MINISTRY OF EDUCATION AND SCIENCE OF UKRAINE

KHARKIV NATIONAL
UNIVERSITY OF RADIO ELECTRONICS

RADIOTEKHNIKA

All-Ukrainian
interdepartmental scientific and technical collection

ISSN 0485-8972
eISSN 2786-5525

Founded in 1965

I S S U E 2 1 2

Kharkiv
Kharkiv National
University of Radio Electronics
2023

UDC 621.3

The collection is included in the List of scientific professional publications of Ukraine, category «Б», technical and physical-mathematical sciences (approved by orders of the Ministry of Education and Science from 17.03.2020 № 409; from 02.07.2020 № 886; from 24.09.2020 № 1188) by specialties: 171 – Electronics; 172 – Telecommunications and Radio Engineering; 173 – Avionics; 125 – Cybersecurity; 151 – Automation and Computer-Integrated Technologies; 152 – Metrology and Information-Measuring Equipment; 153 – Micro- and Nanosystem Technology; 163 – Biomedical Engineering; 105 – Applied Physics and Nanomaterials.

Website: rt.nure.ua

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

The authors are responsible for the content of the article.

Editorial Team

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

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

Recommended by the Scientific and Technical Council of Kharkiv National University of Radio Electronics, protocol № 4 dated 28.03.2023.

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

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ РАДІОЕЛЕКТРОНІКИ

РАДІОТЕХНІКА

Всеукраїнський
міжвідомчий науково-технічний збірник

ISSN 0485-8972
eISSN 2786-5525

Засновано в 1965 р.

В И П У С К 2 1 2

Харків
Харківський національний
університет радіоелектроніки
2023

УДК 621.3

Збірник включено до Переліку наукових фахових видань України, категорія "Б", технічні та фізико-математичні науки (затверджено наказами МОНУ від 17.03.2020 № 409; від 02.07.2020 № 886; від 24.09.2020 № 1188) за спеціальностями: 171 – Електроніка; 172 – Телекомунікації та радіотехніка; 173 – Авіоніка; 125 – Кібербезпека; 151 – Автоматизація та комп'ютерно-інтегровані технології; 152 – Метрологія та інформаційно-вимірювальна техніка; 153 – Мікро- та наносистемна техніка; 163 – Біомедична інженерія; 105 – Прикладна фізика та наноматеріали.

Сайт: rt.nure.ua

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

За зміст статті відповідальні автори.

Рекомендовано Науково-технічною радою Харківського національного університету радіоелектроніки, протокол № 4 від 28.03.2023.

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

CONTENT

SYSTEMS AND METHODS OF INFORMATION PROTECTION

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS

RADAR AND RADIONAVIGATION

MEANS OF TELECOMMUNICATIONS

ЗМІСТ

СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.01**

А.О. ГАПОН, В.М. ФЕДОРЧЕНКО, канд. техн. наук,
О.В. СЄВЄРІНОВ, канд. техн. наук

МЕТОДИ ТА ЗАСОБИ СТАТИЧНОГО ТА ДИНАМІЧНОГО АНАЛІЗУ КОДУ

Вступ

Проблема якості коду є однією з базових, що впливають на якість програмного продукту. Існують засоби контролю та покращення його якості за допомогою як ручних, так і автоматизованих підходів. Процедура перегляду коду - це ручний підхід перевірки якості коду. Ця практика є корисною та необхідною під час розробки програмного продукту, але у той же час може займати багато часу та не прибирає людський фактор. Автоматизовані засоби аналізу коду є ключовими інструментами, які доповнюють ручний перегляд коду та забезпечують високу якість коду та, відповідно, його безпечність.

Статичний аналіз коду

Статичний аналіз програмного коду використовується з початку 1960-х років для оптимізації роботи компіляторів. Пізніше це виявилося корисним для інструментів налагодження, а також для фреймворків розробки програмного забезпечення. Зростає кількість інструментів, які дозволяють використовувати статичний аналіз коду, деякі з яких є інструментами з відкритим кодом і дозволяють аналізувати кілька різних мов програмування.

Інструменти статичного аналізу використовуються для створення звітів і виявлення певних відхилень від встановлених стандартів якості коду. Однак ці інструменти не дозволяють автоматично модифікувати вихідний код. Рішення змінити спосіб структурування раніше написаного коду залишається в руках розробників програмного забезпечення.

Інструменти статичного аналізу коду допомагають розробникам програмного забезпечення, створюючи звіт, де вказується причина певного дефекту, а також те, як цей дефект можна виправити.

Проте залишається відкритим питання, як підійти до недоліків, усунути виділені недоліки та переробити вихідний код.

Коли справа стосується інструментів статичного аналізу коду, великою проблемою є те, що існує велика кількість інструментів, які забезпечують цей тип аналізу. Кількість інструментів постійно збільшується, і необхідно класифікувати існуючі інструменти відповідно до мови програмування, яку вони підтримують, і типів дефектів, які вони виявляють.

Процес статичного аналізу коду корисний не лише для оптимізації роботи компілятора (що було початковою метою), але й для виявлення невідповідностей і можливих дефектів. Таким чином, можна створити інструменти, які допоможуть розробникам зрозуміти поведінку програми та виявити різні дефекти програми без її виконання. Інструментами, які використовуються для статичного аналізу коду, є програми, які пояснюють поведінку інших програм [1].

Статичний аналіз коду значно швидший, ніж звичайне тестування, і може виявити будь-який дефект, видимий у вихідному коді програми. Якщо порівняти інструменти для статичного аналізу коду з ручним переглядом коду розробниками програмного забезпечення, можна зробити висновок, що перегляд коду за допомогою інструменту також є набагато швидшим та ефективнішим. Однак щоб статичний аналіз коду виявив будь-який дефект, він повинен бути видимим у вихідному коді.

Стандартний цикл перевірки коду включає чотири основні фази:

1) встановити цілі;
2) запустити інструмент статичного аналізу;
3) переглянути код;
4) зробити рефакторінг.

Окрім стандартного циклу на рис. 1 показано кілька потенційних зворотних зв'язків або незначних ітерацій між стандартними кроками циклу, що робить цикл складнішим, ніж стандартна процедура.

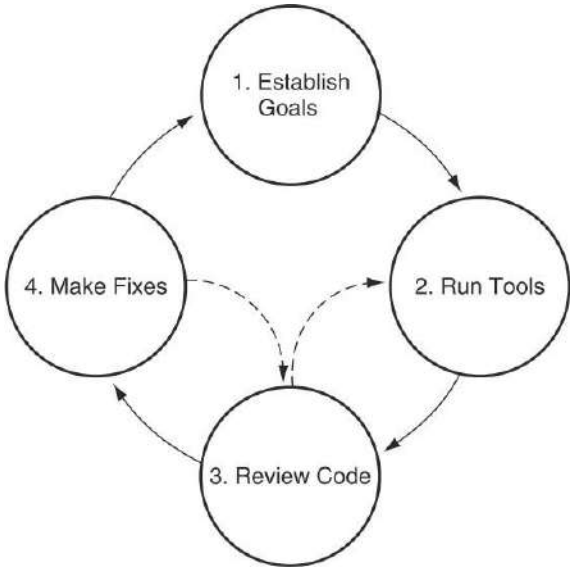


Рис. 1. Цикл огляду коду

Завдання, які вирішуються програмами статичного аналізу коду, можна розділити на три категорії:

1. Виявлення помилок у програмах.

2. Рекомендації щодо оформлення коду. Деякі статичні аналізатори дозволяють перевіряти, чи відповідає вихідний код, прийнятий у компанії, стандарту оформлення коду. Мається на увазі контроль кількості відступів у різних конструкціях, використання пробілів/символів табуляції тощо.

3. Підрахунок метрик. Метрика програмного забезпечення – це міра, що дозволяє отримати чисельне значення певної якості програмного забезпечення або його специфікацій. Існує велика кількість різноманітних метрик, які можна підрахувати, використовуючи інші інструменти.

Серед недоліків статичного аналізу коду можна виділити:

1. Статичний аналіз, як правило, слабкий у діагностиці витоків пам'яті та паралельних помилок. Щоб виявляти такі помилки, фактично необхідно віртуально виконати частину програми. Це дуже складно реалізувати. Також подібні алгоритми вимагають дуже багато пам'яті та процесорного часу. Як правило, статичні аналізатори обмежуються діагностикою найпростіших випадків. Більш ефективним способом виявлення витоків пам'яті та паралельних помилок є використання інструментів динамічного аналізу.

2. Програма статичного аналізу попереджає про підозрілі місця. Це означає, що насправді код може бути абсолютно коректний. Це називається хибно позитивними спрацьовуваннями. Зрозуміти, чи вказує аналізатор на помилку чи видав хибне спрацювання, може лише програміст. Необхідність переглядати помилкові спрацьовування забирає робочий час і послаблює увагу до тих ділянок коду, де насправді містяться помилки.

Методи та засоби статичного аналізу коду

Одним із основних алгоритмів статичного аналізу є data-flow analysis (аналіз потоку даних). Завдання такого аналізу – визначити в кожній точці програми деяку інформацію про дані, якими оперує код. Інформація може бути різною, наприклад, за типом даних або значенням. Залежно від цього, яку інформацію потрібно визначити, можна сформулювати завдання аналізу потоку даних.

Наприклад, якщо потрібно визначити, чи є вираз константою, і навіть значення цієї константи, то вирішується завдання поширення констант (constant propagation). Якщо необхідно визначити тип змінної, можна говорити про завдання поширення типів (type propagation). Якщо необхідно зрозуміти, які змінні можуть вказувати на певну область пам'яті (зберігати ті самі дані), то йдеться про завдання аналізу синонімів (alias analysis). Існує безліч інших завдань аналізу потоку даних, які можуть використовуватись у статичному аналізаторі. Як і етапи побудови моделі коду, ці завдання також використовуються в компіляторах [2].

У теорії побудови компіляторів описані рішення задач внутрішньопроцедурного аналізу потоку даних (відстежити дані необхідно у межах однієї процедури/функції/методу). Рішення спираються на теорію алгебраїчних решіток та інші елементи математичних теорій. Вирішити задачу аналізу потоку даних можна за поліноміальний час, тобто за прийнятний для обчислювальних машин час, якщо умови задачі задовольняють умовам теореми про дозвіл, що на практиці відбувається далеко не завжди.

Symbolic execution (символьне виконання) – цей метод передбачає абстрактний рух шляхами програми, що імітує її виконання залежно від вхідних даних, що супроводжується зміною стану програми у різних точках. Суть методу символьного виконання полягає в розбитті безлічі вхідних даних на класи еквівалентності, що дозволяє оперувати при аналізі не окремими вхідними значеннями (число яких може бути дуже великим та експоненційно зростати залежно від кількості вхідних аргументів) та їх перебором, а цілими класами еквівалентності, число яких може бути і не кінцевим, але не перевищує загальну кількість комбінацій окремих вхідних значень. Однак, як правило, кількість класів еквівалентності комбінацій вхідних даних виявляється значно нижчою від числа всіх можливих комбінацій вхідних даних, що різко збільшує можливості аналізатора з обробки шляхів виконання.

Abstract interpretation (абстрактна інтерпретація) – це теорія звукової апроксимації семантики комп'ютерних програм, заснована на монотонних функціях над впорядкованими множинами, особливо решітками. Його можна розглядати як часткове виконання комп'ютерної програми, яка отримує інформацію про свою семантику (наприклад, потік управління, потік даних) без виконання всіх обчислень.

Його основне конкретне застосування – формальний статичний аналіз, автоматичне вилучення інформації про можливе виконання комп'ютерних програм; такі аналізи мають основні застосування:

1) всередині компіляторів для аналізу програм, щоб вирішити, чи можуть бути застосовані певні оптимізації або перетворення;

2) для налагодження чи навіть сертифікації програм проти класів помилок.

Для мови програмування чи специфікації абстрактна інтерпретація складається із надання кількох семантик, пов'язаних відносинами абстракції. Семантика – це математична характеристика можливої поведінки програми. Найбільш точна семантика, що дуже точно описує фактичне виконання програми, називається конкретною семантикою.

Мета статичного аналізу полягає в тому, щоб в якийсь момент отримати семантичну інтерпретацію, що обчислюється. Наприклад, можна вибрати уявлення про стан програми, що маніпулює цілісними змінними, забувши фактичні значення змінних і зберігши тільки їх знаки (+, – або 0). Для деяких елементарних операцій, таких як множення, така абстракція не втрачає точності: щоб отримати знак твору достатньо знати знак операндів. Для інших операцій абстракція може втратити точність: наприклад, неможливо дізнатися знак суми, операнди якої відповідно позитивні і негативні.

Необхідно знайти компроміс між точністю аналізу та її розв'язаністю (обчислюваністю) чи зручністю читання (обчислювальними витратами).

Насправді певні абстракції адаптуються як до властивостей програми, які потрібно проаналізувати, так і до набору цільових програм.

Динамічний аналіз коду

Динамічний аналіз коду – це метод аналізу програми безпосередньо під час виконання. Звідси випливає, що з вихідного коду в обов'язковому порядку має бути отриманий файл, що виконується, тобто не можна в такий спосіб проаналізувати код, що містить помилки компіляції або збірки. Динамічний аналіз виконується за допомогою набору даних, що подаються на вхід досліджуваної програми. Тому ефективність аналізу безпосередньо залежить від якості та кількості вхідних даних для тестування. Саме від них залежить повнота покриття коду, яка буде одержана за результатами тестування [3].

Використовуючи динамічне тестування, можна отримати такі метрики та попередження:

1. Ресурси, що використовуються: час виконання програми в цілому або її окремих модулів, кількість зовнішніх запитів (наприклад, до бази даних), кількість використовуваної оперативної пам'яті та інших ресурсів.

2. Ступінь покриття коду тестами та інші метрики програми.

3. Програмні помилки: розподіл на нуль, розіменування нульового покажчика, витоку пам'яті, "стан гонки".

4. Детектування деяких вразливостей.

До основних переваг динамічного аналізу коду відносять:

1. Можливість проводити аналіз програми без необхідності доступу до вихідного коду. Тут варто зробити застереження, так як програми для динамічного аналізу розрізняють за способом взаємодії з програмою, що перевіряється. Наприклад, поширений спосіб проведення динамічного аналізу шляхом попереднього інструментування вихідного коду. У цьому випадку доступ до коду програми, що перевіряється, буде необхідний.

2. Можливість виявлення складних помилок, пов'язаних із роботою з пам'яттю: вихід за обсяги масиву, виявлення витоків пам'яті.

3. Можливість проводити аналіз багатопоточного коду безпосередньо в момент виконання програми, тим самим виявляти потенційні проблеми, пов'язані з доступом до ресурсів, що розділяються; можливі deadlock ситуації.

4. У більшості реалізацій поява хибних спрацьовувань неможлива, оскільки виявлення помилок відбувається під час виконання програми, таким чином, виявлена помилка не є передбаченням, зробленим з урахуванням аналізу моделі програми, а констатацією факту її виникнення.

Перерахуємо недоліки, які притаманні динамічному аналізу коду:

1. Не можна гарантувати повного покриття коду, тобто, швидше за все, відсоток коду програми, який був проаналізований у процесі динамічного тестування, не буде рівним ста відсоткам.

2. Практично не виявляються помилки логічного типу. Наприклад, з погляду динамічного аналізатора, завжди справжня умова не є помилкою, оскільки така некоректна перевірка просто зникає ще на етапі компіляції програми.

3. Тяжко локалізувати місце з помилкою у вихідному коді.

4. Більш висока складність використання порівняно зі статичним аналізом, оскільки для досягнення більшої ефективності динамічного аналізу програмі, що тестується, потрібно подання достатньої кількості вхідних даних, щоб отримати більш повне покриття коду.

Динамічне тестування найбільш важливо в тих областях, де головним критерієм є надійність програми, час відгуку або споживані ресурси. Це може бути, наприклад, система реального часу, що управляє відповідальною ділянкою виробництва, або сервер бази даних. У таких областях будь-яка допущена помилка може бути критичною.

Методи та засоби динамічного аналізу коду

Метод *інструменталізації.* Контрольно-вимірювальні прилади виконують вимірювання програми. Це та сама основа, на якій ми можемо відслідковувати, усувати неполадки, налагоджувати, профільтрувати та розуміти, як працюють програми та чому вони працюють певним чином.

На практиці інструменталізація може бути такою ж простою, як запис часу, в який відбувається виконання функції, та його реєстрація, щоб допомогти звузити місце, де існує вузьке місце у продуктивності [4].

Якщо програма споживає більше пам'яті, ніж очікувалося, отримання зразків використання пам'яті з часом (або зразків розподілу пам'яті та показників складання сміття) також може надати цінну інформацію для відстеження витоку пам'яті. Запис вхідних та вихідних даних функцій або цілих систем (наприклад, корисного навантаження запиту та відповіді служби HTTP) може допомогти у налагодженні програм з несподіваними вхідними даними або неправильною логікою. Це лише деякі приклади використання, які показують, наскільки важливим є інструментування для виявлення та вирішення проблем за допомогою моніторингу та усунення несправностей.

Є два способи додати інструменталізацію у програми: вручну чи автоматично. На продуктивність працюючого коду впливає як сам факт впровадження інструментування, так і те, як саме команда реалізує інструментування.

Один із способів написання коду приладу – ручна інструменталізація. Написати додатковий код для виконання вимірювань разом із рештою коду програми. Однак інструментування фрагментів коду вручну може бути важким завданням. І в міру того, як проєкти та команди збільшуються в розмірах, стає дедалі важче стандартизувати, що і як потрібно використовувати. Крім того, все, що робиться вручну і повторюється, схильне до помилок: або про цей метод забувають у деяких місцях, або реалізовують неправильно.

Більшість служб моніторингу та усунення несправностей надають спеціалізовані автоматизовані засоби інструментування коду: для мов, середовищ виконання та фреймворків.

Автоматичне інструментування зазвичай реалізується шляхом додавання проміжного програмного забезпечення, яке обертає певні важливі фрагменти коду логікою інструментування. Типовим прикладом є проміжне програмне забезпечення для HTTP-запиту, яке вимірює час витрат на отримання відповіді, а також інформацію про запит та відповідь, таку як код стану та корисні дані.

Підтримка автоматичного інструментування на різних мовах програмування може відрізнятися. Динамічні інтерпретовані мови часто використовують такі методи для додавання автоматичного інструментарію, у той час як мови, що компілюються за допомогою байт-коду, такі як Java, дозволяють модифікувати байт-код під час виконання для досягнення того ж ефекту.

Рекомендується використовувати автоматичну інструменталізацію. Це часто забезпечує набагато краще охоплення з розумними значеннями за умовчанням, ніж самостійна робота, і практично не вимагає часу для реалізації та обслуговування. Ручну інструменталізацію варто залишити для окремих випадків, якщо такі є, які не охоплені автоматичним підходом.

Серед засобів інструменталізації можна виділити:

1. Трасування коду – отримання інформаційних повідомлень про виконання програми протягом його роботи.

2. Налагодження програми та (структурована) обробка винятків – відстеження та виправлення помилок програмістів у додатку ще на стадії його розробки.

3. Профілювання – набір методик відстеження продуктивності коду, включаючи вимірювання.

4. Лічильники продуктивності – це компоненти, що дозволяють відстежувати рівень продуктивності програми.

5. Реєстратори подій – компоненти, які дозволяють отримувати сповіщення та відстежувати ключові події під час виконання програми.

Інший метод це – *динамічне тестування*. Динамічний аналіз відноситься до вивчення фізичної реакції системи на змінні, які не є постійними та змінюються з часом. При динамічному тестуванні програмне забезпечення має бути скомпільоване та запущене. Воно включає роботу з програмним забезпеченням, введення вхідних значень і перевірку відповідності виводу очікуваним шляхом виконання конкретних тестових випадків, які можна виконати вручну або з використанням автоматизованого процесу. Це відрізняється від статичного тестування.

Процес та функції динамічного тестування у розробці програмного забезпечення можна розділити на модульне тестування, інтеграційне тестування, системне тестування, приймальне тестування та регресійне тестування.

Модульне тестування – це тест, який фокусується на правильності основних компонентів програмного забезпечення. Модульне тестування відноситься до категорії тестування білої скриньки. У всій системі контролю якості модульне тестування має бути завершено групою продуктів, а потім програмне забезпечення передається до відділу тестування.

Інтеграційне тестування використовує визначення того, чи правильно підключені інтерфейси між різними модулями у процесі інтеграції всього програмного забезпечення.

Тестування програмної системи, яка завершила інтеграцію, називається системним тестом, і мета тесту – переконатися, що правильність та продуктивність програмної системи відповідають вимогам, зазначеним у її специфікаціях. Тестувальники повинні дотримуватися встановленого плану тестування. При тестуванні надійності та простоти використання програмного забезпечення його введення, виведення та іншу динамічну робочу поведінку слід порівнювати зі специфікаціями програмного забезпечення. Якщо специфікація програмного забезпечення неповна, системний тест більше залежить від досвіду роботи та суджень тестувальника, такого тесту недостатньо. Системний тест – це тестування «чорної скриньки».

Приймальне тестування – це останній тест перед введенням програмного забезпечення в експлуатацію. Це пробний процес перевірки програмного забезпечення покупцем. У реальній роботі компанії це зазвичай реалізується шляхом звернення до замовника із проханням спробувати чи випустити бета-версію програмного забезпечення. Приймальне тестування – це тестування методом «чорного ящика».

Метою регресійного тестування є перевірка та зміна результатів приймального тестування на етапі обслуговування програмного забезпечення. Наприклад, обробка скарг клієнтів є здійсненням регресійного тестування.

Роль статичного та динамічного аналізу у безпеці

Зараз безпека є ключовим питанням розробки програмного забезпечення для різноманітних пристроїв та є важливим аспектом конкурентоспроможності бізнесу. Розробка безпеки в продукті на ранніх етапах має вирішальне значення для зменшення як ризику, так і вартості подальших загроз безпеки. Інструменти статичного та динамічного аналізу коду відіграють важливу роль у повному наборі інструментів програмного забезпечення та допомагають прискорити розробку безпечного програмного забезпечення.

Атака на підприємство може знизити продуктивність, зв’язати ресурси, зашкодити довірі та знизити прибутки. Оскільки більшість сучасних загроз спрямовані на прикладний рівень, аналіз безпеки коду є обов’язковим для будь-якої конкурентоспроможної організації. Аналіз додатків шукає в програмному забезпеченні такі вразливості, як бекдори додатків або зловмисний код, щоб їх можна було виправити, перш ніж їх виявлять і використають хакери.

Висновки

Статичний та динамічний аналізи коду є важливими інструментами аналізу коду, які доповнюють один одного та виводять рівень безпеки коду на новий рівень. При поєднанні аналізу цих підходів формується новий підхід – гібридний аналіз коду, оскільки поєднуючи

ці підходи, можна з'ясувати більш повну картину, ніж використовуючи їх окремо, бо аналізується повна поведінка програмного забезпечення.

Список літератури:

1. OWASP [Електронний ресурс]. Режим доступу: //owasp.org/www–community/controls/Static_Code_Analysis
2. Citeseerx[Електронний ресурс]. Режим доступу: https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.394.5540
3. Techopedia [Електронний ресурс]. Режим доступу: https://www.techopedia.com/definition/30958/dynamic–application–security–testing–dast
4. Contrast security[Електронний ресурс]. Режим доступу: https://www.contrastsecurity.com/glossary/dynamic–application–security–testing

Надійшла до редколегії 04.03.2023

Відомості про авторів:

Гапон Андрій Олександрович – Харківський національний університет радіоелектроніки, аспірант кафедри безпеки інформаційних технологій; Україна; e-mail: gapon.andrei@gmail.com; ORCID: https://orcid.org/0000-0003-2560-7426

Федорченко Володимир Миколайович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри електронних обчислювальних машин, факультет комп'ютерної інженерії та управління, Україна; e-mail: volodymyr.fedorchenko@nure.ua, ORCID: https://orcid.org/0000-0001-7359-1460

Сєвєрінов Олександр Васильович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри безпеки інформаційних технологій, факультет комп'ютерної інженерії та управління, Україна; e-mail: oleksandr.sievierinov@nure.ua, ORCID: https://orcid.org/0000-0002-6327-6405

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.02**

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

АТАКИ БІЧНИМИ КАНАЛАМИ НА CRYSTALS-KYBER, КОНТРЗАХОДИ ТА ПОРІВНЯННЯ З АЛГОРИТОМ СКЕЛЯ (ДСТУ 8961-2019)

Вступ

Kyber [1] – це постквантовий алгоритм на основі решітки, заснований на складності задачі M-LWE. Kyber пропонує безпечну схему шифрування з відкритим ключем (PKE) проти атаки вибраного відкритого тексту (CPA) і захищений механізм інкапсуляції ключів проти атаки вибраного шифротексту (CCA). Kyber має три рівні безпеки: Kyber512, Kyber768 і Kyber1024, де кількість постквантових бітів безпеки становлять 107, 166 і 232 відповідно.

У даній роботі проведено систематичне дослідження атак бічними каналами (SCA) і атак із впровадженням помилок (помилками) (FIA) на структуровані схеми на основі решітки, з основним акцентом на механізмі інкапсуляції ключів Kyber (KEM), який є одним із провідних кандидатів в процесі стандартизації NIST для постквантової криптографії (PQC).

Враховуючи широкий спектр відомих атак, одночасний захист від усіх атак потребує впровадження індивідуальних засобів захисту/протидії. У роботі представлено ряд індивідуальних контрзаходів, здатних забезпечити захист/пом'якшення щодо існуючих SCA/FIA для Kyber KEM [1].

Проведена на платформі на основі ARM Cortex-M4 оцінка продуктивності показує, що представлені спеціальні контрзаходи спричиняють розумні витрати на продуктивність. Тому, можна сказати, що у роботі наводено аргументи на користь використання спеціальних контрзаходів у реальних реалізаціях схем на основі решітки, або окремо, або як підсилення загальних контрзаходів, таких як маскування.

1. Атаки бічними каналами на Kyber KEM

Атаки бічними каналами на Kyber KEM можна класифікувати наступним чином: на основі мети зловмисника та на основі доступу до входів/виходів цілі.

На основі мети.

Такі атаки можна умовно розділити на дві категорії: атаки відновлення повідомлень і атаки відновлення ключа [2].

Атаки відновлення повідомлення намагаються відновити повідомлення m з дійсного зашифрованого тексту ct, що відповідає дійсному обміну ключами між двома легітимними сторонами. Знання m призводить до відновлення спільного секрету або ключа сеансу K, тим самим ставлячи під загрозу чи компрометуючи конфіденційність відповідного цільового сеансу.

Атаки на відновлення ключа намагаються відновити довгостроковий приватний ключ sk. Відновлення sk призводить до відновлення всіх сеансових ключів K, отриманих з його допомогою. В умовах, коли sk оновлюється для кожного обміну ключами, ефективність атак відновлення ключа та атак відновлення повідомлень є однаковою.

На основі доступу до входів/виходів цілі.

Такі атаки також можна поділити на дві категорії: атаки з відомим шифротекстом (KCA) і атаки з обраним шифротекстом (CCA).

Атаки з відомим шифротекстом – це атаки, які припускають, що зловмиснику відомі лише зашифровані тексти. У такому випадку зловмисник може лише пасивно спостерігати (проводити моніторинг) за цільовим пристроєм, не маючи можливості встановити з ним зв'язок.

Атаки з обраним шифрованим текстом припускають здатність супротивника встановити зв'язок із ціллю. Це застосовується у випадку, коли зловмисник може звертатися до цільового пристрою декапсуляції за допомогою обраних ним зашифрованих текстів.

1.1. Атаки відновлення ключа – відомий шифротекст

Атаки відновлення ключа у випадку з відомим шифротекстом зазвичай спрямовані на витік із операцій, які безпосередньо маніпулюють секретним модулем s в рамках процедури декапсуляції. Поліноміальне множення на основі NTT, що використовується в процедурі дешифрування, можна використовувати для відновлення ключа. Primas та ін. у роботі [3] представили SCA, націлену на NTT, покладаючись на методи на основі Soft-Analytical Side-Channel Attack (SASCA) [4] для відновлення ключа. Їм вдалося відновити весь ключ в одному захопленому сліді бічного каналу (Power/EM) зі схеми Ring-LWE, що працювала на мікроконтролері ARM Cortex-M4. Атака була спрямована на витік з операції $INTT(\hat{u}' \circ s)$ у процедурі дешифрування. Метою є відновлення вхідних даних $(\hat{u}' \circ s)$, що призводить до відновлення приватного ключа s.

Атака проходить у дві фази. По-перше, фаза профілювання використовується для створення шаблонів для проміжних операцій. На етапі атаки ці шаблони зіставляються з відповідними сегментами в отриманому сліді атаки, а результати об'єднуються за допомогою відомого алгоритму поширення переконань (Belief Propagation) [5] для відновлення приватного ключа.

У атаки є кілька недоліків [2]:

- необхідність тривалого етапу профілювання (з понад 100 мільйонами шаблонів;
- детальне знання реалізації INTT;
- вимога відносно високого відношення «сигнал-шум» (SNR) для успішного відновлення ключа.

1.2. Атаки відновлення ключа – обраний шифротекст

Зловмисник може створювати обрані зашифровані тексти, які під час декапсуляції мають здатність посилювати секретно-залежний витік з кількох операцій у рамках процедури декапсуляції. Далі наведено різні типи CCA (атак з обраним шифротекстом), які є застосовними для відновлення ключа [2].

Атаки бічними каналами на основі Оракула.

Це основна категорія атак відновлення ключа у випадку атак з обраним шифротекстом. Вони застосовуються наступним чином: зловмисник робить запит на процедуру декапсуляції за допомогою власноруч створених шифротекстів. Ці шифротексти створені таким чином, що розшифроване повідомлення m' дуже тісно пов'язане з цільовою частиною секретного ключа або, у деяких випадках, із усім секретним ключем. Зловмисник використовує витік із операцій, що займаються обробкою розшифрованого повідомлення, щоб відновити його, таким чином реалізуючи практичний оракул бічного каналу. Така інформація, отримана через кілька ретельно створених зашифрованих текстів, розкриває повний приватний ключ.

Нижче наведено три основні підкатегорії CCA (атак з обраним шифротекстом) з використанням бічних каналів на основі оракула [2]:

1. Атаки бічними каналами з перевіркою відкритого тексту на основі Оракула (Оракул_ПВТ).

Основна ідея цих атак полягає в створенні зашифрованих текстів, щоб обмежити кількість можливих розшифрованих повідомлень. Крім того, значення розшифрованого повідомлення також залежить від єдиного цільового коефіцієнта секретного ключа для вибраних зашифрованих текстів. Витік із бічного каналу з операцій, що займаються обробкою повідомлення використовується для створення екземпляра оракула перевірки відкритого тексту (PC) для відновлення ключа. D'Anvers та ін. у роботі [6] представили атаку бічними каналами з перевіркою відкритого тексту на основі оракулу на схеми PQC, такі як LAC і RAMSTAKE, яка використовує часовий бічний канал для кодів виправлення помилок з непостійним часом. Згодом Ravi на ін. у роботі [7] узагальнили атаку на всі KEM на основі LWE/LWR у другому раунді процесу NIST, включаючи Kyber KEM, використовуючи бічний канал електромагнітної енергії.

Повне відновлення ключа можна виконати за $\approx 2k-4k$ запити для всіх наборів параметрів Kyber KEM. Однією з головних переваг є те, що атака може використовувати витік із усієї процедури повторного шифрування (рядок 19), і, отже, може працювати в налаштуваннях низького відношення «сигнал-шум» і недорогому обладнанні для атак бічними каналами [2].

2. Атаки бічними каналами шляхом збою розшифрування на основі Оракула (Оракул_ЗР)

Ця категорія атак працює шляхом виконання запитів до пристрою декапсуляції за допомогою обережно пошкоджених зашифрованих текстів, так що викликані через пошкодження збої розшифрування залежать від секретного ключа. Тому оракул бічного каналу, який здатний виявити помилки дешифрування, може відновити секретний ключ. Першу подібну атаку бічним каналом шляхом збою розшифрування на основі оракула було запропоновано у роботі Guo та ін. [8] на Frodo KEM, використовуючи непостійне за часом виконання операції порівняння зашифрованого тексту для виявлення помилок дешифрування. Згодом Bhasin та ін. [9] узагальнили атаку на Kyber KEM і продемонстрували, що витік потужності/електромагнітного випромінювання з блоку порівняння зашифрованого тексту може бути використаний для виявлення збоїв дешифрування для відновлення ключа.

Кілька робіт [10, 11] показують, що простого відновлення знака вихідного шуму $d[0]$, тобто $d[0]>0$ або $d[0]<0$ для кількох дійсних шифротекстів достатньо для того, щоб повністю відновити секретний ключ. Помилка дешифрування може бути легко ідентифікована за допомогою витоку бічними каналами з будь-якої операції в рамках процедури повторного шифрування, а також операції порівняння зашифрованого тексту.

3. Атаки бічними каналами шляхом повного розшифрування на основі Оракула (Оракул_ПР).

Щодо цього Xu та ін. у роботі [12] запропонували нову техніку побудови обраних зашифрованих текстів, які одночасно надають 256 біт інформації про приватний ключ. У роботі створюється $ct=(u,v)\in(R_q^k\times R_q)$ так, що

$$u_i=\begin{cases}U\cdot x^0 & \text{if } \ i=0\\ 0 & \text{if } \ 1\le i\le k\text{-}1\end{cases}, \tag{1}$$

$$v=V\cdot\left(\sum_{i=0}^{i=n-1}x^i\right), \tag{2}$$

де $(U,V)\in\square^+$. Зловмисник може вибрати кортежі (U,V) так, що розшифроване повідомлення матиме наступний вигляд:

$$m_i'=F(s_0[i]) \qquad \text{if } \ 1\le i\le n\text{-}1 \tag{3}$$

де кожен біт повідомлення m_i' залежить від відповідного секретного коефіцієнта $s_0[i]$. Крім того, зловмисник може вибрати (U,V) так, щоб кожен біт повідомлення m_i' однозначно ідентифікував відповідний секретний коефіцієнт $s_0[i]$. Таким чином зловмисник ефективно розпаралелив атаку Оракул_ПВТ. Оскільки для відновлення ключа потрібен доступ до повного розшифрованого повідомлення (тобто до оракула повного розшифрування), у роботі [13] пропонується використовувати витік з операції кодування повідомлення під час повторного зашифрування, який можна використовувати для відновлення 256 біт в одному сліді.

Таким чином, повне відновлення ключа можливе шляхом всього лише шести запитів для Kyber512. Подібним чином у роботах Ravi та ін. [14] і Ngo та ін. [15, 16] демонструється можливість використання витоку з операції декодування повідомлення у випадку з обраним шифротекстом для повного відновлення ключа приблизно в 6 – 20 слідах для таких схем, як Kyber і Sabre.

Показані атаки демонструють, що зловмисник може використовувати обрані шифротексти для отримання витоку з різних операцій у рамках процедури декапсуляції для відновлення ключа.

Націлювання на операцію NTT (Витік_з_NTT).

Хоча, для атак з відомим шифротекстом подібні атаки можуть витримувати лише шум зі стандартним відхиленням σ в діапазоні 0,5 – 0,7. Нещодавно Hamburg та ін. [17] про демонстрували, що чутливість цих атак до відношення «сигнал-шум» (SNR) може бути значно покращено у випадку з обраним шифротекстом.

Ідея полягала у створенні обраних шифротекстів, щоб подавати рідкісні вхідні дані $(\hat{u}' \circ \hat{s})$ до екземпляра INTT у процедурі дешифрування. Повідомляється, що це покращує ефективність алгоритму поширення переконань (Belief Propagation), дозволяючи більше шуму під час вимірювань, навіть при націленні на замасковані реалізації. Вони демонструють низку атак відновлення ключа зі складністю сліду в діапазоні від k до $2k$, де k є розміром модуля в Kyber KEM ($k = \{2,3,4\}$). Покращена атака може витримувати набагато більше шуму з $\sigma \leq 2.2$, демонструючи значне покращення атак NTT, якщо вони виконуються у випадку з обраним шифротекстом

1.3. Атаки відновлення повідомлення – відомий шифротекст

Подібні атаки можна розділити на дві категорії:

Націлювання на операції кодування та декодування повідомлень (Витік_з_Кодування_Декодування).

Атаки відновлення повідомлень переважно спрямовані на дві операції, які безпосередньо виконують дії над конфіденційним повідомленням m: операцію $Encode$ в зашифруванні та операцію $Decode$ в розшифруванні. І операції кодування, і операції декодування виконують дії над повідомленням по одному біту за раз, і це побітове маніпулювання конфіденційним повідомленням служить основним джерелом витоку для таких атак. Першу таку атаку продемонстрували Amiet та ін. [18], спрямована вона на операцію кодування повідомлень у NewHope KEM, KEM на основі Ring-LWE на мікроконтролері ARM Cortex-M4. Різницю у вазі Хеммінга коефіцієнтів полінома повідомлення $m[i] = [q/2]$ для $m_i = 1$ та $m[i] = 0$ для $m_i = 0$ можна легко розрізнити за допомогою атаки бічним каналом, що дозволяє повністю відновити окремі біти повідомлення в одному сліді. Згодом Sim та ін. [9] узагальнили техніку атаки, щоб націлити її на всі KEM на основі решітки, що беруть участь у процесі стандартизації NIST, включаючи Kyber KEM.

Пізніше у роботі [8] було представлено нові атаки, які використовують витік з операції декодування повідомлення в процедурі розшифрування для відновлення повідомлення. Незважаючи на те, що в роботі було продемонстровано наявність витоку з окремих бітів повідомлення, вдалося отримати лише 81 % успіху для відновлення одного байта повідомлення Kyber, тоді як обладнання з більш високим SNR потенційно могло б виконати ідеальне відновлення повідомлення з одного сліду.

Націлювання на операцію NTT (Витік_з_NTT).

У роботі [19] демонструється, що операція NTT також може бути ціллю для атаки відновлення повідомлень. Ідея полягала у тому, щоб відновити вхідні дані для екземпляру NTT через ефемерний секрет r, знання якого можна використати для відновлення повідомлення m із шифротексту ct. У роботі [19] також пропонується значне покращення оригінальної атаки з роботи [3] шляхом зменшення кількості шаблонів з одного мільйона до лише 213 шаблонів, а також надання кількох вдосконалень, таких як використання вдосконаленого алгоритму поширення переконань (Belief Propagation) для відновлення повідомлень. Було також показано, що ця атака може бути застосована навіть при використанні маскувальних контрзаходів, хоча й за наявності високого відношення «сигнал-шум».

1.4. Атаки відновлення повідомлення – обраний шифротекст

Націлювання на захищені операції кодування та декодування повідомлень (Витік_з_захищеного_Кодування_Декодування)

Атаки відновлення повідомлень, націлені на операції кодування та декодування повідомлень, здатні відновити все повідомлення в одному сліді у випадку атак з відомим шифротекстом. Однак подібним атакам можна легко запобігти за допомогою простих контрзаходів перетасування, які рандомізують порядок кодування/декодування окремих бітів повідомлення. Хоча перетасування не усуває джерело витоку бічним каналом, воно не дає зловмиснику можливості відновити правильний порядок бітів повідомлення, тим самим перешкоджаючи відновленню повідомлення з одного сліду.

Однак у роботі Ravi та ін. [14] показано, що зловмисник може порушити контрзахід перетасування у випадку з обраним шифротекстом, використовуючи властивість пластичності зашифрованого тексту схем на основі LWE/LWR. Враховуючи цільовий шифротекст $ct = (u, v)$, зловмисник спочатку піддає ct процедурі декапсуляції, щоб відновити окремі біти повідомлення m' через бічні канали, а потім обчислює його вагу Хеммінга (HW). Після цього зловмисник подає пошкоджений (спотворений) зашифрований текст $ct^* = (u, v + q/2 \cdot x^0)$, тобто $q/2$, доданий до першого коефіцієнта v. Це має ефект інвертування першого біта повідомлення m_0', що призводить до спотвореного повідомлення m''. Різниця в HW m' і m'' (збільшення або зменшення) може бути використана для відновлення значення інвертованого біта повідомлення m_0. Таким чином, можливе повне відновлення повідомлення за 257 запитів для Kyber KEM.

2. Атаки помилками на Kyber KEM

У даному пункті представлено короткий огляд атак із впровадженням помилок на KEM на основі структурованої решітки, з основним акцентом на атаках, які застосовуються до Kyber [2].

2.1. Атаки відновлення ключа і відновлення повідомлення – відомий шифротекст

Повторне використання nonce (Повторне_використанн_Nonce).

Ця категорія охоплює атаки на процедури генерації ключів та інкапсуляції, де зловмисник може спостерігати лише помилкові виходи від цілі. У роботі Ravi та ін. [20] запропоновано першу практичну атаку помилками, що застосована до KEM на основі решітки, таких як Kyber, NewHope і Frodo. Запропонована атака випливає зі спостереження, що початкове число, яке використовується для вибірки секрету, і помилки (похибки) для екземплярів LWE відрізняються лише на один байт, тобто s і e обираються з того самого початкового числа $seed_B$, але з різними nonce (випадковими числами) $coins_s$ і $coins_e$, які, в свою чергу, відрізняються одним байтом. Те ж саме стосується процедури зашифрування.

Таким чином, зловмисник може використовувати помилки для примусового повторного використання nonce, тобто так, що $coins_s = coins_e$, щоб створити екземпляри LWE у формі $t = A \cdot s + s = A \cdot (s + 1)$, які можна тривіально розв'язати за допомогою елімінації Гауса. Автори продемонстрували практичність повторного використання nonce за допомогою введення (ін'єкції) електромагнітного збою (EMFI) на мікроконтролері ARM Cortex-M4. У той час, як атака призводить до повного відновлення ключа та відновлення повідомлення у випадку атаки «Людина посередині» (MITM), вона вимагає введення кількох цільових помилок до процедур генерації ключів та інкапсуляції для практичних атак.

У роботі [21] Valencia та ін. провели більш загальне дослідження сприйнятливості CPA-захищених схем на основі LWE/LWR до атак помилками. Вони пропонують різні атаки, націлені на кілька операцій у рамках процедур генерації ключів, шифрування та дешифрування. Однак ці атаки розглядають моделі помилок, такі як обнулення цілих поліномів, чого важко досягти на практиці.

2.2. Атаки відновлення ключа – обраний шифротекст

Однією з головних проблем, пов'язаних із націлюванням на процедуру декапсуляції шляхом помилок, є те, що вона містить внутрішній захист від помилок, тобто FO-перетворення (Fujisaki-Okamoto transformation) для виявлення недійсних зашифрованих текстів із дуже високою ймовірністю. Це представляє значну проблему для виконання атак помилками.

Націлювання на перевірку рівності шифротексту (Пропуск_Перевірки_Шифротексту).

Однією з очевидних цілей в процесі виконання процедури декапсуляції є пропуск перевірки рівності зашифрованого тексту шляхом помилок. У роботі [22] було показано, що безпеку відносно CCA кількох схем, включаючи Kyber, можна легко зламати через одну цільову помилку. Аналіз реалізації операції перевірки рівності зашифрованого тексту в програмній реалізації Kyber KEM з бібліотеки pqm4 [23] показує наступне: масив T містить конфіденційний сумісно використовуваний секрет $\bar{K}'$, отриманий із розшифрованого повідомлення m' після розшифрування. Якщо порівняння зашифрованого тексту не вдається (недійсний/зловмисний зашифрований текст), псевдовипадкове значення z записується в T за допомогою операції умовного переміщення. Згодом T використовується для отримання остаточного спільного секрету K.

Таким чином, процедура декапсуляції записує конфіденційний сумісно використовуваний секрет до T (за умови успішної декапсуляції), перш ніж перевірити дійсність зашифрованого тексту. Таким чином, простий пропуск наступної операції умовного переміщення гарантує використання конфіденційного сумісно використовуваного секрету $\bar{K}'$ для генерування спільного секрету K, навіть якщо операція порівняння зашифрованого тексту дала збій. У роботі [22] показано, що така вразливість може бути використана через прості збої у частоті і згодом може призвести до відновлення ключа за рахунок кількох тисяч запитів із застосуванням атаки обраного зашифрованого тексту.

Атака помилками з обраним шифротекстом.

За винятком перевірки рівності зашифрованого тексту, у процедурі декапсуляції немає інших тривіальних цілей для помилок. Проте Pessl та Prokop у роботі [24] запропонували першу загальну атаку помилками з обраним шифротекстом, яка працює шляхом введення цільових помилок в операцію декодування повідомлення в рамках процедури дешифрування, так що результуючий успіх/невдача декапсуляції може бути використаний для отримання важливої інформації про приватний ключ.

Основна ідея атаки полягає в наступному. Зловмисник надсилає дійсний зашифрований текст ct для декапсуляції та вводить одну помилку, щоб пропустити додавання з $q/2$ під час декодування одного коефіцієнта полінома повідомлення $m'[i]$. Це має непрямий ефект спотворення $m'[i]$ приблизно на $q/4$. Це призводить до інверсії m'_i (збій декапсуляції) лише тоді, коли відповідний коефіцієнт шумового компонента $d[i] < 0$. Однак, коли $d[i] \geq 0$ у m'_i немає жодних змін. Це допомагає зловмиснику створити єдину лінійну нерівність, використовуючи $d[i]$, а зловмисник, який здатний побудувати $5k - 7k$ таких лінійних нерівностей, може виконати повне відновлення ключа для Kyber KEM.

Хоча атаку було продемонстровано за допомогою збою частоти на мікроконтролері ARM Cortex-M4, для атаки все ще потрібно вводити цільову помилку пропуску в процедуру декодування повідомлення. Таким чином, цій атаці можна запобігти, просто застосувавши перемішування для операції декодування повідомлення.

Згодом Hermelink та ін. у роботі [11] запропонували покращення атаки з роботи [24], застосувавши дещо інший підхід. Замість того, щоб використовувати дійсний зашифрований текст ct, автори пропонують подати спотворені зашифровані тексти так, що один коефіцієнт другого компонента зашифрованого тексту $v[i]$ спотворюється на $q/4$. Під час подання спотвореного зашифрованого тексту після дешифрування вводиться помилка, щоб виправити однобітове спотворення в зашифрованому тексті, що зберігається в пам'яті. Якщо внесене спотворення призвело до правильного дешифрування ($d[i] \geq 0$), то введена помилка виправ-

ляє спотворення в зашифрованому тексті, забезпечуючи успішну декапсуляцію. Однак, якщо початкове спотворення призвело до помилки дешифрування ($d[i] < 0$), то це призводить до помилки декапсуляції, навіть після виправлення спотворення в збереженому шифротексті шляхом введення помилки. Ця інформація про d, отримана приблизно за $5k-7k$ таких запитів, може відновити повний приватний ключ.

На відміну від атаки з [24], атака з [11] не має часових обмежень для ін'єкції (введення) помилки, оскільки їй потрібно лише ввести помилку інверсії біта в пам'яті в будь-який час між операцією дешифрування та порівняння зашифрованого тексту. Проте введення точних одиночних помилок інверсії бітів у пам'яті потребує детальної інформації про цільовий пристрій, а також про реалізацію та розширене профілювання цільового пристрою. Нещодавно Del-vaux [25] покращив атаку [11] шляхом розширення поверхні атаки до кількох операцій у рамках процедури декапсуляції, а також роботи з різноманітними більш вільними моделями помилок, такими як довільна інверсія бітів, помилки встановлення у 0, випадкові помилки та помилки пропуску інструкцій. Однак атаки, що покладаються на більш вільну модель помилки, можуть вимагати понад $100k$ запитів з обраним зашифрованим текстом для повного відновлення ключа, залежно від практичності моделі помилки.

На рис. 1, 2 наведено всі SCA та FIA відповідно, що застосовуються для Kyber KEM [2]:

Рис. 1. Атаки бічними каналами (SCA) на Kyber KEM

Рис. 2. Атаки помилками (FIA) на Kyber KEM

3. Захист Kyber KEM від атак бічними каналами та помилками

У попередньому розділі було представлено детальний огляд різних SCA та FIA на Kyber KEM. У цьому розділі будуть надані ряд контрзаходів, які можна використовувати для захисту від згаданих атак [2].

3.1. Захист від атак бічними каналами та помилками з обраним шифротекстом

Підхід ґрунтується на стратегії виявлення, щоб перевірити/визначити, чи є отриманий зашифрований текст шкідливим. У разі виявлення шкідливого/зловмисного тексту ціль може просто відхилити зашифрований текст і змінити/оновити пару відкритий-приватний ключ, повторно запустивши процедуру генерації ключа. Перевагою цього підходу є те, що таке виявлення запобігає подальшому розкриттю приватного ключа.

Перевірка нормальності шифротексту.

Детальний аналіз шифротекстів, які використовуються в Оракул_ПВТ і Оракул_ПР, показує, що більшість коефіцієнтів зашифрованого тексту мають фіксоване значення 0. Однак коефіцієнти дійсного зашифрованого тексту рівномірно розподілені в діапазоні $[0, q]$, враховуючи, що обидва компоненти зашифрованого тексту є по суті екземплярами LWE. Тому пропозицією є проведення статистичного аналізу зашифрованого тексту перед використанням у процедурі дешифрування.

У якості статистичних даних для виявлення перекосу в отриманому зашифрованому тексті було вирішено обрати середнє значення та стандартне відхилення коефіцієнтів зашифрованого тексту. Для заданого полінома $x \in R_q$ позначаємо середнє (μ) і стандартне відхилення (σ) коефіцієнтів x як $\mu(x)$ і $\sigma(x)$ відповідно. Було виконане емпіричне моделювання, щоб обчислити середнє значення та стандартне відхилення $\mu(u)$ і $\sigma(u)$ для окремих поліномів компонента зашифрованого тексту u, а також $\mu(v)$ і $\sigma(v)$ для компонента зашифрованого тексту v, що відповідає дійсним шифротекстам для Kyber KEM. Отримані значення для середнього та стандартного відхилення всіх чотирьох статистичних показників:

$$(\mu(\mu(u)), \sigma(\mu(u))) = (1663, 60) \quad (4)$$

$$(\mu(\mu(u)), \sigma(\mu(u))) = (959, 27)$$

$$(\mu(\mu(v)), \sigma(\mu(v))) = (1560, 60) \quad (5)$$

$$(\mu(\mu(v)), \sigma(\mu(v))) = (957, 27)$$

На основі стандартного відхилення σ для кожного з цих показників можна вибрати прийнятний діапазон для них. Наприклад, якщо вибрано довжину $6 \cdot \sigma$, то прийнятний діапазон для $\mu(u)$ становить $[\mu(\mu(u)) + 6 \cdot \sigma, \mu(\mu(u)) - 6 \cdot \sigma]$. Чим менший допустимий діапазон, тим вища ймовірність хибних спрацьовувань, тобто визначення дійсного зашифрованого тексту недійсним. Однак великий допустимий діапазон збільшує ймовірність пропусків, що призводить до прийняття спотвореного зловмисного зашифрованого тексту як дійсного.

За допомогою емпіричного моделювання було визначено, що довжина 6σ як для середнього, так і для стандартного відхилення призводить до ймовірності $\approx 2^{-22}$ для відхилення правильного зашифрованого тексту.

Хоча такий контрзахід здатний виявляти перекошені зашифровані тексти, обрані зашифровані тексти, що використовуються в Оракул_ЗР [6], а також ті, що використовуються в атаках помилками з обраним шифротекстом [11, 24, 25], є рівномірно випадковими. Це пояснюється тим, що подібні атаки передбачають додавання невеликих помилок до одного коефіцієнта дійсного зашифрованого тексту. Така дія не вносить помітного перекосу в коефіцієнти, тим самим перешкоджаючи контрзаходу Перевірки_нормальності_шифротектсу.

Перевірка нормальності поліному повідомлення.

Цей контрзахід ґрунтується на аналізі коефіцієнтів полінома зашумленого повідомлення $m' = (v' - u' \cdot s)$, отриманого під час дешифрування отриманого зашифрованого тексту ct. Для дійсних зашифрованих текстів ми спостерігаємо, що коефіцієнти m' розподіляються відповідно до дуже вузького розподілу Гауса поблизу $q/2$ або 0, тобто $m[i] = q/2 \pm \delta$ для $m_i = 1$ і $m[i] = 0 \pm \delta$ для $m_i = 1$.

У той час як атака помилками Pessl та ін. [26] додає $q/4$ до $m'[i]$ через помилки, атаки [9, 11, 25] явно додають $q/4$ до цільового коефіцієнта дійсного зашифрованого тексту. Таким чином, усі ці атаки працюють шляхом прямого/опосередкованого додавання $q/4$ до одного з цільових коефіцієнтів полінома повідомлення $m'[i]$. Це гарантує, що принаймні один поліном повідомлення, тобто $m'[i]$ не буде в межах очікуваного діапазону, що відповідає діапазону дійсного зашифрованого тексту

На основі цього спостереження можна запропонувати перевірити розподіл коефіцієнтів полінома повідомлення для отриманого зашифрованого тексту. Нехай допустимий діапазон буде $(q/2 \pm L \cdot \sigma)$ і $(0 \pm L \cdot \sigma)$, де $L \in \mathbb{Z}^+$ залишається на вибір розробника. Чим більший діапазон прийняття $L \cdot \sigma$, тим меншою є ймовірність відкинути дійсний зашифрований текст. Однак вибір меншого діапазону підвищує шанси невиявлення зловмисного обраного зашифрованого тексту. Тому для покращення безпеки важливо вибрати правильне значення для L. При $L = 6$ після більш ніж 2^{25} дійсних декапсуляцій не було виявлено хибнопозитивного результату [2].

