = \{j_1, j_2\}$, $\{u_2,v_2\} = \{j_1, j_3\}$. Звідси з урахуванням рівностей (2) випливає, що цей алгоритм вірно знаходить усі елементи набору $(g_1,...,g_k)$.

Далі, згідно з результатами п. 2, середнє число операцій, потрібних для побудови множини M на кроці 1) алгоритму, дорівнює $O(N \log N)$. Для побудови таблиці T на кроці 2) треба виконати $1/2 \cdot N(N-1)$ операцій. Нарешті, на кроці 3) для знаходження множини $\{u_1,v_1\}$ треба відшукати в таблиці T елемент $c_1 - c_2$, а у випадку його відсутності – елемент $c_2 - c_1$ (хоча б один з цих двох елементів обов'язково є в таблиці), що потребує не більше ніж $N(N-1)$ операцій. Таку ж кількість операцій треба виконати для знаходження множини $\{u_2,v_2\}$. Отже, середня трудомісткість кроку 3) не перевищує $k\left(3+1/(N-1)+1/(N-2)+3+2N(N-1)\right) = O(kN^2)$, і такою ж є оцінка середньої трудомісткості алгоритму в цілому.

Переконаємося зараз у справедливості першої частини твердження.

Позначимо: p_N ймовірність того, що випадкова множина M не задовольняє умові однозначності; m_N – число усіх множин потужності N, які задовольняють умові б) з п. 1 та

не задовольняють умові однозначності; $\tilde{m}_N$ – число усіх впорядкованих наборів $(z_1, \dots, z_N)$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H, й таких, що існує принаймні дві різні множини $\{u_1, v_1\}$, $\{u_2, v_2\}$ чисел від 1 до N такі, що $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$. З наведених означень випливає, що

$$p_N = m_N \binom{r-1}{N}^{-1} |H|^{-N} = \frac{\tilde{m}_N}{N!} \binom{r-1}{N}^{-1} |H|^{-N}. \quad (4)$$

При цьому

$$\tilde{m}_N \le \sum_{\{\{u_1,v_1\},\{u_2,v_2\}\}} \tilde{m}_N(u_1, v_1; u_2, v_2), \quad (5)$$

де $\tilde{m}_N(u_1, v_1; u_2, v_2)$ – число впорядкованих наборів $(z_1, \dots, z_N)$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H, й таких, що $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, а підсумування здійснюється за всіма множинами $\{\{u_1, v_1\}, \{u_2, v_2\}\}$ такими, що $\{u_1, v_1\}$ та $\{u_2, v_2\}$ є різними підмножинами множини чисел від 1 до N

Запишемо суму у правій частині нерівності (5) у вигляді $\tilde{m}_{N,1} + \tilde{m}_{N,2}$, де $\tilde{m}_{N,1}$ та $\tilde{m}_{N,2}$ є сумами чисел $\tilde{m}_N(u_1, v_1; u_2, v_2)$ за всіма такими підмножинами $\{\{u_1, v_1\}, \{u_2, v_2\}\}$, що $\{u_1, v_1\} \cap \{u_2, v_2\} = \varnothing$ та $|\{u_1, v_1\} \cap \{u_2, v_2\}| = 1$ відповідно.

Помітимо, що за умови $\{u_1, v_1\} \cap \{u_2, v_2\} = \varnothing$ число $\tilde{m}_N(u_1, v_1; u_2, v_2)$ дорівнює добутку двох чисел, перше з яких є кількістю всіх наборів $(z_{u_1}, z_{v_1}, z_{u_2}, z_{v_2})$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H та задовольняють умові $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, а друге є кількістю всіх впорядкованих наборів довжини $N-4$, елементи яких належать різним ненульовим суміжним класам групи G по підгрупі H. Перше з цих двох чисел не перевищує кількості розв'язків $(z_{u_1}, z_{v_1}, z_{u_2}, z_{v_2})$ рівняння $z_{u_1} - z_{v_1} = z_{u_2} - z_{v_2}$, яка дорівнює $|G|^3$, а друге дорівнює $(r-5)_{N-4} |H|^{N-4}$, де $(r-5)_{N-4}$ позначає число розміщень з $r-5$ по $N-4$. Таким чином, кожен доданок у сумі, що дорівнює $\tilde{m}_{N,1}$, не перевищує $|G|^3 (r-5)_{N-4} |H|^{N-4}$, в той час як число доданків дорівнює $\frac{1}{2}\binom{N}{2}\binom{N-2}{2} \le \frac{N^4}{8}$. Отже, $\tilde{m}_{N,1} \le \frac{N^4}{8} |G|^3 (r-5)_{N-4} |H|^{N-4}$.

Аналогічно доводиться, що кожен доданок в сумі, яка дорівнює $\tilde{m}_{N,2}$, не перевищує $|G|^2 (r-4)_{N-3} |H|^{N-3}$, в той час як число доданків є $\frac{1}{2}\left(\binom{N}{2}^2 - \binom{N}{2}\binom{N-2}{2} - \binom{N}{2}\right) \le \frac{N^4}{8}$.

Отже, $\tilde{m}_{N,2} \le \frac{N^4}{8} |G|^2 (r-4)_{N-3} |H|^{N-3}$.

Підставляючи наведені оцінки у формулу (4), отримаємо, що

$$p_N \le \frac{\tilde{m}_{N,1} + \tilde{m}_{N,2}}{N!} \binom{r-1}{N}^{-1} |H|^{-N} \le$$

$$\le \frac{N^4 |G|^3}{8} \frac{(r-5)_{N-4} |H|^{N-4}}{(r-1)_N |H|^N} + \frac{N^4 |G|^2}{8} \frac{(r-4)_{N-3} |H|^{N-3}}{(r-1)_N |H|^N} =$$

$$= \frac{N^4 |G|^3}{8|H|^4 r^4} \frac{r^4}{(r-1)(r-2)(r-3)(r-4)} + \frac{N^4 |G|^2}{8|H|^3 r^3} \frac{r^3}{(r-1)(r-2)(r-3)}.$$

Нарешті, враховуючи рівність $r = |G| \cdot |H|^{-1}$, отримаємо, що кожен з двох доданків у наведеній сумі не перевищує $\frac{N^4}{8|G|} \cdot \frac{1}{(1-4r^{-1})^4}$, що у свою чергу, не перевищує $\frac{N^4}{2|G|}$ через умову $r \ge 14$.

Таким чином, справедлива нерівність $p_N \le \frac{N^4}{|G|}$, що й треба було довести.

Зауважимо, що у випадку $G = (\mathbf{GF}(2)^n, \oplus)$ отриману оцінку ймовірності p_N можна декілька підсилити, враховуючи той факт, що $\tilde{m}_{N,2} = 0$. Дійсно, як випливає з означення цього параметра у викладеному доведенні, справедлива рівність $|\{u_1, v_1\} \cap \{u_2, v_2\}| = 1$, оскільки в протилежному випадку множини $\{u_1, v_1\}$ та $\{u_2, v_2\}$ співпадають внаслідок рівності $z_{u_1} + z_{v_1} = z_{u_2} + z_{v_2}$. Таким чином, якщо $G = (\mathbf{GF}(2)^n, \oplus)$, то за умови твердження $p_N \le \frac{N^4}{2|G|}$.

Отриманий результат показує, що у випадку, коли N є не надто великим числом (а саме, $N < \delta |G|^{1/4}$, $\delta \in (0,1)$), і множина M формується за рівноймовірною схемою, описана криптосистема може бути зламана з ймовірністю не менше ніж $1-\delta$ за $O(kN^2)$ операцій в середньому. Для підвищення стійкості криптосистеми треба збільшити параметр N, що негативно відіб'ється на її практичності, оскільки таблицю з N елементів множини M, треба зберігати в пам'яті.

4. Атака Стройка-ван Тілбурга

В даному пункті наведено узагальнення на випадок криптосистеми над скінченною абелевою групою атаки, запропонованої в [5] та вдосконаленої в [6]. Зауважимо, що у відзначених публікаціях для опису атаки використовуються помічені графи та їх групи автоморфізмів, що не є обов'язковим і, на нашу думку, ускладнює викладення. Нижче наведено більш простий опис атаки, аналогічної за сутністю запропонованій в [6]. Показано, що складність атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи G. Основним результатом цього пункту є твердження про те, що зазначений стабілізатор є з високою ймовірністю тривіальним у випадку, коли випадкова множина M формується за рівноймовірною схемою (див. п. 1). Звідси випливає, що у зазначеному випадку трудомісткість наведеної атаки дорівнює $O(kN^2\gamma(N))$, де $\gamma(N)$ – трудомісткість пошуку елемента в масиві, який складається з N елементів групи G.

Атака, що розглядається, проводиться в два етапи, на першому з яких відновлюється множина M за допомогою алгоритму з п. 2. На другому етапі відновлюється набір $(g_1, ..., g_k)$. З цією метою для кожного $i \in \overline{1,k}$ на вхід оракула зашифрування подається повідомлення e_i поки вперше не буде отримано N різних шифротекстів $c_1, ..., c_N$. На підставі формули (1) невідомий елемент g_i задовольняє системі рівнянь

$$x + z_{\pi(j)} = c_j, \ j \in \overline{1,N}, \tag{6}$$

де $M=\{z_1,...,z_N\}$, π – деяка підстановка на множині $\overline{1,N}$. Неважко знайти один з можливих розв'язків системи рівнянь (6), перебираючи всі значення $\pi(1)$ та обчислюючи $x=c_1-z_{\pi(1)}$, $z_{\pi(2)}=c_2-x$, …, $z_{\pi(N)}=c_N-x$. Оскільки обчислені таким чином елементи $z_{\pi(1)}$, $z_{\pi(2)}$, …, $z_{\pi(N)}$ є попарно різними, то необхідною й достатньою умовою правильності вибору значення $\pi(1)$ є приналежність елементів $z_{\pi(2)}$, …, $z_{\pi(N)}$ множині M, тобто умова

$$c_j-c_1+z_{\pi(1)}\in M,\ j\in\overline{2,N}. \tag{7}$$

Якщо ця умова виконується, то шуканий розв'язок g_i системи рівнянь (6) визначається за формулою $g_i=x=c_1-z_{\pi(1)}$.

Позначимо $I(M)=\{x\in G: x+M=M\}$ стабілізатор множини M в групі зсувів абелевої групи G. Відзначимо, що $I(M)$ є найбільшою за включенням підгрупою I групи G такою, що M є об'єднанням деяких суміжних класів G по I. Зокрема, число $|I(M)|$ є дільником числа $N=|M|$.

Твердження 2. Множина розв'язків системи рівнянь (6) має вигляд $x_0+I(M)$, де x_0 – довільний фіксований розв'язок цієї системи рівнянь.

Доведення. Якщо x є розв'язком системи (6), то для деяких підстановок π та π' на множині $\overline{1,N}$ мають місце рівності $x+z_{\pi(j)}=c_j=x_0+z_{\pi'(j)}$, $j\in\overline{1,N}$. Отже, $x-x_0+z_{\pi(j)}=z_{\pi'(j)}$, $j\in\overline{1,N}$, звідки випливає, що $x-x_0\in I(M)$. Навпаки, якщо $y\in I(M)$, то $x_0+y=c_j-z_{\pi'(j)}+y$ і, отже, $(x_0+y)+z_{\pi'(j)}=c_j$, $j\in\overline{1,N}$, тобто x_0+y задовольняє системі рівнянь (6) для підстановки $\pi=\pi'$.

Твердження доведено.

Таким чином, на підставі викладеного можна запропонувати наступний *алгоритм, який реалізує розглянуту атаку на криптосистему*:

1) відновити множину M за допомогою алгоритму, наведеного в п. 2;

2) для кожного $i\in\overline{1,k}$:

– подавати вектор e_i на вхід оракула зашифрування поки вперше не буде отримано N різних шифротекстів $c_1,...,c_N$;

– знайти шляхом перебору значення $\pi(1)\in\overline{1,N}$, для якого виконується умова (7) та покласти $g_i^{(0)}=c_1-z_{\pi(1)}$.

З означення криптосистеми і твердження 2 випливає, що наведений алгоритм завжди знайде певний набір $(g_1^{(0)},...,g_k^{(0)})\in G^k$ такий, що сукупність усіх шуканих наборів $(g_1,...,g_k)$ визначається за формулою

$$\{(g_1^{(0)},...,g_k^{(0)})+(x_1,...,x_k):(x_1,...,x_k)\in I(M)^k\}. \tag{8}$$

Іншими словами, кожен набір, який належить множині (8), може використовуватися (з погляду супротивника, який має доступ тільки до оракула зашифрування) в ролі частини $(g_1,...,g_k)$ секретного ключа криптосистеми. І навпаки, будь-який набір з останньою властивістю належить множині (8).

Таким чином, вся інформація про набір $(g_1,...,g_k)$, яку може отримати супротивник, маючи доступ до оракула зашифрування, полягає в тому, що цей набір належить множині (8), і наведений вище алгоритм дозволяє знайти один з елементів цієї множини.

Як видно з опису алгоритму, його середня трудомісткість складає $O(kN^2\gamma(N))$ операцій, де $\gamma(N)$ – складність пошуку одного елемента групи G в масиві, що використовується для зберігання множини M. Зокрема, $\gamma(N)=O(N)$ для довільної групи G і $\gamma(N)=O(\log N)$ у випадку, коли $G=(\mathbf{GF}(2)^n,\oplus)$ і елементи множини M розташовані в масиві в лексикографічному порядку. Крім того, середня трудомісткість побудови всієї множини (8) дорівнює $O(kN^2\gamma(N)+k\,|\,I(M)\,|^k)$, що при $k\geq 3$ є величиною порядку kN^k.

Знаючи набір $(g_1^{(0)},...,g_k^{(0)})$ і множину M, можна реалізувати на криптосистему атаку на основі відомого шифротексту, викладену в [5, 6] для випадку класичної криптосистеми Рао – Нама.

Дійсно, нехай c є шифрованим текстом вигляду (1), отриманим з деякого невідомого відкритого тексту $(m_1,...,m_k)$. Тоді

$$c=m_1g_1+\cdots+m_kg_k+z=m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}+z', \tag{9}$$

де $z'=z+(m_1g_1+\cdots+m_kg_k)-(m_1g_1^{(0)}+\cdots+m_kg_k^{(0)})$. Оскільки набір $(g_1,...,g_k)$ належить множині (8), а $I(M)$ є підгрупою групи G, то $(m_1g_1+\cdots+m_kg_k)-(m_1g_1^{(0)}+\cdots+m_kg_k^{(0)})\in I(M)$ і, отже, $z'\in M$, оскільки $z\in M$. Перебираючи усі значення z', можна знайти з формули (9) (можливо, у варіантах) значення $m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}$ за $O(N\theta(k,N))$ операцій, де $\theta(k,N)$ позначає складність перевірки приналежності елемента групи G підгрупі, породженій елементами $g_1^{(0)},...,g_k^{(0)}$. При цьому, якщо $I(M)=\{0\}$, то $(g_1,...,g_k)=(g_1^{(0)},...,g_k^{(0)})$ і на підставі умов (а), (б) з означення криптосистеми (п. 1) значення $m_1g_1^{(0)}+\cdots+m_kg_k^{(0)}$, а отже, і невідомий відкритий текст, можна відновити однозначно.

Отримаємо зараз верхню оцінку ймовірності події $I(M)\neq\{0\}$, вважаючи, що M є випадковою множиною, яка формується за рівноймовірною схемою (див. п. 1).

Твердження 3. Нехай $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$, де $\exp G$ – експонента (максимальний порядок елементів) групи G. Тоді для випадкової множини M потужності N, яка формується за рівноймовірною схемою, справедлива нерівність

$$\mathbf{P}(I(M)\neq\{0\})\leq\frac{1}{2}\left(\frac{N}{2}\right)^{\frac{N}{2}}\left(1-\frac{r}{N}\right)^{-N}|\,G\,|^{2-\frac{N}{2}}, \tag{10}$$

де $r=|\,G\,||\,H\,|^{-1}$.

Доведення. Нехай $I(M)\neq\{0\}$. Тоді існує ненульовий елемент $a\in G$ такий, що $a+M=M$, звідки випливає, що M є об'єднанням деяких циклів підстановки $t_a(x)=x+a$, $x\in G$.

Позначимо l порядок елемента a групи G. Підстановка t_a є добутком $\frac{|G|}{l}$ циклів, кожен з яких має довжину l, а кількість усіх множин потужності N, які є об'єднанням циклів цієї підстановки, дорівнює біноміальному коефіцієнту $\binom{\frac{|G|}{l}}{\frac{N}{l}}$, якщо l ділить N, та нулю – у протилежному випадку. При цьому, оскільки множина M формується за рівноймовірною схемою, то

$$\mathbf{P}(I(M)\neq\{0\})\le\sum_{a\in G\setminus\{0\}}\mathbf{P}(a+M=M)\le|G||H|^{-N}\binom{r-1}{N}^{-1}\binom{\frac{|G|}{l}}{\frac{N}{l}}. \qquad (11)$$

Далі, з умови $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$ випливає, що $\frac{N}{l}\le\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2l}$ і, отже,

$$\binom{\frac{|G|}{l}}{\frac{N}{l}}\le\binom{\frac{|G|}{l}}{\left\lceil\frac{N}{2}\right\rceil}\le\frac{\left(\frac{|G|}{2}\right)^{\left\lceil\frac{N}{2}\right\rceil}}{\left\lceil\frac{N}{2}\right\rceil!}\le\frac{\left(\frac{|G|}{2}\right)^{\frac{N}{2}+1}}{\left\lceil\frac{N}{2}\right\rceil!}.$$

Підставляючи зазначену оцінку в формулу (11) та використовуючи рівність $r=|G||H|^{-1}$, отримаємо, що

$$\mathbf{P}(I(M)\neq\{0\})\le|G||H|^{-N}\frac{N!}{(r-1)_N}r^N\frac{|H|^N}{|G|^N}\frac{\left(\frac{|G|}{2}\right)^{\frac{N}{2}+1}}{\left\lceil\frac{N}{2}\right\rceil!}=$$

$$=N(N-1)\cdots\left(\left\lceil\frac{N}{2}\right\rceil+1\right)2^{-\frac{N}{2}-1}\frac{1}{(1-r^{-1})(1-2r^{-1})\cdots(1-Nr^{-1})}|G|^{2-\frac{N}{2}}\le$$

$$\le\frac{1}{2}\left(\frac{N}{2}\right)^{\frac{N}{2}}\left(1-\frac{r}{N}\right)^{-N}|G|^{2-\frac{N}{2}}.$$

Таким чином, справедлива нерівність (10). Твердження доведено.

Наслідок. Нехай $G=(\mathbf{GF}(2)^n,\oplus)$ і N є непарним числом. Тоді $I(M)=\{0\}$. Якщо ж N є парним і $N<2^{n-k-1}$, а множина M формується за рівноймовірною схемою, то

$$\mathbf{P}(I(M)\neq\{0\})\le2^{-n\left(\frac{N}{2}-2-\frac{N}{2n}(\log n+1)\right)-1}. \qquad (12)$$

Доведення. Якщо $N=|M|$ є непарним числом, то рівність $a\oplus M=M$ є неможливою для будь-якого ненульового елемента $a\in\mathbf{GF}(2)^n$. Отже, в цьому випадку $I(M)=\{0\}$. Для парного N нерівність (12) є безпосереднім наслідком формули (10) та нерівності $\left(1-\frac{r}{N}\right)^{-N}\le2^N$, яка випливає з оцінки $N<2^{n-k-1}=\frac{r}{2}$ (зауважимо також, що в даному випадку $\exp G=2$ і нерівність $\left\lceil\frac{N}{2}\right\rceil<\frac{|G|}{2\exp G}$ напевно виконується).

Наслідок доведено.

Висновки

Основним результатом статті є аналітичні оцінки ефективності двох атак на симетричну криптосистему, яка узагальнює відому еквівалентну схему Рао – Нама [6] на випадок скінченної абелевої групи. Секретними ключами такої криптосистеми над групою G є набори $((g_1,\dots,g_k),M,\sigma)$, де $(g_1,\dots,g_k)\in G^k$, $M=\{z_1,\dots,z_N\}\subseteq G$, $\sigma:G\to G'$ – епіморфізм груп,

ядро якого співпадає з підгрупою H, породженою елементами $g_1,...,g_k$. При цьому вважається, що виконані умови (а), (б) з п. 1.

Перша атака, запропонована в цій статті, дозволяє зламувати зазначену криптосистему з ймовірністю не менше ніж $1-\delta$ за $O(kN^2)$ операцій в середньому за умови $N < \delta |G|^{1/4}$, де $\delta \in (0,1)$.

Друга атака являє собою узагальнено-спрощений варіант атаки Стройка-ван Тілбурга [5, 6] і за певних умов (див. твердження 3) дозволяє зламувати зазначену криптосистему за $O(kN^2\gamma(N))$ операцій, де $\gamma(N)$ – трудомісткість пошуку елемента в масиві, який складається з N елементів групи G.

На відміну від [5, 6], при аналізі ефективності атаки Стройка-ван Тілбурга не використовуються деякі "зайві" поняття (поміченого графу та його групи автоморфізмів), що суттєво спрощує викладення. Показано, що складність цієї атаки залежить від потужності стабілізатора множини M у групі зсувів абелевої групи G, який є тривіальним з ймовірністю, зазначеній у формулюванні твердження 3.

Для підвищення стійкості розглянутої криптосистеми до наведених атак треба збільшити параметр N, що негативно відіб'ється на її практичності, оскільки треба зберігати в пам'яті таблицю з N елементів множини M.

Список літератури:

1. Rao T.R.N., Nam K.H. Private-key algebraic code encryption // IEEE Trans. on Inform Theory, 1989. P. 829 – 833.
2. McEliece R.J. A public-key cryptosystem based on algebraic coding theory // Prog. Rep., Jet Prop.Lab., California Inst. Technol, 1978. P. 114 – 116.
3. Jordan J.P. A variant of public-key cryptosystem based on Goppa codes // Sigact news, 1983. P. 61 – 66.
4. Rao T.R.N. Cryptosystems using algebraic codes // Int. Conf on Computer Systems & Signal Processing, 1984.
5. Struik R., van Tilburg J. Thr Rao-Nam scheme is insecure against a chosen plaintext attack // Advances in Cryptology-CRYPTO'87. Proc. Springer, 1988. P. 445 – 457.
6. Van Tilburg J. Security-analyses of a class of cryptosystems based on linear error-correcting codes. PhD Thesis. Technische Universiteit Eindhoven, 1994.
7. Bagheri K., Eghlidos T., Sadeghi M.-R., Panario D. Lattice based join encryption, encoding, and modulation scheme // arXiv: 1906.06280v1[cs.IT] 4 Juni 2019. P. 1 – 30.
8. Колчин В.Ф., Севастьянов Б.А., Чистяков В.П. Случайные размещения. Москва : Наука, 1976. 223 с.
9. Гельфонд А.О. Исчисление конечных разностей. Москва : Наука, 1967. 376 с.
10. Сачков В.Н. Введение в комбинаторные методы дискретной математики. Москва : Наука, 1982. 384 с.

Надійшла до редколегії 25.02.2021

Відомості про авторів:

Олексійчук Антон Миколайович – д-р техн. наук, доцент, Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "КПІ", професор кафедри Кібербезпеки; Україна; e-mail: alex-dtn@ukr.net; ORCID: https://orcid.org/0000-0003-4385-4631

Шевчук Ольга Сергіївна – Інститут спеціального зв'язку та захисту інформації Національного технічного університету України "КПІ", інженер кафедри Кібербезпеки; Україна; e-mail: olia13511@gmail.com; ORCID: https://orcid.org/0000-0002-2866-439X

УДК 004.056.5 **DOI:10.30837/rt.2021.2.205.03**

О.О. КУЗНЕЦОВ, д-р техн. наук, А.В. КОНОНЧЕНКО

СТЕГАНОГРАФІЧНІ МЕТОДИ В ВЕКТОРНІЙ ГРАФІЦІ

Вступ

Приховування інформаційних повідомлень у різні надмірні дані (так звані контейнери, носії, cover-файли, тощо) вивчається стеганографією [1 – 3]. Тобто метою стеганографічних методів є приховування факту існування інформаційних повідомлень [4, 5]. Контейнери (зображення, аудіо, відео, тексти, тощо) передаються відкритими каналами (наприклад, через Інтернет), але не викликають ні в кого підозр. В той же час вповноважена особа, що знає таємний ключ, може витягти приховані дані, тобто відновити інформаційне повідомлення [1 – 3].

Найбільш вдалим для проведення стеганографічних перетворень є формат векторної графіки SVG [6], який завдяки своїй структурі дозволяє легко маніпулювати об’єктами, з яких складається. Його широка підтримка різними платформами також дозволяє підвищити рівень скритності при проведенні передачі секретних даних шляхом передачі звичайних на перший погляд файлів медіа.

Існуючі наукові публікації з методів приховування інформаційних повідомлень стосуються переважно растрової графіки, а роботи щодо векторної в основному полягають у кодуванні відстаней між геометричними об’єктами, що містяться у зображенні, або створенні додаткових точок [7 – 10]. Усі методи вбудовування у векторні зображення мають вразливість до атак афінного перетворення [11, 12].

Найпоширенішими видами афінних перетворень є операції перенесення, повороту, зсуву та масштабування з можливими варіаціями (зсуву за осями абсцис та ординат, масштабування пропорційне та непропорційне, зі стисненням та із розширенням) [13, 14].

Більшість методів вбудовування інформації у векторні зображення забезпечують одноразову стійкість до афінних перетворень, при цьому при повторному накладенні операцій зміни положення об’єктів, повідомлення може зруйнуватися взагалі. Запропоновані у [11, 12] способи приховування даних можуть реалізувати більший рівень стійкості до різного роду перетворень при їх багаторазовому проведенні.

Метою цієї статті є вивчення технік приховування повідомлень в векторні зображення з [11, 12] та експериментальні дослідження стійкості до афінних атак.

Афінні перетворення для реалізації стеганоатак

Афінні перетворення найбільш імовірно можуть бути використані для реалізації стеганоатак, тобто як спосіб руйнування прихованого повідомлення. Так, у загальному випадку афінне перетворення відбувається наступним чином [13, 14]:

$$\begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} ax+by+c \\ dx+ey+f \\ 1 \end{bmatrix}. \quad (1)$$

При відповідних коефіцієнтах над координатами точок, що піддаються операції, відбувається перетворення, що призводить до зміни положення геометричного об’єкта на полотні. Так, основними афінними перетвореннями є перенесення, поворот, зсув та масштабування (рис. 1). При цьому зсув може відбуватися за осями абцис та ординат, а масштабування може бути пропорційним та непропорційним, зі стисненням та розширенням.

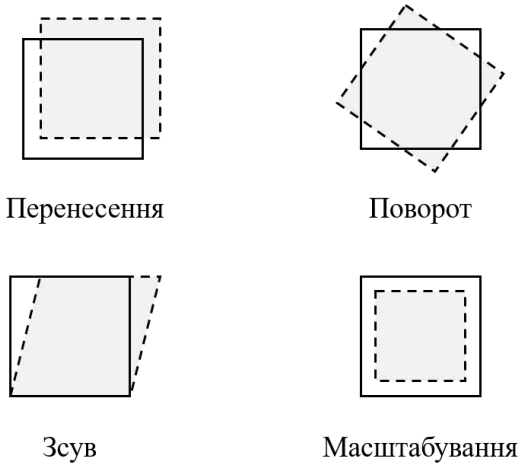


Рис. 1. Основні види афінних перетворень

Вираз (1) для різних випадків може бути записаний наступним чином [12]:

- операція перенесення:

$$\begin{bmatrix} 1 & 0 & c \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+c \\ y+f \\ 1 \end{bmatrix}; \tag{2}$$

- операція повороту:

$$\begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \\ = \begin{bmatrix} x\cos(\alpha) - y\sin(\alpha) \\ x\sin(\alpha) + y\cos(\alpha) \\ 1 \end{bmatrix}; \tag{3}$$

- операція зсуву за віссю абсцис:

$$\begin{bmatrix} 1 & b & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x+by \\ y \\ 1 \end{bmatrix}; \tag{4}$$

- операція зсуву за віссю ординат:

$$\begin{bmatrix} 1 & 0 & 0 \\ d & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ dx+y \\ 1 \end{bmatrix}; \tag{5}$$

- операція масштабування:

$$\begin{bmatrix} a & 0 & 0 \\ 0 & e & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} ax \\ ey \\ 1 \end{bmatrix}; \tag{6}$$

- операція майже афінного перетворення з додаванням шуму:

$$\begin{bmatrix} a+n_1 & b+n_2 & 0 \\ d+n_3 & e+n_4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x+n_5 \\ y+n_6 \\ 1 \end{bmatrix} = \\ = \begin{bmatrix} (a+n_1)(x+n_5)+(b+n_2)(y+n_6) \\ (d+n_3)(x+n_5)+(e+n_4)(y+n_6) \\ 1 \end{bmatrix}. \quad (7)$$

Більшість відомих стеганографічних технік над векторними зображеннями не забезпечують стійкості до поширених видів афінних перетворень при їх багаторазовому застосуванні на контейнері. Наведені результати останніх досліджень Кінзяревого [11, 12] пропонують стійкість до афінних перетворень при їх повторному накладанні. Саму тому тематикою роботи є аналіз запропонованих технік та їх експериментальні дослідження.

Методика досліджень

В роботах [11, 12] запропоновано два методи приховування інформації у векторні зображення (побітовий метод та метод паттернів). Для проведення експериментальних досліджень в роботі програмно реалізовано обидва методи із рекомендованими автором параметрами.

Дослідження стійкості до афінних перетворень виконувалось програмним чином, використовуючи формули (2) – (7). Спочатку реалізовується операція вбудовування інформації довжиною 2400 біт у контейнер. Потім над закодованим контейнером відбувається певне афінне перетворення і виконується спроба вилучити повідомлення. Для наступного перетворення використовуватиметься стегоконтейнер, отриманий на попередньому кроці. Таким чином забезпечується багаторазове накладання операції зміни положення координат на одне й теж ж саме зображення.

Згідно з формулами (2)-(7) для перетворень використовуються такі коефіцієнти:

- перенесення – $(c, f) \in [-500, 500]$;
- поворот – $\alpha = 1$;
- зсув за віссю абсцис – $b = 0,01$;
- зсув за віссю ординат – $d = 0,01$;
- масштабування для стиснення – $a = e = 0,99$;
- масштабування для розширення – $a = e = 1,01$;
- майже афінне перетворення повороту з додаванням шуму:

$$a = \cos(\alpha),\ b = -\sin(\alpha),\ d = \sin(\alpha),\ e = -\cos(\alpha),\ \alpha = 1,\ n_1 = n_4 = n_6 = -0,0001,$$
$$n_2 = n_3 = n_5 = 0,0001.$$

Такі ж самі дослідження було проведено у дисертаційній роботі [12]. Нашою метою була вибіркова перевірка отриманих результатів.

При проведенні наших експериментів операції приховування та вилучення виконуються з різними наборами значень кількості біт, що вбудовуються в одну криву, точності розрахунку координат та похибки відтворення. Вихідні дані для проведення експериментів наведено у табл. 1.

Таблиця 1

Вхідні дані для експериментальних досліджень

Номер експерименту	Біт/крива	Точність	Похибка
e_1	40	5	0,0002
e_2	40	5	0,0004
e_3	40	6	0,0002
e_4	40	6	0,0004
e_5	60	5	0,0002
e_6	60	5	0,0004
e_7	60	6	0,0002
e_8	60	6	0,0004
e_9	80	5	0,0002
e_{10}	80	5	0,0004
e_{11}	80	6	0,0002
e_{12}	80	6	0,0004

Накладання перетворень виконувались 10 раз на кожен попередній контейнер. Зміна кроку для побітового методу дорівнює 0.0005, для побітового методу початковий крок дорівнює 0,0005, а довжина паттерну – 4.

Результати експериментів

Отримані результати зображено на рис. 2 – 15, на яких відображені графіки залежності відсотка втрачених біт від експерименту при проведенні різних перетворень. По шкалі абсцис наведено кількість послідовно виконаних тих самих перетворень.

Графіки на рис. 2, 3 відповідають перетворенню перенесення. Кожний контейнер набував нового вигляду шляхом послідовного додавання до координат значень від -500 до 500 з кроком 100. Побітовий метод продемонстрував високі показники стійкості (рис. 2) – з 12 експериментів лише при першому та дев'ятому експериментах було втрачено невеликий відсоток інформаційних бітів. Метод паттернів у більшості випадків також стійкий до перенесення, при цьому при п'ятому, дев'ятому та десятому експериментах показники втрат досягають 30 – 40 %.

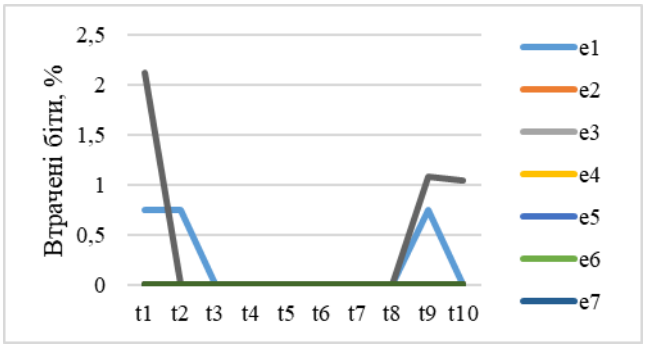


Рис. 2. Результати вилучення інформації з контейнерів після перетворення перенесення побітовим методом

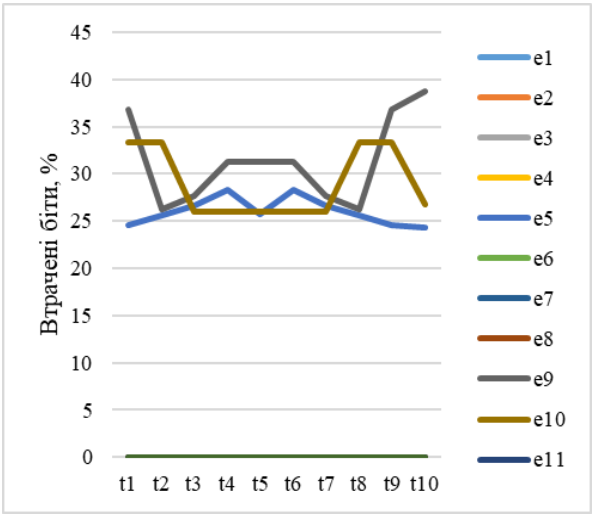


Рис. 3. Результати вилучення інформації з контейнерів після перетворення перенесення методом пат тернів

На рис. 4, 5 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні перетворення повороту. У цьому випадку контейнер поступово повертався на один градус десять разів. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 45 %, а для методу паттернів – 100 %.

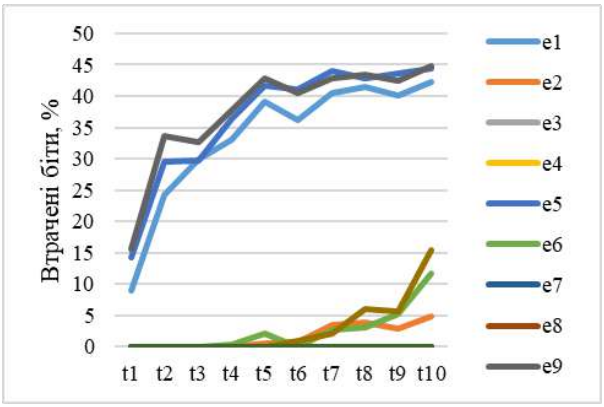


Рис. 4. Результати вилучення інформації з контейнерів після перетворення повороту побітовим методом

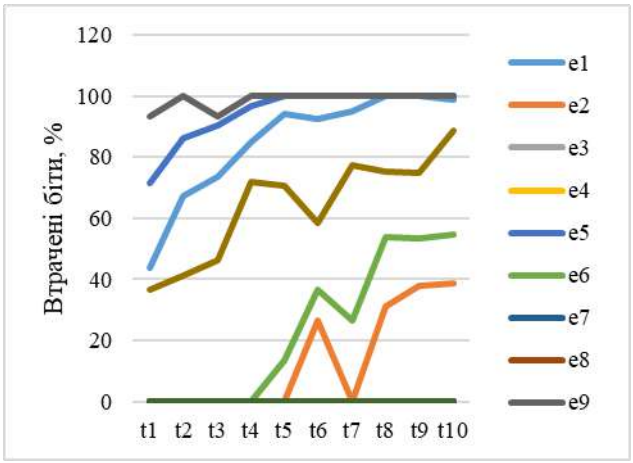


Рис. 5. Результати вилучення інформації з контейнерів після перетворення повороту методом паттернів

На рис. 6, 7 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні перетворення зсуву за віссю абсцис. При цьому *x*-координата змінюється у 0,01 раз, забезпечуючи поступовий зсув контейнера. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 45 %, а для методу паттернів – 100 %.

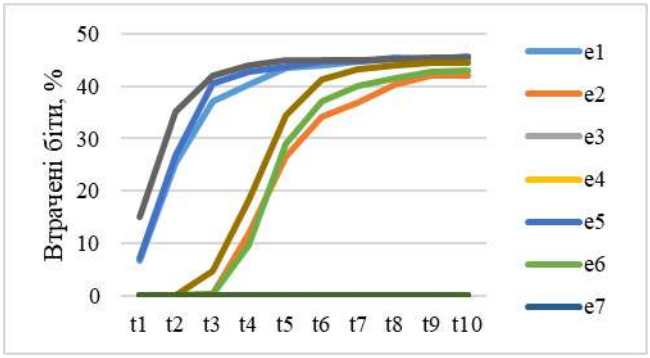


Рис. 6. Результати вилучення інформації з контейнерів
після перетворення зсуву за віссю абсцис побітовим методом

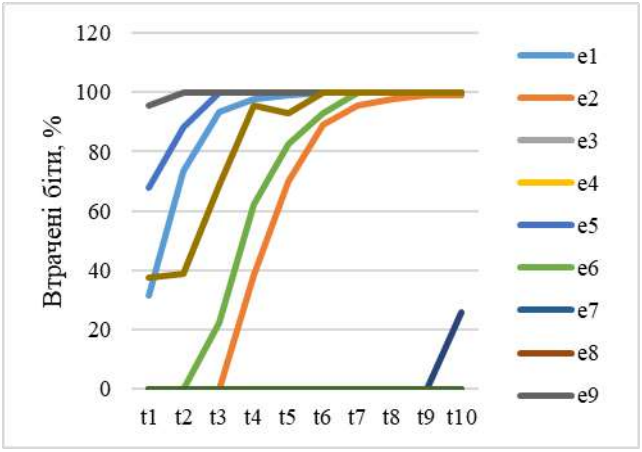


Рис. 7. Результати вилучення інформації з контейнерів
після перетворення зсуву за віссю абсцис методом паттернів

На рис. 8, 9 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні перетворення зсуву за віссю ординат. При цьому y-координата змінюється у 0,01 раз, забезпечуючи поступовий зсув контейнера. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 46 %, а для методу паттернів – 100 %.

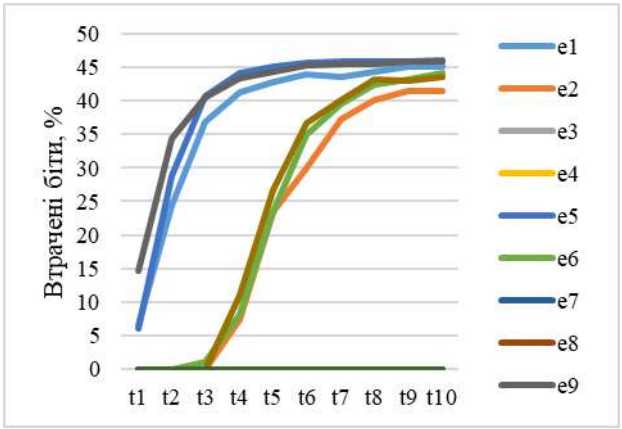


Рис. 8. Результати вилучення інформації з контейнерів
після перетворення зсуву за віссю ординат

побітовим методом

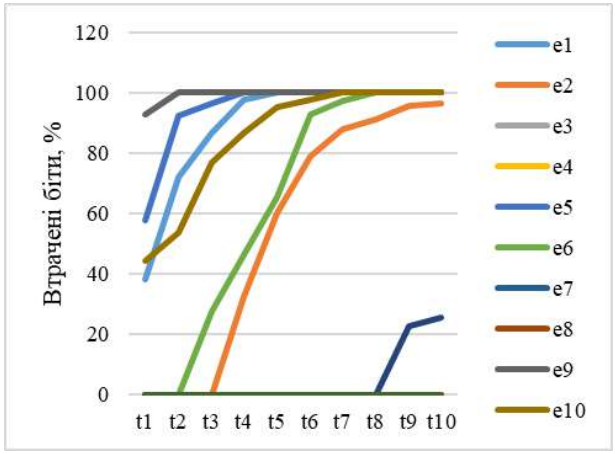


Рис. 9. Результати вилучення інформації з контейнерів
після перетворення зсуву за віссю ординат методом паттернів

На рис. 10, 11 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні перетворення масштабування для стиснення. При цьому кожна координата змінюється у 0,99 раз, забезпечуючи поступове стиснення контейнера. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 44 %, а для методу паттернів – 100 %.

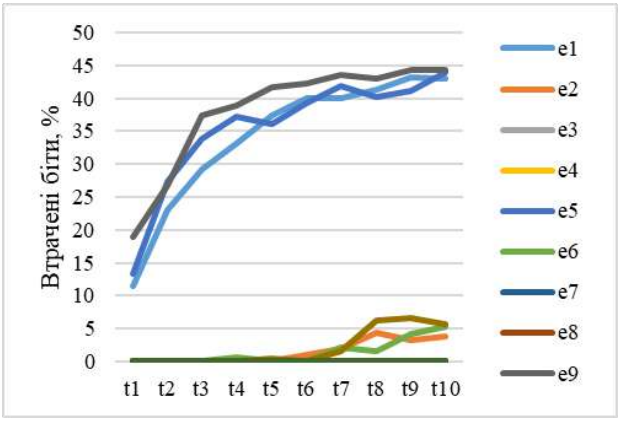


Рис. 10. Результати вилучення інформації з контейнерів
після перетворення масштабування для стиснення побітовим методом

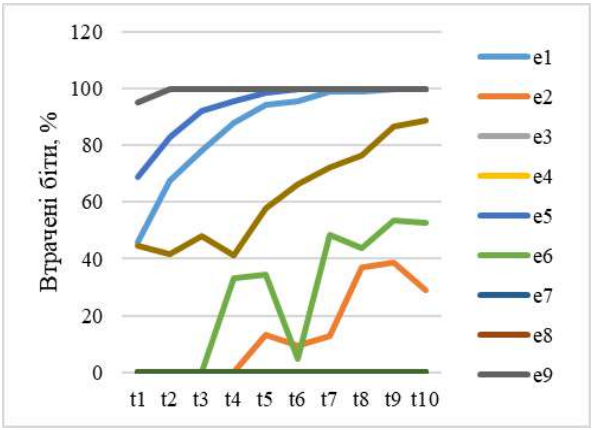


Рис. 11. Результати вилучення інформації з контейнерів
після перетворення масштабування для стиснення методом паттернів

На рис. 12, 13 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні перетворення масштабування для розширення. При цьому кожна координата змінюється у 1,01 раз, забезпечуючи поступове розширення контейнера. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 45 %, а для методу паттернів – 100 %.

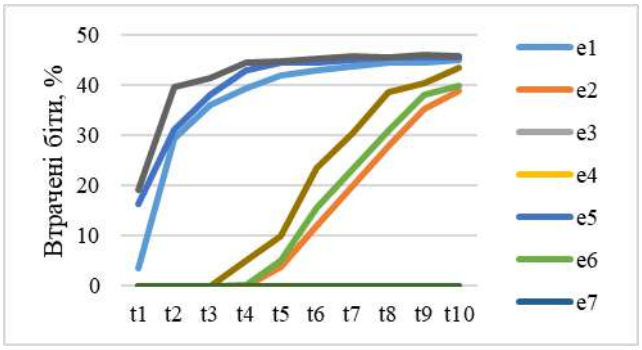


Рис. 12. Результати вилучення інформації з контейнерів
після перетворення масштабування для розширення побітовим методом

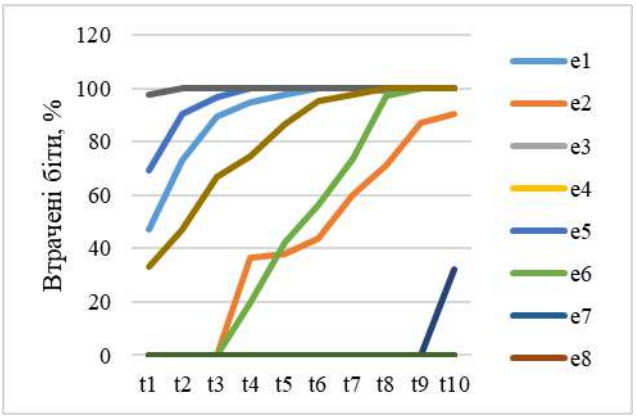


Рис. 13. Результати вилучення інформації з контейнерів
після перетворення масштабування для розширення методом паттернів

На рис. 14, 15 відображені графіки залежності відсотка втрачених біт від експерименту при проведенні майже афінного перетворення повороту. У цьому випадку контейнер поступово повертався на один градус десять разів, і до кожної координати додавалось певне значення, яке відповідає за шум. Для побітового методу при найнесприятливіших параметрах досягається втрата бітів у приблизно 33 %, а для методу паттернів – 100 %.

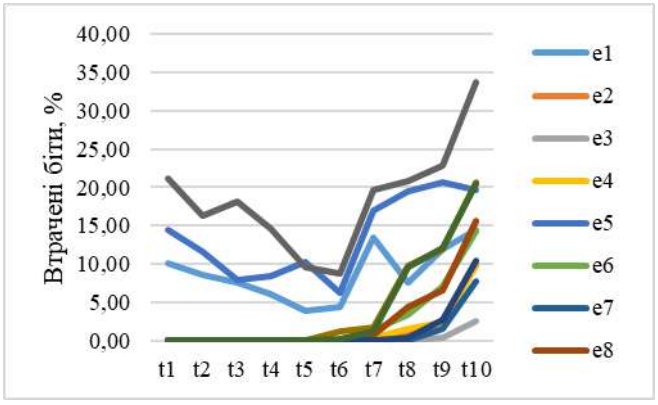


Рис. 14. Результати вилучення інформації з контейнерів
після майже афінного перетворення повороту побітовим методом

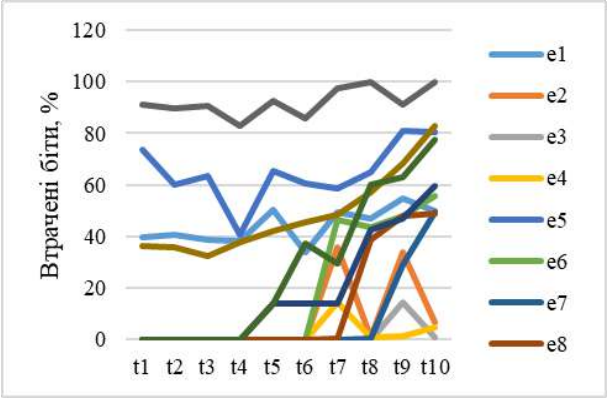


Рис. 15. Результати вилучення інформації з контейнерів
після майже афінного перетворення повороту методом паттернів

Таким чином, проведені експериментальні дослідження показали, що методи приховування інформаційних повідомлень у векторні зображення з [11, 12] дійсно забезпечують деяку стійкість до афінних атак.

Отримані результати відрізняються від вперше представлених у роботах [11, 12] більш високими показниками втрачених біт навіть при меншій кількості разів проведення перетворень над контейнером. Це можна пояснити особливостями програмної реалізації, використовуваними контейнерами, більшим об'ємом відстежуваних винятків при вбудовуванні/вилученні повідомлення тощо.

У загальному випадку отримані результати за допомогою програмно розроблених у даній роботі методів співпадають із представленими результатами О. Кінзерявого [11, 12]. Виключенням є сто відсоткова втрата блоків при застосуванні методу паттернів при таких перетворення як поворот, зсув за осями абсцис та ординат, масштабування для стиснення та розширення, а також при деяких параметрах майже афінного перетворення повороту. В обох реалізація найбільш сприятливим є параметр точності у розмірі 6, при якому не залежачи від кількості біт, що кодуються в криву, та похибки відтворення, досягається нульовий (або майже нульовий) відсоток втрачених або неправильно вилучених біт.

За допомогою програмно розроблених методів можлива реалізація розглянутих методів у різноманітних веб-застосунках для виконання факту перевірки ключових даних. Такими даними можуть бути мітки часу або певна величина, яка формується шляхом відстежування усіх елементів на сторінці і може бути використана для запобігання зловмисного впровадження надбудованих конструкцій, що можуть бути застосовані для перехоплення персональних даних користувачів (така технологія обману носить назву клікджекінгу). Мітки часу можуть бути використані в якості підтвердження існування певних даних у деякий момент.

Окрім цього існує практика створення компонентів DOM-структури із SVG зображень і цілком можливою є ситуація, коли вся веб-сторінка або її складова частина складається із елементів векторної графіки. Цей факт може бути використаний не тільки для збереження даних, а й у якості підтвердження прав авторства, володіння тощо.

Висновки

Результати досліджень показників стійкості до афінних перетворень показали, що метод паттернів програє побітовому. Це пов’язано у першу чергу із алгоритмом вилучення інформації методу паттернів. Так, можлива ситуація втрати цілих блоків інформаційних бітів, коли не знаходиться перша частина послідовності (перший паттерн), а звідси – неможливим стає знаходження наступного коефіцієнту побудови кривої, при якому відбулося б об’єднання сегментів. Така залежність присутня в обох методах, однак побітовий у випадку, коли не виконується умова відтворення, вилучають двійковий «0» (тобто це може бути навіть неправильно вилучений біт), і процес декодування продовжується.

Дослідження стійкості до афінних перетворень також показали, що найбільш сприятливими для приховання та вилучення є випадки, в яких застосовувалась точність 6 для розрахунку координат, при цьому похибка та кількість біт, що приходяться на одну криву, – неважливі.

Практичне застосування методів приховування інформації у векторній графіці полягає у можливості збереження таких даних як мітки часу, права авторства або володіння, значень контрольної суми компонентів DOM-структури для аналізу стану веб-застосунку у випадках застосування анти-клікджекінгових технологій тощо.

Список літератури:

1. Fridrich J. Steganography in Digital Media: Principles, Algorithms, and Applications. Illustrated Edition. Cambridge; New York: Cambridge University Press, 2009. 466 p.
2. Yahya A. Steganography Techniques // Steganography Techniques for Digital Images / ed. Yahya A. Cham: Springer International Publishing, 2019. P. 9–42.
3. Yahya A. Introduction to Steganography // Steganography Techniques for Digital Images / ed. Yahya A. Cham: Springer International Publishing, 2019. P. 1–7.
4. Manoj I.V.S. Cryptography and Steganography // IJCA. 2010. Vol. 1, № 12. P. 63–68.
5. Menon N., Vaithiyanathan. A survey on image steganography // 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy). 2017. P. 1–5.
6. Basic Shapes – SVG 1.1 (Second Edition) [Electronic resource]. URL: https://www.w3.org/TR/SVG11/shapes.html (accessed: 12.05.2021).
7. Doncel V.R., Nikolaidis N., Pitas I. An Optimal Detector Structure for the Fourier Descriptors Domain Watermarking of 2D Vector Graphics // IEEE Transactions on Visualization and Computer Graphics. 2007. Vol. 13, № 5. P. 851–863.
8. Wang X. et al. Reversible Data-Hiding Scheme for 2-D Vector Maps Based on Difference Expansion // IEEE Transactions on Information Forensics and Security. 2007. Vol. 2, № 3. P. 311–320.
9. Wu D., Wang G., Gao X. Reversible Watermarking of SVG Graphics // 2009 WRI International Conference on Communications and Mobile Computing. 2009. Vol. 3. P. 385–390.
10. Peng F. et al. Reversible Data Hiding in Encrypted 2D Vector Graphics Based on Reversible Mapping Model for Real Numbers // IEEE Transactions on Information Forensics and Security. 2019. Vol. 14, № 9. P. 2400–2411.
11. Kinzeryavyy O. et al. Steganographic Method of Bitwise Information Hiding in Point-Defined Curves of Vector Images // Advances in Computer Science for Engineering and Education / ed. Hu Z. et al. Cham: Springer International Publishing, 2019. P. 478–486.
12. Kinzeryavyy O. Steganographic methods for hiding data into vector images that are resistant to active attacks based on affine transformations: Thesis. 2015.
13. Coste A. Image Processing : Affine Transformation, Landmarks registration, Non linear Warping. 2012.
14. Weisstein E.W. Affine Transformation [Electronic resource]: Text. Wolfram Research, Inc. URL: https://mathworld.wolfram.com/AffineTransformation.html (accessed: 24.05.2021).

Надійшла до редколегії 11.03.2021

Відомості про авторів:

Кузнецов Олександр Олександрович – д-р техн. наук, професор, Харківський національний університет імені В.Н. Каразіна, професор кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: kuznetsov@karazin.ua, ORCID:https://orcid.org/0000-0003-2331-6326

Кононченко Ганна Володимирівна – Харківський національний університет імені В.Н. Каразіна, студент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: akononpro@gmail.com, ORCID: https://orcid.org/0000-0002-8101-6500

УДК 004.056.55 **DOI:10.30837/rt.2021.2.205.04**

М.В. ЄСІНА, канд. техн. наук, Б.С. ШАХОВ

АНАЛІЗ АПАРАТНИХ РЕАЛІЗАЦІЙ АЛГОРИТМІВ ЕЛЕКТРОННОГО ПІДПИСУ QTESLA, CRYSTALS-DILITIUM І MQDSS НА РІЗНИХ РІВНЯХ БЕЗПЕКИ

Вступ

Як відомо, існуючі алгоритми криптографії з відкритим ключем, що засновані на RSA та еліптичних кривих, надають гарантії безпеки, які супроводжуються складністю; також можна казати про неможливість вирішення завдань цілочисельної факторизації і дискретного логарифму. Але експерти прогнозують, що створення квантового комп'ютера зможе зламати класичні криптографічні алгоритми. Через цю майбутню проблему Національний інститут стандартів і технологій (NIST) разом із провідними вченими у галузі криптографії розпочав відкритий процес стандартизації алгоритмів з відкритим ключем для квантових атак. Метою цієї роботи є аналіз та порівняння трьох апаратних реалізацій кандидатів 2-го раунду конкурсу NIST на алгоритм електронного підпису.

Криптографія з відкритим ключем – це фундаментальний протокол безпеки для всіх форм цифрового зв'язку, дротового або бездротового. Криптографія з відкритим ключем має три основні криптографічні функції, а саме: шифрування з відкритим ключем, електронні підписи (ЕП) і обмін ключами. Алгоритми криптографії з відкритим ключем на основі RSA і еліптичних кривих забезпечують гарантії безпеки, засновані на складності вирішення завдань цілочисельної факторизації і дискретного логарифму. Пітер Шор показав, що квантові комп'ютери можуть факторизувати цілі числа за поліноміальний час, роблячи традиційні алгоритми криптографії з відкритим ключем неефективними. Після цього криптографи стали шукати надійні альтернативи, такі як криптографічні алгоритми на основі решіток та математичних кодів.

З історичної точки зору, у 1997 р. NIST запросив рекомендації у громадськості для визначення заміни стандарту шифрування даних (DES), Advanced Encryption Standard (AES) [4]. Відтоді відкриті криптографічні оцінки стали способом вибору криптографічних стандартів. Наприклад, NESSIE (2000-2002), eSTREAM (2004-2008), CRYPTREC (2000-2002), SHA-3 (2007-2012) і CAESAR (2013-2019) прийняли цей підхід. У цих оцінках головним параметром була безпека. Продуктивність у програмному забезпеченні, продуктивність у прикладних специфічних інтегральних схемах (ASIC), продуктивність у FPGA та можливість реалізації з використанням обмежених ресурсів (невеликих мікропроцесорів та малопотужних апаратних засобів) є вторинними критеріями. У конкурсі AES Rijndael мав найшвидшу ASIC і другу найшвидшу реалізацію FPGA з тими ж гарантіями безпеки, що й конкуренти [5].

У роботі описується порівняння апаратного забезпечення трьох алгоритмів підпису (qTesla, Crystals-Dilitium, MQDSS), які зокрема є кандидатами 2-го раунду конкурсу NIST PQC, а алгоритм Crystals-Dilitium – фіналістом цього конкурсу.

1. Постквантова криптографія

Експерти та вчені у галузі криптографії почали розробляти постквантові криптографічні алгоритми, які зможуть вистояти проти атак постквантових комп'ютерів. Безперечно, алгоритми постквантової криптографії поділяються на класи. У цій роботі розглядаються два алгоритми на основі алгебраїчних решіток (qTesla та Crystal-Dilitium) і один алгоритм на основі багатовимірної криптографії (MQDSS) [1 - 8].

Криптографія на основі решіток заснована на підході побудови алгоритмів асиметричного шифрування з використанням задач теорії решіток, тобто задач оптимізації на дискретних адитивних підгрупах, що задані у множині $\mathbb{R}^n$. Такі алгоритми забезпечують найкращу продуктивність, але є найменш консервативними серед усіх. Криптографія на основі решіток ґрунтується на вирішенні наступних обчислювально важких задач:

- Задача знаходження найкоротшого вектору (SVP, Shortest Vector Problem) – знайти в заданому базисі решітки ненульовий вектор по відношенню до визначеної нормалі.
- Задача знаходження ідеального найкоротшого вектору (ISVP, Ideal Shortest Vector Problem) не вважається NP-складною. Однак не існує відомих решіток, що засновані на методі редукції, значно більш ефективних на ідеальних структурах, чим на загальних.
- Задача знаходження (приблизно) найкоротшого незалежного вектору (SIVP, Shortest Independent Vector Problem), в якій є базис решітки B і необхідно знайти n лінійно незалежних векторів.
- Задача знаходження найближчого вектору (CVP, Closest Vector Problem) – знаходження вектору за заданим базисом і деяким вектором, який не належить решітці, при цьому максимально схожий за довжиною з заданим вектором.

Навіть з квантовим комп'ютером SVP представлений багаточленом у ступені n. Інші криптографічні алгоритми на основі алгебраїчних решіток засновані на проблемі короткого цілочисельного рішення (SIS).

Багатовимірна криптографія або багатовимірна криптографія відкритого ключа – це загальний термін, що описує асиметричні криптографічні схеми, побудовані на вирішенні рівнянь заснованих на багатовимірних поліномах над кінцевим полем F. Безпека багатовимірної криптографії заснована на припущенні, що вирішення системи квадратичних багаточленів над кінцевим полем F, в загальному випадку, є NP-повною задачею в сильному сенсі або просто NP-повна. Ці схеми часто вважаються гарними кандидатами для постквантової криптографії також через те, що вони пропонують найкоротші підписи.

Існують й інші класи постквантових криптографічних алгоритмів, такі як: криптографія на основі математичних кодів, ЕП на основі гешування та інші криптографічні методи. Але у даній роботі розглядаються лише алгоритми, що засновані на багатовимірній криптографії та криптографії на основі решіток, через те, що вони є найбільш використовуваними, що можна зрозуміти з табл. 1, де показано, що алгоритми ЕП цих двох класів мають найбільше представників у конкурсі NIST.

Таблиця 1

Математика ЕП і KEM, що представлені на конкурс PQC

Математика	Проблема, що вирішується	Підписи	KEM	Всього
Решітки	Пошук найкоротшого вектору, найбільш близького вектору	5 (3)	23 (9)	28 (12)
Коди	Декодування випадкового лінійного коду	3 (0)	17 (7)	20 (7)
Багатовимірна поліноміальна	Розв'язання багатовимірних квадратичних рівнянь	8 (4)	2 (0)	10 (4)
Геш	Стійкість до атаки знаходження прообразу геш-функції	3 (2)	0 (0)	3 (2)
Ізогенії	Пошук ізогенного відображення між еліптичними кривими з однаковою кількістю точок	0 (0)	1 (1)	1 (1)
Інші	–	2 (0)	5 (0)	7 (0)
Всього	–	21 (9)	48 (17)	69 (26)

Наведена кількість алгоритмів для оцінки NIST PQC раунду 1 і раунду 2 (всередині фігурних дужок).

1.1. Електронні підписи

Поточний процес стандартизації PQC NIST консолідує невразливих кандидатів після кожного наступного раунду. Кожен кандидат у конкурсі PQC NIST реалізує одну з трьох функцій: шифрування відкритого ключа, електронний підпис і механізм інкапсуляції ключа (KEM). Табл. 1 показує кількість представлень PQC раунду 2 і підсумовує їх функціональність і математичну складність.

NIST розраховує стандартизувати кілька алгоритмів, щоб забезпечити різні компроміси залежно від програми (швидкість і потужність пам'яті тощо). Пряме порівняння гарантій безпеки, що забезпечуються алгоритмами PQC-кандидатів, є складним завданням у відсут-

ність стандартної платформи квантових обчислень, відмінностей у математичних функціях, що лежать в основі алгоритмів, і складного алгоритму специфікації порівняно з попередніми криптографічними конкурсами [2].

Електронні підписи PQC працюють над принципом, що відправник підписує повідомлення за допомогою закритого ключа, а одержувач перевіряє підпис, використовуючи відкритий ключ відправника. Ці алгоритми використовують три функції:

- *crypto_sign_keypair* генерує відкритий ключ *pk* та секретний ключ *sk*;
- *crypto_sign* приймає *sk* і повідомлення *m* плюс його довжина *mlen* і видає підпис *sm,* доданий до повідомлення;
- *crypto_sign_open* приймає *pk, sm* та довжину *smlen* і видає повідомлення *m.*

1.2. Класифікація безпеки алгоритмів PQC

NIST виділяє п’ять категорій потужності безпеки:
Перший рівень безпеки ⇒ еквівалентний пошуку ключа AES-128.
Другий рівень безпеки ⇒ еквівалентний пошуку колізій SHA-256/SHA3-256.
Третій рівень безпеки ⇒ еквівалентний пошуку ключа AES-192.
Четвертий рівень безпеки ⇒ еквівалентний пошуку колізій SHA-384/SHA3-384.
П’ятий рівень безпеки ⇒ еквівалентний пошуку ключа AES-256.
Рівні безпеки PQC алгоритмів другого раунду NIST наведені у табл. 2.

2. Оцінка апаратних засобів PQC

Алгоритми ЕП PQC, їх реалізація і характеристики безпеки узагальнені в табл. 2. Розглядається програмна реалізація із використанням Xilinx Virtex-7 FPGA.

Розглядаються такі показники продуктивності: затримка, область і область затримки. Затримка – це час, необхідний системі для отримання вихідного сигналу з моменту подачі вхідного сигналу. Пропускна здатність – це максимальна швидкість, з якою можна забезпечити виведення. Мінімальна кількість тактових циклів між двома послідовними входами є інтервалом ініціювання і є мірою пропускної здатності [2].

Таблиця 2

Алгоритми ЕП PQC, які досліджуються

Алгоритм	Вирішувальна задача	Примітив PQC	Підтримуваний рівень безпеки (розмір відкритого ключа в байтах)				
			1	2	3	4	5
Crystals-Dilithium [6]	Решітки	Алгоритм ЕП	**1184**	1472	1760	–	–
MQDSS [7]	Багатовимірна криптографія	Алгоритм ЕП	–	**62**	–	88	–
qTESLA [8]	Решітки	Алгоритм ЕП	**1504** 14880	–	3104 2976 39712	–	–

Алгоритми, що не підтримують визначений рівень безпеки, позначені символом "–".

У табл. 3 і 4 представлені службові дані апаратних засобів і синхронізації для реалізації алгоритмів підпису, верифікації і генерації ключової пари PQC, відповідно, при синтезі без будь-яких додаткових обмежень (затримки).

Таблиця 3

Безпека в порівнянні з площею залежно від часу алгоритмів підпису PQC,
без оптимізації (тобто базового рівня)

Алгоритм	Рівень безпеки	Тригери	LUT	Тактовий цикл (нс)	Затримка
Алгоритми електронного підпису					
Crystals-Dilithium	1	25926	133461	10	609828
qTELSA	1	41978	232582	10	125374
MQDSS	2	35263	193320	15	49365597

Серед алгоритмів підпису рівня безпеки 1 Crystals-Dilitium і qTesla має затримки менше мільйона циклів. Crystals-Dilitium (для генерації підпису) є хорошим кандидатом для пристроїв Інтернету речей. Він має рівень безпеки 1 і є другим найшвидшим і другим найменшим серед алгоритмів підпису.

Таблиця 4

Співвідношення безпеки і площі та часу алгоритмів верифікації PQC
без оптимізації (тобто базового рівня)

Алгоритм	Рівень безпеки	Тригери	LUT	Тактовий цикл (нс)	Затримка
Алгоритми електронного підпису					
Crystals-Dilithium	1	20865	108878	10	5380
qTELSA	1	29875	168570	10	71223
MQDSS	2	26423	147359	15	25124450

Серед алгоритмів підпису qTesla і Crystals-Dilitium мають затримку менше мільйона циклів. Crystals-Dilitium є найшвидшим серед алгоритмів підпису. Crystals-Dilithium (для створення підпису) є хорошими кандидатами для пристроїв Інтернету речей. Жодна з верифікацій безпеки рівня 5 в цьому дослідженні не має низької затримки і, отже, не підходить для серверів. Всі алгоритми з низькою затримкою мають безпеку рівня 1.

У табл. 5 наведено та розглянуто критичні функції і цикли, які призводять до високої затримки.

Таблиця 5

Критичні функції алгоритмів електронного підпису

Алгоритм	Критична функція	# Цикли
Алгоритми електронного підпису		
Crystals-Dilithium	*expand_mat*	2
MQDSS	*crypto_sign*	2
qTELSA	*sparse_mul16*	1

Використовується розв'язка циклу для підпису і верифікації, результати наведені в табл. 6 і 7 відповідно. Crystals-Dilitium – найшвидший алгоритм підпису. Результати підпису з використанням конвеєрного циклу представлені в табл. 8. Серед алгоритмів з рівнем безпеки 1 Crystals-Dilitium – найшвидший алгоритм підпису.

Таблиця 6

Співвідношення безпеки і площі порівняно із синхронізацією алгоритмів підпису PQC,
після завершення циклу

Алгоритм	Рівень безпеки	Тригери	LUT	Тактовий цикл (нс)	Затримка
Алгоритми електронного підпису					
Crystals-Dilithium	1	158313	584742	10	18525
qTELSA	1	97235	328106	10	59854
MQDSS	2	45135	230273	15	49365597

Crystals-Dilitium посідає найбільше місце серед алгоритмів підпису. Серед алгоритмів підпису PQC розгортання циклу не зменшує затримки MQDSS. Розгортання циклу зменшує затримку всіх алгоритмів підпису PQC. Однак це також призводить до збільшення площі. Серед інших алгоритмів високої безпеки MQDSS (підпис, рівень безпеки 2) може використовуватися для пристроїв Інтернету речей, оскільки забезпечує відносно низьке апаратне навантаження.

Таблиця 7

Залежність безпеки від площі залежно від часу алгоритмів верифікації PQC,
після завершення циклу

Алгоритм	Рівень безпеки	Тригери	LUT	Тактовий цикл (нс)	Затримка
Алгоритми електронного підпису					
Crystals-Dilithium	1	108154	388991	10	5380
qTELSA	1	77567	247726	10	36423
MQDSS	2	93945	323734	15	25084906

Розгортання циклу забезпечує суттєве скорочення затримок. За винятком MQDSS, алгоритм не має затримки понад 1 мільйон циклів. Подібно підпису, розгортання циклу зменшує затримку верифікації PQC. Це поставляється з додатковим обладнанням. Для пристроїв Інтернету речей, де високий рівень безпеки не потрібен, можна використовувати алгоритм рівня безпеки 1 з низьким рівнем службових даних, наприклад, qTesla (підпис). З іншого боку, Crystals-Dilitium (підпис), що забезпечує безпеку рівня 1, може використовуватися в серверах через його низьку затримку.

Результати підпису з використанням конвеєрного циклу представлені в табл. 8.

Таблиця 8

Співвідношення безпеки і площі порівняно із синхронізацією алгоритмів ЕП PQC, після конвеєризації циклу

Алгоритм	Рівень безпеки	Тригери	LUT	Тактовий цикл (нс)	Затримка
Алгоритми електронного підпису					
Crystals-Dilithium	1	146076	1327355	10	155166
qTELSA	1	112657	346020	10	63736
MQDSS	2	47441	270713	15	25825918

Подібно до розгортання циклу, конвеєризація також зменшує загальну затримку для алгоритмів підпису PQC. Основна відмінність з таблицею 6 пов'язана з алгоритмом підпису MQDSS. Хоча при розгортанні циклу не вдалося змінити його затримку, конвеєрна обробка може скоротити затримку на 50 %. Конвеєрний цикл зменшує затримку алгоритмів підпису PQC, збільшуючи апаратну область. Поліпшення затримки порівняно з розгортанням циклу не погоджено. Жоден з алгоритмів підпису не забезпечує низьку затримку після конвеєрної обробки [2].

3. Поглиблені дослідження з використанням трьох підписів PQC

У цьому розділі розглянуто три алгоритми PQC на основі підпису і проаналізовано апаратні реалізації на різних рівнях безпеки. Крім того, розглянуто вплив різних оптимізацій конструкції як для процедури підпису, так і для процедури верифікації. Всі досліджувані реалізації виконані на платі Xilinx Artix-7 FPGA. Це платформа, яку NIST розглядає в якості платформи порівняння [2, 3].

3.1. Порівняння за рівнем безпеки 1

Порівнюється область і продуктивність трьох алгоритмів для рівня безпеки 1. Аналізовані параметри – кількість тригерів, кількість LUT, затримка і LAP. Порівняння частин підпису та верифікації наведено в табл. 9 і 10 [2 – 4].

Таблиця 9

Аналіз апаратних реалізацій компонентів "crypto_sign" алгоритмів PQC на основі підпису

Алгоритм	FF	LUT	Такти (нс)	Затримка	LAP
qTesla	26299	126732	12.65	537092	6.8×10^{10}
Crystals-Dilithium	27132	123655	8.738	485963	6.0×10^{10}
MQDSS	21841	106035	17.05	34502428	3.6×10^{10}

У той час як накладні витрати на розмірі більш-менш однакові (MQDSS займає на 83 % менше місця, ніж інші), загальна затримка і LAP для MQDSS надзвичайно високі. MQDSS займає в 71 раз більше затримки і в 61 раз більше LAP порівняно з Crystals-Dilitium. На рівні безпеки 1, Crystals-Dilitium є найкращим алгоритмом для "crypto_sign", оскільки займає найменший LAP.

Таблиця 10

Аналіз апаратних реалізацій "crypto_sign_open" компонентів алгоритмів PQC на основі підпису

Алгоритм	FF	LUT	Такти (нс)	Затримка	LAP
qTesla	17780	87067	12.58	80422	7.0×10^{9}
Crystals-Dilithium	14712	63863	8.738	149950	9.5×10^{9}
MQDSS	23072	117097	17.045	25686731	3.0×10^{12}

Crystals-Dilitium має найменший розмір, тоді як qTesla має найменшу затримку і показує кращу продуктивність. На рівні безпеки 1 qTesla є найкращим алгоритмом для "crypto_sign_open", оскільки займає найменший LAP [2].

3.2. Порівняння за рівнем безпеки 3

У цьому розділі порівнюється вартість і продуктивність впровадження апаратного забезпечення як для "crypto_sign", так і для "crypto_sign_open" для трьох алгоритмів на основі підпису на рівні безпеки 3. Результати для підпису та верифікації наведено у табл. 11 та 12 відповідно.

Таблиця 11

Аналіз апаратних реалізацій "crypto_sign" компонентів алгоритмів PQC на основі підпису на рівні безпеки 3

Алгоритм	FF	LUT	Такти (нс)	Затримка	LAP
qTesla	26011	126311	12.65	347655	4.3×10^{10}
Crystals-Dilithium	27308	123933	8.738	826832	1.0×10^{11}
MQDSS	24748	123170	16.378	119353597	1.4×10^{13}

У той час як накладні витрати на розмір більш-менш однакові для всіх алгоритмів, загальна затримка і LAP для MQDSS надзвичайно високі. На рівні безпеки 3 qTesla є найкращим алгоритмом для "crypto_sign", оскільки займає найменший LAP [2, 3].

Таблиця 12

Аналіз апаратних реалізацій "crypto_sign_open" компонентів алгоритмів PQC на основі підпису

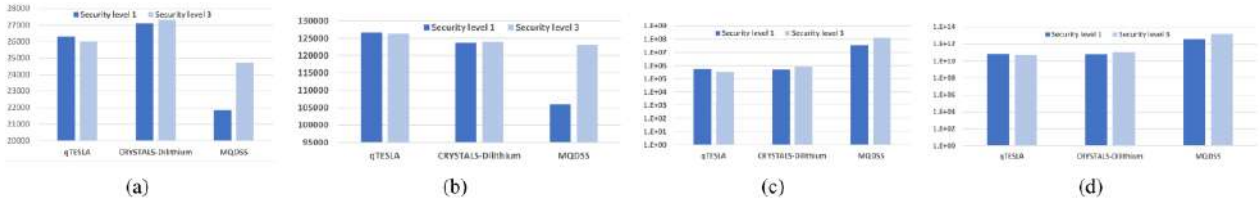
Алгоритм	FF	LUT	Такти (нс)	Затримка	LAP
qTesla	17754	86142	12.65	201027	1.7×10^{10}
Crystals-Dilithium	14783	63980	8.738	297592	1.9×10^{10}
MQDSS	23149	117574	17.045	87861777	1.0×10^{13}

Crystals-Dilitium має найменший розмір, тоді як qTesla має найменшу затримку. На рівні безпеки 3 qTesla є найкращим алгоритмом для "crypto_sign_open", оскільки займає найменший LAP. З іншого боку, MQDSS має найгірші показники LAP [2].

3.3. Співставлення між різними рівнями безпеки

У цьому розділі порівнюються службові дані області та продуктивність алгоритмів на різних рівнях безпеки [2].

Порівнюється область (у термінах тригерів і LUT), продуктивність (у термінах затримки) і LAP для реалізації "crypto_sign" частини трьох алгоритмів для рівнів безпеки 1 і 3. Результати показано на рис. 1. Для qTesla і Crystals-Dilitium площа накладних витрат на рівнях безпеки 1 і 3 аналогічна. Для MQDSS, накладні витрати на рівень безпеки 3 на 13 – 16 % більше, ніж на рівень безпеки 1. Для qTesla затримка знижується на рівні безпеки 3 порівняно з рівнем 1. Для двох інших алгоритмів затримка різко збільшується при більш високому рівні безпеки. Порівнюючи як площу, так і продуктивність, можна помітити, що тільки qTesla має більш низькі LAP на рівні безпеки 3 порівняно з рівнем 1. qTesla на рівні безпеки 3 має найнижчий LAP і забезпечує найвищу безпеку серед усіх альтернатив, що обговорюються в цьому підрозділі.



(a) (b) (c) (d)

Рис. 1. Порівняння (*a*) тригерів, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації компонента "crypto_sign" трьох алгоритмів PQC на різних рівнях безпеки

Для qTesla і Crystals-Dilitium різниця в кількості тригерів не є істотною, в той час як для MQDSS кількість тригерів для реалізації на рівні безпеки 3 збільшується на 13% порівняно з рівнем безпеки 1. Для qTesla і Crystals-Dilitium різниця в кількості LUT не є істотною, в той час як для MQDSS кількість тригерів для реалізації на рівні безпеки 3 збільшується на 16 % порівняно з рівнем безпеки 1. Оскільки затримка MQDSS на два порядку більша, ніж qTesla або Crystals-Dilitium, цей графік зображено в логарифмічній шкалі. qTesla на рівні безпеки 3 і Crystals-Dilitium на рівні безпеки 1 мають найменшу затримку. qTesla на рівні безпеки 3 має найкращий LAP, а MQDSS на обох рівнях безпеки має найгірший LAP.

Далі порівнюється область, продуктивність і LAP для верифікації підпису, тобто компонент "crypto_sign_open" трьох алгоритмів на двох рівнях безпеки. Результати показано на рис. 2. Для всіх алгоритмів службові дані області залишаються однаковими для двох рівнів безпеки. З іншого боку, послідовно затримка і LAP вища на рівні безпеки 3 порівняно з рівнем безпеки 1. qTesla має найменший LAP на обох рівнях безпеки.

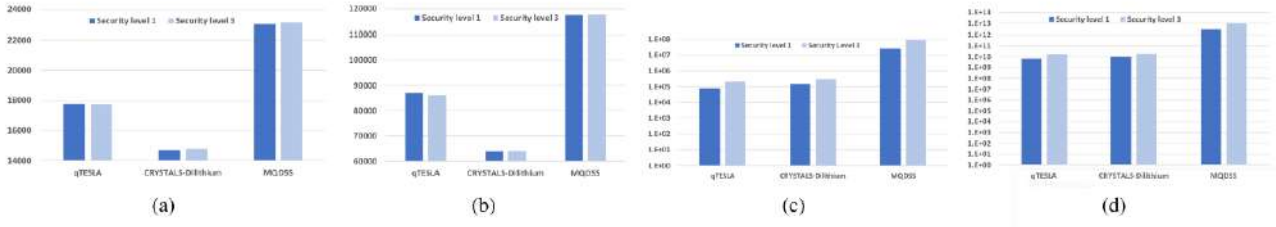


Рис. 2. Порівняння (*a*) FF, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації компонента "crypto_sign_open" трьох алгоритмів PQC на різних рівнях безпеки

У той час як на рис. 1 Crystals-Dilitium займає найбільшу кількість тригерів, при реалізації "crypto_sign_open", потрібна найменша кількість тригерів. Порівняно з рис. 1 різниця в кількості LUT не є значною для жодного алгоритму. qTesla на рівні безпеки 1 вимагає найменшої затримки. MQDSS вимагає найбільших затримок для "crypto_sign" і "crypto_sign_open." qTesla на рівні безпеки 1 має найкращий LAP, а MQDSS на обох рівнях безпеки має найгірший LAP [2].

3.4. Оптимізація

У цьому підрозділі проаналізовано, як різні методи оптимізації (розбиття циклу і конвеєрування циклу) допомагають зменшити загальну затримку алгоритмів PQC. Результати реалізації "crypto_sign" і "crypto_sign_open" показані на рис. 3 і 4 відповідно.

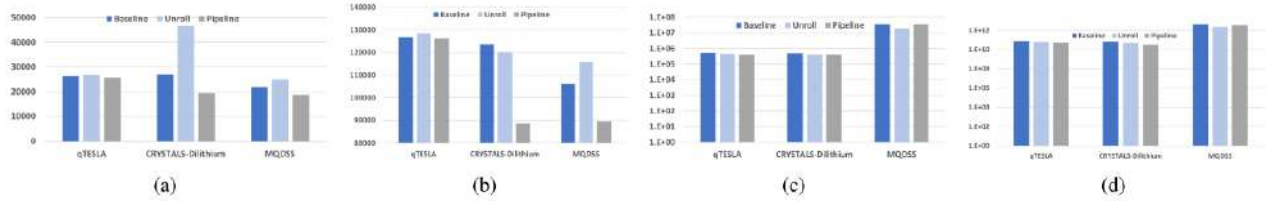


Рис. 3. Порівняння (*a*) FF, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації "crypto_sign" трьох алгоритмів PQC на рівні безпеки 1 для різних оптимізацій

Конвеєрний цикл дає найменшу кількість тригерів по всьому циклу, в той час як розгін циклу вимагає максимальної кількості тригерів. Конвеєрний цикл дає найменшу кількість LUT для всіх конструкцій. За винятком Crystals-Dilitium, кількість LUT збільшується з розмотуванням циклів. Розгортання циклу і конвеєрна обробка зменшують затримку порівняно з базовою лінією. Crystals-Dilitium і qTesla мають найменше значення LAP при конвертуванні циклів. Для MQDSS найменше значення LAP виходить, коли цикли не розгорнуті [2, 3].

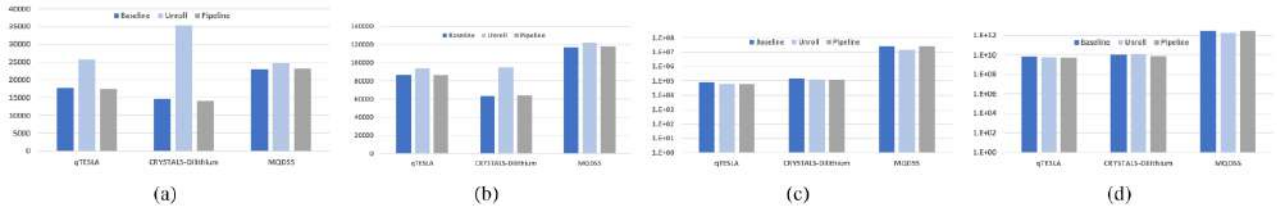


Рис. 4. Порівняння (*a*) FF, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації компонента "crypto_sign_open" трьох алгоритмів PQC на рівні безпеки 1 для різних оптимізацій

Оптимізація з конвеєрними циклами займає найменшу кількість тригерів і LUT у всіх конструкціях, в той час як розукрупнення циклу вимагає максимальної кількості тригерів і LUT. Оптимізація за допомогою розмотування циклів і конвеєрів дозволяє скоротити затримки порівняно з базовою реалізацією. За винятком MQDSS, два інших алгоритми мають найменше значення LAP при виконанні конвеєризації циклу. Для MQDSS найменше значення LAP виходить, коли цикли не розгорнуті [2].

3.5. Наслідки апаратного оптимізованого Keccak

Keccak – сімейство sponge-функцій, що використовуються три алгоритми другого раунду NIST PQC. Протягом багатьох років дослідники розробляли апаратно оптимізовані варіанти Keccak. У цьому підрозділі розглянемо вплив одного з них і спостерігаємо, як змінюється розмір і накладні витрати на продуктивність. Використовується реалізація Keccak [3]. Автори [3] розвивали впровадження функції *KeccakF1600_StatePermute* з окремими функціями, що використовуються для внутрішніх операцій як θ, ρ, π, χ та ι. На рис. 5 порівнюються параметри для реалізації "crypto_sign", а на рис. 6 порівнюється "crypto_sign_open". Для забезпечення узгодженості всі реалізації належать до рівня безпеки 1. На рис. 7, 8 порівнюються накладні витрати з розмотуванням циклів. Для чесного порівняння використовуються одні й ті ж директиви для обох реалізацій. Аналогічне порівняння для конвеєризації циклу показано на рис. 9, 10 [3].

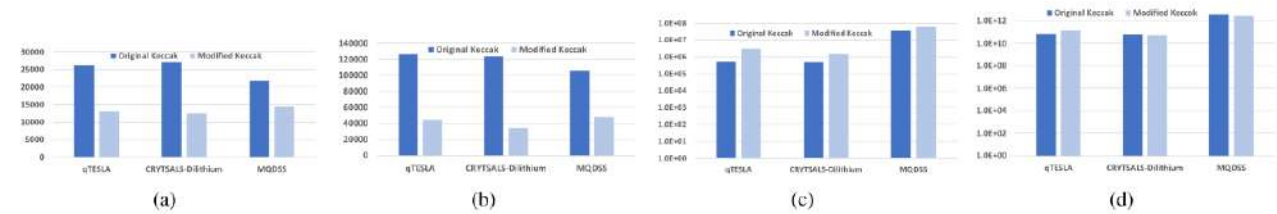


Рис. 5. Порівняння (*a*) FF, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації компонента "crypto_sign" трьох алгоритмів PQC на рівні безпеки 1 для двох версій Keccak

Кількість тригерів і LUT значно зменшується порівняно з початковим Keccak. Crystals-Dilitium має 54 % зниження кількості тригерів і 73% зниження LUT. З іншого боку, затримка збільшується з модифікованим Keccak. Найсильніше постраждав qTesla, який має збільшення затримки в 5,8 разів, в той час як MQDSS несе тільки збільшення затримки в 1,7 рази. З іншого боку, значення LAP зменшується для MQDSS і Crystals-Dilitium і збільшується тільки для qTesla. Для MQDSS LAP скорочується на 21 % [3].

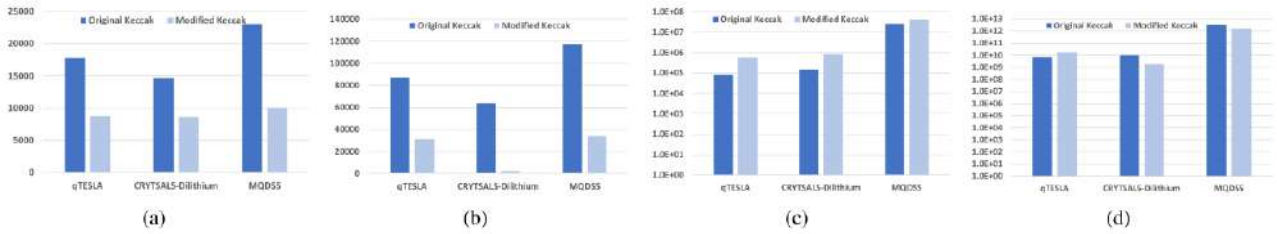


Рис. 6. Порівняння (*a*) тригерів, (*b*) LUT, (*c*) затримки і (*d*) LAP для реалізації компонента "crypto_sign_open" трьох алгоритмів на рівні безпеки 1 для різних версій Keccak

Як і на рис. 5, кількість тригерів і LUT значно зменшується порівняно з вихідним Keccak. MQDSS має максимальне 56 % зниження кількості тригерів і Crystals-Dilitium має близько 96 % зниження LUT. Затримка збільшується з модифікованим Keccak. Найсильніше це стосується qTesla, яка має збільшення затримки в 7 разів, в той час як MQDSS має збільшення затримки в 1,7 рази. LAP значно зменшується для всіх алгоритмів, крім qTesla. Crystals-Dilitium має 82 % зниження LAP порівняно з вихідним Keccak.

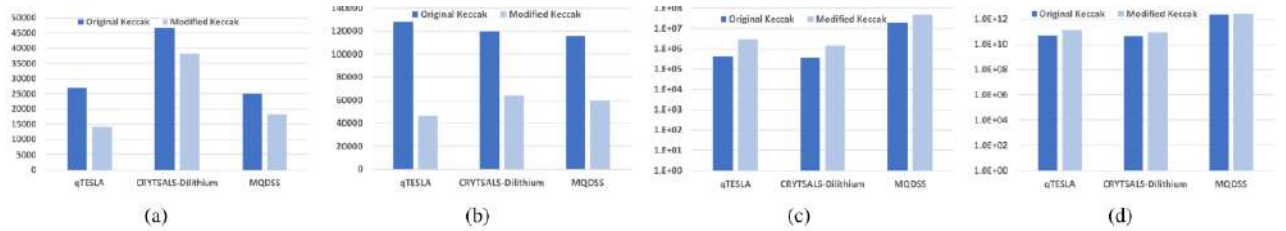


Рис. 7. Порівняння (*a*) тригери, (*b*) LUT, (*c*) Latency і (*d*) LAP для реалізації компонента "crypto_sign" трьох алгоритмів PQC на рівні безпеки 1 для різних версій Keccak після завершення циклу

Кількість тригерів і LUT значно зменшується порівняно з оригінальною версією Keccak. qTesla має приблизно 48 % зниження кількості тригерів і 64% скорочення кількості LUT. З іншого боку, затримка збільшується з реалізацією модифікованого Keccak. Найсильніше вражає qTesla, яка має збільшення затримки в 7,15 разів. На відміну від рис. 5, значення LAP послідовно збільшуються для всіх трьох алгоритмів. Для MQDSS збільшення LAP мінімально – 20 % [3].

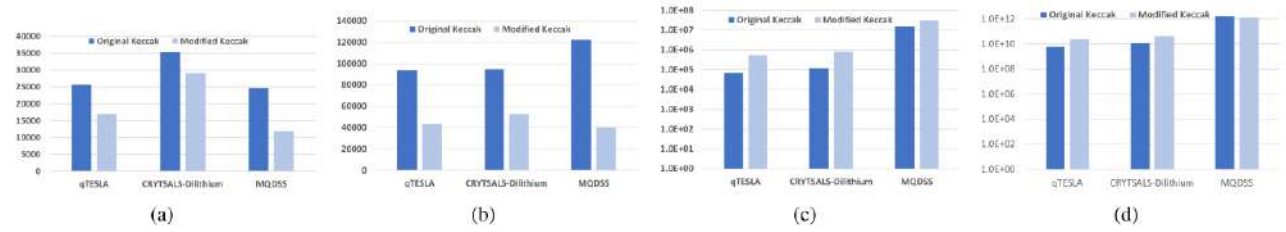


Рис. 8. Порівняння (*a*) тригери, (*b*) LUT, (*c*) LAP і (*d*) LAP для реалізації компонента "crypto_sign_open" трьох алгоритмів PQC на рівні безпеки 1 для різних Keccak після розгортання циклу

Кількість тригерів і LUT значно зменшується порівняно з початковим Keccak. MQDSS має приблизно 52 % зниження кількості тригерів і 61 % зменшення кількості LUT. З іншого боку, затримка збільшується з модифікованим Keccak. Найсильніше постраждав qTesla, який має збільшення в 8,4 рази. LAP збільшується для qTesla і Crystals-Dilitium, але знижується на 28 % для MQDSS.

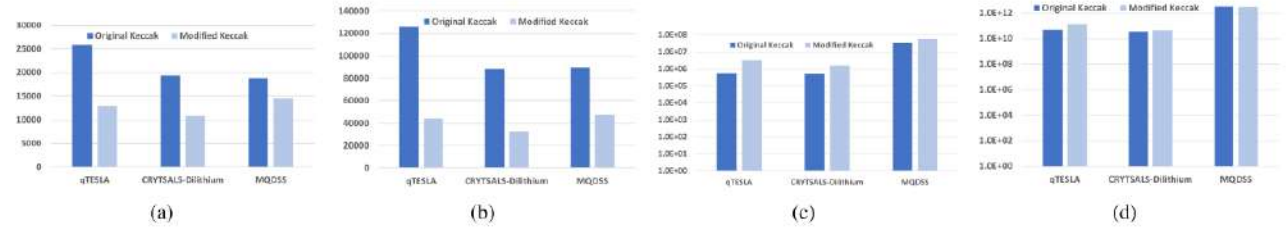


Рис. 9. Порівняння (*a*) тригери, (*b*) LUT, (*c*) затримка і (*d*) LAP для реалізації компонента "crypto_sign" трьох алгоритмів PQC на рівні безпеки 1 для різних версій Keccak, після конвеєризації циклу

Скорочення кількості тригерів і LUT значно порівняно з оригінальним Keccak. qTesla має 50 % зниження кількості тригерів і 65 % скорочення кількості LUT. З іншого боку, затримка збільшується з реалізацією модифікованого Keccak. Найсильніше це торкнулося

qTesla, що має збільшення затримки в 7,7 разів. Значення LAP збільшуються для qTesla і Crystals-Dilitium, але зменшуються на 5 % для MQDSS. В цілому, відносне збільшення LAP менше, ніж при розгортанні циклу на рис. 7.

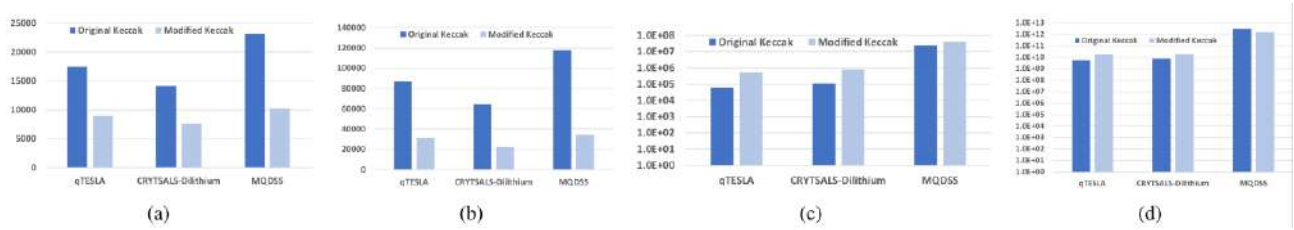


Рис. 10. Порівняння (*a*) тригери, (*b*) LUT, (*c*) затримка і (*d*) LAP для реалізації "crypto_sign_open" з трьох алгоритмів на рівні безпеки 1 з використанням різних версій Keccak, після конвеєризації циклу

Кількість тригерів і LUT значно зменшується порівняно з базовим Keccak. MQDSS має 56 % зниження кількості тригерів і 70 % скорочення кількості LUT. Хоча затримка MQDSS збільшується на 67 % при модифікованому Keccak, загальне значення LAP зменшується на 52 % [2].

Висновки

1. Наразі однією із важливих проблем сучасної криптографії є створення стандартів асиметричних криптоперетворень ЕП, які були б безпечними у постквантовий період. Вирішення цієї проблеми здійснюється в процесі міжнародного конкурсу NIST, завданням якого є розробка такого ЕП, який би був стійким як до квантових, так і до класичних атак.
2. Були проведені та проаналізовані специфічні тематичні дослідження з використанням трьох алгоритмів підпису – qTesla, Crystals-Dilitium і MQDSS. Відзначено розмір і продуктивність накладних витрат, коли апаратні реалізації цих алгоритмів виконуються для різних рівнів безпеки, а також для різних оптимізацій.
3. Алгоритми ЕП qTesla і Crystals-Dilitium мали найменший LAP при конвеєрному циклі, в той час як MQDSS мав найменший LAP при розгортанні циклу. Майбутні дослідження визначать, яка оптимізація підходить для якого алгоритму.
4. Навіть, якщо одні й ті самі директиви оптимізації застосовуються до двох різних версій Keccak, модифікована функція Keccak зменшує загальні витрати області. Однак у цьому випадку LAP знижуються не так сильно (фактично LAP збільшується в більшості випадків), порівняно з базовою лінією (без оптимізації). Зменшення LAP більше для конвеєрного циклу порівняно з розмотуванням циклу.
5. У цьому дослідженні було використано одну і ту ж плату (Virtex або Artix) для реалізації варіантів алгоритмів PQC як з маленьким розміром, так і з низькою затримкою. У майбутньому буде вивчено, які архітектури FPGA ідеально підходять для реалізації з маленьким розміром/низькою затримкою.

Список літератури:

1. PQC Standardization Process: Third Round Candidate Announcement. July 22, 2020. [Electronic resource]. Access mode: https://csrc.nist.gov/News/2020/pqc-third-round-candidate-announcement.
2. Kanad Basu NIST Post-Quantum Cryptography – A Hardware Evaluation Study / Kanad Basu, Deepraj Soni, Mohammed Nabeel, Ramesh Karri // Access mode: https://eprint.iacr.org/2019/047.pdf.
3. K. Gaj Comprehensive comparison of hardware performance of fourteen round 2 SHA-3 candidates with 512-bit outputs using field programmable gate arrays / K. Gaj, E. Homsirikamol, M. Rogawski // 2nd SHA-3 Candidate Conference, Santa Barbara, August, pp. 23–24, 2010.
4. J. Nechvatal Report on the development of the advanced encryption standard (AES) / J. Nechvatal, E. Barker, L. Bassham, W. Burr, M. Dworkin, J. Foti, E. Roback // 2001. Access mode: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4863838/.
5. K. Aoki Fast implementations of AES candidates. / K. Aoki, H. Lipmaa // AES Candidate Conference, pp. 106–120, 2000.
6. L. Ducas Crystals-Dilithium: Digital signatures from module lattices / L. Ducas, T. Lepoint, V. Lyubashevsky, P. Schwabe, G. Seiler, D. Stehle // Access mode: https://pq-crystals.org/dilithium/index.shtml, 2018.

7. Ming-Shing Chen From 5-pass MQ-based identification to MQ-based signatures / Ming-Shing Chen, A. Hulsing, J. Rijneveld, S. Samardjiska, P. Schwabe // Access mode: https://eprint.iacr.org/2016/708.pdf.

8. Erdem Alkim The Lattice-Based Digital Signature Scheme qTesla / Erdem Alkim, Paulo S. L. M. Barreto, Nina Bindel, Juliane Kramer, Patrick Longa, Jefferson E. Ricardini // Access mode: https://eprint.iacr.org/2019/085.pdf.

Надійшла до редколегії 07.04.2021

Відомості про авторів:

Єсіна Марина Віталіївна – канд. техн. наук, Харківський національний університет імені В.Н. Каразіна, старший викладач кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, Україна; e-mail: rinayes20@gmail.com; ORCID: https://orcid.org/0000-0002-1252-7606

Шахов Богдан Сергійович – Харківський національний університет імені В.Н. Каразіна, студент кафедри безпеки інформаційних систем і технологій, факультету комп'ютерних наук, Україна; e-mail: bogdanshahov2000@gmail.com

УДК 004.056 **DOI:10.30837/rt.2021.2.205.05**

В.В. ВИЛИГУРА

АНАЛИЗ ФОРМАЛЬНЫХ МОДЕЛЕЙ УПРАВЛЕНИЯ ДОСТУПОМ И ОСОБЕННОСТИ ИХ ПРИМЕНИМОСТИ ДЛЯ БАЗ ДАННЫХ

Введение

Обеспечить безопасность информационной системы (ИС) легче, если есть четкая модель того, что нужно защищать и кому и что разрешено делать [1]. Поэтому неотъемлемой частью любого проекта по созданию или оценке безопасности ИС, в том числе и баз данных (БД), как отмечается в работе [2], является наличие *модели безопасности*, под которой понимается формальное представление политики безопасности [3, 4]. Основная цель создания политики безопасности ИС и описания ее в виде формальной модели – это определение условий, которым должно подчиняться поведение системы, определение показателей и критерия безопасности, а также проведение формального доказательства соответствия защищаемой системы этому критерию при соблюдении установленных правил и ограничений [4].

Формальная модель политики безопасности, представленная в виде математических выражений, схем, диаграмм, алгоритмов, играет важную роль в процессах разработки и исследования информационных систем в целом и БД в частности, так как обеспечивает системный подход. Модели безопасности позволяют решить задачи, возникающие в ходе разработки и исследования защищенных систем. Так, используя эти модели, заказчики могут в четко определенной форме формулировать требования к создаваемым защищенным ИС, которые соответствуют политике безопасности, принятой в организации, а также оценить соответствие защищенных систем своим потребностям. Разработчики на основе моделей безопасности составляют спецификацию политики безопасности разрабатываемой системы. Эксперты используют модели безопасности в качестве эталонов в ходе анализа адекватности реализации политики безопасности в защищенных системах. На основе этих моделей они создают методики оценки защищенности конкретных ИС, осуществляют сертификацию разработанных систем по требованиям защиты информации. Благодаря формальным моделям можно, опираясь на объективные и неопровержимые положения математической теории, доказать безопасность системы. Разработка соответствующей модели – это первый шаг на пути к созданию теоретических основ обеспечения безопасности в ИС [5].

В данной работе рассмотрим основные положения наиболее распространенных моделей безопасности, основанных на контроле доступа субъектов к объектам, и общем критерии безопасности ИС, который формулируется следующим образом: информационная система безопасна тогда и только тогда, когда субъекты не имеют никаких возможностей нарушать (обходить) установленную в ней политику безопасности.

Модели управления доступом

Наиболее распространенной парадигмой моделей обеспечения безопасности данных в целом и моделей управления доступом, в частности, является субъектно-объектная абстракция. Формально субъектно-объектная модель ИС это тройка: $\langle S, O, Op \rangle$, в которой $S = \{s_1, s_2, ..., s_m\}$ – субъекты (пользователи, процессы); $O = \{o_1, o_2, ..., o_n\}$ – объекты; $Op = \{op_1, op_2, ..., op_L\}$ – множество операций над объектами. Под *субъектом* понимается активная сущность ИС, которая может изменять состояние системы посредством порождения процессов над объектами, в том числе, порождать новые объекты и инициировать порождение новых субъектов. Под *объектом* понимается пассивная сущность ИС, процессы над которой могут в определенных случаях быть источником порождения новых субъектов [6]. Для баз данных такими объектами могут быть следующие: таблицы, представления (англ. view), домены (типы), атрибуты, кортежи, процедуры, функции, триггеры, синонимы, про-

фили и некоторые другие. Активность понимается как возможность выполнять операции над объектами (пассивными сущностями).

Остановимся более подробно на моделях управления доступом:

- модели безопасности на основе дискреционной политики;
- модели безопасности на основе мандатной политики;
- модели безопасности на основе ролевой политики.

Модели безопасности на основе дискреционной политики

Работы по моделям дискреционного доступа к информации в ИС появились еще в 60 – 70-х годах прошлого столетия. Они достаточно широко освещены в научной литературе. Наиболее известные из них – модель ADEPT-50 [7], пятимерное пространство Хартсона [8], модель Хариссона – Руззо – Ульмана [9], модель Take-Grant [10]. Авторами этих моделей был внесен значительный вклад в теорию безопасности компьютерных систем. Их работы заложили основу для последующего создания и развития защищенных ИС.

В теоретическом и практическом плане наибольшее развитие и применение получили дискреционные модели, основанные на *матрице доступа* – таблице (M), описывающей права доступа субъектов (S) к объектам (O), строки которой соответствуют субъектам доступа $s_1, s_2, ..., s_m$, столбцы – объектам доступа $o_1, o_2, ..., o_n$, а в ячейках (элементах матрицы $M[s_i, o_j]$) записываются разрешенные операции (виды доступа) $op_1, op_2, ..., op_L$ соответствующего субъекта над соответствующим объектом.

В приведенной на рис. 1 матрице доступа M виды доступа op_l ($l = 1...L$) допускают следующие операции (виды доступа): чтение (r_d), запись с модификацией (w), запись без модификации (только с новой записью или дописыванием в файл) (a), запуск объекта на выполнение (e). Однако, как отмечается в [5], при необходимости элементы матрицы могут содержать указатели на процедуры. Эти процедуры исполняются при каждой попытке доступа к заданному объекту. Тем самым решение о доступе может приниматься на основании более сложных зависимостей не столь очевидных, как в простой матрице доступа.

	o_1	o_2	…	o_i		o_n
s_1	r_d	r_d, w		e		r_d
s_2	r_d, a	-		r_d		e
…						
s_i	r_d	-		-		r_d
…						
s_m	r_d, w	-		e		e

Рис. 1. Матрица доступа M

Данная модель предполагает, что все попытки доступа к объектам перехватываются и проверяются специальным управляющим процессом. Таким образом, субъект s_i получит инициируемый им доступ op_l к объекту o_j только в случае, если элемент матрицы $M[s_i, o_j]$ имеет значение op_l.

По принципу управления доступом выделяются два подхода [2, 6]:

- принудительное управление доступом;
- добровольное управление доступом.

Принцип принудительного управления доступом основывается на парадигме доверенных субъектов. Согласно этому принципу право создания и изменения матрицы доступа имеют только субъекты администратора системы, который при регистрации для работы в

системе нового пользователя создает с соответствующим заполнением новую строку матрицы доступа, а при возникновении нового объекта, подлежащего избирательному доступу, образует новый столбец матрицы доступа.

Как можно заметить, принудительный способ обеспечивает жесткое централизованное управление доступом. Как итог, ему не хватает гибкости и точности настройки системы разграничения доступа в зависимости от потребностей и полномочий пользователей, которые имеют наиболее полное представление о содержимом и конфиденциальности объектов (активов), как их владельцы.

Принцип добровольного управления доступом основывается на понятии владения объектами. *Владельцем* объекта o_j доступа называется субъект s_i, инициализировавший процесс, в результате которого объект возник в системе. Тем самым, в дополнение к основным положениям субъектно-объектной модели вводится специальное отображение множества объектов на множество субъектов доступа, называемое владением, ставящее в каждый фиксированный момент времени каждому объекту системы подмножество субъектов доступа, инициализированных пользователем-владельцем объекта.

Добровольное управление доступом выражается следующим правилом: *права доступа к объекту определяют (устанавливают) их владельцы*. Из данного правила следует, что формирование элементов матрицы доступа осуществляют субъекты, являющиеся владельцами соответствующих объектов.

На практике в большинстве случаев, в том числе и для БД, применяется комбинированный способ управления доступом, когда определенная часть полномочий на доступ к объектам устанавливается администратором (привилегированным пользователем), а другая часть – владельцами объектов.

Во многих ИС права владения объектами могут передаваться. В результате при добровольном управлении доступом реализуется полностью децентрализованный принцип управления процессом разграничения доступа. Такой подход обеспечивает гибкость формирования правил разграничения доступа для конкретной совокупности пользователей к активам, но приводит к усложнению общего контроля состояния безопасности активов в системе [6]. Это, в свою очередь, требует дополнительного исследования условий и процессов распространения прав доступа.

В теоретическом плане впервые данная проблема была исследована Харрисоном, Руззо и Ульманом, они разработали специальную формальную модель дискреционного доступа (сокращенно модель HRU) [9].

В данной модели дополнительно к субъектам S, объектам O (для того чтобы включить в область действия модели и отношения между субъектами, принято считать, что все субъекты одновременно являются и объектами: $S \subseteq O$) и конечному набору прав доступа $R = (r_1, r_2, ..., r_K)$ ($M[s,o] \subseteq R$) вводится пространство состояний системы $Q = (S, O, M)$. Пространство состояний системы образуется декартовым произведением множеств составляющих ее объектов, субъектов и прав: $S \times O \times M$. Любой элемент матрицы M содержит набор прав субъекта s_i к объекту o_j, принадлежащих множеству прав доступа R. Поведение системы во времени рассматривается как последовательность состояний $\{Q_v\}$, каждое последующее состояние является результатом применения некоторой команды ($\alpha_\tau \in C$, где C – конечный набор команд) к предыдущему: $Q_{v+1} = \alpha_\tau(Q_v)$.

Критерий безопасности в модели HRU формулируется следующим образом.

Система является безопасной относительно права r_k ($k = 1...K$; для простоты часто в дальнейшем вместо r_k будем использовать обобщенное обозначение $r \in R$), если для заданного начального состояния $Q_0 = (S_0, O_0, M_0)$ не существует применимой к Q_0 последова-

тельности команд, в результате которой право r будет занесено в элемент матрицы M, в которой оно отсутствовало в начальном состоянии Q_0.

Другими словами, это означает, что субъект никогда не получит право доступа r к объекту, если он не имел его изначально. Если же право r оказалось в ячейке матрицы M, в которой оно изначально отсутствовало, то говорят, что произошла утечка права r [11].

Из критерия безопасности следует, что для данной модели ключевую роль играет выбор значений прав доступа и их использование в условиях команд. По сути, данная модель описывает не только доступ субъектов к объектам, а распространение прав доступа от субъекта к субъекту, поскольку именно изменение содержания элементов матрицы доступа определяет возможность выполнения команд, в том числе команд, модифицирующих саму матрицу доступа, которые потенциально могут привести к нарушению критерия безопасности.

Однако Харрисон, Руззо и Ульман доказали, что в общем случае не существует алгоритма, который может для произвольной системы, ее начального состояния $Q_0 = (S_0, O_0, M_0)$ и общего права r решить, является ли данная конфигурация безопасной. Помимо проблем с неопределенностью распространения прав доступа в системах на основе модели HRU была подмечена еще одна серьезная проблема – уязвимость по отношению к атаке с помощью «троянского коня». В таких моделях контролируются только операции доступа субъектов к объектам, а не потоки информации между ними. Поэтому, например, когда некоторая троянская программа переносит информацию из доступного некоторому пользователю объекта в объект, доступный злоумышленнику (например, из одной таблицы в другую), то формально никакое правило дискреционной политики безопасности не нарушается, но утечка информации происходит.

Таким образом, дискреционная модель HRU в своей общей постановке не дает гарантий безопасности системы. Однако именно она послужила основой для целого класса моделей политик безопасности таких, например, как модель типизованной матрицы доступа (модель Type Access Matrix – TAM [12]), модель TAKE-GRANT [10] и некоторых других, которые используются для управления доступом и контроля за распространением прав в различных системах [4]. Развитие моделей дискреционного управления доступом заключается преимущественно в построении всевозможных модификаций модели HRU, а также в поиске минимально возможных ограничений, которые можно наложить на описание системы, чтобы вопрос ее безопасности был вычислительно разрешимым [11].

Так, для смягчения условий, в которых можно проводить формальное доказательство безопасности, а также для введения контроля за порождением объектов, как отмечается в работе [6], была предложена модель TAM. Модель типизованной матрицы доступа является обобщением модели HRU, которую можно рассматривать как частный случай TAM с одним единственным типом для всех объектов и субъектов [4]. С другой стороны, любую систему TAM можно выразить через систему HRU, введя для обозначения типов специальные права доступа, а проверку типов в командах заменив проверкой наличия соответствующих прав доступа. Тем не менее введение строгого контроля типов ($T = (t_1, t_2, ..., t_\Theta)$, с одним из которых создается любой объект (включая субъекты)) в модель HRU позволило доказать критерий безопасности систем для более приемлемых ограничений, что существенно расширило область ее применения [13]. Так, используя специальные типы, например «файл», «программа» для объектов; «user», «administrator», «auditor» для субъектов, в ИС на основе TAM можно организовать эффективный контроль порождения субъектов с нейтрализацией проблемы троянских программ [6].

Еще одной моделью, имеющей важное теоретическое значение в исследовании процессов распространения прав доступа в системах, основанных на политике дискреционного доступа, является модель TAKE-GRANT. Данная модель ориентирована на анализ путей распространения прав доступа в системах дискреционного управления доступом. Исходя из основных положений субъектно-объектной парадигмы построения ИС, модель TAKE-

GRANT использует аппарат теории графов для моделирования системы разграничения доступа и процессов ее изменения.

Состояние системы описывается соответствующим ему графом доступов ($G = (S, O, E)$), в котором множество вершин – это множество объектов O субъектов S доступа, причем $S \subseteq O$, а множество ребер – это множество E установленных прав доступа (x, y, α) субъекта x к объекту y с правом α из конечного набора прав $\alpha \subseteq R(r_1, r_2, ..., r_K) \bigcup \{t, g\}$, в том числе с двумя особыми правами – правом *take* (t – право брать права доступа у какого-либо объекта по отношению к другому объекту) и правом *grant* (g – право предоставлять права доступа к определенному объекту другому субъекту). При этом, в отличие от модели HRU, в модели TAKE-GRANT возможно наличие прав доступа не только у субъектов к объектам, но и у объектов к объектам. Основная цель модели TAKE-GRANT – определение и обоснование алгоритмически проверяемых условий проверки возможности утечки права доступа по исходному графу доступов, соответствующего некоторому состоянию системы [13]. Используя правила преобразования «*take*», «*grant*», «*create*» и «*delete*», можно воспроизвести состояния, в которых будет находиться система в зависимости от распределения и изменения прав доступа. Следовательно, можно проанализировать возможные угрозы для данной системы [14].

Модель TAKE-GRANT, как и ее расширенная модель [15], в которой $\alpha \subseteq R(r_1, r_2, ..., r_K) \bigcup \{t, g\} \{r_d, w\}$ (где r_d (*read*) – право на чтение или информационный поток на чтение, w (*write*) – право на запись или информационный поток на запись), играет важную методологическую роль, предоставляя теоретико-графовый инструмент анализа систем разграничения доступа с точки зрения санкционированного и несанкционированного со стороны определенных субъектов распространения прав доступа в рамках дискреционной политики [6].

Основные решения этих моделей нашли применение при предоставлении пользователям прав на выполнение тех или иных операций с теми или иными объектами СУБД (в большей мере реляционных). А именно, дискреционная модель разграничения доступа в СУБД основывается на следующих основных положениях.

1. Все субъекты S и объекты O БД должны быть идентифицированы, то есть каждой сущности (активной, пассивной) должен быть присвоен уникальный идентификатор.

2. Для каждого объекта o_j БД должен быть определен пользователь-владелец $s^{owner_j} = s_i \in S$.

3. Для субъектов S должны быть определены права доступа (иначе называемые привилегиями, или полномочиями) к разным объектам O. Например, в стандарте SQL определены такие типы привилегий доступа, как [16, 17]: SELECT, INSERT, DELETE, UPDATE, REFERENCES, USAGE, TRIGGER, EXECUTE, UNDER. Информация о привилегиях (аналог матрицы доступов M) сохраняется в системном каталоге СУБД в виде полномочий, выраженных с помощью некоторого языка описания. Эта информация используется системой для принятия решения о выполнении запрошенных субъектом операций над данными. При этом для принятия решения о том, выполнять запрошенные субъектом операции над данными или нет, система должна уметь аутентифицировать запрашивающий субъект.

В общем случае набор привилегий зависит от реализации СУБД (определяется производителем). Достаточно часто выделяют такие категории привилегий, как: *системные* (это право субъекта $s_i \in S$ выполнять некоторые административные действия с объектами O базы данных и с самой БД, например, такие как: создание базы данных, создание / удаление / изменение таблицы, создание/удаление представления, процедуры, триггера и т. д.) и *объектные* (права (разрешения) субъекта $s_i \in S$ на выполнение определенного действия с кон-

кретным объектом o_j схемы БД, таким как конкретная таблица, представление, последовательность, процедура, функция и т. д.).

Следует отметить еще одну особенность, присущую традиционным СУБД, связанную с особой ролью в обеспечении безопасности данных, наряду с привилегиями, представлений. Создавая представление и давая субъекту разрешение на доступ к нему, а не к исходной таблице, можно тем самым ограничить доступ субъекта, позволив ему обращаться только к определенным столбцам и строкам, на что еще в свое время обратил внимание автор монографии [5]. Таким образом, представления позволяют осуществлять контроль над тем, какие данные доступны тому или иному субъекту [18].

4. Владелец объекта s^{owner_j} должен обладать правом определения прав доступа к объекту другим субъектам $s_i \in S$.

Субъект получает/теряет определенные права (привилегии) двумя способами: первый связан с созданием объекта o_j и его владением, второй – путем передачи / отзыва определенных прав одним субъектом – другому [16]. Для этих целей, например, в стандарте SQL предусмотрены команды (операторы) GRANT/REVOKE, которые позволяет одному пользователю назначать (предоставлять) определенные права другому или отзывать их у другого. При этом «даритель» (лицо, предоставляющее право – англ. grantor) привилегий не лишается. С каждой привилегией ассоциировано право передачи (использование конструкции WITH GRANT OPTION), при котором имеется возможность передавать не только права на указанные действия, но и право передавать эти права другим субъектам. Назначенные привилегии могут быть в любой момент отозваны. Отзыв одной привилегии нередко связан с выполнением ряда каскадных действий по отмене иных привилегий: если отзывается привилегия, назначенная другим субъектам с правом передачи, она должна быть отозвана и у этих субъектов. Конструкции CASCADE, RESTRICT оператора REVOKE определяют, каким образом должна производиться отмена привилегий. Так, например, конструкция CASCADE отменяет привилегии не только для субъекта, который непосредственно упоминался в операторе GRANT при предоставлении ему соответствующих привилегий, но и для всех субъектов, которым этот субъект, воспользовавшись правом передачи WITH GRANT OPTION, предоставил привилегии.

Поскольку в процессе назначения привилегий и передачи прав на них собственно привилегии зачастую перекрываются и зависимости между теми, кто их предоставлял и теми, кто их получал, становятся все более сложными, структуры привилегий полезно представлять в виде графов (подобно теоретико-графовому инструменту модели TAKE-GRANT), называемых диаграммами назначения (англ. *grant diagrams*) [16]. СУБД поддерживают внутреннее представление подобных диаграмм для отслеживания событий назначения/отзыва привилегий.

5. В системе существует привилегированный пользователь $s^{pv} \in S$, обладающий правом полного доступа к любому объекту. При этом следует отметить, что реализация этого положения не должна позволять такому пользователю использовать свои полномочия незаметно для реального владельца объекта (что зачастую сегодня не всегда выполняется).

Подводя итог, можно сделать выводы о достоинствах и недостатках дискреционной политики управления доступом и моделях безопасности, построенных на ее основе.

К достоинствам дискреционной политики доступа можно отнести относительную простоту и гибкость реализации системы управления доступом. Это подтверждается действующими ИС, безопасность которых обеспечивается выполнением требований именно данной политики. Например, классическая модель HRU до сих пор широко используется при проведении формальной верификации корректности построения систем разграничения доступа в высоко защищенных автоматизированных системах [11, 19].

К недостаткам дискреционных моделей относятся:

– статичность установленных правил управления доступом. Данные модели, как правило, не учитывают динамику изменений состояний ИС;

– невозможность контролировать в полной мере потоки информации между объектами. Это обостряет проблему, связанную с уязвимостью по отношению к атаке с помощью вредоносных программ вида троянский конь;

– сложность отслеживания предоставляемых субъектам привилегий при их большом количестве;

– отсутствие механизма управления доступом к конфиденциальной информации. Как следствие, возможна утечка конфиденциальных данных.

Кроме этого, при использовании дискреционных моделей возникает вопрос о задании правил управления (распространения прав) доступом и анализа их влияния на безопасность ИС. В общем случае при использовании таких моделей достаточно сложно (алгоритмически трудноразрешимая задача) проверить, приведут ли действия субъекта, руководствующегося некоторым набором правил, к нарушению безопасности или нет.

Все это предопределило дальнейший поиск и разработку других моделей управления доступом.

Модели безопасности на основе мандатной политики

Исследователи и критики дискреционной политики, понимая, что основная проблема подобных моделей заключается в отсутствии контроля за информационными потоками, стали анализировать каким образом ее можно разрешить. Их внимание привлекло решение подобных задач в секретном делопроизводстве, в котором критерием безопасности является невозможность получения информации из документов определенного уровня безопасности (уровня / грифа секретности) субъектом доступа, чей уровень безопасности (допуска), ниже, чем уровень безопасности соответствующих документов.

Разграничение доступа и порядок работы с конфиденциальными документами организуются на основе парадигмы градации доверия определенным группам сотрудников в отношении секретов определенной степени важности. С этой целью вводится система уровней безопасности (уровней конфиденциальности/секретности). Сотрудники с самым высоким уровнем безопасности (уровнем доверия, допуска), могут работать с документами самой высокой степени (грифа) секретности. На более низком уровне безопасности (доверия, допуска), вводятся ограничения в отношении работы с документами более высокого уровня безопасности (секретности) и т. д. Соответственно, все сотрудники получают допуск к работе с секретными документами определенного уровня, а документы снабжаются специальной меткой (грифом секретности), отражающей требования к уровню безопасности при работе с ними.

Таким образом, если в дискреционных моделях управление доступом происходит путем предоставления субъектам полномочий для осуществления определенных операций над конкретными объектами, то мандатные модели управляют доступом неявным образом – с помощью назначения всем сущностям системы (субъектам, объектам) уровней безопасности, которые определяют все допустимые взаимодействия между ними. Следовательно, мандатное управление доступом не различает сущностей, которым присвоен одинаковый уровень безопасности, и на их взаимодействия ограничения отсутствуют [4]. То есть мандатный подход к разграничению доступа, основываясь только лишь на парадигме ранжированного доверия, без учета специфики других характеристик субъектов и объектов, приводит в большинстве случаев к избыточности прав доступа для конкретных субъектов в пределах соответствующих классов безопасности, что противоречит самому понятию разграничения доступа. Поэтому в тех ситуациях, когда управление доступом требует более гибкого подхода, мандатный принцип разграничения доступа дополняется дискреционным внутри соответствующих классов безопасности. В теоретических моделях для этого вводят матрицу доступа, раз-

граничивающую разрешенный по мандатному принципу доступ к объектам одного уровня безопасности [6].

Наиболее широкое распространение среди моделей мандатного управления доступа (многоуровневой защиты) получила модель Белла – ЛаПадулы [20]. Основными элементами данной модели являются:

– S – множество субъектов;

– O – множество объектов;

– $R = \{read, write, append, execute\} = \{r_d, w, a, e\}$ – множество видов доступа (a=*append* – доступ на запись в конец объекта);

– $B = \{b \subseteq S \times O \times R\}$ – множество возможных множеств текущих доступов в системе;

– Λ_L – решетка уровней безопасностей L (например, $L = \{U, SU, C, S, TS\}$, где U – *unclassified* (*несекретно*), SU – *sensitive but unclassified* (*чувствительно, но несекретно*), C – *confidential* (*конфиденциально*), S – *secret* (*секретно*), TS – top secret (*совершенно секретно*), $U < SU < C < S < TS$). Решеткой Λ_L называется алгебраическая система вида $(L, \leq, \sqcup, \otimes)$, где $\leq$ – оператор, определяющий частичное нестрогое отношение порядка для уровней безопасностей L; $\sqcup$ – оператор наименьшей верхней границы; $\otimes$ – оператор наибольшей нижней границы.

– $M = \{M_1, M_2, ..., M_c\}$ – множество возможных матриц доступов, где $c = n \cdot m \cdot 2^p$, $p = |R|$, $M_\eta[s, o] \subseteq R$ – права доступа субъекта s к объекту o, $\eta = 1..c$;

– $(f_s, f_o, f_c) \in F = L^S \times L^O \times L^S$ – тройка функций (f_s, f_o, f_c), задающих соответственно: $f_s : S \to L$ – уровень безопасности (доступа) субъектов; $f_o : O \to L$ – уровень безопасности объектов; $f_c : S \to L$ – текущий уровень безопасности (доступа) субъектов, при этом для любого $s \in S$ выполняется неравенство $f_c(s) \leq f_s(s)$;

– $V = B \times M \times F$ – множество состояний системы;

– Q – множество запросов системе;

– D – множество ответов по запросам, например $D = \{yes, no, error\}$;

– $W \subseteq Q \times D \times V \times V$ – множество действий системы, где четверка $(q, d, v^*, v) \in W$ означает, что система по запросу q с ответом d перешла из состояния v в состояние v^*;

– $T = \{1, 2, ..., t, ...\}$ – множество дискретных моментов времени;

– X – множество функций $x : T \to Q$, задающих все возможные последовательности запросов к системе;

– Y – множество функций $y : T \to D$, задающих все возможные последовательности ответов (решений – decision) системы по запросам;

– Z – множество функций $z : T \to V$, задающих все возможные последовательности состояний системы.

Основные свойства (правила, обеспечивающие разграничение доступа) безопасности модели Белла – ЛаПадулы:

1. *Свойство простой безопасности* (*simple security property – ss*). Субъект на уровне безопасности $l \in L$ может проводить операцию чтения только в отношении объектов своего или более низкого уровня. Это свойство также известно, как правило *no read up* (NRU) – нет чтения вверх.

2. *Свойство* * (**-property*). Субъект с заданным уровнем безопасности $l \in L$ может осуществлять запись только в объекты своего или более высокого уровня. Это свойство также известно, как правило *no write down* (NWD) – нет записи вниз.

3. *Свойство дискреционной безопасности* (*discretionary security – ds*). Права дискреционного доступа субъекта к объекту определяются на основе матрицы доступа M. Термин «дискреционная» безопасность уместен в контексте конкретных решений этой модели, поскольку в модель включена возможность изменять M (структуру разрешений) [20].

Характеризуя понятия безопасного состояния, Белл и ЛаПадула предложили следующее определения.

Состояние $(b, m, f) \in V$ называется *безопасным по чтению* (или *просто безопасным*) тогда и только тогда, когда для каждого субъекта, осуществляющего в этом состоянии доступ чтения к объекту, уровень безопасности этого субъекта доминирует над уровнем безопасности этого объекта. *Состояние* ($(b, m, f) \in V$) называется *безопасным по записи* (или **-безопасным*) тогда и только тогда, когда для каждого субъекта, осуществляющего в этом состоянии доступ записи к объекту, уровень безопасности этого объекта доминирует над уровнем безопасности этого субъекта. *Состояние системы* $(b, m, f) \in V$ *безопасно* тогда и только тогда, когда оно безопасно и по чтению, и по записи. А *система* $\Sigma(Q, D, W, z_0) \subset X \times Y \times Z$ (где z_0 – начальное состояние системы) называется *безопасной* тогда и только тогда, когда она обладает *s*-свойством, *-свойством, *ds*-свойством одновременно. Другими словами система $\Sigma(Q, D, W, z_0)$ безопасна тогда и только тогда, когда ее начальное состояние z_0 безопасно и все состояния, достижимые из z_0 путем применения конечной последовательности запросов из Q, безопасны.

Графическое представление модели Белла – ЛаПадулы показано на рис. 2.

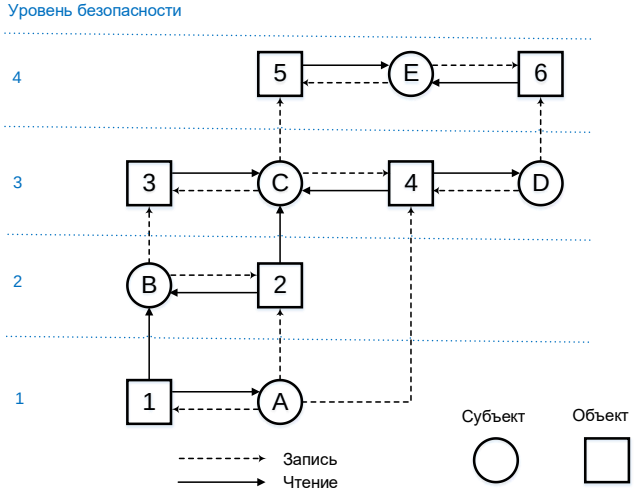


Рис. 2. Многоуровневая модель безопасности Белла – ЛаПадулы

На рис. 2 сплошная стрелка от объекта к субъекту показывает, что субъект осуществляет чтение объекта (информационный поток идет от объекта к субъекту). Пунктирная стрелка от субъекта к объекту показывает, что субъект осуществляет запись в объект (информационный поток идет от субъекта к объекту). Таким образом, направления информационных потоков указываются стрелками. При этом, как видно из рис. 2, например, субъект B может читать данные из объекта 1, но не может считывать данные из объекта 3.

Для СУБД особенно большое внимание методам мандатного управления доступом стало уделяться в начале 1990-х годов. Это было связано, как отмечается в работе [21], с тем фактом, что согласно требованиям Министерства обороны США любая используемая в этом ведомстве СУБД должна была поддерживать принципы мандатного управления доступом. Поэтому разработчикам СУБД пришлось вступить в соперничество за скорейшую разработку методов такого управления.

Мандатная модель управления доступа в СУБД основывается на следующих основных положениях:

1. Все субъекты S и объекты O БД должны быть идентифицированы.

2. Должна быть определена решетка уровней безопасностей L.

3. Каждому объекту БД $o \in O$ должен быть присвоен уровень безопасности (метка конфиденциальности), задающий установленные ограничения на доступ к данному объекту.

4. Каждому субъекту БД $s \in S$ должен быть присвоен уровень безопасности – уровень доступа, задающий уровень полномочий данного субъекта.

5. Субъект $s \in S$ может получить доступ к объекту БД $o \in O$ только в случае, когда уровень доступа субъекта позволяет предоставить ему данный доступ к объекту с заданным уровнем конфиденциальности, и реализация доступа не приведет к возникновению информационных потоков от объектов с высоким уровнем конфиденциальности к объектам с низким уровнем конфиденциальности.

Отличие мандатного управления доступа от дискреционного заключается в том, что в мандатной модели контролируются не операции, выполняемые субъектом над объектом, а потоки информации, которые могут быть только двух видов: либо от субъекта к объекту (запись), либо от объекта к субъекту (чтение).

В реляционной модели в качестве структуры, обладающей меткой, естественно выбрать кортеж (строку). Такое решение позволяет обеспечить достаточную избирательность доступа. Местом хранения самих меток может быть выбран соответствующий атрибут кортежа, определенный на домене уровней безопасностей L. При этом должен быть решен вопрос с механизмом формирования и изменения этого атрибута.

Один из подходов к построению системы управления доступом в реляционных СУБД на основе мандатной политики описан в группе патентов [22 – 24]. Предлагаемый подход обеспечивает контроль доступа на уровне строк (англ. *Row Level Security* – RLS) в таблице реляционной БД. В публикациях такой механизм упоминается под различными названиями: детальный контроль доступа (англ. *Fine Grained Access Control* – FGAC), виртуальная приватная база данных (англ. *Virtual Private Database* – VPD) и некоторых других [25, 26]. Суть детального контроля доступа (FGAC) состоит в следующем. Таблица базы данных содержит столбец метки безопасности, в строки которого записываются конкретные значения меток, определенные в иерархии уровней безопасностей L. Когда субъект $s \in S$ запрашивает доступ к строке (являющейся в рассматриваемом случае объектом доступа $o \in O$) таблицы, механизм безопасности сравнивает уровень безопасности субъекта с уровнем безопасности (меткой конфиденциальности), указанном в соответствующем атрибуте строки. Если уровень безопасности субъекта s преобладает над уровнем безопасности, указанном в соответствующем атрибуте строки, субъекту предоставляется доступ к строке.

Метка определяет уровень безопасности для субъекта в многоуровневой схеме безопасности и определенные привилегии для доступа к данным БД. Кроме того, метка безопасности может определять категории безопасности в рамках того уровня безопасности, к которому субъекту разрешен доступ. Примером категории безопасности может быть некоторый проект, к которому субъект допущен. Например, некоторому субъекту может быть разрешено просматривать данные, обозначенные определенными уровнями безопасности, такими как уровни безопасности: U, SU, C, S, TS. Этому же субъекту также может быть разрешен доступ к данным, относящимся к определенным категориям, таким как, например проекты $A1$, $A2$ и $A3$. Значение, хранящееся в метке, формируется некоторым способом, который позволяет понятно выразить информацию об уровне и категории безопасности для системы безопасности. Примером метки после такого преобразования (кодирования) может быть следующая: $SA1$, где $'S'$ указывает уровень безопасности – секретно, а $'A1'$ задает категорию безопасности, которая является идентификатором проекта, к которому субъект допущен.

Доступ к соответствующей строке разрешен только в том случае, если безопасность субъекта преобладает над безопасностью строки, в которой выполняются оба следующих условия:

– уровень безопасности, указанный меткой безопасности субъекта, больше или равен уровню безопасности, указанному меткой безопасности строки;

– категории безопасности, связанные с меткой безопасности строки, являются надлежащим подмножеством категорий безопасности, связанных с меткой безопасности субъекта.

Реализации технологии мандатного доступа в различных СУБД могут отличаться. Например, в СУБД Oracle она накладывается на реализацию дискреционной модели. А именно, возможность доступа к данным проходит проверку, содержащую два этапа: сначала проверяются стандартные ограничения дискреционного доступа. Затем для субъектов, прошедших проверку первого этапа, проверяется возможность доступа к объектам, базирующаяся на ограничениях мандатной модели. Механизм VPD в Oracle позволяет регламентировать доступ к частям таблицы.

Начиная с версии 8.1.7 в Oracle появилось другое средство – Oracle Label Security (OLS). В результате, реализация технологии мандатного доступа, основывающаяся на механизме OLS, стала опираться не только на дискреционную модель доступа (вначале проверяются права субъекта на выполнение соответствующей операции над таблицей), но и на механизм VPD (если у субъекта есть соответствующие привилегии, проверяется, не прикреплены ли к таблице какие-либо политики VPD). И только после всего этого проверяется наличие политик Oracle Label Security, назначенных защищаемой таблице: сравниваются метки, присвоенные отдельным строкам, с авторизацией меток пользователей, разрешая или запрещая доступ. Важным этапом развития по отношению к механизму виртуальных частных баз данных в OLS стала возможность формировать составную метку доступа.

В интерпретации Oracle (механизм OLS) метка безопасности – это тройка:

$$\lambda = \langle \lambda_1, \lambda_2, \lambda_3 \rangle, \tag{1.2}$$

где λ_1 – элемент линейно упорядоченного множества Θ_1 (множество уровней чувствительности / секретности). Задание в метке доступа уровня секретности является обязательным. Теоретически допускается до 10 000 уровней. Пример различных способов задания уровней секретности приведен в табл. 1;

Таблица 1

Пример различных способов задания уровня секретности

Числовая форма	***Длинная форма***	***Краткая форма***
40	HIGHLY_SENSITIVE	HS
30	SENSITIVE	S
20	CONFIDENTIAL	C
10	PUBLIC	P

λ_2 – подмножество элементов из множества Θ_2 (множество отделений (англ. compartments)). Отделения определяют области, которые описывают чувствительность помеченных данных, обеспечивая более тонкий уровень детализации в пределах уровня. Пример различных способов задания отделений приведен в табл. 2.

Таблица 2

Пример различных способов задания отделения

Числовая форма	***Длинная форма***	***Краткая форма***
85	FINANCIAL	FINCL
65	CHEMICAL	CHEM
45	OPERATIONAL	OP

Отделения не являются обязательными. Метка может содержать ноль или более отделений, то есть не все метки могут иметь отделения. OLS позволяет определять до 10 000 отделений [27];

λ_3 – элемент иерархически упорядоченных элементов (дерева) множества Θ_3 (множество групп (англ. groups)). Все данные, относящиеся к определенному отделению, могут

иметь группу этого отделения в метке. Группы полезны для контролируемого распространения данных и своевременного реагирования на организационные изменения. Группы иерархичны, то есть группа может быть связана с родительской группой. Пример различных способов задания групп приведен в табл. 3.

Таблица 3

Пример различных способов задания группы

Числовая форма	*Длинная форма*	*Краткая форма*	*Родительская группа*
1000	WESTERN_REGION	WR	
1100	WR_SALES	WR_SAL	WR
1200	WR_HUMAN_RESOURCES	WR_HR	WR
1300	WR_FINANCE	WR_FIN	WR
1310	WR_ACCOUNTS_PAYABLE	WR_AP	WR_FIN
1320	WR_ACCOUNTS_RECEIVABLE	WR_AR	WR_FIN

Группы необязательны. Метка может содержать ноль или более групп. Oracle Label Security позволяет определять до 10 000 групп.

В действительности метка безопасности реализуется добавлением специального столбца, содержащего значение, интерпретируемое как тройка $\lambda = \langle \lambda_1, \lambda_2, \lambda_3 \rangle$. Символьные строковые представления меток используют следующий синтаксис:

```
LEVEL:COMPARTMENT1,...,COMPARTMENTn:GROUP1,...,GROUPn.
```

Пример допустимых меток:

```
SENSITIVE:FINANCIAL,CHEMICAL:EASTERN_REGION,WESTERN_REGION
CONFIDENTIAL:FINANCIAL:VP_GRP
SENSITIVE
HIGHLY_SENSITIVE:FINANCIAL
SENSITIVE::WESTERN_REGION
```

С текстовой строкой, представляющей метку, связывается так называемый числовой тег метки (англ. *label tag*). Именно этот тег метки, а не текстовая строка, сохраняется в столбце метки защищаемой таблицы. В отличие от простых меток, соответствующих уровням конфиденциальности информации, составные метки имеют более сложные правила упорядочения и обеспечивают большую функциональность систем защиты информации.

Подводя итог, можно сделать следующие выводы.

Модель Белла – ЛаПадулы сыграла огромную роль в развитии теории компьютерной безопасности, и ее положения были введены в качестве обязательных требований к системам, обрабатывающим информацию, содержащую государственную тайну, в стандартах защищенных ИС. Однако при практической реализации модели Белла – ЛаПадула возникает ряд проблем, например, таких как [28]:

1) завышение уровня безопасности (для некоторой информации может быть определен уровень безопасности выше необходимого);

2) запись вслепую (например, в ситуации, когда субъект производит запись объекта с более высоким уровнем безопасности, операция не нарушает правила NWD, однако после ее завершения субъект не может проверить правильность выполнения записи объекта путем выполнения контрольного чтения, так как это нарушает правило NRU);

3) привилегированные субъекты. Эта проблема связана с работой администратора (системного, базы данных), которая подразумевает выполнение в системе таких критических операций, как добавление и удаление пользователей, восстановление системы после сбоев, аварий, установка программного обеспечения, устранение ошибок. Однако такие операции не вписываются в рамки модели, что означает невозможность осуществления правильного администрирования без нарушения правил данной модели.

Расширения модели Белла – ЛаПадулы [29, 30], связанные с поиском условий и ограничений, повышающих ее безопасность, также не снимают всех недостатков мандатного доступа. В частности, мандатный доступ отчасти снимает проблему троянских программ – толь-

ко с точки зрения опасных потоков «сверху вниз». В пределах же одного класса безопасности

вопросы доступа решаются, как и в дискреционных моделях, – на основе матрицы доступа. Следовательно, для полного устранения проблемы троянских программ в системах мандатного доступа также требуется более тщательный и детализированный контроль информационных потоков.

В целом же основным недостатком многоуровневых моделей является невозможность управления доступом к конкретным объектам на основе учета индивидуальных особенностей каждого из субъектов.

Таким образом, оба рассмотренные выше подхода не в полной мере могут эффективно и гибко управлять безопасным доступом к данным. Следовательно, оба подхода как бы предполагают поиск различных компромиссов между эффективностью, гибкостью и безопасностью. Очевидно, что оптимальное решение вопросов безопасности должно вырабатываться с применением обоих видов моделей.

Модели безопасности на основе ролевой политики

В основе рассмотренных моделей безопасности лежат отношения между отдельным субъектом и объектом доступа, определяемые либо внешним фактором (дискреционный доступ), либо уровнем безопасности (мандатный доступ).

Вместе с тем, анализ различных ИС, и в первую очередь информационных систем организационного управления, показывает, что в реальной жизни все данные системы принадлежат некоторой организации, а не конкретному пользователю (субъекту). Сотрудники этой организации выполняют определенные функциональные обязанности не от своего личного имени, а в рамках некоторой должности, которую можно трактовать как определенную роль, представляющую собой некоторую обобщенную сущность, выражающую определенный тип функций и статус сотрудника (подчиненность, права и полномочия).

Таким образом, политика разграничения доступа в таких ИС должна строиться на основе функционально-ролевых отношений, складывающихся в предметной области. Концепция управления доступом на основе ролей связана с многопользовательскими системами, впервые появившимися в 1970-х годах [31]. Несколько позже появилось формальное выражение – управление доступом на основе ролей (англ. *Role-Based Access Control* – RBAC), используемое сегодня. Основой ролевых моделей является введение в субъектно-объектную модель ИС дополнительной категории активных сущностей – ролей. Ролевая модель определяет особый тип политики, основанный на компромиссе между гибкостью управления доступом, характерной для дискреционных моделей, и жесткостью правил контроля доступа, присущей мандатным моделям. В ролевой модели классическое понятие субъект разделяется на две составляющие: пользователь и роль. *Пользователь* – это человек, работающий с системой и выполняющий определенные служебные обязанности. *Роль* – это активно действующая в системе абстрактная сущность, с которой связывается определенный набор полномочий (привилегий), необходимых для осуществления определенной деятельности.

Управление доступом при использовании ролевой политики осуществляется в два этапа:

1. Создание ролей и определение их полномочий (прав доступа к объектам).
2. Назначение ролей пользователям системы.

Следует отметить, что пользователь может быть ассоциирован с несколькими ролями. Данная возможность значительно упрощает администрирование сложных систем.

Управление доступом в ролевых системах требует разбиения процесса функционирования системы и работы пользователя на сеансы, в каждом из которых выделяются фазы [6]:

– авторизации пользователя в текущем сеансе с одной или несколькими разрешенными для него ролями;

– разрешения/запрещения доступа пользователю к объектам системы в рамках полномочий соответствующих ролей, с которыми этот пользователь авторизован в текущем сеансе.

Нетрудно видеть, что ролевые модели сочетают в себе как мандатный подход к организации доступа – через определенную агрегацию субъектов и объектов доступа (распределение ролей между сеансами и пользователями, а также полномочий между ролями) и тем самым обеспечивают жесткость правил разграничения доступа, так и дискреционный подход (связывание с ролью определенного набора полномочий на объекты системы), обеспечивающий гибкость в настройке системы разграничения доступа в ИС для конкретной предметной области.

Ролевая модель описывает систему в виде следующих множеств [31]:

– U – множество пользователей;

– R – множество ролей;

– P – совокупность полномочий на доступ к объектам (реализованная, например, в виде матрицы доступа);

– C – множество сеансов работы пользователей с системой.

Ролевые отношения устанавливаются следующими отображениями множеств сущностей системы:

– $PA \subseteq P \times R$ – отображение множества полномочий на множество ролей, задающее для каждой роли установленный набор полномочий (отношение типа «многие ко многим»);

– $UA \subseteq U \times R$ – отображение множества пользователей на множество ролей, определяющее набор ролей, доступных данному пользователю (отношение типа «многие ко многим»).

Управление доступом в системе осуществляется на основе введения следующих функций:

– $user : C \to U$ – функция, которая ставит в соответствие каждому сеансу $c_i \in C$ одного пользователя $u \in U$: $u = user(c_i)$ (не меняется в течение сеанса);

– $roles : S \to 2^R$ – функция, которая ставит в соответствие каждому сеансу c_i набор ролей из множества R, доступных в данном сеансе: $roles(c_i) = \{r \mid (user(c_i), r) \in UA\}$ (набор ролей может меняться со временем);

– $permissions : C \to P$ – функция, которая ставит в соответствие каждому сеансу c_i набор доступных в нем полномочий: $permissions(c_i) = \bigcup_{r \in roles(c_i)} \{p \mid (p, r) \in PA\}$, иначе говоря, совокупность полномочий всех ролей, доступных в данном сеансе.

Взаимосвязь пользователей, ролей, полномочий (привилегий) и сеансов показана на рис. 3.

Основное правило (критерий безопасности) ролевого доступа определяется следующим образом: система считается безопасной, если и только если любой пользователь $u \in U$ в системе, работающий в сеансе $c \in C$, может осуществлять действия, требующие полномочий $p \in P$, только в том случае, если $p \in permissions(c)$.

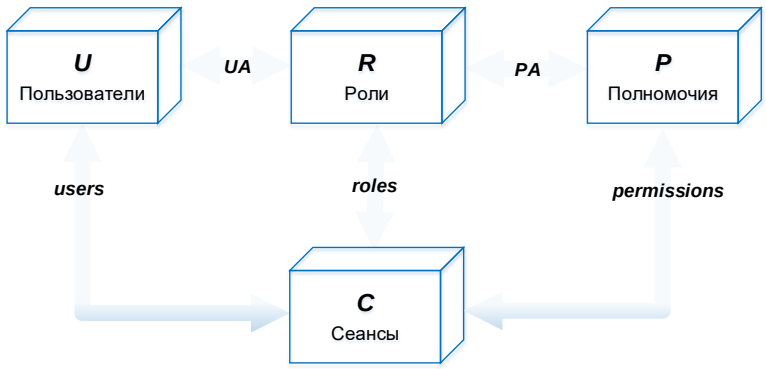


Рис. 3. Взаимосвязь ролей, полномочий, пользователей и сеансов

На практике управление доступом в ИС при использовании ролевой модели осуществляется главным образом не с помощью назначения новых полномочий ролям, а путем задания отношения *UA* (определения ролей, доступных данному пользователю) и функции *roles*, определяющей доступный в сеансе набор ролей. Поэтому многочисленные интерпретации ролевой модели [2, 4, 6, 31]:

– с иерархической организацией системы ролей ($PH \subseteq R \times R$);

– с взаимоисключающими на любые (все) сеансы ролями (модель статического распределения обязанностей);

– с взаимоисключающими на один сеанс ролями (модель динамического распределения обязанностей);

– с количественными ограничениями по ролям; с группированием ролей и полномочий, различаются видом функций *user*, *roles* и *permissions*, а также ограничениями, накладываемыми на отношения *PA* и *UA*.

Использование ролевой модели позволяет повысить эффективность администрирования сложных ИС. Поэтому данный подход является востребованным, в том числе и для СУБД.

Большое число пользователей, статус которых требует различных полномочий (привилегий) для доступа к ресурсам базы данных, создает значительный объем монотонной и утомительной работы администратору БД (АБД). В связи с этим, например, в стандарт SQL была добавлена концепция ролей как именованного набора привилегий [18]. В большинстве современных реализаций SQL роли могут быть предоставлены отдельным идентификаторам пользователей точно так же, как и отдельные привилегии. Кроме того, большинство реализаций СУБД поставляется с набором предопределенных ролей. Например, привилегии, которые обычно необходимы для работы АБД, часто предоставляются поставщиком СУБД в виде роли.

Приобретаемые преимущества от использования ролей в СУБД.

– Роли могут существовать до того, как будут определены учетные записи пользователей. Например, можно создать роль для отдела обработки заказов, вместо того, чтобы все его сотрудники совместно использовали некоторый общий идентификатор пользователя (*user-id*). Когда в отдел приходит новый сотрудник, одна инструкция GRANT с указанием роли предоставляет ему все необходимые для работы в отделе привилегии.

– Роли сохраняются, когда учетные записи пользователей удаляются (например, при увольнении). АБД, удаляя учетную запись некоторого пользователя (*user-id*), больше не беспокоится о потере прав. Роль может быть легко передана другому человеку, принятому на эту должность.

– Роли поддерживают стандартные привилегии. При применении ролей в организации легко обеспечить одинаковые привилегии для всех людей, выполняющих одинаковую работу.

– Роли устраняют рутинную работу по предоставлению привилегий отдельным пользователям. Роли позволяют предоставлять множество привилегий одной простой командой. При добавлении/удалении привилегий к/из роли все изменения немедленно отражаются на всех пользователях, которым предоставлена данная роль.

Создание и назначение привилегий при помощи ролей выполняется достаточно просто с помощью операторов CREATE, GRANT и REVOKE. Например, в стандарте SQL для создания роли используется следующий оператор:

```
CREATE ROLE name_role;.
```

Назначить роли привилегии можно путем выполнения, например, такого оператора:

```
GRANT name_ privilege1, name_ privilege2,… ON name_object TO
name_role;.
```

Предоставить роль пользователям можно так:

```
GRANT name_role TO user_1, user_2, …;.
```

Однако следует заметить, что поддержка ролей в разных реализациях SQL несколько отличается. К тому же не все СУБД поддерживают механизм ролей. Среди современных СУБД ролевая модель разграничения доступа поддерживается, например, в Microsoft SQL Server, Oracle, IBM DB2, PostgreSQL, MySQL 8.0 и некоторых других.

В заключение следует отметить, что ролевые модели позволяют реализовать гибкие, динамически изменяющиеся в процессе функционирования ИС правила разграничения доступа, эффективность которых особенно заметно проявляется при организации доступа к ресурсам сложных информационных систем с большим количеством пользователей и объектов, в том числе благодаря возможности построения иерархий ролей. Оперировать ролями гораздо удобнее, чем субъектами, поскольку это более соответствует распространенным технологиям обработки информации, предусматривающим разделение обязанностей и сфер ответственности между пользователями.

Вместе с тем, в ролевых моделях нет строгих доказательств безопасности системы в соответствии с определенными формализованными критериями. Такой подход позволяет получать простые и понятные правила контроля доступа, которые легко могут быть применены на практике, но лишает систему теоретической доказательной базы. В отличие от других рассмотренных ранее моделей ролевая модель практически не гарантирует безопасность с помощью формального доказательства, а только определяет характер ограничений, соблюдение которых и служит критерием безопасности системы. Поэтому безопасность ролевых моделей основывается на контрольных механизмах дискреционных или мандатных моделей, средствами которых регулируется доступ ролевых субъектов к объектам системы. К недостаткам ролевого разграничения доступа также следует отнести: возможность внесения избыточности (дублирования) при предоставлении пользователям прав доступа, сложность конструирования ролей.

Анализируя разнообразие формальных моделей и множество подходов к их реализации достаточно сложно определиться, какая из них более предпочтительна. И это естественно, так как каждая из них имеет свои преимущества, которые необходимо использовать именно в конкретной ситуации, точно так же, как свести при этом к минимуму недостатки каждой из них. При этом следует понимать, что сама модель безопасности не обеспечивает защиту, а только предоставляет принцип построения защищенной ИС, реализация которого и должна обеспечить заложенные в модели свойства безопасности. Безопасность системы в равной степени определяется тремя факторами [4]: свойствами самой модели (одной или нескольких), ее (их) адекватностью угрозам, воздействующим на систему, и тем насколько она (они) корректно реализована(ы).

В сложившейся ситуации, принимая во внимание научно-практические достижения в области информационной безопасности, современное состояние развития информационных систем, квалификацию злоумышленников, положения и рекомендации различных нормативно-правовых актов, целесообразным представляется проведение дальнейших исследований, результатом которых являлась бы некоторая методология, учитывающая, в том числе, возможности комплексного использования рассмотренных выше моделей, для обоснования, построения и оценки безопасности проектируемых и эксплуатируемых защищенных ИС.

Выводы

1. Проведенный анализ формальных моделей управления доступом позволил выявить их основные достоинства и слабые стороны, которые в зависимости от конкретной ситуации, целесообразно разумно использовать, в том числе и комплексно. При этом следует исходить из того, что сама модель безопасности не обеспечивает защиту, а только предоставляет принцип построения защищенной ИС, БД реализация которого должна обеспечить заложенные в модели свойства безопасности. Безопасность системы в равной степени определяется: свойствами самой модели (одной или нескольких), ее (их) адекватностью угрозам, воздействующим на систему, и тем насколько она (они) корректно реализована(ы).

2. К достоинствам моделей безопасности, построенных на основе дискреционной политики управления доступом, следует отнести ее относительную простоту и гибкость. Модели безопасности на основе дискреционной политики целесообразно применять при проведении формальной верификации корректности построения систем разграничения доступа в хорошо защищенных информационных системах и базах данных. Однако следует учитывать, что этим моделям свойственны недостатки, ограничивающие их применение, а именно: статичность установленных правил управления; невозможность контролировать в полной мере потоки данных между объектами; сложность отслеживания предоставляемых субъектам привилегий при их большом количестве и неконтролируемой передачи (распространении прав) от одного субъекта к другому; отсутствие механизма управления доступом к конфиденциальной информации.

3. Важным выгодным отличием мандатного управления доступа от дискреционного является то, что в мандатной модели контролируются не операции, выполняемые субъектом над объектом, а потоки информации, которые могут быть только двух видов: либо от субъекта к объекту (запись), либо от объекта к субъекту (чтение). Основной целью мандатного разграничения доступа к объектам является предотвращение утечки информации из объектов с высоким уровнем безопасности к субъекту с низким уровнем безопасности (противодействие созданию каналов передачи информации «сверху вниз»). Однако, несмотря на то, что модели безопасности на основе мандатной политики доступа играют значимую роль в теории информационной безопасности и их положения введены в качестве обязательных требований к системам, обрабатывающим информацию, содержащую государственную тайну, а также в стандартах защищенных ИС, при практической реализации этих моделей может возникнуть ряд проблем, таких как: завышение уровня безопасности; запись вслепую; проблема привилегированных субъектов, выполняющих операции, не вписывающиеся в рамки модели (это означает невозможность осуществления правильного администрирования без нарушения правил данной модели). К тому же модель мандатного управления доступом, решая проблему троянских программ, снимает ее только с точки зрения опасных потоков «сверху вниз», но не в пределах одного класса безопасности, когда ее приходится решать аналогично дискреционным моделям, на основе матрицы доступа. Следовательно, для полного устранения проблемы троянских программ в системах мандатного доступа требуется более тщательный и детализированный контроль информационных потоков. В целом же основным недостатком многоуровневых моделей является невозможность управления доступом к конкретным объектам на основе учета индивидуальных особенностей каждого из субъектов.

В отношении различных современных СУБД следует заметить, что сегодня они имеют отличающиеся реализации технологии мандатного доступа.

4. Модели безопасности на основе ролевой политики позволяют реализовать гибкие, динамически изменяющиеся в процессе функционирования информационных систем, баз данных, правила разграничения доступа, эффективность которых особенно заметно проявляется при организации доступа к ресурсам систем с большим количеством пользователей и объектов, в том числе благодаря возможности построения иерархий ролей. Вместе с тем, в ролевых моделях нет строгих доказательств безопасности системы в соответствии с определенными формализованными критериями. Такой подход позволяет получать простые и понятные правила контроля доступа, которые достаточно легко можно применить на практике, но лишает систему теоретической доказательной базы. К недостаткам ролевого разграничения доступа также следует отнести возможность внесения избыточности (дублирования) при предоставлении пользователям прав доступа, сложность конструирования ролей.

Список литературы:

1. Tanenbaum A. S., Herbert Bos H. Modern Operating Systems. Fourth edition. Pearson, 2015. 1136 p.
2. Смирнов С. Н. Безопасность систем баз данных. Москва : Гелиос АРВ, 2007. 352 с.
3. Cunha M. M., Oliveira E. F., Tavares A. J., Ferreira L. G. Handbook of Research on Social Dimensions of Semantic Technologies and Web Services. Hershey, PA: IGI Global, 2009. 1180 p.

4. Зегжда Д. П., Ивашко А. М. Основы безопасности информационных систем. Москва : Горячая линия – Телеком, 2000. 452с.
5. Хоффман, Л. Дж. Современные методы защиты информации. Москва : Сов. радио, 1980. 264 с.
6. Гайдамакин Н. А. Теоретические основы компьютерной безопасности. Екатеринбург : Изд-во Уральск. ун-та, 2008. 212 с.
7. Weissman C. Security controls in the ADEPT-50 time-sharing system // Proceedings of the November 18-20, 1969, fall joint computer conference. 1969. P. 119-133.
8. Hartson H. R., Hsiao D. K. A Semantic Model for Database Protection Languages. Systems for Large Data Bases. North-Holland, Amsterdam : Publishing Co., 1976. P. 27-42.
9. Harrison M. A., Ruzzo W. L., Ullman J. D. Protection in Operating Systems // Communications of the ACM, 1976. № 19(8). P. 461–471.
10. Lipton R. J., Snyder L. A linear time algorithm for deciding subject security // Journal of the ACM (JACM), 1977. № 24(3). P. 455-464.
11. Цирлов В. Л. Основы информационной безопасности автоматизированных систем. Ростов-на-Дону : Феникс, 2008. 173 с.
12. Sandhu R. S. The Typed Access Matrix Model // Proceedings of IEEE Symposium on Security and Privacy, Oakland, California, May 4-6, 1992, P. 122-136.
13. Девянин П. Н. Модели безопасности компьютерных систем. Управление доступом и информационными потоками. 2-е изд. Москва : Горячая линия – Телеком, 2013. 338 с.
14. Скакун В. В. Защита информации в базах данных и экспертных системах. Минск : БГУ, 2015. 140 с.
15. Frank J., Bishop M. Extending the take-grant protection system // Technical Report, Department of Computer Science, University of California at Davis, 1996. 14 p. URL: https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.51.907&rep=rep1&type=pdf. (accessed on: 04.02.2021).
16. Garcia-Molina H., Ullman J. D., Widom J. Database Systems. The Complete Book, 2th ed. Pearson Prentice Hall, 2009. 1203 p.
17. ISO/IEC 9075-2:2016 Information technology. Database languages. SQL. Part 2: Foundation (SQL/Foundation). URL: https://www.iso.org/standard/63556.html. (accessed on: 04.02.2021).
18. Грофф Д. Р., Вайнберг П. Н., Оппель Э. Д. SQL : полное руководство, 3-е изд. ; пер. с англ. Москва : Вильямс, 2015. 960 с.
19. Марков А. С., Цирлов В. Л., Барабанов А. В. Методы оценки несоответствия средств защиты информации. Москва : Радио и связь, 2012. 192 с.
20. Bell D. E., LaPadula L. J. Secure Computer Systems: Unified Exposition and Multics Interpretation (MTR-2997 Rev. 1). Bedford, Mass.: MITRE Corp., 1976. 129 p.
21. Date C. J. An Introduction to Database Systems, 8th ed. New York, USA : Pearson Education, Inc., 2004. 983 p.
22. Patent US 2004/0044655A1, United States, Row-level security in a relational database management system / Curt Cotner, Gilroy, CA (US); Roger Lee Miller, San Jose, CA (US); International Business Machines Corporation, Armonk, NY (US). N 10/233,397; Mar. 4, 2004.
23. Patent 8,131,664 B2, United States, Row-level security in a relational database management system / Curt Cotner, Gilroy, CA (US); Roger Lee Miller, San Jose, CA (US); International Business Machines Corporation, Armonk, NY (US). N 12/242,241; Mar. 6, 2012.
24. Patent 8.478,713 B2, United States, Row-level security in a relational database management system / Curt Cotner, Gilroy, CA (US); Roger Lee Miller, San Jose, CA (US); International Business Machines Corporation, Armonk, NY (US). N 15/343,568; Jan. 16, 2018.
25. Кайт Т. Oracle для профессионалов ; пер. с англ. СПб. : ООО «ДиаСофтЮП», 2003. 672 с.
26. Нанда А., Фейерштейн С. Oracle PL/SQL для администраторов баз данных ; пер. с англ. СПб : Символ-Плюс, 2008. 496 с.
27. Oracle Database 19c. Administrator's Guide. Understanding Data Labels and User Labels. URL: https://docs.oracle.com/en/database/oracle/oracle-database/19/olsag/understanding-data-labels-and-user-labels.html#GUID-2C0383D3-4AA5-4263-B938-827E2CCC40C0 (accessed on: 04.02.2021).
28. Щербаков А. Ю. Современная компьютерная безопасность. Теоретические основы. Практические аспекты. Москва : Книжный мир, 2009. 352 с.
29. McLean J. The specification and modeling of computer security // Computer. 1990. № 23(1). P. 9-16.
30. McLean J. Security models. Encyclopedia of software engineering, Vol. 2. Wiley, 1994. P. 1136-1145.
31. Sandhu R.S., Coyne E. J., Feinstein H. L., Youman C. E. Role-based access control models // IEEE Computer. 1996. № 2. P. 38-47.

Поступила в редколлегию 15.03.2021

Сведения об авторе:
Вилигура Владислав Викторович – Харьковский национальный университет имени В.Н. Каразина, аспирант кафедры безопасности информационных систем и технологий, факультета компьютерных наук; Украина, e-mail: viligura93@gmail.com; ORCID: https://orcid.org/0000-0002-1137-2382

УДК 004.056.55 **DOI:10.30837/rt.2021.2.205.06**

V.A. KULIBABA, PhD student

PROCESSES AND METHODS FOR SELECTING SYSTEM-WIDE PARAMETERS AND ANALYSIS OF RESISTANCE AGAINST THIRD-PARTY CHANNEL ATTACKS FOR THE KEY ENCAPSULATION MECHANISM DSTU 8961:2019

Introduction

In recent years, there has been significant progress in the creation of quantum computers. If scalable quantum computers are implemented, it will jeopardize the security of most widely used public key cryptosystems. The most vulnerable are key schemes, i.e. digital signatures, based on factorization, discrete logarithms and elliptic curve cryptography. The main task now is to develop, evaluate, research and standardize asymmetric crypto transformations at the international level, including key encapsulation mechanisms (KEM), resistant to attacks by violators of the post-quantum period. Another important task is to further study the already adopted national standards of ACS and PIC on resistance to attacks by third-party channels, in particular to assess the dependence of the conversion time using the private key on the structure of the key bits.

The main efforts of the international cryptographic community to develop standardize and implement new post-quantum crypto transformations are centered around the NIST US competition – NIST PQC Standardization Process, which began in December 2016 [1].

Of the 82 submitted candidates, 69 were admitted to the 1st round of the competition. In January 2019, based on open discussion and feedback from the cryptographic community, NIST selected 26 algorithms for the second round [2], including 17 asymmetric encryption and/or KEM.

The state of development and standardization of key encapsulation protocols at the international level and in Ukraine

Of the NIST evaluation criteria, the most important is the algorithm security criterion. For the KEM algorithms, NIST in the program statement of the competition put forward requirements for "semantic security" of algorithms in terms of resistance to attacks with adaptively selected ciphertext, which is equivalent to the security model IND-CCA2. Given that meeting the more stringent requirements of the IND-CCA2 model for some algorithms may affect performance [3], NIST has also adopted algorithms that provide protection against attacks with selected ciphertext in the IND-CPA.

In July 2020, the 2nd stage of the competition ended and the start of the 3rd round was announced [3]. Fifteen candidates advanced to Round 3, of which 7 were selected as finalists and 8 as alternative candidates. In particular, the following KEM algorithms were selected as finalists: Classic McEliece, SABER, CRYSTALS-KYBER, NTRU. Alternative candidates were: BIKE, FrodoKEM, HQC, NTRU Prime, SIKE. Of the 4 main finalists, NTRU, CRYSTALS-KYBER, SABER are based on algebraic lattices, which provided a basis for the assumption of including at least one of them in the standard.

In the report [3] the second most important requirement is speed, and for the candidates of the 3rd round will be considered as the speed of key generation, forward and reverse transformations, and the spatial complexity of public keys, digital signatures and ciphertexts computation. For KEM algorithms, the key generation time is considered to be on a par with the forward and reverse conversion times because a large number of applications use a new key pair for each session to provide perfect forward security. As a result, safety and performance requirements are currently the main ones considered for NIST's decision in the 3rd round of the PQC competition.

In [4], NIST recommended that developers focus on developing parameters for the stability levels 1–5 defined in [5], i.e. for the stability level of 128 bits of quantum and 256 bits of classical security. Despite this, at the national level it was substantiated [6] and set the task to develop an al-

gorithm for asymmetric transformation of the KEM type and parameter sets that would provide 7 levels of stability, i.e. 256 bits of quantum and 512 bits of classical security. In [7], the main algorithms for generating system-wide parameters, encryption and decryption for the advanced NTRU Prime IIT Ukraine KEM algorithm was substantiated.

Further issues of constructing system-wide parameters for the 7th level of stability, proving the correctness of the algorithm, as well as cryptographic stability are considered in detail in [7-9]. Subsequently, based on studies conducted in [6-10], the national standard DSTU 8961:2019 [11] was adopted.

The main parameters of DSTU 8961:2019 and the method of generation parameters for 7th stability level

The main parameters of the algorithm presented in Table 1 and Table 2. The whole set of parameters can be found in [8,10].

Table 1

General parameters

Definition	Description
$(Z/q)[x]$	Ring of polynomials. All the coefficients reduced by module *q*
$N \geq \max(3, 2t)$	Order of polynomial. Should be a prime number for which the polynomial $x^n - x - 1$ is irreversible. The order determines the number of its coefficients
$P = x^n - x - 1 \in (Z/q)[x]$	Monic polynomial of N degree, irreducible over the field $(Z/q)[x]$, by which polynomials are reduced – elements from R/q
$Z[x]/(x^n - x - 1)$	Ring of polynomials $Z[x]$ with a module $x^n - x - 1$
$(Z/3)[x]/(x^n - x - 1)$	Ring of polynomials $(Z/3)[x]$ with a module $x^n - x - 1$
$(Z/q)[x]/(x^n - x - 1)$	Ring of polynomials $(Z/q)[x]$ with a module $x^n - x - 1$
$p = 3$	Smaller module, all the coefficients reduced by this module in the *R/3* polynomial
$q \geq 48t + 3$	Larger module, all the coefficients reduced by this module in the *R/q* polynomial
$t \geq 1$	A natural number, the number of nonzero elements of a polynomial depends on this parameter.
$\lambda \in \{256, 384, 512\}$	The level of crypto strength of classical security
$m \in R/3$	Private message. The number of 0, 1, -1 greater or equal *t*.
$e \in R/q$	Encrypted message.
$G \in R/3$	Random *t*-small element (polynomial), reversible in *R/3* field. The number of 1 and -1 are not necessarily equal. The secret parameter used to calculate the public key
$f \in R/q; f = (1+3F) \bmod q$	A small polynomial, irreducible in *R/q*, is a secret key.
$F \in R/3$	A random polynomial that identifies a private key.
$dF = 2t/3$	Number of 1 and -1 in polynomial F
$df = 2t; df = dr$	Number of 1 and -1 in secret key, not necessarily equal.
$dg = \begin{cases} dg_1 = n/3+1 \\ dg_{-1} = n/3 \end{cases}$	Number of 1 and -1 in polynomial *G*

Table 2

Additional parameters

Definition	Description
$qBits = \lceil \log_2 q \rceil$	Number of bits in *q*
$r \in R/3$	Blinding polynomial, random t-*small*
$dr = n/3$	The number of 1 and -1 in Blinding polynomial
b	Random component (salt), to add to the message.
db=256	Length of the random component (bits).
$bLen = db/8$	Length of the random component (octets).
$\max MLen = \frac{\frac{3(N-1)}{2} - db}{8} - 1$	Maximum length of message for encrypt(bytes).
$Hlen = \begin{cases} 256, \lambda = 256 \\ 512, \lambda = 384 \| 512 \end{cases}$	Hash value length (bits)

The comprehensive algorithm description of calculation general parameters presented in [7, 8, 13]. A simplified sequence of steps can be represented as follows:

Step 1. Select prime number N.

As a prime number, prime N are chosen for which the order is N-1 or (N-1) / 2.

$$2^{\lambda} < (3/2^{N}) \tag{1}$$

Step 2. Formation of key space for private keys

To specify the key space, it is necessary to determine the number of nonzero N / 3 (1 and -1) elements in the polynomial *F = F1 * F2 + F3* and the keys G.

Then the maximum number of nonzero elements in the key defined[8] by the polynomial F = F1 * F2 + F3 taking into account the number (1) and (-1) is equal to

$$2d_1 * 2d_2 + 2d_3 = 4d_1d_2 + 2d_3 \tag{2}$$

In order to find the keys on the polynomials F1, F2, F3 was approximately the same complexity we choose

$$d_1 \approx d_2 \approx d_3$$

In [13] it is proposed to calculate the value according to formulas (3,4):

$$d_1 = \left\lceil \frac{-1 + \sqrt{1 + \frac{8N}{3}}}{4} \right\rceil, d_2 = \left\lceil \frac{\left\lceil \frac{N}{3} \right\rceil - d_1}{2d_1} \right\rceil, d_3 = \max\left(\left\lceil \frac{d_1 + 1}{2} \right\rceil, \left\lceil \frac{N}{3} - 2d_1d_2 \right\rceil \right) \tag{3}$$

$$d_g = \frac{N}{3} \tag{4}$$

Step 3. Calculation of security parameter taking into account key space and attack meeting in the middle (upper security boundaries)

To calculate the security parameter taking into account the key space and the attack of the meeting in the middle[8], the number of keys is determined taking into account their form of representation and the attack of the meet-in-the middle[14].

To determine the minimum prime number that provides the desired stability λ, the next inequality(5) is used [13]

$$2^{\lambda} \le \sqrt{\frac{\binom{N}{d_1 \quad d_1} * \binom{N}{d_2 \quad d_2} * \binom{N}{d_3 \quad d_3}}{N}} \tag{5}$$

If $\lambda <$ *required,* then choose greater prime N, and go to step 2.

Step 4. Calculate the maximum number of non-zero elements in the message (d_m)

During encryption of the data of encoded message, converted into a form of small polynomial, it should contain a number of non-zero elements, defined by general parameter d_m to prevent attacks. However, in case of the number of such non-zero elements is larger than some threshold, then the probability of reselecting a mask and multiplying by a blinding polynomial will be high[8]. Therefore, this parameter significantly affects the performance of the encryption algorithm.

The following condition(6) is a sufficient to eliminate the decryption error:

$$1 - \frac{\sum_{i=d_m}^{N-2d_m-1}\left(\sum_{j=d_m}^{N-d_m-i}\left(\binom{N}{i}\binom{N-i}{j}\right)\right)}{3^N} \le 2^{-10} \tag{6}$$

Step 5. Calculate q. It should be prime and satisfy conditions from [13]

The analysis showed that the value of the modulus q affects the probability of decryption error and is used in assessing the security of the lattice.

To calculate the value of q, which provides the maximum probability of error, which is determined by stability, the inequality(7) can be used:

$$q \ge 24(2d_1d_2 + d_3) + 3 \tag{7}$$

Step 6. Calculate T_{mintm}. Calculate minimum value $0 < r \le N$, which satisfy conditions(8):

$$T_{MITM}(N, r, d_g) > \lambda \text{ and } T_{MITM}(N, r, d_m) > \lambda \tag{8}$$

If at least one of the conditions fails, select greater prime N and go to step 2.

Step 7. Calculate the size of the lattice:

$$S = 2N - r \tag{9}$$

Step 8. Calculate $T_{lattice}$ - the number of operations for the construction of Korkin-Zolotarev-reduced basis[13] of a complete lattice of dimension S by formula(10):

$$T_{Lattice} = 2^{E(S,m,\beta)}, \tag{10}$$

where $E(S, m, \beta) = 0,366098\beta + 0,000784314\beta^2 + 0,875$

This value should be greater the value that corresponds to the securely level λ. If, $T_{Lattice} < \lambda$, select greater prime N, and go to step 2. Else, while $T_{Lattice} > T_{MITM}$, increment *r*, and go to step 6.

Calculated values of parameters N, t, q are present in Table 3. They can be applicable for key encapsulation and direct encryption. Highlighted rows are used in standard DSTU 8961:2019[10].

Table 3

Calculated values of parameters for different stability levels of standard DSTU 8961:2019

	N	t	q
SKELYA 256/128	881	159	7673
	883	168	8089
	907	160	7727
	907	183	8807
	953	132	6343
	953	171	8237
	967	171	8243
	971	101	4871
	971	198	9551
	977	120	5783
	977	162	7817
	991	194	9349
	997	112	5393
	1013	149	7177
	1019	139	6691
	1021	112	5393
	1021	183	8819
SKELYA 384/192	1201	192	9221
SKELYA 512/256	1259	210	10103
	1283	214	10289
	1289	215	10331
	1291	215	10331
	1297	216	10453
	1301	217	10427
	1303	217	10429
	1307	218	10499
	1319	220	10567
	1321	220	10597
	1327	221	10613
	1361	227	10957
	1373	229	11057
	1381	230	11059
	1399	233	11213
	1409	235	11299
	1423	237	11383
	1471	255	12251

Comparative analysis of sets of parameters $\lambda \in (256, 384, 512)$ of the standard DSTU 8961 on criteria of stability and complexity

Table 4

Performance metrics for different stability levels

Stability level λ	Encryption	Decryption
256	89224	102982
384	138237	146128
512	163658	188235

There is a direct relationship between the level of stability and the time of direct and inverse transformation. The Fig. 1 shows a graph of this dependence. It is obvious that with increasing stability level, the complexity of transformations increases, so it is important to choose a sufficient level of stability based on available computing power.