Якщо є принаймні один коефіцієнт за межами цього допустимого діапазону, ми просто позначаємо зашифрований текст як дійсний і оновлюємо пару відкритий-приватний ключ. Хоча цей контрзахід вимагає розшифрувати принаймні один обраний шифртекст для успішного виявлення, однак CCA вимагають принаймні від кількох десятків до кількох тисяч запитів для відновлення ключа.

3.2. Захист операції порівняння шифротексту (Захист_Порівняння_Шифротексту)

У даному пункті пропонуються два рівні захисту для операції порівняння зашифрованого тексту, на яку спрямована атака Пропуск_Перевірки_Шифротексту Xagawa та ін. [22]. На першому рівні додається захист від пропуску операції порівняння зашифрованого тексту, для чого використовується лічильник динамічного циклу, щоб відстежувати кількість порівнюваних байтів зашифрованого тексту в операції порівняння зашифрованого тексту. Якщо кількість байтів зашифрованого тексту для порівняння дорівнює d, тоді лічильник циклу l ініціалізується випадковим значенням $k \cdot d$, де $k \in \mathbb{Z}^+$ вибирається випадковим чином під час кожного виконання. Потім l зменшується на k для кожного байта зашифрованого тексту, що порівнюється. Таким чином, якщо $l = 0$ після порівняння зашифрованого тексту, то можна впевнитись, що всі байти були порівняні. Використання такого динамічного лічильника циклу додає додатковий рівень захисту від атак помилками вищого порядку, які також намагаються збити лічильник циклу.

Щоб захиститись від пропуску операції умовного переміщення, слід записувати спільно використовуваний секрет $\bar{K}'$ в тимчасову змінну tmp замість T. $\bar{K}'$ копіюється в T, якщо порівняння зашифрованого тексту вдалось ($l = 0$), у іншому випадку z копіюється в T. Перевірка того, чи $l = 0$ виконується для кожного байта, скопійованого в T. Це просте виправлення реалізації гарантує, що пропуск операції умовного переміщення не розкриє жодної інформації про спільно використовуваний секрет $\bar{K}'$ для недійсних зашифрованих текстів, тим самим перешкоджаючи відновленню ключа.

3.3. Захист операцій кодування/декодування повідомлення (Перемішування_Кодування_Декодування)

SCA, що націлені на процедури кодування (Encode) і декодування (Decode) повідомлення, є дуже потужними, враховуючи, що зловмисник може виконати відновлення повідомлення в одному сліді. Крім того, витік від цих операцій також можна використовувати як Оракул_ПР для відновлення ключа [12]. Ravi та ін. у роботі [14] показали, що контрзаходи

перетасування для операцій кодування та декодування повідомлень можуть бути зламані у випадку статичного ключа через CCA за кількість від кількох сотень до кількох тисяч запитів. Хоча перетасування не забезпечує конкретного захисту, воно обґрунтовано збільшує зусилля зловмисника щодо відновлення повідомлення, як показано в [16]. Крім того, у ефемерних умовах перетасування забезпечує конкретний захист, оскільки всі вищезгадані атаки вимагають від кількох сотень до кількох тисяч запитів для відновлення повідомлення [14]. Таким чином, реалізація контрзаходу перетасування для операцій кодування та декодування Kyber KEM не буде зайвою.

3.4. Захист NTT (Перемішування_NTT, Маскування_NTT)

SCA, націлені на NTT, також здатні виконувати відновлення ключа та повідомлення в одному сліді [3, 19] або дуже невеликій кількості слідів [17] (атаки Витік_з_NTT). У роботі [26] запропоновано низку загальних контрзаходів перетасування та маскування з різним ступенем деталізації для захисту NTT від вищезгаданих атак із одним слідом. Ґрунтуючись на передбачуваному рівні загрози від потенційного зловмисника та прийнятній продуктивності, розробник може вибрати відповідний контрзахід перетасування для операції NTT.

3.5. Захист вибірки секретів і помилок (похибок) (Збиткове_Порівняння)

У роботі [20] продемонстровано, що повторне використання nonce може бути викликане помилками в процедурі генерації ключа та шифрування Kyber KEM для атак відновлення ключа та відновлення повідомлення відповідно. Таким чином, тривіальним захистом від цієї атаки може бути виконання надлишкового обчислення процедури вибірки. Хоча це не забезпечує повного захисту, це значно підвищує складність для зловмисника.

4. Експериментальна оцінка

У цьому пункті показано практичну оцінку ефективності представлених контрзаходів при інтеграції в оптимізовану реалізацію програмного забезпечення Kyber, що працює на мікроконтролері STM32F4 на основі процесора ARM Cortex-M4.

4.1. Цільова платформа та деталі реалізації

Цільовою платформою для процесора ARM Cortex-M4 є плата STM32F4DISCOVERY, на якій розміщено мікроконтролер STM32F407, а тактова частота становить 24 МГц. Контрзаходи були реалізовані на оптимізованій для M4 реалізації Kyber, доступній у загальнодоступній бібліотеці pqm4 [23] – середовищі порівняльного аналізу схем PQC на мікроконтролері ARM Cortex-M4. Оптимізована для M4 реалізація Kyber базується на високошвидкісній реалізації з ефективним використанням пам’яті, запропонованою Ботросом, Каннвішером і Швабе в [27]. Усі реалізації було скомпільовано за допомогою компілятора arm-none-eabi-gcc-7.3.1 з використанням прапорів компілятора -O3 -mthumb -mcpu=cortex-m4 -mfloat-abi=hard -mfpu=fpv4-sp-d16

4.2. Результати експериментів

У табл. 1 представлено дані про рівень витрат продуктивності через заходи протидії перемішування та маскування (Перемішування_NTT, Маскування_NTT) проти атак витоку з NTT (Витік_з_NTT) на Kyber KEM, який працює на пристрої ARM Cortex-M4. Подібні експерименти надані у роботі [2], головною метою було відтворення та перевірка отриманих у роботі результатів. Н/з – незахищена реалізація, З – реалізація з контрзаходами, %В – рівень затрат продуктивності у відсотках.

Таблиця 1

Продуктивність контрзаходів Shuffled_Masked_NTT для Kyber KEM порівняно з оптимізованими незахищеними реалізаціями на пристрої ARM Cortex-M4. Дані отримано на мікроконтролері STM32F407VG, встановленому на платі STM32F407DISCOVERY, що працює на частоті 24 МГц. Числа позначають кількість тактових циклів $\times 10^3$

Схема	Такти ($\times 10^3$)								
	KeyGen			Encaps			Decaps		
	Н/з	З	%В	Н/з	З	%В	Н/з	З	%В
Контрзахід перемішування									
Kyber512	457.3	782.4	71.1	552.1	970.6	75.8	511.3	1022.6	100.0
Kyber768	748.8	1237.7	65.3	903.6	1483.7	64.2	842.5	1506.4	78.8
Kyber1024	1188.2	1841.5	54.9	1378.5	2125.6	54.2	1297.5	2125.9	63.9
Контрзахід маскування									
Kyber512	457.3	729.4	59.5	552.1	898.2	62.7	511.3	933.6	82.6
Kyber768	748.8	1155.4	54.3	903.6	1386.1	53.4	842.5	1399.4	66.1
Kyber1024	1188.2	1732.4	45.8	1378.5	1998.8	45.0	1297.5	1993.0	53.6

На пристрої ARM Cortex-M4 можна спостерігати вплив на продуктивність у діапазоні 46 – 71 % для генерації ключів, 45 – 76 % для інкапсуляції та 53 – 100 % для процедури декапсуляції для всіх наборів параметрів Kyber KEM.

У табл. 2 наведено дані про рівень витрат продуктивності через контрзаходи Перевірка_нормальності_шифротексту, Перевірка_нормальності_поліному_повідомлення, Захист_Порівняння_Шифротексту і Перемішування_Кодування_Декодування для Kyber KEM, який працює на пристрої ARM Cortex-M4 [2]. Н/з – незахищена реалізація, З – реалізація з контрзаходами, %В – рівень затрат продуктивності у відсотках.

Таблиця 2

Продуктивність спеціальних контрзаходів SCA-FIA для Kyber KEM порівняно з оптимізованою незахищеною реалізацією на пристрої ARM Cortex-M4. Дані отримано на мікроконтролері STM32F407VG, встановленому на платі STM32F407DISCOVERY, що працює на частоті 24 МГц. Числа позначають кількість тактових циклів $\times 10^3$

Схема	Такти ($\times 10^3$)		
	Decaps		
	Н/з	З	%В
Перевірка нормальності шифротексту			
Kyber512	511.4	689.4	34.8
Kyber768	842.4	1029.4	22.2
Kyber1024	1298.1	1492.8	15.0
Перевірка нормальності поліному повідомлення			
Kyber512	511.4	676.0	32.2
Kyber768	842.4	1005.8	19.4
Kyber1024	1298.1	1474.7	13.6
Захист операції порівняння шифротексту			
Kyber512	511.4	579.4	13.3
Kyber768	842.4	909.8	8.0
Kyber1024	1298.1	1364.3	5.2
Захист операцій кодування/декодування повідомлення			
Kyber512	511.4	520.1	1.7
Kyber768	842.4	854.2	1.4
Kyber1024	1298.1	1312.4	1.1

Як видно з табл. 2, на пристрої ARM Cortex-M4 ці контрзаходи створюють дуже розумний рівень витрат продуктивності в діапазоні 15 – 35 %, 13 – 32 %, 5 – 13 % і 1 – 2 % для різних наборів параметрів Kyber KEM.

Отримані дані показано у вигляді діаграми на рис. 3 – 8.

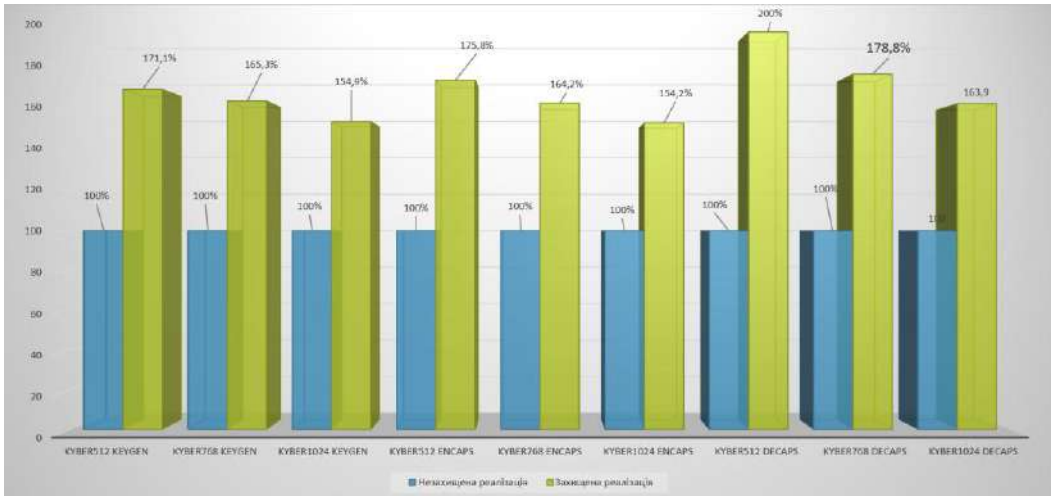


Рис. 3. Оцінка контрзаходу перемішування

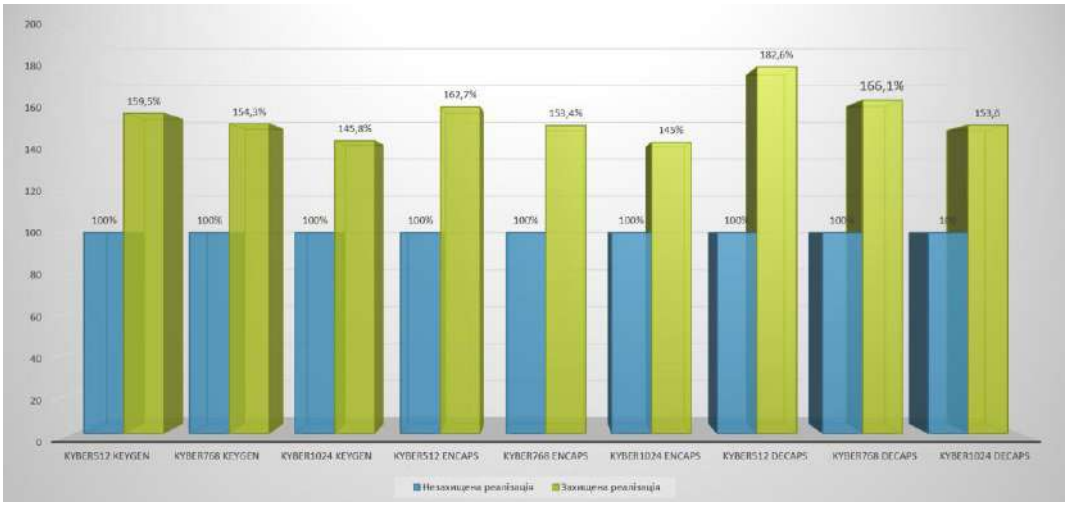


Рис. 4. Оцінка контрзаходу маскування

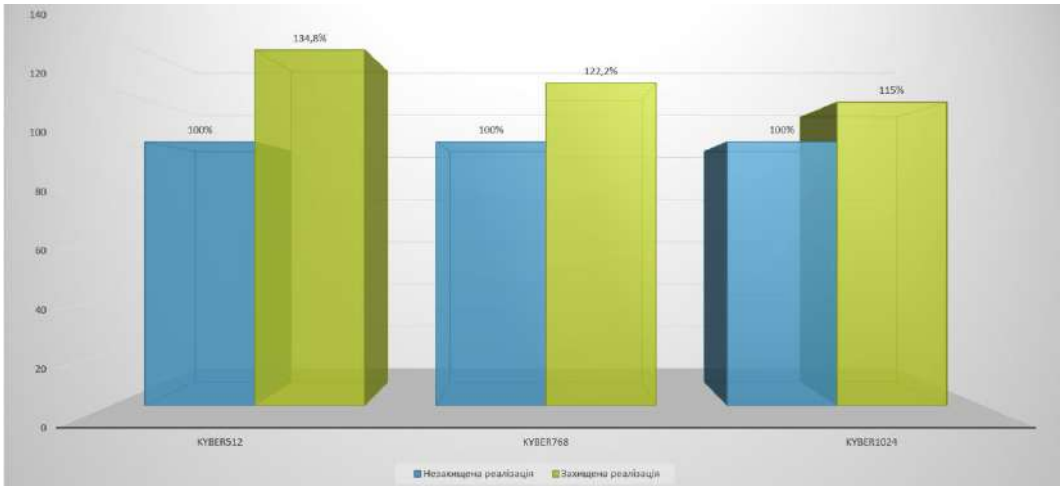


Рис. 5. Оцінка контрзаходу перевірки нормальності шифротексту

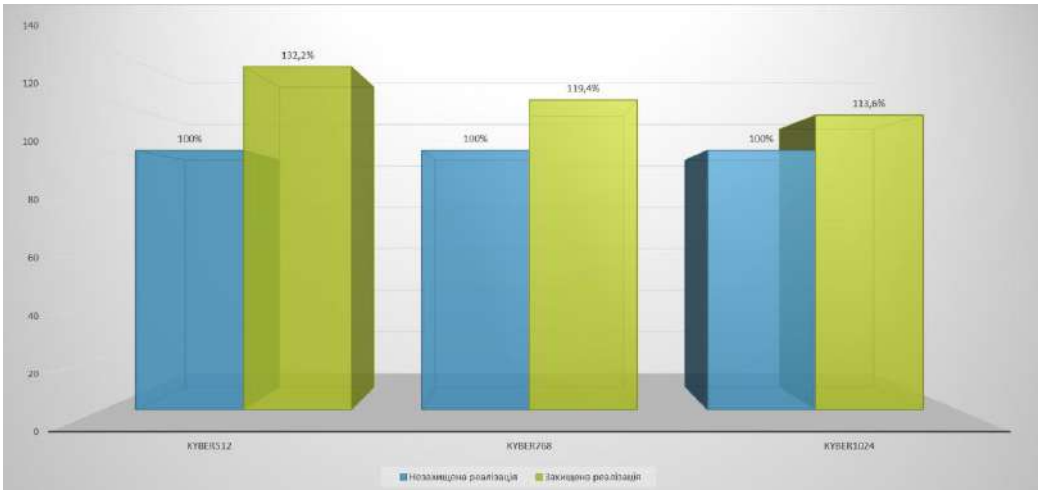


Рис. 6. Оцінка контрзаходу перевірки нормальності поліному повідомлення

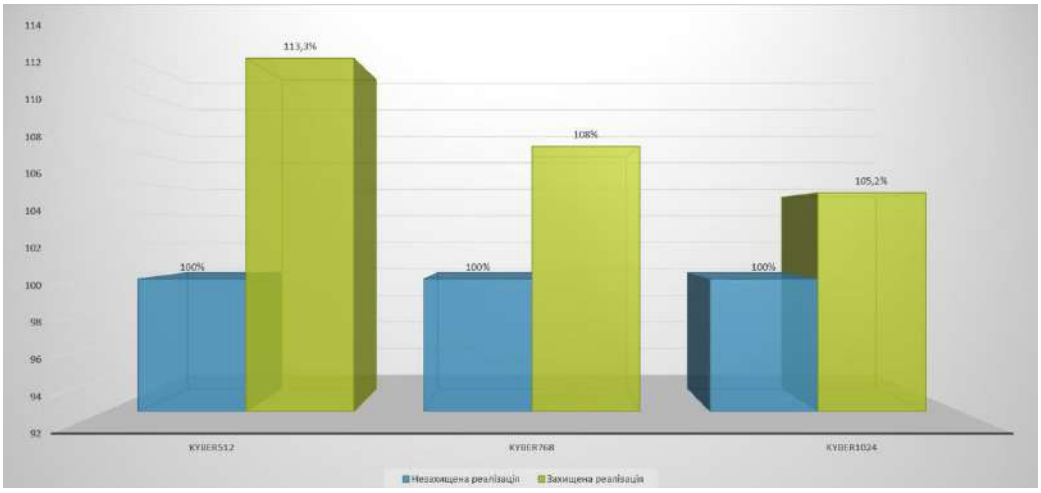


Рис. 7. Оцінка контрзаходу захисту операції порівняння шифротексту

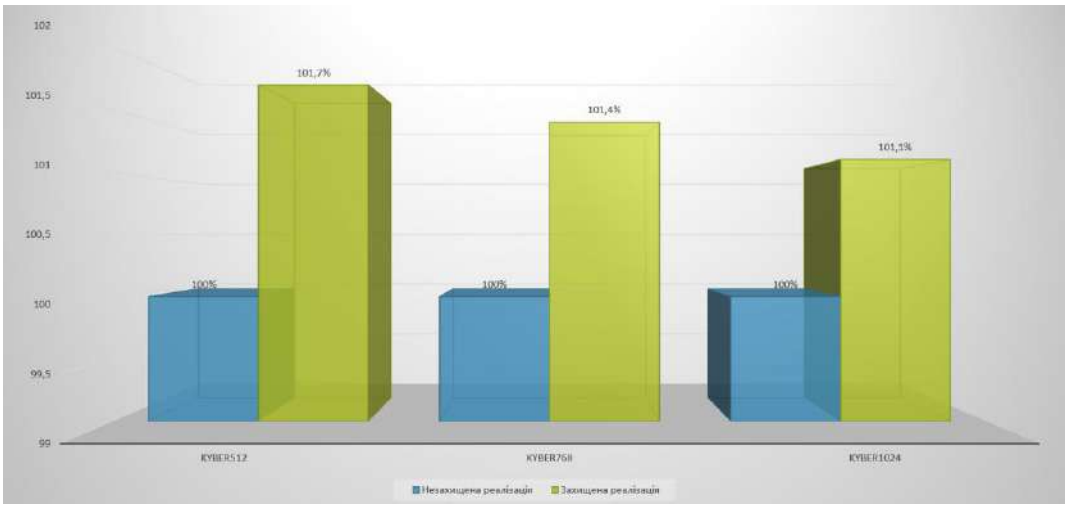


Рис. 8. Оцінка контрзаходу захисту операцій кодування/декодування повідомлення

Зауважимо, що оптимізовані реалізації NTT на цільових пристроях реалізовано на чистому ассемблері (оптимізовано для M4 та оптимізовано для NEON), але контрзаходи реалізовано на основі реалізацій NTT/INTT на основі C, що призводить до високих витрат продук-

тивності. Таким чином, можна отримати значно менші витрати продуктивності за умови, що захищені NTT/INTT реалізовані на ассемблері.

5. Порівняння рівнів захищеності CRYSTALS-KYBER та СКЕЛЯ (ДСТУ 8961-2019)

У табл. 3 надана оцінка рівнів захищеності алгоритмів CRYSTALS-KYBER [1] та СКЕЛЯ (ДСТУ 8961-2019) [28] за шкалою NIST та за Національною шкалою.

Як видно з даних у таблиці СКЕЛЯ (ДСТУ 8961-2019) [28] надає значно вищі рівні безпеки для кожного з режимів роботи, порівняно з алгоритмом Crystal-Kyber. Також однією зі значних переваг алгоритму Скеля можна вважати забезпечення захисту від спеціальних атак.

Таблиця 3

Порівняння рівнів захищеності CRYSTALS-KYBER та СКЕЛЯ

Алгоритм	Рівень захищеності NIST	Національний рівень доказової стійкості (класична / квантова)	Забезпечення захисту від спеціальних атак
KYBER512	1	-	-
KYBER768	3	0-й рівень – 128біт / 64біт	-
KYBER1024	5	1-й рівень – 256біт / 128 біт	-
СКЕЛЯ–КЕМ 256/128	5	1-й рівень – 256біт / 128 біт	+
СКЕЛЯ–КЕМ 384/192	7	2-й рівень – 384біт / 192біт	+
СКЕЛЯ–КЕМ 512/256	9	3-й рівень – 512біт / 256біт	+

На рис. 9 показано дані з табл. 3 у вигляді діаграми.

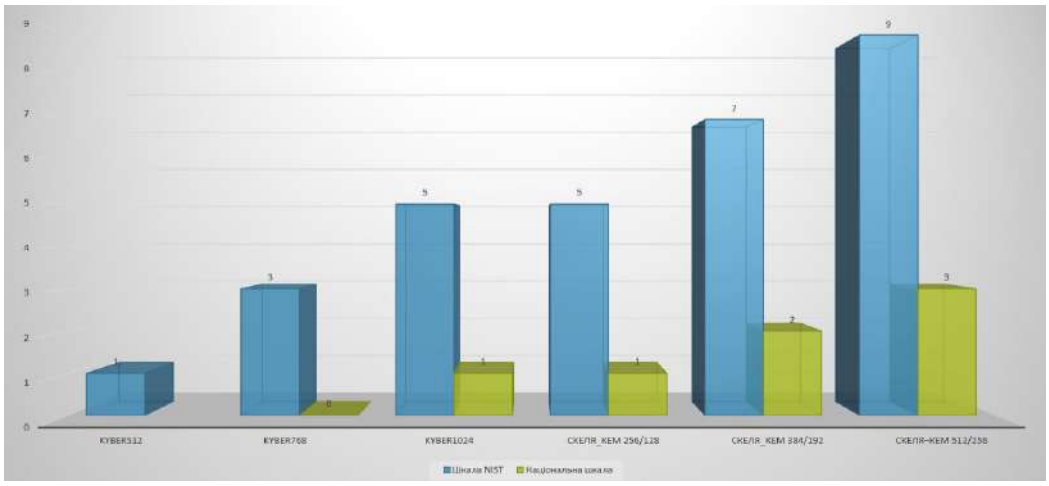


Рис. 9. Рівні безпеки алгоритмів за шкалою NIST та Національною шкалою

Висновки

Представлено систематичне дослідження атак бічними каналами (SCA) і атак помилками (із впровадженням помилок) (FIA) на структуровані схеми на основі решітки, з основним акцентом на Kyber.

Враховуючи невідповідність загальних контрзаходів, таких як маскування, для захисту від широкого спектру відомих атак, у даному пункті також представлено низку спеціальних контрзаходів для захисту від відомих SCA та FIA на Kyber.

Оцінка продуктивності показує, що представлені спеціальні контрзаходи мають розумний рівень витрат продуктивності для оцінюваної платформи. Тому, можна сказати, що у роботі наведено аргументи на користь використання спеціальних контрзаходів у реальних

реалізаціях схем на основі решітки, або окремо, або як підсилення загальних контрзаходів, таких як маскування.

Також у роботі надано оцінку рівнів захищеності алгоритмів CRYSTALS-KYBER [1] та СКЕЛЯ (ДСТУ 8961-2019) [28] за шкалою NIST та за Національною шкалою.

Як видно з даних, отриманих під час порівняння, СКЕЛЯ (ДСТУ 8961-2019) [28] надає значно вищі рівні безпеки для кожного з режимів роботи, порівняно з алгоритмом Crystal-Kyber. Також однією зі значних переваг алгоритму Скеля можна вважати забезпечення захисту від спеціальних атак.

Список літератури:

1. Roberto Avanzi, Joppe Bos, Léo Ducas, Eike Kiltz, Tancrède Lepoint, Vadim Lyubashevsky, John M. Schanck, Peter Schwabe, Gregor Seiler, Damien Stehlé. CRYSTALS-Kyber Algorithm Specifications And Supporting Documentation (version 3.02). URL: https://pq-crystals.org/kyber/data/kyber-specification-round3-20210804.pdf.

2. PRASANNA RAVI, ANUPAM CHATTOPADHYAY, ANUBHAB BAKSI. Side-channel and Fault-injection attacks over Lattice-based Post-quantum Schemes (Kyber, Dilithium): Survey and New Results. 2022. URL: https://eprint.iacr.org/2022/737.pdf.

3. Robert Primas, Peter Pessl, and Stefan Mangard. Single-Trace Side-Channel Attacks on Masked Lattice-Based Encryption. 2017. URL: https://eprint.iacr.org/2017/594.pdf.

4. Nicolas Veyrat-Charvillon, Benoît Gérard, and François-Xavier Standaert. Soft Analytical Side-Channel Attacks. 2014. URL: https://eprint.iacr.org/2014/410.pdf.

5. Judea Pearl. Fusion, propagation, and structuring in belief networks. 1986. URL: https://ftp.cs.ucla.edu/pub/stat_ser/r42-reprint.pdf.

6. Jan-Pieter D'Anvers, Marcel Tiepelt, Frederik Vercauteren, and Ingrid Verbauwhede. Timing attacks on Error Correcting Codes in Post-Quantum Schemes. 2019. URL: https://eprint.iacr.org/2019/292.pdf.

7. Prasanna Ravi, Sujoy Sinha Roy, Anupam Chattopadhyay, and Shivam Bhasin. Generic Side-channel attacks on CCA-secure lattice-based PKE and KEM schemes. 2019. URL: https://eprint.iacr.org/2019/948.pdf.

8. Qian Guo, Thomas Johansson, and Alexander Nilsson. A key-recovery timing attack on post-quantum primitives using the Fujisaki-Okamoto transformation and its application on FrodoKEM. 2020. URL: https://eprint.iacr.org/2020/743.pdf.

9. Shivam Bhasin, Jan-Pieter D'Anvers, Daniel Heinz, Thomas Pöppelmann, and Michiel Van Beirendonck. Attacking and Defending Masked Polynomial Comparison for Lattice-Based Cryptography. 2021. URL: https://eprint.iacr.org/2021/104.pdf.

10. Jan-Pieter D'Anvers, Daniel Heinz, Peter Pessl, Michiel van Beirendonck, and Ingrid Verbauwhede. Higher-Order Masked Ciphertext Comparison for Lattice-Based Cryptography. 2021. URL: https://eprint.iacr.org/2021/1422.pdf.

11. Julius Hermelink, Peter Pessl, and Thomas Pöppelmann. Fault-enabled chosen-ciphertext attacks on Kyber. 2021. URL: https://eprint.iacr.org/2021/1222.pdf.

12. Zhuang Xu, Owen Pemberton, Sujoy Sinha Roy, David Oswald, Wang Yao, and Zhiming Zheng. Magnifying Side-Channel Leakage of Lattice-Based Cryptosystems with Chosen Ciphertexts: The Case Study of Kyber. 2020. URL: https://eprint.iacr.org/2020/912.pdf.

13. Zhuang Xu et al. Magnifying Side-Channel Leakage of Lattice-Based Cryptosystems With Chosen Ciphertexts. 2021. URL: https://ieeexplore.ieee.org/document/9591340.

14. Prasanna Ravi, Shivam Bhasin, Sujoy Sinha Roy, and Anupam Chattopadhyay. On Exploiting Message Leakage in (few) NIST PQC Candidates for Practical Message Recovery and Key Recovery Attacks. 2020. URL: https://eprint.iacr.org/2020/1559.pdf.

15. Kalle Ngo, Elena Dubrova, Qian Guo, Thomas Johansson. A Side-Channel Attack on a Masked IND-CCA Secure Saber KEM Implementation. 2021. URL: https://tches.iacr.org/index.php/TCHES/article/view/9079/8666.

16. Kalle Ngo, Elena Dubrova, and Thomas Johansson. Breaking Masked and Shuffled CCA Secure Saber KEM by Power Analysis. 2021. URL: https://eprint.iacr.org/2021/902.pdf.

17. Mike Hamburg, Julius Hermelink, Robert Primas, Simona Samardjiska, Thomas Schamberger, Silvan Streit, Emanuele Strieder, and Christine van Vredendaal. Chosen Ciphertext k-Trace Attacks on Masked CCA2 Secure Kyber. 2021. URL: https://eprint.iacr.org/2021/956.pdf.

18. Dorian Amiet, Andreas Curiger, Lukas Leuenberger, and Paul Zbinden. Defeating NewHope with a Single Trace. 2020. URL: https://eprint.iacr.org/2020/368.pdf.

19. Peter Pessl and Robert Primas. More Practical Single-Trace Attacks on the Number Theoretic Transform. 2019. URL: https://eprint.iacr.org/2019/795.pdf.

20. Prasanna Ravi, Debapriya Basu Roy, Shivam Bhasin, Anupam Chattopadhyay, and Debdeep Mukhopadhyay. Number "Not Used" Once - Practical fault attack on pqm4 implementations of NIST candidates. 2018. URL: https://eprint.iacr.org/2018/211.pdf.

21. Felipe Valencia, Tobias Oder, Tim Güneysu, and Francesco Regazzoni. Exploring the Vulnerability of R-LWE Encryption to Fault Attacks. 2018. URL: https://dl.acm.org/doi/10.1145/3178291.3178294.

22. Keita Xagawa, Akira Ito, Rei Ueno, Junko Takahashi, and Naofumi Homma. Fault-Injection Attacks against NIST's Post-Quantum Cryptography Round 3 KEM Candidates. 2021. URL: https://eprint.iacr.org/2021/840.pdf.

23. Matthias J. Kannwischer, Joost Rijneveld, Peter Schwabe, and Ko Stoffelen. PQM4: Post-quantum crypto library for the ARM Cortex-M4. 2019. URL: https://github.com/mupq/pqm4.

24. Peter Pessl and Lukas Prokop. Fault Attacks on CCA-secure Lattice KEMs. 2021. URL: https://eprint.iacr.org/2021/064.pdf.

25. Jeroen Delvaux. Roulette: Breaking Kyber with Diverse Fault Injection Setups. 2021. URL: https://eprint.iacr.org/2021/1622.pdf.

26. Prasanna Ravi, Romain Poussier, Shivam Bhasin, and Anupam Chattopadhyay. On Configurable SCA Countermeasures Against Single Trace Attacks for the NTT - A Performance Evaluation Study over Kyber and Dilithium on the ARM Cortex-M4. 2020. URL: https://eprint.iacr.org/2020/1038.pdf.

27. Leon Botros, Matthias J. Kannwischer, and Peter Schwabe. Memory-Efficient High-Speed Implementation of Kyber on Cortex-M4. 2019. URL: https://eprint.iacr.org/2019/489.pdf.

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

Надійшла до редколегії 07.03.2023

Відомості про авторів:

Дерев'янко Ярослав Андрійович - АТ «Інститут інформаційних технологій», науковий співробітник-консультант; Україна; e-mail: yarik0009258@gmail.com; ORCID: https://orcid.org/0000-0002-3290-3373

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

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.03**

М.В. ЄСІНА, канд. техн. наук, А.А. КРАВЧЕНКО, С.О. КРАВЧЕНКО

ОГЛЯД ЗАГРОЗ БЕЗПЕЦІ ТА ЦІЛІСНОСТІ ДАНИХ У ХМАРНИХ ОБЧИСЛЕННЯХ

Вступ

Хмарні обчислення – це технологія, яка швидко набирає популярності та розвитку, поєднує у собі декілька підходів та моделей з надання та управління IT сервісами. Згідно з визначенням Національного інституту стандартів і технології (NIST) США, хмарні обчислення – це модель забезпечення повсюдного та зручного доступу на вимогу, через мережу до спільного пулу обчислювальних ресурсів, що підлягають налаштуванню (наприклад, до комунікаційних мереж, серверів, засобів збереження даних, прикладних програм та сервісів) і які можуть бути оперативно надані та звільнені з мінімальними управлінськими затратами та зверненнями до провайдера [1]. Хмарні обчислення розглядаються як одна з найуспішніших обчислювальних технологій, здатних вирішити цілу низку проблем, що стоять перед людством.

Хмарні обчислення мають декілька ключових особливостей, як-то надійність, широкий мережевий доступ, масштабованість інфраструктури, гнучкість, незалежність від місця розташування, економія на масштабах і економічна ефективність та стійкість [2, 3].

Через зростаючу популярність і широку експлуатацію послуг хмарних обчислень виникає необхідність високого рівня безпеки. У сьогоднішніх реаліях люди використовують технології хмарних обчислень у великих обсягах, наприклад на роботі, в особистих цілях та інше, так як вони мають велику довіру до цих технологій. Щоб запобігти втраті довіреної інформації провайдери послуг мають забезпечити її цілісність.

Стаття присвячена огляду загроз безпеці та цілісності технологій хмарних обчислень, тому що це є важливим аспектом даної технології. Зазначимо, що безпека хмарних обчислень означає захист даних, тоді як цілісність – їх надійність. Безпека та цілісність даних є основною проблемою користувачів, пов'язаною із хмарними обчисленнями.

Основна частина

В роботі [4] визначено безпеку хмарних обчислень як «піддомен комп’ютерної безпеки, мережевої безпеки та, ширше, інформаційної безпеки. Це стосується широкого набору політик, технологій і елементів керування, які застосовуються для захисту даних, додатків і відповідної інфраструктури хмарних обчислень».

Коли організація вирішує зберігати дані або розміщувати додатки в публічній хмарі, вона втрачає можливість мати фізичний доступ до серверів, на яких зберігається її інформація [8]. Як наслідок, потенційно конфіденційні дані піддаються ризику інсайдерських атак. Згідно зі звітом Альянсу хмарної безпеки за 2010 р. [5], внутрішні атаки є однією з семи найбільших загроз у хмарних обчисленнях. Тому постачальники хмарних послуг повинні забезпечити проведення ретельних перевірок співробітників, які мають фізичний доступ до серверів у дата-центрі. Крім того, центри обробки даних рекомендується часто моніторити на предмет підозрілої активності. Існує чотири основні аспекти безпеки в хмарі, за які відповідають як постачальники, так і клієнти [6]:

- обмеження доступу. Оскільки в хмарі всі ресурси доступні через Інтернет, дуже важливо переконатися, що лише належні користувачі матимуть доступ до потрібних їм інструментів протягом визначеного часу;
- захист даних. Організації повинні розуміти, де розташовано їхні ресурси, і застосувати відповідні елементи керування для захисту даних та інфраструктури, де вони розміщені;

- відновлення даних. У разі порушення безпеки надзвичайно важливо мати надійне рішення для резервного копіювання даних і план їхнього відновлення;
- план реагування. У разі атак організаціям потрібен спеціальний план, який дасть їм змогу зменшити наслідки та запобігти ураженню інших систем.

Загрози цілісності

Як і у будь-якій іншій системі, загрози безпеки для технологій хмарних обчислень можна поділити на загрози конфіденційності, цілісності та доступності. Загалом цілісність даних означає захист даних від несанкціонованого видалення, модифікації чи фальсифікації [10].

Предметом дослідження цієї статті є саме загрози цілісності, які розглянуто нижче:

1. *Несанкціонований доступ*.

Ця атака направлена на безконтрольні зміни даних, на які не зможе впливати авторизований користувач. Зловмисник, який проводить несанкціонований доступ з послідовною зміною даних, може провести атаку ззовні або зсередини організації-власника хмари. Це найсерйозніша атака, якщо це станеться, то витік даних відбуватиметься шляхом використання старого обладнання та повторного використання драйверів [11].

2. *Блокування даних*.

Ця загроза утворюється при переході від одного постачальника послуг до іншого. Так як різні постачальники надають різні послуги, тому при переході може трапитись втрата даних користувачів або їх блокування. У хмарі немає правил або умов щодо того, як зберігати дані, це залежить від постачальника хмарних послуг (CSP) [12]. Зазвичай, дані будуть розкидані по всьому серверу та системам. В ідеальній моделі міграція додатків від одного хмарного провайдера до іншого повинна бути простою, що є ще одним викликом для додатків хмарних обчислень, але оскільки кожен хмарний провайдер використовує окрему стандартну мову для своїх систем, це наразі неможливо.

3. *SQL-ін'єкції*.

SQL-ін’єкції націлені на SQL-сервери, які запускають уразливі програми баз даних. Хакери використовують вразливі місця веб-серверів і вводять шкідливий код, щоб обійти вхід і отримати несанкціонований доступ до серверних баз даних. У разі успіху хакери можуть маніпулювати вмістом баз даних, отримати конфіденційні дані, віддалено виконувати системні команди або навіть взяти під контроль веб-сервер для подальшої злочинної діяльності [13].

4. *Атака "людина посередині" (MiMA)*

MIMA зазвичай виникає, коли різні користувачі хмари спілкуються один з одним або спільно використовують ресурси з хмарного середовища [14]. Недостатнє шифрування може зробити користувачів вразливими до атаки "людина посередині", яка є непрямою атакою [15]. TLS – криптографічний протокол, який дозволяє клієнт-серверному додатку [16] запобігти підслуховуванню будь-якої конфіденційної інформації, що відбувається на HTTPS, який використовує TLS. Якщо людина отримує доступ до невідомої мережі і виконує свою роботу в HTTP, зловмисник, який виступає в ролі посередника, потім скористається цим, перехопивши всі конфіденційні дані через HTTPS-пакети.

5. *DDoS-атака*.

Мабуть, у сучасних технологіях це є найбільш серйозною проблемою, оскільки не може бути усуненою повністю. На сьогодні є лише деякі методи пом’якшення наслідків, які допомагають зменшити ризики та послідовності таких атак. DDoS-атаки націлені на веб-сайти та сервери, порушуючи роботу мережевих сервісів з метою виснаження ресурсів програми. Зловмисники, які стоять за цими атаками, наводнюють сайт несанкціонованим трафіком, що призводить до погіршення функціональності сайту або взагалі виводить його з ладу [17].

6. *Атаки на автентифікацію*.

Атаки на автентифікацію складно класифікувати як саме атаки на цілісність, але вони можуть спровокувати такі загрози, тому слід їх зазначити. Нижче наведено декілька відомих атак на автентифікацію:

- атака на відтворення.

Ця атака відбувається, коли невідома особа переглядає трафік даних, а потім надсилає комунікаційні дані на своє місце, як оригінального відправника. Щоб запобігти цій атаці, зазвичай впроваджують мітки часу та порядкові номери [18];

- атака грубої сили або атака за словником.

Це базова атака, при якій зловмисник перебирає всі можливі комбінації для пароля, щоб отримати доступ до даних користувачів. Чим більше довжина пароля, тим більше часу знадобиться зловмиснику, щоб вгадати правильний пароль [19];

- фішингова атака.

Йдеться про те, як зловмисник перебирає всі можливі способи атаки на жертву, підбираючи всі комбінації коду та паролів. Знов-таки, чим складніший код, тим більше часу знадобиться зловмиснику для його підбору [20], при цьому витрачений час буде зростати не лінійно.

7. *Атака відкату.*

Такі атаки можуть виникати під час оновлення системи у випадку, якщо постачальник у цей момент надає старе програмне забезпечення для користування. Це може спровокувати втрату даних, що зберігаються у цій системі. Відкат також відбувається без належного видалення старих даних користувача та оновлення системи до нової версії [21].

8. *Атака підробки тегів.*

Ця атака відбувається, якщо нечемний продавець обманює своїх клієнтів, показує неправильний штрих-код чи дає неправильне посилання. Якщо користувач сканує його на своєму пристрої, то зловмисник отримує доступ до всіх конфіденційних даних, що призводить до можливих ризиків шахрайства та витоку приватної інформації [21].

9. *Візантійська атака*.

Ця атака відбувається на різні частини хмарних обчислень шляхом зупинки або виходу з ладу систем. Це станеться, коли запит буде некоректно проходить через систему [21].

10. *Атака на систему доменних імен (DNS)*.

Ця атака відбувається, якщо систему атакує якесь шкідливе програмне забезпечення. DNS перетворює доменні імена на IP-адреси, і користувач не може бачити, наскільки правильно відбувається перетворення. Кожного разу, коли відкривається невідома веб-сторінка, зловмисник може легко отримати доступ до персональної інформації, що використовується на серверах [22].

11. *Сніфферні атаки*.

Атака відбувається, коли користувач натискає на деякі SOAP (Simple Object Access Protocol) – повідомлення або шкідливе посилання. Після того, як натиснуте посилання буде активоване, програма перехоплює потік пакетів в мережі і отримує доступ до персональних даних користувачів, таких як паролі, реквізити банківських рахунків тощо, які не є зашифрованими [23]. Залежно від обсягу даних та їх характеру, внаслідок цієї атаки втрати даних можуть бути непомітними або стати причиною великих проблем для цілої організації.

Методи забезпечення цілісності даних у хмарному середовищі

Проаналізувавши наведені вище загрози, перейдемо до методів забезпечення цілісності та запобігання атак на цілісність у хмарних сховищах. Як було зазначено, не завжди можна повністю усунути загрози, але майже завжди можна вжити заходи для їх запобігання та зменшення ймовірних втрат. Існує декілька механізмів та схем, які запропоновано для захисту володіння даними та їх цілісності в середовищі хмарних обчислень. Нижче наведено декілька механізмів та схем [21].

1. *Пом'якшення наслідків підробки тегів та атак витоку даних.*

Щоб запобігти цій атаці, існує схема, запропонована Yun Zhu та ін. [24], відома як Cooperative Provable Data Possession (CPCP), яка використовується в поєднанні з двома

іншими (Homomorphic Verifiable Response та Hash Index Hierarchy), що забезпечує прозору перевірку даних та надійний захист.

Сутність даного методу виглядає так: перед тим, як клієнт надсилає інформацію до CSP, він створює тег виклику, а надсилає його постачальнику хмарних послуг пізніше. Клієнт кидає виклик провайдеру хмарних послуг, перевіряючи цілісність даних за допомогою довіреної третьої сторони (TTP) [21].

2. *Послаблення атаки відкату*.

У запропонованій схемі захист від атаки відкату в хмарному середовищі здійснюється шляхом застосування методу геш-дерева Меркла [25, 26]. У цьому методі тег блоку даних та значення його лічильника оновлюються щоразу, коли оновлюються нові дані. Якщо зловмисник хоче змінити дані, значення лічильника також зміниться.

3. *Пом'якшення наслідків візантійського збою та зловмисних атак на дані*.

Рішенням для даної проблеми є запропонована Browers та ін. [27] криптосистема HAIL (High Availability and Integrity Layer) Protocol. Цей протокол гарантує, що дані користувача зберігаються неушкодженими і можуть бути безпечно отримані з серверів. Для забезпечення гарантії доступності даних використовується коригувальний код Erasure [21].

4. *Захист цілісності даних за допомогою шифрування*.

Даний метод є найкращим для загального захисту даних у хмарному середовищі та використовується у будь-якій системі. Оскільки технології продовжують розвиватися, а старі технології стають вразливіше, з'являтимуться нові методи злому шифрів, а також фатальні недоліки в старих методах шифрування. Хмарні провайдери повинні постійно оновлювати своє шифрування, оскільки дані, які вони зазвичай містять, є особливо цінними [7]. Перед тим як зберігати дані на сервері, слід зашифрувати їх та обчислити геш-значення даних. Це гарантуватиме, що дані не були змінені [28].

5. *Техніка доказового володіння даними (PDP)*.

Техніка PDP використовує протокол відповіді на запит для перевірки цілісності даних, що зберігаються на хмарному сервері. Цей метод використовує симетричне шифрування, наприклад MAC або будь-яке інше. Файл заповнюється метаданими перед зберіганням або надсиланням його на хмарний сервер. Після надсилання файлу до постачальника хмарних послуг користувач все одно зберігає метадані файлу, щоб перевірити його цілісність. Після цього користувач видаляє локальну копію файлу та перевіряє докази володіння файлом сервером за допомогою протоколу відповіді на виклик [29].

6. *Техніка доведення можливості вилучення (POR)*.

Метод Proof of Retrievability (POR) використовується для віддаленої перевірки даних, які зберігаються у постачальника хмарних послуг, за допомогою ключа автентифікації. У цьому методі дані не потрібно отримувати з CSP, і користувач також не зберігає оригінальну копію файлу локально. Користувач зберігає свій файл у CSP разом із ключем автентифікації. Потім користувач може перевірити цілісність даних за допомогою ключа автентифікації, не отримуючи файл з CSP [29, 9].

Висновки

Розглянуто поняття безпеки даних у хмарних обчисленнях та декілька з тих поширених загроз, що заважають забезпечити цілісність інформації, що там зберігається. Зараз технології хмарних обчислень – це те, чим люди користуються, майже не акцентуючи на це увагу, бо у сьогоденні це – явище, що зустрічається майже усюди та у кожній компанії. Не дивлячись на те, що хмарні технології зазвичай впроваджують найсучасніші технології безпеки, на жаль, абсолютно позбутись усіх ризиків неможливо. Також розглянуто можливі методи вирішення проблем, мінімізування можливих вразливостей від атак та забезпечення цілісності у хмарних обчисленнях.

Список літератури:

1. P. Mell and T. Grance. The NIST Definition of Cloud Computing // National Institute of Standards and Technology. 2009. Vol.53,no.6. P. 50.

2. Reese G. (2009) Cloud Application Architectures: Building Applications and Infrastructure in the Cloud. Sebastopol. California : O'Reilly Media.

3. Buyya R., Yeo C.S., Venugopal S., Broberg J., and Brandic I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility // Future Generation Computer Systems. 2009. 25 (6). P. 599 – 616.

4. Cloud computing security. Режим доступу: http://en.wikipedia.org/wiki/Cloud_ computing_security.

5. Top Threats to Cloud Computing v1.0 Cloud Security Alliance.

6. Що таке безпека в хмарі?, Microsoft. Режим доступу: https://www.microsoft.com/uk-ua/security/business/security-101/what-is-cloud-security.

7. Rukavitsyn Andrey, Borisenko Konstantin, Holod Ivan, Shorov Andrey. The method of ensuring confidentiality and integrity data in cloud computing // 2017 XX IEEE International Conference on Soft Computing and Measurements (SCM). 2009. P. 272 – 274.

8. Yunchuan Sun, Junsheng Zhang, Yongping Xiong, and Guangyu Zhu – Data Security and Privacy in Cloud Computing.

9. M. S. Giri, B. Gaur, D. Tomar. A Survey on Data Integrity Techniques in Cloud Computing.

10. Yunchuan Sun, Junsheng Zhang zhangjs, Yongping Xiong, and Guangyu Zhu (2014. Data Security and Privacy in Cloud Computing.

11. Dissanayaka, Akalanka Mailewa, Susan Mengel, Lisa Gittner, and Hafiz Khan. Vulnerability prioritization, root cause analysis, and mitigation of secure data analytic framework implemented with mongodb on singularity linux containers // Proceedings of the 2020 the 4th International Conference on Compute and Data Analysis. 2020. P. 58 – 66.

12. A. Jyoti, M. Shrimali, S. Tiwari, and H. P. Singh. Cloud computing using load balancing and service broker policy for IT service: a taxonomy and survey // Ambient Intell. Humaniz. Comput., vol. 11, no. 11, pp. 4785 – 4814, Nov. 2020, doi: 10.1007/s12652-020-01747-z.

13. Te-Shun Chou. Security threats on cloud computing vulnerabilities // International Journal of Computer Science & Information Technology (IJCSIT) Vol. 5, No 3, June 2013, pp. 84 – 85.

14. Ramandeep Kaur, Pushpendra Kumar Pateriya. A Study on Security Requirements in Different Cloud Frameworks // International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Vol.3, Iss.1, March 2013, pp.134 – 135.

15. Y. Chen, L. Li, and Z. Chen. An Approach to Verifying Data Integrity for Cloud Storage // 2017 13th International Conference on Computational Intelligence and Security (CIS), Dec. 2017, pp. 582 – 585, doi: 10.1109/CIS.2017.00135.

16. H. Mohapatra. Handling of Man-In-The-Middle Attack in WSN Through Intrusion Detection System // Int. J. Emerg. Trends Eng. Res., vol. 8, no. 5, pp. 1503 – 1510, May 2020, doi: 10.30534/ijeter/2020/05852020.

17. What is a DDoS attack? Режим доступу – https://www.microsoft.com/en-us/security/business/security-101/what-is-a-ddos-attack.

18. Lai Cheng-I., Alberto Abad, Korin Richmond, Junichi Yamagishi, NajimDehak, and Simon King. Attentive filtering networks for audio replay attack detection // ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6316 – 6320. IEEE, 2019.

19. Shetty, Roshan Ramprasad, Akalanka Mailewa Dissanayaka, Susan Mengel, Lisa Gittner, Ravi Vadapalli, and Hafiz Khan. Secure NoSQL based medical data processing and retrieval: the exposome project // Companion Proceedings of the10th International Conference on Utility and Cloud Computing, pp. 99 – 105. 2017.

20. Mailewa Dissanayaka, Akalanka, Roshan Ramprasad Shetty, Samip Kothari, Susan Mengel, Lisa Gittner, and Ravi Vadapalli. A review of MongoDB and singularity container security in regards to hipaa regulations // Companion Proceedings of the10th International Conference on Utility and Cloud Computing, pp. 91 – 97. 2017.

21. Survey on various data integrity attacks in cloud environment and the solutions // IEEE Conference Publication. Режим доступу – https://ieeexplore.ieee.org/abstract/document/6528889.

22. Thapa, Suman, and Akalanka Mailewa. The Role of Intrusion Detection/Prevention Systems in Modern Computer Networks: A Review // Conference: Midwest Instruction and Computing Symposium (MICS), vol. 53, pp. 1 – 14. 2020.

23. S. Sudalai and S. S., A Survey on Cloud Security Issues and Challenges with Possible MeasuresA Survey on Cloud Security Issues and Challenges with Possible Measures. 2016.

24. Y. Zhu, H. Hu, G. Ahn, and M. Yu. Cooperative Provable Data Possession for Integrity Verification in Multicloud Storage // IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 12, pp. 2231 – 2244, Dec. 2012, doi: 10.1109/TPDS.2012.66.

25. J. Feng, Y. Chen, D. H. Summerville, and K. Hwang. Fair Non-repudiation Framework for Cloud Storage: Part II // Cloud Computing for Enterprise Architectures, Z. Mahmood and R. Hill, Eds. London: Springer, 2011, pp. 283 – 300.

26. J. Feng, Y. Chen, D. Summerville, W. Ku, and Z. Su. Enhancing cloud storage security against roll-back attacks with a new fair multi-party nonrepudiation protocol // 2011 IEEE Consumer Communications and Networking Conference (CCNC), Jan. 2011, pp. 521 – 522, doi: 10.1109/CCNC.2011.5766528.

27. H. Lin and W. Tzeng. A Secure Erasure Code-Based Cloud Storage System with Secure Data Forwarding // IEEE Trans. Parallel Distrib. Syst., vol. 23, no. 6, pp. 995 – 1003, Jun. 2012, doi: 10.1109/TPDS.2011.252.

28. R. V. Rao and K. Selvamani. Data Security Challenges and Its Solutions in Cloud Computing // Procedia Comput. Sci., vol. 48, pp. 204 – 209, Jan.2015, doi: 10.1016/j.procs.2015.04.171.

29. K. N. Sevis and E. Seker. Survey on Data Integrity in Cloud // 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), Jun. 2016, pp. 167 – 171, doi: 10.1109/CSCloud.2016.35.

Надійшла до редколегії 11.03.2023

Відомості про авторів:

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

Кравченко Софія Олександрівна – Харківський національний університет імені В. Н. Каразіна, студентка кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: sofiya.krav@gmail.com

Кравченко Анастасія Андріївна – Харківський національний університет імені В. Н. Каразіна, студентка кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: aakrav02@gmail.com

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.04**

М.В. ЄСІНА, канд. техн. наук, В.В. ОНОПРІЄНКО, канд. техн. наук, А.В. ТОЛОК

МОДЕЛІ ЗАГРОЗ ДЛЯ ХМАРНИХ ПОСЛУГ

Загальний огляд хмарних послуг

Хмарні сервіси – це низка IT-додатків і ресурсів, які включають програмне забезпечення, інфраструктуру та платформи, розміщені у сторонніх провайдерів і надаються на вимогу організаціям та окремим клієнтам через Інтернет. Їх також можна назвати хмарними обчисленнями.