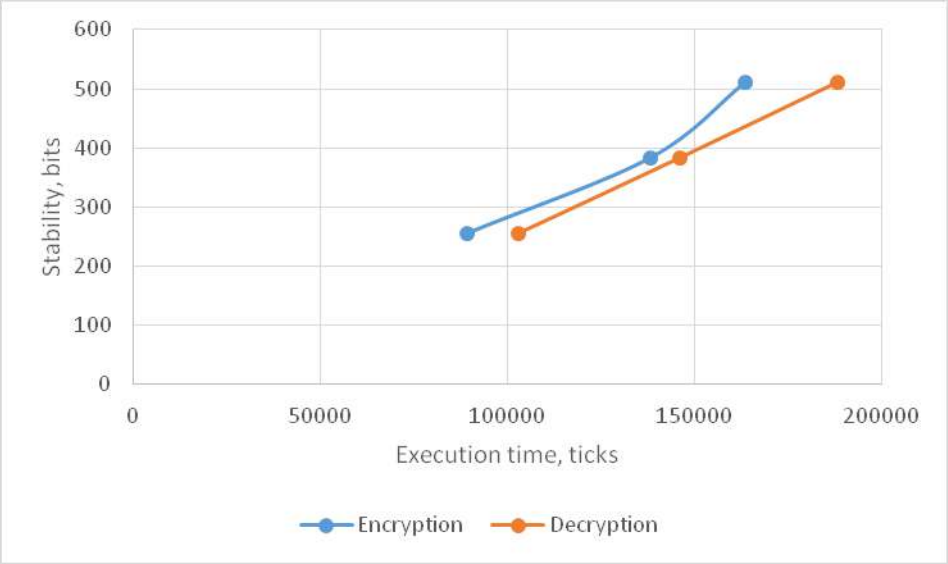


Fig. 1. Encryption and decryption time

Analysis of the stability of the standard DSTU 8961:2019 against side-channel attacks

An important issue is the stability of the national standard DSTU 8961:2019 against side-channel attacks, as well as against attacks on implementation. This paper considers the analysis of the dependence of the time of direct and inverse transformations (encapsulation/decapsulation of keys) on the structure of the long-term key, namely on the number of units in the long-term key.

The hash function defined in the national standard DSTU 7564:2014 was used as a hash function. The algorithm defined in DSTU 8845:2019 is used as an algorithm of symmetric streaming transformation.

For the experiment, 10,000 keys were generated and sorted by increasing number of ones. For each key, 100 calls were made to each of the tested functions and the average value of the execution time in CPU clocks for such a key was calculated. In Fig. 2-3 shows graphs of the number of CPU cycles from the key number. Random deviations can be caused by other processes in the operating system and do not depend on the number of units in the key. However, some optimizations of the implementation still possible and the number of cycles may depend on implementation.

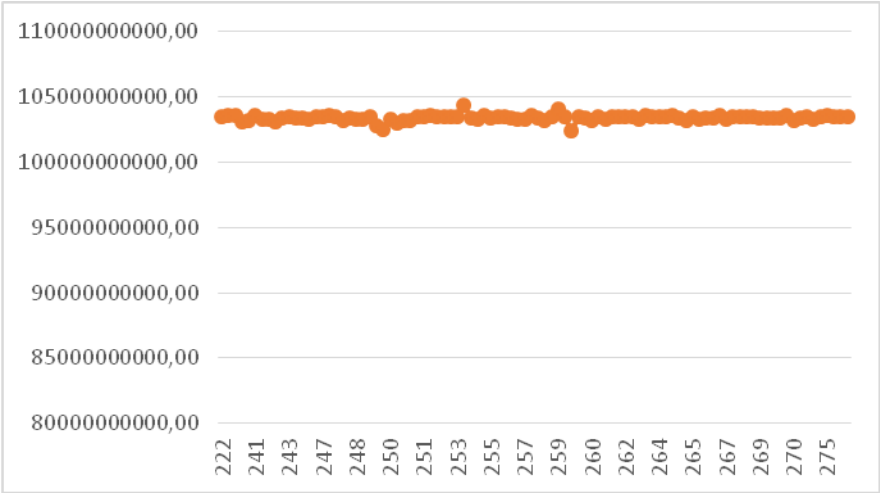


Fig. 2. Fragment of the graph of the dependence of the encapsulation execution time on the number of units in the key for the standard DSTU 8961:2019

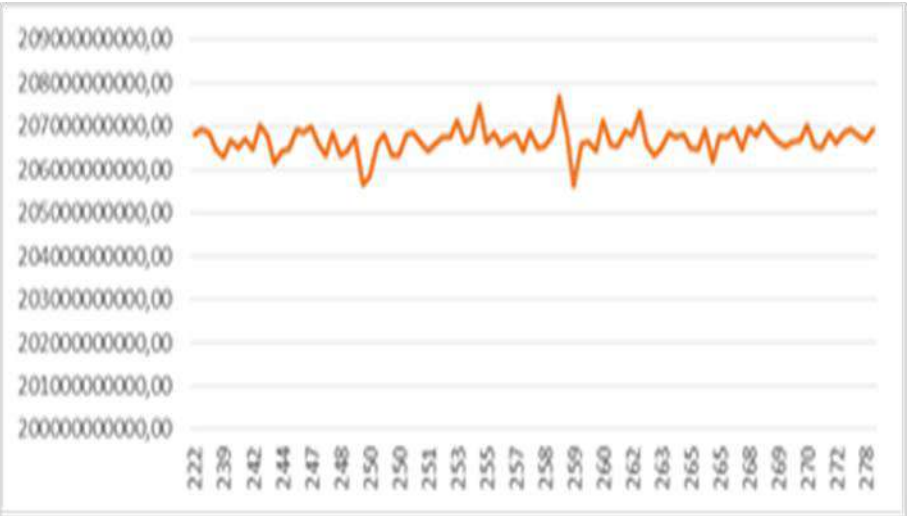


Fig. 3. Fragment of the graph of the dependence of the time of decapsulation on the number of units in the key for the standard DSTU 8961:2019

Table 5

The values of the correlation coefficients between the execution time for the functions of encapsulation / decapsulation of keys and the number of units in the key for the standard DSTU 8961:2019

	KeyNumber	CodeKem	DecodeKem
KeyNumber	1.000000	0.106736	0.153254
CodeKem	0.106736	1.000000	0.768122
DecodeKem	0.153254	0.768122	1.000000

The correlation coefficients between the execution time for the functions of encapsulation / decapsulation of keys and the number of units in the key for the standard DSTU 8961: 2019 are in the range [0.106 ... 0.153], which indicates the practical independence of execution time converted from the number of units in the key.

Conclusions

The main problems with asymmetric cryptography are the development of stable asymmetric crypto transformations such as KEM against both classical and quantum attacks, as well as the construction of system-wide parameters of 5-7 levels of stability [7,8].

Of particular relevance is the provision of cryptographic resistance to attacks by third-party channels, which requires a fundamentally new approach to the implementation of testing of cryptographic solutions of constant time, as well as analysis of existing standards on vulnerability to this class of attacks and countermeasures.

The study of the national standard DSTU 8961:2019 on resistance to attacks by third-party channels was performed. The KEM algorithms set out in DSTU 8961:2019 are protected from attacks by third-party channels in case of correct and accurate implementation, which has been confirmed experimentally.

The stability level has direct impact on the performance, so it is important to choose a sufficient level of stability based on device, e.g. for smart cards and tokens it might be reasonable to choose lower level of stability. The article presents the dependency between performance and stability level.

The simultaneous usage of keys encapsulation end encryption allows to setup symmetric encryption keys such as AES keys to increase speed of secure data transfer across communication channels.

References:

1. Post-Quantum Cryptography – Project Overview (2016) // Electronic resource. Access mode: https://csrc.nist.gov/projects/post-quantum-cryptography.
2. Gorjan Alagic Status Report on the First Round of the NIST Post-Quantum Cryptography Standardization Process. NISTIR 8240 / Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Yi-Kai Liu // Electronic resource. Access mode: https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8240.pdf.
3. Gorjan Alagic Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process. NISTIR 8309 / Gorjan Alagic, Jacob Alperin-Sheriff, Daniel Apon, David Cooper, Quynh Dang, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Yi-Kai Liu // Electronic resource. Access mode: https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8309.pdf.
4. National Institute of Standards and Technology Submission Requirements and Evaluation Criteria for the Post-Quantum Cryptography Standardization Process. // Electronic resource. Access mode: https://csrc.nist.gov/CSRC/media/Projects/Post-QuantumCryptography/documents/call-for-proposals-final-dec-2016.pdf.
5. Gorbenko I.D., Kachko O.G, Alekseychuk A.N., Kuznetsov O.O., Gorbenko Yu.I., Onoprienko V.V., Yesina M.V., Candiy S.O. Algorithms of asymmetric encryptation and encapsulation of keys of post-quantum period of 5-7 levels of stability and their applications // Radiotekhnika. 2019. Is. 198. P. 5-18. DOI:10.30837/rt.2019.3.198.01.
6. Gorbenko I.D., Kachko O.G., Esina M.V. General statements and analysis of the end-to-end encryption algorithm NTRU Prime IIT Ukraine // Radiotekhnika. Kharkov : KNURE, 2018. Is. 193. P. 5-16.
7. Gorbenko I. D., Alekseychuk A.N., Kachko O.G., Yesina M.V., Stelnik I.V., Kandy S.O., Bobukh V. A., Ponomar V.A. Calculation of general parameters for NTRU Prime Ukraine of 6-7 levels of stability // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 4. P.327-340. DOI: 10.1615/TelecomRadEng.v78.i4.40.
8. Gorbenko I.D., Kachko O.G., Gorbenko Yu.I., Stelnik I.V., Kandyi S.O., Yesina M.V. Methods of building general parameters and keys for NTRU Prime Ukraine of 5th–7th levels of stability. Product form // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 7 P. 579-594. DOI: 10.1615/TelecomRadEng.v78.i7.30.98.
9. Kachko O.G, Gorbenko Yu.I., Yesina M.V, Akolzina O. Asymmetric encryption algorithm optimization based on using NTRU Prime mathematics // Radiotekhnika. 2017. Issue 191. P. 5-10.
10. DSTU 8961:2019 Information Technology. Cryptographic information protection. Asymmetric encryption and key encapsulation algorithms.
11. DSTU 7564:2014 Information Technology. Cryptographic information protection. Hashing function.
12. DSTU 8845:2019 Information Technology. Cryptographic information protection. Symmetric flow transformation algorithm.
13. Choosing Parameters for NTRUEncrypt. J. Horstein, J.Pipher, J.Schanck, J.Silverman, W. Whyte, Z. Zhang, https://eprint.iacr.org/2015/708.pdf
14. Nick Howgrave Graham NTRU Cryptosystems Technical Report. Report #4, Version 2. A Meet-In-TheMiddle Attack on an NTRU Private key / Nick Howgrave Graham, Joseph H. Silverman, William Whyte [Electronic resource]. Access mode.

Надійшла до редколегії 09.03.2021

Відомості про автора:

Кулібаба Владислав Андрійович – Харківський національний університет імені В.Н. Каразіна, аспірант кафедри безпеки інформаційних систем і технологій, факультету комп'ютерних наук, Україна, E-mail: vlad.kulibaba1994@gmail.com

УДК 003.026:004.056 **DOI:10.30837/rt.2021.2.205.07**

Д. В. ГАРМАШ

ВЛАСТИВОСТІ БАГАТОВИМІРНОГО АЛГОРИТМУ RAINBOW ТА ЙОГО ЗДАТНІСТЬ ПРОТИСТОЯТИ РІЗНОМАНІТНИМ МЕТОДАМ КРИПТОАНАЛІЗУ І АТАЦІ СТОРОННІМИ КАНАЛАМИ

Вступ

Багатовимірні квадратичні схеми є перспективним рішенням для потреби квантових систем, стійких до атак від квантового комп'ютера. Однак, оскільки цей клас відносно молодий і багато схем цього класу були порушені в минулому, існує дуже мало їх реалізацій, особливо на вбудованих мікроконтролерах. Щоб оцінити, чи можуть ці схеми колись замінити чинні стандарти, необхідно знати, наскільки ефективно їх можна впровадити на різних платформах. У процесі цієї роботи дано теоретичне введення до багатовимірних квадратичних схем. Потім впроваджуються схеми, які певний час витримували атаки: Unbalanced Oil and Vinegar (UOV), Rainbow та enTTS. Особлива увага приділяється виявленню усіх загальних моментів схеми Rainbow.

1. Загальні положення щодо схеми ЕП RAINBOW

Наразі криптосистеми, що засновані на квадратичних поліномах, пройшли за останні 10 років суттєвий розвиток та визнання. Теоретичною основою конструкцій Oil-Vinegar є доведена теорема, згідно з якою вирішення (визначення) набору багатоваріантних поліноміальних рівнянь над кінцевим полем є експоненційно складною проблемою, хоча це є у загальному випадку як необхідною, так і достатньою умовами [2].

Цей напрямок досліджень пов'язаний з появою конструкції Мацумото та Імаї, в тому числі з використанням рівняння лінеаризації [1]. Далі Патарін та його співробітники доклали великих зусиль для розробки безпечних багатоваріантних криптосистем. Один з конкретних напрямків, яким займались Патарін та його співробітники, пов'язаний з рівняннями лінеаризації Dragon, Oil and Vinegar, Unbalanced Oil-Vinegar [1]. Побудова механізму ЕП Rainbow на основі Oil and Vinegar, Unbalanced Oil-Vinegar грунтується на тому, що певні квадратичні рівняння можна легко розв'язати, якщо є можливість вгадувати декілька варіантів [1].

Нехай k буде кінцевим полем. Ключовою конструкцією є відображення (карта) F від k^{o+v} до k^{o}:

$$F(x_1,...,x_0,x_1^{'},...,x_v^{'}) = F(x_1,...,x_0,x_1^{'},...,x_v^{'}),...,F_0(x_1,...,x_0,x_1^{'},...,x_v^{'}) \tag{1}$$

і кожна F_l у формі

$$F(x_1,...,x_0,x_1^{'},...,x_v^{'}) = \sum a_{l,k,j}x_k x_j + \sum b_{l,k,j}x_k^{'}x_j^{'} + \sum c_{l,k}x_k + \sum d_{l,j}x_j^{'} + c_l, \tag{2}$$

де x_i, $i = 1,...,o$ це Oil значення та $x_j^{'}$, $j = 1,...,v$ значення Vinegar у кінцевому полі k.

Потрібно звернути увагу на схожість наведеної вище формули з рівняннями лінеаризації. Такий тип поліномів називається "поліномом Oil-Vinegar". Причина, по якій вона називається схема "Oil-Vinegar", пов'язана з тим, що в квадратичному вимірі змінні Oil та Vinegar не змішуються повністю. Це дозволяє легко знайти одне рішення для будь-якого рівняння виду

$$F(x_1,...,x_0,x_1^{'},...,x_v^{'}) = (y_1,...,y_o), \tag{3}$$

коли $(y_1,...,y_o)$ дано. Щоб знайти одне рішення, потрібно лише випадковим чином вибрати значення для Vinegar змінних та підключити їх до рівнянь вище, що дасть набір o лінійних рівнянь з o змінними. Це має, з імовірністю, близькою до 1, дати рішення. Якщо цього не

сталося, можна спробувати ще раз, вибравши різні значення для Vinegar змінних, поки не вдасться знайти рішення [4].

Це сімейство криптосистем розроблено спеціально для схем підписів, де потрібно лише знайти одне рішення для даного набору рівнянь, а не унікальне рішення. Застосовуючи відображення (карту F), ми «приховуємо» її, складаючи її з лівої та правої сторін за двома оборотними афінними лінійними відображеннями L_1 та L_2. Оскільки L_1 знаходиться на k^o, а L_2 на k^{o+v}, це генерує квадратичне відображення (карту)

$$F^{-} = L_1 \circ F \circ L_2 \tag{4}$$

від k^{o+v} до k^o.

Збалансована схема Oil-Vinegar характеризується тим, що $o = v$, але її удосконалили Кіпніс та Шамір, використовуючи матриці, що відносяться до білінійних форм, визначених квадратичними поліномами [3].

Для незбалансованої схеми Oil-Vinegar, $v > o$, показано, що конкретна атака має складність приблизно $q^{v-o-1}o^4$, коли $v \approx o$. Це означає, що якщо o не надто велике (менше ніж 100) і дане фіксоване поле розміром q, тоді $v - o$ має бути досить великим, але також не надто великим, щоб забезпечити безпеку схеми.

Однак слід зауважити, що в цій схемі документ, що підписується, є вектором у k^o, а підпис – вектором у k^{o+v}. Це означає, що підпис має принаймні вдвічі більший розмір документа, і при великому $v + o$ система стає менш ефективною.

В рамках статті пропонується конструкція, яка використовує конструкцію Oil-Vinegar кілька разів, так що в підсумку підпис буде лише трохи довшим за документ. Отже, ця схема набагато ефективніша. Її називають схемою Rainbow.

2. Застосування різноманітних методів крипто аналізу проти алгоритму RAINBOW

Представляється короткий криптоаналіз схеми підпису Rainbow, розглянувши його для наведеного вище прикладу. Є кілька способів атак, з якими будуть мати справу користувачі алгоритму. Для тих методів, де використовуються квадратні форми, слід пам'ятати, що теорія квадратних форм над скінченними полями відрізняється, коли характеристика дорівнює 2, у порівнянні з випадком, коли характеристика є непарною [6].

2.1. Метод зниження рангу

Метод зниження рангу використовується для розбиття схеми підпису біраціональної перестановки Шаміра. Причина, по якій ця атака може спрацювати, полягає в тому, що простір, що охоплюється поліноміальними компонентами шифру схеми Шаміра, складається з прапора пробілів:

$$V1 \subset V2 \subset \cdots \subset Vt, \tag{5}$$

де V_t – простір, охоплений поліноміальними компонентами шифру, кожна V_i є власною підмножиною V_{t+1}, а ранг відповідної білінійної форми, що відповідає елементам у $V_{t+1} - V_i$, занадто більший, ніж у V_i, а різниця розмірів між V_i та V_{t+1} рівно 1. Завдяки цим властивостям, зокрема останньому, це дозволяє легко знайти цей прапор просторів, а саме всі V_i, спочатку знайшовши V_{n-1}, потім V_{n-2} і так далі шляхом зменшення рангу [8]. Але цей метод атаки вже не може працювати проти цієї схеми. Причиною цього є те, що, в нашому випадку, існує також такий прапор просторів, що кількість компонентів – це точно кількість рівнів, розмірність кожного компонента прапора точно відповідає розміру V_{i+1}, $i = 1,...,u-1$, але

різниця в розмірах останніх двох великих просторів – це точно $O_u - 1$, яка була обрана спеціально для досить великого числа 11, на відміну від випадку Шаміра, коли воно дорівнює 1.

Властивість, наведена вище, якраз і є причиною того, що атака більше не може працювати. Тут не можна використовувати метод зниження рангу через те, що $O_u - 1 = 11$ і більше не 1. „Останній товстий рівень Oil" дозволяє схемі протистояти атаці зниження рангу [7].

2.2. Метод атаки на Oil-Vinegar схеми

Аналіз показав, що дія L_1 полягає у змішуванні всіх поліноміальних компонентів F. Отже, кожен компонент шифру F тепер належить до верхнього рівня поліномів Oil-Vinegar, а саме всі вони є елементами P_4. Це багаточлени Oil-Vinegar з 22 змінними Vinegar та 11 змінними Oil [1]. Для цього випадку можна застосувати метод для незбалансованої схеми підпису Oil-Vinegar, щоб спробувати атакувати систему, що дозволить відокремити змінні верхнього шару Oil-Vinegar. Для цього нам потрібно розділити верхній (або кінцевий) рівень з 11 змінних Oil та 22 змінних Vinegar. Відповідно до криптоаналізу складність атаки цього першого кроку становить $q^{22-11-1} \times 11^4 > 2^{90}$.

2.3. Метод Мінранка

Існує два абсолютно різних способи використання методу Мінранка. Перший – пошук полінома, асоційована матриця якого має найнижчий ранг серед усіх можливих варіантів. Цей набір поліномів із 6 змінними Vinegar та 6 Oil належить до першого рівня, тобто $P1$, і позначався $F\tilde{}_1$. Для цього спочатку ми прив'язуємо до кожного полінома білінійну форму, яка має матрицю розміром 33×33. Потім ми можемо використовувати лінійні комбінації матриць, пов'язаних із компонентами F, для виведення полінома, пов'язана з яким матриця має ранг 12 [3]. В цьому випадку, щоб атакувати систему, проблемою стає пошук матриці рангу 12 серед групи з 27 матриць розміром 33×33. З методу Мінранка ми знаємо, що складність пошуку такої матриці становить $q^{12} \times 27^3$, що набагато більше, ніж 2100.

Інша можливість – це пошук поліномів, що відповідають поліномам у другому останньому рівні, а саме той, який належить P_3 і походить від лінійних комбінацій $F\tilde{}_i$, $i < 4$. У цьому випадку метод Міранка однозначно не може бути використаний , оскільки вони взагалі мають ранг 22. Одним із шляхів, безсумнівно, є випадковий пошук. Оскільки розмірність P_3 дорівнює 16, це стає проблемою пошуку елемента в підпросторі розмірності 16 в загальному просторі розмірності 27. Отже, такий випадковий пошук потребує щонайменше q^{11} пошуків, щоб знайти його, але нам також потрібно визначити, чи дійсно рейтинг нижче 22 для кожного пошуку. У цьому випадку загальна складність повинна бути не менше $q^{11} \times (22 \times 33^2 / 3) > 2^{100}$. Ця ідея атаки насправді пов'язана з іншим методом атаки, і наведений вище аргумент пояснює, чому цей метод більше не може працювати [8].

З останніх результатів електронного друку в цьому напрямку, де вивчаються дуже загальна система, яка називається STS, ми знаємо, що їх метод може бути застосований і до нашого випадку. Відповідно до їх оцінки, безпека нашої системи становить принаймні $27 \times 33^3 \times (2^8)^{12} \times 5 > 2^{100}$.

2.4. Атака за допомогою структури багатошаровості.

Для випадку криптосистеми Мацумото – Імай Патарін зрозумів, що якщо шифр складається з декількох незалежних паралельних «гілок», можна виконати поділ змінних таким чином, що всі поліноми в шифрі виведені як лінійні комбінації поліномів над кожною групою змінних. Ця властивість насправді може бути використана для атаки на систему. На перший погляд, можна подумати, що рівні виглядають як різні «гілки». Тим не менше, слід усвідомити, що рівні жодним чином не є «незалежними», оскільки кожен з них будується на попере-

ньому. Простіше кажучи, можна сказати, що всі рівні злипаються, і ми ніяк не можемо зробити будь-якого розділення змінних. Це зрозуміло, коли розглядаються поліноми останнього рівню P_4. Тому атака з використанням властивості паралельних незалежних гілок у тут не може працювати. Подібним чином можна стверджувати, що атака з використанням системних систем також не може працювати тут, оскільки немає гілок і все насправді «склеєно» [2].

2.5. Загальні методи

Іншими методами, які можуть бути використані для атаки на нашу схему підписів, є ті, які безпосередньо вирішують поліноміальні рівняння, наприклад метод XL та різні його узагальнення, або такі, що використовують основи Грубонера. Безумовно, дуже складно вирішити набір з 27 рівнянь із 33 змінними, оскільки для цього набору рівнянь існує надто багато рішень. Загалом, набагато краще розв'язувати рівняння лише з однією змінною. Через характер проєктування системи можна здогадатися про значення для будь-якого набору змінних $v_1 = 6$, і ми маємо ймовірність $1 / e < 1 / 2.71828 < 0.37$ отримати унікальне рішення. Тепер задача стає проблемою вирішення набору з 27 квадратних рівнянь із 33 змінними. Ми повинні думати про це так, ніби це сукупність випадково вибраних квадратних рівнянь. Відповідно до того, що прийнято вважати, для вирішення цього набору рівнянь складність становить щонайменше $23 \times 27 > 281$.

З цього ми робимо висновок, що загальна складність атаки на наш приклад становить принаймні 280 [3].

2.6. Загальний аналіз безпеки

На основі цього можна побачити, що для атаки на систему можна підійти до неї або з верхнього рівня, або сформувати нижній рівень. Безпека нижнього рівня залежить від того, наскільки ефективно можна використовувати метод Minrank. Загалом складність атаки дорівнює $q^{(v2-1)}o^3{}_u - 1 if v_1 > o_1$, якщо $v_1 > o_1$, або $q^{2v1}o_u{}^3 - 1$, якщо $v_1 \le o_1$. З цього можна отримати, що не можна дозволити $v_2 = o_1 + v_1$ бути занадто малим. З останніх результатів електронного друку [WBP], безпека системи становить принаймні $(n - v_1) \times n^3 \times (q)^{o1+v1} \times u$, що, безсумнівно, вимагає, щоб $o_1 + v_1$ не був малим.

Що стосується випадку атаки зверху, метод атаки для незбалансованого методу Oil-Vinegar говорить, що $v_u - 1 - o_u - 1$ не може бути занадто малим. Також щоб уникнути випадкових атак пошуку $o_u - 1$ не повинно бути занадто малим [4].

3. Здатність алгоритму RAINBOW протидіяти атаці сторонніми каналами

Криптографічні системи повинні бути захищені від широкого кола атак, включаючи атаки сторонніми каналами. Атака сторонніми каналами належить до фізичної атаки, яка являє собою будь-яку атаку, засновану на інформації, отриманій в результаті фізичної реалізації криптографічних систем, а не на грубій силі чи теоретичних недоліках криптографічних алгоритмів. Основним принципом атаки бічного каналу є те, що інформація бічного каналу, така як споживання енергії, електромагнітні витоки, інформація про синхронізацію або навіть звук, може забезпечити додаткові джерела інформації про секрети в криптографічних системах, наприклад криптографічні ключі, часткова інформація про стан, повна або часткові звичайні тексти, які можна використовувати для розбиття криптографічних систем. Загальні класи атаки бічних каналів включають аналіз синхронізації, аналіз потужності, електромагнітний аналіз, аналіз несправностей, акустичний криптоаналіз, аналіз залишків даних та атаки аналізу молоткових рядів.

Атаки аналізу несправностей мають на меті маніпулювати екологічними умовами криптографічних систем, таких як напруга, годинник, температура, випромінювання, світло і вихровий струм, щоб генерувати несправності під час секретних обчислень, наприклад

множення та інверсії в кінцевому полі, і спостерігати за пов'язаною поведінкою, яка може допомогти криптоаналітику зламати криптографічні системи. Атаки аналізу несправностей можна спроєктувати, просто підсвітивши транзистор лазерним променем, що змушує деякі біти приймати неправильні значення. Ідея використання несправності, індукованої під час секретного обчислення, для вгадування секретного ключа практично спостерігалася в реалізаціях RSA, що використовують китайську теорему про залишки.

Атака аналізу потужності може надати детальну інформацію, спостерігаючи за енергоспоживанням криптографічних систем, що приблизно поділяється на простий аналіз потужності (SPA) та аналіз диференціальної потужності (DPA). У сімействі атак аналізу потужності DPA представляє особливий інтерес і є статистичним тестом, який вивчає велику кількість сигналів енергоспоживання для отримання секретних ключів.

Можна виділити наступні атаки:

- атака диференціального аналізу потужності на SFLASH;
- атака на секретні ключі від модуля SHA-1 схем SFLASH.
- атака стороннього каналу на enTTS, яка використовує диференціальний аналіз потужності та аналіз несправностей для атаки двох афінних перетворень та центральної трансформації карти. Цей метод показує, що можна отримати всі секретні ключі enTTS.

Оскільки конструкція Rainbow включає дві афінні перетворення та перетворення центральної карти, такі методи мають потенціал для отримання її секретних ключів. Таким чином, обговорюється захист від можливої атаки бічного каналу для Rainbow, а контрзаходи описані нижче:

- Нехай це повідомлення і кожен елемент у полягає в $GF((2^4)^2)$;
- Береться випадковий вектор $y'(y_0', y_1', ..., y_{25}')$, кожен елемент якого полягає в $GF((2^4)^2)$;
- Обчислюється $y'' = y' + y$;
- Обчислюється $\bar{y}' = Ay' + b$ та $\bar{y}'' = Ay''$, де A – матриця 26x26, b – вектор розміру 26;
- Обчислюється $\bar{y} = \bar{y}' + \bar{y}''$, що еквівалентно $\bar{y} = Ay + b$;
- Розраховано перше афінне перетворення; тоді ми беремо випадкові байти для Vinegar-змінних;
- Двічі перевіряються випадкові байти для захисту від атак аналізу несправностей;
- Обчислюються багатовимірні поліноміальні оцінки та розв'язування систем лінійних рівнянь до завершення перетворення центральної карти;
- $\bar{x}(\bar{x_0}, \bar{x_1}, ..., \bar{x_{42}})$ – це результат трансформації центральної карти; після цього береться два випадкових вектори $\bar{x}'$ та $\bar{x}''$, де $\bar{x} = \bar{x}' + \bar{x}''$, та елементи полягають в $GF((2^4)^2)$;
- Обчислюється $\bar{x}' = C\bar{x}'$ та $\bar{x}'' = C\bar{x}'' + d$, де C – матриця 43x43, b – вектор розміру 43;
- Обчислюється $\bar{x} = \bar{x}' + \bar{x}''$, що еквівалентно $x = Cx + d$;
- $x(x_0, x_1, ..., x_{42})$ це схема підпису Rainbow для $y(y_0, y_1, ..., y_{25})$.

Використовується аналіз несправностей для атаки випадкових байтів у центральних перетвореннях карти; таким чином ми двічі перевіряємо випадкові байти для захисту від атак аналізу несправностей. Також використовується аналіз диференціальної потужності для атаки модуля SHA-1; таким чином, ми беремо метод захисту афінних перетворень. Однак зазначений вище контрзахід є теоретичним; потрібна можливість впровадити та перевірити це на апаратному забезпеченні.

Висновки

1. Постквантова криптографія – частина криптографії, яка залишається актуальною і при появі квантових комп'ютерів і квантових атак. Так як за швидкістю обчислення традиційних криптографічних алгоритмів квантові комп'ютери значно перевершують класичні комп'юте-

рні архітектури, сучасні криптографічні системи стають потенційно вразливими до криптографічних атак. Більшість традиційних криптосистем спирається на проблеми факторизації цілих чисел або завдання дискретного логарифмування, які будуть легко розв'язані на досить великих квантових комп'ютерах, що використовують алгоритм Шора.

2. Багато криптографів ведуть розробку алгоритмів, незалежних від квантових обчислень, тобто стійких до квантових атак. Ці задачі розглянуті на другому етапі конкурсу NIST США.

3. Схема підпису Rainbow виглядає надійною проти великої кількості методів криптоаналізу та проти атак сторонніми каналами.

4. У зв'язку з можливістю появи потужного квантового комп'ютера актуальними є завдання створення постквантових алгоритмів ЕП. В цьому напрямі уже розпочато дослідження, в певній мірі визначено математичні основи, на яких можуть бути побудовані постквантові алгоритми ЕП. Для цього можна застосувати схему Rainbow.

5. Реалізація квантово-захищених алгоритмів вимагає великих матеріально-технічних ресурсів. Вказане пов'язане з великими довжинами ключів та загальних параметрів. Сучасний рівень розвитку техніки дозволяє оптимістично ставитися до можливості ефективної реалізації квантово-захищених алгоритмів.

6. Мультиваріативні квадратичні перетворення можуть бути застосованими для розроблення постквантового стандарту ЕП. Вони вже були використані для побудови схем підпису, але всі спроби побудувати надійну схему поки не були успішними. Попередній аналіз показав, що мультиваріативні квадратичні перетворення можуть вирішити проблему захищеності від атак на основі квантових комп'ютерів, але для цього ще потрібно провести величезний обсяг досліджень та робіт, а також вкласти значні ресурси.

7. Попередній аналіз показує, що розміри загальних параметрів та ключів не викликають сумнівів відносно криптографічної стійкості стандарту, розробленого на основі мультиваріативного квадратичного перетворення. Але залишається проблема просторової складності, яка пов'язана зі значними довжинами загальних параметрів та відкритих ключів.

Список літератури:

1. Lili Chen, Stephen Jordan, Yi-Kai-Liu, Dustin Moody, Rene Peralta, Ray Perlner, Daniel Smith-Tone. Report on Post – Quatum Cryptography. Nistir 8105 (draft).

2. Інтернет-ресурс. Режим доступу http://www.nkj.ru/archive/articles/5309/

3. Інтернет-ресурс. Режим доступу http://www.win.tue.nl/diamant/symposium05/abstracts/wolf.pdf

4. Горбенко Ю.І. Методи побудування та аналізу, стандартизація та застосування криптографічних систем : монографія ; заг. ред. І.Д. Горбенко. Харків : Форт, 2015. 959 с.

5. Потій О.В, Горбенко Ю.І., Ганзя Р.С., Пономар В.І. // Матеріали V міжнар. наук.-техн. конф. «Захист інформації і безпеки інформаційних систем». Львів, 2016, 02.06 – 03.06. С. 52.

6. Reinier Broker. Constructing supersingular elliptic curves // J. Comb. Number Theory, (3): pp. 269–273, 2009.

7. McGrew D., Curcio M. Hash-Based Signatures draft-mcgrew-hash-sigs-00[Электронный ресурс]. Режим доступу: https://tools.ietf.org/html/draftmcgrew-hash-sigs-00.

8. Lili Chen, Stephen Jordan, Yi-Kai-Liu, Dustin Moody, Rene Peralta, Ray Perlner, Daniel Smith-Tone. Report on Post – Quantum Cryptography. NISTIR 8105 (DRAFT). https://www.google.com.ua/search.

9. Bernstein D. J. Grover vs. McEliece // N. Sendrier, editor, Post-Quantum Cryptography, Third International Workshop, PQCrypto 2010, Darmstadt, Germany, May 25-28, 2010. Proceedings, volume 6061 of Lecture Notes in Computer Science, pages 73–80. Springer, 2010.

Надійшла до редколегії 02.04.2021

Відомості про автора:

Гармаш Дмитро Васильович – Харківський національний університет імені В.Н. Каразіна, аспірант кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук, Україна, e-mail: dmitriy.garmash96@icloud.com

УДК 003.026:004.056 **DOI:10.30837/rt.2021.2.205.08**

Г.А. МАЛЄЄВА

АНАЛІЗ ЗАХИЩЕНОСТІ ПОСТКВАНТОВОГО АЛГОРИТМУ ЕЛЕКТРОННОГО ПІДПИСУ RAINBOW ВІД ПОТЕНЦІЙНИХ АТАК

Вступ

Багатовимірна криптографія на основі відкритого ключа є кандидатом для постквантової криптографії, і це дозволяє генерувати особливо короткі підписи та швидку перевірку. Схема підписів Rainbow, запропонована Дж. Діном та Д. Шмідтом, є такою багатовимірною криптосистемою і вважається захищеною від усіх відомих атак. Ця схема підпису може бути реалізована просто та ефективно за допомогою лінійних методів алгебри над невеликим кінцевим полем і, зокрема, створює коротші підписи, ніж ті, що використовуються в RSA та інших постквантових підписах схеми [1]. У другому раунді NIST PQC пропонуються захищені набори параметрів Rainbow і проаналізовано кілька атак на них [1]. Зокрема, атака Rainbow-Band-Separation (RBS) [2] є найкращою серед відомих атак на Rainbow з певним набором параметрів і є важливою. Метою статті є спроба зрозуміти точну безпеку Rainbow від атаки RBS за допомогою F_4.

1. Порівняльний аналіз ЕП на основі MQ-перетворень за критерієм стійкість – складність

При порівняльному аналізі згідно [3] використані системи безумовних та умовних критеріїв оцінки та порівняльного аналізу електронних підписів (ЕП) на основі MQ-перетворень. Причому, якщо відповідні ЕП задовольняють безумовним критеріям, то в подальшому вони оцінюються та порівнюються за умовними критеріями. Таким чином, спочатку для алгоритмів ЕП обчислюються сукупності часткових безумовних критеріїв та інтегральні критерії для кожного з них. Потім для тих ЕП, що пройшли випробовування на першому етапі, обчислюються сукупності часткових умовних критеріїв та інтегральні умовні критерії для кожного з ЕП. Далі ранжування здійснюється з використанням безумовних та умовних інтегральних критеріїв. Перевага віддається алгоритмам ЕП, що пройшли відбір за безумовними критеріями, а також, що мають кращі показники щодо інтегральних умовних критеріїв. Таку методику запропоновано називати раціональною, тобто не оптимальною. Зрозуміло, що окремо ЕП на основі MQ-перетворень оцінюються щодо захищеності від атак сторонніми каналами. Після цього приймається рішення щодо переваг певного алгоритму ЕП на основі MQ-перетворення.

Для порівняння ЕП використані наступні характеристики та відповідні показники:

1) $I_{ст.}$ – рівень стійкості ЕП щодо класичних та квантових атак стійкості;
2) $l_{в.к.}$ – розмір відкритого ключа ЕП (байтів);
3) $l_{о.к.}$ – розмір особистого ключа ЕП (байтів);
4) $l_{рез.}$ – розмір підсумкового ЕП (байтів);
5) $T_{кл.}$ – швидкість (складність) створення ключової пари ЕП (тактів роботи);
6) $T_{пр.}$ – швидкість (складність) вироблення ЕП (тактів);
7) $T_{зв.}$ – швидкість (складність) перевірки ЕП (тактів).

Необхідно відмітити, що при необхідності число характеристик та відповідно показників може бути розширено.

Для визначення важливості тієї чи іншої характеристики (показника) використовуються експертні оцінки. При цьому, при порівнянні, певною перевагою та об’єктивністю є врахування та порівняння усієї множини кандидатів. В цьому якраз полягає сутність та особливість методу ранжування та застосування експертних оцінок.

У табл. 1 наведено експертні оцінки для вказаних характеристик, які були отримані від спеціалістів-криптологів.

Таблиця 1

Експертні оцінки характеристик криптографічних алгоритмів методом ранжування

Показники / Експерти	$I_{ст.}$	$l_{в.к.}$	$l_{о.к.}$	$l_{рез.}$	$T_{кл.}$	$T_{пр.}$	$T_{зв.}$
1	7	5	3	2	1	4	6
2	6	7	1	3	2	4	5
3	5	6	1	2	3	4	7
4	5	6	1	4	2	3	7
5	6	2	1	4	3	5	7
W	0,207	0,186	0,05	0,107	0,079	0,143	0,228

Із табл. 1 видно, що стійкість дійсно є важливою характеристикою для експертів, а реально це інтегральний безумовний критерій. Вона приймається за умови, що відповідні алгоритми ЕП пройшли відбір за інтегральним безумовним критерієм. Також враховано, що вона легко коригується за допомогою вибору параметрів. Крім того, видно, що складність (час) зворотного криптографічного перетворення, тобто час перевірки підпису, є більш важливою характеристикою.

Нижче наводяться результати аналізу та оцінки кожного із алгоритмів ЕП за допомогою методу ранжування та визначається алгоритм, що має найкращу модифікацію для кожного із алгоритмів, а також наводиться вибір більш перспективного алгоритму ЕП на основі MQ-перетворення. Відповідні дані наведені в табл. 2 і обрані найперспективніші кандидати.

Варто зазначити, що для характеристик 2) – 5) чим менше їх значення, тим краще, а для характеристик 1), 6) та 7) чим воно більше, тим краще, оскільки тим більш захищеним є алгоритм і більша швидкодія вироблення та перевірки ЕП.

Для аналізу було прийнято рішення провести порівняльний аналіз відповідно до кожного алгоритму, тобто обрати модифікації з найбільшою перевагою, і порівняти безпосередньо модифікації.

У табл. 2 представлено зведені характеристики щодо обраних модифікацій.

На рис. 1 зображено зведену гістограму відносної переваги алгоритмів ЕП на основі MQ-перетворень.

Таблиця 2

Зведені характеристики кращих модифікацій механізмів ЕП на базі MQ-перетворень другого етапу конкурсу NIST PQC

Характеристики / Модифікація	$I_{ст.}$	$l_{в.к.}$	$l_{о.к.}$	$l_{рез.}$	$T_{кл.}$	$T_{пр.}$	$T_{зв.}$
8-63-256	2	15,872	32	319	39,421,493	26,714,796	15,123,202
Ia	1	152,064	100,250	64	1,302,000,000	601,000	350,000
31-48	2	62	32	32,882	2,957,276	266,840,340	191,666,288
128	1	417,408	14,208	48	1,398,800,000	3,172,000,000	19,656,000

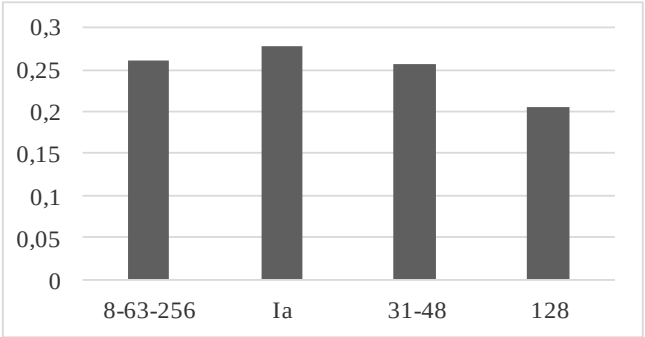


Рис. 1. Відносна перевага механізмів ЕП конкурсу NIST PQC за методом ранжування (серпень 2019, другий семінар)

Таким чином, беручи до уваги результати, які були поданими у [3], можна зробити висновок, що Rainbow та LUOV мають значну перевагу над іншими алгоритмами, як за умови надання мінімального рівня захисту, так і максимального. Ці два алгоритми ЕП в обмеженій групі лише MQ-перетворень можна розглядати щодо використання у постквантовий період як найбільш перспективні. Але необхідно відмітити, що як Rainbow, так і LUOV алгоритми можуть забезпечити обмежені рівні безпеки. Як правило обмеження пов'язані з забезпеченням максимум п'ятого рівня безпеки – мається на увазі 256 біт захищеності від класичних атак та 128 біт квантових атак.

2. Атака RAINBOW-BAND-SEPARATION (RBS)

Атака Rainbow-Band-Separation відновлює секретний ключ Rainbow, розв'язуючи певні системи квадратичних рівнянь, а його складність оцінюється за відомим показником, який називається ступенем регулярності. Однак, як правило, ступінь регулярності більша, ніж ступінь розв'язання в експериментах, і точної оцінки отримати неможливо. Попередні методи оцінки [1, 4] для складності атаки RBS використовують ступінь регулярності як її показник за припущенням, що система квадратичних рівнянь, розв'язана в атаці, є напіврегулярною. Для напіврегулярної системи ступінь регулярності задається як ступінь D_{reg} першого члена, коефіцієнт якого не позитивний у ряді потужностей

$$\frac{(1-t^2)^m}{(1-t)^n}, \qquad (1)$$

де m і n – числа рівнянь і змінних відповідно. Оскільки загальноприйнята квадратична система, що вирішується в прямій атаці, часто є напіврегулярною, то при оцінці складності прямої атаки використовується ступінь регулярності [1].

У роботі [5] запропоновано новий показник складності атаки Rainbow-Band-Separation за допомогою алгоритму F_4, який дає більш точну оцінку порівняно з показником, що використовує ступінь регулярності. Цей показник виводиться двома змінними рядами потужності

$$\frac{\prod_{i=1}^{m}(1-t_1^{d_{i1}}t_2^{d_{i2}})}{(1-t_1)^{n_1}(1-t_2)^{n_2}}, \qquad (2)$$

що збігається з однозмінним рядом потужностей при $t_1=t_2$, виводячи ступінь регулярності. Крім того, показано залежність між атакою Rainbow-Band-Separation за допомогою гібридного підходу та атакою HighRank. Розглядаючи це відношення та показник, ми отримали нову оцінку складності для атаки Rainbow-Band-Separation. Отже, завдяки цьому, можна зрозуміти точну безпеку Rainbow від атаки Rainbow-Band-Separation за допомогою алгоритму F_4.

3. Опис атаки RBS на схему підпису RAINBOW

Нехай m і n – натуральні числа. Позначимо через F кінцеве поле порядку q. Елемент $(f_1, ..., f_m)$ F $[x_1, ..., x_n]^m$ називається поліноміальною системою і дає відображення $F^n \to F^m$ на $a \to (f_1(a),...,f_m(a))$, яке називають поліноміальним відображенням (картою).

Багатовимірна схема підпису відкритого ключа складається з наступних трьох алгоритмів.

Генерація ключів: будуються дві обернені лінійні карти $S:F^n \to F^n$ і $T:F^m \to F^m$ випадковим чином і легко обернена квадратична карта $F: F^n \to F^m$, яку називають центральною картою, а потім обчислюється $P:=T \circ F \circ S$. Відкритий ключ подається у вигляді P. Кортеж (T, F, S) – секретний ключ.

Генерація підписів: для повідомлення $b \in F^m$ обчислюємо $b'=T^{-1}(b)$. Далі ми можемо обчислити елемент a' з $F^{-1}(\{b'\})$, оскільки F легко обернений. Отже, ми отримуємо підпис

$$a=S^{-1}(a') \in F^n.$$

Перевірка: перевіряється, чи $P(a)=b$ має місце. Для натуральних чисел v, o_1 і o_2, нехай $x=\{x_1, ..., x_v\}$, $y=\{y_1, ..., y_{o1}\}$ і $z=\{z_1, ..., z_{o2}\}$ будуть трьома змінними множинами і $n=v+o_1+o_2$, і $m=o_1+o_2$. Центральна карта $F=(f_1, ..., f_m)\in F[x, y, z]^m$ Rainbow

$$\begin{cases} f_1 = g^{(1)}(\mathbf{x}) + \sum_{i=1}^{o_1} l_i^{(1)}(\mathbf{x}) y_i, \\ \quad \vdots \\ f_{o_1} = g^{(o_1)}(\mathbf{x}) + \sum_{i=1}^{o_1} l_i^{(o_1)}(\mathbf{x}) y_i, \\ f_{o_1+1} = g^{(o_1+1)}(\mathbf{x},\mathbf{y}) + \sum_{i=1}^{o_2} l_i^{(o_1+1)}(\mathbf{x},\mathbf{y}) z_i, \\ \quad \vdots \\ f_{o_1+o_2} = g^{(o_1+o_2)}(\mathbf{x},\mathbf{y}) + \sum_{i=1}^{o_2} l_i^{(o_1+o_2)}(\mathbf{x},\mathbf{y}) z_i, \end{cases} \tag{3}$$

де $g^{(j)}$ та $l^{(j)}_i$ – випадковим чином обрані квадратичні многочлени та лінійні многочлени відповідно. Тоді за алгоритмом генерації підписів, наведеним вище, ми можемо легко обчислити елемент a' у попередньому зображенні будь-якого елемента $b'=(b'_1, ..., b'_{o1+o2})$ у F^m під F наступним чином.

1. Випадково обрати $a_v{}' = (a_1', ..., a_v')$ як x.
2. Вирішити систему лінійних рівнянь

$$f_1(\mathbf{a}'_v, \mathbf{y}) = b'_1, \ldots, f_{o_1}(\mathbf{a}'_v, \mathbf{y}) = b'_{o_1}.$$

Нехай $\mathbf{a}'_{o_1} = (a'_{v+1}, \ldots, a'_{v+o_1})$ є одним із її рішень, якщо воно існує. В іншому випадку повернутись до кроку 1.

3. Вирішити систему лінійних рівнянь

$$f_{o_1+1}(\mathbf{a}'_v, \mathbf{a}'_{o_1}, \mathbf{z}) = b'_{o_1+1}, \ldots, f_{o_1+o_2}(\mathbf{a}'_v, \mathbf{a}'_{o_1}, \mathbf{z}) = b'_{o_1+o_2}.$$

Нехай $\mathbf{a}'_{o_2} = (a'_{v+o_1+1}, \ldots, a'_{v+o_1+o_2})$ є одним із її рішень, якщо воно існує. В іншому випадку повернутись до кроку 1.

4. Отримати елемент $\mathbf{a}' = (a'_1, \ldots, a'_{v+o_1+o_2})$ у попередньому зображенні b'.

Нехай (v, o_1, o_2) – набір параметрів Rainbow, покладемо $n=v+o_1+o_2$ і $m=o_1+o_2$. Для відкритого ключа Rainbow $P=(p_1, ..., p_m)$ атака RBS відновлює свій секретний ключ (T, F, S) наступним чином. За визначенням (3) центральної карти $F=(f_1, ..., f_m)$ кожна матриця, відповідна f і має такий вигляд:

$$M_{f_i} = \begin{cases} \begin{pmatrix} *_{v\times v} & *_{v\times o_1} & 0_{v\times o_2} \\ *_{o_1\times v} & 0_{o_1\times o_1} & 0_{o_1\times o_2} \\ 0_{o_2\times v} & 0_{o_2\times o_1} & 0_{o_2\times o_2} \end{pmatrix} \text{ if } 1 \le i \le o_1, \\ \begin{pmatrix} *_{v\times v} & *_{v\times o_1} & *_{v\times o_2} \\ *_{o_1\times v} & *_{o_1\times o_1} & *_{o_1\times o_2} \\ *_{o_2\times v} & *_{o_2\times o_1} & 0_{o_2\times o_2} \end{pmatrix} \text{ if } o_1+1 \le i \le o_1+o_2. \end{cases} \tag{4}$$

Тут $*_{k\times l}$ означають k на l матриці над F. Аналогічно, матриці, відповідні S і T, можна записати наступним чином.

Матриці M_{p1}, ..., M_{pm}, що відповідають відкритим поліномам p_1, ..., p_m, задаються як

$$(M_{p_1}, \ldots, M_{p_m}) = (M_S M_{f_1}{}^t M_S, \ldots, M_S M_{f_m}{}^t M_S) M_T. \tag{5}$$

Існує вектор n на 1 $s=(\lambda_1, ..., \lambda_{v+o1}, 0, ..., 0,1)$ такий, що $s\cdot M_S=(0,...,0,1)$ Тоді для $i=1,...,m$, маємо

$$\mathbf{s} \cdot M_S M_{f_i}{}^t M_S \cdot {}^t\mathbf{s} = (0, \ldots, 0, 1) \cdot M_{f_i} \cdot {}^t(0, \ldots, 0, 1) = 0.$$

Оскільки кожен M_{pk} є лінійною комбінацією $M_S M_{f1}{}^t M_S{}^t;:::; M_S{}^t M_S M_{fm}{}^t$, отримуємо

$$\mathbf{s} \cdot M_{p_k} \cdot {}^t\mathbf{s} = 0, \quad k = 1, \ldots, m. \tag{6}$$

Існує вектор m на 1 $t=(1,0 ...,0, \lambda_{v+o1+1}, ..., \lambda_{v+o1+o2})$ такий, що $M_T \cdot {}^t t = {}^t(1,0, ..., 0)$. Потім,

помноживши рівняння (5) на ${}^t t$, отримаємо

$$M_{p_1} + \sum_{i=1}^{o_2} \lambda_{v+o_1+i} M_{p_{o_1+i}} = M_S M_{f_1} {}^t M_S. \quad (7)$$

де e_k це n на 1 вектор (0;...;0;1;0;...;0). Тут вилучаємо випадок $k=n$, оскільки рівняння (7) для $k=n$ випливає з рівняння (6).

Оскільки s = (λ_1, ..., λ_{v+o1}, 0, ..., 0,1), зрозуміло, що рівняння (6) і (7) є $n+m-1$ квадратичними рівняннями в n змінних $\lambda_1,...,\lambda_n$ і будуються з відкритого ключа $p_1,...,p_m$. Вирішивши ці квадратичні системи, зловмисник може відновити частину секретного ключа S і T, а сааме – s і t. Атака RBS може відновити S і T, повторюючи подібні обговорення, як описано вище (детальніше див. [2]).

Оскільки складність розв'язання квадратичної системи домінує в одній з атаки RBS, достатньо оновити лише систему. Квадратична система, що складається з рівнянь (6) та (7), називається домінуючою системою RBS.

З досліджень [5] можна зробити висновок, що домінуюча система RBS є нерегулярною та дворівневою.

4. Показник складності розв'язання дворівневої системи поліномів

Для дворівневої поліномічної системи $(h_1,...,h_m)$ у $F[x_1,...,xn_1,y_1,...,y_{n2}]^m$, де $deg\ Z^2_{\geq 0}\ h_i=(d_{i1}, d_{i2})$,

$$\sum_{(d_1,d_2)\in\mathbb{Z}^2_{\geq 0}} a_{(d_1,d_2)} t_1^{d_1} t_2^{d_2} = \frac{\prod_{i=1}^{m}(1-t_1^{d_{i,1}} t_2^{d_{i,2}})}{(1-t_1)^{n_1}(1-t_2)^{n_2}}, \quad (8)$$

і $D_{bgd}=D_{bqd}(h_1,...,h_m)$ визначається як мінімальне значення наступного набору, якщо воно існує, – $\{d_1+d_2 \mid a_{(d_1,d_2)} < 0\}$.

Двохзмінний ряд у (8) розглядається як двозначна версія гільбертового ряду.

Зауваження 1. Для дворівневої системи зазначимо, що однозмінний ряд потужності

$$\frac{\prod_{i=1}^{m}(1-t^{\deg f_i})}{(1-t)^n}. \quad (9)$$

що виводить D_{reg}, збігається з двома змінними рядами потужності (8), коли $t=t_1=t_2$. Отже, якщо ми визначимо D'_{bgd} як мінімальне значення набору $\{d_1+d_2 \mid a_{(d_1,d_2)} \leq 0\}$, де a (d_1, d_2) – коефіцієнт ${t^{d1}}_1 {t^{d2}}_2$ у ряді (8) і він існує, тоді $D'_{bgd} \leq D_{reg}$. D'_{bgd} часто менший, ніж ступінь вирішення для деяких наборів параметрів Rainbow. Таким чином, ми не використовуємо D'_{bgd} як відповідний показник. З іншого боку, термін t^{Dreg} у ряді (9) часто має від'ємний коефіцієнт, який виводить один із ${t^{d1}}_1 {t^{d2}}_2$ у ряді (8), де $d_1+d_2=D_{reg}$. А саме, відношення $D_{bgd} \leq D_{reg}$ часто дотримується (див. табл. 3 та 4).

Виходячи з цього, можна побачити, що введений показник D_{bgd} щільно наближає ступінь розв'язання домінуючої системи RBS, ніж ступінь регулярності. Показник D_{bgd} для домінуючої системи RBS з набором параметрів (v,o_1,o_2) задається мінімальною сумарною мірою показників, коефіцієнт яких від'ємний у двовимірній потужності

$$\frac{(1-t_1t_2)^{v+o_1+o_2-1}(1-t_1^2)^{o_1+o_2}}{(1-t_1)^{v+o_1}(1-t_2)^{o_2}}. \quad (10)$$

У табл. 3 порівнюється показник D_{bgd} та ступінь регулярності D_{reg} для домінантних систем RBS з $v=o_i$ та $v \lesssim 2o_i$.

Таблиця 3

D_{bgd} проти D_{reg} для домінуючої системи RBS

q=256	Експеримент			Теорія	
(v; o_i)	d_{slv}	d_{tim}	d_{mem}	D_{bgd}	*Невеликий залишок*
(4; 3)	4	4	4	4	4
(5; 3)	4	4	4	4	5
(6; 3)	4	4	4	4	5
(6; 4)	4	4	4	4	5
(7; 4)	4	4	4	4	6
(8; 4)	4	4	4	5	6
(8; 5)	5	5	5	5	6
(9; 5)	5	5	5	5	6
(10; 5)	5	5	5	5	7
(10; 6)	6	6	6	6	7
(11; 6)	6	6	6	6	7
(12; 6)	6	6	6	6	7
q=16	Експеримент			Теорія	
(v; o_i)	d_{slv}	d_{tim}	d_{mem}	D_{bgd}	*Невеликий залишок*
(3; 3)	3	3	3	3	4
(4; 4)	4	4	4	4	5
(5; 5)	4	4	4	4	5
(6; 6)	5	5	5	5	6
(7; 7)	5	5	5	5	6
(8; 8)	5	6	6	6	7
(9; 9)	6	6	6	6	7

Експериментальні ступені d_{slv}, d_{mem} і d_{tim} в алгоритмі F_4 та теоретичних ступенях D_{bgd} (із серії (10)) та D_{reg} (при k=0) для домінантної системи RBS з $v \lesssim 2o_i$ або $v=o_i$ (i=1,2). Запропонований показник D_{bgd} збігається з d_{slv} у випадках, за винятком (q, v, o_i) = (256,8,4), (16,8,8). Ступінь регулярності D_{reg} завжди більша, ніж d_{slv}, за винятком (q, v, o_i) = (256,4,4).

Експериментальні ступені d_{slv}, d_{mem} і d_{tim} в алгоритмі F_4 та теоретичних ступенях D_{bgd} та D_{reg} гібридного підходу щодо домінуючих систем RBS у змінних $\{\lambda_1, ..., \lambda_{v+o1+o2}\}$ для (q,v,o_1,o_2)=(256,10,5,5) і (16,8,8,8). Цілі числа k_1 і k_2 – це кількість змінних, зафіксованих гібридним підходом у $\{\lambda_1, ..., \lambda_{v+o1}\}$ і $\{\lambda_{v+o1+1}, ..., \lambda_{v+o1+o2}\}$ відповідно. Ступінь регулярності D_{reg} завжди більша, ніж ступінь вирішення d_{slv}. Пропонований показник D_{bgd} щільно наближає d_{slv}, ніж D_{reg}, і є верхньою межею d_{slv}.

Крім того, у табл. 4 порівнюється показник D_{bgd} та ступінь регулярності D_{reg} для гібридного підходу при атаці RBS проти наборів параметрів Rainbow (q,v,o_1,o_2)=(256,10,5,5) та (16,8,8,8). Тут k_1 і k_2 – це кількість змінних, зафіксованих гібридним підходом у $\{\lambda_1, ..., \lambda_{v+o1}\}$ і $\{\lambda_{v+o1+1}, ..., \lambda_{v+o1+o2}\}$, де $\lambda_1, ..., \lambda_{v+o1+o2}$ – змінні домінуючої системи RBS (див. вирази (6) та (7)). Тоді показник D_{bgd} задається мінімальною сумарною мірою показників, коефіцієнт яких від'ємний у двох змінних рядах потужності

$$\frac{(1-t_1t_2)^{v+o_1+o_2-1}(1-t_1^2)^{o_1+o_2}}{(1-t_1)^{v+o_1-k_1}(1-t_2)^{o_2-k_2}}. \qquad (11)$$

Таблиця 4

D_{bgd} проти D_{reg} для гібридного підходу в домінантній системі RBS

(256; 10; 5; 5)		Експеримент			Теорія	
k_1+k_2	$(k_1; k_2)$	d_{slv}	d_{tim}	d_{mem}	D_{bgd}	*Невеликий залишок*
0	(0; 0)	5	5	5	5	7
1	(1; 0)	5	5	5	5	6
	(0; 1)	4	4	4	5	6
2	(2; 0)	4	5	5	5	6
	(1; 1)	4	4	4	4	6
	(0; 2)	4	4	4	4	6
3	(3; 0)	4	4	4	4	6
	(2; 1)	4	4	4	4	6
	(1; 2)	3	4	4	4	6
	(0; 3)	3	3	3	3	6
4	(4; 0)	4	4	4	4	5
	(3; 1)	3	4	4	4	5
	(2; 2)	3	3	3	3	5
	(1; 3)	3	3	3	3	5
	(0; 4)	2	2	2	2	5
(16; 8; 8; 8)		Експеримент			Теорія	
k_1+k_2	$(k_1; k_2)$	d_{slv}	d_{tim}	d_{mem}	D_{bgd}	*Невеликий залишок*
0	(0; 0)	5	6	6	6	7
1	(1; 0)	5	5	5	5	6
	(0; 1)	5	5	5	5	6
2	(2; 0)	5	5	5	5	6
	(1; 1)	5	5	5	5	6
	(0; 2)	5	5	5	5	6
3	(3; 0)	4	5	5	5	6
	(2; 1)	4	5	5	5	6
	(1; 2)	4	5	5	5	6
	(0; 3)	4	4	4	5	6
4	(4; 0)	4	4	4	4	6
	(3; 1)	4	4	4	5	6
	(2; 2)	4	4	4	4	6
	(1; 3)	4	4	4	4	6
	(0; 4)	4	4	4	4	6

Зауваження 2. Як правило, перший доданок, що має від'ємний коефіцієнт у ряді потужностей (9), дає комплекс Кошу, якщо вона існує. Отже, очікується, що D_{bgd} надає комплекс

Кошу, а саме алгоритми на основі підпису, які повертають генератори модуля сизигії, повинні обчислювати до ступеня D_{bgd}.

Атака RBS за допомогою гібридного підходу при $k_2=o_2$ стає схожою на атаку HighRank [6]. Оскільки центральна карта Rainbow має матриці низького рангу M_{f1}, ..., M_{fo1} (див. вираз (4)), можна отримати квадратичний многочлен нижчого рангу, знайшовши лінійну комбінацію $M_{p1},...,M$. Для матриць o_2+1 з $M_{p1},...,M_{pm}$ атака HighRank відновлює такий квадратичний многочлен, знаходячи лінійну комбінацію, підпростір ядра якого має розмір одиниці. З іншого боку, атака RBS з використанням гібридного підходу при $k_2=o_2$ фіксує o_2 значення λ_{v+o1+1}, ..., $\lambda_{v+o1+o2}$ для отримання лінійної комбінації

$$M_{p_1} + \sum_{j=1}^{o_2} \lambda_{v+o_1+j} M_{p_{a_1+j}}$$

матриць o_2 +1 $M_{p1},M_{po1+1},...,M_{po1+o\,2}$ і розв'язує систему лінійних рівнянь (7) в $v+o_1-k_1$ змінних. Отже, атака має квадратичний поліном нижчого рангу і схожа на атаку HighRank.

Висновки

1. При порівнянні ЕП перевага віддається алгоритмам ЕП, що пройшли відбір за безумовними критеріями, а також, що мають кращі показники щодо інтегральних умовних критеріїв, оскільки така методика є більш раціональною.

3. При порівнянні рекомендується використовувати наступні характеристики та відповідні показники: $I_{ст.}$ – рівень стійкості ЕП щодо класичних та квантових атак стійкості; $l_{в.к.}$ – розмір відкритого ключа ЕП (байтів); $l_{о.к.}$ – розмір особистого ключа ЕП (байтів); $l_{рез.}$ – розмір підсумкового ЕП (байтів); $T_{кл.}$ – швидкість (складність) створення ключової пари ЕП (тактів роботи); $T_{пр.}$ – швидкість (складність) вироблення ЕП (тактів); $T_{зв.}$ – швидкість (складність) перевірки ЕП (тактів). Ранжування за цими показниками необхідне для визначення переваг певного алгоритму ЕП на основі MQ-перетворень.

4. Для наведених характеристик 2) – 5) у табл. 1, чим менше їх значення, тим краще, а для характеристик 1), 6) та 7), чим воно більше, тим краще, оскільки тим більш захищеним є алгоритм і більша швидкодія вироблення та перевірки ЕП.

5. Алгоритми ЕП Rainbow та LUOV в класі MQ-перетворень мають значну перевагу над іншими алгоритмами, як за умови надання мінімального рівня захисту, так і максимального. Ці два алгоритми ЕП в обмеженій групі лише MQ-перетворень можна розглядати щодо використання у постквантовий період як найбільш перспективні.

6. Необхідно відмітити, що як Rainbow, так і LUOV алгоритми можуть забезпечити обмежені рівні безпеки. Як правило, обмеження пов’язані з забезпеченням максимум п’ятого рівня безпеки. Мається на увазі 256 біт захищеності від класичних атак та 128 біт квантових атак.

7. Оскільки атака Rainbow-Band-Separation (RBS) відновлює секретний ключ Rainbow, розв’язуючи певну дворівневу багаточленну систему, можна використовувати D_{bgd} для оцінки складності цієї атаки.

8. Згідно з експериментами [5], що використовують F_4 для зменшених наборів параметрів Rainbow у другому раунді NIST PQC, показник D_{bgd} більше наближає ступінь розв'язання, ніж ступінь регулярності D_{reg}, який використовувався раніше. Тоді відношення $D_{bgd} \leq D_{reg}$ дотримується завжди.

9. Крім того, атака RBS може звести отриману поліномну систему до лінійної системи, використовуючи гібридний підхід із спеціальними налаштуваннями. Тоді ця атака стає схожою на атаку HighRank.

Список літератури:

1. Ding J., Chen M.-S., Petzoldt A., Schmidt D., Yang B. Y. Rainbow – Algorithm Specification and Documentation. Specification document of NIST PQC 2nd round submission package (2019)

2. Ding J., Yang B.-Y., Chen C.-H. O., Chen M.-S. and Cheng C.-M. New differential-algebraic attacks and reparametrization of Rainbow // Bellovin, S.M., Gennaro, R., Keromytis, A.D., Yung, M. (eds.) ACNS 2008, LNCS, vol. 5037, pp. 242–257. Springer (2008).

3. Кудряшов І. С., Малєєва Г. А. Аналіз властивостей електронних підписів на базі MQ перетворень / Ін-т кібернетики імені В. М. Глушакова НАН України ; Кам'янець-Подільський нац. ун-т імені Івана Огієнка // Математичне та комп'ютерне моделювання / Кам'янець-Подільський нац. ун-т імені Івана Огієнка. Кам'янець-Подільський, 2019. (Технічні науки: зб. наук праць; 19). С. 69-74.

4. Thomae E. A Generalization of the Rainbow Band Separation Attack and its Applications to Multivariate Schemes // IACR Cryptology ePrint Archive (2012). https://eprint.iacr.org/2012/223.

5. Nakamura S., Ikematsu Y., Wang Y., Ding J., Takagi T. New Complexity Estimation on the Rainbow-Band-Separation Attack. Specification document of NIST PQC.

6. Coppersmith D., Stern J., Vaudenay S. Attacks on the birational signature scheme // Stinson D.R. (ed.) CRYPTO 1994, LNCS vol. 773, pp. 435–443. Springer (1994).

Надійшла до редколегії 05.03.2021

Відомості про автора:

Малєєва Ганна Андріївна – Харківський національний університет радіоелектроніки, аспірант кафедри безпеки інформаційних технологій, факультет комп'ютерної інженерії та управління; Україна; e-mail: hanna.malieieva@nure.ua

УДК 621.391:519.2 **DOI:10.30837/rt.2021.2.205.09**

Є.В. КОТУХ, канд. техн. наук, О.В. СЄВЄРІНОВ, канд. техн. наук,
А.В. ВЛАСОВ, канд. техн. наук, Л.С. КОЗІНА, А. О. ТЕНИЦЬКА, К.О. ЗАРУДНА

МЕТОДИ ПОБУДОВИ ТА ВЛАСТИВОСТІ ЛОГАРИФМІЧНИХ ПІДПИСІВ

Вступ

У статті запропонований огляд властивостей перспективного напряму розвитку криптографічних систем на основі логарифмічних підписів і покриттів кінцевих груп, який належить до постквантової криптографії. Актуальний стан цього напряму й праці останніх років дають підстави припускати, що завдання факторизації елемента кінцевої групи в теорії побудови криптосистем на основі неабелевих груп з використанням логарифмічних підписів є обчислювально складні, що потенційно забезпечує необхідний рівень криптографічного захисту перед атаками, що використовують можливості квантових обчислень. У роботі наведено основні визначення логарифмічних підписів і покриттів кінцевих груп, їх класифікацію, властивості. У рамках аналізу підходів розглянуто практичні випадки використання властивостей логарифмічних підписів для забезпечення криптостійкості базових платформ для побудови криптосистем MST3.

Разом зі зростанням практичних можливостей використання квантових обчислень зростає загроза класичним схемам шифрування та електронного підпису, які використовують як основу класичні математичні проблеми, що долаються обчислювальними можливостями квантових комп'ютерів. Цей факт мотивує дослідження фундаментальних теорем, що стосуються математичних та обчислювальних аспектів постквантових криптосистем-кандидатів. У роботі представлено логарифмічні підписи як особливий тип факторизації в кінцевих групах, розглянуто їх властивості та методи побудови.

Нехай G – кінцева група. Логарифмічний підпис α для групи G є послідовністю підмножин $A_i \subseteq G$ виду $\alpha = [A_1, \ \dots, \ A_s]$, таких, що для кожного елемента g групи G є лише одна факторизація $(*)$ $g = a_q \cdot a_2 \cdot \ \dots \ \cdot a_s$, де $a_i \in A_i$ для $i = 1, \dots, s$. Множини A_i називають блоками. Розмір списку блоків позначають через $r_i := |A_i|$. Для спрощення ми називаємо елементи $A_i \cup \ \dots \ \cup A_s$ елементами логарифмічного підпису α . За певних умов ми розглядаємо впорядкування елементів блоку, тоді для $k_i = 0, \dots, r_i - 1$ позначаємо через a_{ik_i} кожний $(k_i + 1)$-й елемент блоку A_i. Вектор $(r_1, \dots, r_s)$ називають типом α, а

$$\ell(\alpha) = \sum_{i=1}^{s} r_i -$$

довжиною логарифмічного підпису. Множину логарифмічних підписів групи позначають через $\Lambda(G)$. Із визначення одержуємо певні властивості логарифмічних підписів. Через одноманітність факторизації (*) маємо $\prod_{i=1}^{s} r_i = |G|$, і тоді r_i ділить $|G|$ для всіх i. Це також показує, чому зазначені послідовності називають логарифмічними підписами. Логарифмічна функція перетворює добуток на суму, логарифмічні підписи скорочено відображають усі елементи групи довжини $r_1 \ + \ \dots \ + \ r_s$, де група містить у собі $r_1 \cdot \ \dots \ \cdot \ r_s$ елементів.

Якщо $e_G \in A_1 \cap \dots \ \cap A_s$, тоді ми можемо стверджувати, що α є нормалізованою. З огляду на одиничність факторизації ми навіть маємо $\bigcap_{i=1}^{s} A_i = \{e_G\}$ для нормалізованих логарифмічних підписів із більше ніж одним блоком. Легко помітити, що в абелевих $|A_i \cap A_j| \leq 1$ для $i \neq j$. У [1] доведено, що для $|G| = \prod_{j=1}^{t} p_j^{b_j}$ (p_j – просте число) є нижня границя довжини для будь-якого логарифмічного підпису групи G, одержаного таким чином:

$$\ell(\alpha) \geq \sum_{j=1}^{t} b_j \cdot p_j.$$

Логарифмічний підпис α називають мінімальним, якщо $\ell(\alpha)$ досягає нижньої границі, тобто кожен блок має простий ступінь або ступінь 4. У кількох статтях порушене питання існування мінімальних логарифмічних підписів у кінцевих полях [1, 2]. Було запропоновано мінімальні логарифмічні підписи для всіх кінцевих груп.

Приклад 1. Нехай $n \in \mathbb{N}$. Для циклічної групи $(\mathbb{Z}_{2^n}, +)$ послідовність виду $\alpha = \left[[0, 2^{n-1}], [0, 2^{n-2}], \ \ldots\ , \ [0,2], [0,1]\right]$ є нормалізованим логарифмічним підписом типу $(2, \ \ldots\ , 2)$. Обчислення факторизації елемента еквівалентне обчисленню його двійкового відображення, зокрема, якщо $n = 4$, $9 = 1001$ має факторизацію $2^3 + 0 + +0 + 2^0$.

Розглянемо можливість обчислення факторизації елемента групи для зазначеного логарифмічного підпису й певного елемента групи. Для прикладу, атака повним перебором за допомогою пошуку всіх можливих факторизацій, репрезентованих логарифмічним підписом $\alpha = [A_1, \ \ldots, \ A_s]$ групи G, становить $|G| \times (s-1)$ групових операцій у найгіршому випадку. Такий перебір знаходить правильну факторизацію для будь-якого логарифмічного підпису, але загалом неможливий. Приклад 1 показує, що для певних логарифмічних підписів легко обчислити факторизації. Для практичного використання в криптосистемах MST необхідно визначити логарифмічні підписи, для яких факторизація є обчислювально нездійсненною, а також підписи, для яких є ефективні алгоритми розкладання. Здебільшого, терміни «прості» й «складні» логарифмічні підписи вживають для позначення різниці між логарифмічними підписами, для яких відповідно обчислювально легко та складно одержати факторизації [3, 4]. Факторизація одного логарифмічного підпису становить постійний час, але коли ми вивчаємо питання ефективності обчислень для логарифмічних підписів, то розглядаємо сімейство (G_n, α_n) логарифмічних підписів $\alpha_n = \left[A_1^n, \ \ldots, \ A_{s_n}^n\right]$ груп G_n для $n \in \mathbb{N}$. Далі ми припускаємо, що $|G_n| \leq |G_{n+1}|$ і $\alpha_n \neq \alpha_m$ для $n \neq m$. Крім того, нехай одноманітний опис елементів G_n є таким, що $|g|_2 = \mu$ для всіх $g \in G_n$. Зауважуємо, що дає $\mu_n \geq log|G_n|$ для всіх $n \in \mathbb{N}$.

Робимо одне базове припущення: для зазначеного сімейства $(G_n, \ \alpha_n)$ є детермінований алгоритм поліноміального часу A, такий, що, маючи вхідні значення $(a_1, \ \ldots, \ a_{s_n})$ із $A_1^n \times \ \ldots \ \times A_{s_n}^n$, обчислює добуток $a_1 \cdot \ \ldots \ \cdot a_{s_n}$. Ми ідентифікуємо A з ін'єктованою функцією, що його обчислює.

Визначення 1. Для $n \in \mathbb{N}$ нехай α_n буде як у наведеному прикладі. Тоді сімейство $(G_n, \ \alpha_n)_{n \in \mathbb{N}}$ називають складним, якщо для кожного ймовірного алгоритму поліноміального часу A' для кожного позитивного полінома p і всіх істотно великих n маємо

$$pr\big(A'(g_n, \alpha_n) = A^{-1}(g_n)\big) < \frac{1}{p(\lceil log|G_n| \rceil)},$$

де g_n позначає випадковий елемент, одноманітно вибраний із G_n.

Для сімейства складних логарифмічних підписів $(G_n, \alpha_n)_{n \in \mathbb{N}}$ функція A визначає односторонню функцію, як зазначено в [5]. Серед логарифмічних підписів, що не є складними, ми визначаємо ті, які мають ефективні факторизації для всіх елементів групи.

Визначення 2. Для $n \in \mathbb{N}$ нехай a_n буде логарифмічним підписом групи G_n. Сімейство $(G_n, \alpha_n)_{n \in \mathbb{N}}$ називають простим, якщо є алгоритм A'', що, одержуючи на вході елементи $g_n \in G_n$ і α_n, обчислює факторизацію g_n щодо α_n за поліноміальний час.

Різницю між «нескладний» і «простий» уперше розглянуто в [6], і, на відміну від попередніх визначень (у яких ці два поняття еквівалентні), визначення є більш точним. Логарифмічний підпис, що не є ні простим, ні складним, може бути таким, для якого ми можемо ефективно знайти факторизацію для рівно половини групових елементів. Є алгоритм із заданими g і α_n, що випадково вгадує факторизацію з імовірністю $\frac{1}{|G_n|}$. Отже, алгоритм A' у другому

визначенні дає значно більший шанс знайти факторизацію, як і раніше менший, ніж $\frac{1}{p(\lceil log|G_n| \rceil)}$ (для будь-якого p). Алгоритм A' додатково одержує n на вхід для кодування групи G_n, такий, що A' обчислює G_n. Але твердження, що, маючи α_n на вході, гарантовано A' обчислює групу G_n, є не підтвердженими. Також припустимо, що для вхідних даних $(a_1, \ \dots, a_{s_n})$ алгоритм A знає, яку групову операцію застосовувати. Обчислення добутку елементів s_n, кожний із яких був випадково одноманітно вибраним із різних блоків a_n (щодо A), еквівалентне вибору випадкового елемента з G_n.

Відзначимо, що питання наявності варіантів для складних логарифмічних підписів залишається відкритим. Усі логарифмічні підписи, описані в літературі, репрезентовані простими [1 – 4, 7]. Криптосистеми MST використовують як складні, так і прості логарифмічні підписи для побудови криптосистем відкритого ключа. Криптосистема MST_1 ґрунтується на складних логарифмічних підписах; MST_3, у свою чергу, використовує прості логарифмічні підписи в елементарних абелевих 2-групах. Отже, нас цікавить структура простих логарифмічних підписів, а також груп, для яких можуть існувати логарифмічні підписи. Для подальшого аналізу груп важливими є три параметри. Для $n \in \mathbb{N}$ логарифмічний підпис $\alpha_n = [A_1^n, \ \dots, \ A_{s_n}^n]$ повністю описаний як

$$|\alpha_n| = \sum_{i=1}^{s} |A_i^n|$$

елементів із групи G_n. Можливо, є коротше репрезентування певних логарифмічних підписів, але загалом нам необхідний

$$\sum_{i=1}^{s} |A_i^n| \cdot \mu_n$$

біт, щоб виразити логарифмічний підпис, у якому елемент G_n репрезентований за допомогою μ_n біт. Нехай

$$r_n := max\{|A_i^n|, \ \dots, |A_{s_n}^n|\}.$$

Алгоритм факторизації A' у другому визначенні бере як вхідні дані елемент G_n і логарифмічний підпис α_n для певних $n \in \mathbb{N}$. Довжина цих вхідних даних дорівнює не більше за $(s_n \cdot r_n + 1) \cdot \mu_n$ біт.

Отже, ми пропонуємо використовувати три таких значення $(s_n, \ r_n, \ \mu_n)$ як параметри вимірювання ефективності алгоритму факторизації для α_n. Зазначимо, що ми одержуємо перші два параметри зі структури α_n, а третій параметр є незалежним від α_n та обумовлений лише репрезентуванням елементів G_n.

Зауваження . Для $n \in \mathbb{N}$ нехай (α_n) буде логарифмічним підписом у G_n. Якщо для всіх $g \in G_n$ факторизацію щодо α_n можна одержати за поліноміальний час із трьома параметрами s_n, r_n, μ_n, тоді сімейство $(\alpha_n)_n$ є простим.

Приклад 2. Візьмемо сімейство логарифмічних підписів у групі $G_n = \mathbb{Z}_{2^n}$ із прикладу 1. Три параметри ефективності дорівнюватимуть $s_n = n, r_n = 2, \mu_n = n$. Для заданого елемента $g \in \mathbb{Z}_{2^n}$ за допомогою його двійкового представлення $(g_1, \dots, \ g_n)$, де $g_i \in \{0,1\}$ – його факторизація, яка щодо α дорівнює $(g_1 \cdot 2^{n-1}, \dots, g_n \cdot 1)$, що одержано за допомогою не більше ніж n множень. Отже, маємо лінійний час у n. Тоді $(\alpha_n)_{n \in \mathbb{N}}$ є простою.

Перетворення логарифмічних підписів

Розглянемо певні перетворення логарифмічних підписів, зокрема такі, щоб факторизація щодо вихідного логарифмічного підпису була однаково ефективною щодо перетвореного логарифмічного підпису. Ідея полягає в тому, що алгоритм факторизації одного логарифмічного підпису в певному наборі дає алгоритм факторизації для всіх логарифмічних підписів цього набору. Наведемо стандартний підхід до класифікації логарифмічних підписів відповідно до структури блоків. Ми розглядаємо п’ять перетворень логарифмічних підписів, що не

змінюють властивостей належності до простих або складних логарифмічних підписів. Нехай $\alpha = [A_1, \dots, A_s]$ буде логарифмічним підписом групи G. Ми маємо справу з перетвореннями блоків $A_1, \dots,\ A_s$ логарифмічного підпису α на блоки $B_1, \dots, B_s$, які є результуючою послідовністю $\beta = [B_1, \dots, B_s]$ і також репрезентують логарифмічний підпис G.

Перетворення 1. Нехай φ буде автоморфізмом G і $B_i = \varphi(A_i)$ для $i = 1, \dots, s$. Тоді β також є логарифмічним підписом. І якщо $a_1 \cdot a_2 \cdot \ \dots \ \cdot a_3$ є факторизацією елементів $g \in G$, $\varphi(a_1) \cdot \varphi(a_2) \cdot \ \dots \ \cdot \varphi(a_s)$ є факторизацією $\varphi(g)$ щодо β.

Перетворення 2. Нехай $g_0, \dots, g_s$ будуть елементами групи G і $B_i = g_{i-1}^{-1} A_i g_i$ для $i = 1, \dots, s$. Тоді послідовність β також є логарифмічним підписом G, що називають трансляцією α. Якщо $g_0 = g_s = e_G$, β – сендвіч α. Зазначимо, що, якщо $a_1 \cdot a_2 \cdot \ \dots \ \cdot a_s$ є факторизацією елементів $g \in G$, $g_0^{-1} a_1 g_1 \cdot \ \dots \ \cdot g_{s-1}^{-1} a_s g_s$ є факторизацією $g_0^{-1} g g_s$ щодо β.
Необхідно пам'ятати, що в абелевих групах блоки трансляції β від α мають вигляд $B_i = A_i + h_i$ для елементів $h_1, \dots,\ h_s$ групи G, а отже будь-яка факторизація елемента $g \in G$ щодо α миттєво дає факторизацію $g + \sum_{i=1}^{s} h_i$ щодо β.

Перетворення 3. Для $i = 1, \dots, s$ нехай π_i буде перестановкою в S_{r_i} і $B_i = \left[a_{i\pi_i(1)}, \dots, a_{i\pi_i(r_i)}\right]$ для $j = 1, \dots, r_i$, тобто елементи блоку B_i – перестановка елементів блоку A_i. Тоді β також є логарифмічним підписом. І, якщо $a_{1k_1} \cdot a_{2k_2} \cdot \ \dots \ \cdot a_{sk_s}$ – факторизація елемента $g \in G$, $a_{1\pi_1(k_1)} \cdot a_{2\pi_2(k_2)} \cdot \ \dots \ \cdot a_{s\pi_s(k_s)}$ – факторизація g щодо β.

Перетворення 4. Тепер нехай G буде абелевою групою, π – перестановкою в S_s і $B_i = A_{\pi(i)}$, тобто $b_{ij} = a_{\pi(i)_j}$. Тоді послідовність β є логарифмічним підписом групи G. Її також називають перетворенням блочної перестановки α. Зазначимо, що якщо $a_{1i_1} + a_{2i_2} + \ \dots \ + a_{si_s}$ є факторизацією елемента $g \in G$ щодо α, $a_{\pi(1)(i_1)} \cdot a_{\pi(2)(i_2)} \cdot \ \dots \ \cdot a_{\pi(s)(i_s)}$ є факторизацією g щодо β. Крім того в неабелевих групах β може не бути логарифмічним підписом.

Перетворення 5. Для певних $j \in \{1, \dots, s-1\}$ нехай $B_j = A_j \cdot A_{j+1} = \left[x \cdot y | x \in A_j, y \in A_{j+1}\right]$ та $B_i = A_i$ для $i = 1, \dots, s-1$ і $i \neq j, j+1$. Послідовність $\beta = \left[B_1, \dots, B_{s-1}\right]$ є логарифмічним підписом, одержаним з α в результаті перетворення – злиття двох блоків. І, якщо $a_1 \cdot a_2 \cdot \ \dots \ \cdot a_s$ є факторизацією $g \in G$ щодо α, $a_1 \cdot \ a_{j-1} \cdot \ a \cdot a_{j+2} \dots \cdot a_s$, де $a = a_j \cdot a_{j+1}$, є факторизацією g щодо β. Зворотну операцію називають перетворенням розподілу. Для кожного з п'яти наведених перетворень ми описали, як факторизація елемента щодо логарифмічного підпису негайно приводить до факторизації щодо перетвореного логарифмічного підпису. Якщо ми розглядаємо ці перетворення для родин логарифмічних підписів (α_n), то легко помітити, що перемикання між алгоритмами факторизації (α_n) і перетвореннями (α_n) виконується за поліноміальний час, якщо перетворення є відомим або ефективно обчислюваним. Це справедливо й для нормалізації.

Визначення 3. Нехай (α_n) і (β_n) є родинами логарифмічних підписів для груп G_n з параметрами r_n, s_n, μ_n для (α_n). Тоді ми стверджуємо, що (α_n) перетворюється на (β_n), якщо (β_n) можна обчислити з (α_n) за допомогою перетворень 1 – 5 за поліноміальний час для r_n, s_n, μ_n. Зазначимо, що в такому разі кількість перетворень між (α_n) і (β_n) є кінцевою. $T -$ множина (α_n), визначена як $T(\alpha_n) = \{(\beta_n) | \beta_n\ \Lambda\ (G_n)$ і α_n перетвориться для всіх $n\}$.

Приклад 3. Візьмемо логарифмічний підпис $\alpha_n = \left[[0, 2^{n-1}], \dots, [0,2], [0,1]\right]$ групи $\mathbb{Z}_{2^n}$ із прикладу 1. Для $n > 4$ нехай

$$\beta_n = \left[[1,\ 5], [0,\ 1,\ 2,\ 3], [0,\ 8], [0,\ 16], \dots, [0, 2^{n-1}]\right]$$

І $\gamma_n = \left[\left[0,\ 1, 2, 3, \dots, \ 2^n - 1\right]\right]$. Тоді $(\beta_n) \in T(\alpha_n)$ і $(\gamma_n) \notin T(\alpha_n)$.

Визначення 4. Нехай $g_0 = 1$ і $g_i = \left(\prod_{j=1}^{i} a_{j1}\right)^{-1}$ для $i = 1, \ldots, s$. За допомогою трансляції α в $g_0, \ldots, g_s$ ми одержуємо логарифмічний підпис β, у якому

$$b_{i1} = g_{i-1}^{-1} a_{i1} g_i = \left(\left(\prod_{j=1}^{i-1} a_{j1}\right)^{-1}\right)^{-1} a_{i1} \left(\prod_{j=1}^{i} a_{j1}\right)^{-1} = \left(\prod_{j=1}^{i} a_{j1}\right) \cdot \left(\prod_{j=1}^{i} a_{j1}\right)^{-1} = 1,$$

тобто перший елемент у кожному блоці є нейтральним. Ми вважаємо β нормалізацією α.

В абелевих групах ми можемо нормалізувати логарифмічний підпис, застосовуючи трансляцію $B_i = A_i - a_{i1}$. Тоді перший елемент кожного блоку дорівнюватиме $a_{i1} - a_{i1} = 0$.

Інші класи мають стандартне позначення. Нехай G буде кінцевою групою. Ми називаємо точно-поперечним логарифмічний підпис $\alpha \in \Lambda(G)$, якщо такий ланцюг підгруп виду

$$e_G = G_0 < G_1 < \ldots < G_{s-1} < G_s = G,$$

якщо A_i є поперечною групи G_{i-1} у G_i, тобто $G_{i-1}A_i = G_i$ та $|G_{i-1}||A_i| = |G_i|$. Зазначимо, що блок $A_1 = G_1$ є підгрупою G. Відповідний клас позначають через ε. Якщо α є сендвічем точно поперечної логарифмічної групи, α називають поперечною. Клас поперечних логарифмічних підписів позначимо через T_{LS}.

Усі інші логарифмічні підписи належать до NT_{LS} – класу не поперечних логарифмічних підписів. Ми описуємо два підкласи NT_{LS}. Якщо жодний із блоків не є підмножиною (нетривіальною) підгрупи G, логарифмічний підпис є елементом TNT_{LS} – класу абсолютно непоперечних логарифмічних підписів. Клас TA_{LS} повністю аперіодичних логарифмічних підписів містить усі логарифмічні підписи, що не мають навіть періодичного блоку, тобто об'єднані підмножиною G. Отже, маємо $TA_{LS} \subseteq TNT_{LS} \subseteq NT_{LS}$ і $\varepsilon \subseteq T_{LS}$.

Необхідно пам'ятати, що властивість простоти для точно поперечних логарифмічних підписів породжується в певних групах.

Визначення 5. Нехай (α_n) – сімейство точних логарифмічних підписів, а тестування на належність підгрупи до групи G_n проведено за поліноміальний час μ_n. Тоді (α_n) є простим.

Для простішого читання опустимо індекс n. Нехай

$$e_G = G_0 < G_1 < \ldots < G_{s-1} < G_s = G$$

буде ланцюгом підгруп відповідних логарифмічних підписів $\alpha = [A_1, \ldots, A_s]$. Нехай $g \in G$. Є саме одна факторизація g і вона містить $a_1, \ldots, a_s$, де

$$g \cdot a_s^{-1} \cdot \ldots \cdot a_{k+1}^{-1} = a_1 \cdot \ldots \cdot a_k \in G_k$$

для $k = 1, \ldots, s-1$ та $a_1 \cdot \ldots \cdot a_s = g$. Наступний простий алгоритм знаходить факторизацію g щодо α. Алгоритм є детермінованим. Буде потрібно $O(r_n \cdot s_n)$ раундів, у кожному з яких необхідні одне множення, одне звернення й одне тестування на належність до підгрупи, виконувані за поліноміальний час μ_n, s_n, r_n. Крім того, зрозуміло, якщо μ_n, s_n, r_n є поліноміальними в $\lceil log|G_n| \rceil$ та (α_n) є простим сімейством нормалізованих точно поперечних логарифмічних підписів, тоді існує ефективний тест на належність до підгрупи для всіх підгруп G_i групи G_n: $g \in G_i$, лише якщо у факторизації $g = a_1 \cdot \ldots \cdot a_s$ одержуємо результат $a_{i+1} = \ldots = a_s = e_G$.

Можливо, непоперечні й навіть абсолютно непоперечні логарифмічні підписи є гарними варіантами, щоб бути складними. Проте, в [8] було доведено, що досить легко побудувати прості сімейства логарифмічних підписів для симетричних груп, що чергуються, які належать до класу TNT_{LS}. Отже, класи не мають критеріїв, щоб розрізняти прості й складні логарифмічні підписи.

Якщо немає ефективного тесту на належність до підгрупи для груп G_n, у цих групах також будуть складні логарифмічні підписи, що є поперечними. Очевидно, що для груп, для яких є ефективний тест на належність до підгрупи, точно поперечні логарифмічні підписи є простими. Для формалізації цього зауваження використовуємо наступне визначення.

Визначення 6. Для $n \in \mathbb{N}$ нехай P_n буде булевим предикатом, що можна застосовувати до будь-якого логарифмічного підпису групи G_n. Ми стверджуємо, що P є властивістю, яка породжує простоту для G_n, якщо всі сімейства $(\alpha_n)_{n\in\mathbb{N}}$ логарифмічних підписів (G_n), для яких правильне твердження: $P(\alpha_n)$ правдиве для всіх $n \in \mathbb{N}$, де n є простим.

Приклад 4. Переглянемо знову приклад 1. Для $n \in \mathbb{N}$ нехай P_n =«кожним блоком форми $[0, 2^i]$ для $0 \leq i \leq n''$». Тоді ми стверджуємо в прикладі 4, що P є властивістю породження простоти для $\mathbb{Z}_{2^n}$.

У групах, для яких є ефективний тест на належність до підгрупи, належність до поперечних логарифмічних підписів також є властивістю породження простоти.

Висновки

Криптосистеми MST використовують як складні, так і прості логарифмічні підписи для побудови криптосистем відкритого ключа. Криптосистема MST_3 використовує прості логарифмічні підписи в елементарних абелевих 2-групах. Цікавою задачею є дослідження всіх неабелевих груп з великим порядком в їх функціональних полях та структури простих логарифмічних підписів які можуть бути використані в таких групах. З огляду на результати конкурсів з побудови постквантовых криптосистем важливою задачею є квантовий криптоаналіз реалізацій криптосистем MST_3 та пошук вразливостей повязаних з побудовою логарифмічних підписів та груп, що використовуються у якості платформ для криптосистем.

Список літератури:

1. Gonzáles Vasco M. I. On minimal length factorizations of _nite groups / M. I. Gonzáles Vasco, M. Rotteler, R. Steinwandt // Experimental Mathematics. 2003. Vol. 12 (1). P. 1–12.
2. Singhi N. Minimal logarithmic signatures for _nite groups of Lie type / N. Singhi, N. Singhi, S. Magliveras // Designs, Codes and Cryptography. 2010. Vol. 55 (2). P. 243–260.
3. Magliveras S. New approaches to designing public key cryptosystems using one-way functions and trap-doors in nite groups / S. Magliveras, D. Stinson, T. van Trung // Journal of Cryptology. 2002. Vol. 15. P. 285–297.
4. Lempken W. A public key cryptosystem based on non-abelian _nite groups / W. Lempken, T. van Trung, S.S. Magliveras, W. Wei // Journal of Cryptology. 2009. Vol. 22 (1). P. 62–74.
5. Goldreich O. Foundations of Cryptography: Basic Tools / O. Goldreich // Cambridge University Press. 2001.
6. Nuss A. On group based public key cryptography [Electronic resource] : Phd thesis. Access mode : http://nbn-resolving.de/urn:nbn:de:bsz:21-opus-63659.
7. Blackburn S. R. Cryptanalysis of the MST 3 public key cryptosystem / S. R. Blackburn, C. Cid, C. Mullan // Journal of Mathematical Cryptology. 2009. Vol. 3 (4). P. 321–338.
8. Bohli J. Weak keys in MST / J. Bohli, M. I. Gonzáles Vasco, C. J. M. Martínez, R. Steinwandt // Designs, Codes and Cryptography. 2005. Vol. 37 (3). P. 509–524.
9. Caranti A. The round functions of cryptosystem PGM generate the symmetric group / A. Caranti, F. D. Volta // Designs, Codes and Cryptography. 2006. Vol. 38 (1). P. 147–155.
10. Magliveras S. Algebraic Properties of Cryptosystem PGM / S. Magliveras, N. D. Memon // Journal of Cryptology. 1992. Vol. 5 (3). P. 167–183.
11. Khalimov G. MST_3 Cryptosystem Based on a Generalized Suzuki 2-Groups [Electronic resource] / G. Khalimov, Y. Kotukh, S. Khalimova. Access mode : http://ceur-ws.org/Vol-2711/paper1.pdf
12. Khalimov G., Kotukh Y., Khalimova S. MST3 cryptosystem based on the automorphism group of the hermitian function field' // IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019 – Proceedings, 2019, pp. 865 – 868.
13. Khalimov G., Kotukh Y., Khalimova S. Encryption scheme based on the automorphism group of the Ree function field // 2020 7th International Conference on Internet of Things: Systems, Management and Security, IOTSMS 2020, 2020, 9340192.

Надійшла до редколегії 03.03.2021

Відомості про авторів:

Котух Євген Володимирович – канд. техн. наук, доцент, кафедра комп'ютерних наук, Сумський державний університет, Україна, e-mail: yevgenkotukh@gmail.com

Сєвєрінов Олександр Васильович – канд. техн. наук, доцент, доцент кафедри Безпеки інформаційних технологій, Харківський національний університет радіоелектроніки, Харків, Україна, e-mail: oleksandr.sievierinov@nure.ua

Власов Андрій Володимирович – канд. техн. наук, ст. дослідник, ст. викладач кафедри Безпеки інформаційних технологій, Харківський національний університет радіоелектроніки, Харків, Україна, e-mail: andrii.vlasov@nure.ua

Козіна Лідія Сергіївна – аспірантка, кафедра інформаційних технологій та програмної інженерії НУ «Чернігівська Політехника»

Теницька Альона Олексіївна – студентка, факультет Електроніки та інформаційних технологій, Сумський державний університет, Суми, Україна, e-mail: tenickajaalena@gmail.com

Зарудна Катерина Олександрівна – студентка, факультет електроніки та інформаційних технологій, Сумський державний університет, Суми, Україна, e-mail: zarudnayakatyaa@gmail.com

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

УДК 621.373.826 **DOI:10.30837/rt.2021.2.205.10**

АЛЬ-СУДАНИ ХАЙДЕР АЛИ

ПРИНЦИПЫ ПОСТРОЕНИЯ ГИРОСКОПОВ НА БАЗЕ ФОТОННО-КРИСТАЛЛИЧЕСКИХ ВОЛОКОН С ФОТОННОЙ ЗАПРЕЩЕННОЙ ЗОНОЙ

Введение

До недавнего времени в инерциальных навигационных системах (ИНС) подвижных объектов применялись исключительно механические гироскопы (МГ), принцип работы которых основан на содержании оси тела вращения (ротора) в одном заданном положении инерциального пространства (СК) [1]. Однако стоимость механических гироскопов достаточно высока, поскольку для их работы требуются высокая точность изготовления формы ротора, обеспечение минимального трения подшипников и т.д. Но даже при выполнении этих требований МГ достаточно недолговечные и ненадежные приборы вследствие износа трущихся поверхностей, движущихся деталей в них. Поэтому со временем в механических гироскопах появляется значительная погрешность измерения углов, что, в конечном счете, влияет на надежность, а значит – на оборонные показатели объектов во время движения. Для этого нужны частые поверки приборов с механическими гироскопами, что требует дополнительных финансовых затрат.

Эти недостатки в значительной степени устраняются за счет применения оптических гироскопов (ОП), которые имеют следующие преимущества по сравнению с МГ:

- отсутствие подвижных деталей (статическая структура);
- устойчивость к механическим ускорениям;
- мгновенный запуск (не нужно дополнительное время для раскрутки ротора);
- простота и надежность конструкции;
- значительно большая чувствительность;
- высокая линейность характеристик;
- низкая мощность потребления.

Эффект Саньяка в волоконных гироскопах

Эффект Саньяка получил свое название в честь французского физика Жоржа Марка Мари Саньяка, который открыл его в 1913 году. С 1970-х годов эффект Саньяка нашел основное применение в навигации в качестве фундаментального принципа проектирования волоконно-оптических гироскопов (ВОГ) [2, 3]. Интерференция двух волн происходит из-за того, что волны накладываются, образуя одну волну с амплитудой, которая больше или меньше начальных волн. Для пучка света фотоны мешают друг другу, и по мере выявления волн, которые дают результат, появляется картина, иллюстрирующая различные свойства света. Результат меняется с частотой, фазой и амплитудой света. Интерферометр измеряет различия в интерференционной картине для определения некоторых свойств света. Сегодня ВОГ стали высокочувствительными детекторами, измеряющими вращательное движение в навигации [4, 5]. Саньяк обеспечил первую демонстрацию возможностей оптического эксперимента, способного указывать состояние вращения системы отсчета, делая измерения в рамках этой системы координат. На рис. 1 показана схема его интерферометра.

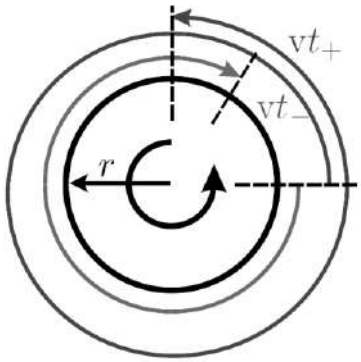


Рис. 1. Влияние вращения на встречные лучи

Физические принципы эффекта Саньяка подробно описаны в работе [6], и их можно разделить следующим образом. Эффект Саньяка может быть легко описан в классических терминах, если предположить, что скорость света равна s по отношению к статическому эфиру.

Эффект Саньяка [7] показывает, что два встречных световых луча, которые имеют различные временные интервалы для прохождения замкнутого пути на вращающемся диске, – это фазовый сдвиг встречных электромагнитных волн, который возникает во вращающемся кольцевом интерферометре (рис. 2). Источник света S выпускает две волны, которые распространяются в противоположных направлениях по траектории радиуса R.

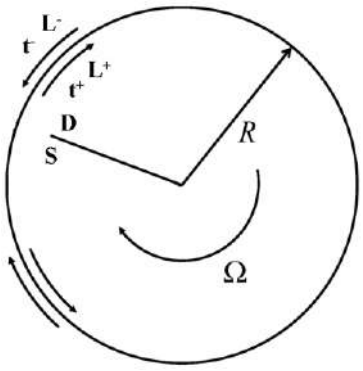


Рис. 2. Изменение времени прохода через вращения

Источник света и датчик D вращаются с угловой скоростью Ω, так что световой луч проходит в направлении вращения большее расстояние L + за большее время t^+. Встречный световой луч проходит более короткое расстояние L- за более короткое время t^-:

$$L^+ = 2\pi R + R\omega t^+, \tag{1}$$

$$L^- = 2\pi R - R\omega t^-, \tag{2}$$

Скорость света в среде

$$c_m = \frac{c}{n}, \tag{3}$$

где c_m – скорость света в среде; n – показатель преломления данной среды.

Свет дольше идет к исходной точке только потому, что приемник будет перемещаться вперед на некоторое расстояние.

Если n~1, c_m~ c, тогда:

$$t^+ = \frac{L^+}{c}, \tag{4}$$

$$t^{-} = \frac{L^{-}}{c}, \tag{5}$$

подстановка выражений (4) и (5) в выражения (1) и (2) дает:

$$L^{+} = \frac{2R\pi}{1-\frac{R\omega}{c}}, \tag{6}$$

$$L^{-} = \frac{2R\pi}{1+\frac{R\omega}{c}}, \tag{7}$$

из выражений (6) и (7) следует:

$$\Delta L = L^{+} - L^{-} = L^{-}, \tag{8}$$

$$= \frac{4R\pi\frac{R\omega}{c}}{1-\frac{R^2\omega^2}{c^2}}, \tag{9}$$

$$\approx \frac{4R^2\pi\omega}{c}, \tag{10}$$

$$\Delta t = \frac{\Delta L}{c} = \frac{4R^2\pi\omega}{c^2}, \tag{11}$$

если волны пройдут путь N раз:

$$\Delta L = \frac{4NR^2\pi\omega}{c} = \frac{4NA\omega}{c} = \frac{LD\omega}{c}, \tag{12}$$

$$\Delta t = \frac{4NR^2\pi\omega}{c^2} = \frac{4NA\omega}{c^2}, \text{`} \tag{13}$$

где A − площадь контура; D − диаметр контура; L − общий путь, который пройден.

Для волны частоты f (длина волны $\lambda=c/f$); $T=1/f$ – время, необходимое для изменения фазы на 2π.

С учетом эффекта Саньяка изменение фазы Φ за час Δt составит:

$$\phi = 2\pi\frac{\Delta t}{T} = 2\pi f\Delta t, \tag{14}$$

после подстановки Δt в (14) и (12) изменение фазы составит:

$$\phi = \frac{8\pi AN\omega}{\lambda c}. \tag{15}$$

Это доказывает, что вращение является частным случаем. Полная разность фаз между двумя встречными лучами в петле. Эффект Саньяка для вращения рассматривается как частный случай. Как известно, в экспериментах с движущейся средой типа Физо проявляется относительное движение между движением инфракрасного луча и средой, стеклом или водой. Этот эксперимент зависит от показателя отражения среды, а коэффициент сопротивления равен нулю, когда показатель преломления равен 1. В исследовании мы будем использовать фотонно-кристаллические волокна (с воздушным сердечником). Они не содержат

такой среды как вода или стекло, но содержат воздух. Можем сделать вывод об обобщении эффекта Саньяка, что прохождение в волоконной петле или в волноводе протяженностью Δl со скоростью *v* добавляет разность фаз в общей разности фаз между двумя встречными пучками в петле. Вклад Δφ не зависит от показателя преломления волновода, а путь длины движения может быть либо линейным или круговым.

Способ построения и работы волоконных оптических гироскопов с фотонно-кристаллическими волокнами

Фотонно-кристаллические волокна [8], также называемые микроструктурными или дырявыми волокнами, представляют собой оптические волноводы нового типа. В ФКВ излучение может передаваться как пустотелым, так и сплошным твердым сердечником, как показано на рис. 3, состоящими из массива цилиндрических воздушных каналов, проходящих вдоль оси волокна, и охватывается микроструктурированной оболочкой. Такая микроструктура обычно изготавливается путем вытяжки из пучка капиллярных трубок и твердых кварцевых стержней. Наряду с обычными волноводные режимы обеспечиваются полным внутренним отражением. Фотонно-кристаллические волокна при определенных условиях могут поддерживать управляемые моды электромагнитного излучения через высокую отражательную способность их оболочки в пределах фотонных полосных зазоров или областей с низкой плотностью фотонных состояний, а также через антирезонансный механизм волновода [9].

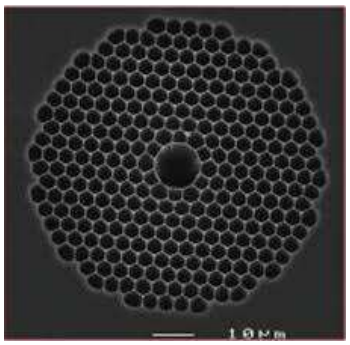


Рис. 3. Изображение поперечного разреза фотонно-кристаллического волокна с фотонной запрещенной зоной (фкв – фзз)

В волоконно-оптическом гироскопе оптическое волокно используется в качестве среды распространения света. ВОГ основаны на эффекте Саньяка. Эффект Саньяка создает разницу оптических фаз Δφ между двумя противоположно направленными волнами в волоконной катушке, которая вращается (оптический путь). Они широко используются в промышленности, там, где приемлемы их возможности по динамическому диапазону и линейности.

Эффект Саньяка возникает во всем контуре, который вращается с угловой скоростью вращения Ω, расстояние между двумя точками A и B изменяется для противоположно движущихся пучков, как показано на рис. 4.

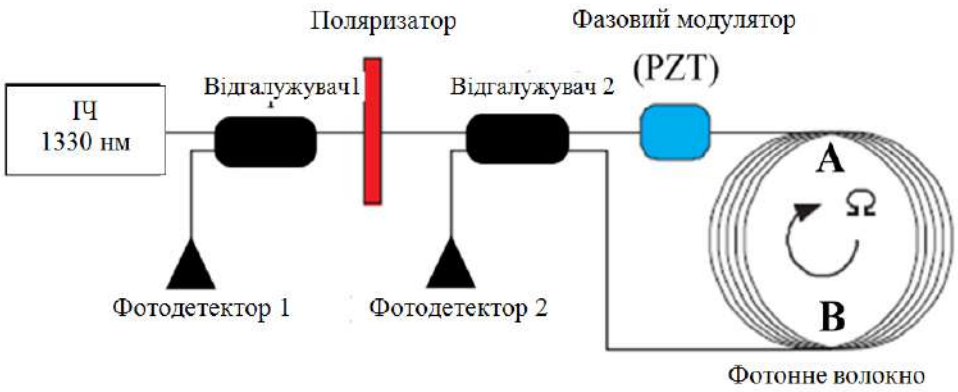


Рис. 4. Схема волоконно-оптического гироскопа

Реакция оптической мощности

$$P = \frac{I_0}{2}(1+\cos\Delta\phi), \tag{16}$$

где $\Delta\phi$ – разность фаз [10].

Разность фаз максимальна при нуле. Чтобы получить высокую чувствительность, этот сигнал должен быть смещен по отношению к рабочей точке с ненулевым наклоном:

$$P = \frac{I_0}{2}[1+\cos(\Delta\phi_s + \Delta\phi_b)], \tag{17}$$

где $\Delta\phi_s$ – фазовый сдвиг.

Фазовый сдвиг должен быть таким же стабильным, как и ожидаемая чувствительность, что значительно лучше, чем 1рад [10].

В волне, движущейся с точки А в точку В в направлении, совпадающем с прямым вращением контура, увеличивается расстояние, за время *dt* точка В смещается на угол $d\varphi = \Omega dt$. Увеличивается время движения светового пучка, равное *dt,* потому что в каждый момент луч направлен вдоль касательной к контуру под одной проекцией с касательной линией скорости:

$$\vec{u} = \vec{u}\times\cos a = \mathrm{W}\times r\times\cos a. \tag{18}$$

Проблема дрейфа фазового сдвига полностью решается с помощью обратного фазового модулятора, расположенного на одном конце катушки, который действует как линия задержки. Волны приобретают сдвиги фаз $\phi_m(t)$, проходя через модулятор в момент времени *t*. Встречные волны проходят через модулятор в разное время; пока CW волна проходит через модулятор в момент времени *t*, CCW волна проходит через модулятор во времени $t-\tau$, где τ – время прохождения цикла $\frac{nl}{c}$. Это приводит к модуляции смещения $\Delta\phi_m$(t) разности фаз:

$$\Delta\phi_m(t) = \phi_m(t) - \phi_m(t-\tau), \tag{19}$$

Таким образом, сигнал помехи

$$P = \frac{I_0}{2}\{1+\cos[(\Delta\phi_s(t) + \Delta\phi_m(t)]\}. \tag{20}$$

Инвариантность скорости света, которая кажется продолжением и сжатием путей для противоположно направленных лучей, можно считать эквивалентной расширению и сужению временных интервалов. Существуют различные механизмы фазовой модуляции. Одним из них является электрооптический модулятор. Пьезоэлектрический преобразователь (PZT) растягивает волокно при приложении внешнего напряжения, удлиняет оптический путь. Инжектированная фаза пропорциональна приложенному напряжению для обоих механизмов. Сигнал на частоты модуляции составляет

$$P(t)\big|_{\omega} = -I_0 J_1(\alpha)\cdot\sin(\Delta\phi_s)\cdot\cos[\omega_m t - \delta]. \qquad (21)$$

Это обобщение показывает принцип проектирования нового волоконно-оптического линейного датчика движения, который имеет высокую чувствительность и высокую стабильность. Датчик состоит из 1 метра фотонно-кристаллического волокна (HC-1330 нм NKT Photonics), которые намотаны вокруг цилиндрической оправки диаметром 8 см. При работе оптического гироскопа волокно, используемое в катушке оптического датчика, является фактически одномодовым при длине волны 1330 нм.

Источник ИК-излучения подключается к различным волокнам, в идеале с малым обратным отражением и малыми потерями.

Сигнал CCW фотодиода состоит из основного сигнала от луча CCW и сигнала, вызванного обратным рассеянием луча CW и отпечатками. Один луч модулируется синусоидально на частоте (36 кГц) с помощью модулятора (ПЗКТМ), которой размещен на волокне, идет к катушке датчика, тогда как другой вход (CW) не модулируется. Общий сигнал на каждом фотодиоде включает модулированный компонент.

Амплитуда модуляции подобрана так, чтобы максимально подавить оптическую частоту, которая является несущей, чтобы сдвигать большую часть низкочастотного дрейфа на кратность сигнала модуляции.

Для начала мы оценили производительность обычного оптического гироскопа с датчиками из фотонно-кристаллического волокна и катушкой из обычного волокна.

Поэтому расчетный дрейф скорости этого волоконно-оптического гироскопа составляет 0,014 °/c, что соответствует минимальному сдвигу частоты вращения ~ 1,8 Гц с полосой детектирования 1 Гц. Необходимо учитывать масштабируемый коэффициент и точность расчетов.

После коррекции различий в масштабируемом коэффициенте и точности измеренный сдвиг между пиками в течение двухминутного интервала составлял 0,027 ° / с, что близко к значению, опубликованном в [11].

Хотя кольцевые резонаторы на базе ФКВ с воздушным каналом, использующим воздушное соединение с основной оптикой системы, показали очень низкий процент потерь, этот метод является непростым. Таким образом, длина пути, который пройден лучом, равна Dl = R.cos (θ + ωt).

Рассуждая аналогично, для встречного луча будет уменьшаться участок пути D2 = R.sin (θ + ωt) (рис. 5).

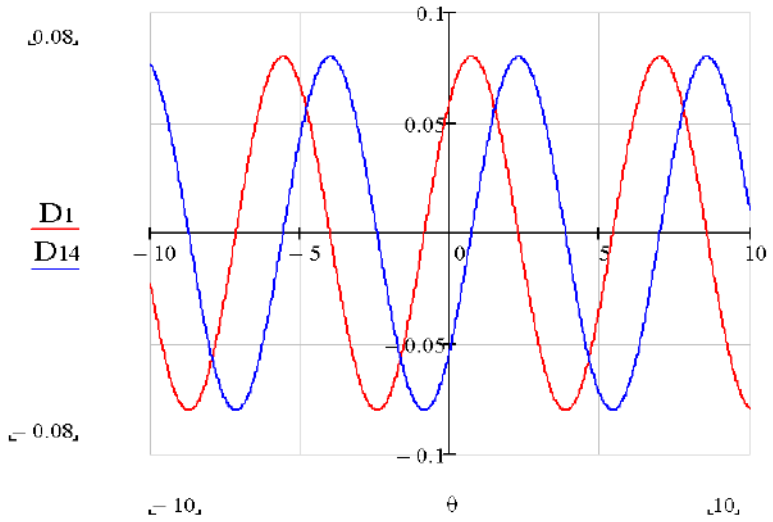


Рис. 5. Длина пути лучей (D_1,D_2)

При неподвижном гироскопе мы регистрируем выходной сигнал вращения для измерения шума. Шум эквивалентной полосы может быть выражен как

$$BW_e = \frac{5}{64.\tau}. \tag{22}$$

Здесь τ – изменение постоянной времени фотодиода (от 100 до 3 мс).

Тогда измеренный дрейф при τ = 3мс даст BW_e = 26.04166 Гц, и при τ = 100 мс составит BW_e = 0,78125 Гц. Диапазон случайных отклонений:

$$R_w = \frac{\sigma}{BW_e^{\frac{1}{2}}}. \tag{23}$$

Для гироскопа с фотонно-кристаллическим волокном (с пустым каналом) R_w=0,055^0/c$^{1/2}$.

Поэтому максимальный дрейф в 120 с составит ~ 1 °/с. Модуляция уменьшает дрейф в 120 с на ~ 17,5 дБ. В 3600 с максимальный дрейф составит 2,3 °/с, и σ = 0,5/с. Измерение выходного сигнала подобного гироскопа показано на рис. 6. При t = 0 скорость вращения ±2,3°/с.

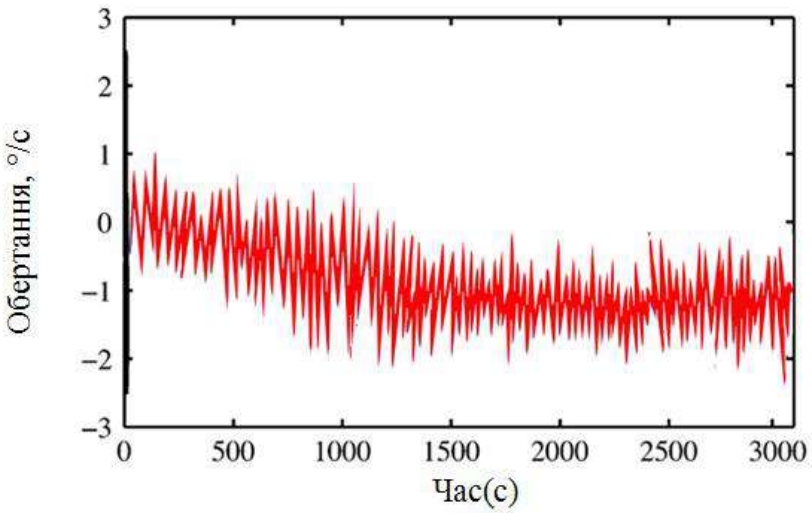


Рис. 6. Максимальный дрейф

Выводы

Гироскопы с фотонно-кристаллическими волокнами являются типом волоконной оптики, содержащей множество новых и улучшенных функций. Ожидается, что благодаря уникальной геометрической структуре использование волокна с воздушной сердцевиной в волоконном гироскопе значительно уменьшит фазовый дрейф и шум, связанные с эффектом Керра, эффектом Фарадея и тепловыми эффектами в чувствительном волокне. Причина в том, что в волокне с воздушным сердечником оптический режим в основном ограничивается воздушным сердечником, тогда как в обычном волокне сигнал полностью перемещается через кремний. Приведенные результаты позволяют ожидать, что использование фотонно-кристаллического волокна вместо обычного обеспечит большую долговременную стабильность, низкую стоимость, более простую конструкцию и меньший уровень шума, что приведет к созданию недорогих инерционных навигационных оптических гироскопов.

Список литературы:

1. Lefevre H. Fiber Optic Gyroscope / M.A. Boston // Artech House. 1993. P. 66−67.
2. Vali V. / R. Shorthill // Appl. Opt. 1977. V.16. P. 290 .
3. W. Leeb, G. Schiffner, and E. Scheiterer. Appl. Opt. 18, 1293 (1979).
4. H. Lefevre. The Fiber−Optic Gyroscope (Artech House, Boston, 1993).
5. Burns, W. K. Optical Fiber Rotation Sensing //Academic Press, Boston, 1994.
6. Post. E.J. Sagnac effect // Rev. Mod. Phys. 1967. Vol. 39. P. 475–493.
7. Sagnac G. / C. R. Acad. // Sci. Paris. 1931. V. 157. P.708.
8. Russell P.S. J. Photonic Crystal Fibres // Science. 2003. V.299. P 358–362.
9. Russell P.S.J. Lightwave Technol. 2006. 24. P.4729.
10. Yablonovitch E. Photonic Band Structure: The Face−Centered−Cubic Case Employing Nonspherical Atoms / T. J. Gmitter, K. M. Leung // Phys. 1991. Rev. Lett. 67. P. 2295−2298.
11. Folkenberg, J. R. Polarization maintaining large mode area photonic crystal fiber / M. D. Nielsen, N. A. Mortensen, C. Jakobsen, H. R. Simonsen // Optics Express, 2004. Vol. 12, №.5. P.956–960. doi: 10.1364/opex.12.000956.

Поступила в редколлегию 07.04.2021

Сведения об авторе:

Аль-Судани Хайдер Али – Харьковский национальный университет радиоэлектроники, аспирант кафедры физических основ электронной техники (ФОЭТ), факультет электронной и биомедицинской инженерии, Украина; e-mail: hadr_2005@yahoo.com

УДК 621.382.232 **DOI:10.30837/rt.2021.2.205.11**

К.С. ЯЦУН

МОДИФІКАЦІЯ АКТИВНОЇ ОБЛАСТІ РЕЗОНАНСНО-ТУНЕЛЬНОГО ДІОДУ

Вступ

У загальному випадку резонансно-тунельний діод являє собою періодичну структуру, яка складається з послідовно розташованих квантових колодязів, розділених потенційними бар'єрами, з електричними контактами до двох крайніх протилежних областей.

Явище резонансного тунелювання було вперше описано в 1958 р. японським дослідником Л. Есакі [1]. Однак експериментальні резонансно-тунельні діоди і транзистори з'явилися лише на початку 90-х років XX століття. Резонансно-тунельний діод являє собою складну періодичну структуру, розміри деяких областей якої складають кілька нанометрів.

Принцип дії резонансно тунельного діоду полягає у тому, що струм досягає максимального значення, коли при поданій напрузі енергія електронів дорівнює енергії дискретного рівня у квантово-обмеженій області. При більш високих або менших напругах енергія електронів стає більшою чи меншою ніж енергія дискретного рівня, і прозорість бар'єру для електронів зменшиться. При цьому струм також зменшиться.

Додавання керуючого електрода до резонансно-тунельного діода перетворює його в резонансно-тунельний транзистор (РТТ) і розширює можливості його застосування. Такі транзистори мають частоти перемикання порядку 10^{12} Гц, що в 100 – 1000 разів вище, ніж у найкращих кремнієвих транзисторів з сучасних інтегральних мікросхем. З точки зору практичного використання найбільш привабливими характеристиками резонансно-тунельних діодів є їх надзвичайно високі швидкості перемикання.

Резонансно-тунельні діоди і транзистори застосовуються як в аналогових, так і в цифрових інтегральних мікросхемах як елементи, що мають вольт-амперну характеристику з ділянкою негативного диференціального опору.

При цьому, найбільш важливими особливостями розглянутих наноелектронних приладів на тунельному ефекті є їх розширені, в порівнянні з традиційними приладами, функціональні можливості. Надзвичайні особливості резонансно-тунельних структур пов'язані з унікальністю їх вольт-амперних характеристик і високою швидкодією. Завдяки їм РТД та РТТ відіграють все більш важливу роль в розробці надшвидкодіючих інтегральних мікросхем з надвисоким ступенем інтеграції.

Резонансно-тунельні структури у сучасній електроніці та мікроелектроніці

Функціонування РТД базується на ефекті резонансного тунелювання носіїв заряду через послідовно розташовані напівпрозорі потенційні бар'єри, розділені квантовими ямами. РТД спочатку були випробувані як детектори випромінювань терагерцового діапазону, а потім як високочастотні генератори [2]. В роботі повідомляється про досягнення частоти генерації 712 ГГц. Був розроблений і створений РТД, максимальна частота генерації якого оцінюється в 2 ТГц. Експериментально спостерігалося перемикання РТД з пікового стану в долині за 1,5 пс.

У даний час ведуться активні роботи по створенню схем, що містять РТД, для таких функціональних пристроїв дециметрового, сантиметрового і міліметрового діапазонів, як генератори фіксованої частоти, частотні модулятори і змішувачі. Необхідно відзначити, що не тільки висока гранична частота, а й інші характеристики РТД, такі як симетрія його N-подібної ВАХ відносно початку координат і знижений рівень шумів, можуть бути практично важливими для його застосування в уже освоєних діапазонах частот.

Великі переваги обіцяє використання РТД в цифрових інтегральних схемах в якості нелінійного навантаження для польових транзисторів. РТД перемикаються швидше ніж транзистори з високою рухливістю електронів (ТВРЕ), тобто не обмежують швидкодії активного

приладу), і вертикально інтегруються в стік активного транзистора, не займаючи при цьому додаткового місця [2, 3].

Крім того, останнім часом з'явилася велика кількість теоретичних і експериментальних робіт, в яких пропонується використовувати резонансно-тунельні структури (РТС) в якості логічних елементів. Монолітний синтез РТС з транзисторними структурами відкриває великі можливості у створенні приладів зі складними логічними функціями. Подібна інтеграція в різних варіантах була здійснена з польовим транзистором і біполярним транзистором. Монолітна (площинна або вертикальна) інтеграція декількох РТС дозволяє формувати компактні осередки багатозначної логіки і багатозначної пам'яті. Такі компактні осередки вже зараз можуть конкурувати зі звичайними транзисторними осередками. У роботі [3] інтегровано шість РТД і транзистор з високою рухливістю електронів на основі InGaAs. В результаті створено аналого-цифровий перетворювач і чотирьохзначний інвертор.

Властива РТД (з навантаженням) бістабільність дозволяє створювати тригерні схеми без використання зворотного зв'язку. Наприклад, описана в [3] ТВРЕ/РТД схема має площу в шість разів, а енергоспоживання в три рази менше, ніж аналогічні ТВРЕ-схеми на частоті 25 ГГц. Завдяки меншому числу елементів скорочуються також затримки на міжз′єднання.

За ТВРЕ/РТД технологією створено: осередки статичної оперативної пам'яті з нановатним споживанням, десятирозрядний зсувний регістр, що працює на частоті 2,5 ГГц, тактовий генератор на 6,5 ГГц, чотирьохбітний АЦП на 2 ГГц та деякі інші пристрої. Ці приклади підтверджують високу перспективність використання РТД в електроніці.

При більш детальному розгляді ефекту проходження електронів через потенційні бар'єри, створювані всією сукупністю заряджених частинок при контакті двох, наприклад напівпровідникових матеріалів, було встановлено кілька фактів. Виявилося, що в структурі з надзвичайно малими розмірами властивості тунельного ефекту залежать від енергії електронів усередині самої структури. Внаслідок чого в наноструктурі з двома потенційними бар'єрами спостерігається різке зростання струму, який протікає через неї, тільки при однаковій (у межах теплового розширення енергетичного стану) енергії електронів в електроді, що поставляє електрони, і окремого вільного енергетичного рівня в наноструктурі. Це явище отримало назву «резонансне тунелювання».

Прилади з РТС (і спеціальним легуванням контактних шарів) мають властивість зберігання інформації (заряду) за кімнатної температури і нульовій доданій напрузі. Дане явище полягає в наявності двох стійких станів системи, які відрізняються профілем зони провідності через відмінності в розподілі заряду і, відповідно, тунельною прозорістю.

Кожному стану відповідає своя гілка ВАХ, перехід між якими відбувається при підвищенні напруги. Такі структури можуть бути використані для виготовлення швидкодіючих енергонезалежних запам'ятовуючих пристроїв. Довгий час серйозною перешкодою для застосування РТД залишалася низька відтворюваність характеристик РТД, яку відносили на рахунок недосконалості технології молекулярно-променевої епітаксії (МПЕ). В останні роки необхідна висока відтворюваність була досягнута. Зараз основною причиною, за якою не відбувається масове впровадження РТД в електроніку, є складність і дорожнеча технології, а також відсутність досвіду розробки схем з РТД у схемотехніці. Для спрощення технології виробництва, перш за все, бажано створити площинні РТС, і недавно стався істотний прорив в даному напрямку [4].

Інша проблема полягає в недостатньо високих статичних характеристиках одержуваних РТД, внаслідок чого вони не у всіх застосуваннях можуть конкурувати з традиційною елементною базою ІС. Найважливішими характеристиками статичної (низькочастотної) ВАХ РТД, що має N-подібний вигляд, є густина пікового струму Jp і відношення густини струму в піку Jp до густини струму в долині (відношення пік-долина) – Jv. Для використання РТД як компонента НВЧ пристроїв визначальним параметром є Jp, який задає частоту і потужність генерації. В даному випадку необхідно збільшення густини пікового струму. З наявних літературних даних інтерес представляють РТД з Jp> 10 А/см.

При використанні РТД в цифрових інтегральних схемах густина пікового струму визначається конкретним проектованим пристроєм. Також густина пікового струму залежить від розмірів елементів, що виготовляються для необхідних потужних режимів. Принциповим в даному випадку є зменшення долинного струму, оскільки він визначає споживану потужність елементу в стані логічного нуля. Мінімізація Iv автоматично збільшує відношення пік / долина.

При цьому вирішується проблема надійності реєстрації логічних рівнів в елементах ІС. Зростання Iv також сприяє поліпшенню інших статичних і динамічних характеристик проектованих пристроїв. Тому основне завдання в області створення приладів з резонансним тунелюванням полягає в отриманні відтвореним чином відносно великих пікових густин струму (10^3 – 10^5 А/см^2) і відносин пік / долина більше 10 при 300 К. Поряд з експериментальними дослідженнями ефекту резонансного тунелювання виникає необхідність в моделях, які дозволили б адекватно описати фізичні процеси, що протікають в структурах і аналізувати залежність електрофізичних характеристик приладів від конструктивно-технологічних параметрів. Такі моделі дозволяють визначити необхідні розміри, підібрати відповідні матеріали на етапі розробки РТД приладів, спрогнозувати їх підсумкові показники. Фізичні моделі спираються на експериментальні дані і, звичайно, оформляються у вигляді комп'ютерних програм-симуляторів.

Повний опис поведінки РТД можливий тільки мовою квантового кінетичного рівняння (ККР), тому що необхідно враховувати процеси розсіювання. Однак надзвичайно високі вимоги симуляторів на основі ККР до обчислювальної потужності, роблять актуальним використання більш простих моделей. Значна їх кількість заснована на розв′язанні стаціонарного рівняння Шредінгера в наближенні однозонного методу ефективної маси.

Моделювання та оптимізація резонансно-туннельного діоду

Тунелювання грає важливу роль у багатьох напівпровідникових діодах. РТД. Важливою характеристикою резонансно-тунельного діоду на основі двохбар'єрної структури з квантовою ямою є наявність падаючої ділянки на ВАХ. Іншими словами поява негативного диференційного опору або негативної диференційної провідності. Якісне пояснення ВАХ базується на енергетичній структурі квантової ями. Вона являє собою дискретну систему енергетичних рівнів розмірного квантування. У загальному випадку, положення квантово-розмірних рівнів та відповідних їм хвильових функцій у ямі можливо розрахувати, розв′язавши стаціонарне рівняння Шредінгера. Хвильові функції частинок, локалізованих у ямі не дорівнюють нулю на границях ями, а плавно спадають у глиб бар'єрів. Таким чином, можливе тунелювання крізь бар'єр.

За відсутності напруги, струм через структуру дорівнює нулю. При додаванні електричного поля потенційний бар'єр для електронів зменшується, і вони можуть тунелювати з емітеру до колектора. Таким чином, буде заповнюватися нижній рівень у ямі, струм буде повільно зростати. При подальшому збільшенні напруги, рівні у ямі рухаються униз по шкалі енергії відносно рівня Фермі. При певній напрузі рівень Фермі попаде у резонанс із підзоною розмірного квантування. Тоді електрони можуть резонансним чином тунелювати у яму, потім вийти з неї, протунелювавши крізь другий бар'єр. Тунельний струм стрімко зросте. При подальшому підвищенні напруги тунельний струм різко падає, утворюючи, таким чином, область з негативним диференційнім опором. Далі, з підвищенням напруги, при співпадінні рівня Фермі з наступним рівнем розмірного квантування у ямі ефект повторюються. Таким чином, можна спостерігати осциляції тунельного струму. Відстань між максимумами буде пропорційна відстані між рівнями енергії частинок в КЯ.

Крайні шари резонансно-тунельного діоду є сильно легованими, і до них приєднуються омічні контакти. У результаті розвитку технології молекулярно променевої епітаксії такі структури стало можливо створювати на основі гетеропереходів. Висота бар'єрів у таких

структурах змінюється при збільшенні змісту Al від 0,25 еВ при x=0,3 до 1,35 еВ при x=1. Типові ширини бар'єрів і ями мають значення 5 – 10 нм.

Оскільки бар'єри мають кінцеву висоту, то електрони, неповністю локалізовані у ямі, та квазірівні енергії у ямі мають кінцеву ширину $\Delta E \approx h/\tau_n$ [5], де час релаксації

$$\tau_n = \frac{1}{\nu_n} \times \frac{4}{|D_1|^2 + |D_2|^2}. \quad (1)$$

Частота класичного руху у ямі $\nu_n = hRqk_n/mL$

D_1, D_2 – амплітуди проходження крізь «перший» 1 та «другий» 2 бар'єри, $|D_1|^2$, $|D_2|^2$ – коефіцієнти проходження крізь 1 та 2 бар'єри, k_n – хвильовий вектор, який відповідає квазірівню енергії чатинки $E_n = h(k_n)^2/2m$ у ямі.

Для прямокутних бар'єрів коефіцієнт проходження крізь всю структуру поряд з енергією квазірівня E_n приблизно можна представити у вигляді

$$|D|^2 = \frac{4 \cdot |D_1 D_2|^2}{\left(|D_1|^2 + |D_2|^2\right)} \times \frac{h^2/\tau_n^2}{|E - E_n|^2 + h^2/\tau_n^2}. \quad (2)$$

Максимальне значення $|D|^2=1$ буде при $E=E_n$ та $|D_1|=|D_2|$ згідно з формулою (2).

При накладанні зовнішнього електричного поля змінюється форма і висота бар'єрів так, що амплітуди проходження D_1, D_2 стають різними, тому загальний коефіцієнт проходження $|D|^2$ буде зменшуватися, згідно формули (2).

Змінюючи висоту або ширину одного або двох бар'єрів можливо отримати, що при заданому значенні напруги коефіцієнт проходження знову стане максимальним. Таким чином, збільшення висоти другого бар'єру фактично повертає рівність коефіцієнтів проходження D_1 та D_2, що приводить до результуючого збільшення коефіцієнта проходження усієї структури згідно (2).

Величина і ширина піку залежності коефіцієнта проходження від енергії визначають вид ВАХ. При низьких температурах, коли $kT \ll \mu$, де μ – рівень Фермі, струм через діод можна записати [6]:

$$I = \frac{q \cdot S \cdot m}{2 \cdot \pi^2 \cdot h^3} \cdot \int_0^{\mu} (\mu - E) \cdot |D(E)|^2 \, dE. \quad (3)$$

Приблизно, при $|D_1|=|D_2|$ з (1) та (2) випливає [5]:

$$I = \frac{q \cdot S \cdot m}{2 \cdot \pi^2 \cdot h^3} \cdot \int_0^{\mu} (\mu - E) \cdot \frac{\left(h^2/\tau_n^2\right) dE}{\left(E - E_n - qV/2\right)^2 + h^2/\tau_n^2}. \quad (4)$$

Основними параметрами ВАХ, що визначають робочі характеристики резонансно-тунельного діоду, є відношення максимального та мінімального (після падаючої ділянки) струму і максимальне значення негативної диференційної провідності – g_{max}.

Інтеграл (4) береться аналітично, звідки можна отримати наступні оцінки:

$$I_{\max} = \frac{emS\mu}{2\pi h^2 \tau_n}, \quad (5)$$

$$g_{\max} = \frac{q^2 mS\mu}{4\pi^2 h^3}. \quad (6)$$

Максимальний струм, як слідує з виразу (5), пропорційний ширині квазірівня ($\sim 1/\tau_n$), а g_{max} не залежить від τ_n. Ці висновки отримані при наближені $|D_1|=|D_2|$, тому реальні значення можуть бути більш складнішим чином залежними від параметрів структури.

При проведенні математичного моделювання були отримані графіки ВАХ резонансно-тунельного діоду, зображені на рис. 1.

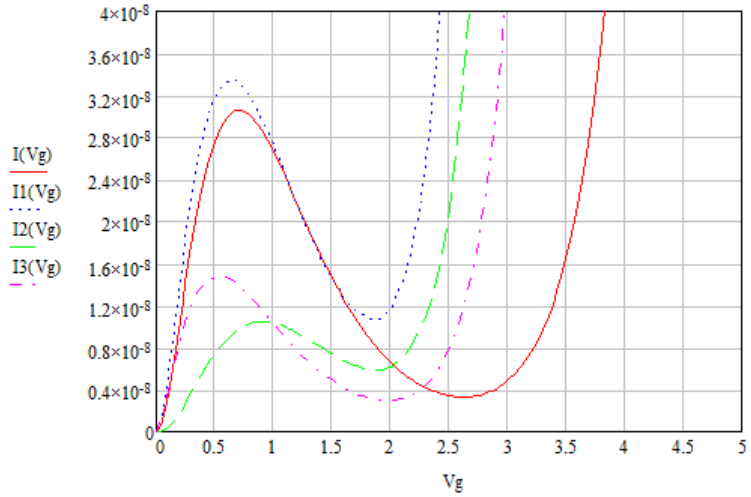


Рис. 1. Вольт-амперні характеристики резонансно-тунельного діоду отримані за результатами математичного моделювання

Висновки

У статті розглянуто структуру та принцип дії резонансно-тунельного діоду. Розраховано коефіцієнти прозорості і відбиття РТС та ВАХ РТД. За допомогою цих розрахунків побудовано графіки ВАХ при різній товщини бар'єру та ширині квантової ями.

Можна зробити висновок, що модифікація активною області резонансно-тунельного діоду можлива при збільшені або зменшені товщини шарів матеріалу а також при зміні складу шарів. У результаті цього змінюється товщина бар'єрів та квантової ями. При зменшені товщини ями або бар'єру електронам потрібно менше енергії для тунелювання і на ВАХ можна спостерігати більше піків струму при меншій напрузі. Це дає перспективи для використання РТД у приладах з низькою напругою живлення.

Список літератури:

1. Chang L.L., Esaki L., Tsu R. Applied Physics Letters. Maryland, 1974. Vol 24. 593 p.

2. Поздняков Д.В. Расчет вольт-амперных характеристик симметричных двухбарьерных резонансно-туннельных структур на основе арсенида галлия с учетом процессов разрушения когерентности электронных волн в квантовой яме // Физика и техника полупроводников. Москва : Наука, 2004. Т.38. Вып. 9. С. 1097–1100.

3. Абрамов И.И. Исследование двухбарьерной резонансно-туннельной структур на основе GaAs/AlAs с использованием комбинированной двухзонной модели // Доклады Бгуир, 2004. Вып.4. С. 42–46.

4. Врубель М.М. О влиянии ширины спейсерных слоев на размеры области бистабильности в вольт-амперных характеристиках двухбарьерных туннельных резонансных диодов // Письма в ЖТФ, 1997. Т. 23. Вып. 21. С. 12–16.

5. Иогансен Л.В. Журнал экспериментальной и теоретической физики. Москва :Наука, 1963. Т.45. Вып. 2. 2207 с.

6. Агарев В.Н. Моделирование резонансного туннелирования в полупроводниковых наноструктурах. ННГУ Фонд образовательных электронных ресурсов, 2008.

Надійшла до редколегії 06.02.2020

Відомості про автора:

Яцун Кирило Сергійович – Харківський національний університет радіоелектроніки, аспірант кафедри мікроелектроніки, електронних приладів та пристроїв (МЕЕПП), факультет електронної та біомедичної інженерії; Україна; e-mail: deadwoldi@gmail.com

УДК 621.373.072.9 **DOI:10.30837/rt.2021.2.205.12**

В.В. РАПИН, д-р техн. наук

ПОГРЕШНОСТЬ МЕТОДОВ МАЛОГО ПАРАМЕТРА ПРИ РЕШЕНИИ УКОРОЧЕННЫХ УРАВНЕНИЙ СИНХРОНИЗИРОВАННОГО АВТОГЕНЕРАТОРА

Введение

Развитие современных систем и устройств радиотехники, связи, радиолокации, навигации и информационно-измерительных комплексов невозможно без широкого использования автоколебательных систем, которые находят многочисленное и разнообразное применение, часто определяя предельные возможности по наиболее важным параметрам. Это обусловлено их способностью выполнять различные функции, такие как усиление и демодуляция амплитудно-модулированных, фазо-модулированных и частотно-модулированных сигналов, умножение и деление частоты, фильтрацию сигналов и различного рода преобразования, например малых изменений частоты в фазовый сдвиг [1 – 9]. Введение фазовой обратной связи в синхронизированных автогенераторах позволило реализовать потенциальные возможности таких устройств, что привело не только к улучшению известных характеристик, например к сокращению длительности переходных процессов, но и появлению совершенно новых свойств в системах синхронизированных автогенераторов [10 – 12]. Появление дифференциальных автогенераторов еще больше расширило сферу применения автоколебательных устройств [13, 14]. Однако их исследование представляет довольно сложную задачу с математической точки зрения. Появление в последнее время таких методов как метод квазималого параметра и комбинированного метода малого параметра позволило существенно продвинуться в этом направлении [15]. Однако еще недостаточно изучены особенности этих методов. Таким образом, целью статьи является исследование погрешностей этих методов и особенности их применения.

Модель синхронизированного автогенератора

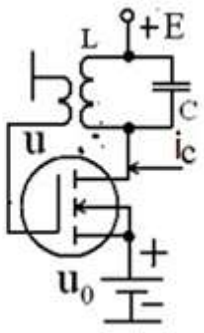


Рис. 1. Схема автогенератора

Рассмотрим, для определенности, синхронизированный на основном тоне одноконтурный LC автогенератор, представленный на рис. 1, для исследования которого использовались указанные методы. Модель автогенератора получена при традиционных упрощающих предположениях: добротность контура автогенератора Q велика, смещение u_0 фиксированное, транзистор является безынерционным элементом с большим входным сопротивлением. Его нелинейная характеристика аппроксимируется полиномом четвертой степени $i = a_0 + a_1 u_y + a_2 u_y^2 + a_3 u_y^3 + a_4 u_y^4$, где u_y управляющее напряжение $u_y = u + u_0,$. В этом случае $u = A\cos(\omega_c t + \varphi)$ напряжение на затворе транзистора, сигналом синхронизации является ток $i_c = I_c \cos(\omega_c t + \varphi_c)$. Используя законы Кирхгофа, дифференциальное уравнение автогенератора можно представить в виде

$$\frac{d^2u}{d\tau^2} - \varepsilon(1 - 2\beta u - 3\gamma u^2 - 4\delta' u^3)\frac{du}{d\tau} + \frac{\omega_0^2}{\omega_c^2}u = R\delta K\frac{di_c}{d\tau}, \qquad (1)$$

где $\tau = \omega_c t$, $\varepsilon = \delta\alpha$ – малый параметр, $\alpha = KR\alpha_0 - 1$, $\alpha_0' = -\alpha_0 + 1/(KR)$, $\beta = \beta_0/\alpha_0'$, $\gamma = \gamma_0/\alpha_0'$, $\delta' = \delta_0/\alpha_0'$, , $\alpha_0' = -\alpha_0 + 1/(KR)$, $\alpha_0 = a_1 + 2a_2u_0 + 3a_3u_0^2 + 4a_4u_0^3$, $\beta_0 = a_2 + 3a_3u_0 + 6a_4u_0^2$, $\gamma_0 = a_3 + 4a_4u_0$, $\delta_0 = a_4$, $\delta = 1/Q$, ω_0 – резонансная частота контура, R – резонансное сопротивление контура, $K = M/L$ – коэффициент положительной обратной связи, L и M – индуктивность контура и взаимная индуктивность.

Пусть $|d\varphi_c / d\tau| << 1$, $\omega_c \approx \omega_0$ и $I_c = const$. Выражение $u = A\cos(\omega_c t + \varphi)$ это решение уравнения (1), где A и φ медленно меняющиеся функции времени. Тогда моделью автогенератора являются укороченные уравнения, полученные из уравнения (1) методом усреднения:

$$\frac{dA}{d\tau} + \frac{\varepsilon}{2}(\frac{3}{4}\gamma A^3 - A) = \frac{\varepsilon}{2}\frac{KRI_c}{\alpha}\cos(\theta^0),$$
$$\frac{d\theta^0}{d\tau} + \frac{\varepsilon}{2}\frac{KRI_c}{\alpha A}\sin(\theta^0) = -(\frac{\Delta\omega}{\omega_0}) - \frac{d\varphi_c}{d\tau}.$$

Далее эти уравнения будут использоваться в более удобной безразмерной форме

$$\frac{dy}{d\tau} + \frac{\varepsilon}{2}(y^3 - y) = \frac{\varepsilon B}{2\alpha}\cos(\theta^0),$$
$$\frac{d\theta^0}{d\tau} + \frac{\varepsilon B}{2\alpha}\sin(\theta^0) = -(\frac{\Delta\omega}{\omega_0}) - \frac{d\varphi_c}{d\tau}. \qquad (2)$$

где $y = A/A_0$ – безразмерная амплитуда, A и $A_0 = \sqrt{4/(3\gamma)}$ это амплитуды сигналов автогенератора, работающего в режиме синхронизации и авономном режиме соответственно, $\theta^0 = \varphi - \varphi_c$, $B_{1(y)} = B/\alpha > 0$, $B_{2(y)} = B/(\alpha y) > 0$ $B = I_c / I_0$, $I_0 = A_0/(KR)$, $\Delta\omega/\omega_0 = (\omega_c - \omega_0)/\omega_0 << 1$.

Для решения уравнений (2) и используются указанные методы, которые основаны на методе малого параметра, но имеют свою специфику, поскольку малый параметр в уравнениях (2) отсутствует.

Метод квазималого параметра

Данный метод представляет собой традиционный метод малого параметра, когда последний вводится в уравнения искусственно.

Преобразуем уравнения с целью выделения членов, создающих трудности при решении. Переходим к переменной составляющей безразмерной амплитуды колебаний путем использования подстановки $y = y_0(1-\Delta)$, где y_0 значение безразмерной амплитуды колебаний при нулевой частотной расстройке, а Δ является переменной составляющей этой амплитуды. Тригонометрические функции разлагаем в ряд Тейлора. Переписываем систему (2), группируя малые нелинейные члены, создающие трудности при решении, умножаем их на некоторое число μ, и делим на число ν, в итоге полагая, что $\mu = \nu = 1$.

$$\frac{d\Delta}{d\tau} + \frac{\varepsilon}{2}(3y_0^2 - 1)\Delta = \frac{\xi}{2}(\theta^0)^2 + \mu\{\xi\sum_{i=3}^{\infty}(-1)^{i-1}\frac{(\theta^0)^{2(i-1)}}{(2(i-1))!} + \frac{\varepsilon}{2}y_0^2\Delta^2(3-\Delta)\}/\nu,$$
$$\frac{d\theta^0}{d\tau} + \xi\theta^0 = -(\frac{\Delta\omega}{\omega_0}) - \frac{d\varphi_c}{d\tau} - \mu\xi\{\sum_{i=2}^{\infty}(-1)^{i-1}\frac{(\theta^0)^{2i-1}}{(2i-1)!} + (\sum_{i-1}^{\infty}\Delta^i)[\sum_{i=1}^{\infty}(-1)^{i-1}\frac{(\theta^0)^{2i-1}}{(2i-1)!}]\}/\nu.$$

Решение этой системы уравнений ищем в виде рядов

$$\Delta = \sum_{i=0}^{\infty}\mu^i\Delta_i \quad \text{и} \quad \theta^0 = \sum_{i=0}^{\infty}\mu^i\theta_i^0$$

Поскольку данное решение является приближенным, а информационным параметром является сдвиг фазы, то для оценки погрешности использовалась невязка по этому параметру. Изменение невязки в полосе синхронизации в зависимости от числа учтенных членов ряда показано на рис. 2, где δ_0 – невязка решения в первом приближении, δ_1 – невязка, учитывающая поправки первого порядка, δ_2 – невязка, учитывающая поправки первого и второго порядка, и δ_3 – невязка, учитывающая поправки всех трех порядков.

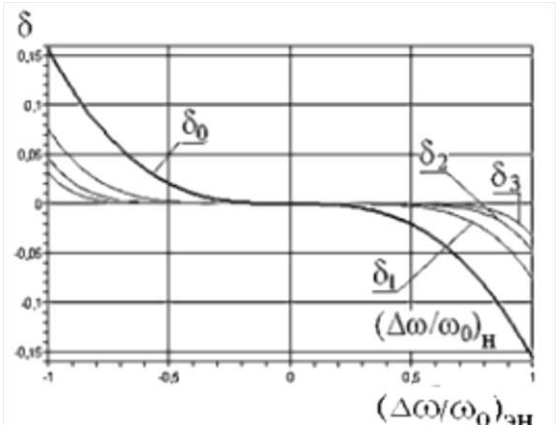


Рис. 2. Изменение невязки
в методе квазималого параметра

$(\Delta\omega/\omega_0)_{эн}$ представляет частотную расстройку. В методе квазималого параметра уравнения первого приближения это линейные дифференциальные уравнения полученные путем линеаризации исходных нелинейных дифференциальных уравнений в окрестности нулевой частотной расстройки. Эти уравнения и вносят основной вклад в сумму ряда, представляющего решение. Для сдвига фазы и стационарного режима это выражение

$$\theta_0 = -\frac{1}{\xi}\left(\frac{\Delta\omega}{\omega_0}\right)_{эн} \qquad (3)$$

где $\xi = \varepsilon B/(2\alpha y_0) = \varepsilon(y_0^2 - 1)/2$

Комбинированный метод малого параметра

Решение уравнений (2) комбинированным методом малого параметра (КМПП) основано на использовании особенности функционирования автогенератора в режиме синхронизации, заключающийся в том, что для автогенератора, синхронизированного на основном тоне, амплитуду колебаний можно считать установившейся при любом мгновенном значении сдвига фазы, т.е. в уравнениях (2) производной $dy/d\tau$ можно пренебречь, и они принимают вид

$$y^3 - y = \frac{B}{\alpha}\cos(\theta^0),$$

$$\frac{d\theta^0}{d\tau} + \frac{\varepsilon B}{2y\alpha}\sin(\theta^0) = -\left(\frac{\Delta\omega}{\omega_0}\right) - \frac{d\varphi_c}{d\tau}.$$

Представляем безразмерную амплитуду колебаний также в виде $y = y_0(1-\Delta)$ и проделываем те же преобразования, что и ранее. Однако малый параметр Δ_s определяется исходя из процедуры аппроксимации фазовой характеристики синхронизированного автогенератора выражением, представляющим первое приближение. Изменение невязки в полосе синхронизации в зависимости от числа учтенных членов ряда, представляющего решение, показано на рис 3.

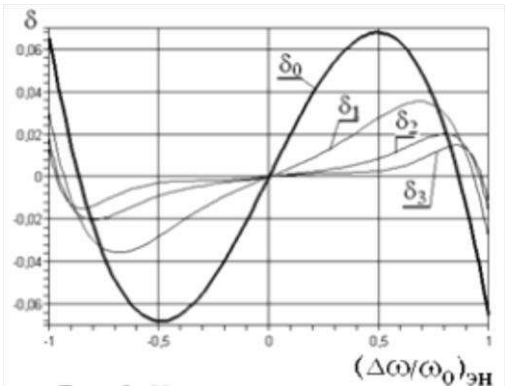


Рис. 3. Изменение невязки
в методе КММП

В комбинированном методе малого параметра уравнения первого приближения получены методом аппроксимации исходных нелинейных дифференциальных уравнений. Для сдвига фазы и стационарного режима это

$$\theta_0 = -\frac{1}{\xi(1-\Delta_s)}\left(\frac{\Delta\omega}{\omega_0}\right)_{эн} \qquad (4)$$

Легко видеть, что в данном случае малое значение невязки получается уже при использовании решения в первом приближении.

Сравнение методов решения укороченных уравнений

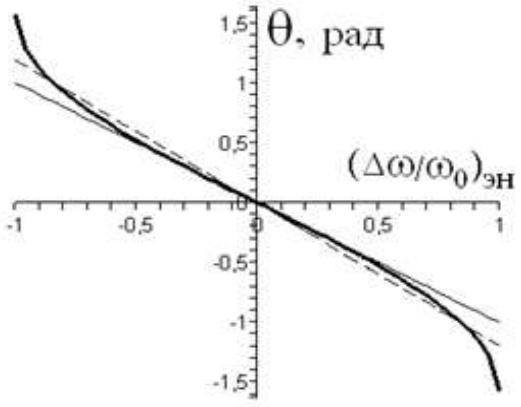


Рис. 4. Фазовая характеристика и первые приближения методов

Сравнение кривых, представленных на рис. 2, 3 говорит о том, что процесс последовательного приближения в обоих случаях является сходящимся, однако комбинированный метод малого параметра позволяет существенно, в 2,5 раза, снизить максимальную величину невязки.

На рис. 4 представлены фазовая характеристика автогенератора и первые приближения рассматриваемых методов. Фазовая характеристика описывается выражением

$$\theta^0 = \arcsin[-(\Delta\omega/\omega_0)_{эн}/\xi].$$

В методе квазималого параметра решение в первом приближении фактически является линеаризацией фазовой характеристики, которая хорошо описывает процессы в достаточно узком диапазоне расстроек в центре полосы синхронизации, поскольку практически совпадает с ней. В центре полосы синхронизации, когда расстройка равна нулю, значение приведенного выражения также равняется нулю, т.е. совпадает со значением фазовой характеристики в этой точке. Очевидно, что погрешность здесь также равна нулю. Погрешность этого метода является непрерывной функцией расстройки $(\Delta\omega/\omega_0)_{эн}$.

Первым приближением в комбинированном методе малого параметра является соотношение (4), которое представляет собой аппроксимацию фазовой характеристики синхронизированного автогенератора и воспроизводит ее с большей точностью во всей полосе расстроек, что существенно уменьшает невязку и позволяет обойтись без поправок при исследовании процесса синхронизации практически во всей полосе синхронизации. Однако особенностью этого метода является поведение погрешности в области малых расстроек. При нулевой расстройке значение этой функции также равно нулю, как и значение фазовой характеристики, что означает нулевое значение погрешности. Рассмотрим теперь значение погрешности при приближении расстройки к нулю справа. Запишем выражение для правостороннего предела:

$$\delta = \lim_{\frac{\Delta\omega}{\omega_0}\to 0}[-\frac{1}{\xi(1-\Delta_s)}(\frac{\Delta\omega}{\omega_0})_{эн} - \arcsin(-\frac{1}{\xi}(\frac{\Delta\omega}{\omega_0})_{эн})]/[\arcsin(-\frac{1}{\xi}(\frac{\Delta\omega}{\omega_0})_{эн})] =$$

$$= \lim_{\frac{\Delta\omega}{\omega_0}\to 0}[-\frac{1}{\xi(1-\Delta_s)}(\frac{\Delta\omega}{\omega_0})_{эн} + \frac{1}{\xi}(\frac{\Delta\omega}{\omega_0})_{эн}]/[-\frac{1}{\xi}(\frac{\Delta\omega}{\omega_0})_{эн}] = \frac{\Delta_s}{1-\Delta_s}$$

Для левостороннего предела получаем аналогичное выражение. Очевидно, что в данном случае предел этой функции при стремлении независимой переменной к нулю существует, но не равен значению этой функции в этой точке. Из этого факта следует, что функция, представляющая погрешность, не является непрерывной. При нулевой расстройке она имеет разрыв. Разрыв такого вида, т.е. с существующим общим пределом, относится к категории устранимых разрывов.

Выводы

Проведен сравнительный анализ двух аналитических методов исследования синхронизированных автогенераторов. Показано, что оба метода могут быть использованы, но метод квазималого параметра требует, как минимум, одну поправку к решению, полученному как первое приближение.

Комбинированный метод малого параметра позволяет обойтись только решением в первом приближении. Однако в данном методе функция, представляющая погрешность, не является непрерывной. Она имеет устранимый разрыв, который можно избежать ее доопределением.

Список литературы:

1. Khokhlov R.V. A Method of Analysis in the Theory of Sinysoidal Self-Oscillations // IRE Trans. Circuit Theory. 1960. Vol. 7, № 4. P. 398-413.
2. Ruthroff C.L. Injection-Locked Oscillator FM Receiver Analysis // The B.S.T.J. 1968. № 7. P. 1653 - 1661.
3. Toyosaku Isobe, Power Amplification for FM and PM Signals with Synchronized IMPATT Oscillators // IEEE Trans. Microwave Theory Tech. 1970. Vol. 18, № 11. P. 906 – 911.
4. Daikoku K., Mizushima Y., Properties of Injection Locking in the non-linear oscillator // Intern. Journ. of Electronics. 1974. Vol. 31, № 3. P. 279-292.
5. Biswas B.N., Ray S.K. Discrimination of a Second-Order Injection Syncyhronized Oscillator Against Interfering Tones // IEEE Trans. Circuits Syst. 1974. Vol. 21, № 3. P. 402- 405.
6. Elwakil A.S., Ozoguz A.S.. On the Generation of Higher Order Chaotic Oscilators via Passive Coupling of Two Identical or Nonidentical Sinusoidal Oscillators // IEEE Trans. Circuits Syst. I. 2006. Vol. 53, № 7. P. 1521 – 1532.
7. Plessas F.C., Papalambrou A., Kalivas G. A 5-GHz Subharmonic Injection-Locked Oscillator and SelfOscillatoing Mixer // IEEE Trans. Circuits Syst. II. 2008. Vol. 55, № 7. P. 633- 637.
8. Zhao L., Xiang L., Liu J., Zhou J. Sampled-data group synchronization of coupled harmonic oscillators subject to controller failure // Proc. CCC 34th Chinese. 2015. P. 2309-2314.
9. Rapin V, Munalo A. Self-oscillator tracking filter with nonlinear feedback // Telecommunications and Radio Engineering. 2019. 78 (2). P.161
10. Rapin V. Synchronized oscillators with the phase-negative feedback // IEEE Trans. on circuits and systems Fundamental theory and applications. 2002. Vol. 49, №. 8. P 1242 – 1245
11. Rapin V. On the phase feedback in the synchronized oscillators // Proc. of 2nd IEEE International Conference on Circuits and Systems for Communications, ICCSC, 2004. June 30, Moscow, Russia.
12. Rapin V. New principle of the phase-locked loop operation // Proc. of 5th IEEE International conference on circuits and systems for communications, 2010. Belgrade, Serbia, November 23-25, P. 145-149
13. Antonio Buonomo, Alessandro Lo Schiavo. Analytical Approach to the Study of Injection-Locked Frequency Dividers // IEEE Trans. Circuits Syst. I. 2013. Vol. 60, No. 1. P. 51- 62.
14. Ahmad Mirzaei, Mohammad E. Heidari, Rahim Bagheri. Saeed Chehrazi, Asad A. Abidi. The Quadrature LC Oscillator: A Complete Portrait Based on Injection Ljcking // IEEE Journal of Solid-State Circuits, 2007. Vol. 42, No. 9. P. 1916-1932.
15. Rapin V.V. Solution of Reduced Equations of Injection-Locked Oscillator // Radioelectronics and Communications Systems 2019. №6. P 271–285.

Поступила в редколлегию 11.03.2021

Сведения об авторе:

Рапин Владимир Васильевич – д-р техн. наук, доцент, Харьковский национальный университет радиоэлектроники, профессор кафедры информационно-сетевой инженерии (ИСИ), факультет инфокоммуникаций, Украина; e-mail: vrapin@ukr.net; ORCID: https://orcid.org/0000-0002-9773-7695

АНТЕНИ ТА ПРИСТРОЇ МІКРОХВИЛЬОВОЇ ТЕХНІКИ

УДК 662.396.67 **DOI:10.30837/rt.2021.2.205.13**

В.В. ДОЛЖИКОВ, д-р физ.-мат. наук

ПРОДОЛЬНОЕ РАСПРЕДЕЛЕНИЕ ИНТЕНСИВНОСТИ ПОЛЯ КРУГЛОЙ СФОКУСИРОВАННОЙ АПЕРТУРЫ

Введение

Одной из характерных особенностей современной теории антенн является резко возросший интерес к изучению структуры поля излучения антенн в их зоне Френеля. Это обусловлено несколькими причинами.

Первая из них – это широкое внедрение в практику систем, в основе которых лежит взаимодействие поля излучения антенны с объектом, находящимся в ее зоне Френеля. К числу таковых относятся системы ближней радиосвязи и радиолокации, беспроводной передачи энергии СВЧ-лучом, антенны с синтезированной апертурой, системы медицинской диагностики и гипертермии, использующие сфокусированные антенны для получения высокого пространственного разрешения, системы беспроводной зарядки мобильных устройств, RFID системы и т. д.

Второй причиной повышенного интереса к зоне Френеля является резкое обострение проблемы ЭМС из-за быстрого роста числа радиоэлектронных средств (РЭС), повышения мощности излучения и чувствительности их приемных устройств, существенно возросших требований к обеспечению нормального функционирования близкорасположенных друг к другу РЭС, что характерно, например, для современных морских судов и летательных аппаратов. К проблеме ЭМС примыкает и важнейшая задача защиты биологических объектов от облучения электромагнитным полем, актуальность которой также усилилась в связи с увеличением числа и мощностей излучения РЭС.

И, наконец, третья причина – это рост электрических размеров L/λ современных антенн, в частности из-за интенсивного освоения все более коротких волн, приводящий к удалению границы дальней зоны ($r_{дз} \approx 2L^2/\lambda$), то есть к увеличению протяженности зоны Френеля и, как следствие, к увеличению числа объектов, попадающих в эту зону.

Изучение структуры и особенностей поля излучающих систем (ИС) в зоне Френеля – задача существенно более трудная, чем анализ их поля в дальней зоне. К таким относятся все задачи, связанные с эволюцией характеристик поля в продольном направлении. Зачастую, даже в простейших задачах, получить результат в аналитической форме затруднительно.

В литературе опубликовано уже немало работ, посвященных исследованию особенностей поля антенн в зоне Френеля [1 – 10]. Однако в большинстве из них приводятся результаты численных расчетов, что не в полной мере удовлетворяет потребности практики.

В работе получены аналитические выражения для основных параметров, характеризующих продольное распределение поля антенны в виде круглой апертуры с равномерным и спадающим возбуждением, сфокусированной как в зону Френеля, так и в дальнюю зону.

Общие соотношения

Рассмотрим плоскую синфазную круглую апертуру с радиусом, равным R. Поместим начало координат в центр апертуры (рис.1). Предположим, что электрическое поле в апертуре линейно поляризовано в направлении x. Тогда x-я компонента напряженности электрического поля в точке $P(r,\theta,\varphi)$ зоны Френеля больших апертур $(2R/\lambda \gg 1)$ определяется формулой Френеля-Кирхгофа [11]:

$$E(r,\theta,\varphi)=\frac{ikE_0(1+\cos\theta)}{4\pi r}e^{-ikr}\int_S A(\rho_1,\varphi_1)e^{i\left[k\rho_1\sin\theta\cos(\varphi-\varphi_1)-\frac{k\rho_1^2}{2r}\left(1-\sin^2\theta\cos^2(\varphi-\varphi_1)\right)\right]}\rho_1 d\rho_1 d\varphi_1\,, \quad (1)$$

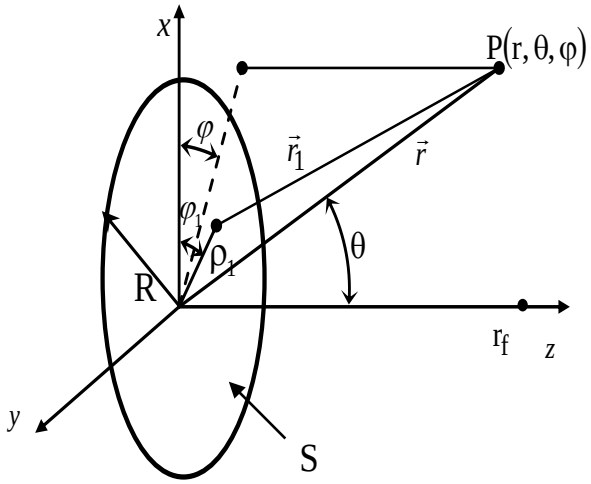


Рис. 1. Геометрия антенны

где E_0 – амплитуда электрического поля на апертуре; $A(\rho,\varphi)$ – функция, описывающая амплитудное распределение возбуждающего поля; $k=2\pi/\lambda$ – волновое число; λ – длина волны в свободном пространстве; r,θ,φ – сферические координаты точки наблюдения; ρ_1,φ_1 – полярные координаты текущей точки на апертуре, S – площадь апертуры.

Если ввести на апертуре дополнительное квадратичное фазовое распределение $k\rho_1^2/2r_f$ (где r_f – фокусное расстояние), то в точке $\theta=0,\ r=r_f$ оно скомпенсирует фазовую ошибку, обусловленную конечностью расстояния до точки наблюдения (второе слагаемое в показателе экспоненты (1)). Поля всех элементарных источников в этой точке будут складываться в фазе. На некоторой части сферы с радиусом r_f (фокальной сфере) угловое распределение поля будет таким же, как и у синфазной апертуры в дальней зоне.

Угловые границы области компенсации определяются из условия

$$\frac{k\rho_1^2}{2r_f}-\frac{k\rho_1^2}{2r_f}\left[1-\sin^2\theta_f^{гр}\cos^2(\varphi-\varphi_1)\right]=\frac{k\rho_1^2}{2r_f}\sin^2\theta_f^{гр}\cos^2(\varphi-\varphi_1)\le\left(kR^2/2r_f\right)\sin^2\theta_f^{гр}=\pi/8\,.$$

Отсюда

$$\sin\theta_f^{гр}=\sqrt{\frac{r_f}{8R^2/\lambda}}=\sqrt{\frac{r_f}{r_{дз}}}=\sqrt{\chi_0}\,,$$

где $r_{д.з}=8R^2/\lambda$ – расстояние до границы дальней зоны, $\chi_0=r_f/r_{дз}$ – нормированное значение расстояния фокусировки, которое связано с числом Френеля N соотношением $\chi_0=1/8N$.

Заметим, что для круглой апертуры область компенсации такая же, как и для линейной антенны.

В приближении малых углов можно считать, что $\sin^2\theta\cos^2(\varphi-\varphi_1)\approx 0,\ (1+\cos\theta)\approx 2$. Введем ряд новых безразмерных переменных: обобщенный угол $\psi=kR\sin\theta$, безразмерную радиальную координату на апертуре $u=\rho_1/R$, нормированную радиальную координату точки наблюдения $\chi=r/r_{дз}$ и обобщенную радиальную координату ζ [7], характеризующую радиальное удаление точки наблюдения от фокальной сферы

$$\zeta=\frac{\pi}{16\chi_0}\left(1-\frac{\chi_0}{\chi}\right). \quad (2)$$

Согласно [9, 11] расстояние до ближней границы зоны Френеля выбрано равным $r_{бл.}=R(2R/\lambda)^{1/3}$, соответственно $\chi_{бл.}=\frac{\sqrt[3]{2}}{8}(\lambda/R)^{2/3}=0.25\sqrt[3]{(\lambda/2R)^2}$. Так, при $R=1\lambda$

$\chi_{бл.} \approx 0.157$ и с ростом радиуса апертуры уменьшается, принимая при $R = 5\lambda$ значение $\chi_{бл.} \approx 0.054$ и при $R = 50\lambda$ значение $\chi_{бл.} \approx 0.012$.

Добавив в показатель экспоненты фокусирующее слагаемое с учетом введенных обозначений, получим из (1) с точностью до множителя $\left(iE_A \pi e^{-ikr}/8\right)$ следующее выражение для поля сфокусированной системы:

$$E(\zeta, \psi, \varphi) = \frac{1}{\chi} F(\zeta, \psi, \varphi), \tag{3}$$

где

$$F(\zeta, \psi, \varphi) = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 A(u, \varphi_1) e^{i2\zeta u^2} e^{iu\psi \cos(\varphi - \varphi_1)} u du d\varphi_1 \tag{4}$$

комплексный множитель круглой апертуры в зоне Френеля. Этот множитель нормирован так, что в фокусе, то есть при $\zeta = 0$ и $\psi = 0$, его значение равно единице.

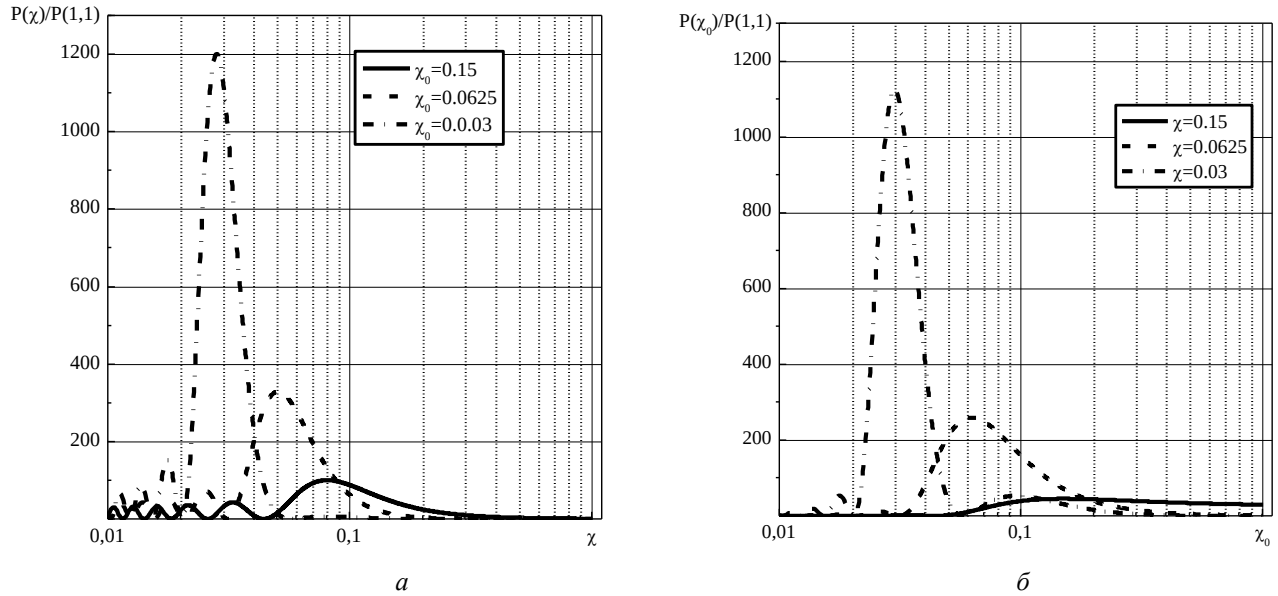
Продольное распределение интенсивности поля (ПРИ) при равномерном возбуждении

Рассмотрим апертуру с равномерным амплитудным распределением $A(\rho, \varphi) = 1$. Для интенсивности поля на фокальной оси $(\psi = 0)$, выполнив интегрирование в (4), получим

$$P(\zeta, \chi_0) = \frac{1}{\chi_0^2} \left[\left(1 - \frac{16\chi_0}{\pi} \zeta \right) \frac{\sin \zeta}{\zeta} \right]^2 \tag{5}$$

Соотношение (5) описывает распределение интенсивности поля вдоль оси круглой равномерно возбужденной апертуры, сфокусированной в фиксированную точку $\chi_0 = \text{const}$. Если же зафиксировано положение точки наблюдения, то (5) определяет зависимость интенсивности в точке наблюдения от расстояния фокусировки.

Рассчитанные по (5) кривые нормированного ПРИ для ближней и дальней фокусировок приведены рис. 2 *а*, *б*.



а *б*

Рис. 2. Зависимость нормированной интенсивности: *а* – от продольной координаты χ для различных координат точки фокусировки, *б* – от координаты точки фокусировки χ_0 для заданных точек на оси

Граница областей ближней и дальней фокусировок условно определена значением $\chi_0 = 0.125$. При таком значении χ расположен последний (дальний от апертуры) максимум продольного распределения интенсивности синфазной апертуры. Критерий $\chi_0 = 0.125$ соответствует числу Френеля $N = 1$, а условия $\chi_0 \le 0.125$ и $\chi_0 \ge 0.125$ (ближней и дальней фокусировок) соответствуют большим $N > 1$ и малым $N < 1$ значениям числа Френеля. Нормировка интенсивности проведена на значение ее на границе дальней зоны синфазной апертуры.