Хмарні сервіси полегшують передачу даних на сервери постачальників послуг і з них, а також на сервери та гаджети клієнтів. Користувачі можуть отримати доступ до хмарних сервісів через комп'ютер з підключенням до Інтернету або віртуальної приватної мережі. Вони дозволяють клієнтам відмовитися від інвестицій в програмне забезпечення та придбання допоміжної мережевої інфраструктури і серверів. Використання хмарних сервісів дозволяє клієнтам отримати доступ до програмного забезпечення, хмарних сховищ, обчислювальних потужностей, IT-інфраструктури та інших послуг без необхідності нести витрати на обслуговування або оновлення програмного та апаратного забезпечення. Постачальники хмарних послуг використовують різні моделі тарифікації, які залежать від спожитих ресурсів. Зазвичай це плани з щомісячною або річною підпискою, які оплачуються за фактом використання.

Хмарні обчислення стали популярною технологією завдяки своїм численним перевагам над традиційними обчисленнями. На відміну від традиційних обчислень, хмарні обчислення дозволяють компаніям отримувати доступ до програмного забезпечення, обладнання та інших послуг віддалено, масштабуючи їх за потреби. Компанії платять лише за ті послуги, які їм потрібні, що може значно зменшити початкові інвестиції та поточні операційні витрати. Крім того, хмарні обчислення є більш безпечними та надійними, ніж традиційні, завдяки можливості віддаленого доступу до даних та високому рівню шифрування і протоколів безпеки, що використовуються постачальниками хмарних послуг.

За даними Gartner [1], традиційні IT-витрати все ще домінують над хмарними. Однак, згідно з прогнозом на 2019 – 2025 роки, витрати на хмарні технології продовжуватимуть зростати, тоді як традиційні витрати на IT продовжуватимуть скорочуватися і зрештою відставатимуть від витрат на хмарні технології з 2025 року.

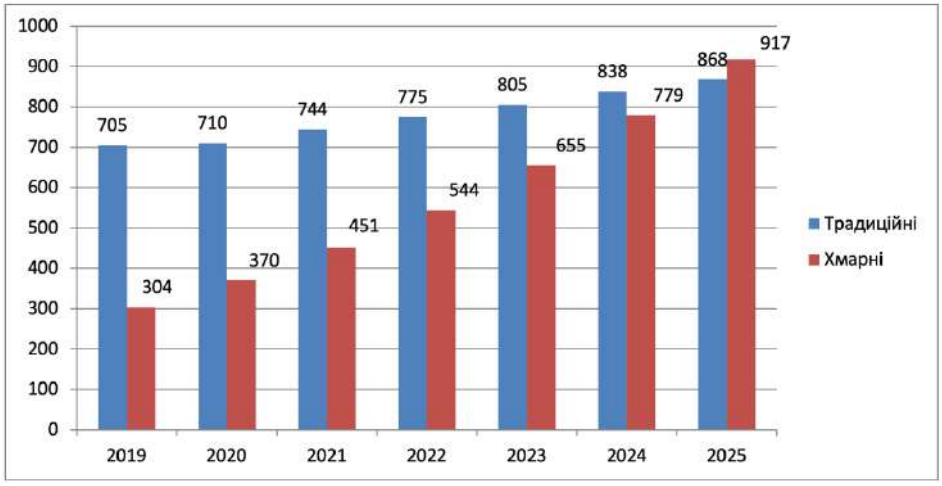


Рис. 1. Порівняння IT-витрат на традиційні та хмарні технології (млрд. $) [1]

Отже, вибір між хмарними та традиційними послугами залежить від потреб індивідуального бізнесу та його відношення до ризиків безпеки. Незалежно від того, який варіант буде

вибраний, важливо забезпечувати безпеку та захист своїх даних, використовуючи кращі практики та відповідні методики безпеки [8].

Укладаючи договір з постачальником хмарних обчислень, потрібно враховувати їх характеристики. Національний інститут стандартів США (NIST) перераховує п'ять основних характеристик хмарних обчислень: самообслуговування за запитом, широкий доступ до мережі, об'єднання ресурсів, швидка масштабованість і узгоджене обслуговування [7].

Моделі розгортання хмарних обчислень вказують на те, як хмарні сервіси стають доступними для користувачів. Існує чотири моделі розгортання, пов'язані з хмарними обчисленнями [2]:

- публічна хмара – цей тип хмарної моделі розгортання підтримує всіх користувачів, які хочуть використовувати обчислювальні ресурси, такі як апаратне (ОС, процесор, пам'ять, сховище) або програмне забезпечення (сервер додатків, база даних) на основі підписки. Найчастіше публічні хмари використовуються для розробки та тестування додатків, некритичних завдань, таких як обмін файлами та електронна пошта;
- приватна хмара – це, як правило, інфраструктура, що використовується однією організацією. Такою інфраструктурою може керувати сама організація для підтримки різних груп користувачів, або ж нею може керувати постачальник послуг, який обслуговує її на місці або за межами організації. Приватні хмари дорожчі за публічні через капітальні витрати, пов'язані з їх придбанням та обслуговуванням. Однак приватні хмари краще вирішують проблеми безпеки та конфіденційності, які сьогодні хвилюють організації;
- гібридна хмара – у цій моделі організація використовує взаємопов'язану приватну та публічну хмарну інфраструктуру. Багато організацій використовують цю модель, коли їм потрібно швидко масштабувати свою IT-інфраструктуру, наприклад, коли вони використовують публічні хмари для доповнення потужностей, доступних у приватній хмарі. Наприклад, якщо Інтернет-магазину потрібно більше обчислювальних ресурсів, він може отримати ці ресурси через публічні хмари;
- суспільна хмара – модель розгортання підтримує спільне використання обчислювальних ресурсів кількома організаціями, які є частиною спільноти. Прикладами можуть бути університети, які співпрацюють у певних галузях. Доступ до хмарного середовища спільноти зазвичай обмежується членами спільноти.

Таблиця 1

Обслуговування та управління різними видами хмарних ресурсів

Вид хмари	Ким обслуговується	Хто є власником	Де знаходиться інфраструктура	У кого є доступ
Публічна	Зовнішнім провайдером	Зовнішній провайдер	У зовнішнього провайдера	У будь-якого користувача
Приватна/ суспільна	Користувачем або зовнішнім провайдером	Користувач або зовнішній провайдер	У зовнішнього провайдера або у користувача	У авторизованого користувача
Гібридна	Користувачем і зовнішнім провайдером	Користувач і зовнішній провайдер	У зовнішнього провайдера і у користувача	У авторизованих і у будь-яких зовнішніх користувачів

Моделі обслуговування хмарних технологій

Існують три основні моделі обслуговування хмарних сервісів [5]: програмне забезпечення як послуга (SaaS), інфраструктура як послуга (IaaS) та платформа як послуга (PaaS). Моделі ціноутворення хмарних сервісів поділяються на моделі з оплатою за використання, на основі підписки та гібридні, що є поєднанням моделей з оплатою за використання та підписки.

Постачальники програмного забезпечення як послуги (SaaS) розміщують додатки, роблячи їх доступними для користувачів через Інтернет. Завдяки SaaS компаніям не потрібно встановлювати або завантажувати будь-яке програмне забезпечення в існуючу IT-інфраструктуру. Модель гарантує, що користувачі завжди використовують найновіші версії програмного забезпечення. Обслуговуванням і підтримкою займається постачальник.

Основна перевага продуктів SaaS полягає в тому, що організації можуть використовувати їх одразу після підписки, оскільки це найпростіша в налаштуванні та експлуатації хмарна модель. Щоб додати користувачів, організаціям достатньо оновити свої існуючі плани або підписки. Їм не потрібно купувати додаткове місце на сервері або ліцензії на програмне забезпечення. Основним недоліком моделі є відсутність контролю. Організації не мають контролю над хмарною інфраструктурою своїх провайдерів. Отже, якщо у провайдера трапляються перебої в роботі, то і у них теж.

Платформа як послуга (PaaS) пропонує платформу для розробки та розгортання програмного забезпечення через Інтернет, надаючи їм доступ до найсучасніших інструментів. PaaS надає фреймворк, який розробники можуть використовувати для створення індивідуальних додатків. Організація або постачальник хмарних послуг керує серверами, сховищами та мережею, а розробники керують додатками. Провайдери PaaS надають більшу частину IT-послуг для організацій, до яких користувачі можуть отримати доступ, якщо у них є підключення до Інтернету та веб-браузер. Вони також допомагають командам розробників працювати разом, незалежно від того, де вони фізично знаходяться. До недоліків можна віднести: відсутність масштабованості та прив'язку до постачальника

Інфраструктура як послуга (IaaS) використовується компаніями, які не хочуть утримувати власні дата-центри. IaaS надає віртуальні обчислювальні ресурси через Інтернет. Хмарний постачальник розміщує компоненти інфраструктури, які зазвичай існують в локальному центрі обробки даних, включаючи сервери, сховища та мережеве обладнання. IaaS полегшує, прискорює та робить більш економічно ефективним управління робочими навантаженнями для організацій, оскільки їм не потрібно купувати, управляти та підтримувати базову інфраструктуру. Хмарна інфраструктура гарантує, що компанії мають доступ до всіх необхідних ресурсів, коли вони їм потрібні. Безпека – найуразливіше місце у цій моделі. У середовищі IaaS організації передають контроль над безпекою хмари сторонньому постачальнику. Тож, навіть, якщо витік даних не вплине безпосередньо на дані компанії, скомпрометована система все одно може поставити під загрозу її діяльність. Деякі організації можуть відчувати простої в роботі з IaaS, які вони не можуть контролювати. Будь-які проблеми, що виникають у провайдера, можуть обмежити доступ компаній до додатків і даних, необхідних для щоденної роботи.

Таблиця 2

Моделі обслуговування за засобами доступу і управління

Моделі обслуговування	Засоби доступу і управління	Вміст
ПЗ як сервіс (SaaS)	Веб-браузер	Хмарні програми: соціальні мережі, офісні застосунки, системи управління вмістом, інтелектуальна обробка даних.
Платформа як сервіс (PaaS)	Хмарне середовище розробки	Хмарна платформа: мови програмування, бібліотеки, утиліти конфігурації композицій сервісів, структуровані дані.
Інфраструктура як сервіс (IaaS)	Система управління віртуальної інфраструктурою	Хмарна інфраструктура: обчислювальні сервера, сховища даних, організація мережевих з’єднань.

Загрози хмарних обчислень

Хмарна атака – це кібератака, націлена на платформи хмарних послуг, наприклад: обчислювальні служби, служби зберігання даних або програмне забезпечення. Хмарні атаки можуть мати серйозні наслідки, такі як витік даних, втрата даних, несанкціонований доступ до конфіденційної інформації та збої в роботі служб. Оскільки все більше організацій і окремих осіб покладаються на хмарні обчислення для зберігання та обробки даних, відповідно збільшується і кількість потенційних цілей для зловмисників. Найбільш значні загрози, які пов’язані з хмарними обчисленнями [3]:

– відмова в обслуговуванні (DDoS): це спроба порушити нормальну роботу системи, перевантаживши її трафіком. У випадку хмарного середовища це зазвичай відбувається шля-

хом одночасного надсилання тисяч і тисяч з'єднань. Ці запити перевантажують сервер і заважають йому обробляти законні запити;

– викрадення облікового запису: це процес, під час якого хмарний обліковий запис фізичної особи або організації викрадається зловмисником. Захоплення хмарних облікових записів є поширеною тактикою в схемах крадіжки персональних даних, коли зловмисник використовує скомпрометований обліковий запис електронної пошти або інші облікові дані, щоб видати себе за власника облікового запису;

– внутрішні загрози: це категорія ризику, яку становлять ті, хто має доступ до фізичних або цифрових активів організації. Такими інсайдерами можуть бути нинішні працівники, колишні працівники, підрядники, постачальники, які мають (або мали) санкціонований доступ до мережі та комп'ютерних систем організації;

– неправильна конфігурація хмари: неправильна конфігурація є проблемою хмарних обчислень, оскільки хмарні середовища можуть бути досить складними, а виявити та виправити помилки вручну може бути важко. Це будь-які збої, прогалини або помилки, які можуть наразити ваше середовище на ризик під час переходу на хмарні технології. Ці кіберзагрози проявляються у вигляді порушень безпеки, що можуть бути використані для несанкціонований доступу до мережі;

– шкідливі файли cookie: зараження файлами cookie в хмарних додатках означає несанкціоновану модифікацію або впровадження шкідливого вмісту в файл cookie, який є невеликим фрагментом даних, що зберігається на комп'ютері користувача. У SaaS та інших хмарних додатках файли cookie часто містять облікові дані, тому зловмисники можуть модифікувати ці файли, щоб отримати доступ до додатків;

– витік даних: це кібератака, під час якої до чутливих, конфіденційних або інших захищених даних було отримано несанкціонований доступ або вони були розголошені. Порушення даних може статися в організації будь-якого розміру, від малого бізнесу до великих корпорацій.

Системи моделювання загроз

Моделювання загроз – це процес визначення, оцінки та зменшення ризиків безпеки в додатку або системі. Використовуючи систему моделювання загроз, ви можете розподілити ресурси для протидії ймовірним загрозам, захисту життєво важливих активів і підтримки безперервності бізнесу. Існують методології та стратегії, які допоможуть зрозуміти, як ваша організація вписується в зростаючий ландшафт загроз і, що ви можете зробити для його захисту.

Існує кілька фреймворків моделювання загроз, які організації можуть використовувати для виявлення потенційних загроз безпеці та вразливостей. Ось деякі з найбільш поширених фреймворків [4, 7]:

1. STRIDE: це підхід до інтеграції на більш ранніх етапах життєвого циклу розробки програмного забезпечення. Як методологія моделювання загроз, фреймворк STRIDE використовується для створення карти додатку на основі його унікальних варіантів використання та бізнес-логіки. Таким чином, його можна використовувати для виявлення та усунення потенційних вразливостей ще до того, як буде написаний хоч один рядок коду. Також можна повертатися до фреймворку STRIDE в будь-який час, поки додаток розробляється або знаходиться у виробництві, і кожного разу, коли випускається новий код, щоб побачити, як він вплине на загальний вектор атак на додаток. Використання моделювання загроз має стати вашим першим кроком на шляху до створення мереж, систем і додатків, які будуть безпечними за своєю суттю. STRIDE – це модель загроз, яку можна використовувати як основу для забезпечення безпечного дизайну додатків [10].

2. DREAD: розшифровується як пошкодження (Damage), відтворюваність (Reproducibility), можливість експлуатації (Exploitability), постраждалі користувачі (Affected users) та можливість виявлення (Discoverability). Це модель оцінки ризиків, яка допомагає пріори-

тезувати загрози відповідно до їх потенційного впливу. DREAD найбільш підходить для малих та середніх організацій, які хочуть швидкий та простий спосіб пріоритезації загроз [9].

3. PASTA: це ризико-орієнтована методологія моделювання загроз, заснована у 2015 р. PASTA дозволяє співпрацювати між розробниками та зацікавленими сторонами бізнесу, щоб по-справжньому зрозуміти ризики, притаманні вашому додатку, ймовірність атаки та вплив на бізнес, якщо відбудеться компрометація. Модель складається з семи етапів, кожен з яких діє як будівельний блок один до одного. Такий підхід дозволяє моделі загроз бути лінійним процесом і використовувати існуючі процеси тестування безпеки у вашій організації, такі як: перегляд коду, аналіз сторонніх бібліотек, статичний аналіз і моніторинг загроз для інфраструктури додатків [7].

4. Attack Trees: дерева атак – це діаграми, які зображують атаки на систему у вигляді дерева. Корінь дерева – це мета атаки, а гілки – шляхи досягнення цієї мети. Кожна мета представляється у вигляді окремого дерева. Таким чином, в результаті аналізу загроз системі створюється набір дерев атак. Використання дерев атак для моделювання загроз є одним з найстаріших і найбільш широко застосовуваних методів. Дерева атак спочатку застосовувалися як окремий метод, а потім були об'єднані з іншими методами та фреймворками [4].

Важливо пам'ятати, що ці фреймворки мають свої переваги та недоліки, тому вибір фреймворку залежить від контексту та потреб конкретної системи.

Таблиця 3

Переваги та недоліки фреймворків

Фреймворк	Переваги	Недоліки
STRIDE	Простий та ефективний у використанні. Допомагає виявити вразливості та загрози на ранніх етапах розробки	Може бути недостатньо детальним. Не надає повного огляду системи. Важко використовувати для складних архітектур
DREAD	Допомагає визначити критичні ризики та рівень загроз, є детальним та точним, може бути використаний для оцінки зроблених вдосконалень у вирішенні загроз	Може бути складним у використанні для новачків. Недостатньо гнучкий для деяких типів систем
PASTA	Надає широкий огляд системи, включаючи аналіз атак. Допомагає знайти слабкі місця в системі. Детальний та комплексний	Використання потребує багато часу. Недостатньо гнучкий для деяких типів системи
Attack Trees	Допомагає виявляти потенційні загрози та визначати критичні елементи системи. Може бути використаний для різних типів систем. Простий та ефективний у використанні	Недостатньо детальний для деяких складних систем. Може використовувати значну кількість ресурсів для виконання

Захист від загроз хмарних послуг

Безпека хмарних послуг стає критичною проблемою, оскільки все більше компаній завершують свою цифрову трансформацію. Хмарні обчислення супроводжують новий робочий світ без кордонів, що сприяє вільному потоку інформації. Це дозволило компаніям бути більш продуктивними і зробило можливою віддалену роботу. Щоб захистити хмарні сервіси від загроз, важливо застосовувати комплексний підхід, який враховує унікальні ризики, пов'язані з хмарними середовищами. Ось кілька найкращих засобів захисту від загроз для хмарних сервісів [11]:

1. Шифрування: Шифруйте дані як у стані спокою, так і під час передачі. Переконайтеся, що ключі шифрування зберігаються належним чином.

2. Мережева безпека: Впроваджуйте засоби контролю мережевої безпеки, такі як брандмауери, системи виявлення/запобігання вторгнення.

3. Контроль доступу: Використовуйте контроль доступу на основі ролей, щоб гарантувати, що користувачі мають доступ лише до тих ресурсів, які їм потрібні.

4. Управління виправленнями: Оновлюйте все програмне забезпечення та системи найновішими оновленнями безпеки, щоб зменшити ризик їх використання зловмисниками.

5. Резервне копіювання та аварійне відновлення: Впровадьте комплексний план аварійного відновлення, щоб мінімізувати вплив будь-якої потенційної втрати даних.

6. Моніторинг загроз: Впровадьте систему моніторингу загроз для виявлення та реагування на інциденти безпеки в режимі реального часу.

Ці засоби захисту є гарною відправною точкою для організацій, щоб захистити свої хмарні сервіси від цілого ряду ризиків безпеки. Однак важливо відзначити, що безпека – це безперервний процес, і організаціям необхідно регулярно переглядати і оновлювати свої системи безпеки, щоб залишатися на крок попереду нових загроз.

Висновки

Хмарні послуги забезпечують користувачам можливість зберігання та обробки даних на віддалених серверах, що дає їм доступ до цих даних з будь-якого місця та пристрою з підключенням до Інтернету. Основною перевагою хмарних послуг є їх гнучкість, ефективність та високий рівень захисту даних.

Обираючи постачальників хмарних послуг, потрібно враховувати основні характеристики хмарних обчислень: самообслуговування за запитом, широкий доступ до мережі, об'єднання ресурсів, швидка масштабованість і узгоджене обслуговування, моделі розгортання (публічна хмара, приватна хмара) та модель обслуговування (SaaS, PaaS, IaaS). Щоб захистити хмарні сервіси, важливо використовувати комплексні підходи. Щодо моделювання загроз слід використовувати фрєймворки моделювання загроз. Важливо пам'ятати, що фреймворки мають свої переваги та недоліки, тому вибір залежить від контексту та потреб конкретної системи.

Список літератури:

1. Cloud Computing Statistics (2023). [Електронний ресурс]. Режим доступу: https://parachute.cloud/cloud-computing-statistics/.

2. Cloud Deployment Models. [Електронний ресурс]. Режим доступу: https://www.geeksforgeeks.org/cloud-deployment-models/.

3. What are Cloud Security Threats? [Електронний ресурс]. Режим доступу: https://www.vectra.ai/learning/cloud-security-threats.

4. Threat Modeling 12 Available Methods. [Електронний ресурс]. Режим доступу: https://insights.sei.cmu.edu/blog/threat-modeling-12-available-methods/.

5. IaaS vs. PaaS vs. SaaS: Cloud Service Model Overview [Електронний ресурс]. Режим доступу: www.intel.com/content/www/us/en/cloud-computing/as-a-service.

6. What is PASTA Threat Modeling?. [Електронний ресурс]. Режим доступу:https://versprite.com/blog/what-is-pasta-threat-modeling/.

7. Essential Cloud Computing Characteristics. [Електронний ресурс]. Режим доступу: https://www.synopsys.com/cloud/insights/essential-cloud-computing-characteristics.html.

8. Cloud Computing vs Traditional Computing. [Електронний ресурс]. Режим доступу: https://www.simplilearn.com/cloud-computing-vs-traditional-computing-article.

9. Threat Modeling with DREAD. [Електронний ресурс]. Режим доступу: https://cyral.com/glossary/threat-modeling-with-dread/.

10. STRIDE Threat Modeling: What You Need to Know. [Електронний ресурс]. Режим доступу: https://www.softwaresecured.com/stride-threat-modeling/.

11. Cloud Infrastructure Security: 7 Best Practices to Secure Your Sensitive Data. [Електронний ресурс]. Режим доступу: https://www.techtarget.com/searchsecurity/definition/cloud-security.

Надійшла до редколегії 11.02.2023

Відомості про авторів:

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

Онопрієнко Віктор Васильович – канд. техн. наук, АТ «Інститут Інформаційних Технологій», Генеральний директор; Україна; e-mail: v25258@gmail.com

Толок Анатолій Вікторович – Харківський національний університет імені В. Н. Каразіна, студент кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; Україна; e-mail: xa12850340@student.karazin.ua

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.05**

Ю.І. ГОРБЕНКО, канд. техн. наук, М.В. ЄСІНА, канд. техн. наук,
В.А. ПОНОМАР, канд. техн. наук, І.Д. ГОРБЕНКО, д-р техн. наук, Є.Ю. КАПТЬОЛ

НАУКОВО-МЕТОДИЧНІ ОСНОВИ АНАЛІЗУ, ОЦІНКИ ТА РЕЗУЛЬТАТИ ПОРІВНЯННЯ ІСНУЮЧИХ ТА ПЕРСПЕКТИВНИХ (ПОСТКВАНТОВИХ) АСИМЕТРИЧНИХ КРИПТОГРАФІЧНИХ ПРИМІТИВІВ ЕЛЕКТРОННОГО ПІДПИСУ, ПРОТОКОЛІВ АСИМЕТРИЧНОГО ШИФРУВАННЯ ТА ПРОТОКОЛІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ

Вступ

Важливою проблемою в криптології є аналіз шляхів зниження ризиків для вразливих криптографічних систем та стану їх розроблення, прийняття та впровадження на міжнародному та національному рівнях постквантових стандартів асиметричних криптоперетворень електронних підписів (ЕП), асиметричних шифрів (АСШ) та протоколів інкапсуляції ключів (ПІК). Тому процеси зниження ризиків для вразливих існуючих стандартизованих криптографічних систем, визначення напрямків розвитку математичних методів та дослідження перспектив їх застосування в ході створення стандартизованих ЕП, АСШ та ПІК є суттєво значимими. Вони зводяться до обґрунтування та визначення математичних методів та механізмів, які дозволять створити перспективні (постквантові) стандартизовані ЕП, АСШ та ПІК [1 – 23].

Підтвердженням наявності ризиків щодо застосування квантових обчислень для злому існуючих асиметричних стандартів криптографічного захисту інформації є прийняття в США закону в формі «Меморандуму про національну безпеку з просування лідерства США в галузі квантових обчислень при одночасному зниженні ризиків для вразливих криптографічних систем» від 04 травня 2022 р. Однак поряд із потенційними перевагами, квантові обчислення також ймовірно будуть становити для економічної та національної безпеки значні ризики.

В подальшому під перехідним періодом будемо розуміти проміжок часу у майбутньому, коли будуть суттєво вдосконалені класичні методи та засоби криптоаналізу, а також будуть створені та застосовуватись для криптоаналізу квантові комп’ютери з обмеженими потужностями. У цей період можуть бути застосованими існуючі стандарти асиметричних та криптографічних перетворень, але з максимально можливими чи збільшеними довжинами загальносистемних параметрів та ключових даних, обмеженою надійністю функціонування. Постквантовий період пропонується визначити як проміжок часу у майбутньому, коли будуть суттєво удосконалені класичні методи та створені квантові комп’ютери з необхідними для успішного криптоаналізу довжинами регістрів (в кубітах) та необхідне для їх реалізації математичне та програмне забезпечення, Вважається, що у постквантовий період квантові комп’ютери будуть працювати з суттєво збільшеною надійністю, тобто з виправленням помилок.

Мета статті – розроблення науково-методичних основ аналізу, оцінки та порівняння існуючих і перспективних (постквантових) асиметричних криптографічних примітивів АСШ, ПІК та ЕП, а також аналіз отриманих практичних результатів.

1. Аналіз стану застосування квантових комп’ютерів для криптоаналізу

Світова цивілізація робить суттєві кроки в науці та практиці, що пов’язані з квантовими обчисленнями; кроки щодо досягнення конкурентної переваги країн в галузі квантової інформаційної науки та практики впровадження квантових технологій. Наукові та практичні дослідження спрямовані перше за все на зниження ризиків, що пов’язані з квантовими комп’ютерами щодо кібербезпеки, економічної та національної безпеки. Визначаються

конкретні дії, які мають зробити технологічно розвинені держави, що розпочинають багаторічний процес переведення вразливих комп'ютерних систем на квантово-стійку криптографію [1, 2, 12, 13]. Здійснюються спроби стимулювання інновацій економіки у сферах від матеріалознавства та фармацевтики до фінансів та енергетики, кібербезпеки тощо. Хоча повний спектр застосування квантових комп'ютерів ще невідомий, проте очевидно, що подальше технологічне та наукове лідерство держав, принаймні частково, буде залежати від здатності країни підтримувати конкурентну перевагу в галузі квантових обчислень та квантової інформаційної науки [1, 2, 6 – 13].

У зв'язку з прийняттям Меморандуму в галузі квантових обчислень Президент США надав півроку на перехід усіх державних органів на постквантову криптографію. Вважається, що квантові обчислення також будуть становити для економічної та національної безпеки США значні ризики. Зокрема, квантовий комп'ютер достатнього розміру та складності – також відомий як криптоаналітично значущий квантовий комп'ютер (КЗКК) – буде здатний зламати більшу частину існуючих стандартизованих криптографічних перетворень з відкритим ключем – АСШ, ПІК та ЕП [1 – 3], що використовується, наприклад, у цифрових системах США та всього світу. Коли він стане доступним для використання, то зможе поставити під загрозу цивільні та військові комунікації, підірвати системи нагляду та контролю за критичною інфраструктурою, а також зруйнувати протоколи безпеки для більшості фінансових операцій в Інтернеті. Переваги та недоліки фізичних реалізацій квантових комп'ютерів:

• масштабованість, тобто можливість створення та управління все більшими і більшими квантовими пристроями зі все більшою кількістю кубітів з використанням фізичних/інженерних ресурсів та керування ними;

• сумісність з різними обчислювальними моделями та простота їх реалізації;

• типовий час декогерентності (тобто скільки часу залишаються збереженими характеристики та вони використані в працездатному стані, а також можуть бути використані квантові особливості, такі як суперпозиції);

• швидкість і точність, з якою вентилі можуть бути застосовано.

Потрібно згодитись, що є проблеми стосовно обґрунтованого вибору фізичної реалізації квантових комп'ютерів.

Високорівневу класифікацію перспективних фізичних реалізацій квантових комп'ютерів можна представити у такому складі [1, 7, 11, 12]:

• квантова оптика, коли інформація зберігається та захищається в станах квантів світла на основі поляризації або в станах з певним числом фотонів, та може бути реалізована в чіпі за допомогою інтегрованої оптики;

• надпровідні системи, коли інформація зберігається та обробляється (захищається) в електричних ланцюгах, які використовують властивості надпровідних матеріалів;

• топологічні системи, коли інформація зберігається та захищається з використанням деяких топологічних властивостей, тобто властивостей, які залежать від «глобальних» (геометричних) властивостей, нечутливих до «локальних» змін – квантових систем;

• іонні пастки, коли інформація зберігається (захищається) та маніпулюється з використанням властивостей іонів (атомів із незникаючим повним електричним зарядом), які обмежені електромагнітними полями;

• квантові спінові системи, коли інформація зберігається та захищається (маніпулюється) у внутрішньому ступені свободи, який називається квантовим спіном. Такі системи можуть бути реалізовані в кремнії як стандартні мікрочіпи, або в менш звичайних системах як алмази з точковими дефектами, відомі як азотно-заміщена (коротше NV) вакансія;

• гази холодних атомів, де нейтральні атоми (а не іони) охолоджуються до значення близького до абсолютного нуля. У той час як іони відштовхуються один від одного через свій електричний заряд, нейтральні атоми цього не роблять, і можуть бути захоплені і організовані в дуже регулярні масиви за допомогою лазерних променів, що створюють так звані оптичні решітки. Атомами можна керувати аж до рівня окремих ділянок в решітці.

Стан створення квантового комп'ютера та можливості вирішення задач криптоаналізу щодо асиметричних криптоперетворень АСШ, ПІК та ЕП можна оцінити наступним чином [1, 7 – 11]:

- IBM розробила та представила квантовий 127-кубітний процесор Eagle. Він прийшов на зміну 65-кубітному квантовому процесору Hummingbird, що відповідає дорожній карті квантових технологій IBM;
- є відомості про наміри IBM представити 433-кубітний процесор Osprey в 2022 р., а 1121-кубітний процесор Condor – в 2023 р.;
- IBM повідомляє про наміри щодо побудови на основі покращених чіпів нової інтегрованої квантової обчислювальної системи IBM Quantum System Two замість існуючої системи IBM Quantum System One;
- компанія D-Wave, що відома розробками псевдоквантових (гібридних) комп'ютерів з великою загальною кількістю кубітів (понад 2000 та понад 5000 кубітів сьогодні), повідомила про наміри представити машину понад 7000 кубітів (2023 – 2024 рр.).

Наведеним даним можна довіряти, але зрозуміло, що фактичний стан розроблення та застосування потужних квантових комп'ютерів є закритим. З іншої сторони, зрозуміло, що створення квантових комп'ютерів здійснюється в умовах суттєвих інвестицій, з випередженням розробки математичних, логічних та програмних основ.

Певний досвід також свідчить. що потрібні інші підходи до підготовки спеціалістів математиків, фізиків, алгоритмістів та програмістів тощо. Навіть використовуючи класичну математику для розробки криптографічно стійких постквантових стандартів, необхідно застосовувати інші математичні методи для асиметричних криптографічних перетворень. Перелік таких методів може зводитися до використання математичних решіток та математичних кодів, математики ізогеній еліптичних кривих, криптографічних перетворень в квадратичних полях (багатовимірних перетворень в квадратичних полях) тощо.

До основних задач криптоаналізу, які можуть бути вирішені на квантовому комп'ютері, необхідно віднести такі [7 – 9]:

- квантовий алгоритм факторизації Шора;
- квантовий алгоритм Гровера пошуку елемента в несортованій базі;
- квантовий алгоритм Шора для розв'язку дискретного логарифму в скінченному полі;
- квантовий алгоритм розв'язку дискретного логарифму в групі точок ЕК Шора тощо.

Окрім класичної криптографії, яка базується на математичних алгоритмах, нині активно розвивається квантова криптографія. Створення квантових комп'ютерів відкриє принципово нові можливості для людства, але при цьому існуючі методи захисту інформації втратять свою ефективність. Не дивлячись на те, що квантові комп'ютери тільки виходять за межі лабораторій, необхідність у використанні квантово безпечної або, як її ще називають, постквантової криптографії є уже сьогодні.

Безпека сучасних інформаційних систем та технологій ґрунтується на стійкості криптографічних перетворень, які використовуються при криптографічному захисті інформації (КЗІ). Криптографічна стійкість КЗІ базується на складності розв'язку певних математичних задач (факторизації великого цілого числа, розв'язку дискретного алгоритму тощо), для таких задач характерна суб'експоненційна або експоненційна складність розв'язку вказаних задач на сучасних (класичних) комп'ютерах. Проте, використовуючи квантові алгоритми Шора та Гровера, певні математичні задачі можна розв'язувати навіть з поліноміальною складністю.

Ідеї використання потужностей квантового середовища висунули Пауль і Фейман. Важливим стало розроблення у 1992 р. Дойчем та іншими першого квантового алгоритму, можливості якого значно перевищували можливості звичайних комп'ютерів. У випадку появи квантового комп'ютера, на якому може бути запущений квантовий алгоритм криптоаналізу Шора або алгоритм пошуку в неупорядкованій базі даних Гровера, можуть виникнути великі загрози у інформаційній сфері відносно забезпечення криптографічної стійкості як для

асиметричних криптоперетворень, так і для певних симетричних. Важливим є не тільки сам факт побудови такого комп'ютера, а й технічні характеристики, якими володітиме квантовий комп'ютер.

В квантовому комп'ютері ключ шифрування передається за допомогою елементарних часток світла – фотонів, тобто внаслідок реалізації квантового протоколу розподілення ключа. Будь-який прилад, за допомогою якого третя сторона спробує перехопити дані, вплине на стан фотона, і ключ стане недійсним. Передавати фотони можна за допомогою виділених оптоволоконних ліній; на обох кінцях такої лінії необхідні спеціальні шифрувальні пристрої. Недоліком квантової криптографії є те, що вона вимагає великих затрат на інфраструктуру. При цьому поки що ключі вдається передавати реально на відстань (50 – 100 км), швидкість їх генерації досить низька, а на передачу фотонів впливає маса зовнішніх факторів.

Постквантова (квантово безпечна) криптографія, як і класична, основана на розв'язку математичних задач. Проте нові алгоритми шифрування повинні бути іншими, щоб їх не могли розв'язати за допустимий час не тільки звичайні, але і квантові комп'ютери.

2. Аналіз стану застосування квантових комп'ютерів для криптоаналізу

Стосовно криптології до основних задач, які можуть бути вирішені на квантовому комп'ютері для криптоаналізу існуючих стандартизованих засобів КЗІ, необхідно, в першу чергу, віднести квантові алгоритми Шора (факторизації, розв'язку дискретного логарифму в скінченному полі та дискретного логарифму в групі точок еліптичних кривих (ЕК)), а також квантовий алгоритм Гровера пошуку елемента в несортованій базі тощо. Нижче наведено їх сутність та складність реалізації.

2.1. Квантовий алгоритм факторизації Шора

Квантовий алгоритм запропонований одним з перших для вирішення задачі факторизації модуля криптоперетворення в кільці, наприклад RSA криптоперетворення. Вирішення вказаної задачі зводиться до факторизації модуля перетворення N, а класичні алгоритми факторизації мають або експоненційну, або суб'експоненційну складність. При цьому вважається, що найкращим за критерієм мінімуму складності факторизації є алгоритм загального решета числового поля та при деяких обмеженнях його модифікації – спеціальні решета числового поля. У той же час алгоритм Шора, що орієнтований на квантовий комп'ютер, має поліноміальну складність. При його застосуванні факторизацію можна здійснити зі складністю

$$O(n^3) \qquad (1)$$

та з використанням $O(n)$ кубітів.

Порівняльний аналіз складності факторизації для класичного та квантового алгоритмів наведено у табл. 1.

Таблиця 1

Порівняльний аналіз класичного та квантового алгоритмів факторизації (RSA)

Розмір модуля N, бітів	Кількість необхідних кубітів $2n$	Складність квантового алгоритму $4n^3$	Складність класичного алгоритму
512	1024	$0.54 \cdot 10^9$	$1.6 \cdot 10^{19}$
3072	6144	$12 \cdot 10^{10}$	$5 \cdot 10^{41}$
15360	30720	$1.5 \cdot 10^{13}$	$9.2 \cdot 10^{80}$

Аналіз даних табл. 1 показує, що для зламу RSA криптосистеми з розміром модуля у 15360 бітів (а це розмір відкритого ключа сертифікату США), необхідно лише $1.5 \cdot 10^{13}$ операцій на квантовому комп'ютері, тоді як з використанням існуючих класичних обчислювальних систем потрібно виконати приблизно 1080 операцій.

Таким чином, якщо з'явиться квантовий комп'ютер з відповідними характеристиками та параметрами, RSA система буде зламана за поліноміальний час.

2.2. Квантовий алгоритм Шора дискретного логарифмування в скінченному полі

Існує декілька методів та алгоритмів дискретного логарифмування в скінченному полі [19]. Вважається, що найкращим класичним алгоритмом дискретного логарифмування в скінченному полі є метод решета числового поля. Для нього складність можна оцінити як суб'експоненційну [19]:

$$O(exp(3^{3/2}(ln(P)ln(ln(P)))^{1/3}). \quad (2)$$

Як слідує із вказаних джерел, класичні методи криптоаналізу дискретного логарифмування в скінченному полі мають суб'експоненційну складність. Алгоритм Шора має поліноміальну складність дискретного логарифмування, що дозволяє вирішити проблему дискретного логарифму в скінченному полі з суттєво меншою складністю [15]. Порівняльний аналіз складності алгоритму дискретного логарифмування в скінченному полі на основі решета числового поля та алгоритму Шора наведено в табл. 2.

Таблиця 2

Порівняльний аналіз класичного і квантового алгоритму дискретного логарифмування в скінченному полі

Розмір модуля перетворення (бітів)	Кількість необхідних кубітів $\approx 3n$	Час квантового алгоритму $\approx n^3$	Час класичного алгоритму
1024	3072	$0.1 \cdot 10^{10}$	$3.3 \cdot 10^{20}$
3072	9216	$2.9 \cdot 10^{10}$	$1.4 \cdot 10^{31}$
15360	46080	$3.6 \cdot 10^{12}$	$5.9 \cdot 10^{56}$

Аналіз даних табл. 2 дозволяє зробити висновок, що збільшення розміру модуля перетворення і, відповідно особистого ключа, при застосуванні квантового алгоритму Шора не забезпечує необхідного збільшення складності дискретного логарифмування в скінченному полі, як при зламі ЕП, так і АСШ (ПІК). Наприклад, для модуля $P \geq 2^{3072}$ складність дискретного логарифмування в скінченному полі складає $1.4 \cdot 10^{31}$, а із застосуванням алгоритму Шора – всього $2.9 \cdot 10^{10}$ операцій. Але, в той же час, при застосуванні квантового алгоритму проблемною є реалізація регістрів зі значним числом кубітів – не менше 9216 кубітів. Очевидно, досягти такого розміру буде ще певний час проблемою.

2.3. Квантовий алгоритм Шора дискретного логарифмування в групі точок еліптичних кривих

Квантовий алгоритм Шора дискретного логарифмування в групі точок еліптичних кривих суттєво схожий за складністю зі складністю алгоритму Шора дискретного логарифмування в скінченному полі. Так, алгоритм Шора для групи точок ЕК має однакові кроки, відмінність в поданні; замість елементів поля потрібно розглядати точки ЕК. Розглянемо їх криптографічну стійкість та зробимо відповідні оцінки для них.

Вважається, що задачі дискретного логарифмування в групі точок еліптичних кривих найбільш ефективно можуть бути вирішені з використанням ρ- та λ-методів Полларда [20, 21]. Для них складність можна оцінити як

$$O(\sqrt{q}), \quad (3)$$

де q – число точок еліптичної кривої.

Визначено, що квантовий алгоритм Шора у загальному випадку має поліноміальну складність вирішення такого класу задач. Він також може бути застосований для розв'язку дискретного логарифмічного рівняння, причому його часова складність може бути оцінена як $O(n^3)$, де n – порядок базової точки ЕК. Деякі оцінки та результати порівняльного аналізу класичних алгоритмів та квантового алгоритму Шора наведено у табл. 3.

Таблиця 3

Порівняльний аналіз складності класичного і квантового алгоритмів дискретного логарифмування групі точок еліптичної кривої (ECC)

Алгоритм розв'язку дискретного логарифмічного рівняння			
Розмір порядку базової точки, бітів	Кількість необхідних кубітів $f(n)=7n+4\log_2 n+10$	Складність квантового алгоритму $360n^3$	Складність класичного алгоритму
163	1210	$1.6\cdot10^9$	$3.4\cdot10^{24}$
256	1834	$6\cdot10^9$	$3.4\cdot10^{38}$
571	4044	$6.7\cdot10^{10}$	$8.8\cdot10^{85}$
1024	7218	$3.8\cdot10^{11}$	$1.3\cdot10^{154}$

Аналіз даних табл. 3 дозволяє зробити висновок, що збільшення розміру порядку базової точки при криптоаналізі з використанням квантового алгоритму не дає суттєвого збільшення криптографічної стійкості криптографічної системи на еліптичних кривих. Також видно, що при збільшенні модуля складність дискретного логарифмування класичними методами в групі точок еліптичної кривої зі збільшенням порядку базової точки збільшується суттєво. Але потрібно взяти до уваги, що реалізація квантового алгоритму пов'язана із застосуванням регістрів з великою кількістю кубітів, яка необхідна для проведення квантової атаки. Наприклад, для базової точки з порядком 2^{571} необхідно використовувати реєстр з довжиною 4016 кубітів. Вважається, що така велика кількість кубітів певний час не може бути реалізована з необхідною надійністю його функціонування.

2.4. Квантовий алгоритм Гровера та його використання

Проблему криптоаналізу, на вирішення якої спрямовано метод Гровера, може бути сформульовано наступним чином. Нехай дано неупорядковану базу даних (список) з N елементів і нехай в ній існує один елемент, що володіє деякою властивістю, яка перевіряється з поліноміальною складністю. Потрібно знайти цей елемент із мінімально можливою складністю і, зрозуміло, за менший час.

Для пошуку скористаємося математичним апаратом узагальненого «парадоксу про день народження». Основними умовами застосування цієї моделі є випадковість та рівноймовірність здійснення запитів, тобто вхідних даних. Тому при виконанні k запитів ймовірностей успіху можна оцінити як k/N [20, 21]. Значить, щоб знайти необхідний елемент з будь-якою наперед заданою ймовірністю, необхідно зробити до бази $O(N)$ запитів. Алгоритм Гровера якраз і дозволяє знайти необхідний елемент з ймовірністю достатньо близькою до 1 за

$$O(\sqrt{N}) \tag{4}$$

кроків, кожен з яких є ітерацією при виконанні процедури. В цілому квантовий комп'ютер дозволяє вирішити цю задачу за

$$O(\sqrt{N}\log N) \tag{5}$$

кроків, використовуючи $\log N$ кубітів, причому $\log N$ кроків необхідно для виконання перетворення Уолша – Адамара [4, 6].

Вирішення цієї задачі може бути виконане з використанням декількох класичних алгоритмів, в яких для підвищення ймовірності успіху процедура повторюється багатократно. При цьому при повторенні такої квантової процедури ймовірність успіху, як правило, збільшується, але після достатньої великої кількості повторень результат знову стає гіршим. Вказане пояснюється тим, що квантова процедура це унітарне перетворення, яке здійснює поворот в комплексному просторі. Внаслідок цього застосування квантового перетворення спочатку, протягом якогось числа ітерацій, можна наближати поточний стан все ближче і ближче до потрібного нам стану, але в подальшому застосування квантового перетворення може оминути потрібний стан і тому віддалити правильне рішення. Для того щоб отримати

при повторюваних квантових перетвореннях результат, що очікується, дуже важливо визначити, коли потрібно зупинитися і провести уточнення.

Наприклад, використовуючи алгоритм Гровера, можна знайти секретний ключ симетричного шифрування чи гешування за $\sqrt{N}$ ітерацій, де N – розмір простору ключів. У якості прикладу в табл. 4 наведено оцінки стійкості симетричних криптографічних систем проти квантового криптоаналізу. Аналіз даних таблиці показує, що стійкість симетричних шифрів при атаці з використанням квантового алгоритму Гровера суттєво зменшується.

Таблиця 4

Стійкість симетричних криптосистем проти квантового криптоаналізу

Вид криптосистеми	Розмір блоку/ключа (біт)	Обсяг пам'яті, яка необхідна для здійснення атаки (блок повідомлення/ключ), кубіт	Стійкість при атаці на	
			блок повідомлення	ключ
AES-128	128/128	128/128	264 (1019,2)	264 (1019,2)
AES-256	128/256	128/256	264 (1019,2)	2128 (1038,4)
DES	64/56	64/56	232 (109,6)	228 (108,4)
TDES	64/168	64/168	232 (109,6)	2134 (1040,2)
ГОСТ-28147	64/256	64/256	232 (109,6)	2128 (1038,4)
Калина-128	128/128	128/128	264 (1019,2)	264 (1019,2)
Калина-512	512/512	512/512	2256 (1076,8)	2256 (1076,8)

Наприклад, алгоритм шифрування DES буде повністю компрометований і не можна говорити про деяку його стійкість, так як оцінка приймає значення 2^{28}. Також із таблиці видно, що навіть для AES-128 та Калина-128 з використанням квантового комп'ютера можна було б знайти секретний ключ за час приблизно 2^{64}. Але значення 2^{64} нині уже може вважатись небезпечним. Видно, що при довжині ключа 256 біт, що стосується AES-256 біт та Калини-2, часова складність роботи алгоритму Гровера становить 2^{128}, що є практично не здійсненним при нинішніх поглядах та можливостях реалізації. Дані відносно шифру ДСТУ 28157-2009 (ГОСТ 28147-89) показують, чому коротка довжина блоку в 64 бітів не може використовуватись, хоча довжина ключа для нього складає 256 бітів, оскільки стійкість на блок повідомлення складає 2^{32}, а при атаці на ключ – 2^{128}. Таким чином, алгоритм Гровера дозволяє реалізувати алгоритм узагальненого парадоксу про день народження.

3. Науково-методичні основи аналізу, оцінки та порівняння існуючих та перспективних (постквантових) стандартизованих асиметричних криптоперетворень

3.1. Призначення та застосування комплексної методики оцінки, аналізу та порівняння криптографічної стійкості та властивостей існуючих та постквантових АСШ, ПІК та ЕП

Комплексна методика безпосередньо призначена для використання при оцінці, аналізі та порівнянні криптографічної стійкості та властивостей існуючих та постквантових АСШ, ПІК та ЕП. Ця методика може бути у явному вигляді застосована при дослідженнях, оцінці та порівнянні існуючих стандартизованих та альтернативних криптографічних примітивів типу АСШ, ПІК та ЕП, включаючи перспективні криптографічні примітиви АСШ, ПІК та ЕП.

Методика аналізу існуючих та постквантових АСШ, ПІК та ЕП є комплексною та визначає три наступні методики оцінки та порівняння:

- криптографічної стійкості існуючих та постквантових АСШ, ПІК та ЕП на основі використання безумовних критеріїв стійкості;
- існуючих та постквантових АСШ, ПІК та ЕП на основі використання умовних критеріїв;
- властивостей існуючих та постквантових АСШ, ПІК та ЕП на основі прагматичних критеріїв.

Ці методики можуть застосовуватись незалежно одна від одної, але основним є застосування їх у вказаній послідовності – спочатку з використанням на основі безумовних критеріїв, потім – на основі умовних критеріїв та при необхідності на основі прагматичних критеріїв.

Методика може бути застосована:

- при обґрунтуванні та розробленні порядку застосування методик оцінки та порівняльного аналізу асиметричних існуючих та перспективних, в тому числі стандартизованих, криптографічних примітивів АСШ, ПІК та ЕП;
- обґрунтуванні та виборі критеріїв та показників оцінки криптографічної стійкості та інших властивостей, в тому числі стандартизованих криптопримітивів типу АСШ, ПІК та ЕП;
- оцінці, аналізі та порівнянні асиметричних існуючих та перспективних постквантових, в тому числі стандартизованих, криптографічних примітивів АСШ, ПІК та ЕП на основі застосування безумовних критеріїв;
- оцінці, аналізі та порівнянні асиметричних існуючих та перспективних, в тому числі стандартизованих, криптопримітивів АСШ, ПІК та ЕП на основі застосування умовних критеріїв;
- оцінці, аналізі та порівнянні асиметричних існуючих та перспективних, в тому числі стандартизованих, криптопримітивів АСШ, ПІК та ЕП на основі застосування прагматичних критеріїв;
- обґрунтуванні та виборі основних методів експертного оцінювання криптографічної стійкості та інших властивостей існуючих та постквантових, в тому числі стандартизованих, криптопримітивів типу АСШ, ПІК та ЕП;
- реалізації методу ієрархій на основі попарних порівнянь та врахуванні особливостей його застосування для оцінки та порівняння властивостей існуючих та постквантових криптографічних примітивів АСШ, ПІК та ЕП;
- обґрунтуванні та виборі для оцінки та порівняльного аналізу існуючих та перспективних криптопримітивів АСШ, ПІК та ЕП, в тому числі стандартизованих, методу вагових коефіцієнтів;
- розробці рекомендації щодо оцінки та порівняння альтернативних криптопримітивів типу АСШ, ПІК та ЕП за прагматичними техніко-економічними та техніко-експлуатаційними критеріями.

3.2. Аналіз стану розроблення та застосування науково-методичних основ комплексної методики оцінки, аналізу

Методики оцінювання та порівняльного аналізу криптопримітивів базуються на використанні системи безумовних та умовних часткових та інтегральних критеріїв, прагматичних критеріїв, а також показників, які дозволяють оцінити ступінь виконання висунутих до криптоперетворення вимог. На наш погляд, основним завданням таких методик є формалізація процесів прийняття рішень відносно виконання висунутих до них вимог, врахування переваг та недоліків криптопримітивів, що є кандидатами на постквантовий стандарт, зменшення впливу суб'єктивних факторів на прийняття рішень, в тому числі несанкціонованого впливу сторонніх організацій тощо. Наприклад, такі методики можуть бути застосованими щодо оцінки та порівняння алгоритмів АСШ, ПІК та ЕП, що є в нашому випадку кандидатами на постквантовий стандарт.

На формальному рівні такі методики оцінки та порівняння алгоритмів АСШ, ПІК та ЕП можуть бути узагальненими (базовими, комплексними). Але, оскільки до названих криптопримітивів висуваються різні вимоги, то для кожного із примітивів вони можуть доповнятись чи спрощуватись та відображати весь спектр висунутих вимог. Також такі методики можуть забезпечити прозорість прийняття рішень, незалежність експертів та допомогти обґрунтувати прийняття відповідних рішень та довіру до них. В подальшому під

методикою для наших досліджень будемо розуміти фіксовану сукупність прийомів практичної діяльності щодо аналізу криптографічної стійкості та властивостей нових, доказовостійких криптоалгоритмів і протоколів, у тому числі у перехідний та постквантовий періоди, що відповідає наведеним вище вимогам та призводить до заздалегідь визначеного результату.

3.3. Обґрунтування та вибір критеріїв та показників оцінки та порівняння рівнів існуючих та перспективних АСШ, ПІК та ЕП на основі комплексної методики

Під критерієм будемо розуміти ознаку, на основі якої здійснюється оцінка, визначення чи класифікація чого-небудь, тобто, будемо розуміти мірило оцінки. Наші попередні дослідження дозволили зробити висновок, що порівняння криптопримітивів можна здійснити з використанням двох сукупностей критеріїв: безумовних та умовних. Такий підхід дозволяє зробити оцінку та порівняння криптопримітивів, що є кандидатами на постквантові стандарти, за частковими та інтегральним умовним критерієм. Такий підхід ґрунтується, в тому числі, і на врахуванні чи використанні експертних оцінок.

На першому етапі перевіряється відповідність криптопримітиву системі часткових безумовних критеріїв, а потім для кожного криптопримітиву на основі часткових обчислюється безумовний інтегральний критерій.

На другому етапі отримуються відповідні оцінки з використанням спочатку системи часткових умовних критеріїв, а потім на їх основі обчислюється інтегральний умовний критерій.

На третьому етапі отримуються відповідні оцінки з використанням системи прагматичних критеріїв.

Такий підхід дозволяє відкинути криптоперетворення, що не відповідають безумовним вимогам, тобто вимогам, які повинні бути виконані безумовно. Причому інтегральний безумовний критерій дозволяє прийняти рішення відносно кожного із криптопримітивів. У нашому випадку це різні криптопримітиви АСШ, ПІК та ЕП.

Застосування часткових умовних критеріїв, а потім на їх основі інтегрального умовного критерію, дозволяє оцінити якість криптопримітиву у широкому сенсі як якість у середньому, а потім і порівняти криптопримітиви, що є кандидатами на постквантовий алгоритм.

До безумовних критеріїв будемо відносити ті, виконання яких для криптопримітиву є обов'язковим. Причому, на наш погляд, для асиметричних криптоперетворень типу АСШ, ПІК та ЕП можна вибрати однакову систему безумовних критеріїв. Але це не виключає можливостей врахування особливостей вимог та відповідно вибору при аналізі та оцінці криптопримітивів додаткових часткових безумовних критеріїв. Розглянемо та виберемо систему часткових безумовних критеріїв, орієнтуючись на вимоги NIST та певні національні нормативно-правові документи.