Согласно кривым, приведенным на рис. 2, *а*, продольное распределение имеет многолепестковый характер и максимальное значение достигается в точке, сдвинутой относительно фокуса в сторону апертуры. При расстояниях фокусировки χ_0 больших, чем $1/16 = 0.0625$ $(N > 2)$ интенсивность монотонно спадает по мере удаления от точки фокусировки в сторону дальней зоны. Если расстояние фокусировки $\chi_0 < 0.0625$, то в области $\chi > 0.0625$ возникают дополнительные осцилляции интенсивности и появляются дополнительные нули и максимумы, величина которых очень мала.

Координаты дополнительных нулей $\chi_n^{(0)}$ определяются выражением

$$\chi_n^{(0)} = \frac{\chi_0}{1 - 16\chi_0 n}, \; n = 1, 2 \ldots . \tag{6}$$

Из (6) следует, что координаты нулей интенсивности и их число зависит от расстояния фокусировки. Так, например (рис. 2, *а* – штрихпунктирная кривая), для $\chi_0 = 0.03$ первый нуль будет при $\chi_1^{(0)} = 0.058$, а второй при $\chi_2^{(0)} = 0.75$.

Согласно рис. 2, *б* максимально возможное значение интенсивности $P_{max}(\chi)$ в заданной точке на фокальной оси достигается при фокусировке именно в эту точку, то есть при $\chi_{0,max} = \chi$. Величина $P_{max}(\chi)$ при этом, согласно (5), равна $P_{max}(\chi) = 1/\chi^2$.

Для характеристики свойств ПРИ в зоне Френеля наиболее часто используются следующие параметры [9, 11]: 1) смещение максимума интенсивности поля (МИП) вдоль фокальной оси относительно точки фокусировки $\Delta\chi_{max}$ (Focal Shift – FS); 2) глубина фокусировки D_f (ширина главного лепестка продольного распределения интенсивности на уровне – 3 дБ (Depth of Focus – DoF); 3) усиление фокусировки G_f (Focusing Gain – FG), под которым понимают отношение интенсивности в максимуме к интенсивности на границе дальней зоны синфазной равномерно возбужденной апертуры.

Смещение максимума интенсивности. Координата МИП определяется из условия равенства нулю первой производной от $P(\chi_0, \zeta)$ по ζ

$$\left(1 - \frac{16\chi_0}{\pi}\zeta\right)\left(\frac{\sin\zeta}{\zeta}\right)^2\left[-\frac{16\chi_0}{\pi} + \left(1 - \frac{16\chi_0}{\pi}\zeta\right)\left(\frac{\cos\zeta}{\sin\zeta} - \frac{1}{\zeta}\right)\right] = 0, \tag{7}$$

Исключив из рассмотрения точки $\zeta = \pm\pi n, \quad n = \overline{1,\infty}$, в которых интенсивность обращается в нуль, после ряда преобразований получим уравнение для нахождения координаты максимума

$$\frac{\cos\zeta}{\sin\zeta} = 1 \Big/ \left[\zeta\left(1 - \frac{16\chi_0}{\pi}\zeta\right)\right]. \tag{8}$$

Исследование функций, стоящих в правой и левой частях уравнения (8), показывает, что

искомый корень, который обозначим $\zeta_{\max}$, лежит в интервале $[-\pi/2,\ 0]$. Уравнение (8) является трансцендентным и в общем случае допускает только численное решение.

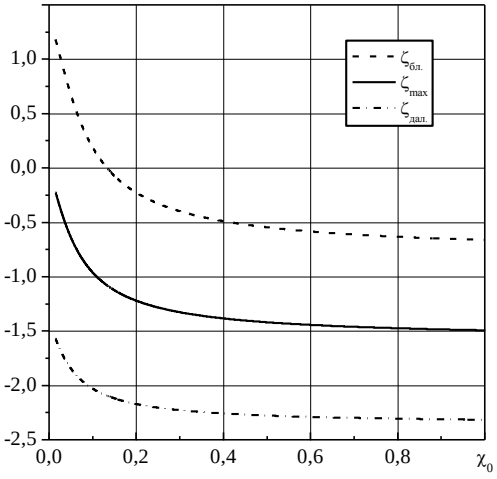


Рис. 3. Зависимость координат максимума, ближней и дальней границ фокальной области от расстояния фокусировки

На рис. 3 представлены зависимости координат максимума $\zeta_{\max}$ главного лепестка (сплошная кривая) и его границ $\zeta_{бл.}$ и $\zeta_{дал.}$ на уровне 0,5 максимума (штриховые кривые), от расстояния фокусировки. На этом и последующих следующих рисунках начальное значение χ_0 взято равным 0.01.

Однако для областей ближней и дальней фокусировок возможно получение приближенного решения в аналитическом виде. Из рис. 3 видно, что для глубокой и мелкой фокусировок $\zeta_{\max}$ мало отличается от 0 и $-(\pi/2)$ соответственно. Следовательно, при определении корня уравнения (8) в случаях дальней и ближней фокусировок можно ввести малый параметр, воспользоваться разложением левой и правой частей (8) в ряд по этому параметру и получить алгебраические уравнения, допускающие аналитическое решение.

Для дальней фокусировки $(\chi << 1)$, положив $\zeta = 0.5\pi - x$, в качестве такого параметра можно взять x, для ближней – модуль переменной ζ.

Опустив промежуточные вычисления, с учетом (2) приведем окончательные выражения для координат главного максимума интенсивности:

$$\chi_{\max} = \frac{2\chi_0}{1+\sqrt{1+12\left(\frac{16\chi_0}{\pi}\right)^2}}, \qquad \chi_0 \le 0.125; \tag{9а}$$

$$\chi_{\max} = \frac{1}{8}\Bigg/\left[1+\frac{1}{8\chi_0}-\frac{4(1+8\chi_0)}{\pi^2(1+8\chi_0)^2-4(1+16\chi_0)}\right], \qquad \chi_0 \ge 0.125. \tag{9б}$$

Формулы (9а), (9б) имеют погрешность $\le 2.3\%$ при $\chi_0 \le 0.125$ и $\chi_0 \ge 0.125$ соответственно. Они позволяют определить положение максимума интенсивности при любых расстояниях фокусировки с приемлемой для практики точностью. Отметим, что для синфазной апертуры $(\chi_0 = \infty)$ координата максимума $\chi_{\max} = 1/8$.

Сдвиг точки максимума $\Delta\chi_{\max} = \chi_0 - \chi_{\max}$ относительно точки фокусировки определится следующими выражениями:

$$\Delta\chi_{\max} = 3\left(\frac{16}{\pi}\right)^2 \chi_0^3\left[1-6\left(\frac{16}{\pi}\right)^2 \chi_0^2+45\left(\frac{16}{\pi}\right)^4 \chi_0^4\right], \qquad \chi_0 \le 0.125, \tag{10а}$$

$$\Delta\chi_{\max} = \chi_0\left[\sqrt{1+12\left(\frac{16\chi_0}{\pi}\right)^2}-1\right]\Bigg/\left[\sqrt{1+12\left(\frac{16\chi_0}{\pi}\right)^2}+1\right], \quad \chi_0 \ge 0.125. \tag{10б}$$

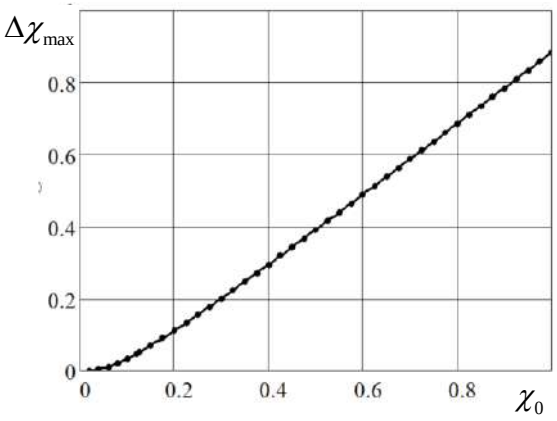


Рис. 4. Зависимость смещения точки максимума от расстояния фокусировки

Погрешность результатов, рассчитанных по формулам (10а), (10б) $\le 3.5\%$ (рис. 4). Максимальная погрешность при $\chi_0 \approx 0.125$. На рис. 4 и далее сплошные или штриховые кривые взяты из работы [10], а точками обозначены результаты расчетов по соотношениям, полученным в данной работе.

Согласно (10а) при приближении точки фокусировки к апертуре смещение точки максимума относительно фокуса убывает пропорционально χ_0^3. Так, при $\chi_0 \to 0$ величина смещения максимума $\Delta\chi_{max} \to 3(16/\pi)^2 \chi_0^3$.

Для $\chi_0 \le 1/96$, близкое к (10а) выражение получено в [12]:

$$\Delta\chi_{max} \approx \frac{\chi_0}{1+\frac{1}{3}\left(\frac{\pi}{16}\right)^2 \chi_0^2}.$$

Усиление фокусировки. Усиление фокусировки $G_f(\chi_0)$ определяется как отношение интенсивности в максимуме к интенсивности на границе дальней зоны при синфазном и равномерном возбуждении:

$$G_f(\chi_0) = P_{max}(\chi_0)/P(\chi = 1, \chi_0 = \infty). \qquad (11)$$

На основании (2), (5) и (9) получим:

$$G_f(\chi_0) = \frac{1}{\chi_0^2}\left(1 + 77.815\chi_0^2 - 1346\chi_0^4\right), \qquad \chi_0 \le 0.125, \qquad (12а)$$

$$G_f(\chi_0) = \left(\frac{16}{\pi}\right)^2 \left(1 + 0.245\frac{1}{\chi_0} + 0.028\frac{1}{\chi_0^2}\right), \quad \chi_0 \ge 0.125. \qquad (12б)$$

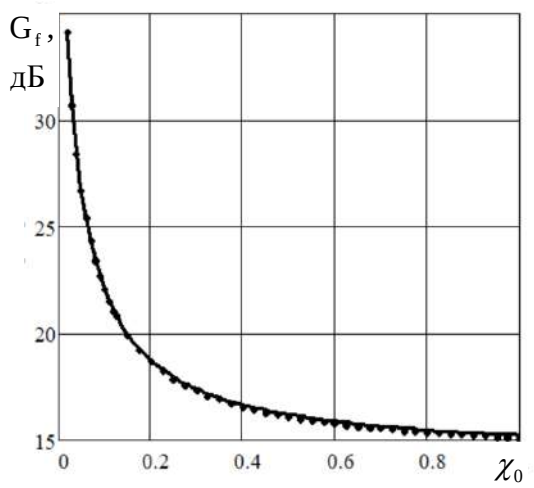


Рис. 5. Зависимость усиления фокусировки от расстояния фокусировки

Погрешность величины $G_f(\chi_0)$, определяемой по (12) не превышает 2% для всей области значений χ_0 (рис.5). При приближении точки фокусировки в пределах зоны Френеля к апертуре усиление фокусировки монотонно растет пропорционально χ_0^{-2}, а при увеличении расстояния фокусировки до бесконечности $G_f(\chi_0 \to \infty) \to (16/\pi)^2$.

Глубина фокусировки. Глубину фокусировки D_χ (продольный размер главного лепестка) принято оценивать шириной главного лепестка продольного распределения интенсивности на уровне 0,5 максимального значения. Точные значения координат ближней и дальней границ на уровне 0,5 находятся численным решением уравнения

$$\mathrm{P}(\zeta)=0.5\,\mathrm{P}_{\max}. \qquad (13)$$

Для получения приближенных формул, справедливых отдельно для мелкой и ближней фокусировок воспользуемся тем, что координаты ближней и дальней границ, выраженные в единицах ζ, незначительно отличаются от величин $\zeta^{(1,2)}=\zeta_{\mathrm{m}}\pm 1$, приведенных на рис. 3. Для значения интенсивности на уровне 0,5 с точностью до членов первого порядка малости имеем

$$\mathrm{P}\left(\zeta_{1,2}\right)=\mathrm{P}\left(\zeta^{(1,2)}\right)+\Delta\zeta_{1,2}\frac{\mathrm{d}}{\mathrm{d}\zeta}\mathrm{P}(\zeta)_{\zeta^{(1,2)}}=0.5\,\mathrm{P}_{\mathrm{m}},$$

тогда $\Delta\zeta_{1,2}=\left[0.5\,\mathrm{P}_{\mathrm{m}}-\mathrm{P}\left(\zeta^{(1,2)}\right)\right]\Big/\mathrm{P}(\zeta)\big|_{\zeta}$ и соответствующие координаты границ лепестка

$$\zeta_{1,2}=\zeta^{(1,2)}+\Delta\zeta_{1,2}.$$

Опустив промежуточные вычисления, приведем окончательные выражения для глубины фокусировки

$$\mathrm{D_f}\left(\chi_0\right)=\chi_0^2\left(\frac{16}{\pi}\right)^2\left(0.594-4.253\chi_0+7.574\chi_0^2\right), \qquad \chi_0\le 0.125, \qquad (14\text{а})$$

$$\mathrm{D_f}\left(\chi_0\right)=\frac{1}{6}-0.017\frac{1}{\chi_0}+0.00068\frac{1}{\chi_0^2}. \quad \chi_0\ge 0.125. \qquad (14\text{б})$$

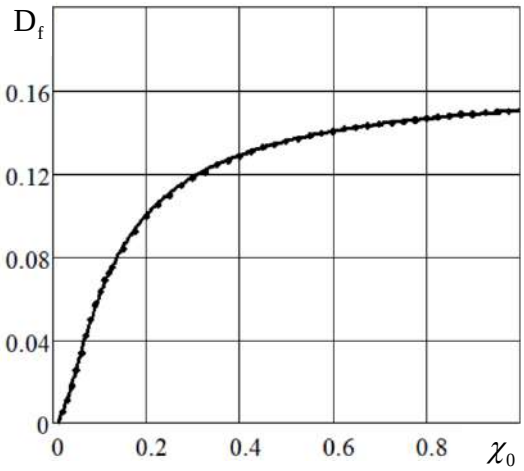


Рис. 6. Зависимость глубины фокусировки от расстояния фокусировки

Рассчитанные по (14а) значения глубины фокусировки отличаются от точных не более, чем на 3,2 %, а по (14б) не более чем на 1,5 % (рис. 6).

Из (14) следует, что глубина фокусировки уменьшается с уменьшением χ_0. Наиболее быстро уменьшение происходит в области глубокой фокусировки – пропорционально χ_0^2. При расстоянии фокусировки, равном расстоянию до границы дальней зоны $(\chi_0=1)$ $\mathrm{D_f}=0.15$, а при $\chi_0\to\infty$ глубина фокусировки, выраженная в единицах расстояния до границы дальней зоны, стремится к значению, равному $1/6\approx 0.17$.

Продольное распределение интенсивности поля при спадающем к краям возбуждении

Рассмотрим апертуру со спадающим амплитудным распределением типа «парабола на пьедестале» $\mathrm{A(u)}=1-(1-\Delta)\mathrm{u}^2$, где Δ – высота пьедестала, которое во многих случаях хорошо аппроксимирует реальное распределение амплитуды [13, 14].

Полагая в (4) $\psi=0$, для продольного распределения интенсивности поля будем иметь

$$\mathrm{P}(\zeta)=\frac{1}{\chi_0^2}\left(1-\frac{16\chi_0}{\pi}\zeta\right)^2\frac{\sin^2\zeta}{\zeta^2}\left\{\Delta+\frac{(1-\Delta)^2}{4}\left[1+\left(\frac{1}{\zeta}-\frac{\cos\zeta}{\sin\zeta}\right)^2\right]\right\}. \qquad (15)$$

На рис. 7 показано продольное распределение нормированной интенсивности при ближней $(\chi_0 \le 0.125)$ и дальней $(\chi_0 > 0.125)$ фокусировках соответственно.

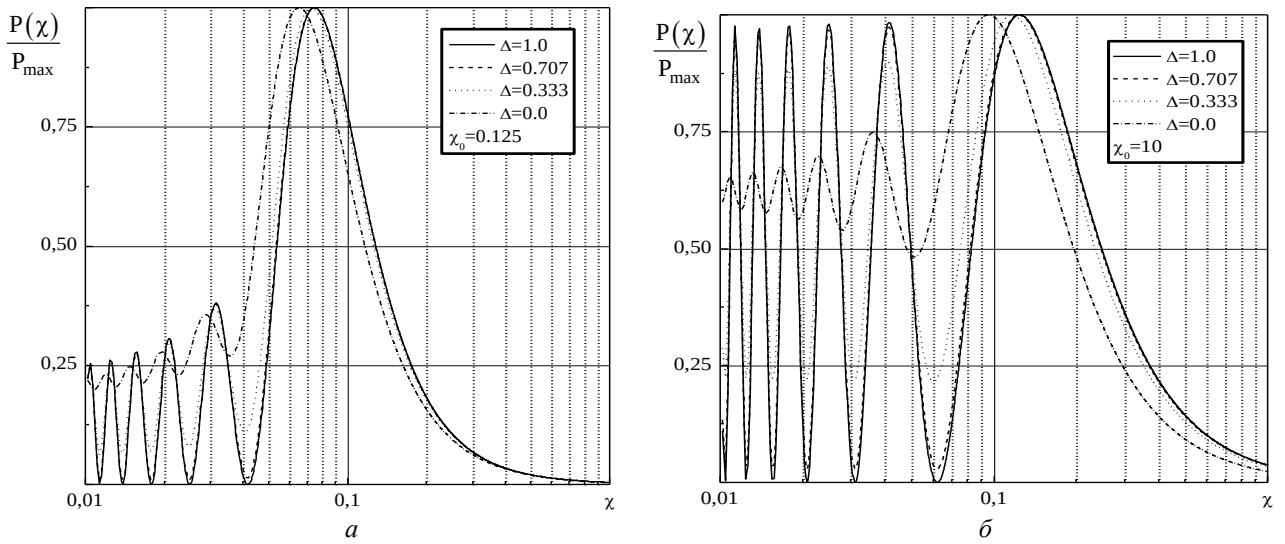


Рис. 7. Продольное распределение интенсивности при неравномерном возбуждении

Видно, что с уменьшением величины пьедестала происходит сглаживание осциллирующего характера продольного распределения: увеличивается средний уровень интенсивности, относительно которого осциллирует ее величина и уменьшаются амплитуды этих осцилляций. При этом имеют место смещение точки максимума интенсивности поля к апертуре по сравнению с случаем равномерного возбуждения, уменьшение величины интенсивности в максимуме, изменение глубины фокусировки, заполнение нулей.

Смещение максимума интенсивности. Координата точки МИП определяется из условия равенства нулю первой производной от $\mathrm{P}(\zeta)$ по ζ, которое приводит к следующему уравнению:

$$\Delta\left[\left(1-\frac{16\chi_0}{\pi}\zeta\right)\frac{\cos\zeta}{\sin\zeta}-\frac{1}{\zeta}\right]-\frac{(1-\Delta)^2}{4}\left[\frac{16\chi_0}{\pi}+2\left(1-\frac{8\chi_0}{\pi}\zeta\right)\left(\frac{1}{\zeta}-\frac{\cos\zeta}{\sin\zeta}\right)^2\frac{1}{\zeta}\right]=0 \qquad (16)$$

Так как аналитическое решение (16) в общем случае невозможно, то аналогично тому, как это было сделано в случае равномерного возбуждения, получено следующее приближенное выражение

$$\Delta\chi_{\max}(\chi_0,\Delta)=\frac{\chi_0\sqrt{1+12\left(\frac{16\chi_0}{\pi}\right)^2\left[1-1.24\left[\frac{(1-\Delta)^2}{12\Delta+2(1-\Delta)^2}\right]^2\right]}-1}{\sqrt{1+12\left(\frac{16\chi_0}{\pi}\right)^2\left[1-\left[\frac{(1-\Delta)^2}{12\Delta+2(1-\Delta)^2}\right]^2\right]}+\frac{6\Delta}{6\Delta+(1-\Delta)^2}} \qquad (17)$$

Погрешность расчетов по (17) не более 7 % (рис.8). Максимальная погрешность при $\chi_0 \approx 0.1$ и нулевом пьедестале.

Из зависимостей смещения максимума при спадающем амплитудном возбуждении, нормированного на смещение при равномерном возбуждении, от высоты пьедестала при различных расстояниях фокусировки, показанных на рис. 9, и соотношения (17) видно, что при любых значениях расстояния фокусировки уменьшение пьедестала ведет к увеличению

смещения максимума. Эффект влияния изменения Δ на величину смещения точки МИП значительно усиливается по мере уменьшения высоты пьедестала. Наиболее сильно он проявляется при $\Delta < 0.25$. Следует также отметить, что если при ближней фокусировке величина смещения точки МИП за счет высоты пьедестала заметно зависит от значения расстояния фокусировки, то при мелкой фокусировке влияние Δ практически одинаково для всех χ_0. Особенно это характерно для $\chi_0 \geq 0.5$ (шрих-пунктирная кривая на рис. 9).

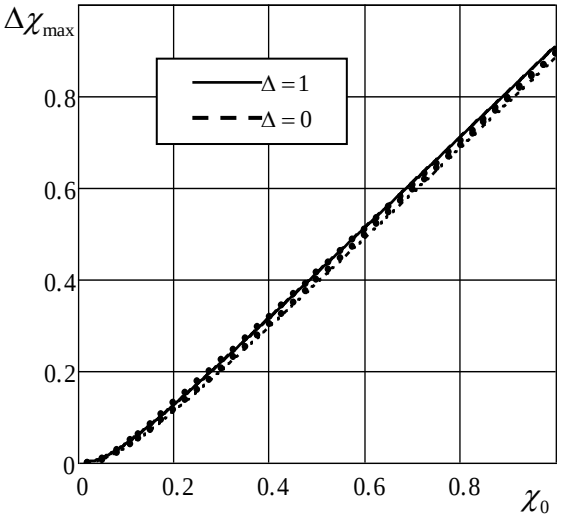


Рис. 8. Зависимость смещения максимума от расстояния фокусировки

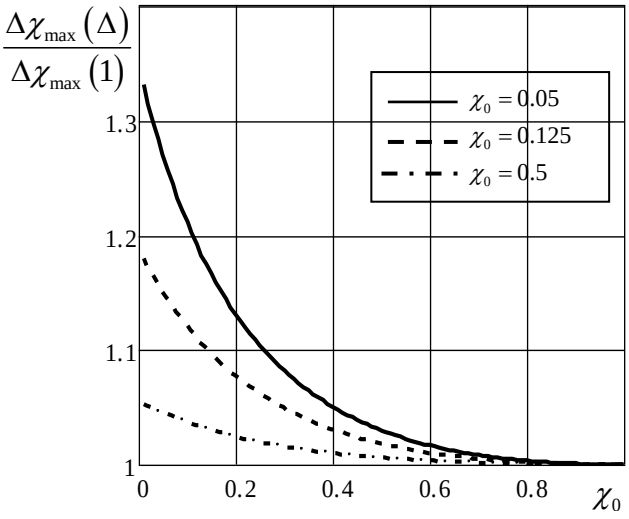


Рис. 9. Зависимость нормированного смещения максимума от расстояния фокусировки

Усиление фокусировки. Величина усиления фокусировки определяется по (11), (15). Соответствующие выражения имеют следующий вид:

ближняя фокусировка ($\chi_0 \leq 0.125$)

$$G_f = \frac{1}{\chi_0^2}\left[\begin{array}{l}\left(1+77.815\chi_0^2-1346\chi_0^4\right)+\left(1+90.784\chi_0^2-1547\chi_0^4\right)(1-\Delta)+\\ +\left(0.25+38.907\chi_0^2-504.598\chi_0^4\right)(1-\Delta)^2\end{array}\right] \qquad (18а)$$

дальняя фокусировка ($\chi_0 \geq 0.125$)

$$G_f = \left(\frac{16}{\pi}\right)^2\left[\begin{array}{l}\left(1+0.245\frac{1}{\chi_0}+0.028\frac{1}{\chi_0^2}\right)-\left(1+0.246\frac{1}{\chi_0}+0.027\frac{1}{\chi_0^2}\right)(1-\Delta)+\\ +\left(0.378+0.092\frac{1}{\chi_0}+0.00469\frac{1}{\chi_0^2}\right)(1-\Delta)^2\end{array}\right]. \qquad (18б)$$

Результаты расчетов по (18а) имеют погрешность не более 4,5 %, а по (18б) – не более чем 3 % (рис. 10, 11).

Согласно (18) с уменьшением пьедестала усиление фокусировки уменьшается как при ближней, так и при мелкой фокусировках.

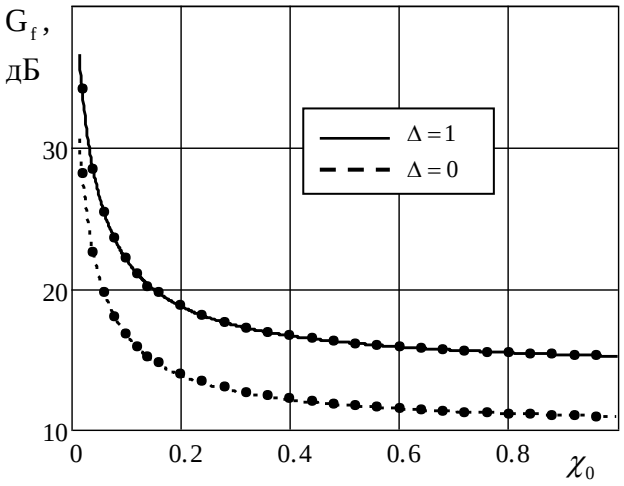


Рис. 10. Зависимость усиления фокусировки от расстояния фокусировки для различных высот пьедестала

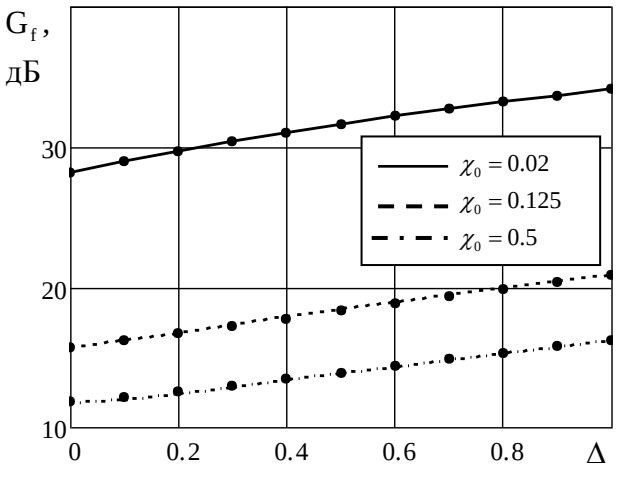


Рис. 11. Зависимость усиления фокусировки от высоты пьедестала для различных значений расстояния фокусировки

Глубина фокусировки. Для определения глубины фокусировки необходимо предварительно определить положение ближней и дальней точек на оси, в которых значение интенсивности равно $0.5P_{max}$. Координаты этих границ удовлетворяют уравнению (13). Они находятся в предположении, что значения их незначительно отличаются от значений при равномерном распределении. Опустив несложные, но громоздкие вычисления, приведем окончательные выражения для глубины фокусировки:

ближняя фокусировка ($\chi_0 \le 0.125$)

$$D_f = \chi_0^2\left(\frac{16}{\pi}\right)^2\left[\left(0.59-4.227\chi_0+7.66\chi_0^2\right)+\left(0.145-2.444\chi_0+10.343\chi_0^2\right)(1-\Delta)^2\right] \qquad (19а)$$

дальняя фокусировка ($\chi_0 \ge 0.125$)

$$D_f = \left(0.1667-0.018\frac{1}{\chi_0}+0.00068\frac{1}{\chi_0^2}\right)-\left(0.021-0.00314\frac{1}{\chi_0}-0.00009\frac{1}{\chi_0^2}\right)(1-\Delta)^2. \qquad (19б)$$

Результаты расчетов по (19а) имеют погрешность не более 4 % и по (19б) не более 3 % (рис. 12, 13).

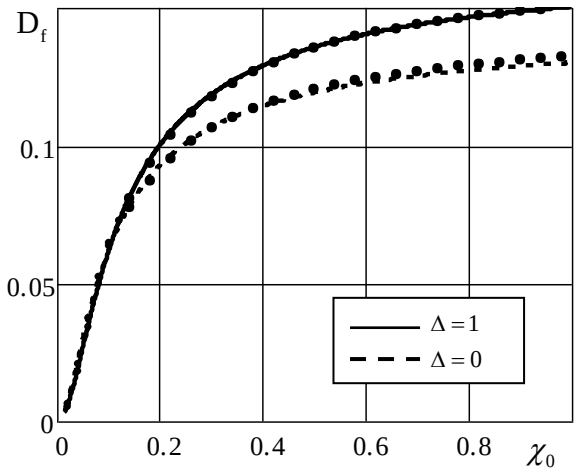


Рис. 12. Зависимость глубины фокусировки от расстояния фокусировки для различных высот пьедестала

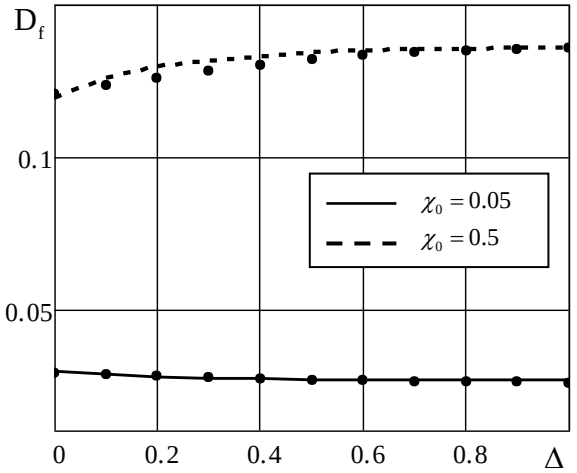


Рис. 13. Зависимость глубины фокусировки от высоты пьедестала для различных расстояний фокусировки

Из (19) следует, что при ближней фокусировке с уменьшением пьедестала продольный размер главного лепестка увеличивается, а при мелкой уменьшается. Так, расширение главного лепестка при нулевом пьедестале для $\chi_0 = 0.02$ составляет 22 %, а его сужение для $\chi_0 = 1.0$ (синфазной апертуры) примерно равно 13 %.

Выводы

Получены аналитические выражения для расчета основных параметров, характеризующих продольное распределение интенсивности поля круглой сфокусированной апертуры: смещения максимума интенсивности относительно точки фокусировки, усиления фокусировки, глубины фокусировки. Рассмотрены случаи равномерного и спадающего амплитудных распределений поля возбуждения. Сравнение с результатами численных расчетов показало, что полученные приближенные соотношения позволяют определить значения упомянутых параметров для любых значений расстояния фокусировки, лежащих как в зоне Френеля, так и в дальней зоне с погрешностью, не превышающей 7 %. Результаты работы будут полезны при расчете поля апертурных антенн в виде круглой сфокусированной апертуры, а также сфокусированных антенных решеток, работающих в зоне Френеля.

Список литературы:

1. Bickmore R. W. and Hansen R. C. Antenna Power Densities in the Fresnel region // Proceedings IRE, **47**, December 1959, pp. 2119-2120.
2. Hu M. K. Fresnel region fields of circular aperture antennas // J. Res.Nat. Bureau Standards, vol. 65D, no. 2, pp. 137–149, Mar.Apr. 1961.
3. Sherman J. W. Properties of Focused Apertures in the Fresnel Region // IRE Transactions on Antennas and Propagation, **10**, 4, July 1962, pp. 399-408.
4. Hansen R. C. Microwave Scanning Antennas. Vol. 1:Apertures, New York, Academic Press, 1964.
5. Hansen R. C. Focal Region Characteristics of Focused Array Antennas // IEEE Transactions on Antennas and Propagation, AP-33, 12, December 1985, pp1328-1337.
6. Kay A. Near-field gain of aperture antennas // IEEE Trans. Antennas Propag., vol. 8, no. 6, 1960, pp. 586–593.
7. Graham W. J. Analysis and Synthesis of Axial Field Pattern of Focused Apertures // IEEE Trans. Antennas Propag., vol. 31, no 4, July 1983, pp. 665-668.
8. Nepa P. Near-Field Focused Antennas for Wireless Communications and Power Transfer // International Spring School on Electromagnetics and emerging technologies for pervasive applications: Internet of Things, Health and Safety. 18th–20th April, 2016, Bologna, Italy.
9. Selvan K.T., Janasvamy R. Fraunhofer and Fresnel Distances // IEEE Antennas and Propagation Magazine. Augest. 2017. P 2-5.
10. Wang W., Gao H., Wu Y., Liu Y. Impact on Focal Parameters for Near‐field‐focused Aperture Antennas // J Numer Model. 2018; e2510, P 1-13. https://doi.org/10.1002/jnm.2510.
11. Silver S. Microwave Antenna Theory and Design. McGraw-Yill, New York,1949. 312 p.
12. Yujun Li, Wolf E. Focal Shift in Diffracted Converging Sphericfl Wave // Opotics Communications. 39, N4. 1981, pp.211-215.
13. Ямпольский В.Г., Фролов О.П. Антенны и ЭМС. Москва : Радио и связь, 1983. 272 с.
14. Balanis, Constantine A. Antenna Theory Analysis and Design, 4-th ed. John Wiley&Sons, Inc., 2016. 1072p.

Поступила в редколлегию 07.03.2021

Сведения об авторе:

Должиков Владимир Васильевич – д-р физ.-мат. наук, профессор, Харьковский национальный университет радиоэлектроники, профессор кафедры компьютерной радиоинженерии и систем технической защиты информации, факультет информационных радиотехнологий и технической защиты информации; Украина, e-mail: vladimir.dolzhikov@nure.ua, ORCID: https://orcid.org/0000-0001-5777-8014

РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

УДК 004.89:621.396 **DOI:10.30837/rt.2021.2.205.14**

В.В. ЖИРНОВ, канд. техн. наук, С.В. СОЛОНСКАЯ, канд. техн. наук

МЕТОД ПРЕОБРАЗОВАНИЯ СИМВОЛЬНЫХ РАДАРНЫХ ОТМЕТОК МАЛОЗАМЕТНЫХ ПОДВИЖНЫХ ОБЪЕКТОВ НА ОСНОВЕ ЭФФЕКТА ТАЛЬБОТА

Введение

В задачах интеллектуальной обработки сигнальной радиолокационной информации используются подходы, основанные на вычислительном интеллекте [1, 2]. Это так называемые нечеткие преобразования, которые устанавливают связь между традиционными преобразованиями и системами нечеткого вывода, а в перспективе и с более мощными средствами вычислительного интеллекта. Предлагаемое нечеткое преобразование – это метод преобразования символьного изображения малозаметных нестационарных радарных отметок на основе эффекта Тальбота [3, 7, 8]. В данной работе показано, как этот подход может использоваться для анализа радиолокационных данных за счет различного представления флуктуаций символьных изображений для разных классов нестационарных радарных отметок. Во-первых, для автоматического обнаружения и распознавания объектов локации из анализа связей и функциональных (семантических) зависимостей между атрибутами (признаками) преобразованных символьных изображений; во-вторых, – для автоматического определения смысловых составляющих символьных изображений радарных отметок.

Преобразование на основе эффекта Тальбота является средством преобразования символьных изображений радарных отметок малозаметных подвижных и малоподвижных воздушных объектов с мерцающими межпериодными флуктуациями, приводящими иногда к полному замиранию сигнала. Это преобразование сводится к установлению определенного соответствия асимптотического равенства восприятия зрительных картин, произвольным образом меняющихся во времени и пространстве, в утверждение об условиях простого равенства восприятия радиолокационных изображений разной частоты флуктуаций.

Радиолокационные данные представляют собой смесь полезных, различных шумовых и мешающих отметок, то есть это нечеткие выборки и множества. Тогда имеется возможность использовать аппарат нечеткого преобразования [4, 5] для обнаружения функциональных зависимостей и связей среди смеси изображений отметок. Полученные зависимости и связи в виде символьных изображений групп сигналов или их следа (пачки) позволяют обнаруживать и распознавать воздушные объекты на фоне ложных сигнальных отметок.

Эффект Тальбота при аппроксимации символьного изображения радарных отметок малозаметных воздушных объектов

В работе проанализирована возможность использования закона Тальбота для разработки и обоснования метода преобразования символьного изображения нестационарных радарных отметок малозаметных подвижных и малоподвижных воздушных объектов (ВО). Радиолокационные изображения радарных отметок малозаметных подвижных воздушных объектов могут иметь мерцающие межпериодные флуктуации, приводящие иногда к полному замиранию сигнала [1, 2]. Такое явление объясняется тем, что вторичное излучение реальных движущихся воздушных объектов рассматривается как вторичное излучение совокупности $n \geq 2$ точек. При изменении положения воздушного объекта блестящие точки перемещаются. Их движение сводится к двум видам: поступательному движению к РЛС и вращательному относительно РЛС. На основании анализа воздействий этих движений изменение (флуктуацию)

огибающей пачки можно пояснить: интерференцией (сложением) когерентных колебаний вторичных отражений от блестящих точек (рис. 1); эффектом Доплера для каждого из блестящих точек (рис. 2); понятием диаграммы обратного вторичного излучения (рис. 3).

Поясним вкратце эти явления. Пусть имеются две связанные блестящие точки летящего объекта, имеющие векторные скорости, одинаково направленные на РЛС, но различные по величине (рис. 1). Тогда центр системы (совокупности блестящих точек) поступательно движется со скоростью $v_{rcp}=\frac{v_{r2}-v_{r1}}{2}$, а вращение с угловой скоростью $\left|\frac{d\theta}{dt}\right|=\frac{\left|v_{r2}-v_{r1}\right|}{l_1}$. Таким образом, при облучении воздушного объекта гармоническим колебанием отраженный сигнал представляет собой результат интерференции двух или больше колебаний, разность фаз которых непрерывно меняется. При этом будут меняться (флюктуировать) амплитуда и фаза результирующего колебания. Спектр сигнала расширяется.

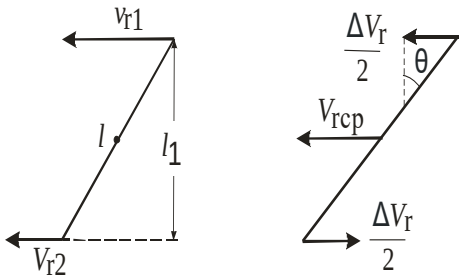


Рис. 1. Движения двух блестящих точек воздушного объекта

Такие же выводы можно получить, основываясь на эффекте Доплера. Если v_{r1} и v_{r2} различны, то отличны и доплеровские частоты $F_{d1}\neq F_{d2}$. Результирующее колебание имеет биения, и как результат появляются флюктуации огибающей пачки импульсов отражений (рис. 2). Приведенная на рисунке величина периода флюктуации огибающей пачки импульсов T обратно пропорциональна разности доплеровских частот ΔF_d блестящих точек подвижного воздушного объекта и при длине волны несущей частоты λ_0

$$T=T_{fl}=\frac{1}{\left|\Delta F_d\right|}=\frac{1}{\left|F_{d1}-F_{d2}\right|}=\frac{\lambda_0}{2\left|\Delta v_r\right|}\ . \qquad (1)$$

К аналогичным выводам придем, заменяя совокупность блестящих точек подвижного воздушного объекта одним излучателем со сложной диаграммой направленностью (ДН) вторичного излучения (рис. 3).

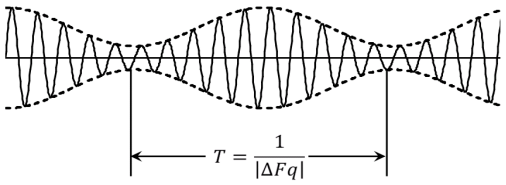


Рис. 2. Биения сигналов отражениями от двух блестящих точек воздушного объекта

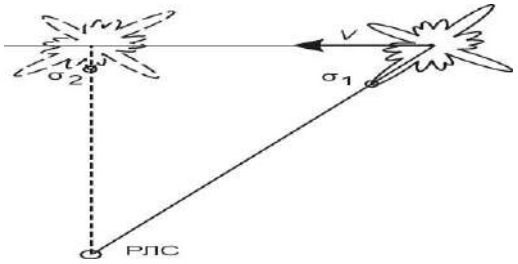


Рис. 3. ДН вторичного излучения подвижного воздушного объекта

Если угловая скорость поворота излучателей равна $\left|\frac{d\theta}{dt}\right|$, а интервал между лепестками диаграммы направленности равен $\Delta\theta$, то средний период флуктуации $T_{fl}\approx\Delta\theta\Big/\left|\frac{d\theta}{dt}\right|$. Сводя воздушный объект к модели из двух блестящих точек (рис. 1), где $l\simeq 20$ì при $\lambda_0=0{,}1$м и

$Cos\theta \simeq 1$ $\Delta\theta = \lambda_0 / 2l\cos\theta \simeq \lambda_0 / 2l$ =1/400рад. Величина угловой скорости поворота излучателей $d\theta/dt$ определяется величинами скорости и радиуса разворота ВО $d\theta/dt = v/R$ и ее величина ограничена с возникающей при развороте перегрузкой $\mu = a/g = v^2/Rg$, которая представляет собой отношение центробежного ускорения a к ускорению земного притяжения g.

Если, например, для самолета μ =3, v =300 м/с, $R \simeq 3$ км и $d\theta/dt \simeq 0{,}1$ рад/с.

Отсюда следует, что минимальный период флюктуации огибающей пачки принятых отражений от самолета $T_{fl.m}$ =1/40 с.

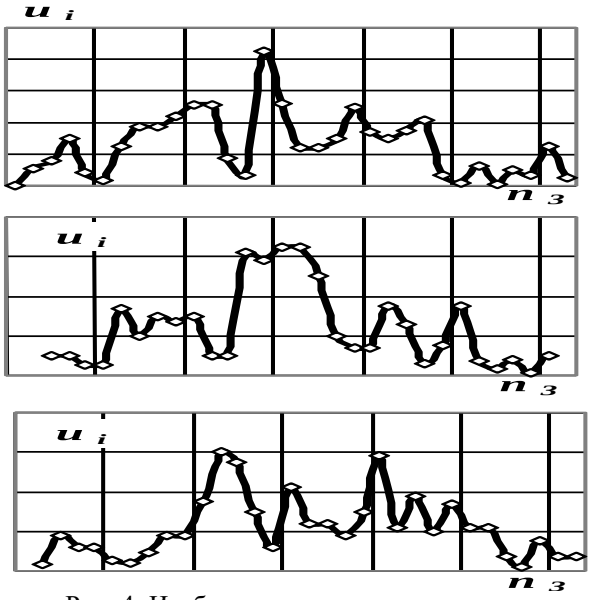


Рис. 4. Изображения отметок ангел-эхо

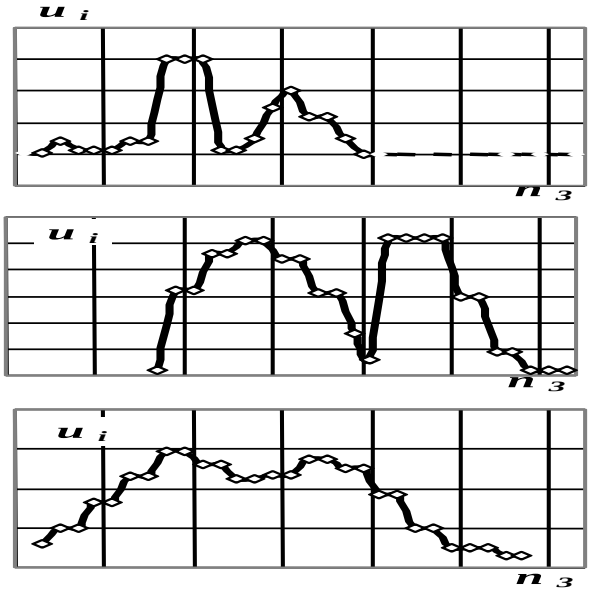


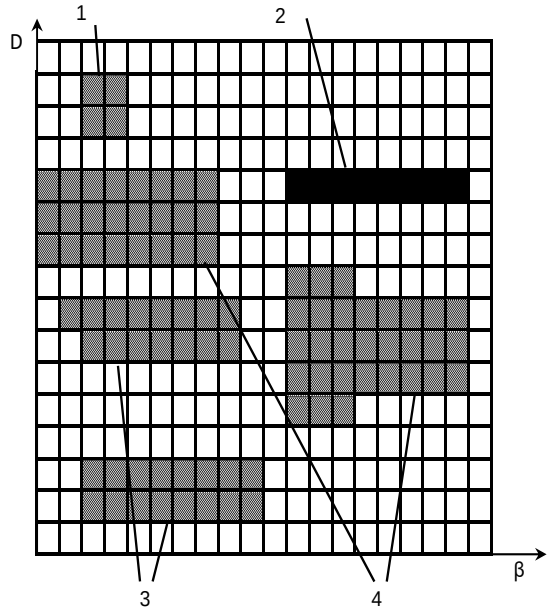
Рис. 5. Изображения отметок самолета

Такая особенность флюктуаций огибающей пачки принятых сигналов является дополнительным признаком для зрительного обнаружения и распознавания малозаметных подвижных и малоподвижных полезных (самолет, вертолет, БПЛ) и мешающих (например, «ангел-эхо») воздушных объектов.

Принцип формирования символьной модели радарных отметок при обнаружении и распознавании малозаметных воздушных точечных объектов

В разработанную модель входят процедуры формализации и анализа символьной модели точечных подвижных объектов на основе алгебры предикатов [8 – 13] и операций, предназначенные для создания предикатной модели процессных знаний при обнаружении и распознавании воздушных точечных объектов. Символьная модель наблюдаемых точечных объектов формируется из набора (пачки) радиолокационных сигналов N информационных ячеек от каждого элемента зоны обзора. Обычно из полученных сигналов формируется карта или матрица данных. В нашем случае формируется символьная модель в виде пачки сигнальных отметок точечных объектов типа самолет, вертолет и БПЛА (рис. 6). Таким образом, обычная база данных превращается в базу знаний, в результате анализа которой можно и нужно получить требуемое решение. Рассмотрим особенности оцениваемых межпериодных информационных потоков. Формируемый массив данных представляет собой матрицу последетекторных амплитуд $\|A\|$ размером $M \times N$. Для таких условий модель символа отметки точечного подвижного объекта будет определяться как совокупность пачки сигналов-отражений от объекта за время облучения его антенной РЛС. Пусть $M = \{q_{11}, q_{12}, ..., q_{ij}, ..., q_{mn}\}$ – множество, представляющее собой матрицу $\|A\|$ размерностью

$M \times N$, состоящее из элементов $k = m \times n$ – значений амплитуд сигналов в элементах обработки зоны обзора РЛС, а B – некоторое из его подмножеств $B \subseteq M$, амплитуды сигналов, которого q_{ij} превышают пороговые значения V_{ij}. Составляем набор логических элементов t_{ij} по следующему принципу: если $q_{ij} \in B$, то $t_{ij} = 1$; если $q_{ij} \notin B$, то $t_{ij} = 0$, $i = \overline{1,m}$, $j = \overline{1,n}$.



1 – импульсная помеха, 2 – точечный объект, 3 – ангел-эхо, 4 – протяженный объект
Рис. 6. Символьные изображения радарных отметок

Предикат $A(x)$ на множестве M, соответствующий множеству B элементов обработки, превысивших порог, с характеристикой $(t_{11}, t_{12}, \cdots, t_{ij}, \cdots, t_{mn})$, запишется формулой:

$$A(x) = t_{11}x^{q_{11}} \vee \ldots \vee t_{mn}x^{q_{mn}} = \bigvee_{i=1,\, j=1}^{mn} t_{ij}x^{q_{ij}} \qquad (2)$$

Выражение $x^{q_{ij}}$ – форма узнавания события, когда $x = q_{ij}$, то $x^{q_{ij}} = 1$.

Предикатная модель процессных знаний о наблюдаемых воздушных или наземных объектах в общем виде – это система n унарных и бинарных предикатов Z_j:

$$M = \left\{ Z_j, \;\; j = 1..n \right\}. \qquad (3)$$

Такая система предикатов (атрибуты или предикатные признаки процесса) позволяет описать ситуацию вокруг анализируемой в данный момент информационной ячейки и процесс формирования символьного изображения отметки из $A(x)$ в течение нескольких циклов зондирования РЛС. Например, для РЛС обзора пространства это могут быть:

– унарный предикат Z_{pij} наличия сигнала в a_{ij} информационной ячейке; i, j – номера элементов зоны обзора РЛС;

– бинарный предикат Z_{dij} ухода сигнала в соседнюю по дальности ячейку a_{ij};

– бинарный предикат Z_{aij} перехода сигнала в смежную по азимуту ячейку, прилегающую к рассматриваемой ячейке.

Эти предикатные признаки формируются по следующим правилам:

$$Z_{pij} = 1, \text{ при } \quad A_{ij} > 0 \qquad (4)$$

$$Z_{dij} = 1, \text{ при } A_{i-1j} > 0 \wedge Z_{pij} = 1 \qquad (5)$$

$$Z_{aij} = 1, \text{ при } Z_{pij} = 1 \wedge A_{ij-1} > 0, \qquad (6)$$

где A_{ij} – предикат события наличия-отсутствия сигнала в соответствующем элементе

На рис. 4 и 5 приведены реальные, экспериментально полученные картины пачек импульсов, отраженных от «ангел-эхо» и самолета. Анализ картин показывает, что радиолокационные изображения нестационарных сигнальных отметок типа отражений от малозаметных подвижных воздушных объектов и естественных мешающих отражений типа «ангел-эхо» могут иметь мерцающие межпериодные флуктуации, приводящие иногда к полному замиранию сигнала. Отсюда появляется возможность использовать закон Тальбота для разработки и обоснования метода преобразования символьного изображения нестационарных

радарных отметок малозаметных подвижных и малоподвижных воздушных объектов, для получения важных символьных (образных) характеристик и смысловых отличительных предикатных признаков для их эффективного обнаружения и распознавания.

Рассмотрим эти возможные символьные преобразования:

1. Для формирования символьного предикатного признака Z_{mij} пачки сигналов используется признак наличия сигнала Z_{aij} в соседних по азимуту ячейках на продолжительности всей длительности пачки. Поскольку могут иметь случаи появления мерцающих межпериодных флуктуаций, приводящих иногда к полному замиранию сигнала, то могут возникнуть сбои при автоматическом формировании символьного предикатного признака Z_{mij} пачки сигналов. Для исключения таких случаев предлагается использовать алгоритм сглаживания мерцающих флюктуаций и основной эффект закона Тальбота. Эти преобразования сводятся к установлению определенного соответствия асимптотического равенства восприятия зрительных картин, произвольным образом меняющихся во времени и пространстве, к утверждению об условиях простого равенства восприятий зрительных (машинных) картин изображений радиолокационных отметок разной частоты флуктуаций.

2. Для формирования частотно-импульсного кода флюктуаций нестационарных радарных отметок типа «ангел-эхо», используя который, можно получить символьную образную характеристику и смысловой отличительный предикатный признак для их эффективного обнаружения и распознавания.

После анализа экспериментальных данных удалось преобразовать формулировку известного обобщенного закона Тальбота [6] и превратить ее из утверждения об условиях асимптотического равенства восприятия зрительных картин, произвольным образом меняющихся во времени и пространстве, в утверждение об условиях простого равенства восприятий зрительных (машинных) картин изображений радиолокационных отметок разной частоты флуктуаций. Преобразованная формулировка обобщенного закона Тальбота имеет следующий вид.

Если отраженный сигнал от малозаметного объекта (МЗО) $A(t)$ и периодически меняющийся во времени сигнал $\{B(t)\}_{\varpi\in(0,\infty)}$ удовлетворяют условию

$$\lim_{\omega\to\infty}\int_{t_1}^{t_2} B_\omega(t)dt = \int_{t_1}^{t_2} A(t)dt, \tag{7}$$

то найдется такая частота $\omega_0 > 0$, что для любого $\omega > \omega_0$ восприятия картины сигнала $B_\omega(t)$ совпадает с восприятием картины сигнала $A(t)$. Здесь t – время, $A(t)$ и $B_\omega(t)$ – изображения графиков (диаграммы) зависимости амплитуды сигнала от времени.

Для отраженного МЗО сигнала $A(t)$ (рис. 7, *a*) можно построить семейство $\{B_\omega(t)\}_{\omega\in(0,\infty)}$ амплитудно-частотной характеристикой (рис. 7, *b*), удовлетворяющей условию (7). Пусть ожидаемый сигнал отметки МЗО $A(t)$ задан на интервале $[0,T_0]$. Полагаем

$$\omega = \frac{a_0 - a}{\varepsilon}, \tag{8}$$

где $a = \max\limits_{0\le t\le T_0} |A(t)|$, $a_0 > a$, $\varepsilon > 0$. Строим последовательность моментов времени $\theta_1, \theta_2, ..., \theta_m$, задавая их равенствами

$$\int_0^{\theta_1}(A(t)+a_0)dt = \varepsilon, \int_0^{\theta_2}(A(t)+a_0)dt = \varepsilon, ..., \int_0^{\theta_m}(A(t)+a_0)dt = \varepsilon. \tag{9}$$

В качестве числа m принимаем наибольшее из натуральных чисел, удовлетворяющее условию $\theta_m \le T_0$. В РЛС обзора воздушного пространства T_0 – это длительность пачки принятых отраженных импульсных сигналов.

В моменты времени $\theta_1, \theta_2, ..., \theta_m$ формируем короткие стандартные импульсы, каждый из которых охватывает площадь ε. Полученную последовательность импульсов, после ее смещения вниз на величину a_0, принимаем в качестве диаграммы сигнала отражений типа «ангел-эхо» $B_\omega(t)$. Диаграмму $B_\omega(t)$ можно с достаточной точностью представить аналитически в виде

$$B_\omega(t) = -a_0 + \varepsilon \sum_{i=1}^{m} \delta(t - \theta_i) \qquad (10)$$

Здесь $\delta(t - \theta_i)$ – функция Дирака. Она задает импульс пренебрежимо малой длительности, возникающий в момент времени θ_i, который охватывает единичную площадь. Диаграмму $B_\omega(t)$ назовем асинхронным частотно-импульсным кодом радиолокационного сигнала отметок отражений типа «ангел-эхо» $A(t)$.

Физический смысл параметра ω следующий. Это минимальная частота следования импульсов в асинхронном частотно-импульсном коде $B_\omega(t)$ сигнала ДМО $A(t)$. Частоту ω можно регулировать практически, изменяя величину площади ε, охватываемой каждым стандартным импульсом. Согласно формуле (8) уменьшение величины ε ведет к увеличению частоты ω. В работе доказано, что множество $\{B_\omega(t)\}_{\omega \in (0,\infty)}$ асинхронных частотно-импульсных кодов сигнала «ангел-эхо» $A(t)$ удовлетворяет условию (7). Это означает, что для этого семейства кодов должно существовать критическое значение $\omega_{кр}$ параметра ω. Величину $\omega_{кр}$ назовем критической частотой следования импульсов в коде $B_\omega(t)$ сигнала $A(t)$.

Поскольку для формирования предикатного признака Z_{mij} пачки сигналов используется признак Z_{aij} соседней по азимуту ячейки, то осуществляется операция прогноза (экстраполяции) данного признака на следующее зондирование на соседнюю по азимуту информационную ячейку с учетом его предыстории. Основой для прогноза значения признака является либо его формирование при выполнении соответствующих условий (6), либо наличие ненулевого уровня признака для предыдущей ячейки, либо выполнение двух этих условий одновременно.

$$Z_{aij+1} = 1 \text{ ïðè } Z_{aij} = 1 \vee (Z_{aij-1} = 1 \wedge A_{ij-1} > 0 \wedge A_{ij} > 0). \qquad (11)$$

Так как в прогнозной формуле (11) осуществляется проверка предикатов наличия сигнала в соседних по азимуту ячейках $A_{ij-1} > 0 \wedge A_{ij} > 0$, то предлагается на этом этапе внести изменения, позволяющие реализовать алгоритм сглаживания мерцающих флюктуаций, используя основной эффект закона Тальбота.

Идея в следующем: если не выполняется условие (11), а именно нет предиката сигнала в текущем зондировании РЛС $A_{ij} = 0$, то на место отсутствующего вставляется замещаемый сигнал с амплитудой, равной амплитуде предыдущей по азимуту ячейки, уменьшенной на величину $c_s \varepsilon$ в соответствии с законом Табольта сигнал согласно формуле (9)

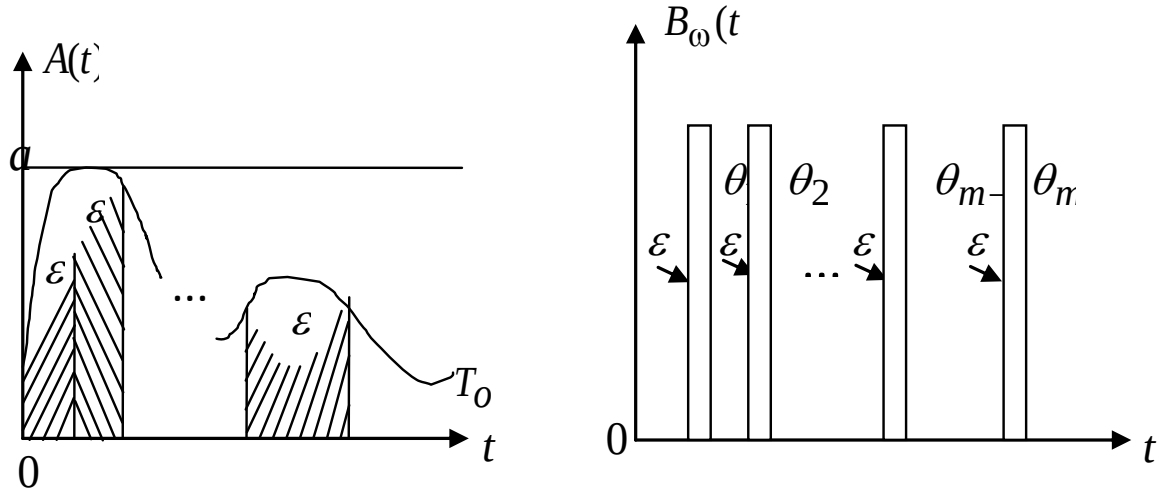
$$A_{ij-1} = 1 \wedge A_{ij-1} = 0 \text{ то } q_{sij} = q_{ij-1} - c_s \varepsilon \ . \qquad (12)$$

Если при этом амплитуда сигнала q_{sij} превышает пороговое значение V_{ij}, то предикат наличия сигнала A_{ij}=1.

Преобразования символьного изображения нестационарных радарных отметок для эффективного обнаружения и распознавания воздушных объектов

Предлагается метод преобразования символьного изображения нестационарных радарных отметок малозаметных подвижных и малоподвижных воздушных объектов для получения важных символьных (образных) характеристик и смысловых отличительных предикатных признаков для их эффективного обнаружения и распознавания. Сформулированы методы автоматического конструирования символьных изображений следования импульсов в асинхронном и синхронном частотно-импульсном коде.

Проведены работы по определению критической частоты $\omega_{кр}$ для различных сигналов нестационарных отражений $A(t)$ и их, асинхронных частотно-импульсных кодов $B_\omega(t)$. В эксперименте использовались записи реальных сигналов «ангел-эхо», полученных на обзорных РЛС сантиметрового диапазона. Определение $\omega_{кр}$ повторялось десятикратно для каждого типа сигнала отражений. Среднеквадратичное отклонение для величины $\omega_{кр}$ составило порядка 5 %. Рассматривались сигналы разной интенсивности от 2 до 48 дБ. При $\omega > \omega_{кр}$ восприятие сигналов $A(t)$ и $B_\omega(t)$ совпадает. Критическая частота $\omega_{кр}$ следования импульсов в асинхронном коде практически не зависит от вида сигнала, на нее влияет амплитуда и нижняя частота флуктуаций сигнала. Критическая частота для разных РЛС строго коррелируется с верхним пределом частоты следования импульсов зондирования.



а *б*

Рис.7. Вид сигнала «ангел-эхо» (*а*), символьное изображение частотно-импульсного кода (*б*), удовлетворяющее условию закона Тальбота

Асинхронные коды преобразованных символьных изображений невозможно ввести в память цифровой вычислительной машины без дополнительной их синхронизации. Для этого необходимо ввести дискретное время и так задержать каждый из импульсов кода, чтобы его передний фронт совпал с ближайшим дискретным моментом времени. Естественно ожидать, что неодинаковое смещение во времени импульсов кода $B(t)$ в процессе его синхронизации может отразиться на характере воспроизведения синхронного кода $C_\psi(t)$. Поэтому возможен случай, когда, несмотря на одинаковое восприятие сигналов $A(t)$ и $B(t)$, картины

$A(t)$ и $C_\psi(t)$ могут быть разными. С другой стороны, ясно, что при увеличении частоты синхронизации ψ положение во времени синхронных импульсов кода $C_\psi(t)$ будет все более приближаться к положению асинхронных импульсов кода $B(t)$. Следовательно, при устремлении к критической частоте синхронизации $\psi_{кр}$ оба кода совпадут.

В ходе работы проведены эксперименты по определению величины $\psi_{кр}$. Частота синхронизации изменялась в пределах от 1 кГц до 20 кГц. Эксперименты показали, что величина $\psi_{кр}$ определяется лишь максимальной частотой $\omega_{\max}$ следования импульсов в сигнале $B(t)$, причем значения $\psi_{кр}$ с высокой точностью укладываются в линейную зависимость от параметра $\omega_{\max}$. Таким образом, имеет место закономерность:

$$\psi_{кр} = \frac{1}{\alpha}\omega_{\max}. \qquad (13)$$

Значения константы дискретности восприятия зрительных картин α, как показывают выполненные опыты, лежат в пределах $\alpha = (2{,}1 + 2{,}8)\cdot 10^{-2}$. При недостаточно высокой частоте синхронизации сигнал $C_\psi(t)$ воспринимается как сигнал $B(t)$, мелькающий на фоне шума.

Выводы

Свойство символьных изображений нестационарных радарных отметок, в том числе, мешающих отражений типа «ангел-эхо», которое проявляется в виде межпериодных мерцающих флуктуаций и приводит иногда к полному замиранию сигнала, предлагается использовать как дополнительный признак для обнаружения и распознавания подобных отметок. Данный признак представляется в виде амплитудно-частотного кода, сформированного с использованием известной закономерности Тальбота для мерцающих картин для данного элемента обработки с учетом предыдущих циклов зондирований РЛС. В результате, можно отказаться от трудоемких операций формирования низкочастотного фильтра межпериодной обработки для защиты от нестационарных мешающих отражений типа "ангел-эхо" и решать эту задачу распознаванием преобразованных символьных изображений с помощью анализа амплитудно-частотного кода отраженного сигнала, формируемого согласно приведенному в статье алгоритму.

Список литературы:

1. Сколник М.И. Справочник по радиолокации : в 2 т. ; пер. с англ. под ред. В.С. Вербы. Москва : Техносфера, 2014. 672 с.
2. Li Jian. Radar Signal Processing and Its Applications / Jian Li, R. Hummel, P. Stoica, E. G. Zelnio. Springer, 2013. 279 p.
3. Talbot H.F. Experiments on light. Phil. Mag., (third series), № 5. 1834. p. 321.
4. Теплов Б.М., Яковлева С.П. О законах пространственного и временного смешения цветов // Зрительные ощущения и восприятия. Т.2. Москва : Соцэкгиз, 1935.
5. Ильин В.Д., Соколов И.А. Символьная модель системы знаний информатики в человеко-автоматной среде // Информатика и ее применения. 2007. Т. 1 №1. С. 66-78.
6. Solonska S., Zhyrnov V. Adaptive semantic analysis of radar data using fuzzy transform (Book Chapter). Springer, 2020, Lecture Notes on Data Engineering and Communications Technologies. Vol 48. P. 157-179.
7. Zhuravlev Yu. I. Analysis of a training sample and classification in one recognition model / Yu. I. Zhuravlev, L. A. Aslanyan, V. V. Ryazanov // Pattern Recognition and Image Analysis: Pleiades Publishing, 2014. Vol. 24, Issue 3. pp 347–352. https://doi.org/10.1134/S1054661814030183.
8. Solonskaya S.V., Zhirnov V.V. Intelligent analysis of radar data based on fuzzy transforms // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radiotekhnika). 2018. 77 (15), pp. 1321-1329.|Scopus|0.69|

9. Solonskaya S.V., Zhirnov V.V. Signal processing in the intelligence systems of detecting low-observable and low-doppler aerial targets // Telecommunications and Radio Engineering (English translation of Elektrosvyaz and Radio-tekhnika). 2018. Vol. 77, Issue 20. P. 1827-1835.
10. Zhyrnov V., Solonska S. PROCESS KNOWLEDGE BOUT OBSERVED OBJECTS IN INTELLECTUAL MONITORING SYSTEMS // Telecommunications and Radio Engineering. 2020. Vol. 79, Issue 18, P. 1599-1607. |Scopus|0.69|. DOI: 10.1615/TelecomRadEng.v79.i18.20.
11. Zhyrnov V., Solonska S. INTELLIGENT SYSTEM FOR DETECTION OF LOW-VISIBLE AIR OBJECTS IN SURVEILLANCE RADARS // Telecommunications and Radio Engineering. 2020. Vol. 79, Issue 17, P. 1513-1519. |Scopus|0.69|. DOI: 10.1615/TelecomRadEng.v79.i17.20.

Поступила в редколлегию 19.02.2021

Сведения об авторах:

Жирнов Владимир Витальевич – канд. техн. наук, Харьковский национальный университет радиоэлектроники, в.н.с. НИЦ интегрированных радиоэлектронных систем и технологий, Украина; e-mail: nauka123@ukr.net

Солонская Светлана Владимировна – канд. техн. наук, доцент кафедры естественных и гуманитарных наук, Харьковский национальный автомобильно-дорожный университет, Украина; e-mail: solonskaya@ukr.net, ORCID: https://orcid.org/0000-0002-8841-7825

УДК 621.396.96, 621.397.48:004.932.2 **DOI:10.30837/rt.2021.2.205.15**

В.М. КАРТАШОВ, д-р техн. наук, В.А.ПОСОШЕНКО, канд. техн. наук,
В.В. ВОРОНИН, канд. техн. наук, В.И. КОЛЕСНИК, А.И. КАПУСТА,
Н.В. РЫБНИКОВ, Е.В. ПЕРШИН

МЕТОДЫ ОБНАРУЖЕНИЯ-РАСПОЗНАВАНИЯ РАДИОЛОКАЦИОННЫХ, АКУСТИЧЕСКИХ, ОПТИЧЕСКИХ И ИНФРАКРАСНЫХ СИГНАЛОВ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Введение

Одна из актуальных задач заключается в защите разнообразных объектов от воздействия беспилотных летательных аппаратов (БПЛА), несущих потенциальную угрозу в военной, хозяйственной и повседневной областях деятельности человека [1, 2]. Относительно невысокая стоимость БПЛА и трудности их наблюдения и контроля приводят к повышению безнаказанности и массовости противоправных действий с их использованием [1 – 3]. Для решения задачи обнаружения, распознавания и измерения координат беспилотных летательных аппаратов используются оптический, инфракрасный, радиолокационный и акустический методы и соответствующие средства [3].

Каждый из известных методов имеет свои достоинства, недостатки и ограничения, характеризуется некоторой областью возможностей, которая определяет множество измеряемых параметров с соответствующими точностными характеристиками и пространственно-временным разрешением, диапазоном дальностей и т. д. Поскольку области возможностей различных методов не совпадают, то появляется предпосылка совместного использования систем различного вида для повышения эффективности их функционирования [4].

Операции измерения координат во всех информационных каналах предшествует задача обнаружения. Известные методы энергетического обнаружения сигналов БПЛА недостаточно эффективны, поскольку операция выполняется, как правило, на фоне разнообразных помех, имеющих определенные структурные сходства с сигналом БПЛА. Вследствие этого задача обнаружения БПЛА на фоне ему подобных объектов на практике во всех рассматриваемых информационных каналах реализуется как задача «обнаружения-распознавания», т.е. при решении задачи обнаружения принимается во внимание наличие некоторых информативных признаков у принимаемого сигнала.

В настоящее время имеется большое количество публикаций, посвященных описанию систем, работающих на разных физических принципах, которые предназначены для обнаружения и наблюдения БПЛА на фоне разнообразных помех и различных объектов, сходных по ряду признаков с БПЛА. В обширной литературе по данному вопросу рассматриваются различные методы приема, обработки информационных сигналов и их последующего интеллектуального анализа. В соответствии с этим актуальной является задача анализа, обобщения и систематизации имеющихся в литературе данных.

Данная статья является обзорной и посвящена анализу возможностей каждого из рассматриваемых методов обнаружения, измерения координат и параметров движения БПЛА, а также комплексной обработке информации, получаемой по каждому из каналов.

Радиолокационный метод

Радиолокационные станции (РЛС) являются традиционным, широко распространенным средством для обнаружения летательных аппаратов. По сравнению с другими методами данная технология позволяет обнаруживать объекты на значительных удалениях (десятки километров), при этом состояние окружающей среды не оказывает определяющего влияния на дальность и основные тактические характеристики станций.

Разработана широкая номенклатура РЛС для решения разнообразных задач при работе по летательным аппаратам [5], имеющим значительную эффективную площадь рассеяния (ЭПР), однако они малопригодны для обнаружения БПЛА. Для наблюдения за специфическими целями потребовалась разработка специализированных РЛС для отделения БПЛА от птиц, имеющих схожие с БПЛА радиолокационные характеристики. Возможности распознавания и деления обнаруживаемых целей на подклассы у классических РЛС достаточно слабые, в то время как у специализированных РЛС по обнаружению БПЛА эти качества являются определяющими.

РЛС, предназначенные для обнаружения БПЛА, являются активными, т.е. они излучают зондирующий сигнал, и получают полезную информацию из сигнала, отраженного от объектов и поступающего на вход. Используются различные виды зондирующего сигнала – импульсный когерентный (в том числе с малой скважностью), непрерывный (в том числе с непрерывным линейно-частотно-модулированным сигналом). Реализуются различные методы обзора – параллельный, а в РЛС кругового обзора – последовательный.

Наиболее распространенная характеристика радиолокационного сигнала РЛС для обнаружения БПЛА – микродоплеровская сигнатура (МДС). МДС широко используется для классификации разнообразных целей при решении различных задач, в том числе в последнее время для классификации БПЛА. Статистические характеристики МДС описывают вращение лопастей винта беспилотного летательного аппарата или вертолета, работу турбин реактивного двигателя, взмах крыльев птиц. В литературе рассматриваются различные методы обработки сигналов при получении сигнатуры, процесса формирования информативных признаков и непосредственно работы классификатора.

Авторы [6, 7] предложили формировать МДС с использованием спектрограммы (на основе кратковременного преобразования Фурье (STFT) [6] и кепстрограммы [7]. Они сосредоточили внимание на формировании признаков, позволяющих выделять из сигнала радара такие характеристики как скорость вращения и наклон лопастей, диаметр и количество винтов, с целью классификации винтовых БПЛА. В [8] предложено использовать кратковременное преобразование Фурье (STFT) и извлекать собственные векторы из корреляционной матрицы МДС в качестве признаков для обучения трех классификаторов: линейного, нелинейного метода опорных векторов и байесовского классификатора, с целью классификации десяти различных винтовых БПЛА и птиц.

В [9] использовалась процедура обработки, аналогичная [8], с последующим применением декомпозиции сингулярных значений (SVD) к спектрограмме. Авторами предложены три основные функции, позволяющие классифицировать скорость цели, периодичность спектра и ширину спектра.

В [10] предложено использовать двумерное регуляризованное комплексное логарифмическое преобразование Фурье и объектно-ориентированную методику уменьшения размерности сигнала с целью повышения информативности и надежности подпространства, специально разработанного для решения задачи распознавания БПЛА на фоне птиц.

В [11] использовался метод эмпирической модовой декомпозиции: полученный сигнал раскладывался на набор осциллирующих колебаний, из которых формировались восемь статистических и геометрических признаков. Использовался нелинейный метод опорных векторов (SVM), который обеспечил более высокие результаты выявления БПЛА, чем в [8].

РЛС кругового обзора для обнаружения БПЛА имеет вращающуюся вкруговую антенну и обеспечивает хорошие поисковые возможности при обнаружении объектов, а также для измерения их координат, но время облучения целей в этом случае получается достаточно малым в силу специфики их работы, что не позволяет формировать МДС. В этом случае классификация объектов осуществляется на основе признаков, описывающих движение цели – траекторной информации.

Авторы [11] использовали выходные данные РЛС кругового обзора, вероятностные модели движения целей, учитывающие направление и скорость движения, а также применяли

фильтр Калмана для сглаживания получаемых оценок. Далее осуществлялся анализ вероятностей описания с помощью используемых моделей наблюдаемых траекторий перемещения БПЛА и птиц. Эффективность предложенных алгоритмов подтверждена результатами компьютерного моделирования и с использованием реальных данных.

В [11] задача классификации БПЛА и птиц по данным РЛС кругового обзора решалась путем формирования девяти поляриметрических параметров, а также использованием алгоритма нахождения ближайшего соседа. Показано, что высокая вероятность правильной классификации может быть достигнута даже без учета признаков, отображающих движение объектов.

В [11] используется двухкоординатная РЛС кругового обзора с непрерывным линейно-частотно-модулированным зондирующим сигналом (LFMCW). Рассматривается двухэтапный процесс автоматической классификации целей: вначале БПЛА отделяются от всех иных объектов (люди, самолеты, автомобили, птицы), а затем БПЛА разделяются на два класса – винтовые и невинтовые. Алгоритм классификации использует набор параметров, основанных на радиолокационной сигнатуре (отношение сигнал-шум (SNR) и траекторная информация (включая информацию о скорости). Используется оптимизированный классификатор SVM, обученный по трем различным схемам. Результаты проведенных экспериментов показали высокую точность решения задачи на обоих этапах, особенно на первом этапе – отделения БПЛА от остального мира.

В последнее время все большее распространение получают методы машинного обучения, применяемые в различных областях, в том числе в области обработки изображений, видео- и аудио-. Применительно к решению разнообразных радиолокационных задач эти достижения скромнее, что связано, прежде всего, с необходимостью использования достаточно представительного массива входных данных для обучения. Кроме того, радиолокационные данные являются достаточно специфическими, и для их использования в системе машинного обучения требуется соответствующая обработка и интерпретация радиолокационных сигналов с целью извлечения необходимой тонкой пространственно-временной радиофизической информации об объектах. Подобная интерпретация сигналов может быть выполнена только соответствующими специалистами.

Несмотря на сложности, методы машинного обучения находят в последнее время все большее применение для обработки радиолокационных данных в различных задачах, в том числе применительно к БПЛА. Основная задача заключается в формировании спектра отраженного сигнала БПЛА, формировании МДС объекта и последующем использовании полученных данных в обучаемой сети.

Авторы [12], используя РЛС с непрерывным частотно-модулированным зондирующим сигналом, получили спектры БПЛА, предложили метод улучшения МДС и провели экспериментальные исследования с различным типами БПЛА, используя GoogleNet.

В [13] с помощью РЛС с непрерывным зондирующим сигналом в S диапазоне были получены функции спектральной корреляции (SCF) БПЛА, отображающие доплеровские особенности целей. Полученные в эксперименте результаты сопоставлялись с четырьмя эталонными SCF, далее осуществлялось их взвешивание и подача на вход глубокой сети доверия (DBN).

В [13] предложен метод обнаружения БПЛА на основе анализа спектра с помощью сверхточной нейронной сети CNN. При этом использовались различные окна для доплеровских спектров целей, на фоне разнообразных помех и при различных отношениях сигнал-шум. Осуществлялось перемещение окна в диапазоне значений доплеровской частоты и решалась задача обнаружения БПЛА на фоне помех.

Аналогичная задача – классификация типов БПЛА путем анализа доплеровских спектров БПЛА с использованием DNN (CNN) решалась в [36]. Эксперименты, проведенные в X диапазоне радиоволн, показали хорошие результаты.

В [13] выполнялось обучение сети непосредственно по отраженному сигналу, поступающему с выхода РЛС и представленному в комплексном виде. Нейронная сеть, состоящая из пяти отдельных ветвей, обучалась на наличие винтов у БПЛА, а затем определялось количество лопастей. Качество решения задачи в значительной степени зависит от отношения сигнал-шум.

Оптический метод

Известно значительное количество публикаций, в которых рассматриваются вопросы обнаружения БПЛА с использованием видеокамер видимого диапазона волн. При этом в качестве детектора, выносящего решение об обнаружении-распознавании БПЛА, используются, как правило, нейронные сети и алгоритмы глубокого обучения, что было обусловлено успехами алгоритмов в классификации изображений по наборам данных базы ImageNet. Предварительно происходит обучение сети [14], как правило, с использованием данных о БПЛА из общедоступной базы ImageNet.

Большое количество экспериментов с использованием сетей глубинной нейронной сети DNN проведено авторами, в которых наилучшие результаты показала сеть VGG-16 [15, 16]. Результаты экспериментов также показывают, что наличие птиц увеличивает количество ложных срабатываний и для повышения эффективности работы алгоритмов предлагается не исключать птиц из видео, используемых в процессе обучения, что позволит обучить систему и выявлять тонкие различия между птицами и БПЛА.

Изображения БПЛА, птиц, а также некоторые фоновые изображения представлены на рис. 1 [35].

Рис. 1. Изображения БПЛА (первый и второй ряд), изображения птиц (третий ряд), фоновые изображения (четвертый ряд)

В работах [17 – 19, 79, 80, 81] рассмотрены новые подходы к обработке изображений, вероятностные характеристики различных объектов.

Авторами [20] предложена сеть на основе сверточной сети VGG в сочетании с RPN для целей обнаружения-распознавания БПЛА и птиц. Авторы подготовили обширную базу изображений птиц, БПЛА и фотографий фона, заимствованных из Интернета, провели обучение сети с использованием этих данных и пришли к выводу, что использование более разнообразных данных в процессе обучения обеспечивает лучшие результаты обнаружения-распознавания объектов.

В [21] предлагается перед основным классификатором расположить U-сеть, выполняющую функции предварительной обработки совокупности последовательных кадров и выделения областей движения в изображениях, которые могут содержать БПЛА. Далее выделенные области обрабатываются основной сетью, выносящей окончательное решение. В статье [22] также используется модуль предварительной обработки изображений, построенный на основе сети глубокого обучения с использованием метода SISR (Single image super-resolution – сверхвысокое изображение одиночного изображения), позволяющий существенно увеличить разрешение изображения, поступающего с видеокамеры. Это позволяет улучшить качество обнаружения и увеличить максимальную дальность, на которой может быть выявлен БПЛА.

В работе [23] при обнаружении БПЛА использовались традиционные методы компьютерной обработки изображений, основанные на сопоставлении исходного и образцового изображений и формировании функции взаимной корреляции. Для уменьшения влияния освещенности на качество обнаружения применялась морфологическая фильтрация, позволяющая расширить возможности обнаружения БПЛА в различных условиях, особенно при достаточно ярких и темных изображениях БПЛА.

В [24] также используются традиционные методы обработки, такие как формирование гистограммы градиентов, для описания малых БПЛА. Кроме того, при проведении машинного обучения применялся каскадный метод классификации для вынесения ряда последовательных оценок на нескольких этапах, характеризующихся все более сложным набором признаков. Если все этапы пройдены успешно, то объект считается обнаруженным. Алгоритм регрессии опорных векторов (SVR) обучался для сетки различных расстояний с целью вынесения оценки дальности до БПЛА при его обнаружении.

Ряд алгоритмов обработки изображений, основанных на методах компьютерного зрения, изложены в работах [25 – 28].

В [29] задача обнаружения БПЛА решается по видеопотоку стационарной видеокамеры. Обработка изображения состоит из этапов обнаружения движущихся объектов и классификации этих объектов с использованием нейросети. При обнаружении БПЛА использованы методы выделения движущихся объектов на неподвижном фоне и анализа истории движения. Для классификации движущихся объектов создана и обучена модель нейросети, позволяющая классифицировать 12 типов подвижных объектов. Проверка работы алгоритма выполнена с использованием экспериментальных видеоданных.

В [30] использовалась сверточная нейронная сеть класса YOLO для детектирования БПЛА на изображениях, которая обеспечивает достаточно быстрое и качественное решение задачи. В статье приведен новый набор данных, который может быть использован для обучения нейронных сетей. Данные получены путем отделения БПЛА от фона на имеющихся изображениях и добавлением изображения БПЛА к различным естественным изображениям с разнообразным и сложным фоном. Созданная база данных позволяет производить глубокое обучение сетей при использовании БПЛА, находящихся в разных условиях и на различных дальностях.

Инфракрасный метод

Видеокамеры инфракрасного диапазона или тепловизионные датчики работают в невидимом диапазоне электромагнитных волн, как правило, в длинноволновом инфракрасном диапазоне с длиной волны 9 – 14 мкм, и принимают излучение, формируемое нагретыми

объектами или отдельными элементами их конструкции. Основным достоинством инфракрасных тепловизионных видеокамер является возможность визуализации окружающего мира и определенных объектов независимо от степени освещения (даже в условиях полной темноты) и погодных условий [31]. Тепловизионные камеры обладают худшим разрешением, чем камеры видимого диапазона, но более устойчивы к изменениям освещенности. Они также более дороги, чем камеры видимого диапазона волн.

Изображения БПЛА, полученные с помощью тепловизионной камеры [35], представлены на рис. 2.

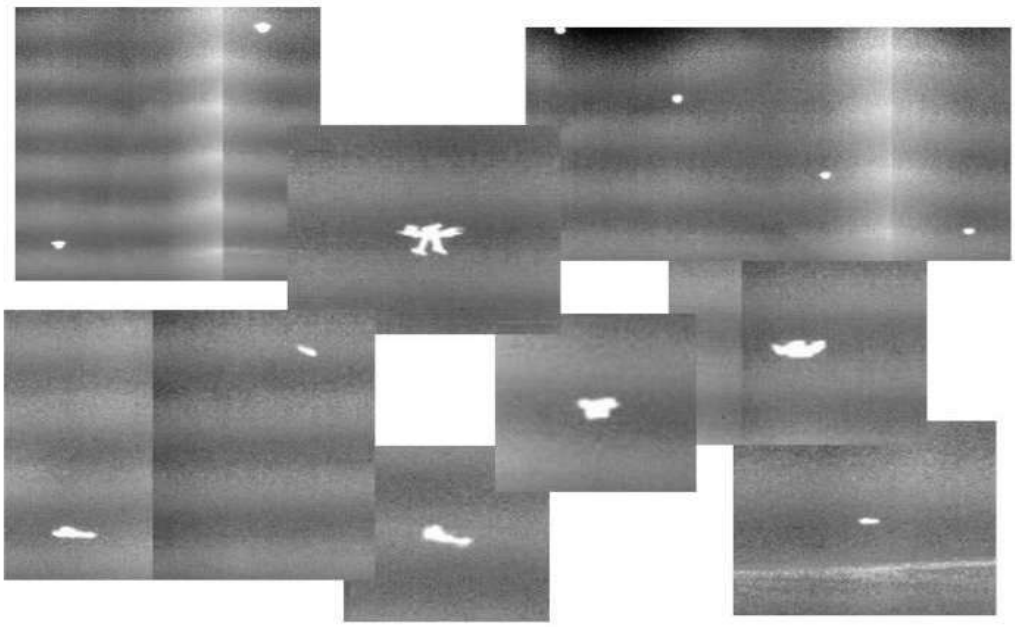


Рис. 2. Изображения БПЛА, полученные с помощью тепловизионной камеры

В [32, 84] исследовалась возможность обнаружения БПЛА при неблагоприятной фоновой обстановке – на фоне деревьев при колышущейся листве. В [33] выполнено сравнение эффективности обнаружения дронов с использованием тепловизионных камер различных участков инфракрасного диапазона – длинноволнового инфракрасного (LWIR), средневолнового (MWIR) и коротковолнового (SWIR). В [34] рассматривается возможность определения угловых координат БПЛА на плоскости и в пространстве при использовании нескольких тепловизионных видеокамер.

В целом, количество экспериментальных и теоретических работ, посвященных обнаружению БПЛА, с использованием тепловизионных видеокамер сравнительно невелико. Хотя возможности использования тепловизионных видеокамер для решения других прикладных задач – обнаружения пешеходов, животных, браконьеров, для распознавания лиц рассматриваются в литературе достаточно часто. Наиболее широко для обработки телевизионных изображений при решении различных задач используются сверточные нейронные сети CNN и методы глубокого обучения [35].

Акустический метод

Значительное внимание при решении задачи обнаружения и измерения координат БПЛА уделяется акустическому информационному каналу, развитию этого направления посвящено значительное число работ [35].

Структура и параметры акустического сигнала, излучаемого БПЛА, зависят от вида объекта, его формы, количества двигателей, количества несущих винтов и т.д. В свою очередь акустическая локационная станция (содар) должна строиться с учетом особенностей струк-

туры принимаемого сигнала. Исследованию особенностей акустического сигнала, формируемого и излучаемого БПЛА, посвящены работы [36 – 46].