До безумовних критеріїв оцінки ПІК можна віднести:

1. Практично реалізований рівень моделі безпеки IK-CPA/CCA2.
2. Криптостійкість (складність криптоаналізу) щодо криптоперетворення ЕП – $W_{ЕП}$, що застосовуються в протоколі ПІК.
3. Криптостійкість (складність криптоаналізу) щодо криптоперетворення інкапсуляції – $W_{ПІК}$, та АСШ – $W_{АСШ}$, що застосовуються в протоколі інкапсуляції та асиметричного шифрування.
4. Криптоживучість ключів щодо криптоперетворення АСШ – $G_{АСШ}$ та АСШ – $W_{АСШ}$ (ЕП – $G_{ЕП}$), що застосовуються в протоколі ПІК (ЕП).
5. Криптоживучість ключів, що застосовуються в протоколі ПІК – $G_{ПІК}$.
6. Захищеність криптопротоколу від раніше переданих повідомлень – $W_{рпп}$.
7. Неспростовності криптоперетворень АСШ – $N_{АСШ}$, що встановлені для криптографічного захисту.

8. Неспростовності криптоперетворень ЕП – $N_{ЕП}$, що встановлені для криптографічного захисту.

9. Новизну ключів АСШ (ЕП) – $W_{кл.}$, що застосовуються в протоколі інкапсуляції ПІК та для АСШ (в кращому випадку використання ключів сеансу).

10. Характеристику ступеню нерозрізнюваності для ключів АСШ, ПІК та ЕП.

Аналіз вимог, що висунуті NIST до часткових безумовних критеріїв асиметричних криптоперетворень типу АСШ, ПІК та ЕП, наш досвід розробки й оцінки властивостей криптоперетворень типу АСШ, ПІК та ЕП тощо, досягнуті результати при практичному розв'язку задач криптоаналізу, в тому числі на основі реалізації алгоритмів квантового криптоаналізу, дозволяють вибрати щонайменше безумовні критерії оцінки АСШ, ПІК та ЕП (в табл. 5).

3.4. Безумовні критерії оцінки криптографічних примітивів

У табл. 5 наведено безумовні критерії оцінки та порівняння АСШ, ПІК та ЕП.

У процесі досліджень число безумовних критеріїв може бути розширеним до переліку, що наведений нижче:

1. Надійність, простота та прозорість математичної бази, що застосовується для АСШ, ПІК та ЕП при криптоперетвореннях. Тобто практична відсутність в порушника можливостей здійснювати відносно ЕП атаки типу «універсальне розкриття» за рахунок недосконалості математичного апарату, що застосовується, чи слабкостей, що можуть бути закладені за рахунок специфічних властивостей загальних параметрів і ключів. При цьому критерієм оцінки надійності математичної бази є той факт, що складність атаки «універсальне розкриття» $I_{ур}$ має експоненційний характер, а критерій ненадійності – суб'експоненційну або поліноміальну складність.

2. Практична захищеність криптоперетворень типу АСШ та ПІК при реалізації алгоритму «семантично безпечного шифрування» від відомих класичних та постквантових атак щодо криптоперетворень АСШ та ПІК, доступу криптоаналітика до 2^{64} обраних шифртекстів, але для моделі безпеки IND-CCA2.

3. Реальна захищеність АСШ, ПІК та ЕП від усіх відомих та потенційно можливих криптоаналітичних атак постквантового періоду. Під захищеністю розуміється той факт, що всі відомі криптоаналітичні атаки типу «повне розкриття» мають експоненційну складність $I_{ес}$, а під критерієм незахищеності – суб'експоненційний $I_{се}$ і нижче характер складності атаки «повне розкриття».

4. Теоретична захищеність криптоперетворень типу АСШ, ПІК та ЕП в постквантовий період проти існуючих силових, аналітичних та спеціальних атак для діючих моделей загроз (мінімум модель EUF-CMA для ЕП та IND-CCA2 для АСШ та ПІК) та складність яких менша, ніж складність атаки типу «повне розкриття».

5. Можливість заміни існуючих стандартизованих криптопримітивів на постквантові АСШ, ПІК та ЕП та застосування в діючих криптосистемах та протоколах в певних умовах та обмеженнях.

6. Статистична безпечність криптоперетворення типу АСШ, ПІК та ЕП. Тобто статистична незалежність результату криптоперетворення (виходу), наприклад АСШ та ПІК (ЕП), від вхідного блоку, що зашифровується (підписується), та особистого ключа, що використовується.

7. Відсутність слабких особистих ключів криптоперетворення типу АСШ та ПІК (ЕП), за яких складність криптоаналітичних атак типу «повне розкриття» та «універсальне розкриття» є меншою, ніж складність атаки «повне розкриття» для інших (не слабких) особистих ключів.

8. Обчислювальна ефективність – складність прямого $I_{пр}$ та зворотного $I_{зв}$ криптоперетворень АСШ, ПІК та ЕП (а також генерування та згортання асиметричних пар ключів) має не вище за поліноміальний характер, а також забезпечення необхідних значень складності

(швидкодії) $I_{пр}$, $I_{зв}$ та $I_{кл}$ при практичному застосуванні в додатках з апаратно-програмною та програмною їх реалізацією.

Таблиця 5

Безумовні критерії оцінки та порівняння АСШ, ПІК та ЕП

Безумовні критерії	Позначення
Надійність, простота та прозорість математичної бази (математичних перетворень), що застосовуються при реалізації постквантових криптоперетворень АСШ, ПІК та ЕП.	W_1
Практична захищеність криптоперетворень типу АСШ та ПІК при реалізації алгоритму «семантично безпечного шифрування» від відомих атак з використанням квантового комп'ютера та доступу криптоаналітика до 2^{64} обраних шифртекстів, але для моделі безпеки IND-CCA2.	W_2
Практична захищеність криптоперетворення типу ЕП від відомих атак з використанням квантового комп'ютера та доступу криптоаналітика до 2^{64} обраних повідомлень, для моделі безпеки EUF-CMA.	W_3
Обґрунтованість реальної стійкості криптоперетворень АСШ, ПІК та ЕП від усіх відомих та потенційно можливих криптоаналітичних атак постквантового періоду на основі використання загальних параметрів та ключів з необхідними розмірами та властивостями (ключі 128 біт квантової безпеки та 256 біт і більше класичної стійкості (безпеки)), включаючи статистичну безпеку.	W_4
Теоретична захищеність криптоперетворень типу АСШ, ПІК та ЕП в постквантовий період проти існуючих силових, аналітичних та спеціальних атак для діючих моделей загроз (мінімум для моделі EUF-CMA для ЕП та IND-CCA2 для АСШ).	W_5
Можливість заміни існуючих стандартизованих криптопримітивів на постквантові та застосування в діючих криптосистемах та протоколах в певних умовах та обмеженнях.	W_6
Обчислювальна ефективність – складність прямого $I_{пр.}$ та зворотного $I_{зв.}$ криптоперетворень АСШ, ПІК та ЕП, а також генерування асиметричних пар ключів $I_{кл.}$ не вище за поліноміальну складність, забезпечення необхідних значень складності (швидкодії) $I_{пр.}$, $I_{зв.}$ та $I_{кл.}$ при практичному застосуванні в додатках з апаратно-програмною та програмною реалізацією.	W_7
Виконання обмежень на мінімальну та максимальну довжини особистих та відкритих ключів, розміри та збитковість шифртексту та ЕП, відсутність слабких особистих ключів для моделей безпеки постквантового періоду EUF-CMA для ЕП та IND-CCA2 для АСШ.	W_8
Обґрунтованість реальної стійкості криптоперетворень АСШ, ПІК та ЕП від усіх відомих та потенційно можливих криптоаналітичних атак постквантового періоду на основі використання загальних параметрів та ключів з необхідними розмірами та властивостями (довжини ключів 256/128, 384/192 та 512/256 біт відповідно класичної стійкості та квантової безпеки (стійкості)).	W_9
Забезпечення захисту від атак на основі сторонніх каналів (наприклад, витоку технічними каналами) та на основі помилок.	W_{10}

9. Виконання обмежень на мінімальну та максимальну довжини особистих та відкритих ключів, розміри та збитковість шифртексту та ЕП, відсутність слабких особистих ключів для моделей безпеки постквантового періоду EUF-CMA для ЕП та IND-CCA2 для АСШ та ПІК.

10. Забезпечення захисту від атак на основі сторонніх каналів (наприклад, вимір складності криптоперетворення, вимір потужності для криптоперетворення, витоку технічними каналами тощо).

11. Забезпечення захисту від атак на основі помилок (наприклад, внесення помилок в процеси прямих та зворотних криптоперетворень, ключів тощо).

Визначається перелік інших безумовних критеріїв, необхідних для дослідження для оцінки АСШ, ПІК та ЕП, наприклад, таких:

$I_{ст.}$ – рівень криптографічної стійкості з використанням безумовних критеріїв;

$l_{в.к.}$ – можливі довжини відкритого ключа;

$l_{о.к.}$ – можливі довжини особистого (секретного) ключа;

$l_{рез.}$ – довжина результату криптоперетворення (збитковість);

$T_{пр.}$ – складність (швидкість) прямого криптоперетворення;

$T_{зв.}$ – складність (швидкість) зворотного криптоперетворення;

$T_{ген.зп.}$ – складність (швидкість) генерування загальних параметрів для відповідного режиму роботи криптоперетворення (у залежності від довжин загальних параметрів та ключів);

$T_{ген.кл.}$ – складність (швидкість) генерування ключа (ключової пари) у залежності від режиму роботи тощо.

З урахуванням наведених вище часткових безумовних критеріїв

$$W_1, W_2, W_3, W_4, W_5, W_6, W_7, W_8, W_9, W_{10}, \quad (6)$$

що наведені в табл. 5, та умови (6) функцію відповідності криптоперетворення вимогам, що викладені вище, запишемо у вигляді інтегрального безумовного критерію:

$$f() = (W_1 \wedge W_2 \wedge W_3 \wedge W_4 \wedge W_5 \wedge W_6 \wedge WW_7 \wedge W_8 \wedge W_9 \wedge W_{10}) \in (1,0), \quad (7)$$

де символ «∧» позначає операцію згідно з (7) кон'юнкції булевих змінних.

Тобто, якість постквантового криптоперетворення АСШ, ПІК та ЕП може бути оцінена з використанням безумовного інтегрального критерію – функції відповідності у вигляді інтегрального безумовного критерію

$$f_{фв}=1, \quad (8)$$

якщо криптоперетворення АСШ, ПІК та ЕП відповідає висунутим вимогам та

$$f_{фв}=0, \quad (9)$$

якщо криптоперетворення АСШ, ПІК та ЕП не відповідає висунутим вимогам.

3.5. Умовні критерії оцінки криптографічних примітивів

Якісну й кількісну оцінку та порівняння криптоперетворень типу АСШ, ПІК та ЕП рекомендується здійснювати з використанням часткових умовних та узагальненого умовного критерію переваги. Розглянемо та визначимо вказані умовні критерії на прикладі АСШ, ПІК та ЕП. У табл. 6 наведено перелік та позначення часткових умовних критеріїв оцінки криптоперетворень типу АСШ, ПІК та ЕП, вимоги до яких висунуті NIST США та ETSI ЄС. Як основні складові узагальненого критерію переваги пропонується використовувати часткові умовні критерії згідно з табл. 6. Ці часткові критерії не є обов'язковими, вони можуть бути змінені, замінені, розширений їх перелік чи взагалі, у залежності від вимог та моделей загроз тощо, відкинуті.

Потрібно підкреслити, що інтегральний умовний критерій є усередним певним чином значенням і ґрунтується на методах експертних оцінок, які розглядаються та обґрунтовуються нижче.

Таблиця 6

Умовні критерії оцінки АСШ, ПІК та ЕП

Умовні критерії	Позначення
Додаткові властивості безпеки: «perfect forward secrecy» (удосконалена пряма безпека); стійкість до атак сторонніми каналами; стійкість до мультиключових атак; стійкість до відмов.	К1
Вимоги до безпеки (стійкості): 1) 128 біт класичної безпеки / 64 біт квантової безпеки (запас стійкості AES-128); 2) 128 біт класичної безпеки / 80 біт квантової безпеки (запас стійкості SHA-256/SHA3-256); 3) 192 біт класичної безпеки / 96 біт квантової безпеки (запас стійкості AES-192); 4) 192 біт класичної безпеки / 128 біт квантової безпеки (запас стійкості SHA-384/SHA3-384); 5) 256 біт класичної безпеки /128 біт квантової безпеки (запас стійкості SHA2-512, SHA3-512).	К2
Додаткові вимоги до стійкості: 1) 512 біт класичної безпеки / 256 біт квантової безпеки (запас стійкості SHA-512/SHA3-512, ДСТУ 7564:2014 – 512 біт); 2) 512 біт класичної безпеки /від 128 до 256 біт квантової безпеки (запас стійкості ДСТУ 7624:2014 (Калина – 512)); 3) 512 біт класичної безпеки / 256 біт квантової безпеки (запас стійкості ДСТУ 7624:2014 (Калина – 512)).	К3

Помилки шифрування. Низький відсоток помилок шифрування ПІК та ЕП.	К4
Можливість багаторазового АСШ, ПІК та ЕП.	К5
Гнучкість: 1) додаткові можливості схеми (оптимізація, неявний обмін ключами тощо); 2) кросплатформеність; 3) можливість розпаралелювання.	К6
Перевірка на коректність. Перевірка правильності опорних та оптимізованих реалізацій.	К7
Перевірка на ефективність: Обчислення часу, що необхідний для генерації ключа, зашифрування, розшифрування, підпису, перевірки підпису, або встановлення ключів (тестування проводиться на оптимізованих версіях).	К8
Умови випробувань: Основні платформи: NIST PQC Reference Platform; Intel x64; Windows або Linux, компілятор GCC. Проведення додаткових тестувань інших умов (8-бітових процесорів, цифрових сигнальних процесорів, виділених CMOS, тощо).	К9
Можливість і умови вільного поширення постквантових криптоперетворень АСШ, ПІК та ЕП.	К10
Рівень довіри до постквантових криптоперетворень АСШ, ПІК та ЕП на різних рівнях застосування.	К11
Перспективність та виправданість застосування постквантових криптоперетворень АСШ, ПІК та ЕП.	К12

3.6. Прагматичні критерії та особливості їх застосування у комплексній методиці

У цьому підрозділі наводиться методика оцінювання за прагматичними критеріями та результати дослідження перспективних стандартизованих криптоперетворень. Вона є третім етапом комплексної методики. Сутність комплексної методики та її третього етапу полягає у наступному.

Відповідно до комплексної методики на перших двох етапах застосовуються умовні методики на основі застосування безумовних та умовних критеріїв. Тобто, на першому етапі спочатку оцінюються та перевіряються криптопримітиви на відповідність системі часткових безумовних критеріїв та на їх основі обчислюється безумовний інтегральний критерій. На другому етапі отримуються оцінки з використанням часткових умовних критеріїв і на їх основі обчислюється інтегральний умовний критерій.

На третьому етапі у залежності від вимог, що висуваються до криптопримітивів, при необхідності потрібно оцінювати та порівнювати альтернативні примітиви за техніко-економічними та техніко-експлуатаційними критеріями (характеристиками). В якості основних рекомендується використовувати такі прагматичні критерії (характеристики) як довжини особистих та відкритих ключів, довжини електронних підписів та довжини блоків, що шифруються, складність (швидкодію) основних прямих та зворотних криптоперетворень ЕП (АСШ, ПІК) тощо, складність генерування (обчислення) ключів та параметрів, а також їх взаємну залежність, у тому числі і у залежності від показників щодо криптостійкості та розмірів параметрів і ключів, а також видів математичних методів, що використовуються для реалізації криптопримітивів тощо. Таким чином, важливістю третього етапу є те, що на ньому здійснюється перевірка відповідності часткових безумовних та умовних критеріїв вимогам, що висунуті щодо них відповідними нормативними документами.

Послідовність оцінювання та порівняння криптопримітивів проводиться у такій послідовності:

1. Аналізу та порівнянню підлягають тільки криптоперетворення, що успішно пройшли тестування згідно з вимогами третього етапу, тобто згідно з безумовними частковими та безумовним інтегральним критеріями.
2. Подальший аналіз проводиться з використанням умовних часткових та інтегрального умовного критеріїв щодо усіх криптопримітивів, що пройшли відбір згідно з безумовними критеріями.
3. Визначається перелік прагматичних критеріїв щодо кожного класу проєктів криптоперетворень – АСШ, ПІК та ЕП.

4. На основі, як правило, експериментальних та в меншій мірі теоретичних оцінок визначаються основні показники щодо техніко-економічних та техніко-експлуатаційних характеристик:

$I_{ст.}$ – рівень криптографічної стійкості АСШ, ПІК (ЕП);

$l_{в.к.}$ – довжина відкритого ключа АСШ, ПІК (ЕП);

$l_{о.к.}$ – довжина особистого ключа АСШ, ПІК (ЕП);

$l_{рез.}$ – довжина АСШ, ПІК (ЕП);

$T_{пр.}$ – складність (швидкість) обчислення АСШ, ПІК (ЕП);

$T_{зв.}$ – складність (швидкість) перевірки АСШ, ПІК (ЕП);

$T_{ген.зп.}$ – складність (швидкість) генерування загальних параметрів АСШ, ПІК (ЕП);

$T_{ген.кл.}$ – складність (швидкість) генерування ключа (ключової пари) АСШ, ПІК (ЕП) тощо з урахуванням особливостей.

5. На основі, як правило, експериментальних та в меншій мірі теоретичних оцінок визначаються залежності необхідних показників між собою щодо їх техніко-економічних та техніко-експлуатаційних характеристик, але з урахуванням криптографічної стійкості.

6. На основі аналізу значень показників, їх залежностей між собою та значень умовних та безумовних критеріїв, що отримані на першому та другому етапах, приймаються рішення про переваги певних кандидатів та розробляються рекомендації щодо прийняття в якості стандартів тих чи інших кандидатів, що підлягали випробовуванням.

7. Наприклад, визначаються:

- залежність довжини відкритого ключа від довжини особистого (закритого) ключа у залежності від математичного методу, який застосовується при побудові асиметричної пари ключа для АСШ, ПІК та ЕП окремо;
- залежність складності генерування відкритого ключа від складності генерування особистого ключа у залежності від математичного методу, який застосовується при побудові асиметричної пари ключа для АСШ, ПІК та ЕП окремо;
- залежність складності генерування загальних параметрів від математичного методу, який застосовується (для АСШ, ПІК та ЕП окремо);
- залежність довжини ЕП від математичного методу, який застосовується при побудові асиметричної пари ключа (для ЕП);
- залежність збитковості АСШ, ПІК від математичного методу, який застосовується (окремо для АСШ та ПІК);
- залежність наведених вище залежностей від виду реалізації (програмна, програмно-апаратна, апаратна тощо).

Очевидно, застосування прагматичної методики необхідне у випадках, коли необхідно забезпечити виконання вимог ТЗ та ТТЗ щодо розмірів ключів та параметрів, складності виконання АСШ, ПІК та ЕП, у залежності від математичного методу та розмірів загальних параметрів та ключів тощо.

Практичні приклади методик з використанням прагматичних критеріїв та показників, та їх використання наведені нижче.

3.7. Приклади критеріїв та вимог NIST.IR 8413 та IT Grundschutz Compedium

Приклади критеріїв та вимог згідно з NIST.IR 8413 [4]:

1. Відповідність моделі безпеки: для АСШ, ПІК – IND-CPA та IND-CCA2; для ЕП – EUF-CMA.
2. Відповідність наборів параметрів категоріям безпеки 1, 2, 3 та 5.
3. Гнучкість, простота та адаптація (відсутність факторів, які могли б перешкодити адаптації) криптоперетворення.
4. Стійкість до атак бічними каналами.
5. Патентна незалежність.
6. Залежність показників криптоперетворення від використовуваного процесора.
7. Розміри параметрів та основних перетворень досліджуваного криптоперетворення.

Приклади критеріїв та вимог згідно з IT Grundschutz Compendium [18]:

1. Забезпечення реалізації захисту від несанкціонованого доступу до ІТ-систем при застосуванні обраного криптографічного алгоритму.

2. Забезпечення реалізації захисту від вразливостей програмного забезпечення або помилок при застосуванні обраного криптографічного алгоритму.

3. Забезпечення запобігання неправильного використання дозволу (авторизації) при застосуванні обраного криптографічного алгоритму.

4. Забезпечення запобігання заперечення (відмови) дій при застосуванні обраного криптографічного алгоритму.

5. Забезпечення реалізації захисту від застосування зловмисного програмного забезпечення при застосуванні обраного криптографічного алгоритму.

6. Забезпечення запобігання відмови в обслуговуванні при застосуванні обраного криптографічного алгоритму.

7. Забезпечення запобігання втрати цілісності конфіденційної інформації при застосуванні обраного криптографічного алгоритму.

8. Забезпечення реалізації визначеної криптографічної концепції при застосуванні обраного криптографічного алгоритму.

9. Забезпечення захисту даних при застосуванні обраного криптографічного алгоритму.

10. Забезпечення використання хмари (при необхідності) при застосуванні обраного криптографічного алгоритму.

11. Забезпечення реалізації механізму виявлення подій, що стосуються безпеки, при застосуванні обраного криптографічного алгоритму.

3.8. Обґрунтування та вибір методів експертного оцінювання та порівняння існуючих та перспективних (постквантових) АСШ, ПІК та ЕП

Під експертними оцінками розуміють метод пошуку і результат застосування методу, що отриманий на основі використання персональної думки експерта або колективної думки групи експертів, а також комплекс логічних і математичних процедур, направлених на отримання інформації від спеціалістів, її аналіз та узагальнення з метою підготовки та вироблення раціональних рішень (рис. 1).

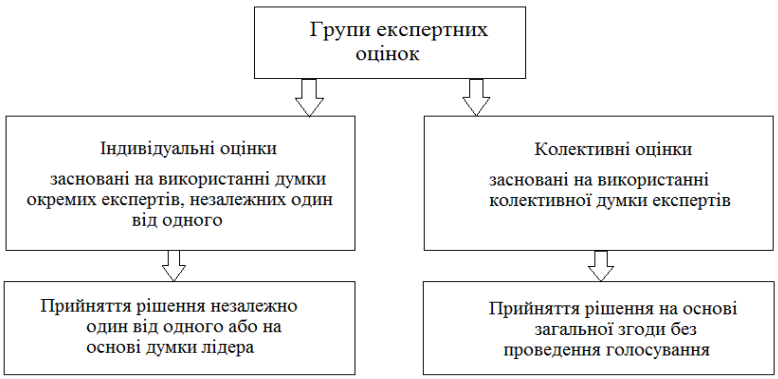


Рис. 1. Схема аналізу груп експертних оцінок

Методи експертних оцінок – це методи організації роботи зі спеціалістами-експертами та обробки думок експертів.

Сутність методів експертних оцінок полягає у покладанні думки спеціаліста або колективу спеціалістів, що заснована на їх знаннях і практичному професійному досвіді, в основу прийнятого рішення, прогнозу, висновку.

Отримали розповсюдження методи експертних оцінок, що виробляються на основі:

- колективних та індивідуальних думок експертів;
- індивідуальних думок експертів;
- колективної роботи групи експертів.

Обґрунтування, вибір та застосування методів експертних оцінок наводяться нижче.

Як правило експертне оцінювання здійснюється у такі етапи:

1) постановка мети дослідження;
2) вибір форми дослідження, визначення бюджету проєкту;
3) підготовка інформаційних матеріалів, бланків анкет, модератора процедури;
4) підбір експертів;
5) проведення експертизи;
6) аналіз результатів (обробка експертних оцінок);
7) підготовка звіту з результатами експертного оцінювання.

В цілому на основі їх аналізу можна зробити висновок, що спільна думка володіє більшою точністю, ніж індивідуальна думка кожного з фахівців. Даний метод застосовують для одержання кількісних оцінок якісних характеристик та вивчення властивостей.

4. Приклади та результати оцінки та порівняння існуючих та перспективних стандартизованих ЕП

4.1. Результати порівняння перспективних механізмів ЕП, що засновані на перетвореннях на алгебраїчних решітках

У табл. 7 наведено характеристики обраних для порівняння алгоритмів (значення швидкості криптоперетворень та генерації ключів наведено в тактах).

Таблиця 7

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_128	3	1 472	3 488	2 693	133 340	109 818	90 328
Вершина_256	5	2 624	5 792	5 345	259 103	233 712	229 669
Вершина_384	7	4 528	9 088	6762	411 040	398 029	317 324
Вершина_512	9	5 824	11 008	10708	643 744	620 989	485 471
Сокіл_128	3	897	4097	666	655 672	139 620	33 696 000
Сокіл_256	5	1 793	8193	1 280	1 338 825	285 714	107 055 000
Сокіл_512	9	3 585	5121	2 515	2 600 053	265 416	28 493 603 229

В порівнянні брали участь проєкти стандартів «Вершина» та «Сокіл», а також алгоритм Dilithium, який за попередніми дослідженнями мав кращі результати серед постквантових алгоритмів підпису, що засновані на перетвореннях на алгебраїчних решітках. Стійкість алгоритмів «Вершини» 128 біт відповідає 3-му рівню стійкості NIST, 256 – 5-му, тому пропорційно для виконання порівняння згідно зі шкалою оцінок попарного порівняння параметрам 384 був наданий 7-й рівень, а 512 – 9-й.

Таблиця 8

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_128	0,0299	0,1395	0,1816	0,0800	0,3006	0,2195	0,2696
Вершина_256	0,0697	0,0655	0,0768	0,0339	0,1583	0,0716	0,1388
Вершина_384	0,1453	0,0327	0,0407	0,0261	0,0975	0,0348	0,0797
Вершина_512	0,2681	0,0233	0,0296	0,0173	0,0479	0,0218	0,0608
Сокіл_128	0,0299	0,2487	0,1212	0,3211	0,0479	0,1466	0,0192
Сокіл _256	0,0697	0,1108	0,0467	0,1989	0,0238	0,1130	0,0146
Сокіл _512	0,2681	0,0406	0,0973	0,0984	0,0143	0,0581	0,0086

У табл. 8 наведено результати досліджень – відносна перевага алгоритмів ЕП, що отримана методом попарних порівнянь за кожною з характеристик.

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

Як видно, найбільшу перевагу має алгоритм «Вершина» з параметрами стійкості 128 біт.

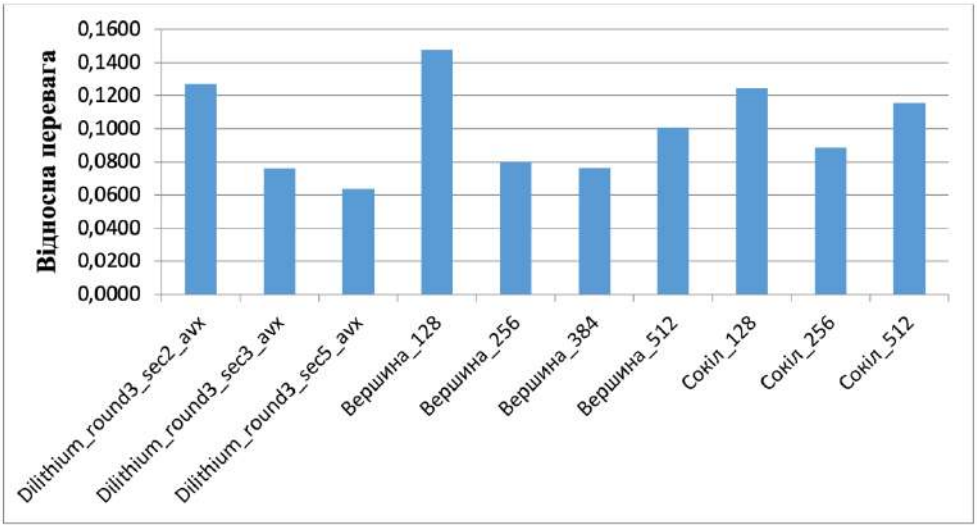


Рис. 2. Переваги алгоритмів ЕП

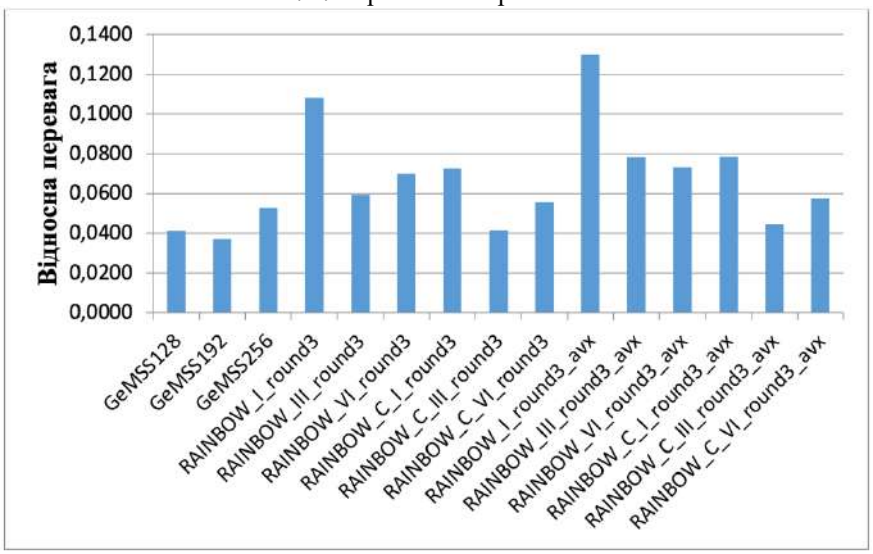


Рис. 3. Переваги алгоритмів ЕП

4.2. Результати кінцевого порівняння перспективних механізмів ЕП

У подальшому порівнювалися алгоритми, що показали кращі результати на попередньому етапі – SPHINCS+_s, «Вершина» та «Сокіл» (через те, що для різних рівнів стійкості перевага у різних алгоритмів), а також Rainbow (оптимізована реалізація зі стандартними параметрами).

У табл. 9 наведено результати досліджень – відносна перевага алгоритмів ЕП, що отримана методом попарних порівнянь за кожною з характеристик.

На рис. 4 відображено гістограму загальної відносної переваги алгоритмів ЕП з урахуванням вагових коефіцієнтів характеристик.

Серед усіх алгоритмів кращий результат у RAINBOW_I_round3_avx (за рахунок малої довжини підпису та великої швидкодії). Але при використані параметрів, що гарантують більшу стійкість, цей алгоритм вже на останньому місці. Якщо ж брати всі можливі параметри алгоритмів, то на першому місці «Вершина» (який в порівнянні з алгоритмами, що засновані на інших математичних апаратах, за сукупністю оцінок обійшов «Сокіл»).

Таблиця 9

Відносна перевага алгоритмів ЕП за кожною з характеристик

Алгоритми	$I_{ст.}$	$l_{в.к.}$	$l_{о.к.}$	$l_{рез.}$	$T_{пр.}$	$T_{зв.}$	$T_{гк.}$
SPHINCS+_128s	0,0155	0,2658	0,2675	0,0189	0,0090	0,0117	0,0164
SPHINCS+_192s	0,0331	0,2289	0,2304	0,0107	0,0090	0,0102	0,0139
SPHINCS+_256s	0,0794	0,1953	0,1984	0,0082	0,0090	0,0079	0,0101
Вершина_128	0,0331	0,0599	0,0664	0,0385	0,2036	0,1525	0,2908
Вершина_256	0,0794	0,0416	0,0433	0,0240	0,1340	0,0713	0,2232
Вершина_512	0,2596	0,0279	0,0274	0,0142	0,0619	0,0305	0,1712
RAINBOW_I_round3_avx	0,0155	0,0117	0,0120	0,2924	0,2864	0,2913	0,0960
RAINBOW_III_round3_avx	0,0331	0,0082	0,0082	0,2043	0,1148	0,1229	0,0500
RAINBOW_VI_round3_avx	0,0794	0,0064	0,0064	0,1746	0,0494	0,0438	0,0263
Сокіл_128	0,0331	0,0770	0,0578	0,1007	0,0619	0,1095	0,0622
Сокіл_256	0,0794	0,0495	0,0337	0,0683	0,0363	0,0898	0,0341
Сокіл_512	0,2596	0,0279	0,0486	0,0451	0,0247	0,0586	0,0057

Тобто, якщо необхідний мінімально задовільний рівень захисту, то кращі результати у Rainbow, а в якості універсального алгоритму краще «Вершина». До того ж, для «Вершини» не були представлені параметри для 1-2 рівнів NIST. Тобто, якщо б були представлені параметри для даних рівнів, то можливо такі параметри обійшли б і Rainbow.

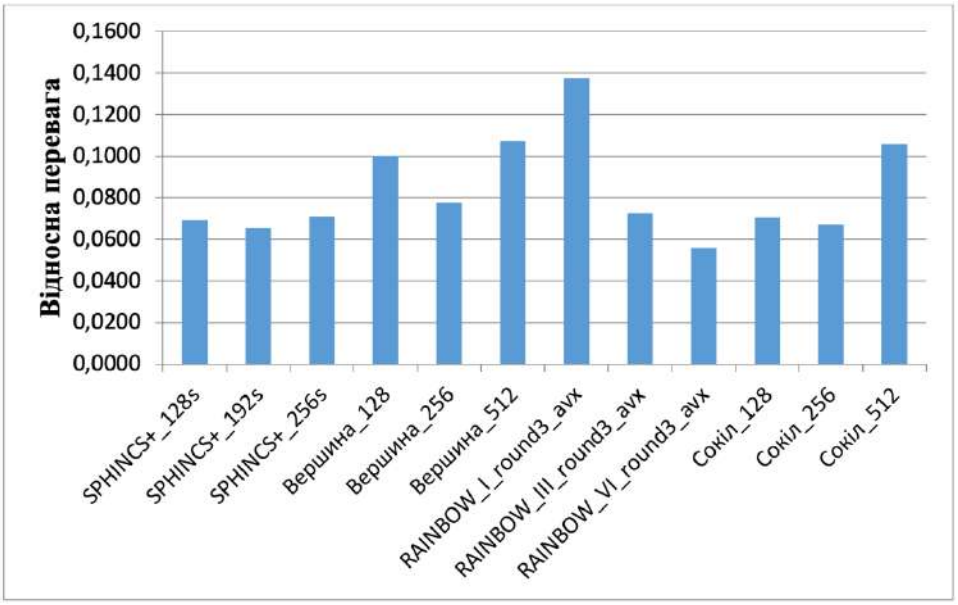


Рис. 4. Переваги алгоритмів ЕП

4.3. Результати повторного порівняння перспективних механізмів ЕП з використанням методу ранжування

Для порівняння методом ранжування варіанти реалізації алгоритмів за їхніми параметрами були розбиті на дві групи рівнів захисту: група параметрів середнього (3-4 рівні) та високого.

На рис. 5 відображено гістограму загальної відносної переваги алгоритмів ЕП (з параметрами високого рівня захисту) з урахуванням вагових коефіцієнтів характеристик з використанням методу ранжування.

При порівнянні методом ранжування з'ясувалося, що кращий результат мають алгоритми, які побудовані на основі перетворень в решеті числового поля.

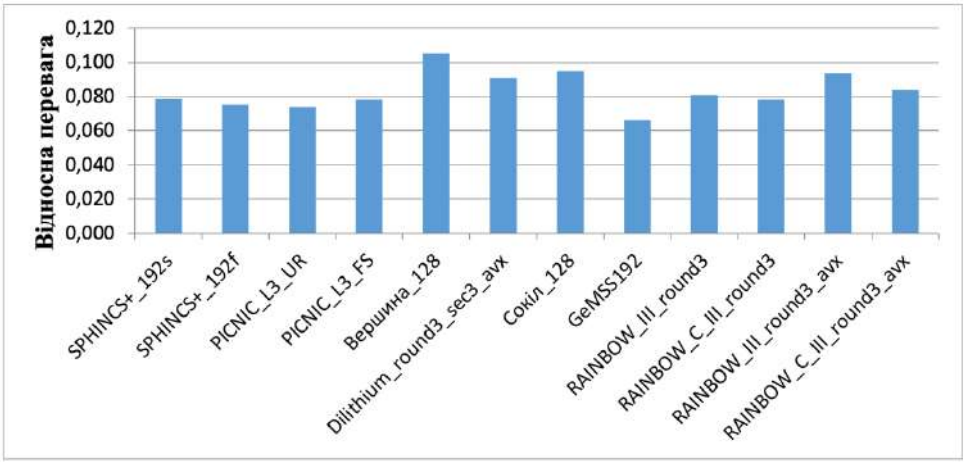


Рис. 5. Переваги алгоритмів ЕП середнього рівня захисту коефіцієнтів характеристик з використанням методу ранжування

4.4. Опис та результати дослідження механізмів ЕП за сукупністю прагматичних

Згідно з комплексною методикою на перших двох етапах застосовуються методики на основі застосування безумовних та умовних критеріїв оцінки ЕП. На третьому етапі залежності від вимог, що висуваються до криптопримітивів, при необхідності потрібно оцінювати та порівнювати альтернативні примітиви за техніко-економічними та техніко-експлуатаційними критеріями (характеристиками) ЕП тощо, які носять локальний характер. В якості основних потрібно використовувати такі критерії (характеристики) як довжини особистих та відкритих ключів ЕП, довжини ЕП, складність (швидкодію) основних – прямих та зворотних криптоперетворень ЕП тощо, складність генерування (обчислення) ключів та параметрів ЕП, а також їх взаємну залежність, у тому числі і у залежності від показників щодо криптостійкості та розмірів параметрів і ключів, а також видами математичних методів ЕП, що використовуються для реалізації криптопримітивів ЕП тощо. Таким чином, важливістю третього етапу комплексної методики оцінки та порівняння ЕП є те, що на ньому здійснюється перевірка щодо відповідності часткових безумовних та умовних критеріїв вимогам, що висунуті щодо них відповідними нормативними документами.

У якості криптопримітивів ЕП виберемо кандидати, що пройшли перший етап відбору. Він є найбільш узагальненим щодо оцінки та порівняння існуючих і перспективних ЕП. На цьому етапі оцінюються можливості виконання прагматичних вимог на основі використання певних математичних методів при криптоперетвореннях ЕП, що пройшли на третій етап міжнародних досліджень.

Послідовність оцінювання та порівняння криптопримітивів ЕП враховують вимоги, а оцінки і порівняння проводяться у наступній послідовності:

1. Аналізу та порівнянню підлягають тільки криптоперетворення ЕП, що успішно пройшли тестування згідно з вимогами другого етапу відповідно до безумовних часткових та безумовного інтегрального критеріїв.

2. Подальший аналіз проводиться з використанням умовних часткових та інтегрального умовного критеріїв щодо усіх криптопримітивів ЕП, що пройшли відбір відповідно до безумовних критеріїв.

3. Визначається перелік необхідних прагматичних критеріїв щодо ЕП. На основі, як правило, експериментальних та в меншій мірі теоретичних оцінок, визначаються основні показники щодо техніко-економічних та техніко-експлуатаційних характеристик, в тому числі з урахуванням NIST.IR 8413 та IT Grundschutz [4, 18]:

$I_{ст.}$ – рівень криптографічної стійкості ЕП;

$l_{в.к.}$ – довжина відкритого ключа ЕП;

$l_{о.к.}$ – довжина особистого ключа ЕП;

$l_{рез.}$ – довжина ЕП;

$T_{пр.}$ – складність (швидкість) обчислення ЕП;

$T_{зв.}$ – складність (швидкість) перевірки ЕП;

$T_{ген.зп.}$ – складність (швидкість) генерування загальних параметрів ЕП;

$T_{ген.кл.}$ – складність (швидкість) генерування ключа (ключової пари) ЕП тощо з урахуванням особливостей.

4. На основі, як правило, експериментальних та в меншій мірі теоретичних оцінок, визначаються залежності необхідних показників між собою щодо їх техніко-економічних та техніко-експлуатаційних характеристик ЕП, але з урахуванням, що вони успішно пройшли перший та другий етапи оцінки та досліджень.

5. На основі аналізу значень показників ЕП, їх залежностей між собою та значень умовних та безумовних критеріїв, що отримані на першому та другому етапах, приймаються рішення про переваги певних кандидатів ЕП.

6. Наприклад, визначаються:

- залежність довжини відкритого ключа ЕП від довжини особистого (закритого) від математичного методу, який застосовується при побудові асиметричної пари ключа для ЕП;
- залежність складності генерування відкритого ключа ЕП від складності генерування особистого ключа у залежності від математичного методу, який застосовується при побудові асиметричної пари ключа для ЕП;
- залежність складності генерування загальних параметрів ЕП від математичного методу, який застосовується;
- залежність збитковості ЕП від математичного методу, який застосовується;
- залежність від виду реалізації (програмна, програмно-апаратна та апаратна тощо).

Очевидно застосування прагматичної методики необхідне у випадках, коли необхідно забезпечити виконання вимог щодо розмірів ключів та параметрів, складності виконання ЕП, у залежності від математичного методу та розмірів загальних параметрів та ключів тощо. Конкретно вони визначаються вимогами до техніко-економічних та техніко-експлуатаційних вимог тощо.

4.5. Дослідження прагматичних оцінок алгоритмів ЕП, що засновані на перетвореннях на алгебраїчних решітках

Для дослідження прагматичних критеріїв було обрано алгоритми, що базуються на криптоперетвореннях на алгебраїчних решітках – Dilithium, «Вершина» та «Сокіл».

На рис. 6 показано залежність довжини відкритого ключа від довжини особистого ключа. Оскільки алгоритм «Вершина» базується на алгоритмі Dilithium, то їх графіки майже співпадають, тільки у «Вершини» він довший за рахунок наявності параметрів, що забезпечує більш високий рівень стійкості. Для алгоритму «Сокіл» графік нерівномірний через особливості кодування особистого ключа для параметрів Сокіл_256 – для цих параметрів особистий ключ довший за інші.

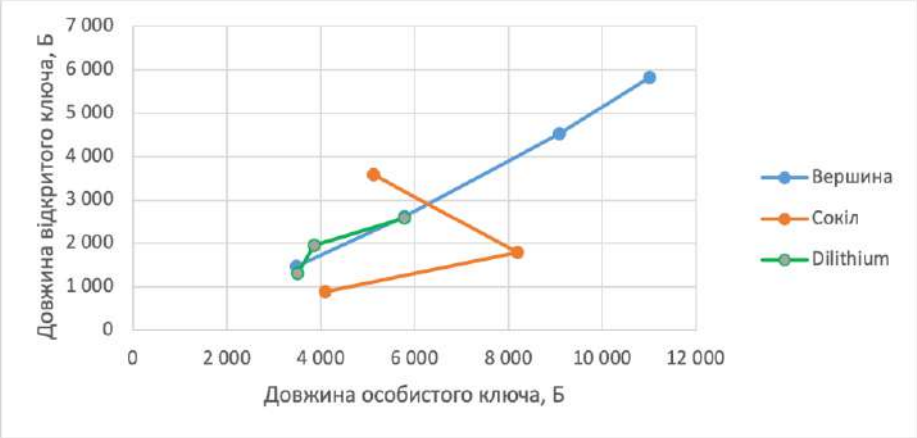


Рис. 6. Залежність довжини відкритого ключа від довжини особистого ключа

На рис. 7 показано залежність довжини підпису від довжини особистого ключа. Залежність та сама, що й для відкритого ключа.

На рис. 8 показано залежність часу підпису від довжини особистого ключа. Графік «Вершини» має як і раніше той же вигляд, що й Dilithium, але за рахунок оптимізацій має більш низькі значення.

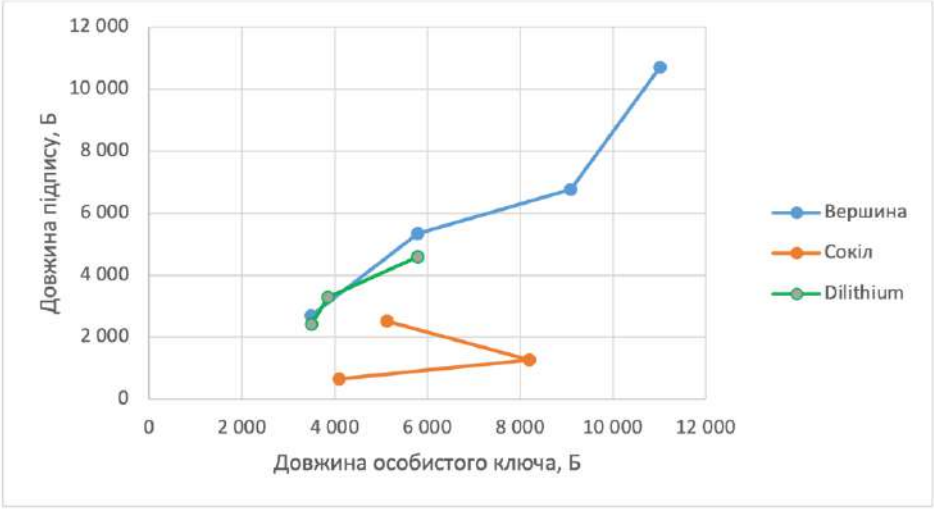


Рис. 7. Залежність довжини повідомлення від довжини особистого ключа

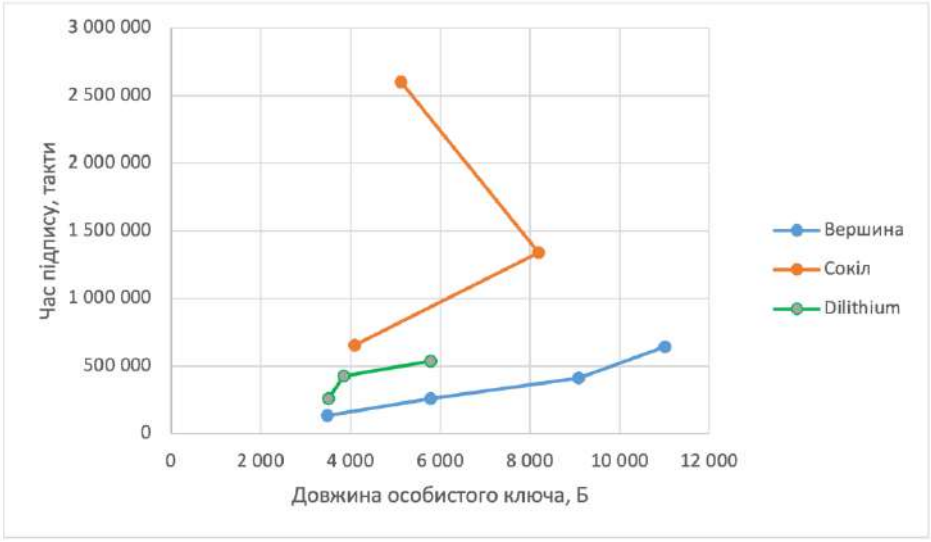


Рис. 8. Залежність часу підпису від довжини особистого ключа

На рис. 9 показано залежність часу перевірки підпису від довжини відкритого ключа.

На рис. 10 показано залежність часу генерації ключів від довжини особистого ключа. Через велику різниці між значеннями було вирішено для наглядності цієї оцінки використовувати в цьому графіку логарифмічну шкалу оцінки (значення вісей мають степені 2, в показниковій прогресії).

За графіками видно, що залежності «Вершини» мають той самий вигляд, що й Dilithium – для розмірів ключів і підпису на рівнях, що відповідають третьому та п’ятому рівням захисту, вони співпадають, а для швидкодії Dilithium має менші значення за рахунок оптимізації. Алгоритм «Сокіл» має кращі показники розмірів ключів, підпису та часу перевірки підпису, але значно програє в швидкодії по самому підпису та генерації ключів. Тому рекомендується використовувати ЕП «Сокіл» в системах, де будуть використовуватися довгострокові ключі, а також кількість операції перевірки підпису буде значно більшою за операції накладання підпису.

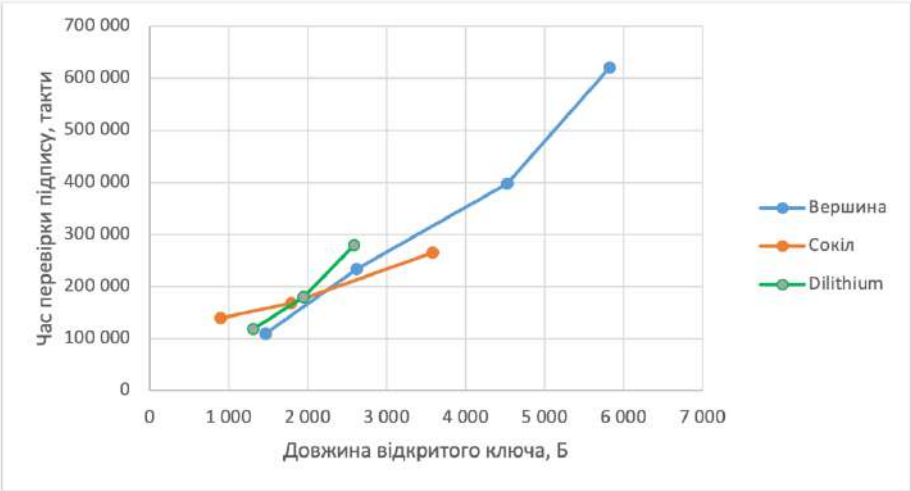


Рис. 9. Залежність часу перевірки підпису від довжини відкритого ключа

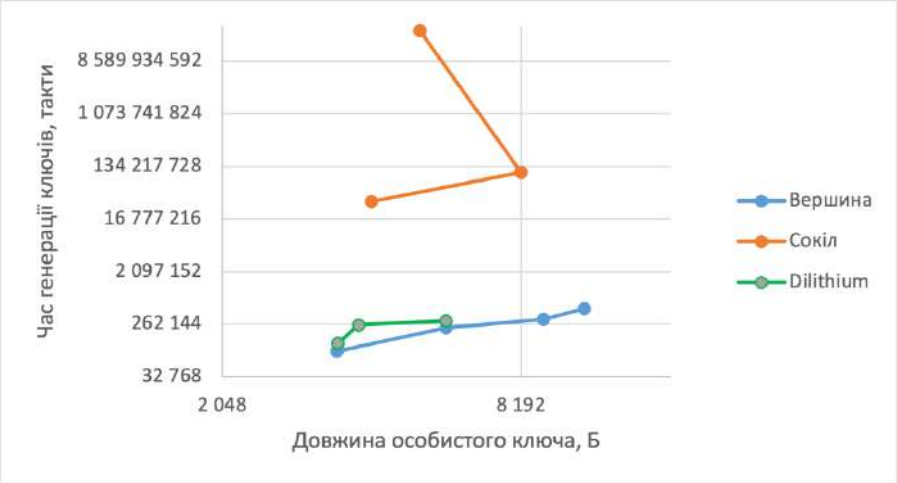


Рис. 10. Залежність часу генерації ключів від довжини особистого ключа

Висновки

1. На міжнародному та національному рівнях суттєво жорсткіші вимоги висунуті стосовно безпеки інформації, техніко-економічних та техніко-експлуатаційних характеристик до інформаційно-комунікаційних систем та інформаційних технологій КЗІ. Це вимоги до надання користувачам в критичних технологіях послуг, цілісності, доступності, конфіденційності, неспростовності відправника та отримувача, криптографічної живучості та санкціонування користувачам дозволу на обмін та обробку інформації взагалі та у критичних технологіях засобом обов'язкового застосування КЗІ.

2. Згідно з суттєво посиленими вимогами системи КЗІ повинні забезпечувати вже захист від класичних, квантових, на основі помилок та атак сторонніми каналами від порушника (криптоаналітика) третього рівня, для якого практично не існує матеріально-технічних та фінансових обмежень, та їх використання на практиці в процесі здійснення кібератак.

3. Під комплексною методикою аналізу криптографічної стійкості існуючих та перспективних (постквантових) АСШ, ПІК та ЕП, у тому числі у перехідний та постквантовий періоди, розуміється фіксована сукупність прийомів практичної діяльності щодо аналізу криптографічної стійкості та властивостей існуючих та перспективних (постквантових) АСШ, ПІК та ЕП, у тому числі у перехідний та постквантовий періоди, шо відповідає вимогам.

4. Методики можуть застосовуватись незалежно одна від одної, але основним є їх застосування у вказаній послідовності – спочатку з використанням на основі безумовних критеріїв, потім – на основі умовних критеріїв та при необхідності – на основі прагматичних критеріїв.

5. Основним завданням комплексної методики є формалізація процесів прийняття рішень відносно виконання висунутих до них вимог, врахування переваг та недоліків криптопримітивів, що є кандидатами на постквантовий стандарт, зменшення впливу суб'єктивних факторів на прийняття рішень, в тому числі несанкціонованого впливу сторонніх організацій тощо. Наприклад, такі методики можуть бути застосованими щодо оцінки та порівняння алгоритмів АСШ, ПІК та ЕП, що є в нашому випадку кандидатами на постквантовий стандарт.

6. До безумовних критеріїв, як мінімум, необхідно віднести:

$I_{ст.}$ – рівень криптостійкості з використанням безумовних критеріїв;

$l_{в.к.}$ – можливі довжини відкритого ключа;

$l_{о.к.}$ – можливі довжини особистого (секретного) ключа;

$l_{рез.}$ – довжину результату криптоперетворення (збитковість);

$T_{пр.}$ – складність (швидкість) прямого криптоперетворення;

$T_{зв.}$ – складність (швидкість) зворотного криптоперетворення;

$T_{ген.зп.}$ – складність (швидкість) генерування загальних параметрів для відповідного режиму роботи криптоперетворення (у залежності від довжин загальних параметрів та ключів);

$T_{ген.кл.}$ – складність (швидкість) генерування ключа (ключової пари) у залежності від режиму роботи тощо.

7. У якості основних рекомендується використовувати такі прагматичні критерії (характеристики) як довжини особистих та відкритих ключів, довжини електронних підписів та довжини блоків, що шифруються, складність (швидкодію) основних прямих та зворотних криптоперетворень АСШ, ПІК та ЕП тощо, складність генерування (обчислення) ключів та параметрів, а також їх взаємну залежність, у тому числі і у залежності від показників щодо криптостійкості та розмірів параметрів і ключів, а також види математичних методів, що використовуються для реалізації криптопримітивів тощо.

8. На основі, як правило, експериментальних та в меншій мірі теоретичних оцінок, визначаються залежності необхідних показників між собою щодо їх техніко-економічних та техніко-експлуатаційних характеристик, але з урахуванням, що вони успішно пройшли перший та другий етапи оцінки та досліджень.

9. Застосування прагматичної методики необхідне у випадках, коли необхідно забезпечити виконання вимог щодо розмірів ключів та параметрів, складності виконання АСШ, ПІК та ЕП у залежності від математичного методу та розмірів загальних параметрів та ключів тощо. Конкретно вони визначаються вимогами до техніко-економічних та техніко-експлуатаційних вимог тощо, наприклад, з урахуванням NIST.IR 8413 та IT Grundschutz.

Список літератури:

1. Закон України «Про захист інформації в інформаційно-телекомунікаційних системах» від 4 червня 2020 року N 681-IX. Режим доступу: https://ips.ligazakon.net/document/z008000?an=4779&ed=2020_06_04.

2. Crystals-Kyber. [Електронний ресурс]. Режим доступу: https://pq-crystals.org/kyber/.

3. Crystals-Dilithium. [Електронний ресурс]. Режим доступу: https://pq-crystals.org/dilithium/index.shtml.

4. NIST IR 8413 Status Report on the Third Round of the NIST Post-Quantum Cryptography Standardization Process / Gorjan Alagic, Daniel Apon, David Cooper, Quynh Dang, Thinh Dang, John Kelsey, Jacob Lichtinger, Carl Miller, Dustin Moody, Rene Peralta, Ray Perlner, Angela Robinson, Daniel Smith-Tone, Yi-Kai Liu. Режим доступу: https://nvlpubs.nist.gov/nistpubs/ir/2022/NIST.IR.8413.pdf.

5. ДСТУ 7624:2014 Інформаційні технології. Криптографічний захист інформації. Алгоритм симетричного блокового перетворення. Режим доступу: https://www.twirpx.com/file/2878521/.

6. ДСТУ 7564:2014 Інформаційні технології. Криптографічний захист інформації. Функція гешування. Режим доступу: http://online.budstandart.com/ua/catalog/doc-page.html?id_doc.66229.

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

8. ДСТУ 8845:2019 Інформаційні технології. Криптографічний захист інформації. Алгоритм симетричного потокового перетворення. Режим доступу: http://online.budstandart.com/ua/catalog/doc-page.html?id_doc=82494.

9. NIST Special Publication 800-208 Recommendation for Stateful Hash-Based Signature Schemes / David A. Cooper, Daniel C. Apon, Quynh H. Dang, Michael S. Davidson, Morris J. Dworkin, Carl A. Miller. Режим доступу: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-208.pdf.

10 Горбенко І. Д., Горбенко Ю. І. Прикладна криптологія : підручник. 2-ге вид. Харків : Форт, 2013. 878 с.

11. Горбенко Ю.І. Методи побудування та аналізу криптографічних систем : монографія. Харків : Форт, 2015. 959 с.

12. ДСТУ ISO/IEC 14888-3:2014 Інформаційні технології – Методи захисту – Цифрові підписи з доповненням. Ч. 3. Механізми, що ґрунтуються на дискретному логарифмі (ISO/IEC 14888-3:2008, IDT). 113 с.

13. ДСТУ 4145-2002 Інформаційні технології. Криптографічний захист інформації. Цифровий підпис, що ґрунтується на еліптичних кривих. Формування та перевірка. К. : Держстандарт України, 2003. 35 с.

14. ДСТУ ISO/IEC 10118-3:2005 Інформаційні технології. Методи захисту. Геш-функції. Ч. 3: Спеціалізовані геш-функції. Режим доступу: https://dnaop.com/html/61829/docA3_ISO_IEC_10118-3_2005.

15. Gorbenko I. D. Algorithms of asymmetric encryption and encapsulation of keys of post-quantum period of 5-7 levels of stability and their application / I. D. Gorbenko, O. G. Kachko, O. M. Aleksiychuk, O. O. Kuznetsov, Yu. I. Gorbenko, V. V. Onoprienko, M. V. Yesina, S. O. Kandy // Радіотехніка. 2019. Вип. 198. С. 5 – 18.

16. 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 – 27.

17. Gorbenko I. D. Generation of general system parameters for Falcon cryptosystem for 256, 384, and 512 security bits / I. D. Gorbenko, S. O. Kandy, M. V. Yesina, Ye. V. Ostryanska // Telecommunications and Radio Engineering, 2022. Vol. 81, Is. 2. P. 49 – 59.

18. IT-Grundschutz-Compendium. Final Draft, 1 February 2022 // Federal Office for Information Security. Germany. Режим доступу: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Grundschutz/International/bsi_it_gs_comp_2022.pdf?__blob=publicationFile&v=2.

19. Горбенко І. Д. Методи, методика та результати порівняльного аналізу електронних підписів згідно з ДСТУ ISO/IEC 14888-3:2014 / І. Д. Горбенко, М. В. Єсіна // Вісник Нац. ун-ту "Львівська політехніка". Сер. "Автоматика, вимірювання та керування". Львів : Нац. ун-т "Львівська політехніка", 2016. № 852. С. 9 – 22.

20. 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.

21. Горбенко І. Д. 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. 2021. Вип. 205. С. 5 – 21.

22. Gorbenko I. D. Calculation of general parameters for NTRU Prime Ukraine of 6-7 levels of stability / I. D. Gorbenko, A. N. Alekseychuk, O. G. Kachko, M. V. Yesina, I. V. Stelnik, S. O. Kandy, V. A. Bobukh, V. A. Ponomar // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 4. P. 327 – 340. DOI: 10.1615/TelecomRadEng.v78.i4.40.

23. Gorbenko I.D. Methods of building general parameters and keys for NTRU Prime Ukraine of 5th–7th levels of stability. Product form / I.D. Gorbenko, O.G. Kachko, Yu.I. Gorbenko, I.V. Stelnik, S.O. Kandyi, M.V. Yesina // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 7. P. 579 – 594. DOI: 10.1615/TelecomRadEng.v78.i7.30.

Надійшла до редколегії 10.02.2023

Відомості про авторів:

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

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

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

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

Каптьол Євген Юрійович – Харківський національний університет імені В. Н. Каразіна, аспірант факультету комп'ютерних наук; співробітник АТ "Інститут Інформаційних Технологій", Україна, e-mail: kaptevg@gmail.com.

УДК 004.056.5 **DOI:10.30837/rt.2023.1.212.06**

С.О. КАНДІЙ

АНАЛІЗ БЕЗПЕКИ ПЕРСПЕКТИВНИХ МЕХАНІЗМІВ ІНКАПСУЛЯЦІЇ КЛЮЧІВ У МОДЕЛІ CORE-SVP

Вступ

Криптографія на решітках є сучасним напрямком досліджень у постквантовій криптографії. Серед фіналістів третього етапу конкурсу NIST PQC [2] більшість перетворень грунтується на структурованих решітках. Зокрема, механізм інкапсуляції ключів CRYSTALS-Kyber [3], який був рекомендований до стандартизації, також грунтується на решітках (проблема Module-LWE [4]). У той же час, в Україні є діючий стандарт ДСТУ 8961:2019 [1], який грунтується на добре відомій проблемі NTRU [5], найкращі методи криптоаналізу якої використовують редукцію решіток.

Оцінка складності редукції решіток для криптографічних схем є давньою проблемою. Асимптотичні оцінки сильно відрізняються від експериментальних значень, тому для вирішення практичних задач був розроблений ряд евристичних методів. По-перше, стандартним засобом при роботі з редукцією решіток є модель core-SVP [6], згідно з якою вартість атаки є вартістю виклику SVP оракула у алгоритмі BKZ [7]. По-друге, робляться евристичні припущення щодо того, як буде змінюватися форма базису протягом процесу редукції. По-третє, робляться еврістичні припущення щодо умови успіху атаки.

За останні роки зроблений значний прогрес у криптографії на решітках. Метою цієї роботи є оцінка безпеки механізмів інкапсуляції ключів CRYSTALS-Kyber та ДСТУ 8961:2019 від прямих атак у моделі coreSVP з врахуванням останніх досягнень в редукції решіток. Для аналізу було обрано дві популярні евристики – GSA [6] та симулятор Чена – Нгуєна [8].

1. Відомості з теорії решіток

Введемо необхідні позначення з теорії решіток [7]. Решітка L з базисом B є множиною цілочисельних комбінацій лінійно незалежних векторів $b_1,...,b_n$: $L(B=b_1,...,b_n)=\left\{\sum_{i=1}^{n}x_ib_i, x_i\in Z\right\}$. Довжиною вектора v є стандартна евклідова норма $\|v\|=\sqrt{v\bullet v}$, де операція $\bullet$ є скалярним добутком та для двох векторів $v=(v_1,...,v_n)$ і $w=(w_1,...,w_n)$ визначена як $v\bullet w=\sum_{i=1}^{n}v_iw_i$. Для чисел i та j з $i<j$ позначення $[i:j]$ означає множину $\{i,i+1,...,j\}$. Для $[1:j]$ використовується позначення $[j]$.

Гамма функція $\Gamma(s)$ визначена для $s>0$ як $\Gamma(s)=\int_0^\infty t^{s-1}e^{-t}dt$. $Ball_n(R)$ означає n-вимірний шар радіуса R та його об'єм $V_n(R)=R^n\pi^{n/2}/\Gamma(n/2+1)$. Для заданного базису $B=(b_1,...,b_n)$ ортогоналізований за Граммом – Шмідтом базис є $B^*=(b_1^*,...,b_n^*)$, де $b_i^*=b_i-\sum_{j=1}^{i-1}\mu_{ij}b_j^*$ для $1\le j<i\le n$, де $\mu_{ij}=(b_i\bullet b_j^*)/\left\|b_j^*\right\|^2$ – коефіцієнти Грамма – Шмідта (ГШ-коефіцієнти), $\left\|b_j^*\right\|$ – довжини векторів Грамма – Шмідта (ГШ-довжини). Детермінант решітки визначений як $\det(L)=\prod_{i=1}^{n}\left\|b_i^*\right\|$ і дорівнює об'єму фундаментального паралелепіпеда $vol(L)$. Ортогональна проекція є відображення $\pi_i: R^n\mapsto span(b_i,...,b_{i-1})^\perp$ для $i\in\{1,...n\}$.

Локальний блок $L_{[i:j]}$ – (проективна) решітка, яка задається наступним чином:

$$L_{[i:j]} = B_i = L(\pi_i(b_i), \pi_i(b+1), ..., \pi_i(b_j)) \quad (1)$$

для $j \in \{i, i+1, ..., n\}$.

У кожній решітці L існує найменший ненульовий вектор. $\lambda_1(L)$ – норма найменшого вектора. Проблема пошуку найменшого вектора (SVP) полягає у пошуку векотора довжини $\lambda_1(L)$. Проблема апроксимації найменшого вектора γ-SVP полягає у пошуку вектора, що має норму, шо менша за $\gamma(n)\lambda_1(L)$, де n - розмірність решітки.

При аналізі часто використовується евристика Гаусса. Якщо задана n-вимірна решітка L та деяка область S, то кількість точок у $S \bigcap L$ приблизно $vol(S)/vol(L)$. Якщо взяти у якості S шар з радіусом R, то число точок можливо апроксимувати як $V_n(R)/vol(L)$, звідки можливо апроксимувати λ_1 як

$$\lambda_1(L) \approx \det(L)^{(1/n)} / V_n(1)^{(1/n)} = \frac{(\Gamma(n/2+1)\det(L))^{1/n}}{\sqrt{\pi}}. \quad (2)$$

Ця оцінка зветься евристикою Гаусса і позначається як $GH(L)$.

1.1. Складні проблеми в теорії решіток

Безпека криптографії на решітках переважно грунтується на проблемах NTRU та LWE (і її різновидах).

Проблема LWE. Нехай $n, q > 0$ – цілі числа, χ – деякий розподіл ймовірностей над множиною цілих чисел Z та s – секретний вектор з рівномірного розподілу над Z_q^n. $L_{s,\chi}$ є розподілом ймовірностей над $Z_q^n \times Z_q$, який отримується наступним чином. Обирається вектор $a \in Z_q^n$ з рівномірного розподілу, значення помилки $e \in Z_q$ з розподілу χ та повертається пара $(a, c) = (a, a \bullet s + e) \in Z_q^n \times Z_q$. Проблема LWE (обчислювальна версія) полягає у тому, щоб для поліноміальної кількості пар (a, c) з розподілу $L_{s,\chi}$ знайти вектор s.

Проблема NTRU. Нехай $n, q > 0$ – цілі числа і задано кільце R_q (на практиці – поле, проте у загальному випадку NTRU визначається для довільних кілець) поліномів ступеня n над кільцем лишків за модулем q. Нехай $f, g \in R_q$ – поліноми з деякого розподілу χ і $h = g/f$. Проблема NTRU (обчислювальна версія) полягає у пошуку поліномів f, g для заданного полінома h.

Проблема LWE має багато узагальнень. Для подальшого аналізу варто згадати про проблеми Ring-LWE [9] та Module-LWE [4]. Загальне визначення цих проблем потребує введення ряду понять з алгебраїчної теорії чисел, тому для простоти викладення наведемо визначення для випадку поля $R_q = Z_q[X]/(X^n+1)$, яке використовується у переважній більшості криптографічних схем на решітках.

Проблема Ring-LWE. Нехай $n, q > 0$ – цілі числа і задано поле $R_q = Z_q[X]/(X^n+1)$. χ – деякий розподіл ймовірностей над R_q та s – секретний поліном з рівномірного розподілу над R_q. $L_{s,\chi}$ є розподілом ймовірностей над $R_q \times R_q$, який отримується наступним чином. Обирається поліном $a \in R_q$ з рівномірного розподілу, поліном помилки $e \in R_q$ з розподілу χ

та повертається пара $(a,c)=(a,a\bullet s+e)\in R_q\times R_q$. Проблема Ring-LWE (обчислювальна версія) полягає у тому, щоб для поліноміальної кількості пар (a,c) з розподілу $L_{s,\chi}$ знайти поліном s.

Проблема Module-LWE визначена аналогічним чином, тільки a,s є векторами поліномів: $a,s\in(R_q)^d$, де $d>0$ деяке ціле число.

1.2. Зауваження щодо Module-LWE

Найкращими відомими методами вирішення проблем Module-LWE та LWE є атаки на основі редукції решіток. Проте, решітки для Module-LWE відрізняються від звичайних LWE решіток додатковою структурованістю, вони складаються з блоків фіксованого розміру. Множина всіх можливих блоків має бієктивне відображення на множину поліномів кільця $R_q=Z_q[X]/\phi(X)$ для деякого незвідного полінома $\phi(X)$. Для довільного кільця $Z_q[X]/\phi(X)$ відображення $rot(f)$ задається як

$$f\in R_q\mapsto\begin{pmatrix} f \bmod \phi(x)\\ f\cdot x \bmod \phi(x)\\ \vdots\\ f\cdot x^{n-1} \bmod \phi(x)\end{pmatrix}\in Z_q^{n\times n}. \tag{3}$$

Відповідно, кожен ідеал $\langle f_1,\ldots,f_{m-1}\rangle\in R_q$ може бути розглянутий як решітка з базисом:

$$\begin{pmatrix} rot(f_1)\\ \vdots\\ rot(f_m)\end{pmatrix}. \tag{4}$$

Решітки такого типу мають назву ідеальних решіток (ideal lattice) і проблема пошуку найменшого вектора (SVP) на них має назву Ideal-SVP.

Модуль M розмірності k над полем R_q задається як множина векторів вигляду $(g_1,\ldots,g_k)$, де кожен g_i належить до певного ідеалу $I_i\in R_q$. Аналогічно до ідеалів кожен модуль може бути розглянутий як решітка з базисом:

$$\begin{pmatrix} rot(f_{11}) & rot(f_{12}) & \cdots & rot(f_{1k})\\ rot(f_{21}) & rot(f_{22}) & \cdots & rot(f_{2k})\\ \vdots & \vdots & \ddots & \vdots\\ rot(f_{m1}) & rot(f_{m2}) & \cdots & rot(f_{mk})\end{pmatrix}. \tag{5}$$

Такі решітки є модульними решітками і проблема пошуку найменшого вектора на них має назву Module-SVP. Для вирішення проблеми Module-LWE зазвичай роблять зведення до проблеми Module-SVP. Будується модульна решітка, достатньо малий вектор якої містить інформацію про вирішення системи рівнянь Module-LWE.

З визначення модульних решіток видно, що вони мають доволі складну алгебраїчну структуру, що задає структурованість базису решітки. Розуміння впливу цієї структурованості на складність пошуку малих векторів є ключовим питанням для оцінки складності криптоаналізу.

Втім, методів використання структурованості модульних решіток існує небагато. Більшість робіт присвячені ідеальним решіткам і не можуть бути узагальнені на випадок модульних решіток.

Для проблеми Ideal-SVP відомі деякі результати [10-12] для квантових комп'ютерів. Зокрема, якщо ідеал є головним (має лише один твірний елемент), то алгоритм [12] дозволяє вирішити задачу Ideal-SVP з фактором апроксимації $\exp(O(\sqrt{n}))$. Ці результати були узагальнені для довільних числових полів у так звані S-Unit атаки [13], які дозволяють вирішувати проблему Ideal-SVP на експоненційний фактор швидше, проте узагальнень S-Unit атак для модульних решіток не відомо.

Взагалом, Module-LWE можливо звести за поліноміальний час до проблеми Ring-LWE за допомогою редукції, що запропонована в [4], і перейти від модульних решіток до ідеальних решіток. Проте, розмірність решітки при цьому значно зростає, що робить такий підхід непрактичним для атаки. До того ж, кількість рівнянь у Module-LWE необхідна більша, ніж є для такої редукції для існуючих криптографічних схем.

Тож, оскільки невідомо методів використання структури модуля для значного зменшення складності пошуку найменшого вектора, то надалі будемо вважати, що використовуються звичайні LWE решітки.

1.3. Алгоритми редукції решіток

Алгоритми редукції решіток на основі заданого базису решітки $(b_1,...,b_n)$ отримують інший базис, у якого ГШ-довжини відносно коротші.

Алгоритм LLL виконує над базисом дві операції: редукція за розміром $b_i \leftarrow b_i - round(\mu_{ij})b_j$ для $j \in [i-1]$ та перестановки b_i і b_{i+1} якщо $\left\|b_{i+1}^*\right\|^2 \le 1/2\left\|b_i^*\right\|^2$ поки відбуваються зміни.

У алгоритмі BKZ (та його варіаціях) фіксується розмір блоку β і відбувається пошук найменшого вектора на решітках $L_{[i,i+\beta'-1]}$ для i від 1 до $n-1$, де $\beta' = \min(\beta, n-i+1)$. Пошук вектора відбувається окремою процедурою (SVP-оракулом).

Для індекса i стандартна реалізація алгоритма BKZ викликає SVP-оракул для решітки $L_{[i,i+\beta-1]}$ і знаходить найкоротший вектор v на цій решітці. Далі BKZ вставляє v у старий базис між b_{i-1} та b_i. Для базису $(b_1,...,b_{i-1},v,b_i,...,b_{min(i+\beta-1,n)})$ застосовується LLL (або BKZ з меншим розміром блока) для отримання нового базису з меншими векторами. Ці процедури складають один раунд алгоритму. У оригінальній версії BKZ алгоритм зупинявся, коли оновлень не відбувалось на протязі $n-1$ раундів.

Наразі існує багато узагальнень BKZ. Особливо варто відмітити слайд редукцію [14], SD-BKZ [15] та G6K [16]. Ці алгоритми мають кращу асимптотичну поведінку. Проте, варто зауважити, що усі узагальнення BKZ роблять поліноміальну кількість викликів до певних алгоритмів пошуку найменшого (або малого) вектора.

Базис $(b_1,...,b_n)$ є HKZ-редукованим, якщо $|\mu_{ij}| \le 1/2$ для всіх i та j і $\pi_i(b_i)$ є найменшим вектором на проєктивній підрешітці $L_{[i:n]}$ для всіх i. ГШ-довжини при цьому можливо оцінити як $\left\|b_i^*\right\| = GH(L_{[i:n]})$.

Мірою якості редукції є кореневий фактор ерміта, який визначений наступним чином:

$$\delta = \left(\|b_0\| / vol(\Lambda)^{1/d}\right)^{1/d}. \quad (6)$$

2. Механізми інкапсуляції ключів

Відповідно до визначення [17] протокол інкапсуляції ключів є трійкою алгоритмів $(Gen, Encaps, Decaps)$, де $Gen: 1^{\lambda} \to (pk, sk)$ – поліноміальний ймовірнісний алгоритм генерації ключової пари. Приймає параметр безпеки 1^{λ} та повертає ключову пару (pk, sk); $Encaps: pk \to (K, C)$ – поліноміальний ймовірнісний алгоритм інкапсуляції ключа. Приймає публічний ключ pk і повертає випадковий ключ K та його інкапсуляцію C; $Decaps: (sk, C) \to \{K, \perp\}$ – детермінований поліноміальний алгоритм декапсуляції ключа. Приймає секретний ключ sk та інкапсуляцію ключа C і повертає ключ K у разі вдалої декапсуляції та символ помилки $\perp$ у разі виникнення помилок.

2.1. Механізм інкапсуляції ключів CRYSTALS-Kyber

Механізм інкапсуляції ключів CRYSTALS-Kyber [2] використовує перетворення у полі $R_q = Z_q[X]/(X^n + 1)$ і грунтується на проблемі Module-LWE. Елементи поля представляються у вигляді поліномів.

У CRYSTALS-Kyber використовуються наступні криптографічні геш функції:

$$PRF: \{0,1\}^{32} \times \{0,1\}^{8} \to \{0,1\}^{*}$$
$$XOF: \{0,1\}^{*} \times \{0,1\}^{8} \times \{0,1\}^{8} \to \{0,1\}^{*}$$
$$H: \{0,1\}^{*} \to \{0,1\}^{32}$$
$$G: \{0,1\}^{*} \to \{0,1\}^{32} \times \{0,1\}^{32}$$
$$KDF: \{0,1\}^{*} \to \{0,1\}^{*}$$

Додатково використовується геш-функція $Parse$, яка перетворює бітову строку на елемент поля з рівномірного розподілу (при умові, якщо вхідні дані з рівномірного розподілу).

Для генерації вектору шуму використовується біноміальний розподіл B_{η} з параметром η. Відповідно, для генерації векторів поліномів з біноміального розподілу використовується функція CBD_{η}.

Схема асиметричного шифрування, що використовується у CRYSTALS-Kyber, зображена на рис. 1.

$KyberKEM.Gen()$:	$KyberKEM.Encaps(pk)$:	$KyberKEM.Decaps(c, sk)$:
1. $z \leftarrow \{0,1\}^{32}$	1. $m \leftarrow \{0,1\}^{32}$	1. $m_1 = KyberPKE.Dec(sk, c)$
2. $(pk, sk_1) = KyberPKE.Gen$	2. $m = H(m)$	2. $(K_1, r_1) = G(m_1 \| h)$
3. $sk = (sk_1, pk, H(pk), z)$	3. $(K, r) = G(m \| H(pk))$	3. $c = KyberPKE.Enc(pk, m_1, r_1)$
4. Повернути (pk, sk)	4. $c = KyberKEM.Enc(pk, m, r)$	4. $K = \begin{cases} KDF(K_1 \| H(c)), c == c_1 \\ KDF(z \| H(c)), c \neq c_1 \end{cases}$
	5. $K = KDF(K \| H(c))$	5. Повернути K
	6. Повернути (c, K)	

Рис. 1. Асиметрична схема KyberPKE

Для отримання протокола інкапсуляції ключів використовується варіант перетворення Фуджісакі – Окамото з неявним відхиленням [18] (рис. 2).

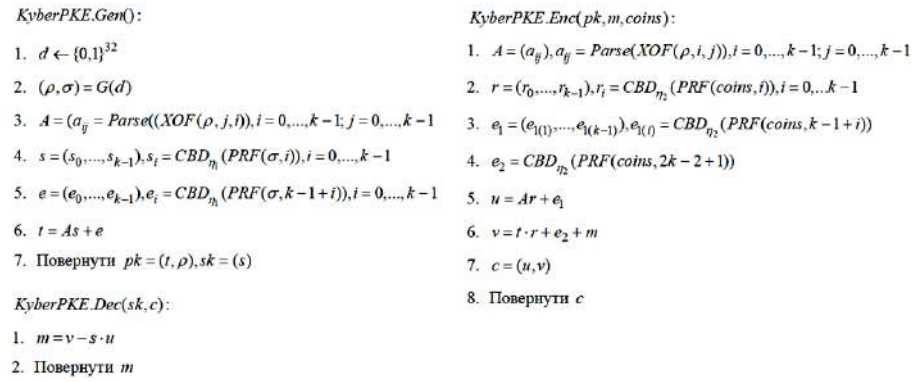


Рис. 2. Механізм інкапсуляції ключів CRYSTALS-Kyber

2.2. Механізм інкапсуляції ключів ДСТУ 8961:2019

ДСТУ 8961:2019 [1] використовує перетворення у полі $R_q = Z_q[X]/(X^n - X - 1)$ і грунтується на проблемі NTRU. Позначимо як R_3 множину усіх поліномів поля R_q, усі коефіцієнти яких належать до множини $\{-1,0,1\}$, як $R_3^{a,b}$ множину усіх поліномів у R_3, що мають кількість ненульових елементів у діапазоні $[a,b]$. Якщо $a = b$, то використовується скорочене позначення $R_3^{t,t} = R_3^t$.

ДСТУ 8961:2019 використовує наступні геш-функції:

$$BPGM : \{0,1\}^L \times R_q \to R_q$$
$$MGF : R_q \to R_3$$
$$H : R_q \to \{0,1\}^\lambda$$
$$KDF : R_q \to \{0,1\}^{K_{len}}.$$

де λ – параметр безпеки, t – загальносистемний параметр, від якого залежить кількість ненульових елементів в поліномах, L – повна довжина повідомлення, K_{len} – довжина ключа інкапсуляції. Додатково використовується бієктивне відображення

$$Pad : \{0,1\}^L \to R_3$$
$$Pad^{-1} : R_3 \to \{0,1\}^L.$$

Механізм інкапсуляції ключів ДСТУ 8961:2019 використовує CPA-to-CCA перетворення власної розробки. На рис. 3 зображено асиметричну схему шифрування, що лежить в основі стандарта і на рис. 4 зображено протокол інкапсуляції ключів ДСТУ 8961:2019.

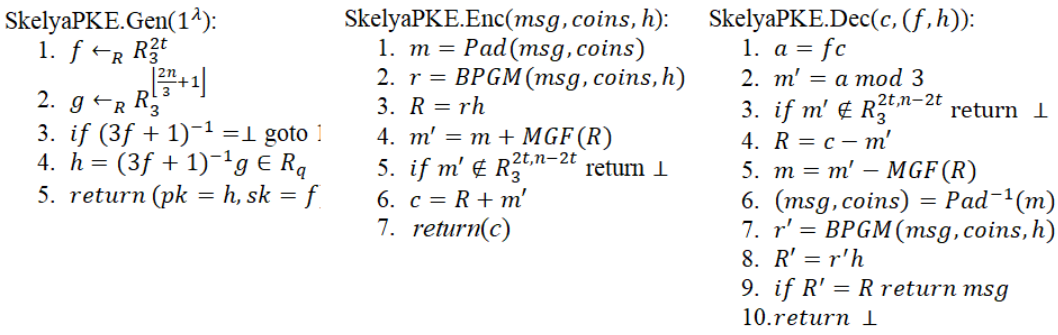


Рис. 3. Асиметрична схема SkelyaPKE

SkelyaKEM.Gen(1^{λ}):
1. $return\ (pk, sk) = SkelyaPKE.Gen(1^{\lambda})$

SkelyaKEM.Encaps($pk = h$):
1. $x \leftarrow_R \{0,1\}^{MsgLen}$
2. $r = BPGM(x, h)$
3. $C_1 = SkelyaPKE.Enc(x, r, pk)$
4. $if\ C_1 = \perp$ goto 1
5. $C_2 = H(r)$
6. $K = KDF(r)$
7. $C = (C_1, C_2)$
8. 8. $return\ (C, K)$

SkelyaKEM.Decaps($C = (C_1, C_2), sk = (f, h)$):
1. $x = SkeylaPKE.Dec(C_1, sk)$
2. $if\ x = \perp$ return $\perp$
3. $r = BPGM(x, h)$
4. $C_2' = H(r)$
5. $C_1' = SkelyaPKE.Enc(x, r, h)$
6. $if\ C_1' = C_1\ \&\&\ C_2' = C_2$
7. $return\ K = KDF(r)$
8. $return\ \perp$

Рис. 4. Механізм інкапсуляції ключів ДСТУ 8961:2019

3. Пряма атака на LWE

Розглянемо більш детально зв'язок проблем NTRU та LWE з теорією решіток.

LWE-рівняння $c - A \cdot s = e \bmod q$ можливо переписати у матричному вигляді наступним чином:

$$B \cdot \begin{pmatrix} * \\ s \end{pmatrix} + \begin{pmatrix} c \\ s \end{pmatrix} = \begin{pmatrix} e \\ s \end{pmatrix}$$
$$B = \begin{pmatrix} qI & -A \\ 0 & I \end{pmatrix} \quad (7)$$

або як

$$B \cdot \begin{pmatrix} * \\ s \\ 1 \end{pmatrix} = \begin{pmatrix} e \\ s \\ t \end{pmatrix}$$
$$B = \begin{pmatrix} qI & -A & c \\ 0 & I & 0 \\ 0 & 0 & t \end{pmatrix}, \quad (8)$$

де t – довільна константа та символ $*$ позначає будь-який вектор. З рівняння видно, що базис B задає решітку, яка містить секретний вектор s з нормою $\sqrt{(n+m)\cdot\sigma^2 + t^2}$.

Нехай $d = n + m + 1$. Якщо $\sqrt{(n+m)\cdot\sigma^2 + t^2} < GH(\Lambda(B)) \approx \sqrt{\frac{d}{2\pi e}} \cdot q^{n/d}$, то B містить вектор, що менший за очікуване значення для випадкових решіток і, отже, має розподілення векторів, що відрізняється від припущень, що використовуються при аналізі BKZ.

Пошук цільового вектора на решітці, що задається базисом B з рівняння (7), зводиться до проблеми BDD, яка формально визначена наступним чином.

Проблема α-BDD. Нехай задано базис B та вектор t і параметр $0 < \alpha < 1/2$, для яких $dist(t, B) < \alpha \cdot \lambda_1(B)$. Необхідно знайти вектор $v \in \Lambda(B)$, який є найближчим до t.

Умова $0 < \alpha < 1/2$ гарантує існування унікального рішення. При $1/2 < \alpha \le 1$ унікальне рішення існує з великою ймовірністю.

LWE з поліноміальною кількістю рівнянь може бути розглянуто як BDD (формула (7)). Асимптотично для будь-якого поліноміально обмеженого $\gamma \ge 1$ існує редукція від $BDD_{1/(\sqrt{2}\gamma)}$ до проблеми $uSVP_{\gamma}$, яка визначена наступним чином:

Проблема $uSVP_{\gamma}$. Нехай задана решітка Λ, для якої виконується $\lambda_2(\Lambda) > \gamma \cdot \lambda_1(\Lambda)$. Необхідно знайти ненульовий вектор $v \in \Lambda$ довжини $\lambda_1(\Lambda)$.

Отже, проблему LWE можливо звести до пошуку малого вектора на решітці з базисом, що задається формулою (8).

4. Пряма атака на NTRU

NTRU решітка має вигляд

$$\Lambda_H^q = \left\{(x, y) \in Z^{2n} \mid Hx - y = 0 \bmod q\right\}, \tag{9}$$

де H є матрицею, у якої i-й стовбець містить коефіцієнти $x^i \cdot h$ для деякого полінома h для $i = 0,...,n-1$. Відповідно базисом NTRU решітки є

$$B = \begin{pmatrix} qI & H \\ 0 & I \end{pmatrix}. \tag{10}$$

Найменшим вектором на решітці є (f, g). Якщо $\|f, g\|$ є меншим за $GH(\Lambda_H^q) \approx \sqrt{n/(\pi e)} \cdot \sqrt{q}$, то решітка містить незвично малий вектор і розподіл малих векторів значно відрізняється від випадкових решіток. Більш того, найменший вектор є не унікальним. Вектори $(x^i \cdot f, x^i \cdot g)$ для $i = 0,...,n-1$ також лежать на решітці.

Якщо не брати до уваги алгебраїчну структуру решітки, то проблему NTRU можливо розглядати як проблему uSVP на решітці з базисом, що задається формулою (9).

Проте, оскільки у NTRU решіток існує багато малих векторів, то це дає додаткову збитковість, яку можна використовувати, на відміну від LWE решіток. У роботах [19, 20] досліджено вплив цієї збитковості. Якщо $q >> n$, то у багатьох випадках можливо пришвидшити редукцію на експоненційний фактор. Параметри криптографічних схем на решітках зазвичай не задовольняють цій умові, проте доволі важко визначити межу, коли прискорення можливо отримати, а коли ні.

Тож, проблеми LWE та NTRU можна звести до проблеми uSVP на решітках з базисами (8) та (10) відповідно.

5. Модель coreSVP

У загальному випадку доволі важко визначити конкретні оцінки часу редукції решітки. Це пов'язано з тим, що асимптотичні оцінки алгоритмів редукції сильно відрізняються від експериментальних даних. Зазвичай алгоритм BKZ та його узагальнення дозволяють отримати менші вектори, ніж зазначено в асимптотичних оцінках.

У роботі [6] запропонована методологія, яка є стандартною методологією для усієї криптографії на решітках. Запропонований підхід дозволяє отримати нижню оцінку часу роботи. Оскільки BKZ та інші алгоритми редукції решіток роблять поліноміальну кількість викликів до SVP-оракула, то автори методології запропонували оцінити складність атаки як час роботи SVP-оракула на решітці розмірності β, де β – найменший розмір блоку для алгоритму BKZ, який дозволяє отримати базис, що містить цільовий вектор.

Визначення мінімальної необхідної розмірності β також є нетривіальною задачею і вирішується за допомогою евристик. У межах цієї роботи використовується еврістичний підхід, що був запропонований в роботі [6]. Сутність евристики полягає у наступному припущенні. Нехай d – розмірність решітки Λ. Якщо під час останнього виклику SVP-оракула для решітки $\Lambda_{[d-\beta:d]}$ буде знайдений вектор $\pi_{d-\beta}(v)$, то з високою ймовірнісю у безисі решітки Λ буде міститися цільовий вектор v після завершення раунду BKZ. Відповідно для мінімальної розмірності β має виконуватися нерівність

$$\left\|\pi_{d-\beta}(v)\right\| < \left\|b^*_{d-\beta}\right\|. \tag{11}$$

У роботі [21] проведено експериментальний аналіз цього підходу для LWE решіток і показано, що така умова вдалого завершення атаки є доволі точною. Тож, проблема зводиться до оцінки значень $\left\|b^*_{d-\beta}\right\|$.

Оскільки теоретичні оцінки доволі сильно відрізняються від експериментальних даних, то використовується ряд евристик. Розглянемо як змінюються логарифми від значень ГШ-норм $\left\|b^*_i\right\|$. На рис. 5 графік значень $\ln\left\|b^*_i\right\|$ для LWE решітки (зліва) та NTRU решітки (праворуч) до та після BKZ редукції з розміром блоку 60.

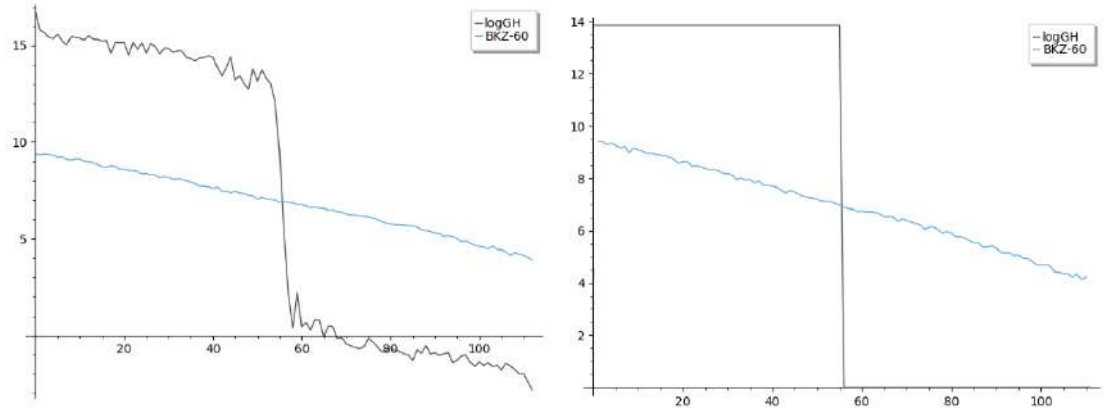


Рис. 5. Графіки ГШ-норм для LWE решітки та NTRU решітки
до та після BKZ редукції з розміром блоку 60

З рис. 5 видно, що після редукції графік ГШ-норм має вигляд прямої. З визначення кореневого фактора ерміта (6) випливає, що $\left\|b^*_0\right\| = \|b_0\| = \delta^d \det(\Lambda)^{1/d}$. Евристика GSA (Geometric Series Assumption) стверджує, що ГШ-норми утворюють геометричну прогресію з параметром δ^2.

Точне значення параметра δ для BKZ не відоме, проте доволі точним наближенням на основі евристики Гауса вважається

$$\delta = \left((\pi\beta)^{1/\beta} \cdot \beta / 2\pi e\right)^{1/2(\beta-1)}. \tag{12}$$

Формула (11) зазвичай використовується при аналізі атак, проте у роботі [7] була запропонована точніша асимптотична оцінка:

$$\delta = \left((\pi\beta)^{1/\beta} \cdot \beta / 2\pi e\right)^{1/2(\beta-1)+\beta/(2n^2)}. \tag{13}$$

Відповідно, перебираючи значення β, можливо знайти найменше, для якого досягається достатньо мале значення δ, щоб (12) або (13) виконувалася.

У табл. 1 – 2 наведено результати експериментальної перевірки оцінок (12) та (13) на решітках малої розмірності для значень $\beta = 20, 50$.

Таблиця 1

Перевірка оцінок (11) та (12) для $\beta = 20$

n	$\log_2 q$	$\log_2 \|b_0\|$	Середньоквадратична помилка для (11)	Середньоквадратична помилка для (12)
50	20	10.50	0.17	0.28
	40	20.34	0.33	0.44
70	20	10.91	0.03	0.11
	40	20.98	0.04	0.04
90	20	11.69	0.48	0.42
	40	21.56	0.35	0.29
110	20	12.10	0.62	0.57
	40	22.21	0.73	0.68
130	20	12.72	0.91	0.92
	40	22.66	0.91	0.86

З табл. 1 – 2 видно, що для розміру блоку 20 оцінка (13) спочатку є більше значення середньоквадратичної помилки, проте з ростом розмірності решітки досягає менших значень помилки, ніж у (12). Проте, для розміру блоку 50 на малих розмірностях не вдалося досягти менших помилок, ніж для (11). З табл. 1 – 2 випливає, що можливо оцінити значення $\|b_0\|$ з точністю до деякого поліноміального фактору.

Таблиця 2

Перевірка оцінок (11) та (12) для $\beta = 50$

n	$\log_2 q$	$\log_2 \|b_0\|$	Середньоквадратична помилка для (11)	Середньоквадратична помилка для (12)
50	20	10.86	0.71	1.55
	40	20.34	0.52	1.36
70	20	10.67	0.53	1.13
	40	20.69	0.51	1.11
90	20	11.12	0.42	0.89
	40	21.07	0.47	0.94
110	20	11.13	0.75	1.14
	40	21.17	0.71	1.10
130	20	11.61	0.62	0.94
	40	22.12	0.63	0.92

На рис. 6 зображено евристику GSA для LWE та NTRU решіток на основі формули (11).

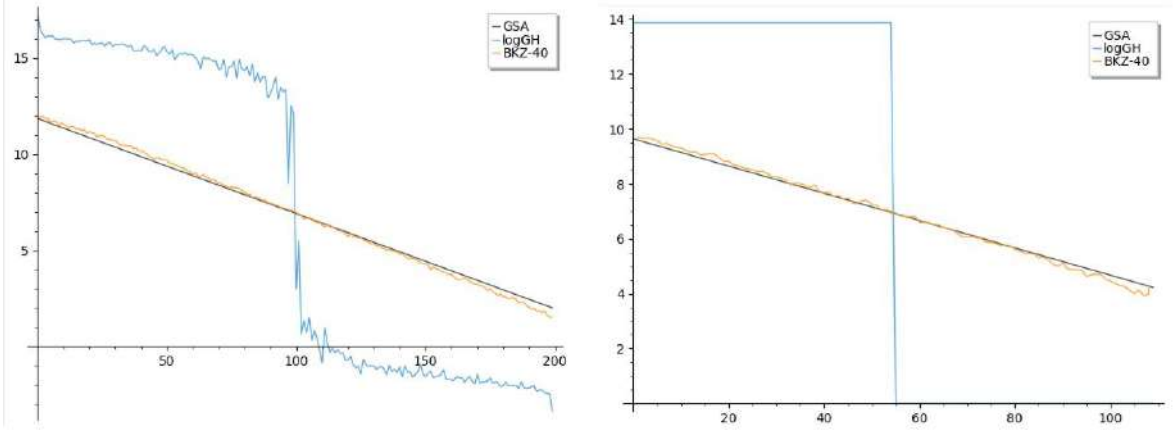


Рис. 6. Евристика GSA

З рис. 6 видно, що евристика GSA добре працює при великих відносно розмірності решітки блоках, проте, якщо розмір блоку є значно меншим за розмірність, то GSA буде виконуватися не для всіх векторів у решітці. Приклад такого явища для LWE та NTRU решіток зображено на рис. 7.

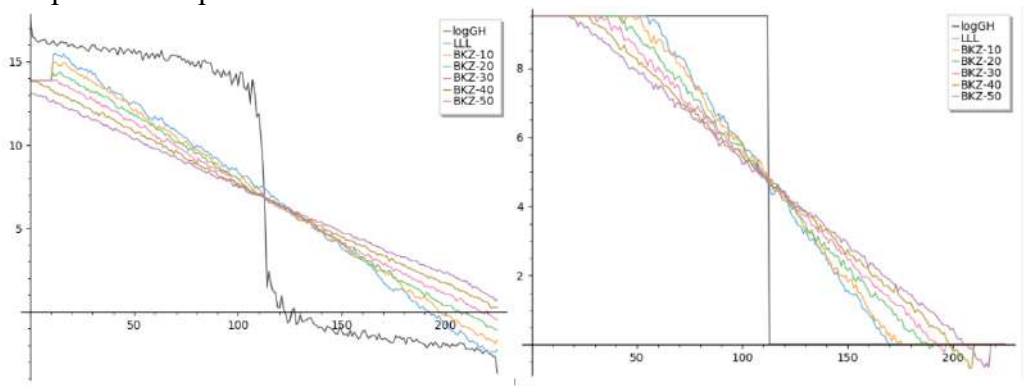


Рис. 7. Приклад ГШ-норм, що частково порушують евристику GSA

Таке явище має місце, оскільки детермінант решітки $\det(\Lambda) = \prod_i \left\| b_i^* \right\|$ є інваріантом. Форми прямої для деяких факторів δ неможливо досягти, оскільки не існує таких векторів, щоб детермінант зберігався. Деяке число перших ГШ-векторів буде мати значення норми q (оскільки LWE решітки та NTRU решітки є q-арними решітками, то вони містять вектори вигляду $(0,...,q,...,0)$ і ці вектори є відносно малими), а останні вектори будуть мати значення норми близьке до 1, оскільки значення детермінанту повинно бути інваріантом. Оцінити кількість векторів з значенням норми q доволі легко. Обчислювати значення вектора згідно з GSA і заміняти його на q, якщо воно є більшим. Такий підхід є евристикою ZGSA. На рис. 8 наведено приклад евристики ZGSA. З рис. 8 видно, що для NTRU решіток ZGSA добре апроксимує значення логарифмів ГШ-норм. Для LWE решіток має місце артефакт. Частина значень, що повинна була мати значення q насправді має значення, відповідають евристиці GSA. Це пов'язано з малою кількістю раундів BKZ (8 раундів). З ростом розмірності і кількості раундів BKZ евристика ZGSA даватиме точніший результат.

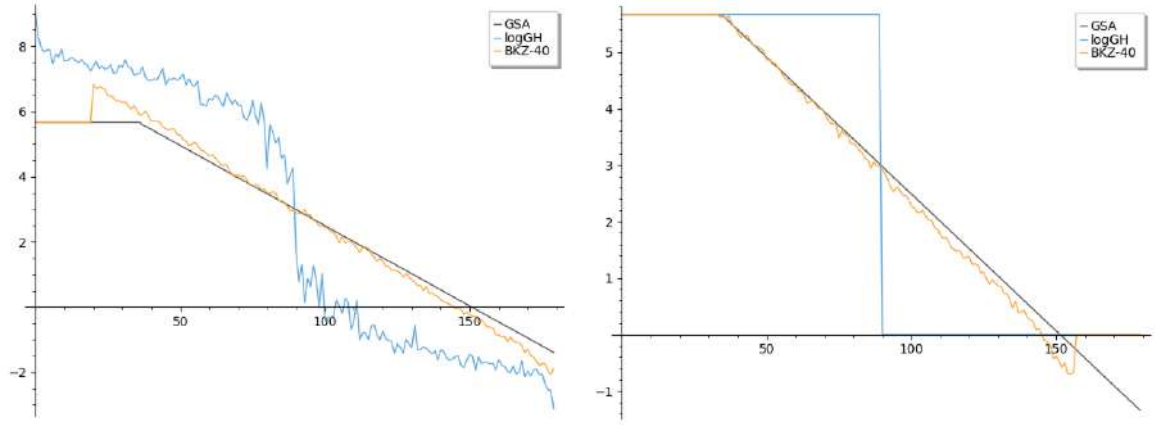


Рис. 8. Евристика ZGSA

Використання ZGSA є стандартним підходом у криптографії на решітках. До недоліків такого підходу можливо віднести те, що важко оцінити наскільки точною буде ця евристика на дійсно великих розмірностях, що відповідають криптографічним значенням параметрів.

Використання ZGSA і евристики (10) дозволяє визначити мінімальне значення β, за якого атака буде успішною. Останнім компонентом для побудови моделі є визначення часу

роботи SVP оракула в залежності від значення β. Двома найпоширенішими підходами до побудови SVP-оракулів є алгоритми просіювання (sieving algorithms) та алгоритми переліку точок (enumeration algorithms) [22].

Алгоритми просіювання грунтуються на суто геометричних ідеях. На початку алгоритму генерується субекспоненційна кількість векторів на решітці. Далі обчислюється множина векторів, що є різницею згенерованих векторів (відбувається "просіювання"). У наступний раунд потрапляють тільки ті вектори, що мають менші довжини, ніж у згенерованих. Процес "просіювання" повторюється ітеративно, поки "просіювання" повертає не пусту множину. Якщо на початку була достатня кількість векторів, то через поліноміальну кількість ітерацій буде знайдено близький до найменшого вектор. Найкращий алгоритм [23] цього класу на сьогоднішній день потребує $\sqrt{3/2}^{\beta/2+o(\beta)}$ операцій. До табл. 3 зведено оцінки роботи найкращих алгоритмів просіювання, згідно з [22].

Зауважимо, що фактор $o(\beta)$ під час аналізу ігнорується, проте він може бути доволі великим. Реальна вартість атаки може коштувати значно більше, тому у табл. 3 зведено як теоретичну оцінку, так і оцінку, що отримана на основі експериментальних даних [22]. Квантові комп'ютери дають відносно невелике прискорення алгоритмів просіювання. Це пов'язано з тим, що операцію «просіювання» важко прискорити за допомогою квантових ефектів.

Таблиця 3

Оцінки часу роботи алгоритмів просіювання

Оцінка	Формула
Алгоритм просіювання на класичному комп'ютері (теоретична оцінка)	$2^{0.29248125036\beta+o(\beta)}$
Алгоритм просіювання на класичному комп'ютері (експерементальна оцінка)	$2^{0.3924062518\beta-5}$
Алгоритм просіювання на квантовому комп'ютері	$2^{0.265\beta+o(\beta)}$

Алгоритми переліку точок є комбінаторним методом вирішення задачі SVP на певній решітці. Для заданного базису $(b_1,...,b_n)$ та ГШ-базису $(b_1^*,...,b_n^*)$ алгоритм переліку полягає у побудові дерева пошуку, у якому вузлами є вектори. Корінь дерева є нульовим вектором. Для кожного вузла $v \in L$ на глибині $k \in 1,...,n$ потомки містять вектори $v+\alpha_{n-k}b_{n-k}(\alpha_{n-k} \in Z)$ з довжиною $\left\|\pi_{n-k}\left(\sum_{i=n-k}^{n} a_i b_i\right)\right\|$ меншою за задану константу $R_{k+1} \in \left(0, \|b_1\|\right]$. Після проходу по всім можливим вузлам алгоритм знаходить вектор на решітці, що менший за радіус пошуку R_n на певній глибині.

Щоб зменшити радіус пошуку, доцільно обирати R_n найближче до норми шуканого вектора. Відповідно, для $\lambda_1(L)$ у якості константи можливо задати $R_n = GH(L)$, тоді кількість векторів, відповідно до евристики Гауса, яку необхідно перебрати, становить

$$FEC(B) = \sum_{k=1}^{n} \frac{V_k(GH(L))}{\prod_{i=n-k+1}^{n} \left\|b_i^*\right\|} \ . \tag{14}$$

Щоб зменшити необхідну кількість векторів, різними авторами були запропоновані евристичні покращення. У роботі [22] запропоновано модель оцінки вартості алгоритмів переліку, у межах якої були запропоновані обмеження на $R_1,...,R_n$. При цьому, на відміну від оригінального алгоритму, така оптимізація робить знаходження необхідного вектора ймовірнісною подією.

Ймовірність p знаходження вектора у шарі з радіусом c, згідно з визначенням, становить

$$p = \Pr_{X \leftarrow S_n}\left[\sum_{i=1}^{l} x_l^2 < R_l^2, l \in 1,...,n\right]. \quad (15)$$

При цьому ціна (кількість векторів) пошуку малого вектора складатиме

$$N = \frac{1}{2}\sum_{k=1}^{n} \frac{vol\{(x_1,...,x_k \in R^k : \sum_{i=1}^{l} x_i^2 < R_l^2, l \in 1,...k\}}{\prod_{i=n-k+1}^{n} \left\| b_i^* \right\|}. \quad (16)$$

З використанням цієї методології у роботі [24] знайдені оптимальні обмеження на $R_1,...,R_n$. У табл. 4 зведено оцінки сучасних алгоритмів переліку з врахуванням описаних вище оптимізацій.

Квантові комп'ютери здатні значно краще прискорити алгоритми переліку, порівняно з алгоритмами просіювання. Прискорення отримується за допомогою застосування варіантів алгоритма Гровера. Алгоритми переліку потребують значно менше пам'яті. Можливість їх реалізації на квантових комп'ютерах ймовірно з'явиться значно раніше, ніж для алгоритмів просіювання, проте, з табл. 4 видно, що асимптотично вони працюють гірше.

Таблиця 4

Оцінки часу роботи алгоритмів переліку

Оцінка	Формула
Класичний комп'ютер, робота [23]	$2^{0.125\beta \cdot \log_2(\beta) - 0.547\beta + 10.4}$
Класичний комп'ютер, робота [34]	$2^{0.1839\beta \cdot \log_2(\beta) - 0.995\beta + 16.25}$
Квантовий комп'ютер	$2^{(0.125\beta \cdot \log_2(\beta) - 0.547\beta + 10.4)/2}$ $2^{(0.1839\beta \cdot \log_2(\beta) - 0.995\beta + 16.25)/2}$

6. Оцінка на основі GSA

Оцінки для CRYSTALS-Kyber у моделі core-SVP для алгоритмів на основі просіювання наведені у табл. 5, з якої видна значна різниця між оцінками безпеки CRYSTALS-Kyber з урахуванням експериментальних результатів та без. Хоча алгоритми просіювання є асимптотично швидшими, проте затрати на пам'ять можуть значно зменшити їх ефективність.

Таблиця 5

Оцінки безпеки CRYSTALS-Kyber
у моделі core-SVP для алгоритмів просіювання

Модель	Класичний комп'ютер (теоретична)	Класичний комп'ютер (експериментальна)	Квантовий комп'ютер
Kyber512	118	155	107
Kyber768	183	243	166
Kyber1024	256	342	232

В табл. 6 наведено оцінки безпеки CRYSTALS-Kyber у моделі core-SVP для алгоритмів переліку. З таблиці видно, що вартість алгоритму переліку зростає асимптотично швидше, проте на квантових комп'ютерах вони мають більше прискорення у порівнянні з алгоритма-

ми просіювання. З урахуванням того, що вони потребують менше пам'яті, можливо для набору параметрів Kyber512 має сенс використовувати саме алгоритми переліку на квантових комп'ютерах.