Экспериментальные исследования структуры и параметров звукового поля БПЛА в виде квадрокоптера показали, что спектры его акустического излучения содержат ярко выраженные гармонические составляющие, имеющие частоты, кратные частоте вращения винта. Основной тон находится в полосе частот 80 – 240 Гц, а количество гармоник может быть от 10 до 40. Спектр сигнала простирается до частот более 10 – 12 КГц [36, 42].

В режиме полета спектральные линии акустического излучения квадрокоптера размываются вследствие различия режимов работы (частоты вращения) имеющихся четырех двигателей при компенсации автоматикой БПЛА воздействия дестабилизирующих факторов, возникающих в процессе полете. Это фактор может являться одним из информационных признаков классификации БПЛА среди других объектов. Расширение спектральных линий проявляется сильнее при увеличении номера гармоники [37, 45]. Указанные особенности акустического сигнала БПЛА наблюдаются на спектрограмме рис. 3 [42].

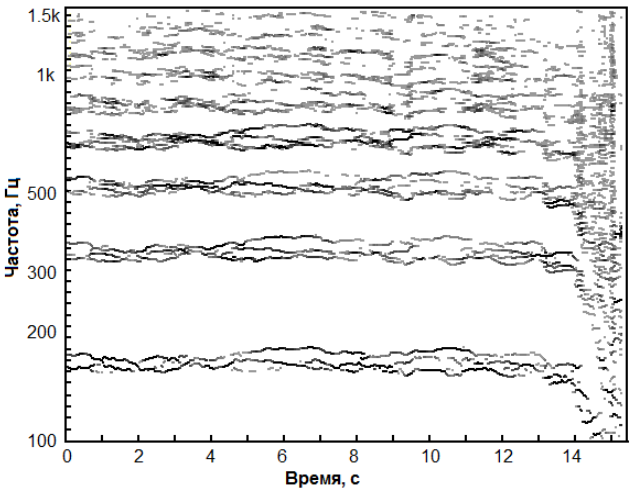


Рис. 3. Спектрограмма акустического сигнала квадрокоптера

Спектральные гармонические составляющие акустического сигнала квадрокоптера шире, чем у моноплана, что объясняется различием режимов работы двигателей в процессе полета или при отработке системой компенсации ветровых возмущений.

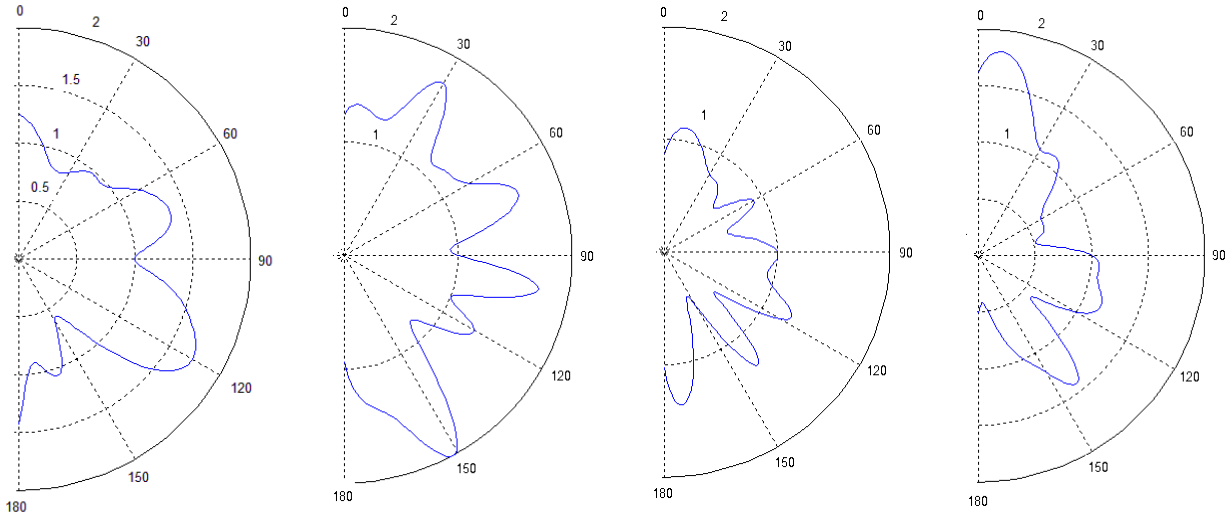
По мере увеличения расстояния, прошедшего акустической волной в атмосфере, происходят изменения в спектре акустического излучения (АИ), сопровождающиеся заметным ослаблением высокочастотных составляющих. Изменения формы спектров АИ БПЛА в реальных условиях наблюдения обусловлены дисперсионными свойствами среды, а также изменчивостью характеристик пространственной направленности излучения в полосе частот [42].

Большое значение для практики имеют диаграммы излучения БПЛА, характеризующие распределение излучаемой акустической энергии по направлениям. В ряде работ рассматривалась пространственная направленность звукового излучения БПЛА, в частности в [39] сделан вывод о том, в первом приближении БПЛА может считаться изотропным источником излучения.

В то же время эксперименты [42, 45, 82] показывают существенную направленность излучения как отдельных элементов конструкции аппарата – винтомоторной группы, электродвигателей квадрокоптера, так и всей конструкции в целом. Показано, что пространственные распределения как отдельных спектральных (гармонических) составляющих, так и полной энергии (во всем диапазоне частот), являются анизотропными.

Нормализованные характеристики пространственной направленности акустического излучения квадрокоптера DJI Phantom 3 в вертикальной плоскости, для первых четырех гармоник лопастной частоты воздушного винта, представлены на рис. 4. Анализ представленных

результатов показывает, что с повышением номера гармоники происходит усложнение формы характеристики направленности: она становится более изрезанной, увеличивается глубина провалов, уменьшается ширина лепестков и происходит изменение направления основного излучения.



а *б* *в* *г*

Рис. 4. Нормализованные характеристики направленности акустического излучения квадрокоптера DJI Phantom 3 в вертикальной плоскости на гармониках лопастной частоты винта: *а* – 1-я гармоника, *б* – 2-я гармоника, *в* – 3-я гармоника, *г* – 4-я гармоника

Как видно из рис. 4, различным ракурсам наблюдения БПЛА соответствуют различные уровни спектральных гармонических составляющих излучения, определяемых характеристиками направленности. Из этого следует, что интенсивность акустического излучения в зависимости от угла наблюдения должна описываться некоторым законом распределения вероятностей, а дальность обнаружения БПЛА является величиной статистической, зависящей от ракурса наблюдения.

Авторами [47 – 49] синтезированы и апробированы на практике алгоритмы обнаружения БПЛА по характерным особенностям спектра и распознавания на основании Mel-Frequency Cepstral Coefficients – MFCC. Основная трудность при решении задач обнаружения и распознавания заключается в различении информационного сигнала от помех, имеющих идентичные с АИ БПЛА спектральные особенности. К таким помехам следует отнести АИ автомобильных двигателей и звуки речи человека.

В публикациях [50 – 52] для обнаружения БПЛА авторы использовали записанные заранее акустические образы, соответствующие различным типам дронов, и вычисление корреляции между сигналами, поступающими на вход и имеющимися в базе. Алгоритм не позволил работать в режиме реального времени и имел ограниченную базу образов.

В [53] для обнаружения БПЛА использовались меп-кепстральные коэффициенты MFCC и алгоритм SVM. В [54] использована приемная акустическая система с одним микрофоном, в которой реализован алгоритм *k*-ближайших соседей, работающий по получаемым спектрам быстрого преобразования Фурье. В работе [55] эти же авторы повысили вероятность правильного решения с 0,83 до 0,86 путем использования нейронной сети и сформировали помеховые фоновые сигналы, соответствующие звукам различных источников, используя базу данных UrbanSound8K [56].

В [57] рассмотрен бинарный алгоритм распознавания, основанный на рекуррентной нейронной сети. Авторы также выполнили расширение набора входных данных, используемых при обработке, путем аддитивного суммирования сигналов дронов с различными фоновыми звуками.

В работе [58] использовалась комплексная система РЛС-содар, данные которой поступали в нейронную сеть с прямой связью.

Существует необходимость в формировании базы звуков БПЛА, которая могла бы быть использована разработчиками и научным сообществом при синтезе, анализе и верификации алгоритмов обнаружения-распознавания дронов.

Ряд других алгоритмов обработки акустических сигналов представлены в работах [53 – 57, 59 – 63].

В статье [64] для решения задачи обнаружения-распознавания БПЛА по акустическому сигналу используется составная модель авторегрессии, которая адекватно описывает корреляционные свойства сигнала на значительных временных интервалах, и обеспечивает повышение спектрального разрешения в области низкочастотных составляющих спектра исследуемого сигнала. Составная модель авторегрессии позволяет выделять низкочастотный пик акустического сигнала, наличие которого является эффективным информационным признаком БПЛА. На практике рекомендуется использовать шестой порядок составной модели авторегрессии.

Проведены экспериментальные исследования с использованием предложенной математической модели, которые показали существенные отличия спектральной плотности мощности (СПМ) АИ БПЛА от СПМ акустических шумов различных источников, что позволило повысить эффективность решения задачи обнаружения-распознавания БПЛА.

Показано, что в задаче распознавания бывает удобнее использовать информацию о частотах пиков спектра, а не сам спектр. Предложен алгоритм определения частот пиков СПМ на базе предложенной модели авторегрессии без вычисления спектра, который позволяет сравнительно просто рассчитывать параметрические оценки частот пиков СПМ и который целесообразно использовать в содарах обнаружения БПЛА при работе в реальном масштабе времени.

Использование методов классического спектрального анализа, для которых характерна работа с сигналами большой длительности, при обнаружении сигнала БПЛА связано с рядом затруднений [65]. Сигнал БПЛА представляет собой набор дискретных гармоник, параметры которых изменяются на сравнительно небольших интервалах времени, в частности, вследствие изменения режимов работы двигателей. Определенная таким образом спектральная плотность приближается по форме к непрерывной. Естественно, при этом теряются характерные особенности формы спектра БПЛА, составляющие основной набор информационных признаков при решении задачи обнаружения-распознавания. Анализ свойств сигнала на небольших промежутках времени позволит выявить его истинную структуру и эффективно решить задачу обнаружения БПЛА на фоне помех.

При обработке акустических сигналов БПЛА необходимо учитывать состояние атмосферы и особенности распространения колебаний в существующих условиях [66 – 69].

Ряд работ посвящен вопросам пеленгования БПЛА по их акустическому излучению, исследованию возможностей различных методов. В [70] показано, что при использовании классических методов для пеленгования БПЛА по их акустическому излучению процесс пеленгования возможен только при наличии в диаграмме направленности содара лишь одного объекта. Наличие в диаграмме направленности антенны (ДНА) нескольких объектов (например, роя БПЛА), приводит к появлению аномальных погрешностей результатов измерений угловых координат. Это обусловлено формированием амплитудно-фазового распределения звукового поля в раскрыве антенны путем суперпозиции волн, полученных от отдельных БПЛА. Вследствие небольших размеров антенной системы содаров обнаружения БПЛА угловое разрешение отдельных БПЛА в этом случае затруднительно.

Неадаптивный метод Бартлетта достаточно прост при реализации системы пеленгации БПЛА с использованием АР, но его угловое разрешение ограничено релеевским пределом, а для повышения пространственного разрешения необходимо увеличивать число элементов микрофонной решетки.

Более высокие характеристики пространственного разрешения источников акустического излучения обеспечивают методы сверхразрешения – Кэйпона, методы MUSIC, а также его разновидности ROOT-MUSIC и ESPRIT.

Недостаток данных методов заключается в формировании выборочных корреляционных матриц, а также выполнении сингулярного разложения или разложения по собственным векторам полученных матриц. При пеленговании малых БПЛА в условиях быстро меняющейся динамической обстановки, это условие накладывает определенные ограничения на временные и вычислительные ресурсы содара.

При использовании методов Бартлетта и Кейпона предполагается использование узкополосных сигналов. Вследствие этого при приеме и обработке широкополосного излучения БПЛА требуется некоторая адаптация данных алгоритмов.

Для определения направления прихода АИ широкополосных сигналов широко используется метод взаимной корреляционной функции (МВКФ) [71, 83, 85, 86]. Сдвиг времени прихода широкополосного акустического сигнала τ к отдельным микрофонам антенной решетки измеряется путем вычисления положения во времени максимумов ВКФ сигналов, принимаемых различными элементами антенной решетки. С целью повышения эффективности работы алгоритма МВКФ при обработке акустических сигналов БПЛА следует принимать меры для уменьшения влияния низкочастотных природных акустических шумов. Эксперименты по определению пеленга БПЛА в широкой полосе частот в открытом пространстве с использованием метода МВКФ показывают хорошее соответствие результатов эксперимента заданным значениям пеленга [72].

Анализ показывает, что сегодня не представляется возможным выделить среди алгоритмов пеленгования БПЛА наилучший, превосходящий все другие методы при различных условиях функционирования. Выбор метода пеленгования БПЛА при разработке содара следует производить с учетом имеющейся априорной информации о наблюдаемых объектах, существующей сигнально-помеховой обстановки, с учетом конфигурации микрофонной антенной решетки, количества элементов и ее геометрических размеров.

Выводы

Радиолокационные наблюдения БПЛА – активно развивающаяся область научных исследований и инженерных разработок. Радиолокационный метод – единственный активный метод (используется излучение зондирующего сигнала), применяемый при наблюдении за БПЛА. Основные показатели качества разрабатываемых и сравниваемых алгоритмов обнаружения БПЛА – вероятность ложной тревоги и вероятность правильного обнаружения. В соответствии с этим при обнаружении целесообразно использовать критерий Неймана – Пирсона, заключающийся, как известно, в максимизации вероятности правильного обнаружения при фиксированном значении вероятности ложной тревоги (порога обнаружения).

Сегодня развиваются два основных метода решения задачи обнаружения-распознавания БПЛА: метод формирования микродоплеровской сигнатуры – МДС и метод, основанный на анализе кинематики движения объекта.

Наиболее часто в литературе рассматривается первый метод, основанный на формировании МДС, параметрами которой являются количество лопастей БПЛА, ракурс облучения объекта, параметры зондирующего сигнала – частота излучения и период следования импульсов, а также время облучения цели. Несмотря на большое количество работ в данной области, все же отмечается недостаток в результатах натурных экспериментов, в результатах анализа этих экспериментов, а следовательно, неудовлетворенность в эффективности и надежности предлагаемых алгоритмов. В значительной части работ эксперименты проводились на небольших дальностях до БПЛА (порядка нескольких сотен метров), либо использовались данные моделирования. Это связано, видимо, с недоступностью соответствующих РЛС, вследствие их цены, для многих исследовательских коллективов, работающих, прежде всего,

в университетах. Отсутствуют и общедоступные наборы данных в Internet, которые могли быть использованы исследователями.

Менее развита область, направленная на обнаружение БПЛА по результатам зондирования, получаемым на выходе РЛС кругового обзора, которые описывают траекторию движения объекта. Перспективность данного направления связывают с возможностью развития методов, основанных на классификации траекторий с использованием методов глубокого обучения.

В целом, по оценкам экспертов, применение методов глубокого обучения к получаемым радиолокационным данным позволит существенно повысить эффективность решения задачи обнаружения-распознавания БПЛА.

Оптическим наблюдениям БПЛА с использованием изображений, получаемых стандартными оптическими камерами RGB, посвящено большое количество публикаций. Обработка изображений осуществляется традиционными методами компьютерного зрения с ручным формированием признаков, описывающих БПЛА, а также с использованием подходов, основанных на глубоком обучении. Вторая группа методов обеспечивает существенно лучшие результаты.

Основные показатели качества рассматриваемых алгоритмов – эффективность решения задачи обнаружения-распознавания (например, по критерию Неймана-Пирсона) и скорость выполнения алгоритмов (предъявляются требования по реализации алгоритмов в реальном масштабе времени).

Малые размеры БПЛА на изображении (малое число пикселей) и свойственные им траектории перемещения учитываются путем использования нескольких последовательных кадров или путем использования методов сверхразрешения.

Выполнено сравнение процедур Faster RCNN [73], SSP [74], YOLO [75]. Алгоритмы YOLO и SSP обеспечивают лучшее быстродействие, Faster RCNN – лучшее качество обнаружения. Известные методы не обеспечивают запросов практики, и требуется улучшать их показатели качества, в частности путем комбинирования и комплексирования известных алгоритмов.

Методы обнаружения БПЛА с использованием оптических камер рассматриваются на ряде Международных конференций, в том числе по системам компьютерного зрения.

Инфракрасные наблюдения, реализуемые с использованием тепловизионных камер, широко используются на практике в системах, предназначенных для обнаружения БПЛА, однако исследования и публикации по данной тематике практически не известны. Это связано, в первую очередь, с дороговизной тепловизионных камер, обладающих высоким разрешением.

Предполагается, что методы обнаружения-распознавания БПЛА, апробированные для видеокамер оптического диапазона, могут быть с успехом использованы применительно к тепловизионным датчикам [35].

Акустические наблюдения БПЛА широко используются на практике, особенно при обнаружении малых аппаратов на небольших расстояниях. Им также посвящено значительное число публикаций. Основным достоинством акустического канала обнаружения БПЛА является его относительная простота, демократичность и информативность.

Недостатки – значительный уровень акустических фоновых шумов, присутствующих практически повсеместно, и небольшая дальность действия содаров по обнаружению БПЛА (до 150 м.).

Перспективные направления улучшения характеристик содаров – использование современных методов обработки сигналов, в частности методов пространственно-временной обработки широкополосного акустического излучения БПЛА [76], а также комплексирование содаров с другими станциями.

Для создания надежных и эффективных алгоритмов для конкретных задач обнаружения БПЛА актуальна задача по созданию представительной общедоступной базы данных, вклю-

чающей звуковые сигналы различных БПЛА. Это может представлять собой серьезную задачу с учетом многообразия используемых дронов, необходимости записи сигналов дронов на различных удалениях, при различных частотах дискретизации, уровнях квантования и т.д.

Перспективным направлением исследований является использование комплексной обработки данных, получаемых с использованием рассмотренных информационных каналов, в единой интегрированной системе обнаружения и распознавания БПЛА [77, 78].

Список литературы:

1. Кошкин Р.П. Беспилотные авиационные системы. Москва : Стратегические приоритеты, 2016. 676 с.
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. 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, Vol. 78, Issue 9. P. 771-781.
4. Карташов В.М. и др. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : ХНУРЭ, 2014. 312 с.
5. Карташов В.М., Ситнік О.В. Радіотехнічні системи : навч. посібник. Харків : Сміт, 2009. 448 с.
6. De Wit J.M., Harmanny R., Premel-Cabic G. Micro-Doppler analysis of small UAVs // Proceedings of the 2012 9th European Radar Conference; Amsterdam, The Netherlands. 31 October–2 November 2012. P. 210–213.
7. Harmanny R., De Wit J., Cabic G.P. Radar micro-Doppler feature extraction using the spectrogram and the cepstrogram // Proceedings of the 2014 11th European Radar Conference; Cincinnati, OH, USA. 11–13 October 2014. P. 165–168.
8. Molchanov P., Harmanny R.I., de Wit J.J., Egiazarian K., Astola J. Classification of small UAVs and birds by micro-Doppler signatures // J. Microw. Wirel. Technol. 2014. 6:435–444.
9. De Wit J., Harmanny R., Molchanov P. Radar micro-Doppler feature extraction using the singular value decomposition // Proceedings of the 2014 International Radar Conference. Lille, France. 13–17 October 2014. P. 1–6.
10. Ren J., Jiang X. Regularized 2D complex-log spectral analysis and subspace reliability analysis of micro-Doppler signature for UAV detection // Pattern Recognit. 2017. 69:225–237.
11. 24. Oh B.S., Guo X., Wan F., Toh K.A., Lin Z. Micro-Doppler mini-UAV classification using empirical-mode decomposition features // IEEE Geosci. Remote Sens. Lett. 2017. 15:227–231.
12. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., Erhan D., Vanhoucke V., Rabinovich A. Going deeper with convolutions // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Boston, MA, USA. 7–12 June 2015. P. 1–9.
13. Mendis G.J., Randeny T., Wei J., Madanayake A. Deep learning based doppler radar for micro UAS detection and classification // Proceedings of the MILCOM 2016-2016 IEEE Military Communications Conference. Baltimore, MD, USA. 1–3 November 2016. P. 924–929.
14. Freund Y., Schapire R.E. Large margin classification using the perceptron algorithm. Mach. Learn. 1999. 37:277–296.
15. Simonyan K., Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv. 20141409.1556.
16. Zeiler M.D., Fergus R. Visualizing and understanding convolutional networks // Proceedings of the European Conference on Computer Vision; Zurich, Switzerland. 6–12 September 2014. P. 818–833.
17. StrelkovaT., KartashovV., Lytyuga A.,StrelkovA. Theoretical Methods of Images Processing in Optoelectronic Systems. Chapter 6 // Developing and Applying Optoelectronics in Machine Vision. Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez. (341p.) USA, Herhey, IGI Global, 2016. P.180-205.
18. StrelkovaT., KartashovV., Lytyuga A.,Strelkov A. Theoretical Methods of Images Processing in Optoelectronic Systems. Chapter 16 // Biometrics: Concepts, Methodologies, Tools, and Applications; Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez. (341p.), IGI Global, 2017. P. 361-381.
19. Developing and Applying Optoelectronics in Machine Vision / O. Sergiyenko, J.C. Rodriguez-Quiñonez // IGI Global, 2016. 341p.
20. Schumann A., Sommer L., Klatte J., Schuchert T., Beyerer J. Deep cross-domain flying object classification for robust UAV detection // Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Lecce, Italy. 29 August–1 September 2017. P. 1–6.
21. Craye C., Ardjoune S. Spatio-temporal Semantic Segmentation for Drone Detection // Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); Taiwan, China. 18–21 September 2019.
22. Vasileios Magoulianitis D.A., Anastasios Dimou D.Z., Daras P. Does Deep Super-Resolution Enhance UAV Detection // Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); Taiwan, China. 18–21 September 2019.

23. Opromolla R., Fasano G., Accardo D. A Vision-Based Approach to UAV Detection and Tracking in Cooperative Applications. Sensors. 2018. 8:3391.

24. Gökçe F., Üçoluk G., Şahin E., Kalkan S. Vision-based detection and distance estimation of microunmanned aerial vehicles. Sensors. 2015. 15:23805–23846.

25. Ivanov M., Sergiyenko O., Mercorelli P., Tyrsa V., Kartashov V., Hernandez W., Sheiko S., Kolendovska M. Individual scans fusion in virtual knowledge base for navigation of mobile robotic group with 3D TVS // Proceedings of 44th Annual Conference of IEEE Industrial Electronics Society (IECON), Washington DC, USA. 2018. P. 3187-3192.

26. Ivanov M., Sergiyenko O., Mercorelli P., Hernandez W., Rodriguez Quinonez J.C., Kartashov V., Kolendovska M., Iryna T. Effective informational entropy reduction in multi-robot systems based on real-time TVS // IEEE International Symposium on Industrial Electronics, June, 8781209, 2019. P. 1162–1167.

27. Oleksandr Sotnikov, Vladimir Kartashov, Oleksandr Tymochko, Oleg Sergiyenko, Vera Tyrsa, Paolo Mercorelli, Wendy Flores-Fuentes. Methods for Ensuring the Accuracy of Radiometric and Optoelectronic Navigation Systems of Flying Robots in a Developed Infrastructure. Chapter 16 // Machine Vision and Navigation; Springer, Cham. P.537–578. Editors: Sergiyenko**,** Oleg, Flores-Fuentes**,** Wendy**,** Mercorelli**,** Paolo.

28. Lindner L., Sergiyenko O., Rivas-López M., Gurko A., Kartashov V.M. Machine vision system for UAV navigation // IEEE, 2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles and International Transportation Electrification Conference, ESARS-ITEC, 2016. P.1–6.

29. Kartashov V., Oleynikov V., Zubkov O., Sheiko S. 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), 9–13 September 2019, Odessa, Ukraine, 4 p.

30. Aker C., Kalkan S. Using deep networks for drone detection // Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); Lecce, Italy. 29 August–1 September 2017. P. 1–6.

31. 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 // Радиотехника. 2020. Вып. 201. С. 120–129.

32. Müller T. Robust drone detection for day/night counter-UAV with static VIS and SWIR cameras // Proceedings of the Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR VIII, International Society for Optics and Photonics; Anaheim, CA, USA. 4 May 2017. P. 1019018.

33. Birch G.C., Woo B.L. Counter unmanned aerial systems testing: Evaluation of VIS SWIR MWIR and LWIR passive imagers. Sandia Rep. 2017.

34. Thomas A., Cotinat A., Gilber M. UAV localization using panoramic thermal cameras // Proceedings of the 12th International Conference on Computer Vision Systems (ICVS); Thessaloniki, Greece. 23–25 September 2019.

35. Samaras S., Diamantidou E., Ataloglou D., Sakellariou N., Vafeiadis A., Magoulianitis V., Lalas A., Dimou A., Zarpalas D., Votis K., Daras P., Tzovaras D. Deep Learning on Multi Sensor Data for Counter UAV Applications // A Systematic Review. Sensors (Basel). 2019 Nov 6. 19(22): 4837. Published online 2019, Nov 6.

36. Massey K., Gaeta R. Noise Measurements of Tactical UAVs. // Georgia Inst. of Technology / GTRI / ATAS, Atlanta. 16th AIAA / CEAS Aeroacoustics Conference. American Institute of Aeronautics and Astronautics, 2010. P. 1–16.

37. 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.

38. Zaslavsky Yu. M., Zaslavsky V. Yu. Acoustic noise of a low flying quadrocopter // NOUSE Theory and Practice. V.5, №3, 2019. P. 21–27.

39. Sinibaldi G., Marino L. Experimental analysis on the noise of the propellers for small UAV // Applied Acoustics, 74 (2013). P. 79–88.

40. Intaratep N., Alexande W. N., Devenport W. J., Grace S. M., Dropkin A. 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.

41. Moshkov P. M., Samokhin V. F. Assessment of the influence of the number of blades and diameter on the noise of the propeller // Vestnik Samarskogo universiteta. Aerokosmicheskaya tekhnika, tekhnologii i mashinostroyeniye, V. 15, No 3, 2016. P. 25–34. (In Rus.).

42. Карташов В.М., Олейников В.Н., Шейко С.А., Бабкин С.И., Корытцев И.В., Зубков О.В., Анохин М.А. Информационные характеристики звукового излучения малых беспилотных летательных аппаратов // Радиотехника. 2017. Вып. 191. С. 181-187.

43. Kartashov V.M., Tikhonov V.A., Voronin V.V. and Tymoshenko L.P. Complex model of random signal in problems of acoustic sounding of atmosphere // Telecommunications and Radio Engineering. 2016. V. 75. Iss. 20. P.1885–1892.

44. Карташов В.М., Харченко О.И., Чумаков В.И. Использование эффекта стохастического резонанса для анализа спектров акустического излучения малых беспилотных летательных аппаратов // Радиотехника. 2019. Вып. 197. С. 100-106.

45. 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. **V**.77 (10). P. 915–924.

46. Карташов В.М., Тихонов В.А., Воронин В.В., Тимошенко Л.П. Комплексные модели случайных сигналов в задачах акустического зондирования атмосферы // Радиотехника. 2016. Вып. 185. С. 81–86.

47. 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. V. 78, Issue 9. P. 759–770.

48. 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.

49. 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), 25-29 Feb. 2020. P.1–4.

50. Chowdhury A.S.K. Master's Thesis. University of Nevada. Las Vegas, NV, USA: 2016. Implementation and Performance Evaluation of Acoustic Denoising Algorithms for UAV.

51. Mezei J., Molnár A. Drone sound detection by correlation // Proceedings of the 2016 IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI); Timisoara, Romania. 12–14 May 2016. P. 509–518.

52. Bernardini A., Mangiatordi F., Pallotti E., Capodiferro L. Drone detection by acoustic signature identification // Electron. Imaging. 2017. P.60–64.

53. Liu H., Wei Z., Chen Y., Pan J., Lin L., Ren Y. Drone detection based on an audio-assisted camera array // Proceedings of the 2017 IEEE Third International Conference on Multimedia Big Data (BigMM); Laguna Hills, CA, USA. 19–21 April 2017. P. 402–406.

54. Kim J., Park C., Ahn J., Ko Y., Park J., Gallagher J.C. Real-time UAV sound detection and analysis system // Proceedings of the 2017 IEEE Sensors Applications Symposium (SAS); Glassboro, NJ, USA. 13–15 March 2017. P. 1–5.

55. Kim J., Kim D. Neural Network based Real-time UAV Detection and Analysis by Sound // J. Adv. Inf. Technol. Converg. 2018. 8:43–52.

56. Salamon J., Jacoby C., Bello J.P. A dataset and taxonomy for urban sound research // Proceedings of the 22nd ACM International Conference on Multimedia. ACM; Mountain View, CA, USA. 18–19 June 2014. P. 1041–1044.

57. 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.

58. Park S., Shin S., Kim Y., Matson E.T., Lee K., Kolodzy P.J., Slater J.C., Scherreik M., Sam M., Gallagher J.C., et al. Combination of radar and audio sensors for identification of rotor-type unmanned aerial vehicles (uavs) // Proceedings of the 2015 IEEE SENSORS; Busan, Korea. 1–4 November 2015. P. 1–4.

59. A. Bernardini, F. Mangiatordi, E. Pallotti, L. Capodiferro, F. Ugo Bordoni. Drone detection by acoustic signature identification // Electronic Imaging, Imaging and Multimedia Analytics in a Web and Mobile World. 2017. P. 60–64.

60. Vasilchenko A., Kartashov V.M. Analysis of influence exerted by longitudinal Doppler effect upon output signal of sodar antenna array // Telecommunications and Radio Engineering. Vol. 66, Issue 9. P. 841–847.

61. Zelnio A.M. Detection of small aircraft using an acoustic array // Electrical Engineering.Wright State University, 2007. 55 p.

62. Kozeruk S. A., Korzhyk A.V. Identification of small aircraft by acoustic radiation // Visnyk NTUU KPI. Series Radiotekhnika Radiobuduvannia. 2019. Iss. 76. P. 15–20.

63. Sadasivan S., Gurubasavaraj M., Sekar S.R. Acoustis siqnature of an unmanned air vehicle exploitation for aircraft localisation and parameter estimation // Eronautical DEF SCI J. 2001 Vol. 51 №3. pp. 279–283.

64. Тихонов В.А., Карташов В.М., Олейников В.М., Леонидов В.И., Тимошенко Л.П., Селезнев И.С., Рыбников Н.В. Обнаружение-распознавание беспилотных летательных аппаратов с использованием составной модели авторегрессии их акустического излучения // Вісник НТУУ «КПІ». Радіотехніка. Радіоапаратобудування. 2020. Вип. №81. С. 38–46.

65. Semenets V. V., Kartashov V.M., Leonidov V. I. Registration of refraction Phenomenon in the Problem of acoustic Sounding of Atmosphere in Airport Zone // Telecommunications and Radio Engineering. 2018. Vol. 77, Iss. 5. P.461–468.

66. Карташов В.М., Тихонов В.А., Воронин В.В. Особенности построения и применения комплексних систем дистанционного зондирования атмосферы // Радиотехника. 2016. Вып. 186. С. 184-185.

67. Карташов В.М., Куля Д.Н., Кушнир М.В., Толстых Е.Г. Выбор модели изменения скорости звука для оптимального линейного фильтра систем радиоакустического зондирования атмосферы // Радиотехника. 2013. №173.С. 63–78.

68. Дистанционные методы и средства исследования процессов в атмосфере Земли ; под ред. Б.Л. Кащеева, Е.Г. Прошкина, М.Ф. Лагутина. Харьков : Бизнес Информ, 2002. 426 с.

69. Карташов В.М., Куля Д.Н., Пащенко С.В. Алгоритм автосопровождения изменений информационного параметра сигнала радиоакустических систем // Восточно-европейский журнал передовых технологий. 2012. №4/9(58). С. 57-61.

70. Олейников В.Н., Зубков О.В, Карташов В.М., Корытцев И.В., Бабкин С.И., Шейко С.А., Селезнев И.С. Экспериментальная оценка эффективности алгоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению // Радиотехника. 2019. Вып. 199. С. 29–37.

71. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И., Левский Н.А., Селезнев И.С. Алгоритмы пеленгации беспилотных летательных аппаратов по их акустическому излучению // Радиотехника. 2019. Вып. 196. С. 22–31.

72. 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.

73. Saqib M., Khan S.D., Sharma N., Blumenstein M. A study on detecting drones using deep convolutional neural networks // Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). Lecce, Italy. 29 August–1 September 2017.

74. Mrunalini Nalamati A.K., Muhammed Saqib N.S., Blumenstein M. Drone Detection in Long-range Surveillance Videos // Proceedings of the 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); Taiwan, China. 18–21 September 2019.

75. Aker C., Kalkan S. Using deep networks for drone detection // Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS); Lecce, Italy. 29 August–1 September 2017. P. 1–6.

76. 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. 2020. V. 79, №9. P.769–780.

77. Карташов В.М., Олейников В.Н., Воронин В.В., Рябуха В.П., Капуста А.И., Рыбников Н.В., Селезнев И.С. Методы комплексной обработки и интерпретации радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов // Радиотехника. 2020. Вып. 202. С. 173-182.

78. Карташов В.М., Олейников В.Н., Леонидов В.И., Воронин В.В., Капуста А.И., Селезнев И.С., Першин Е.В. Комплексная обработка сигналов интегрированной системы наблюдения беспилотных летательных аппаратов с использованием целеуказания // Радіотехніка. 2020. Вип. 203. С. 148-161.

79. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И. Обработка сигналов при пеленгации и определении дальности до малоразмерных БПЛА в оптическом и инфракрасном диапазонах // Радиотехника. 2020. Вып. 202. С. 125-135.

80. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И. Оптико-электронные методы обнаружения воздушных объектов и измерения их координат // Радиотехника. 2020. Вып. 202. С. 153-159.

81. Карташов В.М., Корытцев И.В., Олейников В.Н., Зубков О.В., Шейко С.А., Бабкин С.И. Эффективность детектирования и распознавания изображений дронов по видеопотоку стационарной видеокамеры // Радиотехника. 2020. Вып. 202. С. 136-146.

82. Semenets V.M., Kartashov V.M., Leonidov V.I. Features of Acoustic Noise of Small Unmanned Aerial Vehicles // Telecommunications and Radio Engineering. New York. 2020. Vol. 79, №11. P. 985-995.

83. Kartashov V., Oleynikov V., Koryttsev I., Sheiko S., Zubkov O., Babkin S.. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 25, 2020. Vol. 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). P. 35-39.

84. Koryttsev I., Sheiko S., Kartashov V., Zubkov O., Oleynikov V., Anohin M., Selieznov I. Practical Aspects of Range Determination and Tracking of Small Drones by Their Video Observation // 2020 International Scientific-Practical Conference. Problems of Infocommunications. Science and Technology. Kharkiv, Ukraine. October 6-9, 2020. 5 p.

85. Олейников В.Н., Зубков О.В., Карташов В.М., Шейко С.А., Бабкин С.И., Корытцев И.В. Исследование эффективности обнаружения и распознавания малоразмерных беспилотных летательных аппаратов по их акустическому излучению // Радиотехника. 2018. Вып. 195. С. 209-217.

86. 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.

Поступила в редколлегию 02.04.2021

Сведения об авторах:

Карташов Владимир Михайлович – д-р техн. наук, профессор, Харьковский национальный университет радиоэлектроники, заведующий кафедрой медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail:volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373

Посошенко Виталий Александрович – канд. техн. наук, доцент, Харьковский национальный университет радиоэлектроники, доцент кафедры медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail: vitalii.pososhenko@nure.ua; ORCID: https://orcid.org/0000-0003-0867-9161

Воронин Виталий Валериевич – канд. техн. наук, Светловодский политехнический коледж Центральноукраинского национального технического университета, преподаватель радиотехнических дисциплин; Украина; e-mail:vvvoronin2016@gmail.com; ORCID: https://orcid.org/0000-0002-4495-9024

Колесник Виктория Ивановна – Харьковский национальный университет радиоэлектроники, аспирант кафедры медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail: valeriia.kolesnyk@nure.ua

Капуста Анастасия Игоревна – Харьковский национальный университет радиоэлектроники, аспирант кафедры медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail: anastasiia.kapusta@nure.ua; ORCID: https://orcid.org/0000-0003-2206-1552

Рыбников Николай Владимирович – Харьковский национальный университет радиоэлектроники, аспирант кафедры медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail: mykola.rybnykov@nure.ua; ORCID: https://orcid.org/0000-0003-1340-8788

Першин Евгений Васильевич – Харьковский национальный университет радиоэлектроники, аспирант кафедры медиаинженерии и информационных радиоэлектронных систем, факультет информационных радиотехнологий и технической защиты информации; Украина; e-mail:yevhenii.pershyn@nure.ua; ORCID: https://orcid.org/0000-0002-4573-9381

УДК 621.396.96 **DOI:10.30837/rt.2021.2.205.16**

І.В. СВИД, канд. техн. наук, І.І. ОБОД, д-р техн. наук, О.С. МАЛЬЦЕВ,
М.Г. ТКАЧ, С.В. СТАРОКОЖЕВ, А.О. ГЛУЩЕНКО, В.С. ЧУМАК

МЕТОД ПІДВИЩЕННЯ ЗАВАДОЗАХИЩЕНОСТІ РАДІОЛОКАЦІЙНИХ СИСТЕМ ІДЕНТИФІКАЦІЇ «СВІЙ-ЧУЖИЙ» ПРИ ДІЇ НАВМИСНИХ КОРЕЛЬОВАНИХ ЗАВАД

Вступ

До основних елементів процедури контролю повітряного простору відносяться аналіз повітряної обстановки та прийняття рішень. Слід зазначити, що рішення приймається на основі аналізу відповідним чином підготовленої інформації про стан повітряної обстановки. Вірне рішення може бути прийнято лише тоді, коли є досить повна, точна, достовірна та безперервна інформація про повітряну обстановку в зоні відповідальності. Таким чином, можна стверджувати, що якість прийняття рішень визначається складом та достовірністю інформації, на основі якої уповноважена особа приймає рішення. Достовірність інформації значною мірою визначається інформаційними ресурсами системи контролю повітряного простору, до яких відносяться радіолокаційні системи спостереження [1 – 3]. Радіолокаційні системи спостереження в системі контролю повітряного простору підрозділяються на первинні та вторинні. Первинні радіолокаційні системи спостереження визначають координати виявленого повітряного об’єкта, а вторинні – оцінюють виявлений повітряний об’єкт за ознакою «свій-чужий», тобто є системами ідентифікації (IFF) [4, 5]. Однак, як показано в [6 – 11], системи IFF, побудовані так, що зацікавлена сторона може несанкціоновано використати цей інформаційний ресурс для подальшого визначення координат повітряних об’єктів, з одного боку, та перекручування інформації цього інформаційного ресурсу, з другого боку, що призводить до жахливих результатів [12, 13].

Система IFF забезпечує спостереження за повітряними об'єктами, обладнаними літаковими відповідачами, і забезпечує двосторонній зв'язок за каналом передачі даних між наземними станціями і повітряними об’єктами [14 – 16]. Система IFF відноситься до основних інформаційних джерел як системи контролю повітряного простору, так і системи управління повітряним рухом. Система IFF повинна вирішувати завдання IFF повітряного об'єкту за ознакою «свій-чужий» як в інтересах визначення ступеня небезпеки виявленого повітряного об'єкта, так і при безпосередньому застосуванні зброї. Рішення завдання радіолокаційної IFF за ознакою «свій-чужий» полягає в ухваленні рішення про виявлення повітряного об'єкта системою IFF «свій-чужий». Імітостійка (криптографічна) ідентифікація повітряних об'єктів, що реалізована в існуючих системах IFF, дозволяє однозначно вирішити питання за ознакою «свій-чужий» і є важливою умовою функціонування єдиного інформаційно-комунікаційного простору. Найбільш вразливим місцем в системах IFF, що істотно обмежує завадостійкість та завадозахищеність IFF повітряних об'єктів, є літаковий відповідач [17, 18]. Він побудований за принципом відкритої одноканальної системи масового обслуговування з відмовами, що викликає труднощі при роботі останніх при значних щільностях потоків внутрісистемних завад [19]. Така побудова літакового відповідача викликає суттєві недоліки в безпеці як його, так і безпеки всієї системи IFF. Використання ж єдиної частоти у запитальному каналі систем, що розглядаються, призводить до високої щільності сигналів запиту і, як наслідок, до внутрісистемних завад [20] значної інтенсивності. Зазначені фактори призводять до зниження якості обробки сигнальних даних. Так, в роботі [21] наводиться характеристика середовища щодо оцінки характеристик сучасних вторинних радіолокаційних приймачів спостереження. Основна увага приділяється параметрам, що дають точну характеристику явищ завад, які суттєво обмежують продуктивність даної системи. В роботах [22 – 25] розглядаються питання оптимального виявлення сигналів запиту при однакових рівнях як сигналів запиту, так і завад, які надходять на літаковий відповідач, що представляє собою ідеальний випадок.

Побудова літакового відповідача за принципом одноканальної системи обслуговування сигналів запиту з відмовами визначила значну часову паралізацію літакового відповідача на час обслуговування попереднього сигналу відповіді, що призводить до суттєвих обмежень як до відносної пропускної здатності літакового відповідача, так і до суттєвого зниження як завадостійкості, так і завадозахищеності літакових відповідачів та і усієї системи IFF.

Метод підвищення завадозахищеності радіолокаційних систем ідентифікації повітряних об'єктів за ознакою «свій-чужий»

Системи IFF призначені для вирішення наступних задач:

- визначення координат повітряного об'єкту;
- радіолокаційної ідентифікації повітряного об'єкту за ознакою «свій-чужий»;
- отримання польотної інформації, яка необхідна для контролю управління польотами та наведення повітряного об'єкту;
- диспетчерської ідентифікації повітряного об'єкту.

Виходячи з принципу функціонування систем IFF, антена відповідача є слабкоспрямованою. Це вносить суттєві недоліки в процес функціонування інформаційних систем, що розглядаються

Імовірність завадозахищеності систем IFF може бути визначена так:

$$P_{zz} = 1 - P_v P_{vim} P_{pr}, \tag{1}$$

де P_v – імовірність виявлення сигналів відповіді; . P_{vim} – імовірність вимірювання параметрів сигналів; P_{pr} – імовірність порушення роботи; $P_v P_{vim}$ - скритність; P_{pr} – завадостійкість; $P_{skr} = 1 - P_v$ – імовірність скритної роботи.

Особливістю систем IFF є наявність внутрісистемних завад значної інтенсивності [11]. Дійсно, сусідні системи IFF є джерелами завад для системи IFF, що розглядається. Ці завади проявляються двояко. По-перше, сигнали запиту сусідніх систем IFF закривають літакового відповідача на час обслуговування сигналів запиту, що унеможливлює відповідь на сигнали запиту системи IFF, що розглядається. По-друге, випромінені сигнали відповіді літакового відповідача на запити сусідніх систем IFF є завадою для системи IFF, що розглядається. Ці обставини потребують використовувати несинхронну мережу систем IFF. При такій організації мережі ефективним способом захисту систем IFF від сигнали відповіді, викликаних сусідніми системами IFF, є міжперіодна обробка сигналів. Однак побудова систем IFF на принципі несинхронної мережі та відсутність просторової вибірковості літакового відповідача дозволяє зацікавленій стороні або несанкціоновано отримувати інформацію від літакового відповідача, або подавляти систему IFF шляхом постановки навмисних корельованих завад необхідної інтенсивності.

В роботі [12] показано, що без виключення з обслуговування навмисних корельованих завад, якими є імітовані зацікавленою стороною сигнали запиту, досягнути прийнятних показників завадостійкості систем IFF неможливо. При цьому слід зазначити, що наявність навмисних корельованих завад ускладнює рішення задач як виміру координат повітряного об'єкту, так і передачі польотної інформації.

Відомо, що основою для придушення завад є різниця між корисним сигналом та завадою. В існуючих систем IFF реалізовано принцип обслуговування заявки (сигналу запиту), що визначило реалізацію принципу відкритих одноканальних систем масового обслуговування з відмовами при їх побудові. Сама ж мережа систем IFF реалізована на несинхронному принципі. Несинхронна мережа систем IFF дозволяє ефективно подавляти в апаратурі запитувача внутрішньосистемні завади, які утворені сусідніми системами IFF. Однак саме це дозволяє зацікавленій стороні здійснювати паралізацію систем IFF постановкою навмисними корельованих завад. Отже, така реалізація сучасних систем IFF на принципах відкритої

одноканальної системи масового обслуговування з відмовами ускладнює їх використання в конфліктних ситуаціях.

Дійсно, інтенсивність потоку сигналів запиту в існуючих системах IFF можна визначити з виразу

$$\lambda_c = \sum_{i=0}^{N-1} \lambda_i(T_i) + \lambda_1 + \sum_{j=0}^{M-1} \lambda_j(T_j), \qquad (2)$$

де $\lambda_i(T_i)$ – інтенсивність потоку сигналів запиту от *i*-го запитувача з періодом повторення T_i; λ_1 – інтенсивність потоку хибних сигналів запиту, які утворені з хаотичних імпульсних завад та сумарного потоку сигналів запиту своїх запитувачів та хаотичних імпульсних завад (тобто за рахунок хибної тривоги першого та другого роду); $\lambda_j(T_j)$ – інтенсивність потоку навмисних корельованих завад запитувачами зацікавленої сторони, що пригнічують та несанкціоноване використовують літаковий відповідач, з періодом проходження T_j. Так як в існуючих системах IFF до обслуговування приймаються всі правильно дешифровані сигнали запиту, то це дозволяє зацікавленій стороні пригнічувати системи IFF постановкою навмисних корельованих завад потрібної інтенсивності.

Таким чином, як випливає з принципу обслуговування, побудови та організації мережі, в сучасних системах систем IFF відсутні просторові та часові відмінності між сигналами і навмисними корельованими завадами, що ускладнює захист інформації зазначених інформаційних ресурсів.

Пошук шляхів спадкоємного переходу до систем IFF з захистом інформації [21] призводить до необхідності створення відмінностей між корисними сигналами та навмисними корельованими завадами. Створення просторових відмінностей, хоча і можливо, проте призводить до значних матеріальних витрат і до складності функціонування таких систем. Іншим методом створення відмінностей між корисними сигналами і навмисними корельованими завадами є часова різниця. Саме часовим розбіжностям приділяється основна увага.

Пошук часових відмінностей між корисними сигналами і навмисними корельованими завадами призводить до зміни принципу організації мережі систем IFF. Перехід від несинхронної до синхронної мережі систем IFF дозволяє штучно створити часові відмінності між корисними сигналами та завадами. При реалізації синхронних мереж систем IFF сумарний потік сигналів запиту можна записати як

$$\lambda_c = \sum_{i=0}^{N-1} \lambda_i(T_0(t)) + \lambda_1 + \sum_{j=0}^{M-1} \lambda_j(T_j), \qquad (3)$$

де $T_0(t)$ – період повторення сигналів запиту, єдиний для всіх синхронних мереж систем IFF. Як випливає з (3), часові різниці між корисними сигналами і навмисними корельованими завадами проявляються в часі надходження. Дійсно, так як шкала часу літакового відповідача узгоджена зі шкалами часу всіх елементів синхронних мереж систем IFF, то корисні сигнали запиту надходять на відповідач в синхронні, а навмисними корельованими завадами – в несинхронні моменти часу.

Таким чином, перехід до синхронних мереж систем IFF дозволяє перевести навмисні корельовані завади в несинхронну заваду, методи захисту від якої достатньо вивчені. Зокрема, одним з найбільш ефективних методів захисту від несинхронних імпульсних завад є міжперіодна обробка сигналів.

Перехід до синхронних мереж систем IFF дозволяє розширити методи обслуговування заявок та методи побудови систем. Розглянемо більш докладно ці принципи.

На рис. 1 наведено метод спадкоємного переходу до завадостійких систем IFF на спадкоємній реалізації синхронної мережі.

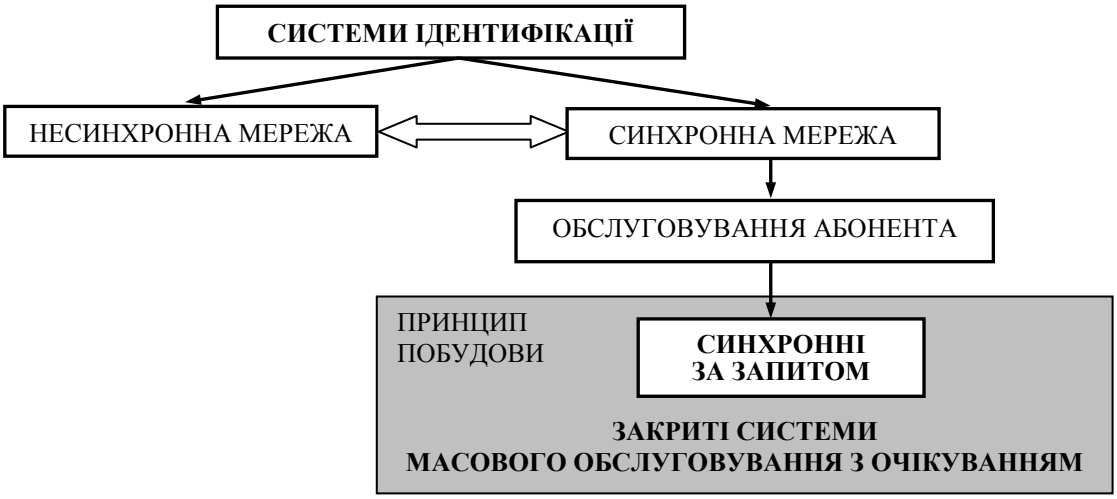


Рис. 1. Метод спадкоємного переходу до завадостійких систем IFF

При такій реалізації синхронної мережі реалізується принцип обслуговування абонента на основі виключення з обслуговування сигналів запиту,які є несинхронними до реалізованої мережі. Це дозволило реалізувати закриту систему масового обслуговування з очікуванням та здійснити спадкоємний перехід до завадостійких систем IFF.

Таким чином, за принципом побудови мережі систем IFF діляться на синхронні і несинхронні. Сучасні систем IFF побудовані за принципом несинхронної мережі. Це зумовлено необхідністю захисту запитувачів від внутрісистемних завад у каналі відповіді. Перехід до синхронної мережі систем IFF дозволяє створити часові відмінності між корисними і навмисними корельованими сигналами запиту зацікавленої сторони. Зауважимо, що при модернізації та розробці систем IFF повинен дотримуватися принцип спадкоємності. З цією метою нами і підкреслюється, що перехід до синхронних мереж має бути здійснений так, що в будь-який момент часу може відбутись зворотний перехід.

В синхронній мережі систем IFF робота всіх елементів системи реалізується на єдиному часі мережі. Кожна з вхідних в синхронній мережі систем IFF може бути побудована за принципом обслуговування абонента (запитувача) або обслуговування мережі (всіх запитувачів). У системах IFF, що реалізують перший принцип літакового відповідача, обслуговується конкретний запитувач, а в системах IFF, реалізованих на другому принципі, – обслуговується вся мережа систем IFF.

Основним методом побудови систем IFF з обслуговуванням абонента є запитальні системи IFF, виконані на базі синхронної мережі, де кожен запитувач формує сигнал запиту в строго певний час мережі систем IFF. На літакового відповідача також формується шкала часу, синхронна з єдиною шкалою часу мережі. Отже, в таких систем IFF вдається виключити можливість обслуговування навмисних корельованих завад літаковим відповідачем, а також значно знизити інтенсивність потоку сигналів запиту.

Реалізація синхронного за запитом методу на базі синхронних мережах може наштовхнутися на складності боротьби з синхронними внутрішньосистемними завадами, зумовленими роботою сусідніх запитувачів. Однак цей недолік усувається, як правило, спільною роботою систем первинної та вторинної радіолокації, а також переходом до зміни принципу обслуговування мережі систем IFF.

У запитувальних системах IFF, реалізованих на базі синхронної мережі, можливо управління потоками сигналів запиту. Наявність у літакового відповідача можливості проведення міжперіодної обробки дозволяє стверджувати, що на обслуговування цієї системи IFF надходять тільки сигнали запиту синхронних мереж. Таким чином, сумарний потік сигналів запиту в синхронних мережах запитувальних систем IFF можна оцінити із наступного виразу

$$\lambda_0 = P_0 \lambda_c \left(T_0(t) \right) . \quad (4)$$

Як випливає з викладеного, управління потоками сигналів запиту в синхронних мережах систем IFF не тільки знижує загальну їх інтенсивність, але і унеможливлює несанкціоноване використання літакового відповідача систем IFF. Усе це дозволяє значно підвищити якість роботи систем IFF та перейти до використання у якості сигналів запиту сигналів з великою базою. Необхідно відмітити, що використання синхронних мереж запитальних систем IFF дозволяє перейти від систем масового обслуговування з відмовами до систем масового обслуговування з очікуванням. В цьому випадку за час спостереження повітряного об'єкту T_n літакового відповідача може обслужити $N =]T_n / T_0 \cdot k[$ запитувачів, де k – необхідне число відповідей конкретному запитувачу, яке потрібне для виконання критерію початку пачки.

Основним методом побудови систем IFF з обслуговуванням мережі систем IFF є запитувальні системи IFF з синхронною відповіддю.

Суть цього методу полягає в наступному. Принцип побудови і функціонування каналу запиту не змінюється, тобто залишається таким, як і при використанні на несинхронній мережі. Змінюється принцип обслуговування запиту (заявки). Є час аналізу T_a, впродовж якого приймаються сигнали запиту літакового відповідача. Сигнали запиту не випромінюються відразу після прийому та дешифрування, а тільки в певний момент, відомий споживачам, що входять в інформаційну мережу. Для випромінювання сигналів запиту необхідно отримати як мінімум один сигнал запиту впродовж інтервалу аналізу.

Час обслуговування усіх заявок, що поступили на інтервалі часу аналізу T_a, вибирається з умови $T_{об} = T_{об} - t_p(t)$, де $t_p(t)$ – відомий, часовий інтервал, який постійно змінюється. Таким чином, в таких системах незалежно від числа заявок на часовому інтервалі аналізу, вони обслуговуються одночасно у момент часу, наведеному вище. При такій реалізації управління потоками сигналів кодуванню підлягає часовий інтервал $t_p(t)$, тоді як при синхронному за запитом методі в синхронних мережах запитувальних систем IFF – T_0. Незалежно від числа запитувачів число сигналів запиту літакового відповідача при синхронному за відповіддю методі визначатиметься часовим інтервалом аналізу. Це значно знижує потік внутрісистемних завад в каналі відповіді.

Оскільки обслуговуванню підлягає будь-який з сигналів запиту, дешифрований в межах інтервалу часу аналізу, то при роботі систем IFF з синхронною відповіддю можливе отримання необхідної інформації від літакового відповідача як за своїм сигналом запиту, так і за сигналами запиту будь-якого запитувача. Це обумовлено тим, що цей метод реалізує принцип обслуговування мережі. Така побудова систем IFF знімає проблему реалізації розосереджених систем IFF, а також часового узгодження сигналів, що поступають за системами первинної та вторинної радіолокації.

Висновки

Розглянутий метод спадкоємного переходу до завадостійких систем IFF дозволяє:

- виключити з обслуговування сигнали запиту, імітовані зацікавленою стороною за рахунок міжперіодної обробки сигналів у літаковому відповідачі;
- знизити інтенсивність сигналів запиту, що обслуговуються літаковим відповідачем;
- перейти до використання у якості сигналів запиту сигналів з великою часовою базою, що загалом дозволяє перейти за спадкоємним принципом до завадозахищених систем IFF як основного інформаційного ресурсу системи контролю повітряного простору.

Список літератури:

1. Stevens. Brian L., Frank L. Lewis, and Eric N. Johnson. Aircraft control and simulation: dynamics, controls design, and autonomous systems. John Wiley & Sons, 2015
2. Маляренко А.С. Системы вторичной радиолокации для управления воздушным движением и государственного радиолокационного опознавания [Справочник]. Харьков : ХУПС, 2007. 78 с.
3. Kim E. and Sivits K. Blended secondary surveillance radar solutions to improve air traffic surveillance // Aerospace Science and Technology. 2015. Vol. 45. P. 203-208.
4. STANAG 4193 Document, Technical Characteristics of IFF Mk X and Mk XII Interrogators and Transponders (Part V). Technical Description of the MkXII System, NATO Standard, 2016.
5. Uzan S. Turan and S. A. Colak. IFF system simulator design based on DSP // 2016 24th Signal Processing and Communication Application Conference, SIU 2016 – Proceedings, 2016. P. 1-4.
6. Benelli G., Giuli D., Mese E. and Pardini S. Characterization of ATC environment for performance evaluation of modern SSR systems // 29th IEEE Vehicular Technology Conference, Arlington Heights, Illinois, USA, 1979. P. 370-377. doi: 10.1109/VTC.1979.1622720.
7. Sharifi-Tehrani O., Sadeghi A. and Razavi S. M. J. Design and Simulation of IFF/ATC Antenna for Unmanned Aerial Vehicle // Majlesi Journal of Mechatronic Systems, vol. 6, no. 1, Jun. 2017.
8. Poornima P., Roja Reddy B. and Anantha Murthy B. G. Design and Simulation of Two-Chain Monopulse Receiver for IFF Radar Application. // 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, India, 2018. P. 1114-1118. doi: 10.1109/RTEICT42901.2018.9012646.
9. Guofeng Jiang; Yangyu Fan; Hongbo Yuan. Assessing the Capacity of Air Traffic Control Secondary Surveillance Radar System // 2019 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC). DOI: 10.1109/CSQRWC.2019.8799146
10. Svyd I., Obod I., Maltsev O., Tkachova T. and Zavolodko G. Improving Noise Immunity in Identification Friend or Foe Systems // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, Ukraine, 2019. P. 73-77. doi: 10.1109/UKRCON.2019.8879812.
11. Obod I., Svyd I., Maltsev O. and Bakumenko B. Comparative Analysis of Noise Immunity Systems Identification Friend or Foe // 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 2020. P. 751-756. doi: 10.1109/ELNANO50318.2020.9088856.
12. Pollack J. and Ranganathan P. Aviation Navigation Systems Security: ADS-B, GPS, IFF // International Conference on Security & Management, SAM'18, International Conference on Security & Management, SAM'18, Las Vegas, Nevada, USA, 2018. P. 129-135.
13. Strelnytskyi O., Svyd I., Obod I., Maltsev O., Voloshchuk O. and Zavolodko G. Assessment Reliability of Data in the Identification Friend or Foe Systems // 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO). Kyiv, Ukraine, 2019. P. 728-731. doi: 10.1109/ELNANO.2019.8783397.
14. Svyd I., Obod I., Maltsev O., Strelnytskyi O., Zubkov O. and Zavolodko G. Method of Increasing the Identification Friend or Foe Systems Information Security // 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT). Lviv, Ukraine, 2019. P. 434-438. doi: 10.1109/AIACT.2019.8847853.
15. Svyd I., Obod I., Maltsev O., Shtykh I. and ZavolodkoG. Model and Method for Detecting Request Signals in Identification Friend or Foe Systems // 2019 IEEE 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM). Polyana, Ukraine, 2019. P. 1-4. doi: 10.1109/CADSM.2019.8779322.
16. Otsuyama T., Honda J., Naganawa J. and Miyazaki H. Analysis of signal environment on 1030/1090MHz aeronautical surveillance systems // 2018 IEEE International Symposium on Electromagnetic Compatibility and 2018 IEEE Asia-Pacific Symposium on Electromagnetic Compatibility (EMC/APEMC). Singapore, 2018. P. 71-71. doi: 10.1109/ISEMC.2018.8394048.
17. Martin Strohmeier. Large-Scale Analysis of Aircraft Transponder Data // IEEE Aerospace and Electronic Systems Magazine (Volume: 32, Issue: 1, January 2017). P. 42 – 44. doi: 10.1109/MAES.2017.160149.
18. David S. and Vitolo A. J. Airborne IFF transponder antenna system with Omni and steerable cardioid patterns, Aug. 1970. P. 279-283.
19. Svyd, I.V., Obod, A.I., Zavolodko, G.E., Melnychuk, I.M., Wójcik, W., Orazalieva, S., Ziyatbekova, G. Assessment of information support quality by "friend or foe" identification systems // Przegląd Elektrotechniczny. 2019. Vol. 1, no. 4. P. 129-133. doi: 10.15199/48.2019.04.22.
20. Svyd I., Maltsev O., Obod I. and Zavolodko G. Fusion Method of Primary Surveillance Radar Data and IFF systems Data // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). Kyiv, Ukraine, 2020. P. 336-340. doi: 10.1109/DESSERT50317.2020.9125040.
21. Semenets V., Svyd I., Obod I., Maltsev O. Tkach M. 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. P. 105-125, 2021. Available: 10.1007/978-3-030-71892-3_5.
22. Svyd I., I. Obod I. Maltsev O. 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. P. 287-306, 2021. doi: 10.1007/978-3-030-71892-3_12.

23. Обод І.І., Свид І.В. Порівняльний аналіз якості виявлення повітряних об'єктів запитальними системами спостереження // Тематичний збірник «Системи обробки інформації». Вип. 9 (90). Харків : ХУПС, 2010. С. 74-76.

24. Obod I., Svyd I., Maltsev O., Vorgul O., Maistrenko G. Zavolodko G. Optimization of the Quality of Information Support for Consumers of Cooperative Surveillance Systems // Radivilova T., Ageyev D., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies. Vol 48. Springer, Cham. P. 133-155, 2020. doi: 10.1007/978-3-030-43070-2_8.

25. Obod I., Svyd I., Maltsev O., Zavolodko G., Pavlova D. Maistrenko G. Fusion the Coordinate Data of Airborne Objects in the Networks of Surveillance Radar Observation Systems // Radivilova T., Ageyev D., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 48. Springer, Cham. P. 731-746, 2020. doi: 10.1007/978-3-030-43070-2_31.

Надійшла до редколегії 03.03.2021

Відомості про авторів:

Свид Ірина Вікторівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, завідувач кафедри мікропроцесорних технологій і систем, Україна; e-mail: iryna.svyd@nure.ua; ORCID: http://orcid.org/0000-0002-4635-6542

Обод Іван Іванович – доктор техн. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри мікропроцесорних технологій і систем, Україна; e-mail: ivan.obod@nure.ua; ORCID: https://orcid.org/0000-0002-9898-0937

Мальцев Олександр Сергійович – Харківський національний університет радіоелектроніки, старший науковий співробітник кафедри мікропроцесорних технологій і систем, Україна; e-mail: aleksandr.maltsev@nure.ua; ORCID: http://orcid.org/0000-0003-1520-9280

Ткач Марія Геннадіївна – Харківський національний університет радіоелектроніки, аспірант кафедри мікропроцесорних технологій і систем, Україна; e-mail: mariia.zavorotna@nure.ua; ORCID: http://orcid.org/0000-0002-4248-7633

Старокожев Святослав Валерійович – Харківський національний університет радіоелектроніки, аспірант кафедри мікропроцесорних технологій і систем, Україна; e-mail: sviatoslav.starokozhev@nure.ua

Глущенко Артем Олександрович – Харківський національний університет радіоелектроніки, аспірант кафедри мікропроцесорних технологій і систем, Україна; e-mail: artem.hlushchenko@nure.ua

Чумак Валерія Сергіївна – Харківський національний університет радіоелектроніки, лаборант кафедри мікропроцесорних технологій і систем, Україна; e-mail: valeriia.chumak@nure.ua; ORCID: http://orcid.org/0000-0002-2403-020X

РАДІОТЕХНІЧНІ ПРИСТРОЇ ТА ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ

УДК 621.396.677.49 **DOI:10.30837/rt.2021.2.205.17**

Ю.Ю. КОЛЯДЕНКО, д-р техн. наук, Н.А. ЧУРСАНОВ

МОДЕЛИ РАСПРОСТРАНЕНИЯ СИГНАЛОВ СЕТЕЙ СВЯЗИ 5 G

Введение

В последние десятилетия во всем мире стремительно развиваются инфокоммуникационные технологии, которые непосредственно влияют на развитие экономики, образования, науки, здравоохранения, культуры и образа жизни человека. По данным Международного Союза Электросвязи (МСЭ) уже в 2012 г. Интернетом пользовался каждый третий житель Земли, и число пользователей неуклонно растет. ИКТ и услуги на базе широкополосного доступа год от года становятся все более доступными для населения, в том числе и по стоимости.

Целью сетей связи пятого поколения является удовлетворение все возрастающих потребностей в мобильной связи [1 – 9]. Сети 5G играют ключевую роль в превращении городов в разумные города, что позволит гражданам и обществу в целом, получить социально-экономические выгоды, которые дает передовая цифровая экономика с интенсивным использованием данных [10].

Технология нового поколения 5G / IMT-2020, как и любая новая технология, привносит свои специфические особенности во все аспекты, касающиеся практики ее применения. Одним из таких особо важных аспектов является электромагнитная совместимость (ЭМС). На этапе подготовки к внедрению радиосетей технологии 5G, названной NewRadio (NR), необходимо заблаговременно позаботиться о принятии мер по эффективной оценки условий ЭМС для этих сетей на основе тщательного анализа особенностей технологии 5G, а правильно и точно оценив эти условия, – успешно обеспечить электромагнитную совместимость радиосредств новых сетей.

Объектом исследования является процесс распространения сигналов миллиметрового диапазона, используемых в технологии нового поколения 5G.

Предмет исследования составляют модели распространения сигналов миллиметрового диапазона, используемых в технологии нового поколения 5G.

Цель работы – разработка модели распространения сигналов миллиметрового диапазона, используемых в технологии нового поколения 5G сети связи 5G.

Постановка задачи

Для внедрения 5G Кабинет министров утвердил план использования радиочастотного ресурса Украины до 2025 г. Эти изменения позволят пользователям получать более стабильные и высокие скорости передачи информации.

МСЭ с 1995 г. стал международным координатором работ по электросвязи, направленных на создание глобального информационного общества. Создав Регламент радиосвязи [11] и разделив поверхность Земли на три региона, МСЭ организовал определенный порядок в частотном пользовании. Однако частотный ресурс принадлежит всему человечеству, исключительно активно используется и по сей день, частотные диапазоны ниже 5 ГГц практически перегружены. Присвоение радиочастот и радиочастотных каналов для радиоэлектронных средств в этих диапазонах осуществляется по технологиям частотно-территориального планирования [12 – 14] с обязательным расчетом ЭМС радиосредств. Поэтому одно из основных направлений по созданию нового поколения мобильной связи 5G – это освоение частотных диапазонов выше 5 ГГц, пока еще недостаточно используемых. Относительно свободные участки спектра есть пока на сверхвысоких частотах, например, на границе диапазонов X и C не занята полоса частот около 1,5 ГГц. Но меньше всего освоен миллиметровый диапазон

(ММД) волн, поэтому именно в этом диапазоне возможно развитие стандарта 5G со скоростями передачи данных от 1 до 10 Гбит/с. Диапазон миллиметровых волн используется пока не очень активно и изучен не полностью. Поэтому представляет интерес исследование возможностей мобильной связи в этом диапазоне волн.

Основная часть

Ключевыми решениями и технологическими компонентами [1 – 4] сети радиодоступа 5G NR являются:

1) использование новых форм сигнала, получивших название Non-Orthogonal Waveform и дающих выигрыш в спектральной эффективности по отношению к OFDM;

2) применение полного дуплекса FD – одновременной передачи и приема в общей полосе частот, преимущественно в коротких соединениях "точка-точка" (D2D, V2V);

3) применение многомерных антенн MIMO, в которых эффективно реализуется режим динамического формирования направленных лучей для передачи (3D / Beamforming), что позволяет увеличить энергетический выигрыш в ожидаемых высоких диапазонах частот и улучшить покрытие и спектральную эффективность в ультраплотных малых сотах;

4) применение малых сот Small Cellco со сверхплотным распределением (один приемопередатчик на каждого пользователя), разгрузка макросотами сети с разделением сред передачи команд управления и пользовательского трафика между макро- и Small-сотами в различных полосах частот (концепция "PhantomCell»).

Радиочастоты для сетей 5G – это одна из главных компонент, оказывает существенное влияние на ЭМС [6, 9, 12].

На Всемирной конференции радиосвязи ВКР-15 были определены новые диапазоны радиочастот для 5G, в том числе диапазоны сантиметровых и миллиметровых волн [12]. В целом этот радиочастотный спектр размещен в трех областях: ниже 1 ГГц, от 1 ГГц до 6 ГГц и выше 6 ГГц (до 100 ГГц).

В качестве главных особенностей этого спектра, с точки зрения ЭМС, можно выделить следующее: возможность использования широкой непрерывной полосы канала (суммарно до 1 – 2 ГГц), малые зоны обслуживания (дальность излучения) в малых (Small) и ультрамалых (UltraSmall) сотах; возможность использования малогабаритных многоэлементных антенн MIMO с узкими лучами как в базовых станциях, так и в абонентских устройствах; различный характер потерь при распространении сигнала, в частности значительное влияние на уровень потерь дополнительных ранее неизвестных в сотовой связи факторов (газы – кислород, водяной пар и др.) [15, 17].

Описание радиоинтерфейсов [1, 6, 9, 12, 15, 17] сетей 5G представлено в табл.1.

Таблиця 1

Описание радиоинтерфейсов сетей 5G

Параметр	NR ниже 6 ГГц	NR выше 6 ГГц
Ширина канала	5 МГц, 10 МГц,5 МГц, 20 МГц, 25 МГц, 30 МГц, 40 МГц, 50 МГц, 60 МГц, 80 МГц 100 МГц.	50 МГц, 100 МГц, 200 МГц и 400 МГц.
Диапазоны	Отдельные полосы 450-3800 МГц, 3800-4200 МГц и 4400-5000 МГц	26,5-29,5 ГГц, 24,25-27,5 ГГц 37-40 ГГц
Задержка (на уровне радиоинтерфейса)	4 мс (Фаза 1) 1 мс (Фаза 2)	4 мс (Фаза 1) 1 мс (Фаза 2)
Пиковые скорости	2 Гбит/с и выше	До 20 Гбит/с

Анализ главных особенностей радиоинтерфейса 5G [1, 6, 9, 12, 15, 17, 18] позволяет указать на ожидаемые особенности процедур оценки условий ЭМС для этих сетей. Эти особенности главным образом касаются учета суммарной помехи от сети при ее особой архитектуре и динамике изменений, выбора новых моделей потерь (моделей канала) при пространст-

венно-распределенном излучении многомерных антенн MIMO и разнородной среде распространения сигнала, а также учета спектральных свойств новых форм сигнала и характера излучения при новых неортогональных методах радиодоступа.

Главными недостатками сигналов миллиметрового диапазона (ММД) являются:

1) сильное затухание миллиметровых волн при распространении;

2) уровень сигнала существенно зависит от влияния гидрометеоров (капли дождя, снег, град, туман) и от присутствия в атмосфере твердых неоднородностей (листья деревьев, стаи птиц, пыль);

3) высокая степень влияния на уровень сигнала препятствий, которые закрывают трассу;

4) наличие зон сильного ослабления сигнала на некоторых частотах из-за ослабления сигналов ММД молекулами кислорода и парами воды.

Как и во всех линиях связи и радиоэлектронных системах, в линиях связи ММД распространяющиеся радиоволны имеют сложную, случайно-детерминированную структуру и подчиняются законам электродинамики, а параметры радиоволн описываются уравнениями Максвелла [19]. Однако такое представление является достаточно сложным, что может привести к усложнению модели, а следовательно, к трудностям работы с ней, к увеличению погрешностей расчетов. Поэтому в качестве модели распространения сигналов в радиолиниях воспользуемся известной моделью, основанной на уравнении передачи [20]:

$$P_{пр} = P_{пер} + G_{пер} + G_{пр} - \eta_{пер} - \eta_{пр} - W, \qquad (1)$$

где $P_{пр}$ – мощность принимаемого сигнала (дБ); $P_{пер}$ – мощность передатчика; $G_{пер}, G_{пр}$ – соответственно коэффициенты усиления передающей и приемной антенн; $\eta_{пр}, \eta_{пер}$ – коэффициенты полезного действия приемного и передающего фидеров; W – потери (ослабление) электромагнитного поля (дБ).

Задачи, связанные с распространением радиоволн в приземной зоне, сложны, поскольку поле около антенны радиоприемника как абонентской станции, так и базовой станции представляет собой суперпозицию сигналов из-за их многолучевого распространения в условиях данной местности. Проблема осложняется влиянием на условия распространения радиоволн передвижения объектов, рассеивающих радиоволны и перемещения самих абонентов в зоне неравномерного поля. Уровень сигнала может изменяться от пиковых значений, превышающих средний уровень на несколько единиц и даже десятков децибел, до десятков децибел ниже среднего в зонах сильного замирания [21].

Для расчета ослабления сигналов при анализе ЭМС и проектировании мобильных сетей связи от 1 до 4G наиболее широко применяется моделирование, основанное на результатах статистической обработки экспериментальных исследований распространения сигналов вдоль земной поверхности. Такие исследования проводились во многих странах мира для различных условий местности. Некоторые из этих моделей являются общепризнанными и рекомендованы МСЭ для использования при проектировании мобильных сетей связи.

Можно выделить два основных типа моделей. Первый тип, где в качестве основных параметров, характеризующих местность и условия распространения сигналов, являются высоты расположения антенн и высоты неровностей местности [22 – 24]. Второй тип – модели ослабления сигналов в статистически однородной среде, где рельеф местности обычно не учитывается [22]. Первый тип модели, хотя и является точным, более пригоден для анализа стационарных систем связи, так как при изменении местоположения абонента будут и изменяться получаемые результаты исследований и при наборе статистики при моделировании в результате получатся средние значения. Кроме того, первый тип модели требует большего числа входных априорных данных и большего времени на проведение исследований. Математические модели распространения радиоволн, построенные на основе экспериментальных данных и описывающие поле в статистически однородной среде (городская территория, пригород, сельская местность, открытое пространство), являются общепризнанными, о чем свидетельствуют Рекомендации ITU и CEPT, и могут быть использованы для расчета зон по-

крытия мобильных сетей связи и оценки их ЭМС. Кроме того, целесообразно выделить в особую категорию модели распространения в пределах зданий. Рекомендуется придерживаться следующего соответствия между типами моделей ослабления и характеристиками среды, в которой используется мобильная сеть связи [22 – 24]:

1. Отдельное помещение, офис внутри здания.

2. Открытое пространство – открытые участки без насаждений.

3. Плотная городская застройка – плотная застройка в основном высокими зданиями выше 20 этажей с малой площадью зеленых насаждений.

4. Городская застройка – многоэтажная административная и жилая застройка, индустриальные районы с зелеными насаждениями.

5. Пригород – одиночные дома, административные здания высотой 1 – 3 этажа. Большие площади зеленых насаждений, парковые зоны с отдельными группами зданий плотной застройки.

6. Сельская местность – открытое пространство с несколькими зданиями, фермы, кустарниковые насаждения, шоссе.

В общем виде ослабление сигнала (помехи) W определяется выражением

$$W = W_{РЭС} + W_{ср}, \tag{2}$$

где $W_{РЭС}$ – ослабление помехи, определяемое характеристиками радиоэлектронного средства (РЭС), при расчете ослабления полезного сигнала можно считать $W_{РЭС} = 0$;

$W_{ср}$ – ослабление сигнала (помехи), определяемое условиями среды распространения радиоволн.

Параметр ослабления энергии помехи $W_{РЭС}$ определяется следующими составляющими, [22 – 24]:

$$W_{РЭС} = W^{(1)}_{РЭС}(\Delta f) + W^{(2)}_{РЭС} + W^{(3)}_{РЭС} + W^{(4)}_{РЭС} + W^{(5)}_{РЭС}, \tag{3}$$

где $W^{(1)}_{РЭС}(\Delta f)$ – ослабление помехи, определяемое избирательностью приемника по основному каналу $S_{пр}(\Delta f)$, шириной спектра основного и внеполосного излучения передатчика и частотным разносом между каналами излучения и приема, дБ; $W^{(2)}_{РЭС}$ – ослабление помехи за счет побочного излучения. Значения $W^{(2)}_{РЭС}$ определяются в результате проведения эксперимента. При отсутствии экспериментальных данных по значениям плотностей потоков мощности ослабление сигнала за счет побочного излучения принимается равным относительному уровню боковых излучений. Уровень боковых излучений определяется коэффициентом подавления по боковым излучениям. Для РЭС мобильных сетей связи он обычно составляет 60 дБ;

$W^{(3)}_{РЭС}$ – ослабление помехи за счет приема по побочным каналам, определяется избирательностью приемника по побочным каналам приема;

$W^{(4)}_{РЭС}$ – ослабление помехи при полных частотных несовпадениях помехи с основным и побочными каналами приема. При этом помеха является продуктом нелинейностей радиоприемного тракта;

$W^{(5)}_{РЭС}$- ослабление помехи, вызванное различием в поляризации возбуждающего поля и антенны приемника, определяется соотношением

$$W^{(5)}_{РЭС} = 10 \cdot \lg(\gamma_{ij}), \tag{4}$$

где γ_{ij} – поправочный коэффициент, который учитывает расхождения поляризаций полезной радиоволны *i*-го РЭС с радиоволной помехи *j*-го РЭС.

В условиях использования сети связи в отдельном помещении или здании ослабление сигнала $W_{ср}$, определяемое условиями среды распространения, рассчитывается согласно выражению

$$W_{ср} = W_{св} + W_{доп}^{(1)} + W_{доп}^{(2)} + W_{доп}^{(3)} + W_{доп}^{(4)} + W_{доп}^{(5)} + W_{доп}^{(6)}, \quad (5)$$

где $W_{св}$ – ослабление в свободном пространстве (дБ)

$$W_{св} = 92{,}4 + 20\lg(f) + 20\lg(d), \quad (6)$$

где d – расстояние между передатчиком и приемником, f – частота;

$W_{доп}^{(1)}$ – дополнительное ослабление, вызванное влиянием стен и перекрытий этажей:

$$W_{доп}^{(1)} = W_{0с} \cdot N_{ст}^{\left(\frac{N_{ст}+2}{N_{ст}+1} - c\right)} + W_{0э} \cdot N_{э}^{\left(\frac{N_{э}+2}{N_{э}+1} - c\right)}, \quad (7)$$

где $W_{0с}$ – ослабление за счет влияния стены или межэтажного перекрытия. Обычно [22 – 24] для стены $W_{0с}$=8,38 дБ и c=0,51, межэтажного перекрытия $W_{0э}$=18,3 дБ и c=0,46, $N_{ст}$ – количество стен; $N_{э}$ – количество межэтажных перекрытий;

$W_{доп}^{(2)}$ – дополнительные потери энергии сигнала при заполнении пространства различными предметами [22 – 24]:

$$W_{доп}^{(2)} = \delta \cdot d, \quad (8)$$

где δ – коэффициент погонного ослабления, учитывающий заполнение пространства предметами. Для почти пустого пространства δ=0,2 дБ/м, для переполненного пространства δ=0,6 дБ/м;

$W_{доп}^{(3)}$ – дополнительное ослабление, вызванное потерей энергии радиоволн при распространении через дожди,

$$W_{доп}^{(3)} = kdK_d Y^a \quad (9)$$

где k – коэффициент, определяющий наличие или отсутствие осадков, Y – интенсивность осадков, мм/ч, K_d – параметр, зависящий от частоты, температуры, поляризации дБч/м2, a – безразмерный параметр, зависящий от частоты, температуры, поляризации.

В табл. 2 приведены значения погонного затухания сигнала в дождях в зависимости от интенсивности осадков и частоты.

Таблица 2

Погонное затухание в дождях

Частота, ГГц	Процент времени года	Интенсивность осадков	Погонное затухание сигнала в дождях, дБ/км
30	1	0,6	0,1
	0,3	2,4	0,43
	0,1	6	1,08
	0,03	12	2,18
	0,01	22	4,02
60	1	0,6	0,63
	0,3	2,4	1,84
	0,1	6	3,73
	0,03	12	6,35
	0,01	22	10,12
90	1	0,6	0,7
	0,3	2,4	1,99
	0,1	6	3,98
	0,03	12	6,7
	0,01	22	10,6

$W_{доп}^{(4)}$ – дополнительное ослабление, вызванное потерей энергии радиоволн из-за тумана:

$$W_{доп}^{(4)} = kdl_T V_T,$$

l_T – удельный погонный коэффициент ослабления сигнала ММД в тумане, V_T – коэффициент содержания воды в атмосфере, который определяется по оптической видимости,

Удельный погонный коэффициент ослабления сигнала ММД в тумане приведен в табл.3.

Таблица 3

Удельный погонный коэффициент ослабления сигнала ММД в тумане

f_0, ГГц	30	43	60	150
l_T, дБ $м^3$ /гкм	0,438	0,876	1,65	7,14

Коэффициент содержания воды в атмосфере приведен в табл.4.

Таблица 4

Коэффициент содержания воды в атмосфере

Оптическая видимость, м	30	50	80	200
V_T, г/ $м^3$	2	1	0,5	0,2

$W_{доп}^{(5)}$ – дополнительное ослабление сигнала при распространении через листья деревьев,

$$W_{доп}^{(5)} = 0{,}2 f^{0,3} r^{0,6},$$

где r - глубина слоя перекрывающего листья, м.

Случайная компонента дополнительного ослабления

$$W_{доп}^{(6)} = W_{сл}^{(м)} + W_{сл}^{(б)}, \qquad (10)$$

где $W_{сл}^{(м)}, W_{сл}^{(б)}$ отображают соответственно медленные и быстрые случайные замирания.

В расчетных задачах по энергетике радиолиний малых расстояний быстрыми замираниями обычно пренебрегают, ибо они характерны для достаточно протяженных, (d >30 – 50 км), преимущественно закрытых, или полузакрытых радиотрасс. Для офисных радиолиний или радиолиний в пределах микрорайона принято считать $W_{сл}^{(б)}(t) \to 0$. Медленные замирания $W_{сл}^{(м)}$ составляют 10 – 16 дБ. Медленные замирания подчиняются случайному логарифмически-нормальному закону, тогда [20]

$$W_{сл}^{(м)}\{u(t)\} = \rho W_{сл}^{(м)}\{u(t-\Delta t)\} + \sqrt{1-\rho^2} \cdot N(0,\sigma), \qquad (11)$$

где ρ – коэффициент корреляции между двумя сечениями случайного процесса изменения $W_{сл}^{(м)}\{u(t)\}$, разнесенных на интервал Δt.

В условиях использования мобильных сетей связи вне здания считают, что основными механизмами распространения радиоволн (РРВ) являются: дифракция, рефракция, распространение в свободном пространстве и вдоль земной поверхности. Данные механизмы могут действовать совместно или в различной комбинации, в зависимости от физико-географических условий. Имеется ряд рекомендаций ITU, которые позволяют учитывать те или иные механизмы РРВ: ITU-R PN.525 – Расчет ослабления в свободном пространстве [23], ITU-R PN.526 – Распространение радиоволн с учетом дифракции [24].

В рекомендациях ITU по учету потерь при РРВ для частот свыше 1 ГГц предлагается использовать полуэмпирическую модель, изложенную в рекомендации ITU-R P.1146 [25]. Эта модель позволяет рассчитывать напряженность поля в точке приема в диапазоне частот 1 – 3 ГГц для стационарных и мобильных систем связи. Недостатком этой модели является то,

что она не позволяет рассчитывать потери при РРВ, если высоты передающей и приемной антенн выше 30 м. В этих случаях рекомендуется использовать модель, изложенную в рекомендации ITU-R P.452-8 [26].

Уровень сигнала на входе приемника в диапазоне 0,03 – 3 ГГц, особенно в случае использования мобильных абонентских станций, зависит от многих факторов: типа местности, высоты зданий и плотности застройки города, высот приемной и передающих антенн, наличия растительности и многих других факторов. Поэтому при расчете получают медианное значение напряженности поля в точке приема с определенной вероятностью. Таким образом, вывод о наличии или отсутствии помехи также имеет вероятностный характер.

Выводы

Технология нового поколения 5G / IMT-2020, как и любая новая технология, привносит свои специфические особенности во все аспекты, касающиеся практики ее применения. Одним из таких особо важных аспектов является ЭМС. На этапе подготовки к внедрению радиосетей технологии 5G необходимо заблаговременно позаботиться о принятии мер по эффективной оценки условий ЭМС для этих сетей на основе тщательного анализа особенностей технологии 5G, а правильно и точно оценив эти условия – успешно обеспечить электромагнитную совместимость радиосредств новых сетей.