Таблиця 6

Оцінки безпеки CRYSTALS-Kyber
у моделі core-SVP для алгоритмів переліку

Модель	Класичний комп'ютер	Квантовий комп'ютер
Kyber512	228	114
Kyber768	394	197
Kyber1024	603	301

Оцінки для ДСТУ 8961:2019 у моделі core-SVP для алгоритмів просіювання наведені у табл. 7, якої видно, що для квантових комп'ютерів ДСТУ 8961:2019 забезпечує необхідний рівень безпеки, проте для класичних комп'ютерів оцінки є дещо нижчими за необхідний рівень безпеки.

Таблиця 7

Оцінки безпеки ДСТУ 8961:2019 у моделі core-SVP
для алгоритмів просіювання

Модель	Класичний комп'ютер (теоретична)	Класичний комп'ютер (експериментальна)	Квантовий комп'ютер
Скеля256	169	224	153
Скеля384	239	319	217
Скеля512	296	397	268

Оцінки для ДСТУ 8961:2019 у моделі core-SVP для алгоритмів переліку наведені у табл. 8. Для алгоритмів переліку ДСТУ 8961:2019 забезпечує заявлені рівні безпеки.

Таблиця 8

Оцінки безпеки ДСТУ 8961:2019 у моделі core-SVP
для алгоритмів переліку

Модель	Класичний комп'ютер	Квантовий комп'ютер
Скеля256	358	179
Скеля384	554	277
Скеля512	722	361

7. Оцінка на основі симулятора Чена – Нгуєна

Евристика ZGSA є доволі спрощеною оцінкою форми ГШ-норм. Більш сучасним підходом є оцінки на основі симуляції. Вперше підхід на основі симуляції був запропонований у роботі [8] (симулятор Чена – Нгуєна).

Симулятор працює наступним чином. Нехай $(l_1,...,l_n)$ є симульовані значення ГШ-довжин $\left\|b_i^*\right\|$ для $i=1,...,n$. Тоді симульовані значення для детермінанта і евристики Гауса будуть $\prod_{j=1}^{n} l_i$ та $GH(L_{[i:i+\beta-1]}) = V_{\beta'}(1)^{-1/\beta'} \prod_{j=i}^{i+\beta'-1} l_i$, де $\beta' = \min\{\beta, n-i+1\}$.

Для симуляції раунда BKZ з розміром блоку β передбачається, що кожен виклик алгоритму переліку знаходить вектор з довжиною $GH(L_{[i:i+\beta-1]})$. Далі знайдений вектор використовується для оновлення поточного блоку. Значення (l_i, l_{i+1}) оновлюються до $(l_i^{'}, l_{i+1}^{'})$ для $i = 1,...,n-1$, де $l_i^{'} = GH(L_{[i:i+\beta-1]})$ та $l_{i+1}^{'} = l_{i+1} \cdot (l_i / l_i^{'})$.

Останні 45 ГШ-довжин модифікуються за допомогою ГШ-довжин HKZ-редукованого базису, який обчислений усередненням експериментально отриманих редукованих базисів.

На рис. 9 зображено порівняння евристики GSA та симулятора Чена – Нгуєна. З рисунку видно, що симулятор Чена – Нгуєна може набагато точніше передбачувати форму ГШ-норм.

Проте, недоліком підходу на основі симуляції є те, що він не враховує структурованість решіток і іноді не може точно врахувати ефекти, що виникають при малих розмірах блоку (відносно розмірності решітки).

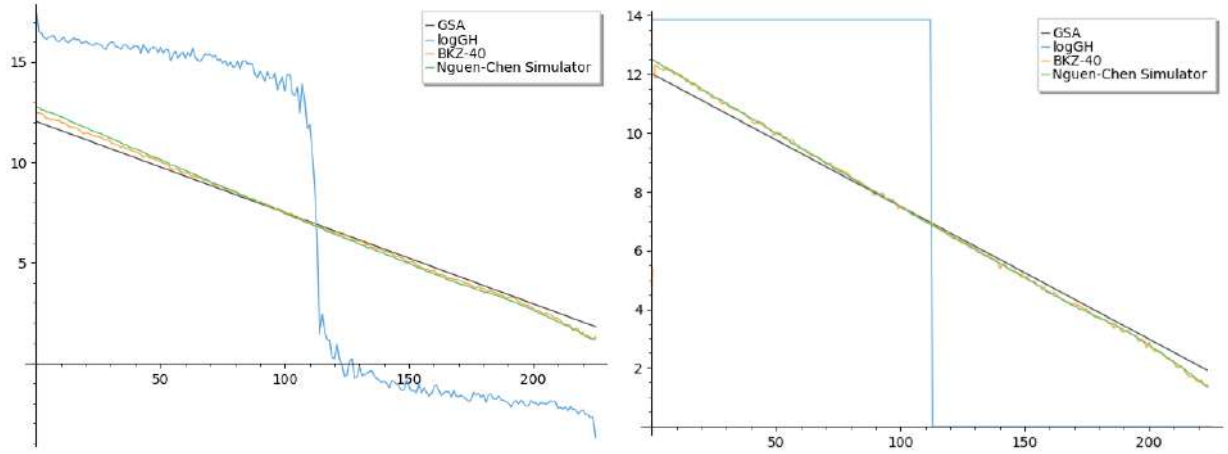


Рис. 9. Порівняння GSA та симулятора Чена – Нгуєна

На рис. 10 наведений приклад такої ситуації для LWE решітки. Симулятор дає на виході GSA-подібний графік, у той час, як перші та останні ГШ-норми не підпорядковуються GSA в результаті редукції.

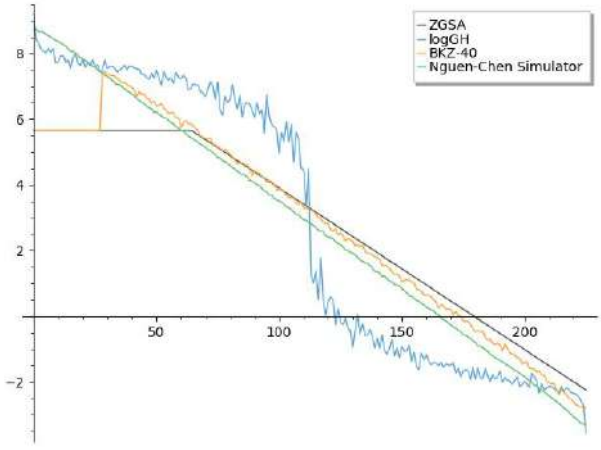


Рис. 10. Приклад неточності симуляції для LWE решітки

Іноді можливі ситуації, коли ані GSA, ані ZGSA, ані симуляція не дають достатньо точних результатів. Приклад такої ситуації наведений на рис. 11.

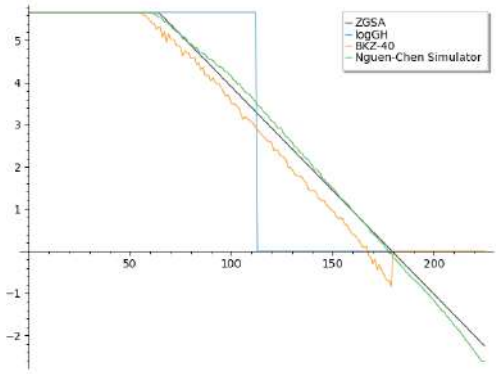


Рис. 11. Неточність евристики ZGSA та симулятора Чена – Нгуєна

Проте, більшість існуючих неточностей стосуються не криптографічних випадків. Для криптографічних випадків $\beta \approx 0.4d$, де d – розмірність решітки. За тих параметрів, що використовуються в криптографії, графік ГШ-норм буде близький до GSA, і у цьому випадку симулятори дуже гарно себе показують.

В табл. 9 наведено оцінки безпеки для CRYSTALS-Kyber на основі симулятора Чена – Нгуєна для алгоритмів просіювання.

Таблиця 9

Оцінки безпеки на основі симулятора Чена – Нгуєна
для алгоритмів просіювання

Модель	Класичний комп'ютер	Класичний комп'ютер (з пам'яттю)	Квантовий комп'ютер
Kyber512	115	151	105
Kyber768	166	220	151
Kyber1024	223	297	202

В табл. 10 наведено оцінки безпеки для CRYSTALS-Kyber на основі симулятора Чена – Нгуєна для алгоритмів переліку.

Таблиця 10

Оцінки безпеки на основі симулятора Чена – Нгуєна
для алгоритмів переліку

Модельг	Класичний комп'ютер	Квантовий комп'ютер
Kyber512	220	110
Kyber768	350	175
Kyber1024	507	253

З таблиць видно, що оцінка безпеки є дещо меншою, що є наслідком врахування симулятором ефектів редукції решіток, які GSA ігнорує.

Оцінки для ДСТУ 8961:2019 у моделі core-SVP для алгоритмів просіювання наведено у табл. 11.

Таблиця 11

Оцінки безпеки ДСТУ 8961:2019 у моделі core-SVP
для алгоритмів просіювання

Модель	Класичний комп'ютер (теоретична)	Класичний комп'ютер (експериментальна)	Квантовий комп'ютер
Скеля256	163	216	148
Скеля384	228	304	206
Скеля512	280	375	254

Аналогічно, оцінки є дещо меншими, ніж при використанні GSA. У табл. 12 зведено оцінки безпеки ДСТУ 8961:2019 для алгоритмів переліку.

Таблиця 12

Оцінки безпеки ДСТУ 8961:2019 у моделі core-SVP
для алгоритмів переліку

Модель	Класичний комп’ютер	Квантовий комп’ютер
Скеля256	343	171
Скеля384	520	260
Скеля512	674	337

На рис. 12, 13 наведено діаграму отриманих оцінок безпеки для алгоритмів просіювання та алгоритмів переліку відповідно.

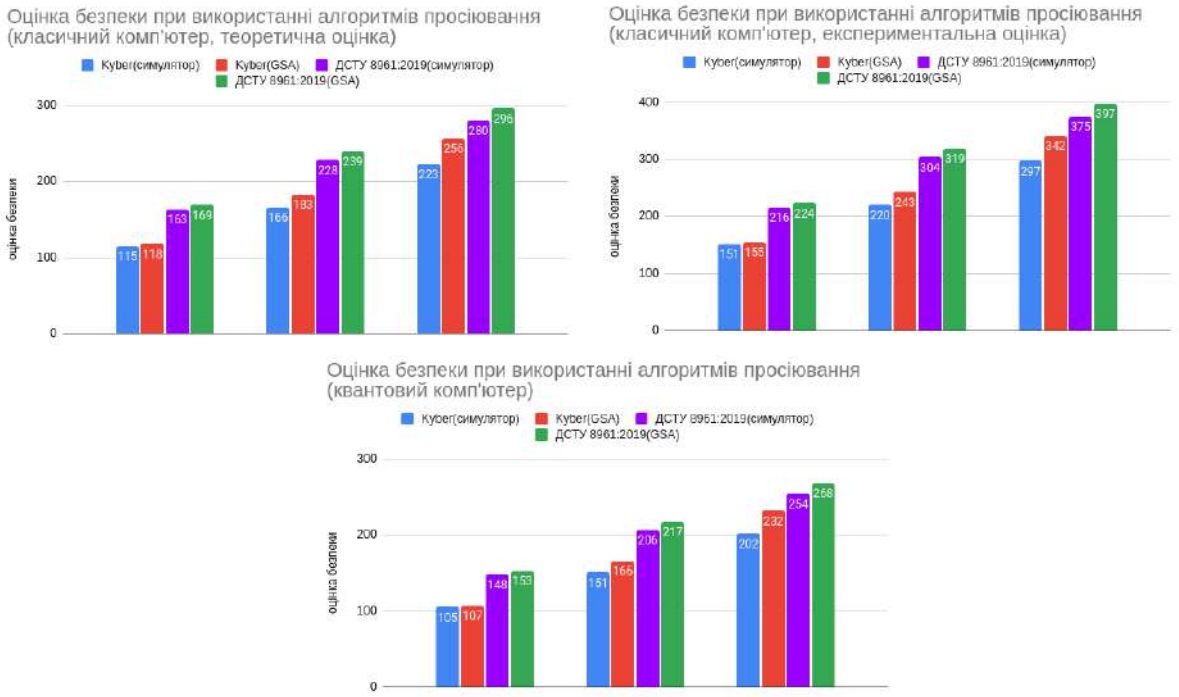


Рис. 12. Безпека CRYSTALS-Kyber та ДСТУ8961:2019 при використанні алгоритмів просіювання

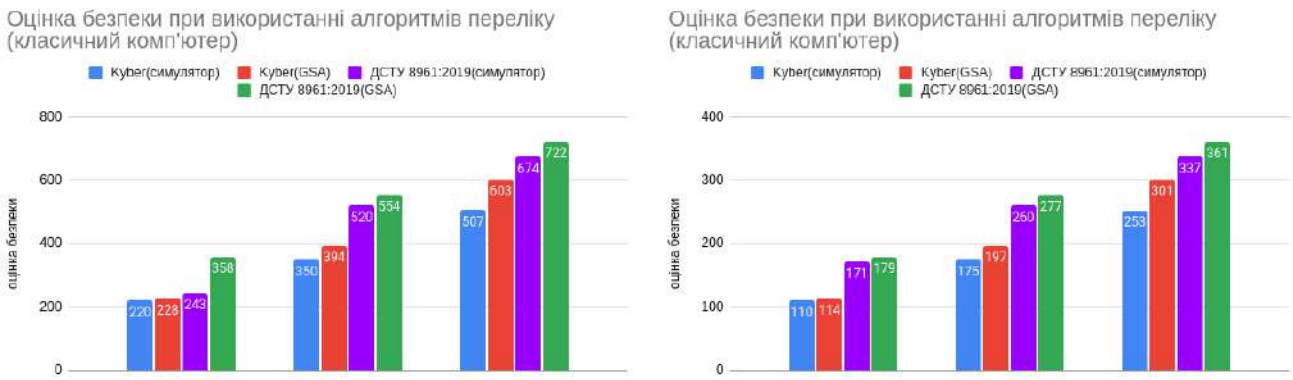


Рис. 13. Безпека CRYSTALS-Kyber та ДСТУ8961:2019 при використанні алгоритмів переліку

Висновки

1. У моделі coreSVP на практиці використовується ряд евристик. Умова вдалого завершіення атаки (нерівність (10)) є евристикою і оцінка ГШ-норм відбувається за допомогою евристики GSA (або ZGSA). Головне питання полягає в тому, наскільки ці евристики точні для решіток високої розмірності. Згідно з результатами роботи [25] умова вдалого завершення атаки дуже гарно апроксимує найменше значення β для LWE решіток. Проте, у тій же роботі автори зазначають, що при певних значеннях параметрів (які не є криптографічними) можливі випадки, коли мінімальне значення β є меншим, ніж передбачає умова вдалого завершення. У той же час, для NTRU решіток (і у деякій мірі для LWE решіток також) до кінця незрозуміло як саме впливає на мінімальне значення β велика кількість малих векторів. У роботі [20] експерментально показано, що вплив у загальному випадку є. Проте, з їх аналізу випливає, що цей ефект є релевантним для більших значень q, ніж використовуються у механізмах інкапсуляції ключів. В цілому, тема встановлення точної умови вдалого завершення атаки є відкритою, але, скоріш за все, потенційні ефекти від уточнення умови вдалого завершення атаки будуть доволі обмеженими.

2. З отриманих оцінок видно, що використання симуляторів дає менші оцінки безпеки, ніж GSA. Перші ГШ-норми та останні ГШ-норми дещо відхиляються від послідовності GSA, і з урахуванням цих ефектів оцінки безпеки є на $\approx 10-20$ біт меншими. Головне питання полягає у тому наскільки точно симулятори дозволяють оцінити норму $b^*_{d-\beta}$. З однієї сторони, на малих розмірностях легко отримати приклади неточних оцінок за допомогою симуляторів. Проте, з іншої – ці неточности пов'язані, як правило, з першими або останніми векторами і для аналізу неважливі. Автори рандомізованого симулятора Чена – Нгуєна наводять приклади ситуацій [26], коли симулятор дає надто оптимістичні оцінки. Тож, існує ймовірність, що менші значення безпеки є результатом саме такого ефекту. Використання симуляторів є перспективним напрямком, проте, необхідні більш точні симулятори, що враховують структурованість LWE та NTRU решіток.

3. Отримані оцінки безпеки показують, що для CRYSTALS-Kyber та ДСТУ8961:2019 загальносистемні параметри для квантових комп'ютерів забезпечують необхідний рівень безпеки, проте для класичних комп'ютерів для деяких наборів параметрів оцінки є нижчими за необхідний рівень.

Список літератури:

1. ДСТУ 8961:2019. Інформаційні технології. Криптографічний захист інформації. Алгоритми асиметричного шифрування та інкапсуляції ключів. Чинний від 21.12.2019. Вид. офіц. Київ : УкрНДНЦ, 2019. 72 с.

2. CRYSTALS – Kyber: a CCA-secure module-lattice-based KEM // Cryptology ePrint Archive, Report 2017/634. [Electronic resource]. Online: https://eprint.iacr.org/2017/634.pdf

3. Lyubachevsky V., Ducas L., Kiltz E. CRYSTALS-Kyber Techn. rep. NIST, 2017. [Electronic resource]. Access mode: https://pq-crystals.org/kyber/ (дата звернення: 21.03.2023)

4. Albrecht M., Deo A. Large Modulus Ring-LWE ≥ Module-LWE // URL: https://eprint.iacr.org/2017/612.pdf (дата звернення: 21.03.2023)

5. Hoffstein J., Pipher J., Silverman H. NTRU: a ring based public key cryptosystem // Algorithmic Nuber Theory. Third International Symposium. 1998. P. 267 – 288.

6. Alkim E., Ducas L., Pöppelmann T., Schwabe P. Post-quantum key exchange – a new hope // URL: https://eprint.iacr.org/2015/1092.pdf (дата звернення: 21.03.2023)

7. Li J., Nguyen P. A Complete Analysis of the BKZ Lattice Reduction Algorithm // URL: https://eprint.iacr.org/2020/1237 (дата звернення: 21.03.2023)

8. Chen Y., Nguyen P. BKZ 2.0: Better Lattice Security Estimates // ASIACRYPT, 2011.

9. Lyubashevsky V., Peikert C., Regev O. On ideal lattices and learning with errors over rings // EUROCRYPT, 2010. P. 1 – 23.

10. Eisenträger K., Hallgren S., Kitaev A., Song F. A quantum algorithm for computing the unit group of an arbitrary degree number field // Proceedings of the forty-sixth annual ACM symposium on Theory of computing. P 293 – 302. ACM, 2014.

11. Campbell P., Groves M., Shepherd D. Soliloquy: A cautionary tale. // ETSI 2nd Quantum-Safe Crypto Workshop. P. 1 – 9.

12. Biasse J., Song F. Efficient quantum algorithms for computing class groups and solving the principal ideal problem in arbitrary degree number fields // ACM-SIAM symposium on Discrete Algorithms, 2017. P. 893 – 902.

13. Bernstein D., Lange T. Non-randomness of S-unit lattices // [Electronic resource]. Online: https://s-unit.attacks.cr.yp.to/spherical.html

14. Gamma N., Nguyen P. Finding short lattice vectors within Mordell's inequality // STOC, 2008. P. 3 – 13.

15. Micciancio D., Walter M. Practical, predictable lattice basis reduction // EUROCRYPT, 2016. P. 56 – 73.

16. Albrecht M., Ducas L., Herold G., Kirshanova E., Postlethwaite E., Stevens M. The General Sieve Kernel and New Records in Lattice Reduction. // URL: https://eprint.iacr.org/2019/089 (дата звернення: 21.03.2023)

17. Dent A. A Designer's Guide to KEMs. Cryptography and Coding // Cryptography and Coding, 2003. Vol 28. P. 29 – 56.

18. Hofheinz D., Hovelmanns K., Kiltz E. A modular analysis of the fujisaki-okamoto transformation // Lecture Notes in Computer Science. 2017. Vol. 10677. P. 341 – 371.

19. Cheon J., Jeong J., Lee C. An algorithm for ntru problems and cryptanalysis of the ggh multilinear map without a low-level encoding of zero // LMS Journal of Computation and Mathematics, 2016. Vol. 19 P. 255 – 266.

20. Kirchner P., Fouque P. Revisiting lattice attacks on overstretched NTRU parameters // STOC, 2017. P. 3 – 26.

21. Albrecht M., Göpfert F., Virdia F., Wunderer T. Revisiting the expected cost of solving uSVP and applications to LWE // ASIACRYPT. 2017. Vol. 10624. P.297 – 322.

22. Bernstein D., Chuengsatiansup C., Lange T., Vredendaal C. NTRU Prime: reducing attack surface at low cost // URL: https://eprint.iacr.org/2016/461 (дата звернення: 21.03.2023)

23. Laarhoven T., Mariano A. Progressive lattice sieving // URL: https://eprint.iacr.org/2018/079.pdf (дата звернення: 21.03.2023)

24. Gama N., Nguyen P., Regev O. Lattice Enumeration Using Extreme Pruning // URL: https://hal.science/hal-01083526/document (дата звернення: 21.03.2023)

25. Albrecht M., Göpfert F., Virdia F., Wunderer T. Revisiting the expected cost of solving uSVP and applications to LWE // ASIACRYPT. 2017. Vol. 10624. P.297 – 322.

26. Bai S., Stehle D., Wen W. Measuring, simulating and exploiting the head concavity phenomenon in BKZ // ASIACRYPT, 2018. P. 389 – 404.

Надійшла до редколегії 15.02.2023

Відомості про авторів:

Кандій Сергій Олегович – Харківський національний університет імені В. Н. Каразіна, аспірант кафедри безпеки інформаційних систем і технологій, факультет комп'ютерних наук; АТ «Інститут Інформаційних Технологій», технік-конструктор, Україна; e-mail: sergeykandy@gmail.com

УДК 621.396:004.056 **DOI:10.30837/rt.2023.1.212.07**

А.М. ОЛЕЙНІКОВ, канд. техн. наук, В.А. ПУЛАВСЬКИЙ, канд. техн. наук,
О.Г. БІЛОЦЕРКІВЕЦЬ

ШЛЯХИ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МЕТОДІВ ТА ЗАСОБІВ ПРОТИДІЇ НЕСАНКЦІОНОВАНОМУ ЗАПИСУ МОВИ ТА ЇХ ПОРІВНЯЛЬНИЙ АНАЛІЗ

Вступ

Небезпечною загрозою в акустичному каналі витоку інформації являються диктофони, або в сучасних реаліях це будь-які звукозаписні пристрої. Зловмиснику достатньо ввімкнути запис на своєму смартфоні для того, щоб несанкціоновано отримати інформацію.

Існує достатньо технічних засобів для того, щоб забезпечити конфіденційність акустичної інформації на об’єкті. Зазвичай технічний засіб являє собою відносно компактний пристрій, який допоможе зберегти приватність вашої розмови. На жаль, деякі виробники таких пристроїв захисту користуються технологіями маркетингу, які недоцільно застосовувати в сфері захисту. Дані дії призводять до того, що покупець, а далі користувач стане жертвою несанкціонованого запису мови. Тому, щоб унеможливити дану ситуацію, користувач повинен розуміти який метод використовується в його засобі захисту.

На сьогодні маємо три основних методи протидії несанкціонованому запису мови: акустичний метод подавлення, електромагнітний метод подавлення та ультразвуковий.

Акустичний метод протидії несанкціонованому запису мови

Найбільш простим і очевидним способом постановки перешкоди запису варто вважати акустичні перешкоди в тій же смузі частот, що й мова, і бажано з близькими кореляційними властивостями. На практиці це означає, що переговори, з погляду безпеки та ефективності захисту від запису, потрібно вести там, де грає голосна музика, транслюється передача чи працює двигун. При цьому через особливості слуху людина здатна селектувати голос співрозмовника, а мікрофон буде насамперед сприймати найбільш голосні звуки, викликаючи спрацьовування системи автоматичного регулювання підсилення (АРП) і зниження коефіцієнта підсилення до значення, при якому шуми і перешкоди задавлять сигнал при наступному відтворенні. Цей спосіб особливо ефективний за умови, що співрозмовник не в змозі вплинути на вибір місця переговорів і підготуватися до них заздалегідь. Для забезпечення високої ефективності акустичного методу протидії несанкціонованому запису мови потрібно провести вибір оптимальних параметрів акустичного завадового сигналу. Для підвищення ефективності акустичного методу потрібно використовувати «мовоподібні» завади. Також слід звернути увагу, що акустичне джерело завади повинно генерувати сигнал в тій самій смузі, що і сам корисний сигнал; це допоможе уникнути ефекту від використання різних додаткових фільтрів. Шляхи підвищення ефективності акустичного методу наведено на рис. 1.

Даний метод вважають малоефективним через те, що співрозмовники під час роботи даної акустичної завади починають говорити гучніше, тим самим збільшуючи амплітуду сигналу, який потребує захисту. Також даний метод негативно впливає на психологічний стан співрозмовників.

Електромагнітний метод протидії несанкціонованому запису мови

Електромагнітний метод оснований на тому, щоб направити на звукозаписний пристрій високочастотний ампплітудо-імпульсний модульований сигнал. Механізм впливу даного методу полягає в тому, що обвідна амплітудно-імпульсного модульованого сигналу протидії знаходиться в смузі частот мовного сигналу, наводять високочастотні струми на елементах плат апарату запису звуку як на «випадкових антенах» і детектуються на будь-якій нелінійності – у підсилювачах, детекторі системи АРП та др.

Рис. 1. Підвищення ефективності акустичного методу

У результаті ці явища призводять до того, що система АРП знижує посилення сигналу мікрофона, для підвищення ефективності методу можна збільшити рівень детектованої перешкоди. Таким чином створюється велика ймовірність, що пристрій звукозапису зовсім припинить запис сигналу з мікрофона. Для підвищення ефективності електромагнітного методу слід застосовувати шляхи, які наведено на рис. 2.

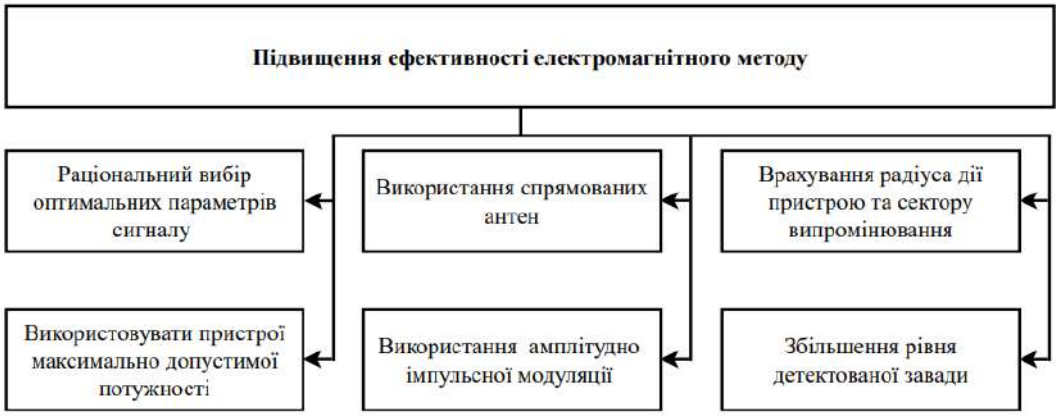


Рис. 2. Підвищення ефективності електромагнітного методу

Електромагнітний метод ефективний майже проти всіх побутових звукозаписних пристроїв, але розвиток схемотехнічної бази елементів та використання екранування сильно знижує ефективність даного методу. На сьогодні більшість сучасних смартфонів мають високий рівень екранування, що суттєво зменшує вплив електромагнітного методу.

Ультразвуковий метод протидії несанкціонованому запису мови

Ультразвуковий метод передбачає опромінення звукозаписного пристрою потужними ультразвуковими коливаннями. Досить велика кількість звукозаписних пристроїв використовує електретні мікрофони, їхня верхня межа смуги пропускання становить 25 – 27кГц [1]. Через це смуга пропускання мікрофону потрапляє в ультразвуковий діапазон і пристрій запису вразливий до потужних ультразвукових коливань, які унеможливлять запис корисного мовного сигналу. Застосовують одночастотний та двочастотний методи захисту на придушення сучасних диктофонів або радіоакустичних закладних пристроїв (РАЗП) з електретними мікрофонами зі смугою пропускання, що дозволяє сприймати ультразвукові сигнали.

Система одночастотного ультразвукового придушення випромінює ультразвукові коливання (УЗК) великої інтенсивності, що впливають безпосередньо на мікрофони диктофону або РАЗП (зазвичай частота випромінювання – трохи більше 20 кГц). Існує два механізми впливу на об'єкт придушення:

- перевантаження підсилювального тракту диктофона або РАЗП;
- спровокована реакція системи автоматичного регулювання рівня запису.

Підвищити ефективність методу можливо за рахунок ультразвукового впливу великої інтенсивності, що призводить до перевантаження мікрофонного підсилювача, який стоїть відразу після акустичного приймача. Відбувається зміщення робочої точки нелінійного елемента підсилювача, що призводить до виникнення значних спотворень мовних сигналів, що записуються, часто до такого рівня, що не піддається розбірливому сприйняттю. Наявність автоматичного регулювання рівня запису (АРРЗ) в диктофоні або РАЗП підвищує ефективність протидії несанкціонованому запису мови внаслідок реакції АРРЗ на УЗ сигнал великої інтенсивності, що призводить до значного зменшення коефіцієнта посилення мікрофонного підсилювача до значення, недостатнього для розбірливого сприйняття мовного сигналу.

Найбільш ефективним є двочастотний метод, який використовує властивість мікрофонного підсилювача як нелінійного елемента. Система двочастотного ультразвукового подавлення випромінює два потужні ультразвукові коливання з частотами, що відрізняються один від одного на 0,3 – 4 кГц. Ці два сигнали збиваються на нелінійному елементі мікрофонного підсилювача, у результаті виходить сигнал з комбінаційною частотою. Цей сигнал різницевої частоти і виступає як перешкода. Для покращення результативності ультразвукового методу слід дотримуватися рекомендацій, вказаних на рис. 3.

Рис. 3. Підвищення ефективності ультразвукового методу

Ультразвуковий метод подавлення є зовсім неефективним, якщо пристрій запису працює лише в мовному діапазоні, має фільтр, що обмежує смугу вхідного сигналу, або пристрій захищено спеціальним матеріалом, який не пропускає ультразвукові коливання.

Аналіз шляхів підвищення ефективності методів протидії несанкціонованому запису мови

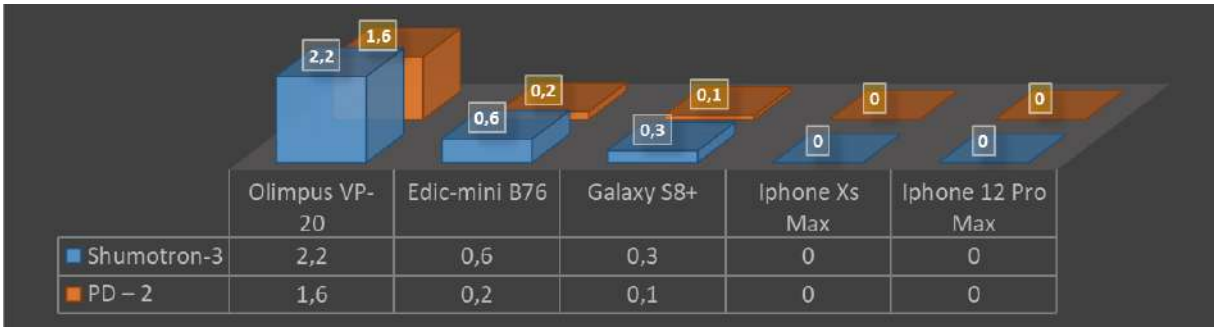
Аналізуючи наведені методи подавлення, отримуємо висновок, що, не знаючи типу звукозаписного пристрою, неможливо забезпечити конфіденційність мовного сигналу. Для підвищення захисту від несанкціонованого запису мовного сигналу пропонується використовувати адаптований акустичний метод [2]. Метод відрізняється від звичайного акустичного тим, що перешкода створюється на основі мови самого співрозмовника, і таку заваду складно відфільтрувати, оскільки вона займає ту саму смугу частот, що і мовний сигнал. Адаптація акустичного методу протидії несанкціонованому запису мови полягає в наступному:

- відстань між джерелом акустичних перешкод і місцем можливого розташування звукозаписного пристрою має бути мінімізовано в кілька разів менше відстані між джерелом мови та звукозаписним пристроєм;
- формувати акустичну перешкоду на основі мовлення співрозмовників. Дані перешкоди корелюють із сигналом, що забезпечує ефективне придушення навіть при малому відношенні сигнал/шум і не піддаються фільтруванню, оскільки займають ту ж смугу частот, що і мовний сигнал. Отримати копію перешкод для очищення стає важче. Шум присутній лише тоді, коли існує мовний сигнал і він відсутній під час пауз в мовленні. Перевагою такого акустичного придушення є те, що в разі наближення джерела перешкоди до звукозаписного пристрою значно підвищується відношення рівня перешкоди до рівня мовного сигналу. Крім того, завдяки утворенню перешкод лише під час існування мовного сигналу, можна позбутися негативного психологічного впливу акустичних перешкод на співрозмовників під час мовчання в паузах між мовленням. Під час ведення розмови таке втручання не привертає уваги співрозмовників і не перешкоджає спілкуванню. Використання особливостей такого акустичного подавлення дозволяє значно підвищити ефективність протидії від несанкціонованого запису на звукозаписувальні пристрої;
- потрібно покращити технічні параметри акустичної системи для випромінювання мовної перешкоди, застосувавши електростатичну акустичну систему випромінювання перешкоди, слід відмовитися від використання традиційних електродинамічних випромінювачів. Дані дії призведуть до підвищення лінійності частотної характеристики акустичної системи, зменшаться її нелінійні спотворення та звузиться діаграма спрямованості акустичної системи. Ці зміни технічних параметрів акустичної системи дозволять максимально наблизити спектральні характеристики перешкод до голосів співрозмовників.

Експериментальне дослідження засобів протидії несанкціонованому запису мови та їх порівняльний аналіз

Результативність адаптованого акустичного методу можемо перевірити в експериментальному дослідженні. Для оцінки ефективності адаптованого акустичного методу подавлення несанкціонованого запису мови з використанням перешкоди, сформованої електростатичним випромінювачем, було проведено експеримент – порівняння технічних параметрів засобів захисту від несанкціонованого запису мови, побудованих з використанням адаптованого акустичного (EST-ST, EST-P), електромагнітного (Шумотрон –3, PD – 2) та ультразвукового (USPD-C, UltraSonic-50) методів подавлення для п'яти сучасних типів звукозаписних пристроїв – цифрових диктофонів та смартфонів (Olimpus VP-20, Edic-mini B76, Galaxy S8+, Iphone Xs Max, Iphone 12 Pro Max.).

На рис. 4 наведено дальність в метрах повного подавлення диктофонів при використанні електромагнітних подавлювачів «ШУМОТРОН-3» та «PD-2».



	Olimpus VP-20	Edic-mini B76	Galaxy S8+	Iphone Xs Max	Iphone 12 Pro Max
Shumotron-3	2,2	0,6	0,3	0	0
PD – 2	1,6	0,2	0,1	0	0

Рис. 4. Електромагнітне подавлення

На рис. 5 наведено дальність в метрах повного подавлення диктофонів при використанні ультразвукових подавлювачів.

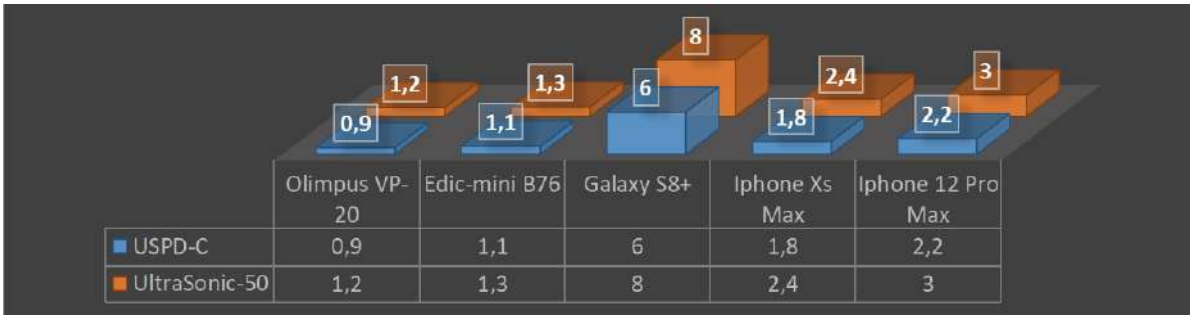


Рис. 5. Ультразвукове подавлення

На рис. 6 наведено дальность в метрах повного подавлення диктофонів при використанні подавлювачів "EST-ST" та "EST-P".

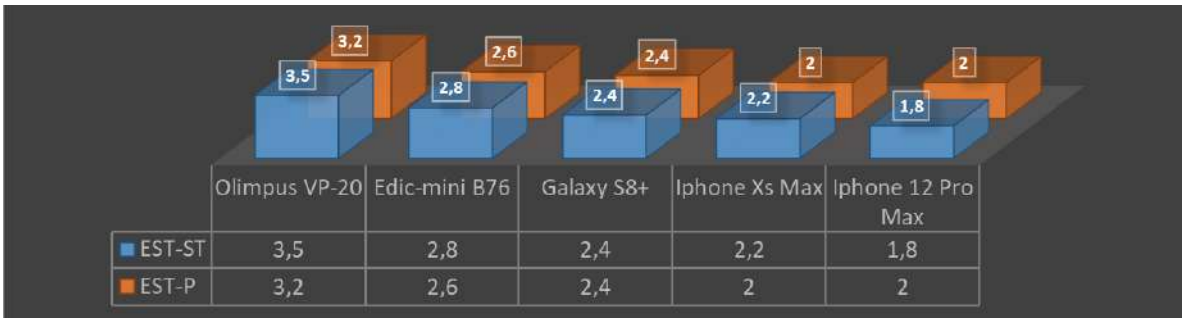


Рис. 6. Подавлення адаптивним методом

Висновок

В ході аналізу виявлено, що жоден метод та засоби на їх основі без апріорного знання типу звукозаписного пристрою не забезпечує гарантованого подавлення несанкціонованого запису мови. На основі цього оптимальним рішенням є адаптація акустичного методу. Аналізуючи результати експерименту (рис. 4 – 6), отримуємо висновок, що адаптований акустичний метод є найбільш ефективним. Ефективність методу забезпечується дальністю подавлення, вона склала від 1,8 до 3,5 м в залежності від пристрою. Для порівняння в ультразвуковому експерименті один із показників склав 0,9 м, а в електромагнітному навіть 0 м. Формування перешкоди по акустичному каналу описаним способом забезпечує універсальність запропонованого методу до будь-якого типу пристрою придушення несанкціонованого запису мови незалежно від методу, що лежить в його основі – електромагнітного, ультразвукового чи акустичного.

Список літератури:

1. Олейников А.Н., Пулавский В.А., Кривенко М.А. Ультразвуковые методы защиты речевой информации // Радиотехника. 2012. Вып. 169. С. 176 – 181.

2. Олейников А.Н., Пулавский В.А., Цыбулевский П.В. Оценка эффективности акустического противодействия несанкционированной записи на диктофон // Современная защита информации. 2010. №1. С. 8 – 16.

Надійшла до редколегії 08.02.2023

Відомості про авторів:

Анатолій Миколайович Олейніков – канд. техн. наук, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації;Україна; email: anatoly.oleynikov@nure.ua; ORCID: https://orcid.org/0000-0002-4458-8833

Володимир Антонович Пулавський – канд. техн. наук, директор фірми «Пулавський»; Україна; email: pulavskiy.v@gmail.com; ORCID: https://orcid.org/0000-0002-9976-9439

Олексій Геннадійович Білоцерківець – Харківський національний університет радіоелектроніки, асистент кафедри мікропроцесорних технологій і систем; Україна; email: oleksii.bilotserkivets@nure.ua; ORCID: https://orcid.org/0000-0002-8514-9650

УДК 621.391: 537.86:519 **DOI:10.30837/rt.2023.1.212.08**

В.А. ТИХОНОВ, д-р фіз.-мат. наук, В.М. БЕЗРУК, д-р техн. наук

ЗАСТОСУВАННЯ ФАКТОРИЗАЦІЇ ДЛЯ ПІДВИЩЕННЯ РОЗДІЛЬНОЇ ЗДАТНОСТІ ПАРАМЕТРИЧНОЇ ОЦІНКИ СПЕКТРАЛЬНОЇ ЩІЛЬНОСТІ ПОТУЖНОСТІ

Вступ

При вирішенні багатьох прикладних задач обробки в галузях радіолокації, радіозв'язку, технічної і медичної діагностики виникає необхідність виконувати спектральний аналіз спостережуваних процесів, які носять випадковий характер. Спектральний аналіз випадкових процесів умовно ділять на два види: непараметричний і параметричний. Непараметричні методи оцінювання спектральної щільності потужності (СЩП) грунтуються на перетворенні Фур'є. Параметричні методи оцінювання СЩП вимагають попереднього оцінювання параметрів стаціонарних моделей лінійного прогнозування, зокрема моделей авторегресії (АР), ковзного середнього (КС), авторегресії – ковзного середнього (АРКС). Оцінки СЩП, що отримані через параметри цих моделей, мають ряд переваг у порівнянні з оцінками СЩП, знайденими непараметричними методами. Зокрема, з'являється можливість отримати більш високу роздільну здатність при більш коротких реалізаціях процесів у порівнянні з непараметричними методами. У цьому випадку підвищення роздільної здатності СЩП пов'язане з неявним застосуванням прогнозування кореляційної функції випадкових процесів.

Природним розвитком у дослідженнях моделей лінійного передбачення є заміна стаціонарності моделей на нестаціонарність і лінійності моделі на нелінійність. Нестаціонарність моделей досліджується порівняно рідко тому, що її важко враховувати. Звичайний метод врахування нестаціонарності моделі полягає в диференціації початкового часового ряду для отримання стаціонарного процесу. Це використання моделі авторегресії проінтегрованого ковзаючого середнього (АРПКС) [1]. Використання дискретної похідної від нестаціонарного процесу, що приводить для деяких часових рядів до стаціонарності моделі, дозволяє в дослідженнях нестаціонарні властивості початкового процесу замінити на стаціонарні характеристики його похідної. Це істотно спрощує вирішення ряду практичних задач, зокрема прогнозування часових рядів. Декомпозиція нестаціонарного процесу з трендом і сезонною складовою [2] дає можливість частково виправити положення з аналізом нестаціонарного випадкового процесу. Методи простору станів широко використовуються для моделювання нестаціонарних часових рядів [3 – 5]. Янг у роботах [6, 7] запропонував використовувати простори станів до нелінійних і нестаціонарних процесів для їх обробки на основі виявлення і оцінювання стохастичних моделей зі змінними в часі параметрами. Якщо лінійність моделі АР єдина, то нелінійність допускає множину варіацій і пов'язана з додатковими складнощами в оцінюванні параметрів моделі. Для розширення можливості застосування лінійної моделі АР була використана множина нелінійних моделей. Запропоновано такі моделі як порогова авторегресійна модель Тонгу [8], експоненціальна авторегресійна модель Одзаки і Хаггана [9, 10], модель Прістлі [11], функціонально-коефіцієнтна авторегресійна модель Чена і Цая [12] та інші. Огляд літератури показує на неослабний інтерес до моделі АР, викликаний, з одного боку, широкими можливостями моделі, а з іншого боку – обмеженнями, що накладаються на цю модель.

Тому актуальним є продовження досліджень АР моделі з метою подальшого підвищення роздільної здатності оцінки СЩП з використанням факторизації СЩП випадкових процесів. Під факторизацією розуміється розкладання багатомодової СЩП на простіші одномодові складові. Це дозволяє більш точно проводити аналіз низькочастотних, середньочастотних і

високочастотних складових СЩП. Основна увага у даній роботі приділена дослідженню завдання підвищення роздільної здатності оцінки СЩП випадкових процесів на основі моделі авторегресії лінійного передбачення шляхом факторизації методом Юла – Уокера і Берга.

Особливості моделі авторегресії

Модель АР раціонально використовувати для описування вузькосмугових випадкових процесів. В основу моделі АР покладено кореляцію відліків випадкового процесу у нинішній момент часу з деяким скінченним або нескінченним числом відліків у попередні моменти часу. Кореляційні зв'язки дозволяють здійснити регресію поточного відліку процесу на попередні відліки. Такий вид регресії називається авторегресією. Вважатимемо у подальшому, що корельований стаціонарний випадковий процес має нульове математичне сподівання. У рівнянні АР поточний відлік представляється зваженою сумою попередніх відліків процесу з деякими коефіцієнтами ваги:

$$x[t] = \sum_{j=1}^{p} \Phi[p, j] x[t - j] + a[t], \tag{1}$$

де $\Phi[p, j]$ – коефіцієнти АР, $a[t]$ – помилки передбачення, що є некорельованими випадковими відліками процесу з дисперсією D_a, p – порядок моделі АР. Перший індекс у коефіцієнтів АР вказує на порядок моделі. Модель (1) оптимальна, якщо

$$E\{a[t]a[t - i]\} = 0, \; i > 0,$$

де $E\{\}$ – оператор статистичного усереднювання. Умова оптимальності дозволяє отримати рівняння і визначити критерії для оцінки параметрів моделі – коефіцієнтів АР і порядку моделі.

Співвідношення для розрахунку оптимальних оцінок коефіцієнтів АР отримують з рівняння (1) і носить назву рівняння Юла – Уокера

$$R[i] - \sum_{j=1}^{p} \Phi[p, j] R[j - i] = 0, \; i = 1, ..., p. \tag{2}$$

У співвідношенні (2) використовуються незміщені оцінки значень кореляційної функції $R[i]$. Параметричне представлення СЩП АР процесу в загальному випадку багатомодове і описується через параметри моделі виразом

$$P(f) = \frac{D_a}{\left| 1 - \sum_{i=1}^{p} \Phi[p, i] e^{-j 2\pi f i T} \right|^2}. \tag{3}$$

Як видно з формули (3), всього декілька параметрів моделі АР несуть повну інформацію про СЩП випадкового процесу.

Метод факторизації СЩП

Факторизація СЩП дозволяє представити багатомодову СЩП, що описується виразом (3), простішими, зокрема одномодовими СЩП. Щоб показати можливість розкладання багатомодового параметричного спектру на простіші одномодові складові, припускаємо, що модель АР представлена мультиплікативною моделлю АР, тобто описується моделлю виду $\mathrm{AP}_1 \times \mathrm{AP}_2 \times \ldots \times \mathrm{AP}_\mathrm{k}$ [13]. Параметри складових моделей АР легко обчислюються по коренях характеристичного рівняння моделі АР

$$c^p - \Phi[p,1] c^{p-1} - ... - \Phi[p, p] = \prod_{i=1}^{p} (c - c[p, i]) = 0, \tag{4}$$

де невідомі $c[p,i]$ є коренями характеристичного рівняння, яке легко вирішується в загальному випадку чисельними методами. У виразі (4) перший індекс у коренів рівняння $c[p,i]$ вказує на порядок вихідної моделі АР.

Розглянемо різні варіанти можливих рішень рівняння (4). Для простоти спочатку розглянемо випадок, коли корені комплексно зв'язані. Замітимо, що комплексні корені характеристичного рівняння (4) попарно комплексно зв'язані, оскільки в лівій частині маємо дійсне рівняння. За наявності цих коренів легко знайти відповідні коефіцієнти АР другого порядку:

$$\Phi[2,1]=c[2,1]+c[2,2], \qquad (5)$$
$$\Phi[2,2]=-c[2,1]c[2,2].$$

Для іншої пари комплексно зв'язаних коренів можна знайти інші пари коефіцієнтів АР. Таким чином, для нових коренів характеристичного рівняння знаходимо нові пари коефіцієнтів АР, що формують одну моду СЩП. Використовуючи комплексно зв'язані корені, отримуємо факторизацію багатомодової СЩП, тобто її розкладання на одномодові СЩП, що відповідають кожній парі коефіцієнтів АР.

Для дійсного кореня характеристичного рівняння (4) зв'язок з коефіцієнтом АР простий

$$\Phi[1,1]=\pm c[1,1]. \qquad (6)$$

Нагадаємо, що в цьому випадку мода в СЩП розташована або на нульовій частоті, або на максимальній частоті. Отже, модель АР p-го порядку з довільним числом мод можна представити у вигляді набору коефіцієнтів АР першого або другого порядків. Параметричне представлення спектру моделі (3), що залежить тільки від коефіцієнтів моделі АР, представляється набором одномодових спектрів. У такий спосіб здійснюється факторизація СЩП, що полягає в розкладанні параметричної багатомодової СЩП на одномодові складові. Запропонований метод факторизації спектрів і оцінювання частоти моди може застосовуватися для підвищення роздільної здатності СЩП. Тоді багатомодовий спектр, оцінювання якого здійснюється по моделі АР p-го порядку, представляється (факторизується) набором одномодових спектрів першого або другого порядків. Таким чином, багатомодовий спектр (3) представляється виразом

$$P(f)=\frac{D_a}{\left|\sum_{i=0}^{p_1}\Phi_1[p_1,i]e^{-j2\pi fiT}\right|^2\left|\sum_{i=0}^{p_2}\Phi_2[p_2,i]e^{-j2\pi fiT}\right|^2\times\ldots\times\left|\sum_{i=0}^{p_k}\Phi_k[p_k,i]e^{-j2\pi fiT}\right|^2}. \qquad (7)$$

У виразі (7) коефіцієнти АР описують одномодові складові СЩП.

Зв'язок частотних характеристик мод СЩП з коренями характеристичного рівняння

Корені характеристичного рівняння можна виразити через центральні частоти мод і їх ширини смуг в СЩП. Для цього використовується зв'язок між коренями характеристичного рівняння і параметрами спектру: частотою мод f_i і її шириною смуги СЩП Δf. Тоді дійсні корені характеристичного рівняння p-го порядку і параметри АР, пов'язані співвідношенням (6), виражаються через ширину смуги моди [14]:

$$c[p,i]=\pm e^{-\pi\Delta f_i T}. \qquad (8)$$

Дійсні корені описують моди на нульовій частоті або на максимальній частоті дискретного спектру, що дорівнює половині частоти дискретизації.

Оскільки модель АР дійсна, то комплексні корені характеристичного рівняння описуються через центральні частоти і ширину смуги частот спектральних мод наступними виразами

$$c[p,i] = e^{-\pi\Delta f_i T - j2\pi f_i T}, \quad c[p,i+1] = e^{-\pi\Delta f_{i+1} T + j2\pi f_{i+1} T}. \tag{9}$$

У простих випадках вузькосмугових СЩП на один дійсний корінь доводиться одна мода на нульовій частоті або на максимальній частоті. На два комплексно-зв'язаних корені характеристичного рівняння припадає два порядки моделі АР або одна мода на ненульовій частоті, або не на максимальній частоті. В якості характеристик сигналів можна використати центральні частоти мод. Якщо корені характеристичного рівняння комплексні, то частота моди визначається так:

$$f = ar\cos\{\Phi[2,1]/(\sqrt{-\Phi[2,2]}\}/2\pi T, \tag{10}$$

а ширина полоси моди визначається виразом

$$\Delta f = \ln(-\Phi[2,2])/2\pi T. \tag{11}$$

Якщо корені дійсні, то модель АР має перший порядок. При цьому частоти мод f=0, якщо Ф[1,1]>0 або f=1/2T, якщо Ф[1,1]<0. Ширина смуги моди в цьому випадку визначається, як випливає з (8), виразом

$$\Delta f = \ln(|\Phi[1,1]|)/\pi T.$$

Зауважимо, що корені характеристичного рівняння, як випливає з (4), повністю характеризують модель, як і коефіцієнти АР. Таким чином, усі моди СЩП можна визначити безпосередньо через корені рівняння, використовуючи співвідношення (8) і (9). Так, для комплексних коренів з (9) – (11) маємо

$$f = ar\cos\{(c[2,1]+c[2,2])/2\exp(\pi\Delta fT)\}/2\pi T, \tag{12}$$

$$\Delta f = -\ln(|c[2,1]c[2,1]|)/2\pi T.$$

Роздільна здатність оцінки СЩП при використанні факторизації

Багатомодові випадкові процеси часто визначаються дією різних чинників, що призводить до формування спектральних мод на різних частотах. У деяких випадках корисно описувати такі складні процеси полімоделлю виду $АР_1 \times АР_2 \times \ldots \times АР_k$ [15]. Така модель характеризується полюсами передаточної функції, які виражаються через комплексно зв'язані пари. Представлення такої системи моделей через множення полюсів відповідає одній із форм її реалізації. Перемноження складових функції передачі відповідає послідовному (каскадному) включенню фільтрів першого і другого порядків з дійсними коефіцієнтами. Така послідовна реалізація часто використовується на практиці, оскільки вона дозволяє ослабити небажані ефекти, пов'язані з помилками округлення.

Оскільки кожна модель АР, що характеризується своїми коефіцієнтами АР, визначає СЩП процесу, то можна цю полімодель представити у вигляді мультиплікативної, тобто послідовної моделі. Представимо мультиплікативну багатомодову модель набором мономоделей із СЩП, що містить одну або декілька мод. Можна факторизувати багатомодову СЩП процесу на окремі моди і проаналізувати вплив різних чинників на формування СЩП.