На Всемирной конференции радиосвязи ВКР-15 были определены новые диапазоны радиочастот для 5G, в том числе диапазоны сантиметровых и миллиметровых волн. В целом этот радиочастотный спектр размещен в трех областях: ниже 1 ГГц, от 1 до 6 ГГц и выше 6 ГГц (до 100 ГГц). В качестве главных особенностей этого спектра, с точки зрения ЭМС, можно выделить следующие: различный характер потерь при распространении сигнала, в частности значительное влияние на уровень потерь дополнительных ранее неизвестных в сотовой связи факторов (газы – кислород, водяной пар и др.).

Разработанная математическая модель распространения сигналов сетей связи 5 G учитывает:

ослабление сигналов в свободном пространстве,

ослабление сигналов, вызванное влиянием стен и перекрытий этажей,

потери энергии сигнала при заполнении пространства различными предметами,

ослабление сигналов, вызванное потерей энергии радиоволн при распространении через дожди,

ослабление сигналов, вызванное потерей энергии радиоволн из-за тумана,

ослабление сигналов при распространении через листья деревьев, медленные и быстрые случайные замирания.

Список литературы:

1. 3GPP TR 22.891. Feasibility Study on New Services and Markets Technology Enablers. Ver. 14.2.0, Sep. 2016.

2. 3GPP TR 38.913. Study on Scenarios and Requirements for Next Generation Access Technologies. Ver. 14.3.0, June 2017.

3. 3GPP TS 28.554. Management and orchestration; 5G end to end Key Performance Indicators (KPI). Ver. 2.0.0, release 15, Sep 2018.

4. 5G PPP Architecture Working Group white paper. View on 5G Architecture. July 2016.

5. Abuarqoub A. Behaviour Profiling in Healthcare Applications Using the Internet of Things Technology / Abuarqoub A., Hammoudeh M. H. // Proceedings of Fourth International Conference on Advances in Information Processing and Communication Technology. 2016. P. 1-4. DOI:https://doi.org/10.15224/978-1-63248-099-6-25

6. Agiwal M. Next generation 5G wireless networks: A comprehensive survey / Agiwal M., Roy A., Saxena N // IEEE Communications Surveys & Tutorials. 2016. №18(3). P. 1617-1655. DOI:https://doi.org/10.1109/COMST.2016.2532458

7. Aijaz A. Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks / A.Aijaz, M.Dohler, A.H.Aghvami, V.Friderikos, Frodigh M. // IEEE Wireless Comm. 2017. 24(2). P. 82-89. DOI:https://doi.org/10.1109/MWC.2016.1500157RP

8. Aijaz A. Shaping 5G for the Tactile Internet / Aijaz A.; Simsek M.; Dohler M. and Fettweis G. // 5G Mobile Communications. Springer International Publishing, pp.677-691, 2017. DOI:https://doi.org/10.1007/978-3-319-34208-5_25

9. Aijaz A. Towards 5G-enabled tactile internet: Radio resource allocation for haptic communications // Proceedings of the 2016 IEEE Wireless Communications and Networking Conference (WCNC), Doha, Qatar, 3-6 April 2016. P. 1-6. DOI:https://doi.org/10.1109/WCNC.2016.7564661

10. Бородин А. С. Сети связи пятого поколения как основа цифровой экономики / А.С. Бородин, А.Е. Кучерявый // Электросвязь. 2017. №5. С. 47-51.

11. Radio Regulations. Ed. ITU. In 4 vol. 2016.

12. Resolution COM 6/20 (WRC-15) Studies on frequency-related matters for International Mobile Telecommunications identification including possible additional allocations to the mobile services on a primary basis in portion(s) of the frequency range between 24.25 and 86 GHz for the future development of International Mobile Telecommunications for 2020 and beyond.

13. Бабков В.Ю. Сети мобильной связи. Частотно-территориальное планирование / В.Ю. Бабков, М.А. Вознюк, П.А. Михайлов. Москва : Горячая линия – Телеком, 2007. 224 с.

14. Тихвинский В.О. Технологии 5G – базис мобильной инфраструктуры цифровой экономики // Электросвязь. 2018. № 3. С. 49–55.

15. Куракова Т.П. Имитация радиоканалов миллиметрового диапазона поколения 5G : дис. … канд. техн. наук ; ФГУП НИИР//http://diss.vlsu.ru/uploads/media/Dissertacija_ Kurakovoi.pdf.

16. Kurakova T. How ITU can help develop future networks/ T. Kurakova, M.Valdburger // ITU News. 2013. № 1. P. 38-41. DOI:https://doi.org/10.1525/aft.2013.41.3.38

17. Молчанов Д.А. Разработка подходов, методов исследования и моделей обеспечения показателей качества обслуживания в беспроводных сетях пятого поколения : дис. … д-ра техн. наук ; Рос. ун-т дружбы народов. Москва, 2019. 306 с.

18. Кременецька Я.А. Аналіз обмежуючих та компенсуючих факторів при розрахунку енергетичної ефективності радіосистем в міліметровому діапазоні /Я.А. Кременецька, С.Ю. Марков, Н.В. Градобоєва, Є.М. Харченко // Телекомунікаційні та інформаційні технології. 2019. №1. С. 12-21. Режим доступу: http://nbuv.gov.ua/UJRN/vduikt_2019_1_4 . DOI: 10.31673/2412-4338.2019.011221

19. Griffiths D. J. Introduction to Electrodynamics. 4th ed. Pearson. Boston, 2013. P. 347.

20. Williams, T. EMC for Product Designers // Elsevier Science & Technology, 2016. P. 513.

21. Коляденко Ю.Ю. Математическая модель взаимодействия элементов системы абонентского радиодоступа // Праці УНДІРТ. Теоретичний та науково-практичний журнал радіозв'язку, радіомовлення і телебачення. 2004. №1 (37). С. 31-35.

22. Recommendation ITU-R PN.452.

23. Recommendation ITU-R PN.525.

24. Recommendation ITU-R PN.526.

25. Recommendation ITU-R P.1146.

26. Recommendation ITU-R PN.452-8.

Поступила в редколлегию 05.03.2021

Сведения об авторах:

Коляденко Юлия Юрьевна – д-р техн. наук, профессор, профессор кафедры инфокоммуникационной инженерии имени В.В. Поповского, Харьковский национальный университет радиоэлектроники, Украина; e-mail: yuliia.koliadenko@nure.ua; ORCID: https://orcid.org/0000-0002-0247-2736

Чурсанов Никита Александрович – Харьковский национальный университет радиоэлектроники, аспирант кафедры инфокоммуникационной инженерии имени В.В. Поповского, Украина; e-mail: mykyta.chursanov@nure.ua; ORCID: https://orcid.org/0000-0002-3568-2633

РАДІОТЕХНІЧНІ СИСТЕМИ

УДК 621.391.82: 004.056.53 **DOI:10.30837/rt.2021.2.205.18**

С.П. СЕРГІЄНКО, канд. фіз.-мат. наук, В.Г. КРИЖАНОВСЬКИЙ, д-р техн. наук ,
Д.В. ЧЕРНОВ, канд. техн. наук, Л.В. ЗАГОРУЙКО, канд. техн. наук

ЕФЕКТИВНІ РЕЖИМИ РОБОТИ РАДІОЗАКЛАДНИХ ПРИСТРОЇВ ДЛЯ ПОТАЙНОГО ЗНІМАННЯ ІНФОРМАЦІЇ У ПОЛІ ШУМОВИХ ЗАВАД

Вступ

Відомі пасивні радіозакладні пристрої, які використовують для несанкціонованого зняття інформації [1], не мають джерела живлення. Такі пристрої використовують енергію зовнішнього джерела радіовипромінювання, рівень напруги яких перевершує термічний потенціал. Пристрій передає інформацію завдяки перетворенню частоти та енергії електромагнітних хвиль, якими вони опромінюються. Складовою частиною такого пристрою є елемент з нелінійною вольт-амперною характеристикою. Зазвичай таким елементом є високочастотний діод або пристрій з подібною вольт-амперною характеристикою. Нелінійність вольт-амперної характеристики $p-n$ переходу або діоду Штоки призводить до генерації коливань високочастотного поля на кратних частотах від частоти зовнішнього джерела опромінювання. Ефективне перетворення енергії високочастотного сигналу можливо при рівні сигналу на діоді від зовнішнього джерела опромінення, який перевищує напругу термічного потенціалу $kT/q = 0,26$ мВ. Частота інформаційного сигналу знаходиться в акустичному діапазоні спектру, що набагато менше ніж частоти електромагнітного опромінювання. Рівень генерації кратних гармонік залежить від зміщення робочої точки вольт-амперної характеристики $p-n$ переходу. Робоча точка зміщується напругою акустичного сигналу. Таким чином відбувається модульована по амплітуді генерація на кратних гармоніках частоти зовнішнього опромінювання. Використання шумового сигналу для забезпечення передачі інформації у різних варіантах пропонується в роботах [2, 3]. В [2] пропонується використовувати проміжний ретранслятор, який використовує енергію маскуючого зашумлення, що генерується зі сторони приймача системи передачі інформації. Зашумлення передається імпульсами, під час яких здійснюється передача від передавача до проміжного ретранслятора і накопичення енергії від сигналу зашумлення в ретрансляторі, а в паузах зашумлення інформація передається від ретранслятора до приймача. Для підтримки такого режиму роботи ретранслятор має дві направлені антени, спрямовані до передача та приймача.

В роботі [3] досліджується безпека фізичного рівня кооперативного неортогонального множинного доступу NOMA (Non-Orthogonal Multiple Access) в жорсткому сценарії, де немає прямого зв'язку від джерела до дальнього пункту призначення, в той час як прямий зв'язок між пристроєм підслуховування і джерелом існує. Для забезпечення безпеки зв'язку в цьому сценарії пропонується схема кооперативного неортогонального множинного доступу з використанням глушіння. Зокрема, глушильний пристрій може пасивно збирати енергію оточуючих джерел радіочастотних сигналів, а зібрана енергія використовується для випромінювання шуму, щоб не дозволити підслуховування. Доданий шум змінює локальний рівень сигналу. Завдяки цьому зовнішній пристрій, що підслуховує, не зможе по рівню потужності визначити адресу призначення інформації.

Для боротьби з несанкціонованим зніманням інформації за допомогою пасивних радіозакладок використовують генератори шуму [4 – 6]. Генератори шуму випромінюють в потенційно небезпечному діапазоні частот радіохвилі, їх рівень потужності перевершує потужність генерації радіозакладного пристрою. Зашумлення об’єктів, що підлягають захисту, у всьому можливому на даний час діапазоні частот технічно складно, і це також призводить до

унеможливлення нормальної роботи систем радіозв'язку навколо об'єктів, що захищаються. Більш прийнятним є захист з використанням зашумлення на деяких вузьких частотних діапазонах, на яких є загроза передачі інформації радіозакладними пристроями. Це може бути на діапазонах роботи стільникового зв'язку, або в діапазоні, в якому починає роботу радіопередавальний невідомий пристрій та рівень випромінювання якого достатній для активізації радіозакладних пристроїв [7]. Потенційно небезпечно передавати інформацію за допомогою NFC-зв'язку. Зловмисники можуть несанкціонованим чином зняти інформацію на кратних гармоніках [8]. Для протидії такому витоку інформації можливо використовувати зашумлення на частотах кратних частоті передачі інформації NFC-зв'язку 13,65 МГц. Такий захист значно простіше використовувати, так як ці діапазони відносяться до промислових частот і тому він не буде заважати роботі інших користувачів.

В статті демонструється можливість роботи радіозакладного пристрою з використання енергії генератору шуму, який повинен заважати несанкціонованому зніманню інформації. Схема, в якій можливе потенційне знімання інформації з використанням генератору шуму, представлена на рис. 1, де: ПП – пристрій, що підслуховує; ФНЧ – фільтр низьких частот; ГШ – генератор шуму для протидії підслуховуванню; ПВЧ – приймач високих частот; ПНЧ – приймач низьких частот. Інформація може передаватися по високочастотних і низькочастотних каналах.

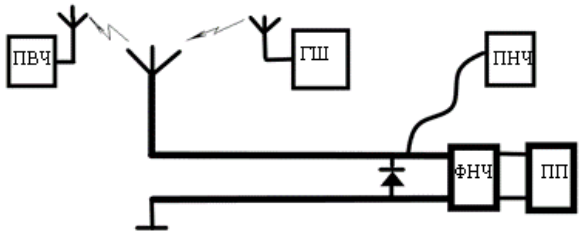


Рис. 1

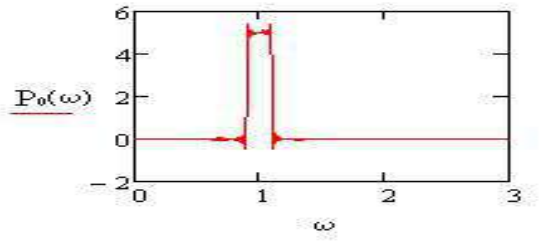


Рис. 2

Моделювання нелінійних перетворень спектру шумових завад

Моделювання потенційної можливості знімання інформації в полі шумових завад ГШ (рис. 1) проводиться на спрощеній моделі [9]. Радіозакладка моделюється діодом, який підключено до кінця довгої лінії з хвильовим опором Z_0. До лінії передачі з іншого кінця підключена широкосмугова антена. Шумовий сигнал модулюється Гаусовим смуговим сигналом зі спектром, як на рис. 2, частота ω – це нормована частота на центральну частоту діапазону.

Для радіозакладних пристроїв частіше використовується дециметровий діапазон радіохвиль. Вибір цього діапазону зумовлений тим, що розмір антени повинен бути невеликим, щоб не викривати скритність пристрою і забезпечити можливість безперешкодного розповсюдження інформаційного сигналу в умовах приміщень. Частота інформаційного акустичного сигналу не перевершує 10 кГц. Після змішення випадкової напруги шумового сигналу з інформаційним акустичним сигналом на нелінійному елементі випадковий сигнал змініть свій спектр і перестане бути ергодичним. Інформація буде закодована в амплітудну модуляцію шумового сигналу.

Рівняння, яке описує зв'язок напруги падаючої хвилі U_F від ГШ, параметри лінії передач (хвильовий опір Z_0), з напругою і струмом нелінійного елемента та напругою відбитої хвилі U_R, має вигляд (1) [10]. Це рівняння справедливо для безінерційного нелінійного перетворення. Таке наближення справедливо в разі, якщо частота сигналу і постійна часу τ знаходяться в співвідношенні $\omega \ll 1/\tau$. Стала τ визначається хвильовим опором лінії Z_0 і ємністю $p-n$ переходу C і дорівнює $\tau = Z_0 C$. Напруга на кінці довгої лінії буде складатися з напруги

падаючої хвилі U_F, напруги відбитої хвилі U_R і напруги зміщення U_0, яка модулює опір діоду:

$$U = U_0 + U_F + U_R. \tag{1}$$

Зв'язок між напругами падаючої хвилі, відбитої хвилі та струмом на кінці довгої лінії і хвильовим опором описується виразом [10]

$$U_F - U_R = I \cdot Z_0. \tag{2}$$

В якості навантаження на кінці довгої лінії встановлено діод, його вольт-амперну характеристику будемо вважати ідеалізованою з експоненціальною залежністю струму від напруги. Враховуючи (1) і (2), отримаємо

$$U_F - U_R = j_0\left(e^{\frac{q(U_0+U_F+U_R)}{kT}} - 1\right)Z_0. \tag{3}$$

З рішення рівняння (3) відносно U_R отримаємо залежність напруги відбитої хвилі $U_R\left(U_F, U_0, Z_0\right)$. Падаючу хвилю (шумовий сигнал, який використовується для протидії несанкціонованому зніманню інформації) описуємо випадковим сигналом з Гаусовим розподілом амплітуди і смуговим спектром. Кореляційна функція такої такого сигналу описується [11] так:

$$r(\tau) = \frac{\sin \Delta\omega\tau \cdot \cos \omega_0\tau}{\pi\Delta\omega\tau}. \tag{4}$$

Спектральна щільність потужності падаючої хвилі

$$P_0(\omega) = \int_{-\infty}^{\infty} r(\tau) e^{-i\omega\tau} d\tau. \tag{5}$$

Графік спектральної щільності потужності падаючої хвилі представлений на рис. 2. Кореляційна функція відбитої хвилі визначається формулою (6); в такому вигляді кореляційна функція нормована на опір довгої лінії з приєднаним нелінійним елементом. Спектр відбитої хвилі находиться заміною $r(\tau)$ на $B(\tau)$ у формулі (5):

$$\begin{aligned} B(\tau) = \int_{-\infty}^{\infty}\int U_R\left(U_{F1}, U_0, Z_0\right) U_R\left(U_{F2}, U_0, Z_0\right) \frac{e^{\frac{-\left(U_{F1}^2 + U_{F2}^2 - 2U_{F1}U_{F2}r(\tau)\right)}{2\sigma^2\sqrt{1-r(\tau)^2}}}}{Z_0 \cdot 2\pi\sigma\sqrt{1-r(\tau)^2}} dU_{F1} dU_{F2} - \\ - \frac{1}{Z_0}\left(\int_{-\infty}^{\infty} \frac{U_R\left(U_F, U_0, Z_0\right) e^{\frac{U_F^2}{2\sigma^2}}}{\sqrt{2\pi\sigma}} dU_F\right)^2 \end{aligned} \tag{6}$$

Спектральна щільність потужності відбитої хвилі для напруги зміщення $U_0 = 3$ і опору довгої лінії $Z = Z_0/\left(kT/qj_0\right) = 1$ має вигляд, представлений на рис. 3. Взаємне співвідношення максимумів спектральної щільності поблизу частот $\omega = 0$, $\omega = 2$ та $\omega = 1$ змінюється в протилежних напрямках в залежності від опору довгої лінії і напруги зміщення. Якщо спектральна щільність поблизу частот $\omega = 0$, $\omega = 2$ збільшується, то спектральна щільність поблизу частоти $\omega = 1$ зменшується, і навпаки. На відмінність від спектральної щільності потужності шумового сигналу який використовується для протидії несанкціонованого зніманню інформації (рис. 2), хвиля, відбита від нелінійного елемента, має у спектру потужності складові, які відсутні у шумовому сигналі падаючої хвилі. Залежність потужності відбитої хвилі від дисперсії шумового сигналу на подвійної частоті і при напрузі $U_0 = -3$ представлена на рис. 4. Потужність відбитої хвилі зростає за експоненціальним законом від дисперсії шуму.

На рис. 5 представлена залежність спектральної щільності потужності від сталої напруги, прикладеної до нелінійного елемента. Розрахунок проводився для опору $Z = 0{,}1$. Вибір опору був обумовлений тим, що зміна напруги зміщення в Вольтах від $+3kT/q$ до $-9kT/q$ призводить до збільшення діапазону зміни диференційного опору діоду, що відображається на кількості максимумів і співвідношенні висот максимумів спектральної щільності. Як видно з графіку, зміна спектральної щільності в залежності від напруги зміщення носить немонотонний характер. Найбільшій рівень нелінійних перетворень спектру відбитої хвилі спостерігається при прямому зміщенні діоду.

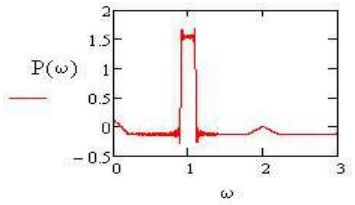


Рис. 3

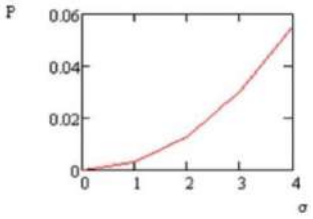


Рис. 4

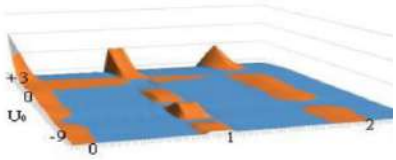
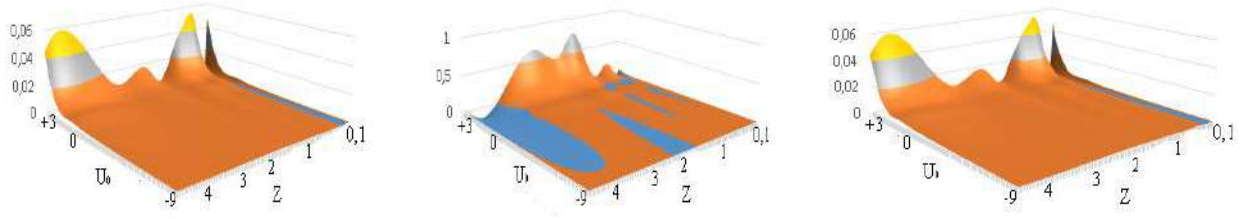


Рис. 5

Для несанкціонованого знімання інформації більш привабливими є спектральні складові відбитої хвилі, які групуються біля нульової частоти та біля подвійної частоти $2\omega_0$ завдяки більшої потужності цих складових, та відсутності в цих діапазонах сигналу зашумлення. Використання частот біля нульової частоти більш привабливе для передачі інформації через провідні лінії електропередачі, металеві елементи будівельних конструкцій, металеві елементи водо-теплопостачання, каналізаційні споруди, тощо. Для передачі інформації через ефір більш підходить спектральний максимум біля частоти $2\omega_0$, де ω_0 – центральна частота смуги зашумлення. В такому випадку простіше забезпечити ефективний прийом та передачу однією антеною на відміну від використання спектральних складових на частотах $3\omega_0$ та вище. Ефективність нелінійного перетворення шумового сигналу залежить від двох параметрів: напруги зміщення, що подається на діод та опору довгої лінії (точніше відношення опору лінії до диференціального опору діоду без зміщення. Диференційний опір діоду при розрахунках без напруги зміщення береться в безрозмірних одиницях $Z_d = Z_0/(kT/qj_0)$ де Z_0 хвильовий опір довгої лінії. $Z_d = 1$ буде в випадку підключення до довгої лінії діоду зі $j_0 \cong 2\cdot 10^{-3}$ А. Ефективність перетворення випадкового сигналу падаючої хвилі буде залежати від рівня сигналу шуму σ, та співвідношення хвильового опору Z_0 та диференційного опору діоду $(kT/qj_0)e^{(qU_0/kT)}$.

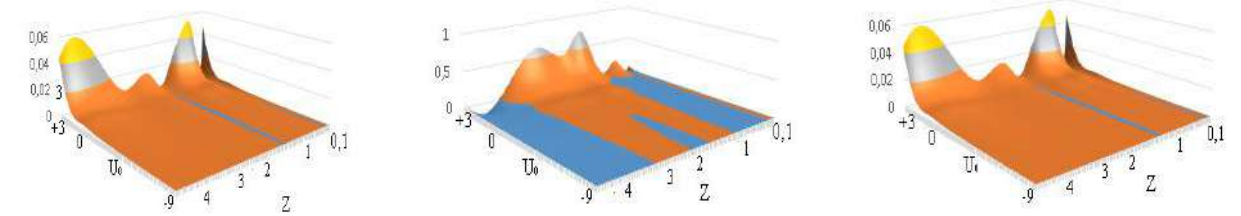
Всі розрахунки велися для рівня дисперсії генератору шуму $\sigma = 3$. Дисперсія виміряється в одиницях термічного потенціалу. Диференційний опір залежить від напруги, яка зміщує робочу точку вольт-амперної характеристики діоду U_0. Таким чином, змінюючи напругу U_0, теоретично можливо змінювати опір діоду від 0 до ∞. Рівень корисного сигналу від пасивної радіозакладки низький порівняно з активними радіозакладними пристроями. Тому для збільшення відстані, на якій можливий впевнений радіоприйом, треба використовувати широкосмугові приймачі. Моделювання такого прийому проводилося з урахуванням енергії відбитої хвилі на всій смуги низькочастотного максимуму $P_0(U_0, Z) = \int_0^{2\Delta\omega} W(\omega, U_0, Z)d\omega$; на спектральному максимумі на частоті падаючої хвилі $P_1(U_0, Z) = \int_{\omega_0-\Delta\omega}^{\omega_0+\Delta\omega} W(\omega, U_0, Z)d\omega$ та максимумі біля подвійної частоті від центральної частоти падаючої хвилі $P_2(U_0, Z) = \int_{2\omega_0-\Delta\omega}^{2\omega_0+\Delta\omega} W(\omega, U_0, Z)d\omega$. Залежність потужностей відбитої хвилі на різних максимумах спектральної щільності від напруги зміщення і хвилевого опору представлена на рис. 6, *а* – $P_0(U_0, Z)$, на рис. 6, *б* – $P_1(U_0, Z)$, на рис. 6, *в* – $P_2(U_0, Z)$.



а *б* *в*

Рис. 6

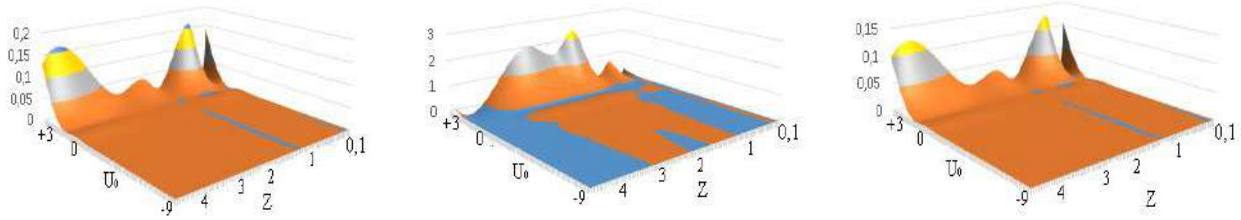
Для передачі інформації можливо використовувати два режими. Бінарний – це коли управляюча напруга приймає два значення. В цьому випадку вибираються два значення напруги, які зміщують робочу точку таким чином, щоб забезпечити максимальну різницю потужності перетвореного сигналу у вибраному діапазону частот. Цей режим більш відповідає передачі інформації цифровим сигналом. Ефективність бінарного режиму модуляції визначається максимумом різниці спектральної щільності потужності при прикладенні до діоду напруги зміщення $P(U_0, Z)$ (умовна логічна одиниця) та при прикладенні зворотної напруги зміщення $P(-9, z)$ (умовний логічний нуль). Передача інформації буде більш надійнішою, якщо логічний перепад $\Delta P_{\text{і}}(U_0, Z) = P_{\text{і}}(U_0, Z) - P_{\text{і}}(-9, Z)$ буде максимальним. Залежність потужності $\Delta P_{\text{і}}(U_0, Z)$ від U_0 і Z представлено на рис. 7. Опір довгої лінії представлений в відносних одиницях $(kT)/(qj_0)$. В $\Delta P_{\text{і}}(U_0, Z)$ різні індекси відноситься до різних максимумів спектральної щільності. На рис. 7, *а* зображена залежність $\Delta P_0(U_0, Z)$ спектральної щільності потужності біля нульової частоти. Відповідно залежність $\Delta P_1(U_0, Z)$ спектральної щільності потужності біля частоти ω_0 – на рис. 7, *б*, а залежність $\Delta P_2(U_0, Z)$ спектральної щільності потужності частоти біля частоти $2\omega_0$ – на рис. 7, *в*.



а *б* *в*

Рис. 7

Другий режим передачі інформації більш придатний для передачі аналогової інформації. В цьому режимі управляючий сигнал складається з двох компонентів – зі сталої напруги, яка переміщує робочу точку у діапазон, в якому диференційна характеристика перетворення енергії падаючої хвилі буде більшою, та змінної напруги, якою є акустичний сигнал. Було получено диференційну залежність спектральної потужності від напруги U_0 та опору довгої лінії Z. На рис. 8 представлено залежність $\left|\frac{\partial P_i(U_0,Z)}{\partial U_0}\right|$ від напруги зміщення та різних опоров довгої лінії. Похідна береться поблизу максимумів біля частот 0, ω_0 і $2\omega_0$.



а *б* *в*

Рис. 8

Ефективність двох режимів найбільша при роботі в режимі прямого зміщення діоду. Залежності $\Delta P_{\text{і}}(U_0, Z)$, яка відповідає бінарному режиму модуляції, і $\frac{\partial P_i(U_0,Z)}{\partial U_0}$, яка відповідає аналоговому режиму модуляції від опору довгої лінії, мають немонотонний характер з декількома максимумами. Аналоговий режим модуляції має більшу ефективність, якщо хвильовій опір довгої лінії буде менший за опір нелінійного елемента, на якому відбувається перетворення шумового сигналу. Це справедливо для всіх трьох максимів спектральної щільності.

Висновки

Показано, що пасивний радіозакладний пристрій з використанням нелінійного елемента в полі радіошумових перешкод може передавати інформацію з використанням енергії цих радіошумових завад. Показана можливість та запропоновано режими передачі аналогової та цифрової інформації. Рівень сигналу завдяки нелінійному перетворенню енергії вузькосмугового сигналу радіоперешкод більший при прямому зміщенні діоду. Показано, що передача аналогового та бінарного сигналу в режимі амплітудної модуляції по високочастотному і низькочастотному каналах більш ефективна при прямому зміщенні діоду для співвідношення опорів діоду Z_d і довгої лінії $Z = Z_d/2$ і $Z = 4Z_d$. Передача інформації бінарним і аналоговим сигналом в смузі частот поблизу максимумів ω_0 ефективна для співвідношення опорів $3{,}5Z_d \geq Z \geq 2{,}5Z_d$.

Список літератури:

1. Энциклопедия промышленного шпионажа ; под общ. ред. Е.В. Куренкова. С.-Петербург : ООО «Изд-во Полигон», 1999. 515с.
2. Kalamkar S. S. and Banerjee A. Secure Communication via a Wireless Energy Harvesting Untrusted Relay // IEEE Transactions on Vehicular Technology. March 2017. Vol. 66, no. 3. P. 2199-2213,
3. Cao et K. al. Energy Harvesting Jammer Enabled Secure Communication for Cooperative NOMA Systems // 2020 International Conference on Wireless Communications and Signal Processing (WCSP), 2020. P. 801-806
4. United States Patent 8665607 Bouza, et al. Anti-eavesdropping device. H05K 7/14; H05K 7/18, March 4, 2014
5. Емельянов С. и др. Проблемные аспекты реализации пространственного и линейного зашумления в системах активной защиты информации // Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні. Харків, 2001. Вип. 2. С. 62-67.
6. Петров А.А. Оценка эффективности систем активной защиты в сетях общего пользования // Системы обробки інформації. Харьков, 2011. №4(94). С.174-178.
7. Патент РФ № 2732486 Способ радиоподавления систем когнитивных систем радиосвязи. Бюл. № 26 17.09.20.
8. Крижановський В.Г., Сергієнко С.П., Чернов Д.В., Крижановський В.В. Підслуховування NFC-зв'язку на частотах вищих гармонік // Радіотехніка. 2021. Вып. 204. С. 99-104.
9. Serhiienko S., Krizhanovski V. Modeling of the potential threat of unauthorized removal of information by a passive radio tab in the rooms protected by noise field // The Fourth International Conference on Information and Telecommunication Technologies and Radio Electronic (UkrMiCo'2019) 09–13 September 2019 Odessa, Ukraine.
10. Нейман Р.Л., Демирчан К.С. Теоретические основы радиотехники. Т.1. Ленинград : Энергоиздат, 1981. 536с.
11. Баскаков С.И. Радиотехнические цепи и сигналы. Москва : Высш. шк., 1983. 536с.

Надійшла до редколегії 09.03.2021

Відомості про авторів:

Крижановський Володимир Григорович – д-р техн. наук, професор, Донецький національний університет імені Василя Стуса (м. Вінниця), професор кафедри радіофізики та кібербезпеки; Україна; email: v.krizhanovski@donnu.edu.ua; ORCID: https://orcid.org/0000-0002-2685-9740

Сергієнко Сергій Петрович – канд. техн. наук, доцент, Донецький національний університет імені Василя Стуса (м. Вінниця), доцент кафедри радіофізики та кібербезпеки; Україна; email: s.serhiienko@donnu.edu.ua; ORCID: https://orcid.org/0000-0001-5515-8946

Чернов Дмитро Вікторович – канд. техн. наук, Донецький національний університет імені Василя Стуса (м. Вінниця), доцент кафедри радіофізики та кібербезпеки; Україна; email: d.chernov@donnu.edu.ua; ORCID: https://orcid.org/0000-0001-7173-0842

Загоруйко Любов Василівна – канд. техн. наук, Донецький національний університет імені Василя Стуса (м. Вінниця), доцент кафедри радіофізики та кібербезпеки; Україна; email: l.zahoruiko@donnu.edu.ua; ORCID: https://orcid.org/0000-0002-6958-8696

УДК 621.396 **DOI:10.30837/rt.2021.2.205.19**

А.А. КУЗНЕЦОВ, д-р техн. наук, А.А. СМИРНОВ, д-р техн. наук, Т.Ю. КУЗНЕЦОВА

ШУМОПОДОБНЫЕ ДИСКРЕТНЫЕ СИГНАЛЫ ДЛЯ АСИНХРОННЫХ СИСТЕМ КОДОВОГО РАЗДЕЛЕНИЯ РАДИОКАНАЛОВ

Введение

Технология прямого расширения спектра (Direct-Sequence Spread Spectrum – DSSS) используется в системах радиосвязи с множественным доступом, системах глобального позиционирования, беспроводных сетях различного назначения и пр. [1 – 3]. Модуляция прямой расширяющей последовательностью делает передаваемый сигнал более широким по полосе пропускания, чем полоса пропускания информационного сигнала [4]. Это позволяет повысить устойчивость к непреднамеренному или намеренному заклиниванию (jamming), реализовать совместное использование одного канала несколькими пользователями, затруднить перехват и т.д. [1 – 3]. Технология расширения спектра прямой последовательностью используется в различных приложениях, например в спутниковых навигационных системах (GPS, Galileo и ГЛОНАСС), в системах множественного доступа с кодовым разделением каналов (CDMA), в сетях IEEE 802.11 и IEEE 802.15.4 и др. [3].

В основе технологии DSSS лежит использование длинных псевдослучайных последовательностей (дискретных сигналов), которые называются расширяющими. Информационные биты модулируются расширяющей последовательностью и сообщение скремблируется. Передаваемый расширенный сигнал похож на ограниченный по полосе белый шум, т.е. приобретает вид шумоподобного. При этом скорость передачи расширенного сигнала значительно выше исходной информационной скорости, т.е. полоса пропускания расширяется кратно длине расширяющей последовательности [2].

На принимающей стороне используется точная копия расширяющего сигнала. Вычисляя корреляцию между принятым сигналом и расширяющей последовательностью, приемник восстанавливает информационное сообщение [2].

Следует заметить, что для мобильных абонентов сложно обеспечить точную координацию линий связи между мобильными устройствами. Это означает, что корреляция может вычисляться для произвольно выбранных начальных точек (для любой копии циклически сдвинутого дискретного сигнала) [3]. В этом случае используют асинхронные техники, т.е. такие наборы расширяющих последовательностей (например, Голда (Gold), последовательности Касами (Kasami) и пр.), которые статистически некоррелированы для произвольно выбранных начальных точек [1 – 3]. В то же время кардинальность (мощность множества) известных наборов расширяющих последовательностей как правило невелика [5, 6].

В статье рассматриваются новые наборы расширяющих последовательностей для асинхронных систем кодового разделения радиоканалов, предложенные и подробно исследованные в [7 – 11]. Эти наборы имеют значительно большую кардинальность. При этом последовательности статистически некорелированы. Функция корреляции многозначна, причем максимальное абсолютное значение $\rho_{\max}$ асимптотически стремится к границе Велча (Welch) [12, 13].

1. Обзор литературы

Наиболее подробно теоретические положения и аспекты практического применения технологии прямого расширения спектра изложены в [2, 3, 14, 15]. Исследованию наборов расширяющих последовательностей посвящено много научных статей. Однако наибольшее практическое использование получили коды Голда [16, 17], а также большой и малый набор последовательностей Касами (Kasami) [3, 18, 19].

Коды Голда используются в глобальной спутниковой навигационной системе GPS для разделения сигналов космических аппаратов, в 3G системе мобильной связи стандарта WCDMA для скремблирования CDMA кодов и т. п. [2 – 4, 20]. Такая популярность кодов Голда объясняется, с одной стороны, исключительной простотой их генерации (используются простейшие переключательные схемы на основе линейных регистров сдвига). С другой стороны, последовательности Голда слабокоррелированы друг с другом, максимальное абсолютное значение функции корреляции очень близко к известной границе Велча (Welch) [12, 13, 21].

К сожалению, кардинальность кодов Голда невелика. Например, для длины последовательностей $N = 2^n - 1$ кардинальность $M = 2^n + 1$. При этом коды Голда имеют трехзначную функцию автокорреляции [4, 16, 17]:

$$\rho_{Gold} = \begin{cases} \dfrac{\varphi(t) - 2}{N}, \\ \dfrac{-1}{N}, \\ \dfrac{-\varphi(t)}{N}, \end{cases}$$

где максимальное абсолютное значение этой функции (а также функции взаимной корреляции)

$$\rho_{Gold\max} = \varphi(t) = \begin{cases} \dfrac{1 + 2^{\frac{n+1}{2}}}{N}, n = 2p + 1, \\ \dfrac{1 + 2^{\frac{n+2}{2}}}{N}, n = 2p. \end{cases} \quad (1)$$

Здесь и далее используем нормированные относительно длины N функции корреляции.

Другой хороший пример расширяющих последовательностей, часто используемый в системах DSSS, - это большой и малый набор последовательностей Касами (Kasami) [3, 18, 19].

Для малого набора кодов Касами (с периодом $N = 2^n - 1$) кардинальность $M = 2^{\frac{n}{2}}$. При этом

$$\rho_{SKasami\max} = \frac{1 + 2^{\frac{n}{2}}}{N}, n = 2p. \quad (2)$$

Для большого набора кодов Касами кардинальность $M = 2^{\frac{n}{2}}(2^{\frac{n}{2}} + 1)$ и

$$\rho_{GKasami\max} = \frac{1 + 2^{\frac{n+2}{2}}}{N}, n = 2p. \quad (3)$$

Таким образом, кардинальность наборов расширяющих последовательностей Голда и Касами сопоставима с периодом последовательностей $M \approx N$. Это вносит ограничения на емкость (Capacity) множественного доступа [22 – 24]. Если каждой паре абонентов выделена одна расширяющая последовательность, тогда число одновременно обслуживаемых пар абонентов ограничено $M \approx N$.

Например, для $n = 10$ период последовательности $N = 1023$ и:

- кардинальность набора кодов Голда $M = 1025$;

- кардинальность малого набора кодов Касами $M = 32$;
- кардинальность большого набора кодов Касами $M = 1056$

и емкость множественного доступа с кодовым разделением каналов в зоне действия одной базовой станции будет ограничена примерно 10^3.

Для повышения емкости множественного доступа нужно либо увеличить N (что приведет к увеличению временных затрат на обработку каждого информационного сообщения). Либо нужны новые наборы расширяющих последовательностей, которые для таких же значений N обладают повышенной кардинальностью $M > N$. При этом расширяющие последовательности должны быть слабокоррелированы друг с другом для произвольно выбранных начальных точек [13].

Таким образом, задача генерации расширяющих последовательностей для асинхронных CDMA состоит в формировании большого числа псевдослучайных последовательностей (обычно рассматриваются двоичные векторы с элементами 1 или -1) с особыми корреляционными свойствами [7, 10], т.е. необходимо максимизировать кардинальность M набора последовательностей длины N, для которых взаимная корреляция не превосходит $\rho_{\max}$.

Известной фундаментальной границей, связывающей M, N и $\rho_{\max}$ является граница Велча [12, 13, 21].

2. Граница Велча

Граница Велча впервые введена в работе [12]. Эта граница устанавливает ограничение на квадрат корреляции $(\rho_{\max})^2$ различных последовательностей при заданной длине N и кардинальности M.

Пусть $\{x_1, \ldots, x_M\}$ являются двоичными векторами длины N (все элементы каждого вектора x_i равны 1 или -1).

Определим

$$\rho_{\max} = \frac{\max\limits_{i \neq j} \langle x_i, x_j \rangle}{N},$$

где $\langle x_i, x_j \rangle$ – скалярное произведение векторов x_i и x_j.

Тогда для $k = 1, 2, \ldots$ справедлива граница (Велча) [12]:

$$(\rho_{\max})^{2k} \geq \frac{1}{M-1}\left[\frac{M}{\binom{N+k-1}{k}} - 1\right].$$

Очевидно, что для $k = 1$ имеем границу на квадрат взаимной корреляции последовательностей x_i:

$$(\rho_{\max})^2 \geq \frac{1}{M-1}\left[\frac{M}{N} - 1\right] = \frac{M-N}{(M-1)N}. \quad (4)$$

Для асинхронных CDMA все последовательности x_i должны быть статистически некоррелированы для произвольно случайных начальных точек. Другими словами, помимо каждой последовательности x_i мы должны рассмотреть также все циклически сдвинутые копии x_i. Тогда в границе (4) нужно заменить M на MN, что дает [13]:

$$(\rho_{\max})^2 \geq \frac{MN-N}{(MN-1)N} = \frac{M-1}{MN-1}. \quad (5)$$

Граница (5) определяет фундаментальный предел, ниже которого квадрат взаимной корреляции между любыми циклически сдвинутыми копиями разных последовательностей опуститься не может.

Для случая $M >> 1$ граница (5) принимает очень простой вид [13]:

$$\left(\rho_{\max}\right)^2 \geq \frac{1}{N} \tag{6}$$

и чаще всего модуль взаимной корреляции сгенерированных сигналов сравнивают именно с величиной $\frac{1}{\sqrt{N}}$. Асимптотически $\lim_{N\to\infty}\left(\frac{1}{N}\right) = 0$.

3. Коды Голда, множества Касами и граница Велча

3.1. Коды Голда и граница Велча

Рассмотрим границу Велча для параметров кодов Голда:

$$N = 2^n - 1,\ M = 2^n + 1 = N + 2.$$

Подставляя в (5), получим:

$$\left(\rho_{\max}\right)^2 \geq \frac{N+1}{(N+2)N-1},$$

причем

$$\lim_{N\to\infty}\left(\rho_{\max}\right)^2 = \lim_{N\to\infty}\left(\frac{N+1}{(N+2)N-1}\right) = 0.$$

Разложение в ряд Лорана (Laurent series) имеет вид

$$\left(\rho_{\max}\right)^2 \approx \frac{1}{N} - \left(\frac{1}{N}\right)^2 + \left(\frac{3}{N}\right)^3 - \left(\frac{7}{N}\right)^4 + \left(\frac{17}{N}\right)^5 - \left(\frac{41}{N}\right)^6 +$$
$$+ \left(\frac{99}{N}\right)^7 - \left(\frac{239}{N}\right)^8 + \left(\frac{577}{N}\right)^9 - \left(\frac{1393}{N}\right)^{10} + O\left(\left(\frac{1}{N}\right)^{11}\right)$$

и это хорошо аппроксимирует границу (6).

Реальные значения взаимной корреляции для кодов Голда соответствуют (1). Например, для нечетных n имеем:

$$\left(\rho_{Gold\max}\right)^2 = \left(\frac{1 + 2^{\frac{n+1}{2}}}{2^n - 1}\right)^2 = \left(\frac{\sqrt{2(N+1)} + 1}{N}\right)^2.$$

Аппроксимация с помощью ряда Пюизо (Puiseux series) дает

$$\left(\rho_{Gold\max}\right)^2 \approx \frac{2}{N} + 2\sqrt{2}\left(\frac{1}{N}\right)^{(3/2)} + \frac{3}{N^2} +$$
$$+ \sqrt{2}\left(\frac{1}{N}\right)^{(5/2)} - \frac{\left(\frac{1}{N}\right)^{(7/2)}}{2\sqrt{2}} + \frac{\left(\frac{1}{N}\right)^{(9/2)}}{4\sqrt{2}} + O\left(\left(\frac{1}{N}\right)^{(11/2)}\right).$$

Используя только первый член ряда, имеем $\left(\rho_{Gold\max}\right)^2 \approx \frac{2}{N}$. Это означает, что для больших значений N взаимная корреляция кодов Голда $\approx \frac{1.4}{\sqrt{N}}$. Это очень близко к границе (6).

В пределе имеем

$$\lim_{N\to\infty}\left(\rho_{Gold\max}\right)^2 = \lim_{N\to\infty}\left(\left(\frac{\sqrt{2(N+1)}+1}{N}\right)^2\right) = 0,$$

т.е. асимптотически (при $N \to \infty$) коды Голда удовлетворяют границе Велча.

3.2. Множества Касами и граница Велча

Рассмотрим границу Велча для малого набора кодов Касами:

$$N = 2^n - 1, \quad M = 2^{\frac{n}{2}} = \sqrt{N+1}.$$

Подставляя в (5), получим:

$$\left(\rho_{\max}\right)^2 \geq \frac{\sqrt{N+1}-1}{\sqrt{N+1}N-1},$$

причем

$$\lim_{N\to\infty}\left(\rho_{\max}\right)^2 = \lim_{N\to\infty}\left(\frac{\sqrt{N+1}-1}{\sqrt{N+1}N-1}\right) = 0.$$

Разложение в ряд Пюизо (Puiseux series) имеет вид

$$\left(\rho_{\max}\right)^2 \approx \frac{1}{N} - \left(\frac{1}{N}\right)^{3/2} + \frac{3}{2}\left(\frac{1}{N}\right)^{5/2} - \left(\frac{1}{N}\right)^3 -$$
$$-\frac{7}{8}\left(\frac{1}{N}\right)^{7/2} + \frac{2}{N^4} - \frac{5}{16}\left(\frac{1}{N}\right)^{9/2} - \frac{2}{N^5} + O\left(\left(\frac{1}{N}\right)^{11/2}\right),$$

что хорошо аппроксимирует границу (6).

Реальные значение взаимной корреляции соответствуют (2), т.е.:

$$\left(\rho_{SKasami\max}\right)^2 = \left(\frac{\sqrt{N+1}+1}{N}\right)^2.$$

Аппроксимация с помощью ряда Пюизо (Puiseux series) дает

$$\left(\rho_{SKasami\max}\right)^2 \approx \frac{1}{N} + 2\left(\frac{1}{N}\right)^{3/2} + \frac{2}{N^2} +$$
$$+\left(\frac{1}{N}\right)^{5/2} - \frac{1}{4}\left(\frac{1}{N}\right)^{7/2} + \frac{1}{8}\left(\frac{1}{N}\right)^{9/2} + O\left(\left(\frac{1}{N}\right)^{11/2}\right).$$

Аппроксимируя по первому члену ряда, получаем $\left(\rho_{SKasami\max}\right)^2 \approx \frac{1}{N}$, что совпадает с (6).

Для большого набора кодов Касами кардинальность

$$M = 2^n + 2^{n/2} = N + 1 + \sqrt{N+1},$$

т.е. граница имеет вид

$$\left(\rho_{\max}\right)^2 \geq \frac{N+\sqrt{N+1}}{(N+1+\sqrt{N+1})N-1} \approx$$

$$\approx \frac{1}{N} - \left(\frac{1}{N}\right)^2 + \left(\frac{1}{N}\right)^{5/2} + \left(\frac{1}{N}\right)^3 - \frac{3}{2}\left(\frac{1}{N}\right)^{7/2} -$$

$$-\frac{1}{N^4} + \frac{23}{8}\left(\frac{1}{N}\right)^{9/2} + O\left(\left(\frac{1}{N}\right)^{11/2}\right) \approx \frac{1}{N},$$

что соответствует (6).

Для реальных значений взаимной корреляции (3) имеем:

$$\left(\rho_{GKasami\max}\right)^2 = \left(\frac{1+2\sqrt{N+1}}{N}\right)^2 \approx \frac{4}{N} + 4\left(\frac{1}{N}\right)^{3/2} + \frac{5}{N^2} +$$

$$+2\left(\frac{1}{N}\right)^{5/2} - \frac{1}{2}\left(\frac{1}{N}\right)^{7/2} + \frac{1}{4}\left(\frac{1}{N}\right)^{9/2} + O\left(\left(\frac{1}{N}\right)^{11/2}\right) \approx \frac{4}{N}.$$

Итоговые результаты оценок кардинальности и модуля корреляции рассмотренных последовательностей сведены в табл. 1.

Таблица 1

Оценка квадрата корреляции расширяющих последовательностей

Параметр	Кардинальность набора, M	Оценка модуля корреляции по первому члену ряда
Коды Голда	$N+2$	$\frac{1.4}{\sqrt{N}}$
Малое множество Касами	$\sqrt{N+1}$	$\frac{1}{\sqrt{N}}$
Большое множество Касами	$N+1+\sqrt{N+1}$	$\frac{2}{\sqrt{N}}$
Новый набор	$(N+1)^2+N+2$	$\frac{2.8}{\sqrt{N}}$

Очевидно, что даже незначительное увеличение кардинальности наборов расширяющих последовательностей приводит к кратному увеличению модуля корреляции. Например, только для малого множества Касами с $M=\sqrt{N+1}$ оценка корреляции сравнима с границей Велча. Для большого множества модуль корреляция возрастает в два раза. Хотя асимптотически (при $N\to\infty$) все наборы из табл. 1 удовлетворяют границе Велча $\lim\limits_{N\to\infty}\left(\rho_{\max}\right)^2 = 0$.

4. Новые наборы расширяющих последовательностей

Новые наборы последовательностей со специальными корреляционными свойствами рассмотрены в [7 – 11]. В частности, в [9, 10] представлена общая концепция генерации таких наборов, в [7, 11] изучены их корреляционные свойства, а в [8] – аспекты реализации. Генерация этих наборов последовательностей реализуется с помощью простейших переключательных устройств (регистров сдвига с линейными обратными связями).

В данной работе мы рассматриваем простейший случай расширяющих последовательностей с пятизначной функцией автокорреляции (для нечетных n) [7, 11]:

$$\rho_{New} = \begin{cases} \dfrac{-1-2^{\frac{n+1}{2}+1}}{2^n-1}; \\ \dfrac{-1-2^{\frac{n+1}{2}}}{2^n-1}; \\ \dfrac{-1}{2^n-1}; \\ \dfrac{-1+2^{\frac{n+1}{2}}}{2^n-1}; \\ \dfrac{-1+2^{\frac{n+1}{2}+1}}{2^n-1}. \end{cases}$$

Максимальное абсолютное значение этой функции (а также функции взаимной корреляции)

$$\rho_{New\max} = \frac{1+2^{\frac{n+1}{2}+1}}{2^n-1},\ n = 2p+1 \tag{7}$$

с параметрами

$$N = 2^n - 1,\ M = 2^{2n} + 2^n + 1 = (N+1)^2 + N + 2.$$

Подставляя эти параметры в (5), получим:

$$\left(\rho_{\max}\right)^2 \geq \frac{(N+1)^2 + N + 1}{((N+1)^2 + N + 2)N - 1}$$

что при аппроксимации рядом Лорана дает

$$\left(\rho_{\max}\right)^2 \approx \frac{1}{N} - \left(\frac{1}{N}\right)^3 + \left(\frac{4}{N}\right)^4 - \left(\frac{9}{N}\right)^5 + \left(\frac{14}{N}\right)^6 -$$
$$- \left(\frac{11}{N}\right)^7 - \left(\frac{18}{N}\right)^8 + \left(\frac{101}{N}\right)^9 - \left(\frac{260}{N}\right)^{10} + O\left(\left(\frac{1}{N}\right)^{11}\right).$$

В пределе

$$\lim_{N\to\infty}\left(\rho_{\max}\right)^2 = \lim_{N\to\infty}\left(\frac{(N+1)^2 + N + 1}{((N+1)^2 + N + 2)N - 1}\right) = 0.$$

Используя реальное значение (7), получим:

$$\left(\rho_{New\max}\right)^2 = \left(\frac{1+2^{\frac{n+1}{2}+1}}{2^n-1}\right)^2 = \left(\frac{\sqrt{8(N+1)}+1}{N}\right)^2,$$

что при аппроксимации рядом Пюизо (Puiseux series) дает

$$\left(\rho_{New\max}\right)^2 \approx \frac{8}{N} + 4\sqrt{2}\left(\frac{1}{N}\right)^{(3/2)} + \frac{9}{N^2} + 2\sqrt{2}\left(\frac{1}{N}\right)^{(5/2)} -$$
$$- \frac{\left(\frac{1}{N}\right)^{(7/2)}}{\sqrt{2}} + \frac{\left(\frac{1}{N}\right)^{(9/2)}}{2\sqrt{2}} + O\left(\left(\frac{1}{N}\right)^{(11/2)}\right).$$

Это означает, что для больших значений N взаимная корреляция таких кодов $\approx \dfrac{2.8}{\sqrt{N}}$ (при аппроксимации первым членом ряда).

В пределе

$$\lim_{N\to\infty}\left(\rho_{\max}\right)^2=\lim_{N\to\infty}\left(\frac{\sqrt{8(N+1)}+1}{N}\right)^2=0,$$

т.е. асимптотически (при $N\to\infty$) новые наборы удовлетворяют границе Велча.

В последней строке табл. 1 приведены значения кардинальности новых наборов и оценки модуля корреляции по первому члену ряда. Очевидно, что в сравнении с другими наборами мы имеем значительное повышение кардинальности. Например, в сравнении с кодами Голда кардинальность увеличивается более чем в N раз. При этом модуль корреляции увеличивается примерно в два раза.

Выводы

Рассмотрены несколько наборов (коды Голда, большое и малое множество Касами, наборы из [7 – 11]) расширяющих последовательностей для возможного использования в асинхронных системах кодового разделения радиоканалов. Асинхронность предполагает использование последовательностей, которые статистически не коррелированны для произвольной циклически сдвинутой копии сигналов. Фундаментальным теоретическим пределом для этой характеристики является известная граница Велча.

Проведено сравнение корреляционных свойств различных наборов с этим фундаментальным пределом. Проведена оценка параметров разных кодов, приведены соответствующая граница и сравнение ее с реальными корреляционными характеристиками кодов. Для аппроксимации использовалось разложение в ряд Лорана и ряд Пюизо. Также оценивались асимптотические свойства, т.е. при $N\to\infty$.

Исследования показали, что асимптотически все рассмотренные наборы кодов удовлетворяют границе Велча, т.е. при $N\to\infty$ квадрат корреляции стремится к нулю.

При аппроксимации первым членом ряда удалось показать различия в корреляционных характеристиках расширяющих наборов. При этом все наборы близки к теоретическому пределу $\frac{1}{\sqrt{N}}$ модуля корреляции.

Новый расширенный набор последовательностей с пятизначной функцией корреляции также асимптотически удовлетворяют границе Велча. При аппроксимации первым членом ряда модуль корреляции ограничен $\frac{2.8}{\sqrt{N}}$, что в два раза превышает значение для кодов Голда. Однако при этом кардинальность нового набора значительно (в N раз) выше. Это существенное преимущество, которое позволит значительно повысить емкость асинхронных CDMA и удешевить услуги связи. Например, при $N=1000$ достигается расширение набора более, чем на три порядка, и это может оказаться существенным.

Новые наборы расширяющих сигналов можно использовать для мягкой емкости (Soft Capacity), т.е. базовая станция может увеличить абонентскую емкость при незначительном снижении качества обслуживания.

Список литературы:

1. Yang S.-M.M. Modern Digital Radio Communication Signals and Systems. Springer International Publishing, 2019.
2. Torrieri D. Principles of Spread-Spectrum Communication Systems. Springer International Publishing, 2018.
3. Spread Spectrum and CDMA: Principles and Applications | Wiley [Electronic resource] // Wiley.com. URL: https://www.wiley.com/en-us/Spread+Spectrum+and+CDMA%3A+Principles+and+Applications-p-9780470091784 (accessed: 01.08.2020).
4. Sklar B., Harris F.J. Digital Communications: Fundamentals and Applications. 3 edition. Hoboken: Prentice Hall, 2020. 1104 p.
5. Khalife J., Kassas Z.M. Navigation With Cellular CDMA Signals – Part II: Performance Analysis and Experimental Results // IEEE Transactions on Signal Processing. 2018. Vol. 66, № 8. P. 2204–2218.
6. Sklar B. Digital Communications: Fundamentals and Applications. Edición: 2. Upper Saddle River, NJ: Prentice Hall, 2017. 1104 p.
7. Kuznetsov A. et al. Formation of Discrete Signals with Special Correlation Properties // 2019 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo). Odessa, Ukraine: IEEE, 2019. P. 1–6.

8. Kuznetsov A. et al. Generators of Pseudorandom Sequence with Multilevel Function of Correlation // 2019 IEEE International Scientific-Practical Conference Problems of Infocommunications, Science and Technology (PIC S T). 2019. P. 517–522.
9. Kuznetsov A. et al. Formation of Pseudorandom Sequences with Special Correlation Properties // 2019 3rd International Conference on Advanced Information and Communications Technologies (AICT). 2019. P. 395–399.
10. Kuznetsov A. et al. Discrete Signals with Special Correlation Properties // Proceedings of the Second International Workshop on Computer Modeling and Intelligent Systems (CMIS-2019), Zaporizhzhia, Ukraine, April 15-19, 2019 / ed. Luengo D. et al. CEUR-WS.org, 2019. Vol. 2353. P. 618–629.
11. Kuznetsov A. et al. Pseudorandom Sequences with Multi-Level Correlation Function for Direct Spectrum Spreading // 2019 IEEE International Conference on Advanced Trends in Information Theory (ATIT). 2019. P. 232–237.
12. Welch L. Lower bounds on the maximum cross correlation of signals (Corresp.) // IEEE Transactions on Information Theory. 1974. Vol. 20, № 3. P. 397–399.
13. Ipatov V.P. Spread Spectrum and CDMA: Principles and Applications. Chichester, UK: John Wiley & Sons, Ltd, 2005.
14. Rao R., Dianat S. Basics of Code Division Multiple Access (CDMA). 1000 20th Street, Bellingham, WA 98227-0010 USA: SPIE, 2005.
15. Buehrer R.M. Code Division Multiple Access (CDMA). Morgan & Claypool Publishers, 2006. 192 p.
16. Gold R. Optimal binary sequences for spread spectrum multiplexing (Corresp.) // IEEE Transactions on Information Theory. 1967. Vol. 13, № 4. P. 619–621.
17. Hamid M., Miller A. Gold Code Generators in Virtex Devices [Electronic resource]. 2000. URL: /paper/Gold-Code-Generators-in-Virtex-Devices-Hamid-Miller/9ce406a10eb3ae90edd8fa20590a0dcd8ed03c86 (accessed: 01.08.2020).
18. Kasami T. Weight Distribution Formula for Some Class of Cyclic Codes. Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1966.
19. Shi M., Krotov D.S., Solé P. A New Approach to the Kasami Codes of Type 2 // IEEE Transactions on Information Theory. 2020. Vol. 66, № 4. P. 2456–2465.
20. GPS explained: Transmitted GPS Signals [Electronic resource] // archive.is. 2012. URL: http://archive.is/eC7C (accessed: 01.08.2020).
21. Massey J.L., Mittelholzer T. Welch's Bound and Sequence Sets for Code-Division Multiple-Access Systems // Sequences II / ed. Capocelli R., De Santis A., Vaccaro U. New York, NY: Springer, 1993. P. 63–78.
22. Stüber G.L. Spread Spectrum Techniques // Principles of Mobile Communication / ed. Stüber G.L. Cham: Springer International Publishing, 2017. P. 449–499.
23. Torrieri D. Chapter 7 Code-Division Multiple Access // Principles of Spread-Spectrum Communication Systems / ed. Torrieri D. Cham: Springer International Publishing, 2015. P. 405–460.
24. Song T., Zhou K., Li T. CDMA System Design and Capacity Analysis Under Disguised Jamming // IEEE Transactions on Information Forensics and Security. 2016. Vol. 11, № 11. P. 2487–2498.
25. Korhonen J., You J. Peak signal-to-noise ratio revisited: Is simple beautiful? // 2012 Fourth International Workshop on Quality of Multimedia Experience. 2012. P. 37–38.
26. Kuznetsov A. et al. Adaptive Pseudo-Random Sequence Generation for Spread Spectrum Image Steganography // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2020. P. 161–165.

Поступила в редколлегию 28.03.2021

Сведения об авторах:

Кузнецов Александр Александрович – д-р техн. наук, профессор, Харьковский национальный университет имени В.Н. Каразина, профессор кафедры безопасности информационных систем и технологий, факультет компьютерных наук; Украина; e-mail: kuznetsov@karazin.ua; ORCID: https://orcid.org/0000-0003-2331-6326

Смирнов Алексей Анатольевич – д-р техн. наук, профессор, Центральноукраинский национальный технический университет, м. Кропивницкий, заведующий кафедрой кибербезопасности и программного обеспечения; Украина; e-mail: dr.smirnovoa@gmail.com; ORCID: https://orcid.org/0000-0001-9543-874X

Кузнецова Татьяна Юрьевна – Харьковский национальный университет имени В.Н. Каразина, научный сотрудник научно-исследовательской части; Украина; e-mail: kuznetsova.tatiana17@gmail.com; ORCID: https://orcid.org/0000-0001-6154-7139

МЕТОДИ ТА АЛГОРИТМИ КРИПТОГРАФІЧНОГО ЗАХИСТУ ІНФОРМАЦІЇ
МЕТОДЫ И АЛГОРИТМЫ КРИПТОГРАФИЧЕСКОЙ ЗАЩИТЫ ИНФОРМАЦИИ
METHODS AND ALGORITHMS OF CRYPTOGRAPHIC PROTECTION OF INFORMATION

УДК 004.056.55

Основні положення та результати порівняння властивостей електронних підписів постквантового періоду на алгебраїчних решітках / *І.Д. Горбенко, О.Г. Качко, О.В. Потій, А.М. Олексійчук, Ю.І. Горбенко, М.В. Єсіна, І.В. Стельник, В А. Пономар* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 5 – 21.

Розглядаються постквантові проєкти стандартів електронних підписів (ЕП) Falcon та Dilithium, які є фіналістами конкурсу NIST США. При їх побудуванні використовується математичний апарат алгебраїчних решіток та відповідні методи. При подальшому дослідженні та порівнянні вказаних постквантових проєктів стандартів ЕП, як з теоретичних, так і практичних позицій, основоположним є обґрунтування вимог до параметрів та ключів, та у цілому обчислення основних показників згідно прийнятих умовних та безумовних критеріїв. Важливим при таких дослідженнях є визначення достатності забезпечення гарантованості їх захищеності від класичних, квантових, спеціальних та атак на основі помилок. Вказане може бути забезпечено, у тому числі, засобом обґрунтованого вибору розмірів загальних параметрів та ключів, та практичного їх побудування згідно прийнятої моделі безпеки. Але при виборі розмірів загальних параметрів та ключів виникає суттєве протиріччя між властивостями проєктів стандартів ЕП Falcon та Dilithium, щодо стійкості та складності перетворень. Так, збільшення розмірів загальних параметрів та ключів приводить до збільшення складності перетворень, і навпаки. Метою цієї статі є: аналіз проблемних питань вибору розмірів параметрів та ключів для постквантових проєктів ЕП, побудованих на основі математичних методів Falcon та Dilithium, та особливості їх реалізації, в тому числі і реалізації згідно прийнятої моделі безпеки. Порівняльний аналіз стійкості та складності проєктів стандартів ЕП Falcon та Dilithium у залежності від розмірів параметрів та ключів, в тому числі для 6 та 7 рівнів безпеки. Розробка пропозицій стосовно рішень щодо прийняття в якості національних постквантових стандартів ЕП на основі математичних методів Falcon та Dilithium. Визначення впливу безумовних, умовних та прагматичних критеріїв на переваги при прийнятті рішення щодо стандартизації ЕП на основі математичних методів Falcon та Dilithium, в тому числі з урахуванням наявності патентів та необхідності отримання ліцензій тощо.

Ключові слова: алгебраїчні решітки; алгоритм; електронний підпис; основні параметри; постквантова криптографія.

Табл. 12. Іл. 2. Бібліогр.: 24 назв.

УДК 004.056.55

Основные положения и результаты сравнения свойств электронных подписей постквантового периода на алгебраических решетках / *И.Д. Горбенко, Е.Г. Качко, А.В. Потий, А.Н. Олексийчук, Ю.И. Горбенко, М.В. Есина, И.В. Стельник, В.А. Пономарь* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 5 – 21.

Рассматриваются постквантовые проекты стандартов электронных подписей (ЭП) Falcon и Dilithium, которые являются финалистами конкурса NIST США. При их построении используется математический аппарат алгебраических решеток и соответствующие методы. При дальнейшем исследовании и сравнении указанных постквантовых проектов стандартов ЭП, как с теоретических, так и практических позиций, основополагающим является обоснование требований к параметрам и ключам, и в целом вычисления основных показателей согласно принятых условных и безусловных критериев. Важным при таких исследованиях является определение достаточности обеспечения гарантированности их защищенности от классических, квантовых, специальных и атак на основе ошибок. Это может быть обеспечено, в том числе, посредством обоснованного выбора размеров общих параметров и ключей, и практического их построения согласно принятой модели безопасности. Но при выборе размеров общих параметров и ключей возникает существенное противоречие между свойствами проектов стандартов ЭП Falcon и Dilithium, по устойчивости и сложности преобразований. Так, увеличение размеров общих параметров и ключей приводит к увеличению сложности преобразований, и наоборот. Цель этой статьи: анализ проблемных вопросов выбора размеров параметров и ключей для постквантовых проектов ЭП, построенных на основе математических методов Falcon и Dilithium, и особенности их реализации, в том числе и реализации согласно принятой модели безопасности. Сравнительный анализ устойчивости и сложности проектов стандартов ЭП Falcon и Dilithium в зависимости от размеров параметров и ключей, в том числе для 6 и 7 уровней безопасности. Разработка предложений относительно решений о принятии в качестве национальных постквантовых стандартов ЭП на основе математических методов Falcon и Dilithium. Определение влияния безусловных, условных и прагматических критериев на преимущества при принятии решения о стандартизации ЭП на основе математических методов Falcon и Dilithium, в том числе с учетом наличия патентов и необходимости получения лицензий и тому подобное.

Ключевые слова: алгебраические решетки; алгоритм; электронная подпись; основные параметры; постквантовая криптография.

Табл. 12. Ил. 2. Библиогр.: 24 назв.

UDC 004.056.55

Basic principles and results of comparison of electronic signatures properties of the postquantum period based on algebraic lattices / *I.D. Gorbenko, O.G. Kachko, O.V. Potii, A.M. Oleksiychuk, Yu.I. Gorbenko, M.V. Yesina, I.V. Stelnyk, V.A. Ponomar* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 5 – 21.

The paper considers post-quantum projects of the Falcon and Dilithium electronic signature standards (ES), which are finalists of the NIST USA competition. The mathematical apparatus of algebraic lattices and appropriate methods are used in their construction. In further study and comparison of these post-quantum ES draft standards, both from a theoretical and practical standpoint, it is fundamental to substantiate the requirements for parameters and keys and in general to calculate the main indicators according to the accepted conditional and unconditional criteria. In such studies, it is important to determine the sufficiency of ensuring the guarantee of their security against classical, quantum, special and error-based attacks. This can be ensured, inter alia, through a reasonable choice of the sizes of common parameters and keys, and their practical construction in accordance with the adopted security model. However, when choosing the sizes of common parameters and keys, a significant contradiction arises between the properties of the draft of the Falcon and Dilithium ES standards, So increasing the size of the general parameters and keys leads to an increase in the complexity of transformations, and vice versa. The purpose of this article consists in analysis of problematic issues of choosing the size of parameter and keys for post-quantum ES projects based on mathematical methods of Falcon and Dilithium, and features of their implementation, including implementation according to the adopted security model. Comparative analysis of the stability and complexity of the Falcon and Dilithium ES draft standards depending on the size of the parameters and keys, including for 6 and 7 security levels. Development of proposals for decisions on the adoption of national post-quantum ES standards based on the mathematical methods Falcon and Dilithium. Determining the influence of unconditional, conditional and pragmatic criteria on the advantages when deciding on the ES standardization based on Falcon and Dilithium mathematical methods, including taking into account the availability of patents and the need to obtain licenses, etc.

Key words: algebraic lattices; algorithm; electronic signature; basic parameters; post-quantum cryptography.

12 tab. 2 fig. Ref: 24 items.

УДК 621.391:519.2

Оцінки ефективності атак на основі підібраних відкритих текстів на криптосистему Рао-Нама над скінченною абелевою групою / *А.М. Олексійчук, О.С. Шевчук* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 22 – 31.

Криптосистема Рао – Нама являє собою симетричну версію кодової криптосистеми Мак-Еліса, запропоновану з метою позбутися слабкостей, притаманних найпершим симетричним кодовим схемам шифрування. Майже одразу після опублікування цієї криптосистеми з'явилися атаки на неї на основі підібраних відкритих текстів, що привело до появи різноманітних удосконалень та модифікацій оригінальної криптосистеми.

Секретним ключем у традиційній схемі Рао – Нама є певна булева матриця та множина двійкових векторів, які використовуються для формування спотворень при зашифруванні. Такі вектори повинні мати різні синдроми, тобто бути різними за модулем коду, породженому рядками зазначеної матриці. В оригінальній роботі Рао і Нама розглянуто два способи формування множини цих векторів, перший з яких полягає у використанні заздалегідь визначених векторів достатньо великої ваги, а другий – у випадковому виборі цих векторів за рівноймовірною схемою. Відомо, що перший варіант не забезпечує належну стійкість криптосистеми Рао – Нама (внаслідок невеликої кількості та простої будови зазначених векторів), проте другий варіант є більш змістовним та потребує додаткових досліджень.

Мета статті – отримання оцінок ефективності (трудомісткості при заданій верхній межі ймовірності помилки) атак на криптосистему, яка узагальнює традиційну схему Рао – Нама на випадок скінченної абелевої групи (зауважимо, що необхідність дослідження подібних версій криптосистеми Рао – Нама обумовлена їх розглядом у нещодавніх публікаціях). Представлено дві атаки, які будуються на основі підібраних відкритих текстів. Перша з них не згадується у відомих авторам цієї статті працях і за певних, точно визначених умов дозволяє відновлювати секретний ключ криптосистеми із квадратичною складністю.

Друга атака являє собою узагальнено-спрощений варіант відомої атаки Стройка-ван Тілбурга. Показано, що складність цієї атаки залежить від потужності стабілізатора множини векторів, яка утворює другу частину ключа, у групі зсувів абелевої групи, над якою розглядається криптосистема Рао – Нама. Отримано оцінку ймовірності тривіальності стабілізатора за умови випадкового вибору цієї множини. З отриманої оцінки випливає, що атака Стройка-ван Тілбурга є в середньому помітно більш ефективною в порівнянні із найгіршим випадком, розглянутим раніше.

Ключові слова: кодова криптографія; криптосистема Рао – Нама; атака на основі підібраних відкритих текстів; атака Стройка-ван Тілбурга.

Бібліогр.: 10 назв.

УДК 621.391:519.2

Оценки эффективности атак на основе подобранных открытых текстов на криптосистему Рао – Нама над конечной абелевой группой / *А.Н. Алексейчук, О.С. Шевчук* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 22 – 31.

Криптосистема Рао – Нама представляет собой симметричную версию кодовой криптосистемы Мак-Элиса, предложенную с целью избавиться от недостатков, присущих первым симметричным кодовым схемам шифрования. Почти сразу после опубликования этой криптосистемы на нее появились атаки на основе подобранных открытых текстов, что привело к появлению различных усовершенствований и модификаций оригинальной криптосистемы.

Секретным ключом в традиционной схеме Рао – Нама являются определенная булева матрица и множество двоичных векторов, используемых для формирования искажений при зашифровании. Такие векторы должны иметь различные синдромы, то есть быть разными по модулю кода, порожденного строками указанной матрицы. В оригинальной работе Рао и Нама рассмотрены два способа формирования множества этих векторов, первый из которых заключается в использовании заранее определенных векторов достаточно большого веса, а второй – в случайном выборе этих векторов по равновероятной схеме. Известно, что первый вариант не обеспечивает надлежащую стойкость криптосистемы Рао – Нама (вследствие небольшого количества и простого строения указанных векторов), однако второй вариант является более содержательным и требует дополнительных исследований.

Цель статьи – получение оценок эффективности (трудоемкости при заданной верхней границе вероятности ошибки) атак на криптосистему, которая обобщает традиционную схему Рао – Нама на случай конечной абелевой группы (заметим, что необходимость исследования подобных версий криптосистемы Рао – Нама обусловлена их рассмотрением в недавних публикациях). Представлены две атаки, которые строятся на основе подобранных открытых текстов. Первая из них не упоминается в известных авторам трудах и при некоторых условиях позволяет восстанавливать секретный ключ криптосистемы с квадратичной сложностью.

Вторая атака представляет собой обобщенно-упрощенный вариант известной атаки Стройка-ван Тилбурга. Показано, что сложность этой атаки зависит от мощности стабилизатора множества векторов, которое образует вторую часть ключа в группе сдвигов абелевой группы, над которой рассматривается криптосистема Рао – Нама. В работе получена оценка вероятности тривиальности стабилизатора при условии случайного выбора этого множества. Из полученной оценки следует, что атака Стройка-ван Тилбурга является в среднем заметно более эффективной по сравнению с худшим случаем, рассмотренным ранее.

Ключевые слова: кодовая криптография; криптосистема Рао – Нама; атака на основе подобранных открытых текстов; атака Стройка-ван Тилбурга.

Библиогр.: 10 назв.

UDC 621.391:519.2

Evaluation of effectiveness of chosen-plaintext attacks on the Rao-Nam cryptosystem over a finite Abelian group / *A.N. Alekseychuk, O.S. Shevchuk* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 22 – 31.

The Rao-Nam cryptosystem is a symmetric version of the McEliece code-based cryptosystem proposed to get rid of the shortcomings inherent in the first symmetric code-based encryption schemes. Almost immediately after the publication of this cryptosystem, attacks on it based on selected plaintexts appeared, which led to the emergence of various improvements and modifications of the original cryptosystem.

The secret key in the traditional Rao-Nam scheme is a certain Boolean matrix and a set of binary vectors used to generate distortions during encryption. Such vectors must have different syndromes, that is, be different modulo of the code generated by the rows of the specified matrix. The original work of Rao and Nam considered two methods of forming the set of these vectors, the first of which consists in using predetermined vectors of sufficiently large weight, and the second is random selection of these vectors according to the equiprobable scheme. It is known that the first option does not provide the proper security of the Rao – Nam cryptosystem (due to the small number and simple structure of these vectors), but the second option is more meaningful and requires additional research. The purpose of this paper is to obtain estimates of the effectiveness (time complexity for a given upper bound of the error probability) of attacks on a cryptosystem, which generalizes the traditional Rao – Nam scheme to the case of a finite Abelian group (note that the need to study such versions of the Rao – Nam cryptosystem is due to their consideration in recent publications). Two attacks, based on selected plaintext, are presented. The first of them is not mentioned in the works known to the authors of this article and, under certain well-defined conditions, it allows recovering the secret key of the cryptosystem with quadratic complexity.

The second attack is a generalized and simplified version of the well-known Struik-van Tilburg attack. It is shown that the complexity of this attack depends on the power of the stabilizer of the set of vectors, which forms the second part of the key, in the translation group of the Abelian group, over which the Rao – Nam cryptosystem is considered. In this paper, a bound is obtained for the probability of triviality of the stabilizer under the condition of random choice of this set. From the obtained bound, it follows that Struik-van Tilburg attack is, on average, noticeably more efficient than the worst case considered earlier.

Key words: code-based cryptography; Rao – Nam cryptosystem; chosen-plaintext attack; Struik-van Tilburg attack.

Ref: 10 items.

УДК 004.056.5

Стеганографічні методи в векторній графіці / *О.О. Кузнецов, Г.В. Кононченко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 32 – 41.

Для приховування інформації застосовуються різні стеганографічні техніки. Зазвичай інформацію приховують у зображеннях, аудіо- та відеофайлах, текстових документах, тощо. В статті розглянуто векторні зобра-

ження, що складаються із різних математичних об'єктів (точки, лінії, криві першого та другого порядку, криві Без'є, вузли, дотичні, керуючі точки, тощо). Техніки приховування інформації змінюють ці математичні об'єкти, наприклад, через кодування координат базових точок. Найбільш вдалим для проведення стеганографічних перетворень є формат векторної графіки SVG, який завдяки своїй структурі дозволяє легко маніпулювати об'єктами, з яких складається. Його широка підтримка різними платформами також дозволяє підвищити рівень скритності при проведенні передачі секретних даних шляхом передачі звичайних на перший погляд файлів медіа. В статті розглянуто два методи (побітовий та метод паттернів) приховування інформації в векторні зображення, вивчено їх особливості, переваги та недоліки. Також досліджено різні афінні перетворення, які можна застосовувати для порушення роботи стеганосистеми. Найпоширенішими видами афінних перетворень є операції перенесення, повороту, зсуву та масштабування з можливими варіаціями (зсуву за осями абсцис та ординат, масштабування пропорційне та непропорційне, зі стисненням та із розширенням). Більшість методів вбудовування інформації у векторні зображення забезпечують одноразову стійкість до афінних перетворень, при цьому при повторному накладенні операцій зміни положення об'єктів, повідомлення може зруйнуватися взагалі. Досліджені в роботі методи реалізують більший рівень стійкості до різного роду перетворень при їх багаторазовому проведенні і проведені експерименти це наочно доводять. Отримані результати показують, що векторні зображення дійсно можуть застосовуватися для приховування інформації, але стійкість проти певних афінних атак не завжди є високою.

Ключові слова: приховування інформації; векторна графіка; стеганографія; афінні перетворення.

Табл. 1. Іл. 15. Бібліогр.: 14 назв.

УДК 004.056.5

Стеганографические методы в векторной графике / *А.А. Кузнецов, А.В. Кононченко* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 32 – 41.

Для сокрытия информации применяются различные стеганографические техники. Обычно информацию скрывают в изображениях, аудио и видео файлах, текстовых документах и тому подобное. В статье рассмотрены векторные изображения, состоящие из различных математических объектов (точки, линии, кривые первого и второго порядка, кривые Безье, узлы, касательные, базовые точки и т.д.). Техники сокрытия информации меняют эти математические объекты, например, через кодирование координат базовых точек. Наиболее удачным для проведения стеганографических преобразований является формат векторной графики SVG, который благодаря своей структуре позволяет легко манипулировать объектами, из которых состоит. Его широкая поддержка различными платформами также позволяет повысить уровень скрытности при передаче секретных данных путем пересылки обычных на первый взгляд файлов медиа. В статье рассмотрены два метода (побитовый и метод паттернов) сокрытия информации в векторные изображения, изучены их особенности, преимущества и недостатки. Также были исследованы различные аффинные преобразования, которые можно применять для нарушения работы стеганосистемы. Наиболее распространенными видами аффинных преобразований являются операции переноса, поворота, сдвига и масштабирования с возможными вариациями (смещения по осям абсцисс и ординат, масштабирование пропорциональное и непропорциональное, со сжатием и с расширением). Большинство методов встраивания информации в векторные изображения обеспечивают единовременную устойчивость к аффинным преобразованиям, при этом при повторном наложении операций изменения положения объектов, сообщение может разрушиться вообще. Исследованные в работе методы реализуют больший уровень устойчивости к различного рода преобразованиям при их многократном проведении, проведенные эксперименты это наглядно демонстрируют. Полученные результаты показывают, что векторные изображения действительно могут применяться для сокрытия информации, но устойчивость против определенных аффинных атак не всегда высока.

Ключевые слова: сокрытие информации; векторная графика; стеганография; аффинные преобразования.

Табл. 1. Ил. 15. Библиогр.: 14 назв.

UDC 004.056.5

Steganographic methods in vector graphics / *A.A. Kuznetsov, Г.В. Кононченко* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 32 – 41.

Various steganographic techniques are used to hide information. Usually, information is hidden in images, audio and video files, text documents, and the like. The article deals with vector images consisting of various mathematical objects (points, lines, curves of the first and second order, Bezier curves, nodes, tangents, base points, etc.). Information hiding techniques alter these mathematical objects, for example, by encoding the coordinates of the base points. The most successful for carrying out steganographic transformations is the SVG vector graphics format, which, due to its structure, makes it easy to manipulate the objects of which it consists. Its broad support across platforms also allows for increased secrecy when transferring sensitive data by sending seemingly ordinary media files. The article discusses two methods (bitwise and the method of patterns) of hiding information in vector images, studied their features, advantages and disadvantages. Various affine transformations that can be used to disrupt the operation of the steganosystem were also investigated. The most common types of affine transformations are the operations of transfer, rotation, shift and scaling with possible variations (offsets along the abscissa and ordinate axes, proportional and non-proportional scaling, with compression and expansion). Most of the methods for embedding information into vector images provide a one-time resistance to affine transformations, while the repeated imposition of operations for changing the position of objects may destroy the message altogether. The methods investigated in the work (bitwise and the method of patterns) implement a higher level of resistance to various kinds of transformations when they are repeated many times, and the

conducted experiments clearly demonstrate this. The results obtained show that vector images can indeed be used to hide information, but the resistance against certain affine attacks is not always high.

Key words: information concealing; vector graphics; steganography; affine transformations

1 tab. 15 fig. Ref: 14 items.

УДК 004.056.55

Аналіз апаратних реалізацій алгоритмів електронного підпису qTesla, Crystals-Dilitium і MQDSS на різних рівнях безпеки / *М.В. Єсіна, Б.С. Шахов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 42 – 52.

Відомо, що існують алгоритми криптографії з відкритим ключем, що засновані на RSA та еліптичних кривих, надають гарантії безпеки, які супроводжуються складністю. Можна казати про неможливість вирішення завдань цілочисельної факторизації і дискретного логарифма. Але експерти прогнозують, що створення квантового комп'ютера зможе зламати класичні криптографічні алгоритми. Через цю майбутню проблему національний інститут стандартів і технологій (NIST) разом із провідними вченими у галузі криптографії розпочав відкритий процес стандартизації алгоритмів з відкритим ключем для квантових атак. Важливою особливістю постквантового періоду у криптографії є суттєва невизначеність щодо вихідних даних для криптоаналізу та протидії в частині можливостей квантових комп'ютерів, їх математичного та програмного забезпечень, а також застосування квантового криптоаналізу до існуючих криптоперетворень та криптопротоколів. В якості основних методів NIST США обрано математичні методи електронного підпису (ЕП), що пройшли суттєвий аналіз та обґрунтування в процесі широких досліджень криптографами та математиками на найвищому рівні. Вони детально описані та пройшли дослідження на першому та другому етапах міжнародного конкурсу NIST США PQC. З історичної точки зору, у 1997 р. NIST запросив рекомендації у громадськості для визначення заміни стандарту шифрування даних (DES), Advanced Encryption Standard (AES). Відтоді відкриті криптографічні оцінки стали способом вибору криптографічних стандартів. Наприклад, NESSIE (2000-2002), eSTREAM (2004-2008), CRYPTREC (2000-2002), SHA-3 (2007-2012) і CAESAR (2013-2019) прийняли цей підхід. У цих оцінках головним параметром була безпека. Продуктивність у програмному забезпеченні, продуктивність у прикладних специфічних інтегральних схемах (ASIC), продуктивність у FPGA та можливість реалізації з використанням обмежених ресурсів (невеликих мікропроцесорів та малопотужних апаратних засобів) є вторинними критеріями. У роботі описується порівняння апаратного забезпечення трьох алгоритмів підпису (qTesla, Crystals-Dilitium, MQDSS), які зокрема є кандидатами 2-го раунду конкурсу NIST PQC, а алгоритм Crystals-Dilitium – фіналістом цього конкурсу. Метою цієї роботи є аналіз та порівняння трьох апаратних реалізацій кандидатів другого раунду конкурсу NIST PQC на алгоритм електронного підпису.

Ключові слова: постквантова криптографія; електронний підпис; qTesla; Crystals-Dilitium; MQDSS.

Табл. 12. Іл. 10. Бібліогр.: 8 назв.

УДК 004.056.55

Анализ аппаратных реализаций алгоритмов электронной подписи qTesla, Crystals-Dilitium и MQDSS на различных уровнях безопасности / *М.В. Есина, Б.С. Шахов* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 42 – 52.

Известно, что существуют алгоритмы криптографии с открытым ключом, основанные на RSA и эллиптических кривых предоставляют гарантии безопасности сопровождающихся сложностью. Исходя из этого, можно говорить о невозможности решения задач целочисленной факторизации и дискретного логарифма.) Но эксперты прогнозируют, что создание квантового компьютера сможет сломать классические криптографические алгоритмы. Из-за этой будущей проблемы национальный институт стандартов и технологий (NIST) вместе с ведущими учеными в области криптографии начал открытый процесс стандартизации алгоритмов с открытым ключом для квантовых атак. Важной особенностью постквантового периода в криптографии является существенная неопределенность относительно исходных данных для криптоанализа и противодействия в части возможностей квантовых компьютеров, их математического и программного обеспечения, а также применение квантового криптоанализа к существующим криптопреобразованиям и криптопротоколам. В качестве основных методов NIST США избрал математические методы электронной подписи (ЭП), прошедшие существенный анализ и обоснование в процессе широких исследований криптографами и математиками на высшем уровне. Они подробно описаны и прошли исследования на первом и втором этапах международного конкурса NIST США PQC. С исторической точки зрения, в 1997 г. NIST запросил рекомендации общественности для определения замены стандарта шифрования данных (DES), Advanced Encryption Standard (AES). С тех пор открытые криптографические оценки стали способом выбора криптографических стандартов. Например, NESSIE (2000-2002), eSTREAM (2004-2008), CRYPTREC (2000-2002), SHA-3 (2007-2012) и CAESAR (2013-2019) приняли этот подход. В этих оценках главным параметром была безопасность. Производительность в программном обеспечении, производительность в приложениях специфических интегральных схем (ASIC), производительность в FPGA и возможность реализации с использованием ограниченных ресурсов (небольших микропроцессоров и маломощных аппаратных средств) являются вторичными критериям. В работе описывается сравнение аппаратного обеспечения трех алгоритмов подписи (qTesla, Crystals-Dilitium, MQDSS), которые, в частности, являются кандидатами 2-го раунда конкурса NIST PQC, а алгоритм Crystals-Dilitium – финалистом этого конкурса. Цель работы – анализ и

сравнение трех аппаратных реализаций кандидатов второго раунда конкурса NIST PQC на алгоритм электронной подписи.

Ключевые слова: постквантовая криптография; электронная подпись; qTesla; Crystals-Dilitium; MQDSS.

Табл. 12. Ил. 10. Библиогр.: 8 назв.

UDC 004.056.55

Analysis of hardware implementations of electronic signature algorithms qTesla, Crystals-Dilitium and MQDSS at different levels of security / *M.V. Yesina, B.S. Shahov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 42 – 52.

It is known, that existing public-key cryptography algorithms based on RSA and elliptic curves provide security guarantees accompanied by complexity. Based on this one can talk about the impossibility to solve problems of integer factorization and discrete logarithm. However, experts predict that the creation of a quantum computer will be able to crack classical cryptographic algorithms. Due to this future problem, the National Institute of Standards and Technologies (NIST), together with leading scientists in the field of cryptography, began an open process of standardizing public-key algorithms for quantum attacks. An important feature of the post-quantum period in cryptography is the significant uncertainty regarding the source data for cryptanalysis and counteraction in terms of the capabilities of quantum computers, their mathematical and software, as well as the application of quantum cryptanalysis to existing cryptotransformations and cryptoprotocols. Mathematical methods of electronic signature (ES) have been chosen as the main methods of NIST USA, which have undergone significant analysis and substantiation in the process of extensive research by cryptographers and mathematicians at the highest level. These methods are described in detail and passed the research at the first stage of the international competition NIST USA PQC. Historically, in 1997, NIST sought public advice to determine the replacement of the data encryption standard (DES), Advanced Encryption Standard (AES). Since then, open cryptographic estimations have become a way of choosing cryptographic standards. For example, NESSIE (2000-2002), eSTREAM (2004-2008), CRYPTREC (2000-2002), SHA-3 (2007-2012) and CAESAR (2013-2019) have adopted this approach. Security was the main parameter in these estimations. Performance in software, performance in application-specific integrated circuits (ASICs), performance in FPGAs, and feasibility with limited resources (small microprocessors and low-power hardware) are secondary criteria. This paper presents the comparison of the hardware of three signature algorithms (qTesla, Crystals-Dilitium, MQDSS), which, in particular, are the candidates for the 2nd round of the NIST PQC competition, and the Crystals-Dilitium algorithm is the finalist of this competition. The objective of this work is to analyze and compare three hardware implementations of candidates for the second round of the NIST PQC contest for an electronic signature algorithm.

Key words: post-quantum cryptography; electronic signature; qTesla; Crystals-Dilitium; MQDSS.

12 tab. 10 fig. Ref: 8 items.

УДК 004.056

Аналіз формальних моделей управління доступом і особливості їх застосування для баз даних / *В.В. Віліґура* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 53 – 70.

Невід'ємною частиною будь-якого проекту по створенню або оцінки безпеки інформаційних систем і баз даних є наявність моделі безпеки. В роботі розглядаються основні положення найбільш поширених моделей безпеки, заснованих на контролі доступу суб'єктів до об'єктів. Проведений аналіз формальних моделей управління доступом виявив, що кожна з них, маючи певні переваги і недоліки, має право на використання. Вирішальним фактором у прийнятті рішення є оцінка конкретної ситуації, яка дозволить зробити правильний вибір. Так, в роботі відзначається, що моделі безпеки на основі дискреційної політики доцільно застосовувати при проведенні формальної верифікації коректності побудови систем розмежування доступу в добре захищених інформаційних системах і базах даних. При цьому однак підкреслюється, що цим моделям властиві певні недоліки, що обмежують їх застосування. В роботі констатується, що незважаючи на те, що моделі безпеки на основі мандатної політики доступу відіграють важливу роль в теорії інформаційної безпеки і їх положення введені в якості обов'язкових вимог до систем, що обробляють секретну інформацію, а також в стандартах захищених систем, при практичній реалізації цих моделей може виникнути ряд проблем: завищення рівня безпеки, запис наосліп, проблема привілейованих суб'єктів, що виконують операції, які не вписуються в рамки моделі. Також робиться висновок про те, що використання моделей безпеки на основі рольової політики дозволяє реалізувати правила розмежування доступу, що динамічно змінюються в процесі функціонування інформаційних систем, баз даних, ефективність яких особливо помітно проявляється при організації доступу до ресурсів систем з великою кількістю користувачів і об'єктів.

Ключові слова: модель безпеки; управління доступом; інформаційна система; база даних.

Табл. 3. Іл. 3. Бібліогр.: 31 назв.

УДК 004.056

Анализ формальных моделей управления доступом и особенности их применимости для баз данных / *В.В. Вилигура* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 53 – 70.

Неотъемлемой частью любого проекта по созданию или оценке безопасности информационных систем и баз данных является наличие модели безопасности. В работе рассматриваются основные положения наиболее распространенных моделей безопасности, основанных на контроле доступа субъектов к объектам. Проведенный анализ формальных моделей управления доступом выявил, что каждая из них, имея определенные пре-

имущества и недостатки, имеет право на использование. Решающим фактором в принятии решения является оценка конкретной ситуации, которая позволит сделать правильный выбор. Так, в работе отмечается, что модели безопасности на основе дискреционной политики целесообразно применять при проведении формальной верификации корректности построения систем разграничения доступа в хорошо защищенных информационных системах и базах данных. При этом, однако, подчеркивается, что этим моделям свойственны определенные недостатки, ограничивающие их применение. В работе констатируется, несмотря на то, что модели безопасности на основе мандатной политики доступа играют значимую роль в теории информационной безопасности и их положения введены в качестве обязательных требований к системам, обрабатывающим секретную информацию, а также в стандартах защищенных систем, при практической реализации этих моделей может возникнуть ряд проблем: завышение уровня безопасности, запись вслепую, проблема привилегированных субъектов, выполняющих операции не вписывающиеся в рамки модели. Делается вывод о том, что использование моделей безопасности на основе ролевой политики позволяет реализовать динамически изменяющиеся в процессе функционирования информационных систем, баз данных правила разграничения доступа, эффективность которых особенно заметно проявляется при организации доступа к ресурсам систем с большим количеством пользователей и объектов.

Ключевые слова: модель безопасности; управление доступом; информационная система; база данных.

Табл. 3. Ил. 3. Библиогр.: 31 назв.

UDC 004.056

Analysis of formal models for access control and specific features of their applicability to databases / *V.V. Vilihura* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 53 – 70.

An integral part of any project to create or assess the security of information systems and databases is the presence of a security model. The paper considers the main positions of the most common security models based on controlling the access of subjects to objects. The analysis of formal models for access control has revealed that each of them, having certain advantages and disadvantages, has the right to be used. The decisive factor in making a decision is an assessment of a specific situation, which will allow one to make the right choice. In this regard, the paper notes that security models based on discretionary policies are advisable to be applied when conducting formal verification of the correctness of building access control systems in well-protected information systems and databases. However, it is emphasized that these models have certain drawbacks that limit their use. The paper states that despite the fact that security models based on the mandatory access policy play a significant role in information security theory and their provisions have been introduced as mandatory requirements for systems that process secret information, as well as in the standards of secure systems, a number of problems may arise in the practical implementation of these models. Among these problems there are the problems associated with overestimating the security level, blind recordings, performing operations that do not fit into the framework of the model by privileged subjects. The paper also concludes that the use of security models based on role-based policy allows one to implement access control rules dynamically changing during the operation of information systems and databases, the effectiveness of which is especially noticeable when organizing access to the resources of systems with a large number of users and objects.

Key words: security model; access control; information system; database.

3 tab. 3 fig. Ref: 31 items.

УДК 004.056.55

Процеси та методи вибору загальносистемних параметрів та аналіз стійкості проти атак сторонніми каналами для алгоритму направленого шифрування та інкапсуляції ключів стандарту ДСТУ 8961:2019 / *В.А Кулібаба* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 71 – 78.

В останні роки відбувся значний прогрес у створенні квантових комп'ютерів. Якщо масштабовані квантові комп'ютери будуть впроваджені найближчим часом, це поставить під загрозу безпеку найбільш широко використовуваних криптосистем з відкритим ключем. Найбільш вразливими є ключові схеми, засновані на факторизації, дискретних логарифмах та криптографії еліптичної кривої. Зараз головним завданням є розробка, оцінка, дослідження та стандартизація асиметричних крипто перетворень на міжнародному рівні, включаючи механізми інкапсуляції ключів та направленого шифрування, стійкі до атак порушників постквантового періоду. Важливою особливістю перехідного та постквантового періоду є застосування нових математичних методів для протидії квантовому криптоаналізу. У роботі розглядаються основні атаки на механізми інкапсуляції ключів та направленого шифрування, а також загальносистемні параметри стандарту ДСТУ 8961:2019, що впливають на стійкість від атак та складність перетворень. Розглядаються методи генерування загальносистемних параметрів 5 та 7 рівнів стійкості – 512 біт класичної та 256 біт квантової безпеки, а також захищеність алгоритму від атак сторонніми каналами. Проаналізовано залежність часу шифрування та розшифрування від рівня стійкості. Наведено результати обчислень загальносистемних параметрів для рівнів стійкості 256/128, 384/192 та 512/256, а також надано рекомендації щодо вибору загальносистемних параметрів в залежності від оточення та обчислювальних можливостей. Наведено обрані та рекомендовані до застосування у стандарті ДСТУ 8961:2019 набори параметрів. Зроблено висновки про можливість застосування стандарту ДСТУ 8961 в постквантовий період.

Ключові слова: загальносистемні параметри; протоколи інкапсуляції ключів; направлене шифрування; алгебраїчні решітки; криптографічна стійкість.

Табл. 5. Іл. 3. Бібліогр.: 14 назв.

УДК 004.056.55

Процессы и методы выбора общесистемных параметров и анализ устойчивости против атак сторонними каналами для алгоритма направленного шифрования и инкапсуляции ключей стандарта ДСТУ 8961:2019 / *В.А Кулибаба* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 71 – 78.

В последние годы произошел значительный прогресс в создании квантовых компьютеров. Если масштабируемые квантовые компьютеры будут внедрены в ближайшее время, это поставит под угрозу безопасность наиболее широко используемых криптосистем с открытым ключом. Наиболее уязвимыми являются ключевые схемы, основанные на факторизации целых чисел, дискретных логарифмах и криптографии на эллиптической кривой. Сейчас главной задачей является разработка, оценка, исследование и стандартизация асимметричных криптопреобразований на международном уровне, включая механизмы инкапсуляции ключей и направленного шифрования, устойчивые к атакам нарушителей постквантового периода. Важной особенностью переходного и постквантового периода является применение новых математических методов для противодействия квантовому криптоанализу. В работе рассматриваются основные атаки на механизмы инкапсуляции ключей и направленного шифрования, а также общесистемные параметры стандарта ДСТУ 8961:2019, влияющие на устойчивость от атак и сложность преобразований. Рассматриваются методы генерирования общесистемных параметров 5 и 7 уровней устойчивости – 512 бит классической и 256 бит квантовой безопасности, а также защищенность алгоритма от атак сторонними каналами. Проанализирована зависимость времени шифрования и расшифровки от уровня устойчивости. Приведены результаты вычислений общесистемных параметров для уровней устойчивости 256/128, 384/192 и 512/256, а также даны рекомендации по выбору общесистемных параметров в зависимости от окружения и вычислительных возможностей. Приведены рекомендованные к применению в стандарте ДСТУ 8961: 2019 наборы параметров. Сделаны выводы о возможности применения стандарта ДСТУ 8961 в постквантовый период.

Ключевые слова: общесистемные параметры; протоколы инкапсуляции ключей; направленное шифрование; алгебраические решетки; криптографическая стойкость.

Табл. 5. Ил. 3. Библиогр.: 14 назв.

UDC 004.056.55

Processes and methods for selecting system-wide parameters and analysis of resistance against third-party channel attacks for the key encapsulation mechanism DSTU 8961:2019 / *V.A. Kulibaba* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 71 – 78.

In recent years, there has been significant progress in the creation of quantum computers. If scalable quantum computers are implemented in the near future, this will jeopardize the security of the most widely used public key cryptosystems. The most vulnerable are public-key schemes based on factorization, discrete logarithms and elliptic curve cryptography. Currently, the main task is to develop, evaluate, study and standardize asymmetric crypto transformations at the international level, including mechanisms of key encapsulation and directional encryption, resistant to attacks by violators of the post-quantum period. An important feature of the transition and post-quantum period is the usage of new mathematical methods to opposite quantum crypto analysis. The paper considers the main attacks on the mechanisms of key encapsulation and directional encryption, as well as system-wide parameters of the DSTU 8961: 2019 standard, which affect the resistance to attacks and the complexity of transformations. Methods for generating system-wide parameters of 5 and 7 levels of stability – 512 bits of classical and 256 bits of quantum security, as well as the protection of the algorithm from attacks by third-party channels are considered. The dependence of encryption and decryption time on the level of stability is analyzed. The results of calculations of system-wide parameters for stability levels 256/128, 384/192 and 512/256 are presented, as well as recommendations for the selection of system-wide parameters depending on the environment and computing capabilities. Sets of parameters selected and recommended for use in the DSTU 8961: 2019 standard are given. Conclusions are drawn about the possibility of applying the DSTU 8961 standard in the post-quantum period.

Key words: system parameters; key encapsulation mechanisms; direct encryption; algebraic lattices; cryptographic stability.

5 tab. 3 fig. Ref: 14 items.

УДК 003.026:004.056

Властивості багатовимірного алгоритму Rainbow та його здатність протистояти різноманітним методам криптоаналізу і атаці сторонніми каналами / *Д.В. Гармаш* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 79 – 84.

Розглядається аналіз сутності та можливості захисту постквантового криптографічного алгоритму Rainbow. Визначаються основні властивості алгоритму Rainbow та загальна суть криптографічних алгоритмів шифрування та електронного підпису на основі мультиваріативних квадратичних перетворень. Наводяться основні положення стосовно протоколів. Наводяться аналізи стосовно здатності захисту алгоритму від різноманітних атак. Досліджується вразливість алгоритму до атаки сторонніми каналами. Розглядаються загальні положення алгоритму. Алгоритм зображується та розглядається з математичної точки зору, також викладається математична суть криптографічних алгоритмів шифрування та електронного підпису на основі мультиваріативних квадратичних перетворень. Вивчається застосування різноманітних методів криптоаналізу против криптографічного алгоритму на основі мультиваріативних квадратичних перетворень Rainbow. Аналізується метод знижен-

ня рангу проти алгоритму Rainbow. Вивчається метод криптоаналізу за допомогою атаки на Oil-Vinegar схему та метод криптоаналізу "метод мінранку". Досліджується атака за допомогою структури багатошаровості.

Ключові слова: Rainbow; криптоаналіз; вразливість; мінранк; схема; алгоритм.

Бібліогр.: 9 назв.

УДК 003.026:004.056

Свойства мультивариативного алгоритма Rainbow и его способность противостоять различным методам криптоанализа и атаке посторонними каналами / *Д.В. Гармаш* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 79 – 84.

Рассматривается анализ сущности и возможности защиты постквантового криптографического алгоритма Rainbow. Определяются основные свойства алгоритма Rainbow и общая суть криптографических алгоритмов шифрования и электронной подписи на основе мультивариативних квадратичных преобразований. Приводятся основные положения относительно протоколов. Приводятся анализы относительно способности защиты алгоритма от различных атак. Исследуется уязвимость алгоритма к атаке сторонними каналами. Рассматриваются общие положения алгоритма. Алгоритм изображается и рассматривается с математической точки зрения, излагается математическая суть криптографических алгоритмов шифрования и электронной подписи на основе мультивариативних квадратичных преобразований. Изучается применение различных методов криптоанализа против криптографического алгоритма на основе мультивариативних квадратичных преобразований Rainbow. Анализируется метод снижения ранга против алгоритма Rainbow. Изучается метод криптоанализа с помощью атаки на Oil-Vinegar схему и метод криптоанализа "метод минранку". Исследуется атака с помощью структуры многослойности.

Ключевые слова: Rainbow; криптоанализ; уязвимость; минранк; схема; алгоритм.

Библиогр.: 9 назв.

UDC 003.026:004.056

Properties of the Rainbow multi-variant algorithm and its ability to resist various crypto-analysis methods and attack by outside channels / *D.V. Harmash* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 79 – 84.

This work presents the analysis of the essence and possibilities of protection of the Rainbow post-quantum cryptographic algorithm. The main properties of the Rainbow algorithm and the general essence of cryptographic encryption and electronic signature algorithms based on multivariate quadratic transformations are determined. The main provisions regarding the protocols are given. Analyses are given regarding the ability to protect the algorithm against various attacks. The vulnerability of the algorithm to attack by third-party channels is investigated. The general provisions of the algorithm are considered. The algorithm is presented and considered from a mathematical point of view, as well as the mathematical essence of cryptographic algorithms for encryption and electronic signature based on multivariate quadratic transformations. The application of various methods of cryptanalysis against cryptographic algorithm based on multivariate quadratic Rainbow transformations is studied. The method of decreasing rank against the Rainbow algorithm is analyzed. The method of cryptanalysis by attacking the Oil-Vinegar scheme and the method of cryptanalysis "minranku method" are investigated. The attack is studied using a multilayer structure.

Key words: Rainbow; cryptanalysis; vulnerability; minrank; scheme; algorithm.

Ref: 9 items.

УДК 003.026:004.056

Аналіз захищеності постквантового алгоритму електронного підпису Rainbow від потенційних атак / *Г.А. Малєєва* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 85 – 93.

Багатовимірна криптографія на основі відкритого ключа є кандидатом для постквантової криптографії, і це дозволяє генерувати особливо короткі підписи та швидку перевірку. Схема підписів Rainbow, запропонована Дж. Діном та Д. Шмідтом, є такою багатовимірною криптосистемою і вважається захищеною від усіх відомих атак. Необхідність проведення досліджень ЦП Rainbow обґрунтовується тим, що назріла необхідність розроблення та прийняття постквантового національного стандарту ЦП, а також тим, що в процесі проведення конкурсу NIST США стосовно математичних основ застосування методу криптографічного перетворення Rainbow отримано перспективні результати. Тому вважається важливим їх врахування та використання в Україні. Схема підпису Rainbow може бути реалізована просто та ефективно за допомогою лінійних методів алгебри над невеликим кінцевим полем і, зокрема, створює коротші підписи, ніж ті, що використовуються в RSA та інших постквантових підписах. У 2-му раунді NIST PQC пропонуються захищені набори параметрів Rainbow і проаналізовано кілька атак на них. При порівнянні ЕП перевага віддається алгоритмам ЕП, що пройшли відбір за безумовними критеріями, а також, що мають кращі показники щодо інтегральних умовних критеріїв, оскільки така методика є більш раціональною. Зокрема, атака Rainbow-Band-Separation (RBS) є найкращою серед відомих атак на Rainbow з певним набором параметрів і є важливою. Атака Rainbow-Band-Separation відновлює секретний ключ Rainbow, розв'язуючи певні системи квадратичних рівнянь, а його складність оцінюється за відомим показником, який називається ступенем регулярності. Однак, як правило, ступінь регулярності більша, ніж ступінь розв'язання в експериментах, і точної оцінки отримати неможливо. У роботі запропоновано новий показник складності атаки Rainbow-Band-Separation за допомогою алгоритму F_4, який дає більш точну оцінку порівняно з показником, що використовує ступінь регулярності.

Мета роботи – порівняльний аналіз ЕП на основі MQ-перетворень за критерієм стійкість-складність та спроба зрозуміти безпеку Rainbow від атаки RBS за допомогою F_4.

Ключові слова: : багатовимірна криптографія; аналіз атак; постквантовий період.

Табл. 4. Іл. 1. Бібліогр.: 6 назв.

УДК 003.026:004.056

Анализ защищенности постквантового алгоритма электронной подписи Rainbow от потенциальных атак / *А.А. Малеева* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 85 – 93.

Многомерная криптография на основе открытого ключа является кандидатом для постквантовой криптографии, и это позволяет генерировать особенно короткие подписи и быструю проверку. Схема подписей Rainbow, предложенная Дж. Дином и Д. Шмидтом, является такой многомерной криптосистемой и считается защищенной от всех известных атак. Необходимость проведения исследований ЦП Rainbow обосновывается тем, что назрела необходимость разработки и принятия постквантового национального стандарта ЦП, а также тем, что в процессе проведения конкурса NIST США математических основ применения метода криптографического преобразования Rainbow получены перспективные результаты. Поэтому считается важным их учет и использование в Украине. Схема подписи Rainbow может быть реализована просто и эффективно с помощью линейных алгебраических методов над небольшим конечным полем и, в частности, создает короткие подписи, в отличие от тех, что используются в RSA и других постквантовых подписях. Во 2-м раунде NIST PQC предлагаются защищенные наборы параметров Rainbow и проанализированы несколько атак на них. При сравнении ЭП предпочтение отдается методам ЭП, которые прошли отбор по безусловными критериям, а также имеют лучшие показатели по интегральным условным критериям, поскольку такая методика является более рациональной. В частности, атака Rainbow-Band-Separation (RBS) является лучшей среди известных атак на Rainbow с определенным набором параметров и важна. Атака Rainbow-Band-Separation восстанавливает секретный ключ Rainbow, решая определенные системы квадратичных уравнений, а его сложность оценивается по известному показателю, который называется степенью регулярности. Однако, как правило, степень регулярности больше, чем степень решения в экспериментах, и точной оценки получить невозможно. В работе предложен новый показатель сложности атаки Rainbow-Band-Separation с помощью алгоритма F_4, который дает более точную оценку по сравнению с показателем, использует степень регулярности.

Цель работы – сравнительный анализ ЭП на основе MQ-преобразований по критерию устойчивость-сложность и попытка понять безопасность Rainbow от атаки RBS с помощью F_4.

Ключевые слова: многомерная криптография; анализ атак; постквантовий период.

Табл. 4. Ил. 1. Библиогр.: 6 назв.

UDC 003.026:004.056

Analysis of security of post-quantum algorithm of Rainbow electronic signature against potential attacks / *G.A. Maleeva* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 85 – 93.

Multidimensional public key cryptography is a candidate for post-quantum cryptography, and it makes it possible to generate particularly short signatures and quick verification. The Rainbow signature scheme proposed by J. Dean and D. Schmidt is such a multidimensional cryptosystem and it is considered to be protected against all known attacks. The need for research on Rainbow ES is justified by the fact that there is a need to develop and adopt a post-quantum national securities standard, and that in the process of the US NIST competition on the mathematical basis of cryptographic transformation method Rainbow, promising results. Therefore, it is considered important to take them into account and use them in Ukraine. The Rainbow signature scheme can be implemented simply and efficiently using linear algebra methods over a small finite field and, in particular, creates shorter signatures than those used in RSA and other post-quantum signatures [1]. In the 2nd round of NIST PQC, protected sets of Rainbow parameters are offered and several attacks on them are analyzed [1]. When comparing ES, preference is given to ES algorithms that have been selected according to unconditional criteria, as well as those that have better indicators for integral conditional criteria, because such a technique is more rational. In particular, the Rainbow-Band-Separation (RBS) attack [2] is the best known Rainbow attack with a certain set of parameters and is important. The Rainbow-Band-Separation attack restores the Rainbow secret key by solving certain systems of quadratic equations, and its complexity is measured by a well-known measure called the degree of regularity. However, as a rule, the degree of regularity is greater than the degree of solution in experiments, and it is impossible to obtain an accurate estimate. The paper proposes a new indicator of the complexity of the Rainbow-Band-Separation attack using F_4 algorithm, which gives a more accurate estimate compared to the indicator that uses the degree of regularity.

The aim of the work is a comparative analysis of ES based on MQ-transformations on the criterion of stability-complexity and an attempt to understand the security of Rainbow against RBS attack using F_4.

Key words: multidimensional cryptography; attack analysis; postquantum period.

4 tab. 1 fig. Ref: 6 items.

УДК 621.391:519.2

Методи побудови та властивості логарифмічних підписів / *Є.В. Котух, О.В. Северинов, А.В. Власов, Л.С. Козіна, А.О. Теницька, Е. О. Зарудна* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 94 –99.

Розвиток та перспективні напрями досліджень у побудові практичних моделей квантових комп'ютерів сприяє пошуку та розробці ефективних криптографічних примітивів. Разом зі зростанням практичних можливостей використання квантових обчислень зростає загроза класичним схемам шифрування та електронного підпису, які використовують як основу класичні математичні проблеми, що долаються обчислювальними можливостями квантових комп'ютерів. Цей факт мотивує дослідження фундаментальних теорем, що стосуються математичних та обчислювальних аспектів постквантових криптосистем-кандидатів. Однією з актуальних проблем є розробка нової квантово стійкої асиметричної криптосистеми. Перспективним напрямом розробки асиметричних криптосистем є використання логарифмічних підписів і покриттів кінцевих груп. Актуальний стан цього напряму й праці останніх років дають підстави припускати, що завдання факторизації елемента кінцевої групи в теорії побудови криптосистем на основі неабелевих груп з використанням логарифмічних підписів є обчислювально складні, що потенційно забезпечує необхідний рівень криптографічного захисту перед атаками, що використовують можливості квантових обчислень. У роботі представлено логарифмічні підписи як особливий тип факторизації в кінцевих групах, розглянуто їх властивості та методи побудови.

Ключові слова: постквантова криптографія, логарифмічні підписи, покриття, неабелеві групи.

Бібліогр.: 13 назв.

УДК 621.391:519.2

Методы построения и свойства логарифмических подписей / *Е.В. Котух, А.В. Северинов, А.В. Власов, Л.С. Козина, А.А. Теницкая, Е. А. Зарудная* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 204. С. С. 94 –99.

Развитие и перспективные направления исследований в построении практических моделей квантовых компьютеров способствуют поиску и разработке эффективных криптографических примитивов. Вместе с ростом практических возможностей использования квантовых вычислений растет угроза классическим схемам шифрования и электронной подписи, которые используют как основу классические математические проблемы, преодолеваются вычислительными возможностями квантовых компьютеров. Этот факт мотивирует исследования фундаментальных теорем, касающихся математических и вычислительных аспектов постквантовых криптосистем-кандидатов. Одной из актуальных проблем является разработка новой квантово-устойчивой асимметричной криптосистемы. Перспективным направлением разработки ассиметричных криптосистем является использование логарифмических подписей и покрытий конечных групп. Актуальное состояние этого направления и работы последних лет дают основания предполагать, что задача факторизации элемента конечной группы в теории построения криптосистем на основе неабелевых групп с использованием логарифмических подписей является вычислительно сложной, потенциально обеспечивает необходимый уровень криптографической защиты перед атаками, использующими возможности квантовых вычислений. Представлены логарифмические подписи как особый тип факторизации в конечных группах, рассмотрены их свойства и методы построения.

Ключевые слова: постквантовая криптография, логарифмические подписи, накрытия, неабелевые группы.

Библиогр.: 13 назв.

UDC 621.391:519.2

Methods of construction and properties of logariphmic signatures / *E.V. Kotukh, O.V. Severinov, A.V. Vlasov, L.S. Kozina, A.O. Tenytska, E.O. Zarudna* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №204. P. C. 94 –99.

Development and promising areas of research in the construction of practical models of quantum computers contributes to the search and development of effective cryptographic primitives. Along with the growth of the practical possibilities of using quantum computing, the threat to classical encryption and electronic signature schemes using classical mathematical problems as a basis, being overcome by the computational capabilities of quantum computers. This fact motivates the study of fundamental theorems concerning the mathematical and computational aspects of candidate post-quantum cryptosystems. Development of a new quantum-resistant asymmetric cryptosystem is one of the urgent problems. The use of logarithmic signatures and coverings of finite groups a promising direction in the development of asymmetric cryptosystems. The current state of this area and the work of recent years suggest that the problem of factorizing an element of a finite group in the theory of constructing cryptosystems based on non-Abelian groups using logarithmic signatures is computationally complex; it potentially provides the necessary level of cryptographic protection against attacks using the capabilities of quantum calculations. The paper presents logarithmic signatures as a special type of factorization in finite groups; it also considers their properties and construction methods.

Key words: post-quantum cryptography, logarithmic signatures, coverings, non-abelian groups.

Ref: 13 items.

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ
ФИЗИКА ПРИБОРОВ, ЭЛЕМЕНТОВ И СИСТЕМ
PHYSICS OF INSTRUMENTS, ELEMENTS AND SYSTEMS

УДК 621.373.826

Принципи побудови гіроскопа на базі фотонно-кристалічних волокон з фотонною забороненою зоною / *Аль-Судані Хайдер Алі* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 100 – 107.

Гіроскоп – пристрій, який дозволяє вимірювати зміну кутів орієнтації зв'язаного з ним тіла обертання відносно інерціальної системи координат. Гіроскопи з фотонно-кристалічними волокнами є свого роду оптичні

гіроскопи, які дають безліч нових і поліпшених характеристик, крім тих, які можуть запропонувати звичайні волоконно-оптичні гіроскопи. У будь-якому випадку властивості оптичного волокна можуть зіграти велику роль у визначенні характеристик гіроскопа. Принцип дії більшості оптичних гіроскопів оснований на ефекті Саньяка (Sagnac) або інтерферометрі Саньяка, суть якого полягає в наступному. Якщо в замкнутому оптичному контурі (інтерферометрі) в протилежних напрямках поширюються дві світлові хвилі, то у випадку нерухомого контура фазові набіги обох хвиль, що пройшли увесь контур в протилежних напрямках, будуть однаковими. При обертанні контура навколо осі, нормальної до площини контуру, фазові набіги хвиль стають неоднаковими, а їх різниця (фазовий зсув) в загальному випадку буде пропорційний кутовій швидкості обертання контура, площі, яку охоплює контур, і частоті електромагнітної хвилі (ЕМХ). Оскільки під час роботи гіроскопа площа і частота ЕМХ залишаються незмінними, фазовий зсув буде пропорційний тільки кутовій швидкості. Використання фотонно-кристалічного волокна для підвищення чутливості є дуже перспективним, воно значно зменшує дрейф через термічні поляризаційні нестійкості і ефект Керра. У статті запропоновано використовувати в порожнистій серцевині оптичного гіроскопу фотонно-кристалічне волокно 1550nmλ, Ø10 мкм замість звичайних волокон.

Ключові слова: оптичний гіроскоп; ефект Саньяка; фотонно-кристалічне волокно.

Іл. 6. Бібліогр.: 11 назв.

УДК 621.373.826

Принципы построения гироскопов на базе фотонно-кристаллических волокон с фотонной запрещенной зоной / *Аль-Судани Хайдер Али* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 100 – 107.

Гироскоп – устройство, которое позволяет измерять изменение углов ориентации связанного с ним тела вращения относительно инерциальной системы координат. Гироскопы с фотонно-кристаллическими волокнами представляют собой своего рода оптические гироскопы, которые дают множество новых и улучшенных характеристик, кроме тех, которые могут предложить обычные волоконно-оптические гироскопы.

В любом случае свойства оптического волокна могут сыграть большую роль в определении характеристик гироскопа. Принцип действия большинства оптических гироскопов основан на эффекте Саньяка (Sagnac) или интерферометре Саньяка, суть которого заключается в следующем. Если в замкнутом оптическом контуре в противоположных направлениях распространяются две световые волны, то в случае недвижимого контура фазовые набеги обеих волн, прошедших весь контур в противоположных направлениях, будут одинаковыми. При вращении контура вокруг оси, нормальной к плоскости контура, фазовые набеги волн становятся неодинаковыми, а их разница в общем случае будет пропорциональна угловой скорости вращения контура, площади, охватывающей контур, и частоте электромагнитной волны (ЭМВ). Поскольку во время работы гироскопа площадь и частота ЭМВ остаются неизменными, фазовый сдвиг будет пропорционален только угловой скорости. Использование фотонно-кристаллического волокна для повышения чувствительности является перспективным, оно значительно уменьшает дрейф через термические поляризационные нестойкости и эфект Керра. В статье предложено использовать в полой сердцевине оптического гироскопа фотонно-кристаллическое волокно 1550nmλ, Ø10 мкм вместо обычных волокон.

Ключевые слова: оптический гироскоп; эффект Саньяка; фотонно-кристаллическое волокно.

Ил. 6. Библиогр.: 11 назв.

UDC 621.373.826

Principles of constructing gyroscopes based on photonic crystal (band-gap) fibers / *Al-Sudani Haider Ali Muse* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 100 – 107.

The gyroscope is a device that makes it possible to measure the change in the orientation angles associated rotation of the body relative to an inertial coordinate system. Photonic crystal fiber gyroscopes are a kind of optical gyroscopes that offer many new features beyond that conventional fiber optic gyroscopes can offer. In any case, the properties of the optical fiber can play a large role in determining the characteristics of the gyroscope. The principle of operation of most optical gyroscopes is based on the Sagnac effect or the Sagnac interferometer, the essence of which is as follows. If two light waves propagate in a closed optical circuit in opposite directions, then in the case of an immovable circuit, the phase incursions of both waves that have passed the entire circuit in opposite directions will be the same. When the contour rotates around an axis normal to the contour plane, the phase incursions of the waves become unequal, and their difference in the general case will be proportional to the angular velocity of the contour rotation, the area covered by the contour, and the frequency of the electromagnetic wave (EMW). Since the area and frequency of the EMW remain unchanged during the operation of the gyroscope, the phase shift will be proportional only to the angular velocity. The use of photonic crystal fiber to increase the sensitivity is very promising; it significantly reduces the drift through thermal polarization, resistance, and the Kerr effect. This article suggests the use of photonic-crystal (hollow-core) fiber in optical gyroscope instead of conventional fibers.

Key words: optical gyroscope; photonic crystal (hollow core) fiber; Sagnac effect.

6 fig. Ref: 11 items.

УДК 621.382.232

Модифікація активної області резонансно-тунельного діоду / *К.С. Яцун* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 108 – 112.

В останні роки зріс інтерес до вивчення мезоскопічних структур. У першу чергу це обумовлено розвитком напівпровідникової технології, яка дозволяє створювати структури із розмірами порядку одиниць та десятків нанометрів. Лінійні розміри таких структур поступаються довжині хвилі де-Бройля електронів, тому транспорт електронів визначається, в основному, їх хвильовими властивостями що, у свою чергу приводить до появи цілого ряду нових ефектів.

До мезоскопічних структур можна віднести резонансно-тунельний діод (РТД), вперше запропонований Есакі та Тсу, і який є одним із перших приладів наноелектроніки. Він складається із шару напівпровідника з доволі вузькою забороненою зоною – шару квантової ями (КЯ), розташованого між двома шарами напівпровідника (бар'єрами) з більш широкою забороненою зоною. Ці шари, у свою чергу, розташовуються між шарами (спейсерами) слабко легованого вузького напівпровідника, за якими слідують сильно леговані шари емітера і колектора. У КЯ виникають один або декілька енергетичних рівнів розмірного квантування. Під дією напруги зміщення струм через РТД проходить лише у тому випадку, якщо у емітері присутні електрони які можуть тунелювати. Резонансне тунелювання відбувається на енергетичний рівень у КЯ, а з нього – у колектор, де спектр енергетичних станів – зонний. РТД має дуже високу швидкодію, наприклад відомо, що нелінійні властивості РТД зберігаються аж до 10^4 ТГц. Також РТД має і інші унікальні властивості: він являється єдиним приладом наноелектроніки який може працювати при кімнатній температурі, а на ВАХ РТД спостерігаються ділянки негативної диференційної провідності (НДП).

У статті досліджується принцип дії резонансно-тунельного діоду та детально розглядаються явища тунелювання у нанофізиці. Проводиться розрахунок моделі вольт-амперної характеристики (ВАХ) двохбар'єрного резонансно-тунельного діоду. Досліджено, як зміна коефіцієнтів прозорості та відбиття потенційного бар′єру прямокутної форми впливають на ВАХ РТД. Це дослідження може бути базовим для подальшого розгляду того, як модифікація активної області резонансно-тунельного діоду впливає на його характеристики. Окрім того, результати досліджень дозволяють якісно оцінювати енергію, необхідну електронам для тунелювання крізь структуру РТД.

Ключові слова: потенційний бар′єр; квантове обмеження; тунелювання; квантова яма; рівняння Шредінгера; негативна диференційна провідність; резонансно-тунельний діод.

Іл. 1. Бібліогр.: 6 назв.

УДК 621.382.232

Модификация активной области резонансно-туннельного диода / *К.С. Яцун* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 108 – 112.

В последние годы значительно возрос интерес к изучению мезоскопических структур. В первую очередь это обусловлено развитием полупроводниковой технологии, которая позволяет создавать структуры с размерами порядка единиц и десятков нанометров. Линейные размеры таких структур уступают длине волны де-Бройля электронов, поэтому транспорт электронов определяется, в основном, их волновыми свойствами что, в свою очередь приводит к появлению целого ряда новых эффектов.

К мезоскопическим структурам можно отнести резонансно-туннельный диод (РТД), впервые предложенный Эсаки и Тсу, и который является одним из первых приборов наноэлектроники. Он состоит из слоя полупроводника с довольно узкой запрещенной зоной – слоя квантовой ямы (КЯ), расположенного между двумя слоями полупроводника (барьерами) с более широкой запрещенной зоной. Эти слои, в свою очередь, располагаются между слоями (спейсерами) слабо легированного узкого полупроводника, за которыми следуют сильно легированные слои эмиттера и коллектора. В КЯ возникают один или несколько энергетических уровней размерного квантования. Под действием напряжения смещения ток через РТД проходит только в том случае, если в эмиттере присутствуют электроны, которые могут туннелировать. Резонансное туннелирование происходит на энергетический уровень в КЯ, а из него – в коллектор, где спектр энергетических состояний – зонный. РТД имеет очень высокое быстродействие, например, известно, что нелинейные свойства РТД сохраняются до 10^4 ТГц. Также РТД имеет и другие уникальные свойства: он является единственным прибором наноэлектроники, который может работать при комнатной температуре, а на ВАХ РТД наблюдаются участки отрицательной дифференциальной проводимости (НДП).

В статье исследуется принцип действия резонансно-туннельного диода и рассматриваются явления туннелирования в нанофизике. Проводится расчет модели вольтамперной характеристики (ВАХ) двухбарьерного резонансно-туннельного диода. Исследовано, как изменение коэффициентов прозрачности и отражения потенциального барьера прямоугольной формы влияют на ВАХ РТД. Это исследование может быть базовым для дальнейшего рассмотрения того, как модификация активной области резонансно-туннельного диода влияет на его характеристики. Кроме того, результаты исследований позволяют качественно оценивать энергию, необходимую электронам для туннелирования через структуру РТД.

Ключевые слова: потенциальный бартер; квантовое ограничение; тунелирование; квантовая яма; уравнение Шредингера; отрицательная дифференциальная проводимость; резонансно-туннельный диод.

Ил. 1. Библиогр.: 6 назв.

UDC 621.382.232

Modification of active region of resonant tunnel diode / *K.S. Yatsun* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 108 – 112.

Interest in the study of mesoscopic structures has grown significantly in recent years. This is primarily due to the development of semiconductor technology, which makes it possible to create structures with sizes of the order of units

and tens of nanometers. The linear dimensions of such structures are inferior to the de Broglie wavelength of electrons, so the transport of electrons is determined mainly by their wave properties, which, in turn, leads to a number of new effects.

Mesoscopic structures include the resonant tunnel diode (RTD), first proposed by Esaki and Tsu, and which is one of the first nanoelectronic devices. It consists of a semiconductor layer with a fairly narrow band gap, a quantum well (QW) layer located between two semiconductor layers (barriers) with a wider band gap. These layers, in turn, are located between the layers (spacers) of weakly doped narrow semiconductor, followed by highly doped layers of the emitter and collector. There are one or more energy levels of dimensional quantization in the QW. Under the action of bias voltage, the current passes through the RTD only if the emitter contains electrons that can tunnel. Resonant tunneling occurs at the energy level in the QW, and from there to the collector, where the spectrum of energy states is band. RTD has a very high speed of action, for example, it is known that the nonlinear properties of RTD persist up to 104 THz. The RTD is also of great power: it is the only device of nanoelectronics that can be used at room temperatures, and on the VAC of the RTD the areas of negative differential conductivity (NDC) are observed.

In this article, the principle of a resonant tunneling diode is revealed, and the phenomena of tunneling in nanophysics are examined in detail. The volt-ampere characteristic (VAC) model of a two-barrier resonance tunnel diode is calculated. The paper investigates how the change of transparency coefficients and the reflection of the potential barrier of a rectangular shape affect the VAC of the RTD. This study can be the basis for further consideration of how the modification of the active region of the resonant tunnel diode affects its characteristics. In addition, the results of the research allow us to estimate qualitatively the energy required by electrons for tunneling through the structure of the RTD.

Key words: potential barrier; quantum constraint; tunneling; quantum well; Schrödinger equation; negative differential conductivity; resonant-tunnel diode.

1 fig. Ref: 6 items.

УДК 621.373.072.9

Похибка методів малого параметру при вирішенні укорочених рівнянь синхронізованого автогенератора / *В.В. Рапін* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 113 – 117.

Розглядається застосування аналітичних методів вирішення укорочених рівнянь синхронізованого автогенератора. Це метод квазімалого параметра та комбінований метод малого параметра. В обох методах використовується класичний метод малого параметра. Особливістю його застосування є те, що в даному випадку він використовується для вирішення нелінійних диференційних рівнянь, які не містять малий параметр. Відмінність вказаних методів полягає в отриманні рівнянь першого наближення.

У методі квазімалого параметра – це лінійні диференціальні рівняння, отримані шляхом лінеаризації вихідних нелінійних диференційних рівнянь в області нульової частотної розстройки. У комбінованому методі малого параметра рівняння першого наближення отримані методом аппроксимації вихідних нелінійних диференційних рівнянь. Звичайно для цього було здійснено ряд перетворювань цих рівнянь. Апроксимація дозволила краще представити вихідні нелінійні диференціальні рівняння лінійними диференційними рівняннями. Це призвело до отримання меншої похибки, яка в обох випадках представлялася у вигляді нев'язки, з якої безпосередньо не представляється можливим отримати відносну похибку і дослідити її особливість.

Дослідження відносної похибки методу квазімалого параметра залежно від частотної розлади показало, що це безперервна функція з нульовим значенням при нульовій частотній розладі.

Для комбінованого методу малого параметру функція, що представляє відносну похибку, має розрив при нульовій частотній розстройці. Однак розрив такого виду, тобто з існуючою загальною межею, відноситься до категорії розривів, які можливо усунути.

Ключові слова: синхронізований автогенератор; укорочені рівняння; методи малого параметра; нев 'язка; похибка.

Іл. 4. Бібліогр.: 15 назв.

УДК 621.373.072.9

Погрешность методов малого параметра при решении укороченных уравнений синхронизированного автогенератора / *В.В. Рапин* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 113 – 117.

Рассматривается применение недавно появившихся аналитических методов решения укороченных уравнений синхронизированного автогенератора. Это метод квазималого параметра и комбинированный метод малого параметра. В обоих методах используется классический метод малого параметра. Особенностью его применения является то, что в данном случае он используется для решения нелинейных дифференциальных уравнений, которые не содержат малый параметр. Отличие указанных методов состоит в получении уравнений первого приближения. В методе квазималого параметра – это линейные дифференциальные уравнения, полученные путем линеаризации исходных нелинейных дифференциальных уравнений в окрестности нулевой частотной расстройки. В комбинированном методе малого параметра уравнения первого приближения получены методом аппроксимации исходных нелинейных дифференциальных уравнений. Конечно, был произведен ряд преобразований этих уравнений. Аппроксимация позволила лучше представить исходные нелинейные дифференциальные уравнения линейными дифференциальными уравнениями. Это обеспечило меньшую погреш-

ность, которая в обоих случаях представлялась в виде невязки, из которой непосредственно невозможно получить относительную погрешность и исследовать ее особенность.

Исследование относительной погрешности метода квазималого параметра в зависимости от частотной расстройки позволило установить, что это непрерывная функция с нулевым значением при нулевой частотной расстройке.

Для комбинированного метода малого параметра функция, представляющая относительную погрешность, имеет разрыв при нулевой частотной расстройке. Однако разрыв такого вида, т.е. с существующим общим пределом, относится к категории устранимых разрывов.

Ключевые слова: синхронизированный автогенератор; укороченные уравнения; методы малого параметра; невязка; погрешность.

Ил. 4. Библиогр.: 15 назв.

UDC 621.373.072.9

Error of small parameter methods in solving shortened equations of a synchronized oscillator / *V.V. Rapin* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 113 – 117.

The paper considers the use of recently appeared analytical methods for solving shortened equations of a synchronized oscillator. These are a quasi-small parameter method and a combined small parameter method. Both methods use the classic small parameter method. A peculiarity of their application is that in this case they are used for solving nonlinear differential equations that do not contain a small parameter. The difference between the above methods is in obtaining the equations of the first approximation. In the quasi-small parameter method, they are linear differential equations obtained by linearizing the original nonlinear differential equations in the area of the zero frequency detuning. In the combined small parameter method, the equations of the first approximation are obtained by approximating the original nonlinear differential equations. Of course, a number of transformations of these equations were made for this. The approximation made it possible to obtain better representation of the original nonlinear differential equations by means of linear differential equations. This representation provided a smaller error, which in both cases was presented as a discrepancy. The discrepancy does not allow obtaining a relative error and investigating its peculiarity.

A study of the relative error of the quasi-small parameter method shows that this error is a continuous function of the frequency detuning with a zero value for a zero frequency detuning.

A function representing relative error has a gap at zero frequency detuning for the combined small parameter method. However, this kind of gap can be eliminated by additional function definition.

Key words: synchronized oscillator; shortened equations; small parameter methods; discrepancy; error.

4 fig. Ref: 15 items.

АНТЕНИ ТА ПРИСТРОЇ МІКРОХВИЛЬОВОЇ ТЕХНІКИ
АНТЕННЫ И УСТРОЙСТВА МИКРОВОЛНОВОЙ ТЕХНИКИ
ANTENNAS AND MICROWAVE DEVICES

УДК 662.396.67

Поздовжній розподіл інтенсивності поля круглої сфокусованої апертури / *В.В. Должиков* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 118 – 128.

Антени мікрохвильового і міліметрового діапазонів, сфокусовані в зону Френеля, які зазвичай називають антенами з фокусуванням в зону Френеля, стають все більш популярними. У порівнянні зі звичайними антенами, сфокусованими в далеку зону (синфазними), вони можуть забезпечити найкращі характеристики при відносно невеликій вартості реалізації в системах зв'язку малого радіусу дії, в пристроях бездротової передачі енергії, в установках дистанційного неруйнівного зондування, в пристроях радіочастотної ідентифікації та багатьох інших. В роботі отримано аналітичні вирази для розрахунку основних параметрів, що характеризують поздовжній розподіл інтенсивності поля антени у вигляді круглої сфокусованої апертури з відносно великим діаметром $\left(2R/\lambda \geq 10\right)$: зміщення максимуму інтенсивності щодо точки фокусування, посилення фокусування, глибини фокусування. Розглянуто випадки рівномірного і спадаючого амплітудних розподілів поля збудження. Знайдені наближені співвідношення дозволяють визначити значення згаданих параметрів для будь-яких значень поздовжньої координати точки фокусування, що лежать як в зоні Френеля, так і в далекій зоні. Порівняння з чисельними розрахунками показало, що похибка одержуваних значень параметрів не перевищує 5 %. Результати роботи будуть корисними при розрахунку поля антен у вигляді круглої сфокусованої апертури, а також сфокусованих антенних решіток, що працюють в зоні Френеля.

Ключові слова: сфокусовані апертурні антени; зона Френеля; посилення фокусування; глибина фокусування.

Іл. 13. Бібліогр.: 14 назв.

УДК 662.396.67

Продольное распределение интенсивности поля круглой сфокусированной апертуры / *В.В. Должиков* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 118 – 128.

Антенны микроволнового и миллиметрового диапазонов, сфокусированные в зону Френеля, которые обычно называют антеннами с фокусировкой в зону Френеля, становятся все более популярными. По сравне-

нию с обычными антеннами, сфокусированными в дальнюю зону (синфазными), они могут обеспечить лучшие характеристики при относительно небольшой стоимости реализации в системах связи малого радиуса действия, в устройствах беспроводной передачи энергии, в установках дистанционного неразрушающего зондирования, в устройствах радиочастотной идентификации и многих других. В работе получены аналитические выражения для расчета основных параметров, характеризующих продольное распределение интенсивности поля антенны в виде круглой сфокусированной апертуры с относительно большим диаметром $\left(2R/\lambda \geq 10\right)$: смещения максимума интенсивности относительно точки фокусировки, усиления фокусировки, глубины фокусировки. Рассмотрены случаи равномерного и спадающего амплитудных распределений поля возбуждения. Найденные приближенные соотношения позволяют определить значения упомянутых параметров для любых значений продольной координаты точки фокусировки, лежащих как в зоне Френеля, так и в дальней зоне. Сравнение с численными расчетами показало, что погрешность получаемых значений параметров не превышает 5 %. Результаты работы будут полезны при расчете поля антенн в виде круглой сфокусированной апертуры, а также сфокусированных антенных решеток, работающих в зоне Френеля.

Ключевые слова: сфокусированные апертурные антенны; зона Френеля; усиление фокусировки; глубина фокусировки.

Ил. 13. Библиогр.: 14 назв.

UDC 662.396.67

Longitudinal distribution of the field intensity of a circular focused aperture / *V.V. Dolzhikov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 118 – 128.

Microwave and millimeter-wave antennas focused in their Fresnel zone, which are usually named as near-field focused (NFF) antennas, are becoming increasingly popular. Indeed, when compared to conventional far-field focused antennas, they can guarantee performance improvement at a relatively limited implementation cost, in short-range communication systems, wireless power transfer arrangements, remote nondestructive sensing setups, and radiofrequency identification apparatus, among many others. In this paper, analytical expressions are obtained for calculating the main parameters characterizing the longitudinal distribution of the circular focused aperture field intensity with a relatively large diameter $\left(2R/\lambda \geq 10\right)$: the displacement of the intensity maximum relative to the focal point, focusing gain and depth of focus. Cases of uniform and decreasing amplitude distributions of the excitation field are considered. The found approximate relations make it possible to determine the values of the above parameters for any values of the longitudinal coordinate of the focal point, lying both in the Fresnel zone and in the far zone. Comparison with numerical calculations showed that the error in the obtained parameter values does not exceed 5%. The results of this paper will be useful when calculating the field of antennas in the form of a circular focused aperture, as well as focused antenna arrays operating in the Fresnel zone.

Key words: focused aperture antennas; Fresnel zone; focal shift; focusing gain; depth of focus.

13 fig. Ref: 14 items.

РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ
РАДИОЛОКАЦИЯ И НАВИГАЦИЯ
RADIOLOCATION AND NAVIGATION

УДК 004.89: 621.396

Метод перетворення символьних радарних відміток малопомітних рухомих об'єктів на основі ефекту Тальбота / *В.В. Жирнов, С.В. Солонська* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 129 – 137.

Розглядається метод перетворення символьних зображень радарних відміток малопомітних рухомих повітряних об’єктів з мерехтливими міжперіодними флуктуаціями, що приводять іноді до повного зникнення сигналу, за допомогою ефекту Тальбота. Ці перетворення зводяться до встановлення певної відповідності асимптотичної рівності сприйняття зорових картин, довільним чином мінливих в часі і просторі, до твердження про умови простої рівності сприйняття зображень радіолокаційних відміток, які мають різні частоти флуктуацій. Пропоноване перетворення символьних зображень – це математично обґрунтований метод перетворення символьного зображення малопомітних радарних відміток на основі ефекту Тальбота. Показано, як цей підхід може використовуватися для інтелектуального аналізу радіолокаційних даних за рахунок перетворення й згладжування невидимих на тлі завад мерехтливих флуктуацій сигналу у видимі символьні зображення. По-перше, для автоматичного виявлення й розпізнавання повітряних об’єктів з аналізу зв’язків та функціональних (семантичних) залежностей між ознаками. По-друге, для прийняття рішення на основі семантичних складових символьних зображень радарних відміток. Експериментально перевірено можливість використання таких перетворень для формування частотно-імпульсних кодів символьних зображень флуктуацій радарних відміток типу ангел-луна як важливої характеристики для їх розпізнавання. Сформульовано алгоритми автоматичного формування символьних зображень в асинхронному та синхронному частотно-імпульсному коді. Символьне зображення, представлене таким кодом, є додатковою ознакою для розпізнавання та відсіювання природних завад типу ангел-луна.

Ключові слова: символьне зображення; нестаціонарна радарна відмітка; ефект Табольта; виявлення; розпізнавання; інтелектуальний аналіз.

Іл. 7. Бібліогр.: 13 назв.

УДК 004.89: 621.396

Метод преобразования символьных радарных отметок малозаметных подвижных объектов на основе эффекта Тальбота / *В.В. Жирнов, С.В. Солонская* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 129 – 137.

Рассматривается метод преобразования символьных изображений радарных отметок малозаметных подвижных воздушных объектов с мерцающими межпериодными флуктуациями, приводящими иногда к полному исчезновению сигнала, с помощью эффекта Тальбота. Эти преобразования сводятся к установлению определенного соответствия асимптотического равенства восприятия зрительных картин, произвольным образом меняющихся во времени и пространстве, к утверждению об условиях простого равенства восприятия изображений радиолокационных отметок, которые имеют разные частоты флуктуаций. Предлагаемое преобразование символьных изображений – это математически обоснованный метод преобразования символьного изображения малозаметных радарных отметок на основе эффекта Тальбота. Показано, как этот подход может использоваться для интеллектуального анализа радиолокационных данных за счет преобразования и сглаживания, невидимых на фоне помех мерцающих флуктуаций сигнала в видимые символьные изображения. Во-первых, для автоматического обнаружения и распознавания объектов локации из анализа связей и функциональных (семантических) зависимостей между признаками. Во-вторых, для принятия решения на основе семантических составляющих символьных изображений радарных отметок. Экспериментально проверена возможность использования таких преобразований для формирования частотно-импульсных кодов символьных изображений флюктуаций радарных отметок типа «ангел-эхо» как важной характеристики для их распознавания. Сформулированы алгоритмы автоматического формирования символьных изображений в асинхронном и синхронном частотно-импульсном коде. Символьное изображение, представленное таким кодом, является дополнительным признаком для распознавания и отсеивания естественных помех типа ангел-эхо.

Ключевые слова: символьное изображение; нестационарная радарная отметка; эффект Табольта; обнаружение; распознавание; интеллектуальный анализ.

Ил. 7. Библиогр.: 13 назв.

UDC 004.89: 621.396

Method for transforming symbolic radar marks of low-noticeable moving objects based on the Talbot effect / *V. Zhyrnov, S. Solonskaya* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 129 – 137.

In this paper a method to transform radar images of moving aerial objects with scintillating inter-period fluctuations, sometimes resulting to complete signal fading, using the Talbot effect is considered. These transformations are reduced to the establishment of a certain correspondence of the asymptotic equality of perception of visual images, arbitrarily changing in time and space, in the statement about the conditions of simple equality of perception of images of radar marks that have different frequencies of fluctuations. It is shown how this approach can be used to analyze radar data by transforming and smoothing scintillating signal fluctuations, invisible in the presence of interference, into visible symbolic images. First, to detect and recognize the aerial objects from the analysis of relations and functional (semantic) dependencies between attributes, second, to make a decision based on semantic components of symbolic radar images. The possibility of using such transformation to generate pulse-frequency code of fluctuations of the symbolic radar angel-echo images as an important characteristic for their recognition has been experimentally verified. Algorithms for generating symbolic images in asynchronous and synchronous pulse-frequency code are formulated. The symbolic image represented by such a code is considered as an additional feature for recognizing and filtering out natural interferences such as angel-echoes.

Key words: symbolic image; non-stationary radar marker; Tabolt effect; detection; recognition; intellectual analysis.

7 fig. Ref: 13 items.

УДК 621.396.96, 621.397.48:004.932.2

Методи виявлення-розпізнавання радіолокаційних, акустичних, оптичних і інфрачервоних сигналів безпілотних літальних апаратів / *В.М. Карташов, В.О. Посошенко, В.В. Воронін, В.І. Колесник, А.І. Капуста, М.В. Рибников, Є.В. Першин* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 138 – 153.

Захист різноманітних об'єктів від впливу безпілотних літальних апаратів (БПЛА), що несуть потенційну загрозу у військовій, господарській і повсякденній областях діяльності людини, – одна з актуальних задач сучасності. У даний час відома велика кількість публікацій, присвячених опису методів і систем, заснованих на різних фізичних принципах, які призначені для виявлення і спостереження БПЛА на тлі наявних перешкод. У них розглядаються канали прийому, способи обробки прийнятих інформаційних сигналів і подальшого їх інтелектуального аналізу. Показано, що відомі методи енергетичного виявлення сигналів БПЛА недостатньо ефективні, оскільки операція виконується, як правило, на тлі перешкод, що мають певні структурні подібності з сигналом БПЛА. Значна увага приділяється методам інтерпретації одержуваних даних з використанням навчаємих нейронних мереж. Оскільки кількість публікацій в даній області постійно збільшується, то актуальним відповідно до цього є завдання аналізу, узагальнення та систематизації наявних в літературі даних.

Стаття є оглядовою і присвячена узагальненню і систематизації відомих методів прийому та обробки радіолокаційних, акустичних, оптичних і інфрачервоних сигналів з метою виявлення-розпізнавання, вимірювання координат і параметрів руху БПЛА.

Ключові слова: безпілотний літальний апарат; виявлення; розпізнавання; радіолокаційна станція; содар; відеокамера; зображення; акустичний сигнал.

Іл. 4. Бібліогр.: 86 назв.

УДК 621.396.96, 621.397.48:004.932.2

Методы обнаружения-распознавания радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов / *В.М. Карташов, В.О. Посошенко, В.В. Воронин, В.И. Колесник, А.И. Капуста, Н.В. Рыбников, Е.В. Першин* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 138 – 153.

Защита разнообразных объектов от воздействия беспилотных летательных аппаратов (БПЛА), несущих потенциальную угрозу в военной, хозяйственной и повседневной областях деятельности человека, – одна из актуальных задач современности. Известно большое количество публикаций, посвященных описанию методов и систем, основанных на разных физических принципах, которые предназначены для обнаружения и наблюдения БПЛА на фоне имеющихся помех. В них рассматриваются каналы приема, способы обработки принимаемых информационных сигналов и последующего их интеллектуального анализа. Показано, что известные методы энергетического обнаружения сигналов БПЛА недостаточно эффективны, поскольку операция выполняется, как правило, на фоне помех, имеющих определенные структурные сходства с сигналом БПЛА. Большое внимание уделяется методам интерпретации получаемых данных с использованием обучаемых нейронных сетей. Поскольку количество публикаций в данной области постоянно увеличивается, то актуальной в соответствии с этим является задача анализа, обобщения и систематизации имеющихся в литературе данных.

Статья является обзорной и посвящена обобщению и систематизации известных методов приема и обработки радиолокационных, акустических, оптических и инфракрасных сигналов с целью обнаружения-распознавания, измерения координат и параметров движения БПЛА.

Ключевые слова: беспилотный летательный аппарат; обнаружение; распознавание; радиолокационная станция; содар; видеокамера; изображение; акустический сигнал.

Ил. 4. Библиогр.: 86 назв.

UDC 621.396.96, 621.397.48:004.932.2

Methods for detection-recognition of radar, acoustic, optical and infrared signals of unmanned aerial vehicles / *V.M. Kartashov, V.A. Pososhenko, V.V. Voronin, V.I. Kolesnik, A.I. Kapusta, N.V. Rybnikov, E.V. Pershin* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 138 – 153.

The protection of various objects against the impact of unmanned aerial vehicles (UAVs), which carry a potential threat in the military, economic and everyday areas of human activity, is one of the urgent tasks of our time. Currently, there are a large number of publications devoted to the description of methods and systems based on different physical principles designed to detect and observe UAVs against the background of existing interference. They consider the reception channels, methods of processing the received information signals and their subsequent intelligent analysis. It is shown, that the known methods of energy detection of UAV signals are insufficiently effective, since the operation is performed, as a rule, against a background of noise that has certain structural similarities with the UAV signal. Considerable attention is paid to the methods for interpreting the obtained data using trained neural networks. Since the number of publications in this area is constantly increasing, the task of analyzing, generalizing and systematizing the data available in the literature is relevant in accordance with this.

The article is an overview and it is devoted to the generalization and systematization of known methods of receiving and processing radar, acoustic, optical and infrared signals for detection-recognition, measurement of coordinates and parameters of UAV movement.

Key words: unmanned aerial vehicle; detection; recognition; radar station; sodar; video camera; image; acoustic signal.

4 fig. Ref: 86 items.

УДК 621.396.96

Метод підвищення завадозахищеності радіолокаційних систем ідентифікації «свій-чужий» при дії навмисних корельованих завад / *І.В. Свид, І.І. Обод, О.С. Мальцев, М.Г. Ткач, С.В. Старокожев, А.О. Глущенко, В.С. Чумак* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 154 – 160.

Проаналізовано принципи побудови і структуру систем ідентифікації «свій-чужий» та виявлено, що у існуючій системі зацікавлена сторона має можливість несанкціоновано використати даний інформаційний ресурс для дальнього визначення координат повітряних об’єктів, з одного боку, та перекручування інформації цього інформаційного ресурсу, з другого боку, що призводить до непередбачуваних наслідків. Показано, що найбільш вразливим місцем в системах ідентифікації «свій-чужий» є літаковий відповідач, який істотно впливає на завадостійкість та завадозахищеність ідентифікаційних систем повітряних об'єктів. Запропоновано метод спадкоємного переходу до завадостійких систем ідентифікації «свій-чужий» на основі синхронних мереж систем ідентифікації, який дозволяє істотно розширити методи обслуговування заявок та методи побудови систем. Така методика побудови ідентифікаційних систем виключає наявну проблему розосереджених ідентифікаційних си-

стем, а також проблему часового узгодження сигналів, що поступають з систем первинної та вторинної радіолокації. Запропонований метод спадкоємного переходу до завадостійких систем ідентифікації «свій-чужий» дозволяє виключити можливість несанкціонованого доступу зацікавленої сторони до ідентифікаційних інформаційних ресурсів, що значною мірою підвищує завадозахищеність ідентифікаційної системи в цілому.

Ключові слова: радіолокаційна система; система ідентифікації «свій-чужий»; управління повітряним рухом; повітряний об'єкт; літаковий відповідач; завадозахищеність; завадостійкість; сигнал запиту; сигнал відповіді; оптимізація; мережева структура; відносна пропускна здатність.

Іл. 1. Бібліогр.: 25 назв.

УДК 621.396.96

Метод повышения помехозащищенности радиолокационных систем идентификации «свой-чужой» при действии преднамеренных коррелированных помех / *И.В. Свид, И.И. Обод, А.С. Мальцев, М.Г. Ткач, С.В. Старокожев, А.А. Глущенко, В.С. Чумак* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 154 – 160.

Проанализированы принципы построения и структура систем идентификации «свой-чужой» и выявлено, что в существующей системе заинтересованная сторона имеет возможность несанкционированно использовать данный информационный ресурс для дальнего определения координат воздушных объектов, с одной стороны, и искажения информации этого информационного ресурса, с другой стороны, что приводит к непредсказуемым последствиям. Показано, что наиболее уязвимым местом в системах идентификации «свой-чужой» является самолетный ответчик, который существенно влияет на помехоустойчивость и помехозащищенность идентификационных систем воздушных объектов. В работе предложен метод наследственного перехода к помехоустойчивых систем идентификации «свой-чужой» на основе синхронных сетей систем идентификации, который позволяет существенно расширить методы обслуживания заявок и методы построения систем. Такая методика построения идентификационных систем исключает имеющуюся проблему рассредоточенных идентификационных систем, а также проблему временного согласования сигналов, поступающих из систем первичной и вторичной радиолокации. Предложенный метод наследственного перехода к помехоустойчивых систем идентификации «свой-чужой» позволяет исключить возможность несанкционированного доступа заинтересованной стороной в идентификационные информационные ресурсы, в значительной мере повышает помехозащищенность идентификационной системы в целом.

Ключевые слова: радиолокационная система; система идентификации «свой-чужой»; управления воздушным движением; воздушный объект; самолетный ответчик; помехозащищенность; помехоустойчивость; сигнал запроса; сигнал ответа; оптимизация; сетевая структура; относительная пропускная способность.

Ил. 1. Библиогр.: 25 назв.

UDC 621.396.96

Method for increasing noise immunity of radar "friend or foe" identification systems under the action of intentional correlated interference / *I.V. Svyd, I.I. Obod, O.S. Maltsev, M.G. Tkach, S.V. Starokozhev, A.O. Hlushchenko, V.S. Chumak* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 154 – 160.

The paper analyzes the principles of construction and structure of "friend or foe" identification systems. It is revealed, that the party, interested in the existing system, has the ability of unauthorized use of this information resource for long-range determination of air objects coordinates, on the one hand, and distortion of information of this information resource, on the other hand, which leads to unpredictable consequences. It is shown, that the most vulnerable place in the "friend or foe" identification systems is the aircraft transponder, which significantly affects noise stability and noise immunity of the identification systems of air objects. The paper proposes a method of hereditary transition to noise-immune "friend or foe" identification systems based on synchronous networks of identification systems, which allows expanding significantly the methods of servicing requests and methods of constructing systems. This method of constructing identification systems eliminates the existing problem of dispersed identification systems, as well as the problem of temporal matching of signals coming from primary and secondary radar systems. The proposed method of hereditary transition to noise-immune "friend or foe" identification systems makes it possible to exclude the possibility of unauthorized access to identification information resources by an interested party, significantly increases the noise immunity of the identification system as a whole.

Key words: radar system; "friend or foe" identification system; air traffic control; air object; aircraft transponder; noise immunity; noise stability; request signal; answer signal; optimization; network structure; relative bandwidth.

1 fig. Ref.: 25 items.

РАДІОТЕХНІЧНІ ПРИСТРОЇ ТА ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ
РАДИОТЕХНИЧЕСКИЕ УСТРОЙСТВА И СПОСОБЫ ТЕЛЕКОММУНИКАЦИИ
RADIO ENGINEERING DEVICES AND TELECOMMUNICATION METHODS

УДК 621.396.677.49

Моделі поширення сигналів мереж зв'язку 5 G / *Ю.Ю. Коляденко, М.О. Чурсанов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 161 – 168.

Технологія нового покоління 5G / IMT-2020, як і будь-яка нова технологія, привносить свої специфічні особливості в усі аспекти, що стосуються практики її застосування. Одним з таких особливо важливих аспектів

є електромагнітна сумісність. На етапі підготовки до впровадження радіомереж технології 5G, названої NewRadio, необхідно завчасно подбати про вжиття заходів щодо ефективної оцінки умов електромагнітної сумісності для цих мереж на основі ретельного аналізу особливостей технології 5G, а правильно і точно оцінивши ці умови – успішно забезпечити електромагнітну сумісність радіозасобів нових мереж.

На Всесвітній конференції радіозв'язку ВКР-15 були визначені нові діапазони радіочастот для 5G, в тому числі діапазони сантиметрових і міліметрових хвиль. Цей радіочастотний спектр розміщений в трьох областях: нижче 1 ГГц, від 1 ГГц до 6 ГГц і вище 6 ГГц (до 100 ГГц). В якості головних особливостей спектра, з точки зору ЕМС, можна виділити наступне: різний характер втрат при поширенні сигналу, зокрема значний вплив на рівень втрат додаткових раніше невідомих в стільниковому зв'язку факторів (гази – кисень, водяна пара та ін.).

Розроблено математичну модель поширення сигналів мереж зв'язку 5 G, яка враховує ослаблення сигналів у вільному просторі, ослаблення сигналів, викликане впливом стін і перекриттів поверхів, втрати енергії сигналу при заповненні простору різними предметами, ослаблення сигналів, викликане втратою енергії радіохвиль при поширенні через дощі, ослаблення сигналів, викликане втратою енергії радіохвиль через туман, ослаблення сигналів при поширення через листя дерев, повільні і швидкі випадкові завмирання.

Ключові слова: мережі зв'язку 5G; електромагнітна сумісність; математична модель поширення сигналів.

Табл. 4. Бібліогр.: 26 назв.

УДК 621.396.677.49

Модели распространения сигналов сетей связи 5 G / *Ю.Ю. Коляденко, Н.А. Чурсанов* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 161 – 168.

Технология нового поколения 5G / IMT-2020, как и любая новая технология, привносит свои специфические особенности во все аспекты, касающиеся практики ее применения. Одним из таких особо важных аспектов является электромагнитная совместимость. На этапе подготовки к внедрению радиосетей технологии 5G, названной NewRadio, необходимо заблаговременно позаботиться о принятии мер по эффективной оценки условий электромагнитной совместимости для этих сетей на основе тщательного анализа особенностей технологии 5G, а правильно и точно оценив эти условия – успешно обеспечить электромагнитную совместимость радиосредств новых сетей.

На Всемирной конференции радиосвязи ВКР-15 были определены новые диапазоны радиочастот для 5G, в том числе диапазоны сантиметровых и миллиметровых волн. Этот радиочастотный спектр размещен в трех областях: ниже 1 ГГц, от 1 ГГц до 6 ГГц и выше 6 ГГц (до 100 ГГц). В качестве главных особенностей спектра, с точки зрения ЭМС, можно выделить следующее: различный характер потерь при распространении сигнала, в частности, значительное влияние на уровень потерь дополнительных ранее неизвестных в сотовой связи факторов (газы – кислород, водяной пар и др.).

Разработана математическая модель распространения сигналов сетей связи 5 G, которая учитывает ослабление сигналов в свободном пространстве, ослабление сигналов, вызванное влиянием стен и перекрытий этажей, потери энергии сигнала при заполнении пространства различными предметами, ослабление сигналов, вызванное потерей энергии радиоволн при распространении через дожди, ослабление сигналов, вызванное потерей энергии радиоволн из-за тумана, ослабление сигналов при распространение через листья деревьев, медленные и быстрые случайные замирания.

Ключевые слова: сети связи 5G; электромагнитная совместимость; математическая модель распространения сигналов.

Табл. 4. Библиогр.: 26 назв.

UDC 621.396.677.49

5 G communication network signal propagation models / *Yu.Yu. Kolyadenko, N.A. Chursanov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 161 – 168.

The next generation 5G / IMT-2020 technology, like any new technology, brings its own specific features to all aspects related to the practice of its application. One of these particularly important aspects is electromagnetic compatibility. At the stage of preparation for the introduction of 5G radio networks, called NewRadio, it is necessary to take early measures to assess effectively the electromagnetic compatibility conditions for these networks based on a thorough analysis of the features of 5G technology. Correct and accurate assessments of these conditions means successful provision of the electromagnetic compatibility of radio equipment of new networks.

The World Radio Communication Conference WRC-15 identified new radio frequency bands for 5G, including centimeter and millimeter wave bands. In general, this RF spectrum is located in three regions: below 1 GHz, 1 GHz to 6 GHz, and above 6 GHz (up to 100 GHz). From the EMC standpoint, the following can be distinguished as the main features of this spectrum: different nature of losses during signal propagation, in particular, a significant influence of additional factors (gases – oxygen, water vapor, etc.) on the level of losses previously unknown in cellular communication.

The mathematical model of signal propagation of 5 G communication networks has been developed which takes into account: the attenuation of signals in free space; attenuation of signals caused by the influence of walls and floor slabs, loss of signal energy, when space is filled with various objects; attenuation of signals caused by loss of energy of radio waves, when propagating through rains; signal attenuation due to loss of radio wave energy due to fog; signal attenuation, when propagating through tree leaves, slow and fast random fading.

Key words: 5G communication networks; electromagnetic compatibility; mathematical model of signal propagation.

4 tab. Ref: 26 items.

РАДІОТЕХНІЧНІ СИСТЕМИ
РАДИОТЕХНИЧЕСКИЕ СИСТЕМЫ
RADIO ENGINEERING SYSTEMS

УДК 621.391.82: 004.056.53

Ефективні режими роботи радіозакладних пристроїв для потайного знімання інформації у полі шумових завад / *С.П. Сергієнко, В.Г. Крижановський, Д.В. Чернов, Л.В. Загорулько* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 169 – 174.

Інформаційна безпека сучасного суспільства знаходиться у постійній протидії і постійному удосконаленні технічних засобів, які використовуються для несанкціонованого знімання інформації, і технічних засобів, які цьому заважають. В роботі проаналізовано приклади методів застосування шумових завад для протидії несанкціонованому зніманню інформації. Проаналізовано та показано можливість несанкціонованого знімання інформації пасивними радіопристроями з використанням шумових завад, які застосовуються для боротьби з підслуховуючими пристроями. Передача несакціоновано знятої інформації можливо по радіохвильовим каналам і по низькочастотним каналам з використанням металевих конструкцій або комунікацій будівель. В якості моделі шумових завад використовувався випадковий вузько-смуговий сигнал з законом розподілу Гауса. Електрична модель пристрою моделювалася довгою лінію з високочастотним діодом на кінці. В якості вольт-амперної характеристики діоду використовувалась ідеалізована експонентна залежність струму від напруги. Отримано спектри відбитої хвилі при різних співвідношеннях опору довгої лінії і диференційного опору діоду та зовнішньої напруги зміщення прикладеної до діоду. Проаналізовано режими і особливості передачі аналогової і цифрової інформації радіозакладним пристроєм з використанням енергії радіошумових завад. В радіозакладному пристрої інформація передається відбитою хвилею, спектр якої спотворюється на нелінійним елементі на кінці довгої лінії. Роботу пристрою проаналізовано у всьому можливому частотному діапазоні, пов’язаному з частотним спектром падаючої шумової завади. Розраховано оптимальні параметри елементів пасивної електричної схеми: опір довгої лінії, диференційний опір діоду, напруга зміщення та режим модуляції в залежності від частотного діапазону, в якому можливий витік інформації.

Ключові слова: пасивні радіозаставні пристрої; радіозашумлення що маскує; нелінійне перетворення спектра шуму; захист інформації.

Іл. 8. Бібліогр.: 11 назв.

УДК 621.391.82: 004.056.53

Эффективные режимы работы радиозакладных устройств для скрытого снимания информации в поле шумовых помех / *С.П. Сергиенко, В.Г. Крыжановский, Д.В. Чернов, Л.В. Загорулько* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 169 – 174.

Информационная безопасность современного общества находится в постоянном противодействии и постоянном совершенствовании технических средств, используемых для несанкционированного съема информации, и технических средств, которые этому мешают. В работе проанализированы примеры методов применения шумовых помех для противодействия несанкционированному съему информации. Показана и проанализирована возможность несанкционированного съема информации пассивными радиоустройствами с использованием шумовых помех, которые применяются для борьбы с подслушивающими устройствами. Передача несанкционированно снятой информации возможна как по радиоволновым каналам, так и по низкочастотным каналам с использованием металлических конструкций или коммуникаций зданий. В качестве модели шумовых помех использовался случайный узкополосный сигнал, имеющий закон распределения Гаусса. Электрическая модель устройства моделировалась длинной линией с высокочастотным диодом на конце. В качестве вольтамперной характеристики диода использовалась идеализированная экспоненциальная зависимость тока от напряжения. Получены спектры отраженной волны при различных соотношениях сопротивления длинной линии, дифференциального сопротивления диода и внешнего напряжения смещения, приложенного к диоду. Проанализированы режимы и особенности передачи аналоговой и цифровой информации радиозакладным устройством с использованием энергии радиошумовых помех. В радиозакладном устройстве информация передается отраженной волной, спектр которой искажается на нелинейном элементе, стоящем на конце длинной линии. Работа устройства проанализирована по всему возможному частотному диапазону, связанному с частотным спектром падающей шумовой помехи. Рассчитаны оптимальные параметры элементов пассивной электрической схемы: сопротивление длинной линии, дифференциальное сопротивление диода, напряжение смещения и режим модуляции в зависимости от частотного диапазона, в котором возможна утечка информации.

Ключевые слова: пассивные радиозакладные устройства; маскирующие радио зашумление; нелинейное преобразование спектра шума; защита информации.

Ил. 8. Библиогр.: 11 назв.

UDC 621.391.82: 004.056.53

Effective modes of operation of radio-bombing devices for covert information gathering in the field of noise interference / *S.P. Serhiienko , V.G. Krizhanovski, D.V. Chernov, L.V. Zahoruiko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 169 – 174.

The information security of modern society is in constant counteraction and constant improvement of technical means used for unauthorized information pickup, and technical means that prevent it. The paper analyzes examples of

methods of applying noise interference to counteract the unauthorized pickup. The possibility of unauthorized pickup by passive radio devices using noise interferences is shown and analyzed using noise interferences, which are used to suppress the eavesdropping devices. The transfer of picked up information is possible both by radio wave and low-frequency channels using metal structures or water pipes. As a model of noise interference, a random narrow-band signal with a Gaussian distribution was used. The electrical model of the device was simulated by a transmission line with a high-frequency diode at its end. The idealized exponential dependence of the diode current on the voltage was used. The reflected wave spectra are obtained for different ratios of the transmission line resistance, the differential resistance of diode, and the external offset voltage at the diode. The modes and features of analog and digital information transmission by the radio tab device using energy of radio noises are analyzed. In the radio tab device, the information is transmitted by reflected wave, the spectrum of which is distorted at a nonlinear element placed at the end of transmission line. An analysis of the device operation was carried out along the full possible frequency range associated with the spectrum of the incident noise interference. The optimal elements parameters for the passive electrical circuit are calculated: the resistance of the transmission line, the differential resistance of the diode, the offset voltage, and the modulation mode, depending on the frequency range in which the leak is possible.

Key words: passive radio embedded devices; masking radio noise; nonlinear noise spectrum transformation; information protection.

8 fig. Ref: 11 items.

УДК 621.396

Шумоподібні дискретні сигнали для асинхронних систем кодового поділу радіоканалів / *О.О. Кузнецов, О.А. Смирнов, Т.Ю. Кузнецова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2021. Вип. 205. С. 175 – 183.

Розглянуто шумоподібні дискретні сигнали (псевдовипадкові послідовності) для асинхронних систем кодового розподілу радіоканалів. Асинхронність передбачає використання послідовностей, статистично некорельованих для довільної циклічно зрушеної копії сигналів, тобто коефіцієнт їх взаємної кореляції для довільно обраних початкових точок близький до нуля. Фундаментальною теоретичною межею для цієї характеристики є відома межа Велча. Проведено порівняння кореляційних властивостей різних множин (коди Голда, послідовності Касамі та ін.) З цією фундаментальною межею. Проведено оцінку параметрів різних кодів, наведено відповідну межу і порівняння її з реальними кореляційними характеристиками кодів. Для апроксимації використовувалося розкладання в ряд Лорана і ряд Пюїзе. Також оцінювалися асимптотичні властивості. Розглянуто нові ансамблі шумоподібних дискретних сигналів для асинхронних систем. Ці коди статистично некорельовані, асимптотично квадрат їх взаємної кореляції для довільно обраних початкових точок прагне до теоретичної межі Велча. При цьому кардинальність (потужність безлічі) нових ансамблів сигналів значно вище, ніж у кодів Голда і множин Касамі. Отже, практичне використання таких шумоподібних дискретних сигналів дозволить підвищити ємність асинхронних систем кодового розподілу радіоканалів і здешевити послуги зв'язку. Крім того, нові набори розширювальних сигналів будуть корисні для реалізації так званої м'якої ємності (Soft Capacity), тобто коли при необхідності базова станція може збільшити абонентську ємність при незначному зниженні якості обслуговування.

Ключові слова: шумоподібні дискретні сигнали; розширювальна послідовність; множинний доступ; пряме розширення спектра; асинхронні системи кодового поділу радіоканалів.

Табл. 1. Бібліогр.: 26 назв.

УДК 621.396

Шумоподобные дискретные сигналы для асинхронных систем кодового разделения радиоканалов / *А.А. Кузнецов, А.А. Смирнов, Т.Ю. Кузнецова* // Радиотехника : Всеукр. межвед. науч.-техн. сб. 2021. Вып. 205. С. 175 – 183.

Рассмотрены шумоподобные дискретные сигналы (псевдослучайные последовательности) для асинхронных систем кодового разделения радиоканалов. Асинхронность предполагает использование последовательностей, статистически некоррелированных для произвольной циклически сдвинутой копии сигналов, т.е. коэффициент их взаимной корреляции для произвольно выбранных начальных точек близок к нулю. Фундаментальным теоретическим пределом для этой характеристики является известная граница Велча. Проведено сравнение корреляционных свойств различных множеств (коды Голда, последовательности Касами и пр.) с этим фундаментальным пределом. Проведена оценка параметров разных кодов, приведена соответствующая граница и сравнение ее с реальными корреляционными характеристиками кодов. Для аппроксимации использовалось разложение в ряд Лорана и ряд Пюизо. Также оценивались асимптотические свойства. Рассмотрены новые ансамбли шумоподобных дискретных сигналов для асинхронных систем. Эти коды статистически некоррелированы, асимптотически квадрат их взаимной корреляции для произвольно выбранных начальных точек стремится к теоретической границе Велча. При этом кардинальность (мощность множества) новых ансамблей сигналов значительно выше, чем у кодов Голда и множеств Касами. Следовательно, практическое использование таких шумоподобных дискретных сигналов позволит повысить емкость асинхронных систем кодового разделения радиоканалов и удешевить услуги связи. Кроме того, новые наборы расширяющих сигналов будут полезны для реализации так называемой мягкой емкости (Soft Capacity), т.е. когда при необходимости базовая станция может увеличить абонентскую емкость при незначительном снижении качества обслуживания.

Ключевые слова: шумоподобные дискретные сигналы; расширяющая последовательность; множественный доступ; прямое расширение спектра; асинхронные системы кодового разделения радиоканалов.

Табл. 1. Библиогр.: 26 назв.

UDC 621.396

Noise-like discrete signals for asynchronous code division radio systems / *A.A. Kuznetsov, O.A. Smirnov, T.Y. Kuznetsova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2021. №205. P. 175 – 183.

This article discusses noise-like discrete signals (pseudo-random sequences) for asynchronous code division systems for radio channels. Asynchrony implies the use of sequences that are statistically uncorrelated for an arbitrary cyclically shifted copy of the signals, i.e. their cross-correlation coefficient for arbitrarily chosen starting points is close to zero. The fundamental theoretical limit for this characteristic is the well-known Welch boundary. In this paper, we compare the correlation properties of various sets (Gold codes, Kasami sequences, etc.) with this fundamental limit. The parameters of different codes are estimated, the corresponding bound is shown and compared with the real correlation characteristics of the codes. For the approximation, the Laurent series expansion and the Puiseau series were used. The asymptotic properties were also estimated. The paper also considers new ensembles of noise-like discrete signals for asynchronous systems. These codes are statistically uncorrelated, asymptotically the square of their cross-correlation for arbitrary starting points tends to the theoretical Welch bound. Moreover, the cardinality (power of the set) of new signal ensembles is much higher than that of Gold codes and Kasami sets. Consequently, the practical use of such noise-like discrete signals will increase the capacity of asynchronous code division systems for radio channels and reduce the cost of communication services. In addition, new sets of spreading signals will be useful for the implementation of the so-called. soft capacity, i.e. when, if necessary, the base station can increase the subscriber capacity with a slight decrease in the quality of service.

Key words: noise-like discrete signals; spreading sequence; multiple access; direct spreading of the spectrum; asynchronous code division radio system.

1 tab. Ref: 26 items.

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 205
Українською, російською, та англійською мовами

СБОРНИК НАУЧНЫХ ТРУДОВ
РАДИОТЕХНИКА
Выпуск 205
На украинском, русском и английском языках

COLLECTION OF SCIENTIFIC PAPERS
RADIOTEKHNIKA
Issue 205
In Ukrainian, Russian and English

Коректор Л.І. Сащенко

Підп. до друку 05.07.2021. Формат 60x90/8. Папір офсет. Гарнітура Таймс. Друк. ризограф.
Ум. друк. арк**.** 10,9. Обл.-вид. арк. 9,36. Тираж 300 прим. Зам. № 464. 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.