Покажемо процес факторизації на прикладі моделі, що використовується при дослідженні роздільної здатності СЩП. Нехай часовий ряд описується мультиплікативною моделлю $АР_1 \times АР_2 \times \ldots \times АР_k$. Враховуючи умову оптимальності, для знаходження коефіцієнтів АР маємо узагальнене нелінійне рівняння типу Юла – Уокера

$$\Phi_k(z)\Phi_{k-1}(z)...\Phi_1(z)R[j] = 0. \tag{13}$$

де z – оператор сдвигу назад.

Ефективний спосіб вирішення цього рівняння продемонстровано нижче на прикладі моделі АР четвертого порядку, що використовується для дослідження підвищення роздільної здатності методом факторизації. Коефіцієнти двох складових моделей другого порядку $\Phi_1[1]$, $\Phi_1[2]$, $\Phi_2[1]$, $\Phi_2[2]$ повинні розраховуватися шляхом вирішення системи нелінійних рівнянь, яку можна отримати з виразу (13):

$$\begin{aligned}
\Phi[4,1] &= \Phi_1[2,1] + \Phi_2[2,1], \\
\Phi[4,2] &= \Phi_1[2,2] - \Phi_1[2,1]\Phi_2[2,1] + \Phi_2[2,2], \\
\Phi[4,3] &= -\Phi_1[2,2]\Phi_2[2,1] - \Phi_1[2,1]\Phi_2[2,2], \\
\Phi[4,4] &= -\Phi_1[2,2]\Phi_2[2,2].
\end{aligned} \tag{14}$$

Ця процедура чисельного вирішення системи нелінійних рівнянь вимагає значного обсягу обчислень і трудомісткого відкидання зайвих коренів. Факторизація дозволяє спростити цю задачу, якщо необхідно розкласти СЩП моделі АР четвертого порядку на одномодові складові, що описуються моделями другого порядку. Для мультиплікативної моделі четвертого порядку $АР_1 \times АР_2$, представленої двома моделями другого порядку зі своїми комплексно-зв'язаними коренями характеристичного рівняння, завдання відшукання коефіцієнтів АР(2) спрощується. За наявності цих коренів відповідні коефіцієнти АР другого порядку мають вид:

$$\begin{aligned}
\Phi_1[2,1] &= c_1[2,1] + c_1[2,2], \\
\Phi_1[2,2] &= -c_1[2,1]c_1[2,2], \\
\Phi_2[2,1] &= c_2[2,1] + c_2[2,2], \\
\Phi_2[2,2] &= -c_2[2,1]c_2[2,2],
\end{aligned} \tag{15}$$

де нижній індекс вказує на номер моделі, рівний в даному випадку один або два.

Таким чином, при факторизації СЩП немає потреби знаходити коефіцієнти АР мультиплікативної моделі четвертого порядку $АР_1 \times АР_2$, Досить знайти комплексно-зв'язані корені і виразити через них відповідні коефіцієнти АР для моделей другого порядку.

Факторизація параметричних спектрів на їх мультиплікативні моделі дає точніше уявлення про ці складові. Ця властивість може бути використана для підвищення роздільної здатності параметричної оцінки спектрів. Якщо дві моди розташовані близько одна від одної, то їх параметрична оцінка СЩП може бути представлена як одна мода, тобто неможливо розділити ці моди. Проте особливості цієї моди поблизу своєї вершини, спостережувані як її розширення, дозволяє шляхом факторизації розділити вершину на дві моди. Факторизація є досить чутливим інструментом аналізу спектрів. Для підтвердження цього ефекту була проведена серія розрахунків і статистичних експериментів.

Розраховувалися коефіцієнти АР(4) для двомодового спектру з близько розташованими модами. Відповідно до теорії побудови мультиплікативних моделей спочатку оцінювалися параметри моделі АР(4) відомими методами параметричного спектрального оцінювання, а потім по ним розраховувалися коефіцієнти АР(2) мультиплікативної моделі. Для первинного оцінювання коефіцієнтів АР можуть бути використані різні методи: незміщені автокореляційні оцінки Юла – Уокера, зміщені автокореляційні оцінки Юла – Уокера, геометричний метод, метод Берга, метод Кея, коваріаційні методи, модифікований коваріаційний метод [16].

Експериментальні дослідження роздільної здатності СЩП

Проведено дослідження роздільної здатності з використанням незміщених автокореляційних оцінок у методі Юла – Уокера і гармонійного алгоритму (методу Берга). Незміщені оцінки в методі Юла – Уокера отримано з використанням незміщених оцінок функції кореляції процесу. У методі Берга коефіцієнти АР отримано рекурсивно алгоритмом Левінсона –

Дарбіна [16], який дозволяє рекурентно обчислювати коефіцієнти АР по коефіцієнтах відбиття гратчатого фільтру (ГФ) $K[n]$. Коефіцієнти відбиття n-ї ланки ГФ обчислювалися згідно з виразом

$$K[n] = \frac{-2\sum_{t=n+1}^{N} a_{n-1}[t]\, d_{n-1}[t-1]}{\sum_{t=n+1}^{N} \left|a_{n-1}[t]\right|^2 + \sum_{t=n+1}^{N} \left|d_{n-1}[t-1]\right|^2}, \tag{16}$$

де $a_{n-1}[t]$, $d_{n-1}[t-1]$ – помилки прямого і зворотного передбачення $(n-1)$-ї ланки ГФ. Помилки прямого передбачення в (16) обчислювались згідно з

$$a_{n-1}[t] = a_{n-2}[t] + K[n-1]\, d_{n-2}[t-1]\ . \tag{17}$$

Аналогічно виражаються помилки зворотного передбачення:

$$d_{n-1}[t] = d_{n-2}[t-1] + K[n-1]\, a_{n-2}[t]. \tag{18}$$

Вирази (17) і (18) витікають із структури РФ, зображеної на рис. 1.

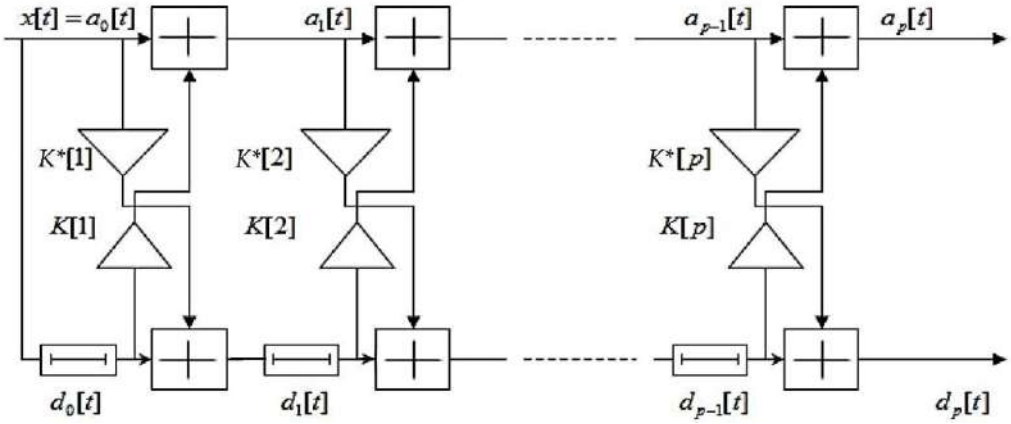


Рис. 1. Структура ГФ

Алгоритм Левінсона – Дарбіна використано для рекурентного обчислення коефіцієнтів АР згідно з співвідношенням

$$\begin{gathered} K^*[n] = K[n] = \Phi[n,n] = -\frac{R[n] + \sum_{i=1}^{n-1} \Phi[n-1,i]\, R[n-i]}{\left(\sigma_{a,n-1}\right)^2}, \quad n = 2,\ldots,p, \\ \Phi[n,i] = \Phi[n-1,i] + \Phi[n,n]\,\Phi[n-1,n-i], \quad \left(\sigma_{a,n}\right)^2 = \left(1-\Phi^2[n,n]\right)\left(\sigma_{a,n-1}\right)^2, \end{gathered} \tag{19}$$

з ініціалізацією

$$\Phi[1,1] = \frac{-R_1}{\left(\sigma_x\right)^2}\ , \quad \left(\sigma_{a,1}\right)^2 = \left(1-\Phi^2[1,1]\right)\left(\sigma_x\right)^2\ ,$$

де σ_x – середньоквадратичне відхилення (СКВ) процесу, $\sigma_{a,n}$ – СКВ помилки передбачення для моделі АР n-го порядку.

Оцінки коефіцієнтів відбиття (16) використовувалися для розрахунку коефіцієнтів АР відповідно до алгоритму Левінсона – Дарбіна (19).

Дослідження проводилися методом статистичного моделювання. При цьому генерувався випадковий процес АР(4). По заданим частотним характеристикам мод СЩП розраховували-

ся корені характеристичного рівняння з (8) і (9). Знайдені у такий спосіб корені характеристичного рівняння використовувалися для розрахунку коефіцієнтів моделі АР(4) при формуванні досліджуваного процесу АР з близько розташованими в СЩП модами.

Наведемо формули, що зв'язують коефіцієнти АР з коренями характеристичного рівняння, що витікає з (4) для $p = 1 \div 4$:

$$
\begin{aligned}
&\Phi[1,1] = \pm c[1],\\
&\Phi[2,1] = c[2,1] + c[2,2];\\
&\Phi[2,2] = -c[2,1]c[2,2],\\
&\Phi[3,1] = c[3,1] + c[3,2] + c[3,3];\\
&\Phi[3,2] = -(c[3,1]c[3,2] + c[3,1]c[3,3] + c[3,2]c[3,3]),\\
&\Phi[3,3] = c[3,1]c[3,2]c[3,3],\\
&\Phi[4,1] = c[4,1] + c[4,2] + c[4,3] + c[4,4];\\
&\Phi[4,2] = -(c[4,3]c[4,4] + c[4,2]c[4,3] + c[4,1]c[4,3] + c[4,4]c[4,1] + c[4,2]c[4,3] + c[4,2]c[4,4])\\
&\Phi[4,3] = c[4,1]c[4,3]c[4,4] + c[4,2]c[4,3]c[4,4] + c[4,1]c[4,2]c[4,3] + c[4,1]c[4,2]c[4,4]\\
&\Phi[4,4] = -c[4,1]c[4,2]c[4,3]c[4,4].
\end{aligned}
\qquad (20)
$$

У виразах (20) перший індекс в квадратних дужках вказує на відповідний порядок моделі згенерованого випадкового процесу. Потім згідно з (1) формувався випадковий процес з використанням породжуючого процесу у вигляді білого гаусівського шуму [17].

Результати дослідження роздільної здатності СЩП з використанням факторизації

Підвищення роздільної здатності шляхом факторизації СЩП проілюстроване на прикладі випадкового процесу АР(4) з параметрами двох спектральних мод СЩП: з центральними частотами $f_1 = 110$, $f_2 = 130$ і шириною смуги частот $df_1 = df_2 = 1$. По заданим частотним параметрам, згідно з (9), розраховувалися чотири попарно комплексно-зв'язані корені характеристичного рівняння. Потім, з використанням виразу (20), знаходилися коефіцієнти АР четвертого порядку. За цими даними розраховувалася теоретична СЩП згідно з (3). Для отримання вибірки випадкового процесу використовувався формуючий фільтр, на виході якого з відліків білого гаусівського шуму по знайденим коефіцієнтам АР(4) формувалася вибірка випадкового процесу згідно з рівнянням (1). Довжина вибірки процесу складала 100 відліків.

Методами Юла – Уокера і Берга оцінювалися коефіцієнти АР(4) і знаходився вибірковий параметричний спектр четвертого порядку відповідно до (3). Для отримання оцінок СЩП шляхом факторизації за вибіркою процесу АР(4) оцінювалися попарно комплексно-зв'язані корені характеристичного рівняння процесу. Потім по цим кореням розраховувалися дві пари коефіцієнтів АР(2), використовуючи вираз (15). Факторизована СЩП знаходилася відповідно до виразу (7). Теоретична СЩП для моделі АР(4) показана на рис. 2. По сформованому процесу АР(4) методами Юла – Уокера і Берга оцінювалися коефіцієнти АР(4). Параметрична оцінка спектру по моделі АР(4), що отримана відповідно до (3) для незміщених оцінок Юла – Уокера, представлена на рис. 3. Аналіз графіку СЩП показує, що спектральні моди в цьому випадку не розділяються.

Відомо, що роздільна здатність параметричної оцінки СЩП, що отримана методом Берга, є вищою ніж оцінки, які отримано методом Юла – Уокера [16]. Проте, навіть, параметрична оцінка спектру згідно з методом Берга не дає хорошої роздільної здатності при частотних параметрах спектру досліджуваного випадкового процесу (рис. 4).

У тих випадках, коли параметричні оцінки СЩП дають для близько розташованих частот слабо помітні моди або спостерігаються нерівномірності графіку СЩП поблизу вершини моди, підвищити роздільну здатність можна застосуванням факторизації мультиплікативної моделі АР. На рис. 5 представлено два графіки СЩП, побудованих по моделях АР1(2) і

АР2(2), отриманих шляхом факторизації. Параметри складових моделей знайдено за оцінками Юла – Уокера згідно з рівнянням (2). Хоча точність оцінок частот мод не така висока, як отримана методом Берга, розділення двох мод шляхом факторизації очевидна. Дослідження показали, що розділення мод покращується при застосуванні факторизації, якщо моди навіть слабо розділяються, тобто провал між модами має малу глибину. У тих випадках, коли замість нерівномірності мод спектральна оцінка дає одну гладку моду, застосування факторизації не приводить до розділення мод.

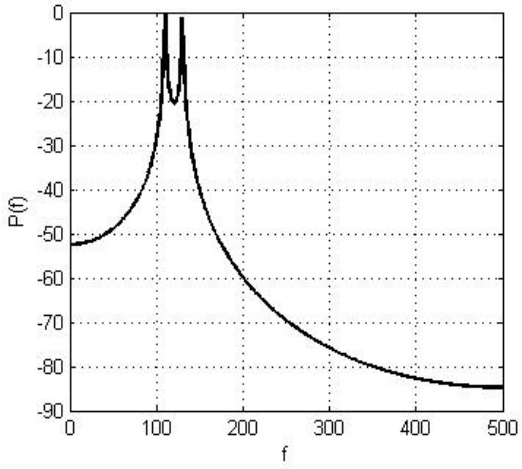


Рис. 2. Теоретична СЩП для моделі АР(4) з параметрами $f_1 = 110,\ f_2 = 130,\ df_1 = df_2 = 1$

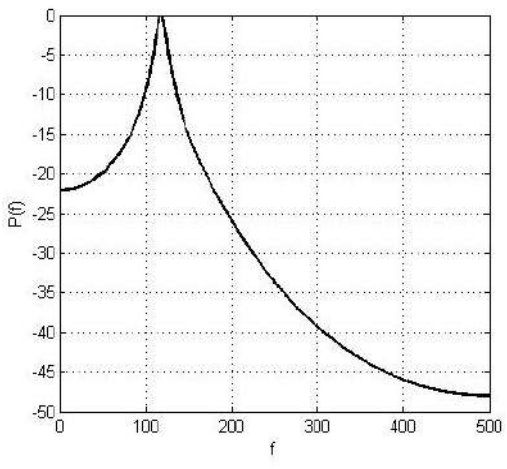


Рис. 3. Параметрична оцінка спектру для моделі АР(4) методом Юла – Уокера при $f_1 = 110,\ f_2 = 130,\ df_1 = df_2 = 1$

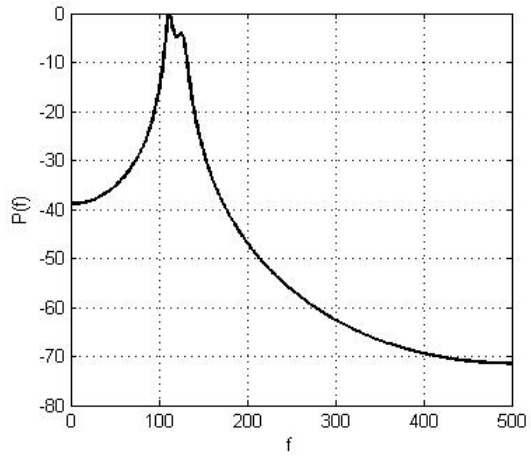


Рис. 4. Параметрична оцінка спектру для моделі АР(4) методом Берга при $f_1 = 110,\ f_2 = 130,\ df_1 = df_2 = 1$

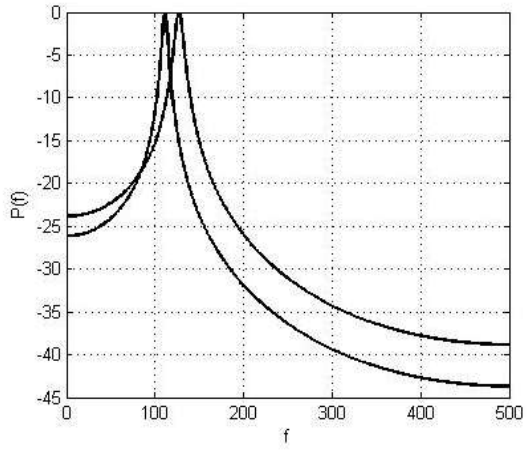


Рис. 5. Мультиплікативне представлення спектру для моделі АР(4) методом Юла – Уокера при $f_1 = 110,\ f_2 = 130,\ df_1 = df_2 = 1$

При факторизації надрозділення за двома оцінками СЩП складових моделей АР(2) зручно використати один графік СЩП. Тому запропоновано побудувати новий графік СЩП так, щоб ліва на графіку частина кривої співпадала з оцінкою СЩП лівої моди до частоти, рівній частоті лівого графіку, плюс половина різниці частот лівої і правої мод, тобто

$$f_{\max} = (f_2 - f_1)/2,$$

де f_1, f_2 – центральні частоти правої і лівої мод. Тоді права частина кривої співпадає з оцінкою СЩП другої моди, починаючи з частоти $f_{\max} + 1.$ Графік такого представлення розділення мод шляхом факторизації показано на рис. 6. Порівняння графіків на рис. 3, 6 демонструє ефект отримання надрозділення шляхом факторизації оцінки СЩП відповідно до методу Юла – Уокера. Роздільну здатність СЩП можна істотно підвищити, якщо для первинного оцінювання коефіцієнтів АР використати геометричний метод, метод Берга, метод

Кея, коваріаційні методи або модифікований коваріаційний метод, які мають помітну перевагу в порівнянні з оцінками Юла – Уокера. Нижче продемонстровано істотне поліпшення роздільної здатності на прикладі методу Берга. На рис. 7 представлено графік теоретичного спектру, отриманого для моделі АР(4) із заданими параметрами СЩП: $f_1 = 110$, $f_2 = 115$, $df_1 = df_2 = 1$.

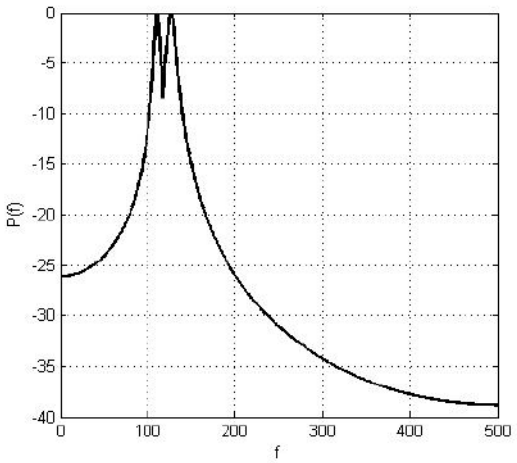


Рис. 6. Мультиплікативна оцінка спектру для моделі АР(4) згідно з методом Юла – Уокера при $f_1 = 110$, $f_2 = 130$, $df_1 = df_2 = 1$

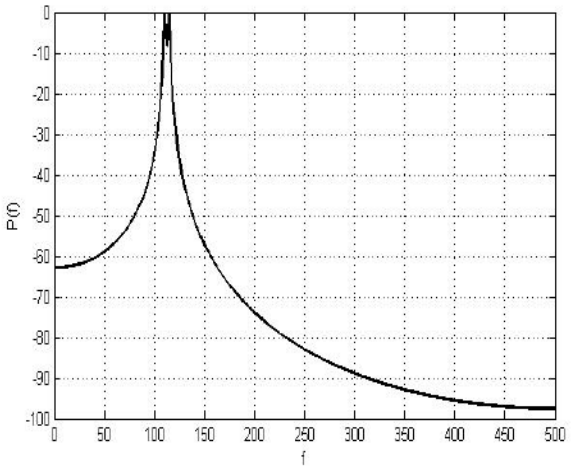


Рис. 7. Теоретична СЩП для моделі АР(4) згідно з методом Юла – Уокера при $f_1 = 110$, $f_2 = 115$, $df_1 = df_2 = 1$

На рис. 8, 11 представлено результати експериментів з дослідження роздільної здатності спектрів для процесу АР(4). Параметрична оцінка спектру моделі АР(4) згідно з методом Юлу – Уокера не дає можливості застосувати факторизації для підвищення роздільної здатності, оскільки графік моди СЩП гладкий без ознак наявності другої моди (рис. 8). На графіку параметричної оцінки спектру для моделі АР(4) згідно з методом Берга (рис. 9) є нерівномірність поблизу моди, що дозволяє розділити дві моди застосуванням методу факторизації. Аналіз графіків на рис. 9 – 11 показує, що для цих частот також можна отримати роздільну здатність на основі факторизації, хоча відстань між частотами, що задаються, істотно менша, у порівнянні з попереднім випадком ($f_1 = 110$, $f_2 = 130$, $df_1 = df_2 = 1$).

Аналіз графіків показує безперечну перевагу мультиплікативного розкладання для підвищення роздільної здатності параметричних оцінок спектрів. Застосування для розрахунків параметрів моделей АР, отриманих методом Берга, сприяє підвищенню спектрального розділення вибіркових оцінок.

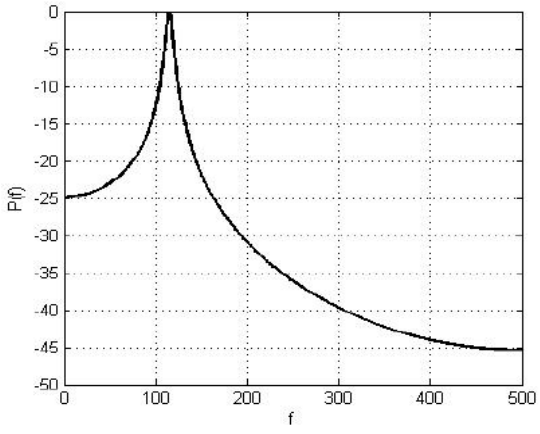


Рис. 8. Параметрична оцінка спектру для моделі АР(4) згідно з методом Юла – Уокера при $f_1 = 110$, $f_2 = 115$, $df_1 = df_2 = 1$

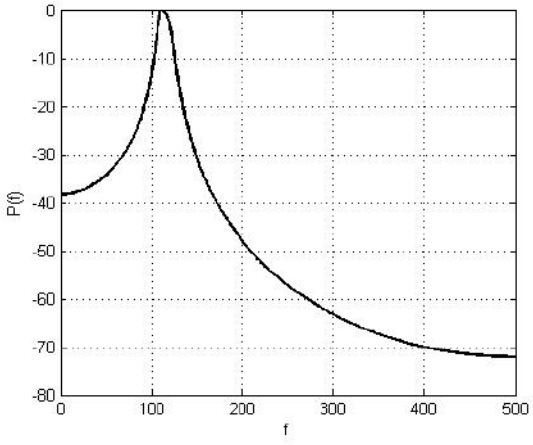


Рис. 9. Параметрична оцінка спектру для моделі АР(4) згідно з методом Берга при $f_1 = 110$, $f_2 = 115$, $df_1 = df_2 = 1$

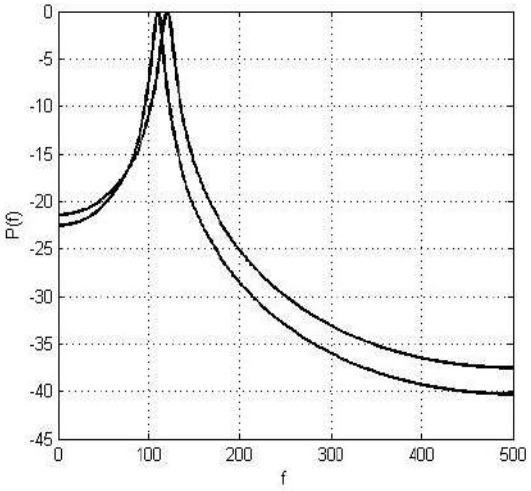


Рис. 10. Мультиплікативне представлення спектру моделі АР(4) згідно з методом Берга при $f_1 = 110,\ f_2 = 115,\ df_1 = df_2 = 1$

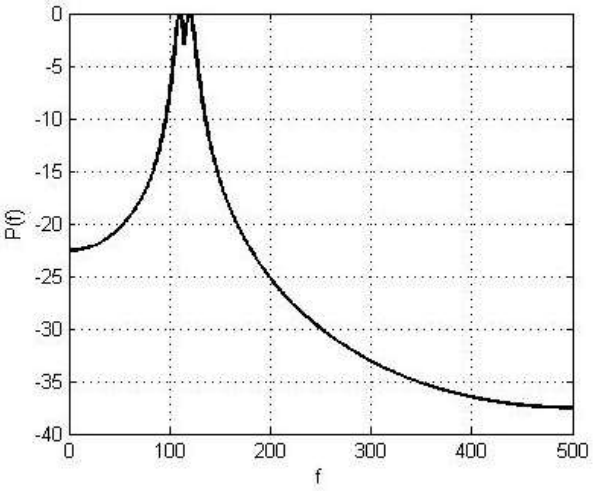


Рис. 11. Мультиплікативна оцінка спектру для моделі АР(4) згідно з методом Берга при $f_1 = 110,\ f_2 = 115,\ df_1 = df_2 = 1$

На рис. 12 і 13 представлено результати розділення двох мод в параметричній оцінці СЩП для різних довжин вибірок процесу при використанні для оцінки коефіцієнтів АР методу Берга. Як видно з графіків, при довжині 200 відліків процесу можна отримати роздільну здатність оціни СЩП для двох мод при наступних заданих параметрах мод $f_1 = 110$, $f_2 = 115$, $df_1 = df_2 = 1$. Використання факторизації дозволяє отримати роздільну здатність оцінки СЩП для цих параметрів мод уже при довжині вибірки 100 відліків процесу (рис. 11).

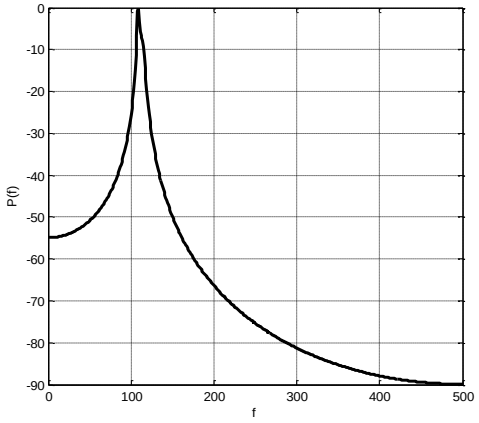


Рис. 12. Мультиплікативне представление спектру для моделі АР(4) методом Берга по вибірці довжиною 150 відліків процесу при $f_1 = 110,\ f_2 = 115,\ df_1 = df_2 = 1$

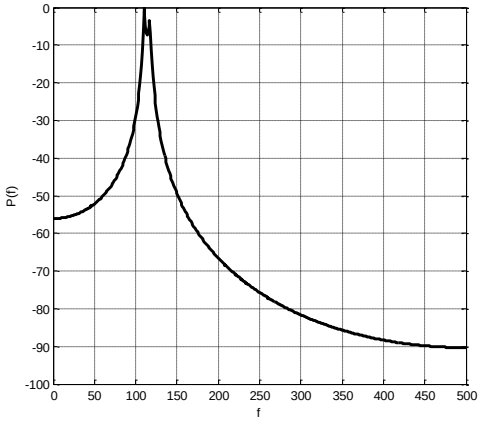


Рис. 13. Мультиплікативна оцінка спектру для моделі АР(4) методом Берга по виборці довжиною 200 відліків процесу при $f_1 = 110,\ f_2 = 115,\ df_1 = df_2 = 1$

На рис. 14 і 15 показано роздільну здатність, отримувану шляхом факторизації мод СЩП для різних ділянок частот. Довжина вибірки складала 100 відліків процесу. Розділення розглядалось на частотах $f_1 = 50,\ f_2 = 65,\ df_1 = df_2 = 1$ (рис. 14) і на частотах $f_1 = 230,\ f_2 = 245,\ df_1 = df_2 = 1$ (рис. 15). Хоча точність оцінки частот трохи відрізнялася від частот, що задавалися, спостерігалось упевнене розділення мод.

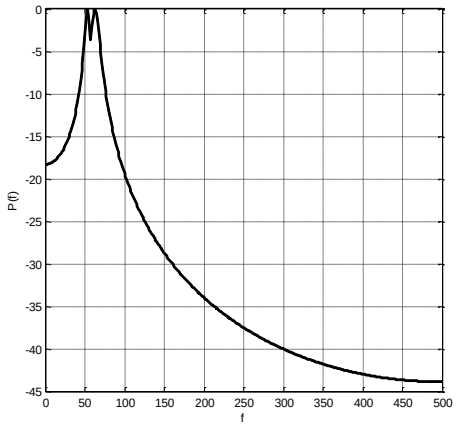


Рис. 14. Мультиплікативне представлення спектру для моделі АР(4) по вибірці 100 відліків процесу методом Берга при $f_1 = 50$, $f_2 = 65$, $df_1 = df_2 = 1$

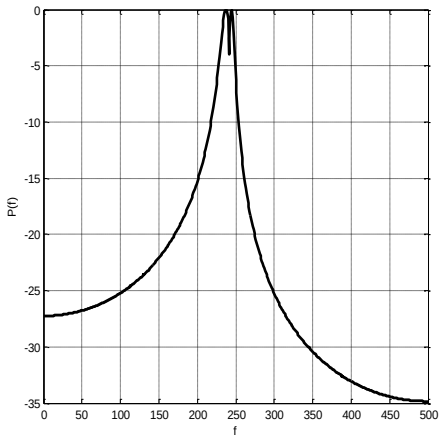


Рис. 15. Мультиплікативна оцінка спектру для моделі АР(4) методом Берга по вибірці 100 відліків при $f_1 = 230$, $f_2 = 245$, $df_1 = df_2 = 1$

Результати досліджень спектрального розділення на основі факторизації з використанням коефіцієнтів АР, що обчислюються методом Берга, показали перевагу в порівнянні з випадком використання методу Юла – Уокера.

Висновки

1. Представлено результати досліджень підвищення спектральної роздільної здатності на основі факторизації СЩП випадкового процесу, яка дозволяє розкласти багатомодову СЩП на простіші одномодові складові. Наведено теоретичне обгрунтування методу факторизації СЩП випадкового процесу на основі моделі АР.

2. Наведено результати експериментальних досліджень підвищення роздільної здатності СЩП випадкового процесу при використанні факторизації. Дослідження виконані шляхом статистичного моделювання на вибірках випадкового процесу, які отримані з допомогою формуючого фільтру. Результати досліджень показують перевагу факторизації з використанням методу Берга в порівнянні з методом Юла – Уокера.

3. Підвищення роздільної здатності оцінки СЩП випадкових процесів має важливе значення при вирішенні прикладних задач обробки сигналів в області радіолокації, радіозв'язку, технічної і медичної діагностики [18]. Це дає можливість більш точно проводити роздільний спектральний аналіз низькочастотних, середньочастотних і високочастотних складових СЩП спостережуваних процесів.

Список літератури:

1. G. Box, G. Jenkins, G.C. Reinsel. Time Series Analysis, Forecasting and Control, 4th ed. Hoboken. USA : Wiley, 2008.

2. Brockwell P.J., Davis R.A. Introduction to Time Series and Forecasting. Springer, 2002.

3. Kitagawa G., Gersch W. Smoothness Priors Analysis of Time Series. New York : Springer, 1996.

4. Drubin J., Koopman S. Time Series Analysis by State Space Methods. Oxford : Oxford Univ. Press, 2008.

5. Hyndman R., Koehler A., Snyde R., Grose S. A state space framework for automatic forecasting using exponential smoothing methods // Int. J. Forecast. 2002. Vol. 18, no. 3. Pp. 439 – 454.

6. Young P.C. Stochastic, dynamic modelling and signal processing: Time variable and state dependent parameter estimation // Nonlinear and Nonstationary Signal Processing. Cambridge : Cambridge Univ. Press, 2000. Pp. 41 – 114.

7. Young P.C. Recursive estimation and time series analysis // Introduction for the Student and Practioner. Berlin : Springer-Verlag, 2011.

8. Tong H. Nonlinear Time Series: A Dynamical Systems Approach. Oxford : Oxford Univ. Press, 1990.

9. Ozaki T. Non-linear time series models for non-linear random vibrations // J. Appl. Probabil. 1980. Vol. 17. Pp. 84–93.

10. Haggan V., Ozaki T. Modeling nonlinear random vibrations using an amplitude-dependent autoregressive time series model // Biometrika. 1981. Vol. 68. Pp. 189 – 196.

11. Priestley M.B. State dependent models: A general approach to nonlinear time series analysis // J. Time Series Anal. 1980. Vol. 1, no 1. Pp. 57 – 71.

12. Chen R., Tsay R.S. Functional-coefficient autoregressive models // Amer. Statist. Assoc. 1993. Vol. 88, no. 21. Pp. 298 – 308.

13. Tykhonov V.A., Kudriavtseva N.V., Chmelar P. Factorization of speech signals parametric spectra using multiplicative linear prediction models // Proceedings Elmar. 57th International Symposium ELMAR-2015, Zadar, 28 – 30 September. 2015. Pp. 124 – 130.

14. Кармалита В.А. Цифровая обработка случайных колебаний. Москва : Мир, 1989.

15. Карташов В.М., Олейников В.Н., Тихонов В.А. и др. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : Компания СМИТ, 2014.

16. Марпл мл. С.Л. Цифровой спектральный анализ и его приложения. Москва : Мир, 1990.

17. Тихонов В.А., Русановский Д.Е., Тихонов Д.В. Генерирование узкополосных имитационных случайных процессов // Радиоэлектроника и информатика. 1999. №4. С. 83 – 85.

18. Омельченко В.А., Безрук В.М., Коваленко Н.П. Распознавание заданных радиосигналов при наличии неизвестных сигналов на основе авторегрессионной модели // Радиотехника. 2001. Вып. 123. С. 195 – 199.

Надійшла до редколегії 05.03.2023

Відомості про авторів:

Тихонов В'ячеслав Анатолійович – д-р фіз.-мат. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; e-mail: vyacheslav.tykhonov@nure.ua; ORCID: http://orcid.org/0000-0002-4618-4787

Безрук Валерій Михайлович – д-р, техн. наук, професор, Харківський національний університет радіоелектроніки, Україна; e-mail: valerii.bezruk@nure.ua, ORCID: https://orcid.org/0000-0003-2349-7788

УДК 537.868 **DOI:10.30837/rt.2023.1.212.09**

Г.Л. КОМАРОВА, канд. техн. наук

ВПЛИВ ФЕРОМАГНІТНОГО РЕЗОНАНСУ НА ПЕРЕТВОРЕННЯ ЕНЕРГІЇ НВЧ СИСТЕМОЮ З ДВОХ ЦИЛІНДРІВ У МЕХАНІЧНУ

Вступ

Відомі наукові розробки, спрямовані на створення перетворювачів електромагнітної енергії НВЧ в механічну енергію, показують, що вони мають надзвичайно малу силу тяги $4 \cdot 10^{-6}$ мН/кВт [1], 0,6 мН/кВт [2, 3], 1,2 мН/кВт [4], 53 мН/кВт [5]. Основними елементами таких перетворювачів є генератор НВЧ, електромагнітна хвиля, металевий екран, діелектричний еліпсоїд, феритовий циліндр або сфера.

Цей недолік не дозволяє використовувати їх у промислових умовах. Тому розробка нових методів перетворення електромагнітної енергії НВЧ на механічну енергію є актуальною.

Метою роботи є удосконалення методу перетворення електромагнітної енергії НВЧ на механічну енергію.

1. Аналіз останніх досліджень та публікацій

Для вдосконалення методу перетворення електромагнітної енергії на механічну використано результати наступних досліджень.

У роботі [4] досліджено силову дію неоднорідної електромагнітної хвилі на діелектричний диск, поміщений у внутрішній частині замкнутого несиметричного металевого об'єму. Рівень силового впливу неоднорідної хвилі становить 1,2 мН/кВт. Не досліджено вплив геометричних розмірів диска на силовий вплив енергії НВЧ, вплив неоднорідної електромагнітної хвилі на систему діелектричних дисків.

У роботі [6] запропоновано метод перетворення енергії неоднорідної електромагнітної хвилі на механічну енергію. Метод перетворення полягає у виготовленні перетворювача з феромагнетику та впливі на нього постійним та неоднорідним електромагнітним полем. Величина напруженості постійного магнітного поля відповідає виникненню феромагнітного резонансу. Технічний результат – збільшення коефіцієнта перетворення електромагнітної енергії на механічну енергію.

У роботі [2] експериментально показано, що електромагнітна хвиля потужністю 10 Вт, що розповсюджується у прямокутному хвилеводі з поперечним перерізом $10 \cdot 23$ мм^2, діє на феритову сферу діаметром 3.55 мм, поміщену в постійне магнітне поле з силою, що дорівнює $(6 \pm 0{,}5) \cdot$мкН. Величина напруженості постійного магнітного поля відповідає виникненню феромагнітного резонансу. Не досліджено силову дію стоячої хвилі на систему феритових сфер.

У роботі [3] методом фізичного моделювання отримано алгоритм обчислення сили, з якою стояча електромагнітна хвиля діє на феритову сферу довільного діаметра, поміщену в постійне магнітне поле. Розмір напруженості постійного магнітного поля відповідає виникненню феромагнітного резонансу. Стояча електромагнітна хвиля, створена падаючою із щільністю потоку 622 кВт/м^2 і довжиною 3,2 см і відбитою від металевого екрану, діє на феритову сферу, резонансний радіус якої дорівнює 4,2634 мм, з силою, що дорівнює 0,12 Н. Не досліджено силову дію стоячої хвилі на систему феритових сфер.

У роботі [5] досліджено перетворення енергії стоячої електромагнітної хвилі на феритовому циліндрі, який поміщений у постійне магнітне поле. Стояча електромагнітна хвиля, створена падаючою із щільністю потоку 622 кВт/м^2 і довжиною 3,2 см і відбитою від металевого екрану. Центр феритового циліндра, довжина якого дорівнює 0,64 м, а резонансний радіус − 2,808 мм розташований від металевого екрану на відстані рівній ($\lambda_o/8 + n \cdot \lambda_o/2$, де $n = 0, 1, 2, 3\ldots$). Стояча електромагнітна хвиля діє на феритовий циліндр із силою, що дорівнює 10,6 Н. Не досліджено силову дію стоячої хвилі на систему феритових циліндрів.

У роботі [7] досліджено поширення електромагнітних хвиль у 3D-решітках магнітних циліндричних нанодротів в умовах магнітного резонансу. Дослідження підтверджують вплив взаємної орієнтації постійного та високочастотних магнітних полів на коефіцієнт поширення. Силовий вплив електромагнітних хвиль на магнітні матеріали в області магнітного резонансу не досліджено.

У роботі [8, 9] розроблено аналітичну теорію електромагнітних явищ у резонансних складних просторових системах малих резонансних однорідних ізотропних магнітодіелектричних сфер. Досліджено вплив геометрії розташування центрів сфер на електромагнітну хвилю, що розсіюється в зовнішній простір. Встановлено, що при певних розмірах між центрами сфер є мінімуми в інтенсивності потужності енергії НВЧ, що розсіюється, системою сфер. Недолік: дослідження проведено у межах квазістаціонарного наближення (електромагнітне поле вздовж перерізу сфери передбачається постійним). Кількісний показник застосування результатів дослідження ($R/\lambda_ф << 1$ де R – радіус циліндра, $\lambda_ф$ – довжина хвилі у фериті).

З аналізу відомих робіт випливає, що для збільшення силового впливу неоднорідної електромагнітної хвилі на перетворювач (феритовий циліндр, який має резонансний радіус) його потрібно замінити на систему феритових циліндрів. Створити умови, по-перше, для резонансу між рухом магнітних моментів доменів фериту і рухом вектора напруженості неоднорідного магнітного поля електромагнітної хвилі і, по-друге, для виникнення просторового і групового резонансів.

Для досягнення поставленої мети необхідно: 1. Розв'язати задачу про дифракцію плоскопаралельної електромагнітної хвилі на системі з двох феритових циліндрів, величини радіусів яких відповідають просторовому резонансу ($R \le 0{,}1\cdot\lambda_o$ де λ_o – довжина хвилі у вільному просторі). 2. Дослідити зміну напруженості магнітного поля НВЧ у середині перерізів феритових циліндрів від відстані між ними. 3. Дослідити силову дію електромагнітної енергії НВЧ на систему із двох феритових циліндрів при феромагнітному резонансі.

2. Постановка задачі про дифракцію плоско-паралельної електромагнітної хвилі на системі, що складається з двох нескінченно довгих феритових циліндрів

На даний час задачу про дифракцію плоскої електромагнітної хвилі на системі нескінченно довгих феритових циліндрів з резонансними розмірами їх радіусів не вирішено.

Фізичну модель падіння плоско паралельної електромагнітної хвилі на систему, що складається з двох нескінченно довгих феритових циліндрів, можна представити так. Є декартіва система координат (x, y, z). У цій системі розташована перша циліндрична система координат – r1,α1,z і друга – r2,α2,z. Координати центру першого циліндра – x1, y1 другого – x2, y2. Плоска поляризована хвиля поширюється вздовж осі x. Вектор магнітної напруженості H змінюється з часом уздовж осі z. Вектор електричної напруженості E змінюється з часом уздовж осі y. Плоска поляризована хвиля падає на циліндри 1 і 2 та індукує в них внутрішні поля. Індуковані поля створюють поля, що розсіюються, які взаємодіють з першим та другим циліндрами.

В основу вирішення поставленої задачі покладено інтегральне рівняння макроскопічної електродинаміки [10, 11] та вирази для складових електромагнітного поля в середині циліндра та розсіяних циліндром в навколишній простір, отримані в результаті розв'язання рівнянь Максвелла і які задовольняють граничним умовам [5, 12].

Інтегральне рівняння макроскопічної електродинаміки має такий вигляд [11]:

$$\boldsymbol{H} = \boldsymbol{H_0} + \sum_{i=1}^{2}\left(\frac{a}{k_o^2}\cdot\int_{S_i}\left(\boldsymbol{H}\cdot\operatorname{grad}\right)\cdot\operatorname{grad}\varphi\cdot\mathrm{d}s_i + \cdot a\cdot\int_{S_i}\boldsymbol{H}\cdot\varphi\cdot\mathrm{d}s_i - b\cdot\int_{S_i}\left[\operatorname{grad}\varphi,\boldsymbol{E}\right]\cdot\mathrm{d}s_i\right), \qquad (1)$$

$$b = \frac{\mathrm{i}\cdot k_o}{2\cdot\pi}\cdot\left(\frac{\varepsilon}{\varepsilon_1}-1\right),\quad a = \frac{k_o^2}{2\cdot\pi}\cdot\left(\frac{\mu}{\mu_1}-1\right),\quad \varphi(|\,\boldsymbol{r}-\boldsymbol{r_1}\,|) = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{(\boldsymbol{r}-\boldsymbol{r_1})^2}\right),$$

де $\boldsymbol{E}$, $\boldsymbol{H}$ – вектори електричної та магнітної напруженості електромагнітної хвилі в середині циліндрів; $\boldsymbol{H}_o$ – вектор магнітної напруженості електромагнітної хвилі, що падає на циліндри, $\varphi(|\boldsymbol{r}-\boldsymbol{r_1}|)$ – функція Гріна; $k_o = 2/\lambda_o$, λ_o – довжина хвилі у вільному просторі; H_0^2 – функція Ханкеля другого роду нульового порядку; $\boldsymbol{r}$ – радіус вектор, проведений в точку спостереження; $\boldsymbol{r_1}$ – радіус вектор, проведений в точку випромінювання (інтегрування); i – уявна одиниця, ε, μ – діелектрична та магнітна проникності середовища циліндрів, ε_1, μ_1 – діелектрична та магнітна проникності зовнішнього середовища. Рівняння (1) записано у системі СГС фізичних величин.

Складові електромагнітного поля (у середині циліндра j (j = 1,2), на яке падає плоско-паралельна електромагнітна хвиля $H_m \cdot e^{i\cdot(\omega\cdot t \mp k_o\cdot x)}$, одержані в результаті розв'язання рівнянь Максвелла, задовольняючі граничним умовам, можна подати в наступному вигляді [5, 12]:

$$H\boldsymbol{в}_{jz}^{\mp} = H_m \cdot h\boldsymbol{в}_{jz}^{\mp}, \quad h\boldsymbol{в}_{jz}^{\mp} = e^{i\left(\omega\cdot t \mp k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} b_n \cdot J_n\left(k\cdot r_j\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)} \qquad (0 < r_j \le R) \tag{2}$$

$$E\boldsymbol{в}_{jr}^{\mp} = H_m \cdot e\boldsymbol{в}_{jr}^{\mp}, \quad e\boldsymbol{в}_{jr}^{\mp} = \frac{\pm 1}{k_o\cdot\varepsilon} \cdot e^{i\cdot\left(\omega\cdot t \mp i\cdot k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} b_n \cdot \frac{n}{r_j} \cdot J_n\left(k\cdot r_j\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)} \tag{3}$$

$$E\boldsymbol{в}_{j\alpha}^{\mp} = H_m \cdot e\boldsymbol{в}_{j\alpha}^{\mp}, \quad E\boldsymbol{в}_{jr}^{\mp} = \frac{i\cdot k}{k_o\cdot\varepsilon} \cdot e^{i\cdot\left(\omega\cdot t \mp i\cdot k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} b_n \left(\frac{n}{k\cdot r_j} \cdot J_n\left(k\cdot r_j\right) - J_{n+1}\left(k\cdot r_j\right)\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)}, \tag{4}$$

де H_m − амплітуда напруженості магнітного поля падаючої плоско-паралельної електромагнітної хвилі на циліндри; $h\boldsymbol{в}_{jz}^{\mp}, e\boldsymbol{в}_{jr}^{\mp}, e\boldsymbol{в}_{jr}^{\mp}$ − коефіцієнти посилення складових електромагнітного поля в середині внутрішнього простору першого циліндрів; $\mp$ − верхні знаки відповідають плоскій електромагнітній хвилі, що розповсюджується вздовж осі x, нижні знаки відповідають електромагнітної хвилі, що поширюється у протилежному напрямку осі x; r_j, α_j – координати точки спостереження складових електромагнітного поля в середині j-циліндра; $J_n(k\cdot r_j)$ – функція Бесселя.

Складові електромагнітного поля, розсіяного в навколишній простір циліндрами, можна представити в наступному вигляді:

$$Hp_{jz}^{\mp} = H_m \cdot hp_{jz}^{\mp}, \quad hp_{jz}^{\mp} = e^{i\cdot\left(\omega\cdot t \mp i\cdot k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} a_n \cdot H_n^2\left(k_o\cdot r_j\right)\cdot e^{i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)} \qquad (R \ge r_j < \infty) \tag{5}$$

$$Ep_{jr}^{\mp} = H_m \cdot ep_{jr}^{\mp}, \quad ep_{jr}^{\mp} = \frac{\pm 1}{k_o\cdot\varepsilon_1} \cdot e^{i\cdot\left(\omega\cdot t \mp i\cdot k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} a_n \cdot \frac{n}{r_j} \cdot H_n^2\left(k_o\cdot r_j\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)} \tag{6}$$

$$Ep_{j\alpha}^{\mp} = H_m \cdot ep_{j\alpha}^{\mp}, \quad ep_{j\alpha}^{\mp} = \frac{i}{\varepsilon_1} \cdot e^{i\cdot\left(\omega\cdot t \mp i\cdot k_o\cdot x_j\right)} \cdot \sum_{n=\infty}^{\infty} a_n \left(\frac{n}{k_o\cdot r_j} \cdot H_n^2\left(k_o\cdot r_j\right) - H_{n+1}^2\left(k_o\cdot r_j\right)\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_j - \frac{\pi}{2}\right)} \tag{7}$$

де $hp_{jz}^{\mp}, ep_{jr}^{\mp}, ep_{j\alpha}^{\mp}$ – коефіцієнти посилення складових електромагнітного поля в середині зовнішнього простору, r_j, α_j – координати точки спостереження складових електромагнітного поля в середині зовнішнього простору, b_n, a_n – постійні коефіцієнти [5, 12].

Рівняння (1) – (7) записано у системі СГС фізичних величин.

2.1. Дослідження явища дифракції плоско-паралельної електромагнітної хвилі на відокремленому циліндрі шляхом вирішення інтегрального рівняння макроскопічної електродинаміки

Розглянемо випадок падіння плоско-паралельної електромагнітної хвилі $H_m \cdot e^{i\cdot(\omega\cdot t - k_o\cdot x)}$ на відокремлений циліндр (другий циліндр відсутній). Рівняння (1) для H_z складової вектора $\boldsymbol{H}$

усередині перерізу циліндра в циліндричній системі координат, після виконання математичних перетворення $(\boldsymbol{H}\cdot\mathrm{grad})\cdot\mathrm{grad}\,\varphi$, $[\mathrm{grad}\,\varphi,\boldsymbol{E}]$ має такий вигляд:

$$Hв_{1z}^{-} = H_{1oz}^{-} + \int_0^{2\cdot\pi}\int_0^{r_i-\delta_i}\left[a\cdot Hв_{1z}^{-}\cdot\varphi_1 - b\cdot Eв_{1\alpha}^{-}\cdot\frac{d\varphi_1}{dr_1} + b\cdot Eв_{1r}^{-}\cdot\frac{d\varphi_1}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1 +$$

$$+\int_0^{2\cdot\pi}\int_{r_i+\delta_i}^{R}\left[a\cdot Hв_{1z}^{-}\cdot\varphi_1 - b\cdot Eв_{1\alpha}^{-}\cdot\frac{d\varphi_1}{dr_1} + b\cdot Eв_{1r}^{-}\cdot\frac{d\varphi_1}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1 \qquad (8)$$

$$H_{1oz}^{-} = H_m\cdot e^{i\cdot(\omega\cdot t - k_o x_1)}$$

$$\varphi_1 = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{(r\cdot\cos\alpha - r_1\cdot\cos\alpha_1)^2 + (r\cdot\sin\alpha - r_1\cdot\sin\alpha_1)^2}\right) \quad (\mathrm{r}\le\mathrm{R})$$

де H_{1oz} – магнітна напруженість плоско-паралельної електромагнітної хвилі, що падає на перший циліндр; r, α – координати точки спостереження; r_1, α_1 – координати точки інтегрування; $\delta_i = \mathrm{r_i}/10^{20}$ – нескінченно мала величина, яка обмежує область інтеграції, де функція $\varphi(|\boldsymbol{r}-\boldsymbol{r_1}|)$ досягає нескінченності; $Eв_{1\alpha}$ та $Eв_{1r}$ – складові вектора електричної напруженості електромагнітного поля в середині циліндрів.

Напруженість магнітного поля НВЧ, розсіяного першим циліндром в навколишній простір, може бути обчислена за допомогою виразу

$$Hp_{1z}^{-} = \int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot Hв_{1z}^{-}\cdot\varphi_1 - b\cdot Eв_{1\alpha}^{-}\cdot\frac{d\varphi_1}{dr_1} + b\cdot Eв_{1r}^{-}\cdot\frac{d\varphi_1}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1, \qquad (9)$$

$$\varphi_1 = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{(r\cdot\cos\alpha - r_1\cdot\cos\alpha_1)^2 + (r\cdot\sin\alpha - r_1\cdot\sin\alpha_1)^2}\right) \quad (\mathrm{r}\ge\mathrm{R})$$

При падінні плоско-паралельної електромагнітної хвилі тільки на відокремлений циліндр електромагнітну хвилю в середині циліндра можна уявити рівнянням (2) або рівнянням (8), якщо в його праву частину підставити вирази (2) – (4).

Слід звернути увагу, що в рівняннях (8) і (9) праві та ліві їх частини містять загальний співмножник H_m, який можна скоротити. Отже, в рівняннях (8) і (9) можна вирази $Hв_{1z}^{-}$, H_{1oz}^{-}, Hp_{1z}^{-}, $Eв_{1r}^{-}$, $Eв_{1\alpha}^{-}$ замінити відповідно на вирази $hв_{1z}^{-}$, $e^{i\cdot(\omega\cdot t-k_o x_1)}$, hp_{1z}^{-}, $eв_{1r}^{-}$, $eв_{1\alpha}^{-}$.

На рис. 1 представлено залежність напруженості магнітного поля електромагнітної хвилі від координати r. $Hв_i$ – залежність z складової напруженості магнітного поля НВЧ усередині феритового циліндра від координати r ($0 < r \le \mathrm{R}$, $\alpha = \pi/2$; $r_i = r/\mathrm{R}$), обчисленої за допомогою виразу (2); $\mathrm{H1в_i}$ – залежність z складової напруженості магнітного поля НВЧ усередині феритового циліндра від координати r ($0 < r \le \mathrm{R}$, $\alpha = \pi/2$; $r_i = r/\mathrm{R}$) обчислена за допомогою виразу (8); $\mathrm{Hp_i}$ – залежність z складової напруженості магнітного поля НВЧ поза феритовим циліндром від координати r ($\lambda_o > r \ge \mathrm{R}$, $\alpha = 0$; $r_i = r/\lambda_o$), обчисленої за допомогою виразу (5); $\mathrm{H1p_i}$ – залежність z складової напруженості магнітного поля НВЧ всередині феритового циліндра від координати r ($\lambda_o > r \ge \mathrm{R}$, $\alpha = 0$; $r_i = r/\lambda_o$) обчислена за допомогою виразу (9).

Параметри розрахунку: P − потужність падаючої електромагнітної хвиля дорівнює 200 кВт, H_m − амплітудне значення падаючої електромагнітної хвилі, яка має ліву кругову поляризацію, дорівнює 28,7 А/м, λ_o = 3,2 см, ε = (12,5 – i·0,0031), μ = (143,5 – i·1,4), Ферит марки 3СЧ17, напруженість магнітного поля при насиченні − 12 Є, насичення намагніченості дорівнює 1600/4π Гс, ширина кривої феромагнітного резонансу − 570 Є [13]. Дійсна частина магнітної проникності обчислена за методикою, представленою у роботі [5]. Уявна частина обчислена методом розв'язання рівняння, руху вектора намагніченості в постійному магнітному полі та полі НВЧ [14]. Ці параметри використовуються для всіх обчислень у цій роботі.

Резонансний радіус R1 циліндра дорівнює 0,4183 мм.

Крива $hв_i$, обчислена за допомогою рівняння (2), та крива $h1в_i$, обчислена за допомогою інтегрального рівняння (8), збігаються з похибкою, що дорівнює 0,02 %. Збігаються і розсіяні феритовим циліндром магнітні поля hp_i та $h1p_i$ з похибкою, що дорівнює $2 \cdot 10^{-9}$ %.

Слід звернути увагу, що напруженість магнітного поля hp_i розсіяна першим циліндром на відстані $\lambda_o/2$ = 16 мм у 110 разів менша за амплітудою, ніж амплітуда падаючої на циліндр електромагнітної хвилі. На зазначеній відстані розсіяна хвиля в об'ємі циліндра, який має радіус R рівний 0,4183 мм, практично не змінюється.

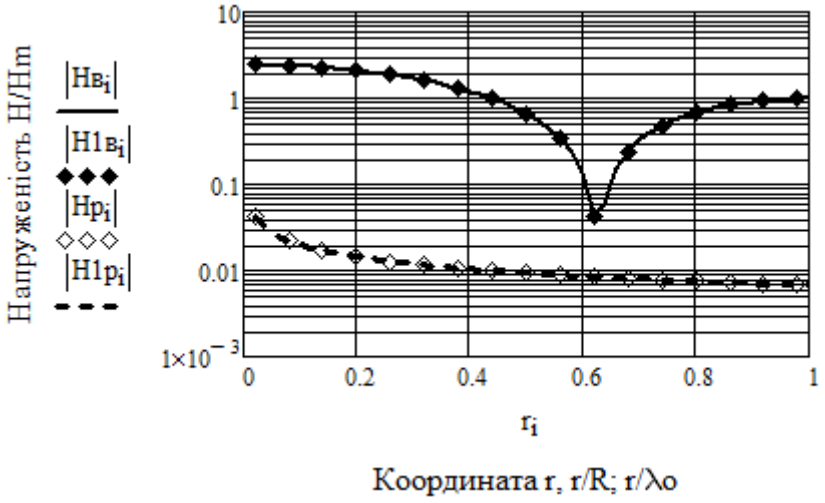


Рис. 1. Залежності напруженості магнітного поля НВЧ від координати r

Збіг кривих $hв_i$, $h1в_i$ і hp_i та $h1p_i$ свідчить, що вирази (2) – (7), отримані на основі рішення системи рівнянь Максвелла і які задовольняють граничним умовам на межі циліндра, можна використовувати при вирішенні інтегрального рівняння (1).

2.2. Дослідження явища дифракції плоско-паралельної електромагнітної хвилі на системі, що складається з двох феритових циліндрів шляхом вирішення інтегрального рівняння макроскопічної електродинаміки

Після виконання математичних перетворень $(\boldsymbol{H} \cdot \mathrm{grad}) \cdot \mathrm{grad}\,\varphi$, $[\mathrm{grad}\,\varphi, \boldsymbol{E}]$ та заміни виразів $Hв_{1z}^{\mp}, H_{1oz}^{\mp}, Hp_{1z}^{\mp}, Eв_{1r}^{\mp}, Eв_{1\alpha}^{\mp}$ на відповідні вирази $hв_{1z}^{\mp}$, $e^{i\cdot(\omega\cdot t \mp i\cdot k_o\cdot x_1)}$, $hp_{1z}^{\mp}$, $eв_{1r}^{\mp}, eв_{1\alpha}^{\mp}$ інтегральне рівняння (1) для обчислення магнітного поля НВЧ у центрі першого циліндра (r_{01} = 0) буде мати наступний вигляд:

$$hв_{1z2}^{\mp} = e^{i\cdot\cdot(\omega\cdot t \mp i\cdot k_o\cdot x_1)} + \int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot hв_{1z2}^{\mp}\cdot\varphi_{11} - b\cdot eв_{1\alpha 2}^{\mp}\cdot\frac{d\varphi_{11}}{dr_1} + b\cdot eв_{1r2}^{\mp}\cdot\frac{d\varphi_{11}}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1, +$$

$$+ \int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot hв_{2z2}^{\mp}\cdot\varphi_{12} - b\cdot eв_{2\alpha 2}^{\mp}\cdot\frac{d\varphi_{12}}{dr_2} + b\cdot eв_{2r2}^{\mp}\cdot\frac{d\varphi_{12}}{r_2\cdot d\alpha_2}\right]\cdot r_2 dr_2 d\alpha_2, \qquad (10)$$

$$\varphi_{11} = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{\left[\left((x_1 - x_1) + r_{01}\cdot\cos\alpha_{01} - r_1\cdot\cos\alpha_1\right)\right]^2 + \left[(y_1 - y_1) + \left(r_{01}\cdot\sin\alpha_{01} - r_1\cdot\sin\alpha_1\right)\right]^2}\right)$$

$$\varphi_{12} = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{\left[\left((x_1 - x_2) + r_{01}\cdot\cos\alpha_{01} - r_2\cdot\cos\alpha_2\right)\right]^2 + \left[(y_1 - y_2) + \left(r_{01}\cdot\sin\alpha_{01} - r_2\cdot\sin\alpha_2\right)\right]^2}\right),$$

де $hв_{1z2}^{\mp}, eв_{1r2}^{\mp}, eв_{1\alpha 2}^{\mp}, hв_{2z2}^{\mp}, eв_{2r2}^{\mp}, eв_{2\alpha 2}^{\mp}$ – коефіцієнти посилення складових електромагнітної хвилі всередині першого і другого циліндрів; x_1, y_1, x_2, y_2 – координати розміщення центру першого і другого циліндрів; r_{01}, α_{01} – координати точки спостереження в площі перерізу

першого циліндра; r_1, α_1 – координати точки інтегрування у першому циліндрі; r_2, α_2 – координати точки інтегрування у другому циліндрі.

Інтегральне рівняння (1) для обчислення магнітного поля НВЧ у центрі другого циліндра ($r_{02} = 0$) мають такий вигляд:

$$hв_{2z2}^{\mp} = e^{i\cdot(\omega\cdot t\mp k_o x_2)} + \int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot hв_{2z2}^{\mp}\cdot\varphi_{22} - b\cdot eв_{2\alpha 2}^{\mp}\cdot\frac{d\varphi_{22}}{dr_2} + b\cdot eв_{2r2}^{\mp}\cdot\frac{d\varphi_{22}}{r_2\cdot d\alpha_2}\right]\cdot r_2 dr_2 d\alpha_2 ,+$$

$$+\int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot hв_{1z2}^{-}\cdot\varphi_{21} - b\cdot eв_{1\alpha 2}^{-}\cdot\frac{d\varphi_{21}}{dr_1} + b\cdot eв_{1r2}^{-}\cdot\frac{d\varphi_{21}}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1 , \qquad (11)$$

$$\varphi_{22} = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{\left[\left((x_2 - x_2) + r_{02}\cdot\cos\alpha_{02} - r_2\cdot\cos\alpha_2\right)\right]^2 + \left[\left(y_2 - y_2\right) + \left(r_{02}\cdot\sin\alpha_{02} - r_2\cdot\sin\alpha_2\right)\right]^2}\right),$$

$$\varphi_{21} = \frac{\pi}{\mathrm{i}\cdot 2}\cdot \mathrm{H}_0^2\left(k_o\cdot\sqrt{\left[\left((x_2 - x_1) + r_{02}\cdot\cos\alpha_{02} - r_1\cdot\cos\alpha_1\right)\right]^2 + \left[\left(y_2 - y_1\right) + \left(r_{02}\cdot\sin\alpha_{02} - r_1\cdot\sin\alpha_1\right)\right]^2}\right),$$

де r_{02}, α_{02} – координати точки спостереження у площі перерізу другого циліндра.

Сума перших двох доданків у правій частині рівнянь (10) представляє електромагнітне поле всередині відокремленого першого циліндра $hв_{1z}^{\mp}(r_1,\alpha_1)$, представлено виразом (2). Третій доданок представляє розсіяну електромагнітну хвилю другим відокремленим циліндром у середині першого циліндра $hp_{2z}^{\mp}(r_{21},\alpha_{21})$ представлено виразом

$$hp_{21z}^{\mp} = e^{i\cdot(\omega\cdot t\mp i\cdot k_o\cdot x_2)}\cdot\sum_{n=\infty}^{\infty} a_n\cdot H_n^2\left(k_o\cdot r_{21}\right)\cdot e^{\pm i\cdot n\cdot\left(\alpha_{21}-\frac{\pi}{2}\right)}, \qquad (12)$$

$$r_{21} = \sqrt{\left[\left(x_1 + r_1\cdot\cos\alpha_1\right) - x_2\right]^2 + \left[\left(y_1 + r_1\cdot\sin\alpha_1\right) - y_2\right]^2}\ ,\ \alpha_{21} = -\arccos\left(\frac{\left(x_1 + r_1\cdot\cos\alpha_1\right) - x_2}{r_2}\right).$$

Представимо рівняння (10) у такому вигляді:

$$hв_{1z2}^{\mp} = e^{i\cdot(\omega\cdot t\mp k_o x_1)} + hp_{2z}^{\mp} + \int_0^{2\cdot\pi}\int_0^{R}\left[a\cdot hв_{1z2}^{\mp}\cdot\varphi_{11} - b\cdot eв_{1\alpha 2}^{\mp}\cdot\frac{d\varphi_{11}}{dr_1} + b\cdot eв_{1r2}^{\mp}\cdot\frac{d\varphi_{11}}{r_1\cdot d\alpha_1}\right]\cdot r_1 dr_1 d\alpha_1 . \qquad (13)$$

Якщо в рівнянні (13) припустити, що $hp_{2z}^{\mp} = 0$ то рішення його буде $hв_{1z}^{\mp}$ (дивись рівняння (8)). Якщо в рівнянні (13) припустити, що $e^{i\cdot(\omega\cdot t\mp k_o x_1)} = 0$, то рішення його представимо в наступному вигляді $h_{21вz}^{\mp}$. Де $h_{21вz}^{\mp}$ z – складова електромагнітного поля в середині першого циліндра, коли на нього падає електромагнітна хвиля $hp_{2z}^{\mp}$ розсіяна другим циліндром. Вираз $h_{21вz}^{\mp}$ потрібно обчислити.

Відповідно до принципу суперпозиції рішення інтегрального рівняння (10) можна подати у вигляді

$$Hв_{1z2}^{\mp} = H_m\cdot hв_{1z2}^{\mp} = H1^{\mp}\cdot hв_{1z}^{\mp} + H2^{\mp}\cdot h_{21вz}^{\mp}, \qquad (14)$$

де $H1^{\mp}$ та $H2^{\mp}$ – невідомі постійні коефіцієнти, що підлягають знаходженню.

2.3. Обчислення z складової електромагнітного поля в середині першого циліндра, коли на нього падає електромагнітна хвиля, розсіяна другим циліндром

Знайдемо невідомий вираз $h_{21вz}^{\mp}$. За умови R << d (d – відстань між центрами циліндрами) розсіяну другим циліндром електромагнітну хвилю $hp_{21z}^{\mp}$ ($r_1 \le R$) (12) можна представити

(в циліндричній системі координат першого циліндра) плоско паралельною електромагнітною хвилею, яка поширюється під кутом α_{21} до напрямку осі x [5, 12, 15]:

$$hn_{21z}^{\mp} = hp_{21z.}^{\mp}(r_1 = 0) \cdot e^{\pm i \cdot k_o \cdot r_1 \cdot (\alpha_{2.1} - \alpha_1)}. \quad (15)$$

Середні величини $\left|hn_{21z}^{\mp}\right|$ і $\left|hp_{21z}^{\mp}\right|$ по площі поперечного перерізу першого циліндра відрізняються на 1,3, 0,48, 0,15, 0,02 % при відстанях *d,* рівних $2 \cdot R$, $\lambda_o/2$, λ_o, $3 \cdot \lambda_o$ відповідно. Параметри обчислення: радіус феритових циліндрів R_{11} дорівнює 3,863 мм ($x_1 = 0$, $y_1 = 0$, $x_2 = d$, $y_2 = d$).

Плоско-паралельна хвиля (15) падає на перший циліндр і індукує в ньому внутрішнє електромагнітне поле [5, 12, 15]:

$$h_{21вz}^{\mp} = hp_{21z}^{\mp}(r_1 = 0) \cdot \sum_{n=\infty}^{\infty} b_n \cdot J_n(k \cdot r_1) \cdot e^{\pm i \cdot n \cdot \left[(\alpha_1 - \alpha_{21}) - \frac{\pi}{2}\right]}, \quad (16)$$

$$e_{21вr}^{\mp} = \pm \frac{hp_{21z}^{\mp}(r_1 = 0)}{k_o \cdot \varepsilon} \cdot \sum_{n=\infty}^{\infty} b_n \cdot \frac{n}{r_1} \cdot J_n(k \cdot r_1) \cdot e^{\pm i \cdot n \cdot \left[(\alpha_1 - \alpha_{21}) - \frac{\pi}{2}\right]}, \quad (17)$$

$$e_{21вα}^{\mp} = \frac{i \cdot k \cdot}{k_o \cdot \varepsilon} hp_{21z}^{\mp}(r_1 = 0) \cdot \sum_{n=\infty}^{\infty} b_n \left(\frac{n}{k \cdot r_1} \cdot J_n(k \cdot r_1) - J_{n+1}(k \cdot r_1) \right) \cdot e^{i \cdot n \cdot \left[(\alpha_1 - \alpha_{21}) - \frac{\pi}{2}\right]}, \quad (18)$$

Рівняння (16) – (18) задовольняють граничним умовам та рівнянням Максвелла.

Складові електромагнітного поля, розсіяного першим циліндром у точках поперечного перерізу другого циліндра, мають такий вигляд:

$$h_{12вz}^{\mp} = hp_{12z}^{\mp} \cdot \sum_{n=\infty}^{\infty} b_n \cdot J_n(k \cdot r_2) \cdot e^{\pm i \cdot n \cdot \left[(\alpha_2 - \alpha_{12}) - \frac{\pi}{2}\right]}, \quad (19)$$

$$e_{12вr}^{\mp} = \pm \frac{hp_{12z}^{\mp}}{k_o \cdot \varepsilon} \cdot \sum_{n=\infty}^{\infty} b_n \cdot \frac{n}{r_2} \cdot J_n(k \cdot r_2) \cdot e^{\pm i \cdot n \cdot \left[(\alpha_2 - \alpha_{12}) - \frac{\pi}{2}\right]}, \quad (20)$$

$$e_{12вα}^{\mp} = \frac{i \cdot k}{k_o \cdot \varepsilon} \cdot hp_{12z}^{\mp} \cdot \sum_{n=\infty}^{\infty} b_n \left(\frac{n}{k \cdot r_2} \cdot J_n(k \cdot r_2) - J_{n+1}(k \cdot r_2) \right) \cdot e^{\pm i \cdot n \cdot \left[(\alpha_2 - \alpha_{12}) - \frac{\pi}{2}\right]}, \quad (21)$$

$$hp_{12z}^{\mp} = e^{i(\omega \cdot t \mp k_o \cdot x_1)} \cdot \sum_{n=\infty}^{\infty} a_n \cdot H_n^2(k_o \cdot r_{12}) \cdot e^{i \cdot n \cdot \left(\alpha_{12} - \frac{\pi}{2}\right)}$$

$$r_{12} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}, \quad \alpha_{12} = \arccos\left(\frac{x_2 - x_1}{r_{12}}\right)$$

де $hp_{12z}^{\mp}$ – коефіцієнт посилення z складової розсіяної електромагнітної хвилі першим відокремленим циліндром в середині другого циліндра.

З рівнянь (10), (13), (14) слідує, що поле всередині першого циліндра можна представити у вигляді суми поля всередині відокремленого першого циліндра, на який падає плоско-паралельна електромагнітна хвиля $H_m \cdot e^{i \cdot (\omega \cdot t - k \cdot x)}$, та поля всередині першого циліндра, на який падає розсіяна другим циліндром хвиля, представлена виразами (16), (18).

Аналогічно обчисленню виразу (14) з інтегрального рівняння (1) і виразів (2) – (7) отримаємо інші складові електромагнітного поля всередині першого та другого циліндрів:

$$Hв_{1z2}^{\mp} = H1^{\mp} \cdot hв_{1z}^{\mp} + H2^{\mp} \cdot h_{21вz}^{\mp}, \quad (22)$$

$$Eв^{\mp}_{1r2} = H1^{\mp} \cdot eв^{\mp}_{1r} + H2^{\mp} \cdot e^{\mp}_{21вr}, \quad (23)$$

$$Eв^{\mp}_{1\alpha 2} = H1^{\mp} \cdot eв^{\mp}_{1\alpha} + H2^{\mp} \cdot e^{\mp}_{21в\alpha}, \quad (24)$$

$$Hв^{\mp}_{2z2} = H1^{\mp} \cdot h^{\mp}_{12вz} + H2^{\mp} \cdot hв^{\mp}_{2z}, \quad (25)$$

$$Eв^{\mp}_{2r2} = H1^{\mp} \cdot eв^{\mp}_{12вr} + H2^{\mp} \cdot eв^{\mp}_{2r}, \quad (26)$$

$$Eв^{\mp}_{2\alpha 2} = H1^{\mp} \cdot eв^{\mp}_{12в\alpha} + H2^{\mp} \cdot eв^{\mp}_{2r\alpha}. \quad (27)$$

У системі рівнянь (22) – (27) дві невідомі величини $H1^{\mp}$ та $H2^{\mp}$. Для знаходження невідомих величин підставимо праві частини виразів (22) – (24) в інтегральне рівняння (10), а праві частини виразів (25) – (27) в інтегральне рівняння (11). Отримаємо таку систему інтегральних рівнянь:

$$b^{\mp}_{1} = H1^{\mp} \cdot a^{\mp}_{11} + H2^{\mp} \cdot a^{\mp}_{12}, \quad (28)$$

$$b^{\mp}_{2\,2} = H1^{\mp} \cdot a^{\mp}_{21} + H2^{\mp} \cdot a^{\mp}_{22}, \quad (29)$$

$$b^{\mp}_{1} = -H_m \cdot e^{\mp i \cdot k_o \cdot x_1}, \qquad b^{\mp}_{2} = -H_m \cdot e^{\mp i \cdot k_o \cdot x_2}, \quad (30)$$

$$a^{\mp}_{11} = j^{\mp}_{1,1} + j^{\mp}_{1,3} + j^{\mp}_{1,5} + j^{\mp}_{1,7} + j^{\mp}_{1,9} + j^{\mp}_{1,11} - hв^{\mp}_{1z}, \quad (31)$$

$$a^{\mp}_{12} = j^{\mp}_{1,2} + j^{\mp}_{1,4} + j^{\mp}_{1,6} + j^{\mp}_{1,8} + j^{\mp}_{1,10} + j^{\mp}_{1,12} - hв^{\mp}_{21вz}, \quad (32)$$

$$a^{\mp}_{21} = j^{\mp}_{2,1} + j^{\mp}_{2,3} + j^{\mp}_{2,5} + j^{\mp}_{2,7} + j^{\mp}_{2,9} + j^{\mp}_{2,11} - hв^{\mp}_{12вz}, \quad (33)$$

$$a^{\mp}_{22} = j^{\mp}_{2,2} + j^{\mp}_{2,4} + j^{\mp}_{2,6} + j^{\mp}_{2,8} + j^{\mp}_{2,10} + j^{\mp}_{2,12} - hв^{\mp}_{2z}, \quad (34)$$

Вирази для невідомих інтегралів $j^{\mp}_{1,1}, j^{\mp}_{1,2}, j^{\mp}_{1,3} \ldots j^{\mp}_{2,12}$ наведено у таблиці.

Розв'язання системи рівнянь:

$$H1^{\mp} = \frac{\Delta 1^{\mp}}{\Delta^{\mp}}, \ldots \quad H2^{\mp} = \frac{\Delta 2^{\mp}}{\Delta^{\mp}}, \ldots\ldots\ldots\ldots \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (35)$$

де

$$\Delta^{\mp} = \begin{vmatrix} a^{\mp}_{11} & a^{\mp}_{12} \\ a^{\mp}_{21} & a^{\mp}_{22} \end{vmatrix} \quad \Delta 1^{\mp} = \begin{vmatrix} b^{\mp}_{1} & a^{\mp}_{12} \\ b^{\mp}_{2} & a^{\mp}_{22} \end{vmatrix} \quad \Delta 2^{\mp} = \begin{vmatrix} a^{\mp}_{11} & b^{\mp}_{1} \\ a^{\mp}_{21} & b^{\mp}_{2} \end{vmatrix}.$$

Підставивши знайдені величини $H1^{\mp}$ і $H2^{\mp}$ у рівняння (22) – (27), отримаємо систему рівнянь для обчислення електричних і магнітних складових електромагнітного поля всередині першого і другого циліндрів.

На рис. 2 наведено залежності напруженості магнітного поля НВЧ від циліндричної координати *r* в середині першого та другого циліндрів, де $H1_i$ – крива залежності магнітного поля НВЧ у середині першого циліндра, обчислена за допомогою рівняння (22); $h1_i$ – крива залежності магнітного поля НВЧ у середині першого циліндра, обчислена з допомогою інтегрального рівняння (10); $H2_i$ – крива залежності магнітного поля НВЧ у середині другого циліндра, обчислена з допомогою рівняння (25); $h2_i$ – крива залежності магнітного поля НВЧ у середині другого циліндра, обчислена з допомогою інтегрального рівняння (11). Криві, представлені на рис. 2, побудовані на підставі чисельного аналізу рівнянь (22), (25) та (10), (11).

Таблица

$$j_{1,1}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{1z}^{\mp}\cdot\varphi_{11}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,1}^{\mp} = \int_0^{2\cdot\pi}\int_0^{R} h_{12\text{в}z}^{\mp}\cdot\varphi_{22}\cdot r_2 dr_2 d\alpha_2$$

$$j_{1,2}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{21\text{в}z}^{\mp}\cdot\varphi_{11}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,2}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{2z}^{\mp}\cdot\varphi_{22}\cdot r_2 dr_2 d\alpha_2$$

$$j_{1,3}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{1\alpha}^{\mp}\cdot\frac{d\varphi_{11}}{dr_1}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,3}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{12\text{в}\alpha}^{\mp}\cdot\frac{d\varphi_{22}}{dr_2}\cdot r_2 dr_2 d\alpha_2$$

$$j_{1,4}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e_{21\text{в}\alpha}^{\mp}\cdot\frac{d\varphi_{11}}{dr_1}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,4}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{2\alpha}^{\mp}\cdot\frac{d\varphi_{22}}{dr_2}\cdot r_2 dr_2 d\alpha_2$$

$$j_{1,5}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{1r}^{\mp}\cdot\frac{d\varphi_{11}}{r_1\cdot d\alpha_1}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,5}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e_{12\text{в}r}^{\mp}\cdot\frac{d\varphi_{22}}{r_2\cdot d\alpha_2}\cdot r_2 dr_2 d\alpha_2$$

$$j_{1,6}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e_{21\text{в}r}^{\mp}\cdot\frac{d\varphi_{11}}{r_1\cdot d\alpha_1}\cdot r_1 dr_1 d\alpha_1 \qquad j_{2,6} = b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{2r}^{\mp}\cdot\frac{d\varphi_{22}}{r_2\cdot dr_2}\cdot r_2 dr_2 d\alpha_2\ \mp$$

$$j_{1,7}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{12\text{в}z}^{\mp}\cdot\varphi_{12}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,7}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{1z}^{\mp}\cdot\varphi_{21}\cdot r_1 dr_1 d\alpha_1$$

$$j_{1,8}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{2z}^{\mp}\cdot\varphi_{12}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,8}^{\mp} = a\int_0^{2\cdot\pi}\int_0^{R} h\text{в}_{21\text{в}z}^{\mp}\cdot\varphi_{21}\cdot r_1 dr_1 d\alpha_1$$

$$j_{1,9}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{12\text{в}\alpha}^{\mp}\cdot\frac{d\varphi_{12}}{dr_2}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,9}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{1\alpha}^{\mp}\cdot\frac{d\varphi_{21}}{dr_1}\cdot r_1 dr_1 d\alpha_1$$

$$j_{1,10}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{2\alpha}\cdot\frac{d\varphi_{12}}{dr_2}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,10}^{\mp} = -b\int_0^{2\cdot\pi}\int_0^{R} e_{21\text{в}\alpha}^{\mp}\cdot\frac{d\varphi_{21}}{dr_1}\cdot r_1 dr_1 d\alpha_1$$

$$j_{1,11}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e_{12\text{в}r}^{\mp}\cdot\frac{d\varphi_{12}}{r_2\cdot d\alpha_2}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,11}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{1r}^{\mp}\cdot\frac{d\varphi_{21}}{r_1\cdot d\alpha_1}\cdot r_1 dr_1 d\alpha_1$$

$$j_{1,12}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{2r}^{\mp}\cdot\frac{d\varphi_{12}}{r_2\cdot d\alpha_2}\cdot r_2 dr_2 d\alpha_2 \qquad j_{2,12}^{\mp} = b\int_0^{2\cdot\pi}\int_0^{R} e\text{в}_{21\text{в}r}^{\mp}\cdot\frac{d\varphi_{21}}{r_1\cdot dr_1}\cdot r_1 dr_1 d\alpha_1$$

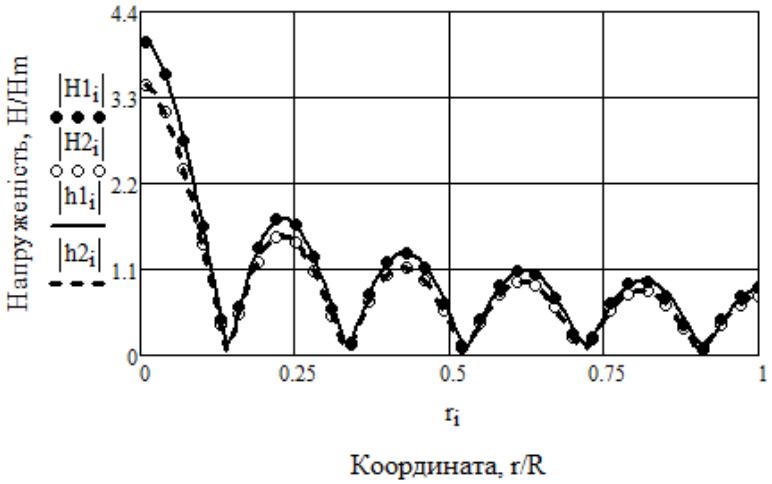


Рис. 2. Залежність напруженості магнітного поля НВЧ від координати r

Параметри розрахунку: резонансний радіус R_5 дорівнює 1,7997 мм, $\alpha_1 = \alpha_2 = \pi/2$.

Хід кривих, наведених на рис. 2, показує, що криві, які характеризують магнітне поле НВЧ у площі перерізу першого циліндра $H1_i$ та $h1_i$, збігаються. Збігаються також і криві $H2_i$ та $h2_i$. Це свідчить, що розроблений метод обчислення електромагнітного поля НВЧ для системи, що складається з плоско-паралельної електромагнітної хвилі та двох феритових циліндрів, задовольняє: рівнянням Максвелла, граничним умовам для першого та другого циліндрів та інтегральним рівнянням макроскопічної електродинаміки.

На рис. 3 представлено залежності відношення модуля z складової напруженості магнітного поля НВЧ у центрі циліндрів до напруженості падаючої плоскої хвилі Hm від відстані d_i між циліндрами, де $H1_i$ і $H2_i$ криві відносяться до першого і другого циліндрів, розташованих уздовж осі *y*, на яких падає хвиля $H_m \cdot e^{i\cdot(\omega\cdot t - ko\cdot x)}$. Криві $h1_i$ b $h2_i$ відносяться до першого і другого циліндрів, що знаходяться в полі стоячої хвилі $H_m \cdot e^{i\cdot(\omega\cdot t - ko\cdot x)} + H_m \cdot e^{i\cdot(\omega\cdot t + ko\cdot x)}$. Координати розташування центрів циліндрів такі: $x_1 = \lambda_o/8$, $y_1 = -d_i/2$, $x_2 = \lambda_o/8$, $y_2 = d_i/2$. Пряма лінія відносяться до відокремленого циліндра. Резонансний радіус R_{11} всіх циліндрів, що розглядаються, дорівнює 3,8631 мм. Криві, що представлені на рис. 2, побудовані виходячи з чисельного аналізу рівнянь (2), (22), (25), (35).

Хід кривих показує, що у системі, що складається з двох циліндрів, виникає груповий резонанс. Зі збільшенням відстані між циліндрами явище групового резонансу слабшає, і напруженість магнітного поля НВЧ усередині циліндрів прагне напруженості магнітного поля в центрі відокремленого циліндра H_i.

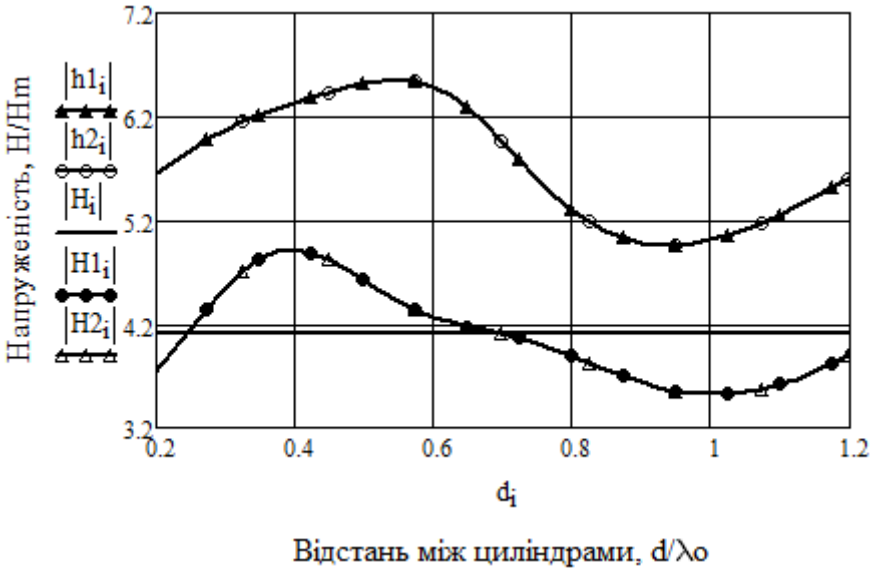


Рис. 3. Залежність напруженості магнітного поля від відстані між циліндрами

3. Дослідження силової дії електромагнітної енергії НВЧ на систему з двох феритових циліндрів

Потенційну енергію, яку одержують два феритові циліндри, розташовані в електромагнітному полі стоячої хвилі на відстані $\lambda_o/2$ один від одного, можна обчислити за допомогою виразу [4]:

$$U = -\mu_o \cdot h \cdot \int_0^{2\cdot\pi R}\int_0 \left|(\mu - 1)\cdot\left(H\text{в}_{1z2}^{-} + H\text{в}_{1z2}^{+}\right)\right| \cdot \left|\mu\cdot\left(H\text{в}_{1z2}^{-} + H\text{в}_{1z2}^{+}\right)\right| \cdot r1 dr1 d\alpha 1 \cdot - \cdot$$

$$- \mu_o \cdot h \cdot \int_0^{2\cdot\pi R}\int_0 \left|(\mu - 1)\cdot\left(H\text{в}_{2z2}^{-} + H\text{в}_{2z2}^{+}\right)\right| \cdot \left|\mu\cdot\left(H\text{в}_{2z2}^{-} + H\text{в}_{2z2}^{+}\right)\right| \cdot r2 dr2 d\alpha 2, \qquad (36)$$

де μ_o – магнітна постійна; h − довжина феритових циліндрів, h = 0,64 м; R − радіус феритового циліндра; $Hв^{-}_{1z2}$ та $Hв^{-}_{2z2}$ – напруженість магнітного поля НВЧ усередині першого та другого циліндрів, коли плоска електромагнітна хвиля поширюється вздовж осі x; H^{+}_{1z2} та H^{+}_{2z2} – напруженість магнітного поля НВЧ усередині першого та другого циліндрів, коли електромагнітна хвиля поширюється у протилежному напрямку осі x. Вираз (36) представлено в СІ.

Сила, з якою стояча електромагнітна хвиля діє на систему феритових циліндрів, може бути обчислена за допомогою рівняння [4]

$$F = gradU\,. \tag{37}$$

На рис. 4 представлено залежність сили від величини резонансного радіусу, де F_i – сила, що діє на відокремлений феритовий циліндр; $F1x_i$ – сила, що діє на перший феритовий циліндр (координати розташування центрів циліндрів такі: x_1=0, y_1=0, x_2= λ_o/2, y_2=0); $F2x_i$ – сила, діюча на другий феритовий циліндр (x_2= λ_o/2, y_2=0, x_1=0, y_1=0); Fy_i – сила, що діє на перший і другий феритовий циліндри, центри яких розташовані вздовж осі y (x_1=0, y_1= – λ_o/2, x_2= 0, y_2= λ_o/2). Довжина феритових циліндрів дорівнює 0,64 м. Криві, що представлено на рис. 5, побудовані на підставі чисельного аналізу рівнянь (37) та (38).

Хід кривих, наведених на рис. 4, показує, що сили, що діють на феритові циліндри, поміщені в стоячу електромагнітну хвилю, збільшуються зі збільшенням розміру резонансного радіуса. Зі збільшенням резонансного радіусу збільшується залежність величини сили від фаз вторинних НВЧ хвиль індукованих першим і другим циліндрами.

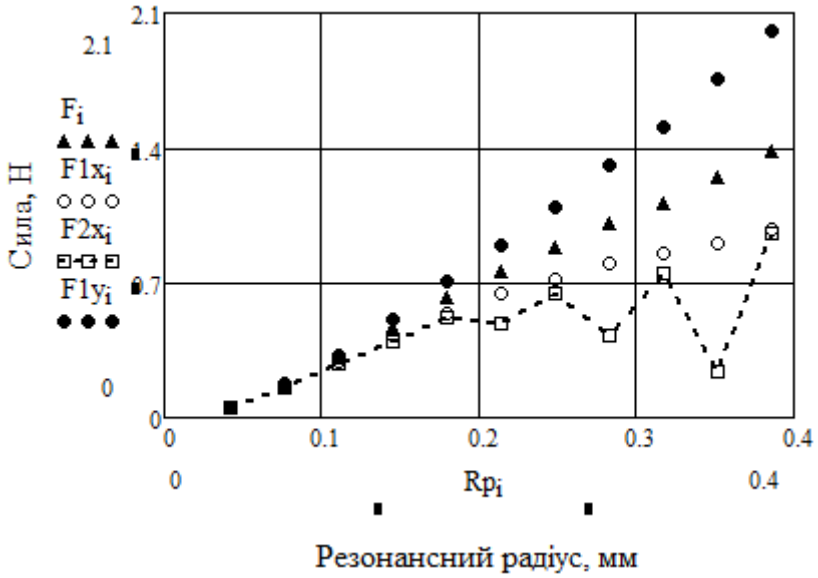


Рис. 4. Залежність сили від величини резонансного радіусу

Рівень сили, що діє на циліндри, центри яких розташовані вздовж осі y, більші, ніж сили, що діють на циліндри, центри яких розташовані вздовж осі x. Наприклад, для феритових циліндрів, резонансний радіус яких R_{11} = 3,863 мм, діють наступні сили: F_i = 1,4 Н, $F1x_i$ = 1 Н, $F2x_i$ = 0,9 Н, $F1y_i$ = 2 Н, $F2y_i$ = 2 Н.

Приведення величини сили показує, що сумарна сила, що діє на систему, що складається з двох феритових циліндрів, центри яких розташовані вздовж осі y і розташовані на відстані λ_o/2, дорівнює 4 Н. Величина цієї сили в 2,8 рази більше в порівнянні з силою, що діє на два відокремлені циліндри.

Результати дослідження явища дифракції на системі, що складається з двох циліндрів, показують, що сила, з якою неоднорідність стоячої електромагнітної хвилі діє на систему

феритових циліндрів, залежить від просторового резонансу, обумовленого взаємним розташуванням їх у просторі.

4. Обговорення результатів досліджень

Замінивши в рівняннях (37), (38) μ-1 і μ на Re(μ-1) та Re(μ), отримаємо чисельну величину для сумарної сили, що діє на перший і другий феритові циліндри $R_{11} = 4{,}28$ Н. Наближене значення сили більше на 0,7 % від сили, обчисленої за допомогою (36), (37). З рівняння (36) випливає, що $(H_c)^2$ пропорційна квадрату напруженості магнітного поля падаючої електромагнітної хвилі $(\eta \cdot H_m)^2$.

З урахуванням сказаного приблизно (з похибкою 3,6 %) сила може бути обчислена за допомогою рівнянь:

$$F = -\mu_o \cdot \mathrm{Re}(\mu - 1) \cdot \mathrm{Re}(\mu) \cdot \frac{d}{dz}\left(\eta \cdot H_m \cdot \cos(2 \cdot \pi \cdot x / \lambda_o)\right)^2 \cdot V, \qquad (38)$$

$$H_c = \sqrt{\frac{1}{\pi \cdot R^2} \int_0^{2\cdot\pi} \int_0^{R} |H\text{в}_{1z}^{-} + H\text{в}_{1z}^{+}|^2 \cdot r1 dr1 d\alpha 1}. \quad \eta_1 = \frac{H_c}{H_m},$$

де H_m – модуль вектора напруженості змінного магнітного поля падає на феритовий стрижень і обертається навколо осі z; V – об`єм феритового циліндра; η – коефіцієнт пропорційності .

У роботі використовуються параметри: потужність генератора НВЧ P_2=200 кВт; площа поперечного перерізу променю, в якому поширюється падаюча електромагнітна хвиля, S_2=0,32 м2; амплітудне значення падаючої електромагнітної хвилі, яка має ліву кругову поляризацію H_m = 28,7 А/м; довжина феритового циліндра дорівнює 0,64 м; радіус феритового циліндра R_{11} = 3,863 мм; об`єм феритового циліндра $V=3{,}2\cdot10^{-7}$ м^3; μ = 143,5; F_i = 1,4 Н, $F1y_i$ = 2 Н, η_i = 0,95, $\eta 1y_i$ = 0,111; координата розміщення центру першого циліндра $x_1 = \lambda_o/8$, іншого циліндра – $x_2 = \lambda_o/8 + \lambda_o/2$.

Відповідно до виразу (38) сила, що діє на відокремлений феритовий циліндр, дорівнює 1,4 Н для двох циліндрів – 4 Н.

У роботі [2] використовуються такі параметри: потужність генератора НВЧ P_1 = 10 Вт; площа перетин хвилеводу S1 = 2,3·1 см^2; амплітудне значення електромагнітної хвилі H_m = 35,26 А/м; $V_1=2{,}34\cdot10^{-8}$ м^3 (радіус кулі R дорівнює 1,775 мм); μ = 143,5, $F= (6 \pm 0{,}5)\cdot10^{-6}$ Н; η = 0,25 [3], координата розміщення центру феритової кулі x = a/4. Відповідно до (39) силу, що діє на феритову сферу, поміщену у прямокутному хвилеводі, можна представити приблизно (з похибкою 3,6 %) таким виразом:

$$F = -\mu_o \cdot (\mu - 1) \cdot \mu \cdot \frac{d}{dz}\left(\eta \cdot H_m \cdot \cos(\pi \cdot x / a)\right)^2 \cdot V1 = 6{,}4 \cdot 10^{-6} \cdot\cdot H. \qquad (39)$$

Збіг результату обчислення сили за спрощеним виразом з величиною сили, виміряної експериментально в межах похибки вимірювання, підтверджує справедливість розробленого алгоритму для дослідження силової дії неоднорідної електромагнітної хвилі на систему, що складається з двох феритових циліндрів та перебувають у стані феромагнітного резонансу.

Висновки

Вирішено завдання про дифракцію плоскополяризованої електромагнітної хвилі на системі, що складається з двох феритових циліндрів, величина радіусів яких відповідає просторовому резонансу. Розроблений метод обчислення електромагнітного поля НВЧ для системи, що складається з плоско-паралельної електромагнітної хвилі та двох феритових циліндрів, задовольняє: рівнянням Максвелла, граничним умовам для першого та другого циліндрів та інтегральним рівнянням макроскопічної електродинаміки. Межі застосування розробленого методу $R \le 0{,}1\cdot\lambda_o$, $d \ge 0{,}1\cdot\lambda_o$. Результати дослідження явища дифракції на системі, що склада-

ється з двох феритових циліндрів, показують, що сумарна сила, з якою неоднорідність стоячої електромагнітної хвилі діє на два феритові циліндри, дорівнює 4 Н. Це в 2,8 рази більше, ніж сила, що діє на один відокремлений циліндр. Отримані результати досліджень можуть бути використані розробниками перетворювачів НВЧ енергії в механічну енергію. Подальше удосконалення методу перетворення електромагнітної енергії НВЧ в механічну енергію, мабуть, полягає у вирішенні задачі про дифракцію плоскополяризованої електромагнітної хвилі на системі, що складається з N феритових циліндрів, дослідженні втрат електромагнітної енергії в системі феритових циліндрів.

Список літератури:

1. Пондеромоторное действие электромагнитного поля (теория и приложения) / Р. А. Валитов, Н. А. Хижняк, В. С. Жилков, [и др.] ; под ред. Р. А. Валитова. Москва : Сов. радио, 1975. 232 с.

2. Мартыненко Л. Г. Влияние ферримагнитного резонанса на преобразование электромагнитной энергии в механическую / Л. Г. Мартыненко, Г. Л. Комарова, В. В. Маличенко // Известия вузов, радиоэлектроника. 2016. Т. 59. №. 10. С. 30 – 36. DOI: 10.3103/S0735272716100046.

3. Комарова Г. Л. Влияние ферримагнитного резонанса на преобразование энергии электромагнитной волны ЖИГ – резонатором в механическую энергию // Радиотехника. 2021. Вып. 207. С. 149 – 158.

4. Measurement of Impulsive Thrust from a Closed Radio-Frequency Cavity in Vacuum / H. White, P March, J. Lawrence et al. // Journal of propulsion and power. Vol. 33, No. 4. July–August 2017. P. 830 – 841. DOI: 10.2514/1.B36120.

5. Мартыненко Л. Г. Влияние ферримагнитного резонанса на преобразование энергии электромагнитной стоячей волны в механическую энергию / Л. Г. Мартыненко, Г. Л. Комарова // Известия вузов, радиоэлектроника. 2020. Т. 63. №. 5. С 290 – 298. DOI: 10.3103/S0735272720050039

6. Мартиненко Л. Г. Спосіб перетворення електромагнітної енергії в механічну / Л.Г. Мартиненко, Г.Л. Комарова, В.В. Маличенко. Патент на винахід України. № 117748. Бюл. № 18 від 25.09.2018.

7. Макеева Г. С. Электродинамический анализ постоянных распространения электромагнитных волн в 3D-решотках магнитных нанопроволок в условиях магнитного резонанса в микроволновом диапазоне / Г. С. Макеева, О. А. Голованов // Радиотехника и электроника. 2016. Т. 61. №. 1. С. 3 – 11. DOI : 10.7868/S0033849415110145

8. Kozar A. I. Resonant degenerate crvstal made of spheres located magnetodielectric medium // International Journal of Electromagnetics and Applications. 2013. Vol. 3, No. 2. Pp. 15 – 19. DOI: 10.5923/j.idea.20130302.02.

9. Козар А.И. Резонансные метакристаллы из малых магнитодиэлектрических сфер : монография. Харьков : ХНУРЭ, 2014. 352 с.

10. Хижняк Н. А. Интегральные уравнения макроскопичной электродинамики. Киев : Наук. думка. 1986, 280 с.

11. Хижняк Н. А. Функция Грина уравнений Максвелла для неоднородных сред // Журнал технической физики. 1958. Т. 28. Вып. 7. С. 1592 – 1609.

12. Никольский В.В. Теория электромагнитного поля. Москва : Высш. шк., 1961. 371 с.

13. Микроволновые ферриты. https://www.domen.ru/mikrovolnovye-ferrity 07.01.2018. Дата доступу: 13.02.20.

14. Гуревич А.Г. Ферриты на сверхвысоких частотах. Москва : Физматгиз, 1960.

15. .Стреттон Дж. А. Теория электромагнетизма. Москва : Гостехиздат, 1948. 539 с.

Надійшла до редколегії 26.02.2023

Відомості про автора:

Комарова Ганна Леонідівна – канд. техн. наук, доцент, Український державний університет залізничного транспорту, доцент кафедри інженерії вагонів та якості продукції, Україна; e-mail: anna.kom3793@gmail.com.ua; ORCID: https://orcid.org/0000-0001-8597-5891

УДК 621.357 **DOI:10.30837/rt.2023.1.212.10**

В.М. БОРЩОВ, д-р техн. наук, О.М. ЛІСТРАТЕНКО, канд. техн. наук,
М.А. ПРОЦЕНКО, канд. техн. наук, І.Т. ТИМЧУК, канд. техн. наук, О.В. КРАВЧЕНКО,
О.В. СУДДЯ, І.В. БОРЩОВ, М.І. СЛІПЧЕНКО, д-р фіз.-мат. наук.

КОМБІНОВАНІ ТЕПЛОПРОВІДНІ ПЛАТИ З ДІЕЛЕКТРИКАМИ З ПОЛІІМІДУ

Вступ

Різноманітність виконань сучасних напівпровідникових приладів, їхнє застосування для широкого спектра завдань, необхідність відведення великої кількості тепла, робота в жорстких умовах експлуатації – все це зумовлює пошук унікальних конструкторських і технологічних рішень для виготовлення електронних пристроїв. Необхідні спеціальні матеріали, що володіють високою технологічністю і виправданою собівартістю при забезпеченні необхідних експлуатаційних характеристик виробів.

Також необхідна оптимізація конструкції і технології їх складання, використання нових матеріалів з поліпшеними теплопровідними властивостями. При цьому застосування нових матеріалів з поліпшеними теплопровідними властивостями часто залишається єдиним прийнятним способом відведення тепла від поверхні напівпровідникових чипів.

Очевидно, що в міру збільшення потужності та робочої температури напівпровідникових приладів роль і використання нових матеріалів з високою теплопровідністю зростатиме. Це є основним стимулом до створення нових перспективних матеріалів з поліпшеними параметрами, зокрема композитних теплопровідних матеріалів [1].

Особливе значення мають теплові та електрофізичні параметри металевих і діелектричних теплопровідних матеріалів різних друкованих плат, зокрема плат на металевих основах. У сучасних потужних напівпровідникових пристроях для виготовлення таких друкованих плат використовуються композитні структури DBC (Direct Bonded Copper) і DPC (Direct Plated Copper), а також комбіновані структури MC PBC (Metal Core PCB).

Застосування плат з DBC і DPC керамікою є одним із найефективніших способів відводу тепла, однак їхнє використання часто призводить до необґрунтованого збільшення вартості електронних модулів і друкованих вузлів на їхній основі через високу ціну такого типу друкованих плат.

Стандартна комбінована структура MC PCB – це структура, що складається з тепловідвідної металевої основи, наприклад, з алюмінію, міді або їхніх сплавів завтовшки 0,5 – 3,2 мм, на якій розміщено тепловідвідний діелектричний шар з високою електричною міцністю завтовшки 100 – 150 мкм, ламінований, наприклад, мідною або алюмінієвою фольгою завтовшки 35 – 350 мкм.

Наразі багатьма компаніями у світі проводяться роботи з підвищення характеристик теплопровідних склеювальних діелектриків і друкованих плат MC PCB на їхній основі. Такі виробники як Ruikai, Totking, (Китай) і Bergquist (США) на основі теплопровідних адгезивів із полімерів із добавками дрібнодисперсних порошків теплопровідних керамік випускають плати MC PCB зі значеннями теплових опорів від 0,45 до 1,42 ºС / Вт. Однак подібні друковані плати на металевих основах мають, як і раніше, істотно високі ціни. Проводилися також дослідження зі створення плат MC PCB на алюмінієвих основах, у яких було застосовано як діелектрик стандартні поліімідні плівки компанії DuPont серій LX і LA порівняно малої товщини близько 20 мкм [2]. Такі комбіновані плати давали змогу забезпечити нижчий тепловий опір усієї системи «*p-n*-перехід – навколишнє середовище», а їхній процес виготовлення доволі простий і економічний.

Тому в даній роботі компанією ТОВ "НВП "ЛТУ" (Україна) запропоновано нові конструктивно-технологічні підходи щодо створення власних плат типу MC PCB з різними варіантами поліімідних (ПІ) діелектриків. Зокрема, тонких теплопровідних ПІ діелектриків, які дають змогу ще більше знизити витрати на матеріали, які використовуються в платах,

і витрати на їх виготовлення. Застосування в комбінованих платах як діелектриків поліімідних плівок дає змогу виділити такі плати за конструкцією в особливу низку, оскільки в них використовують ПІ плівки порівняно малої товщини, близько 0,02 – 0,025 мм проти 0,1 мм у стандартних платах МС РСВ. Хоча поліімідні плівки мають низькі значення теплопровідності близько 0,12 – 0,14 Вт/(м·К), проте їхня мала товщина в платах забезпечує доволі малі теплові опори тепловідвідної системи загалом. При цьому ПІ плівки, не дивлячись на малу товщину, мають високу електричну міцність (до 110 кВ/мм і більше) порівняно з іншими типами плат МС РСВ. А використання інноваційних композитних, зокрема теплопровідних ПІ плівок з підвищеною теплопровідністю в інтервалі від 0,36 до 0,75 Вт/(м·К), дають змогу ще більше зменшити сумарний тепловий опір друкованих плат на металевих основах з тонкими ПІ діелектриками [3 – 6].

Таким чином, метою роботи є розробка нових підходів для створення вдосконалених конструктивно-технологічних рішень комбінованих плат на теплопровідних основах для електронних модулів і друкованих вузлів з використанням як діелектриків у платах тонких поліімідних плівок. Зокрема теплопровідних композитних поліімідних плівок, що випускаються промислово, з теплопровідністю від 0,36 до 0,75 Вт/(м·К).

1. Предмет та методи дослідження

1.1. Теплопровідні властивості поліімідних діелектриків для комбінованих теплопровідних друкованих плат

Оскільки основною метою використання металевих друкованих плат у радіоелектронній апаратурі є поліпшення теплопередачі від електронних теплонавантажених компонентів до системи забезпечення теплового режиму, основним критерієм під час розрахунків доцільно розглядати тепловий опір у системі «напівпровідникова структура – зворотний бік друкованої плати».

Відведення тепла від електронних компонентів здійснюється шляхом початкового розведення теплового потоку тепловідводом і подальшого скидання в навколишнє середовище за рахунок конвекції та випромінювання. Отже, для забезпечення максимального відведення тепла потрібен обґрунтований і ретельний підбір матеріалів з необхідними фізичними параметрами, оптимізація об'ємних і геометричних характеристик друкованої плати.

Теоретичні дослідження теплопровідних властивостей металевих друкованих плат полягають у розробленні теплових моделей пристрою. В основі будь-якої теплової моделі лежить поняття теплового опору. Що нижчий тепловий опір, то більше і швидше відведення тепла. Передача тепла з одного місця (наприклад, напівпровідникового кристала) в інше (навколишнє повітря) визначається товщиною шарів і тепловим опором матеріалів, а також площею їхнього дотику (що більша площа дотику, то більшу кількість тепла може бути передано). Усе це означає, що якщо переходи кристал – теплопровідна основа електронного компонента і теплопровідна основа – тепловідвід плати мають невиправдано високі значення теплового опору, то не буде забезпечено необхідне охолодження кристала.

Величина теплового опору металевих друкованих плат в основному визначається теплопровідністю діелектричного шару між фольгою друкованих провідників і металевою основою плати. За ідентичності таких величин, як площа плати, товщина фольги, товщина і властивості металевої основи, товщина шару припою або клейового з'єднання між теплопровідною основою електронного компонента і контактною площадкою на фользі друкованих провідників, товщина діелектрика і теплопровідність діелектрика будуть визначальними для теплового опору плати.

Для теоретичного дослідження різних варіантів тестових зразків діелектричних ПІ плівок було обрано такі типи поліімідних плівок, які виробляють у промислових масштабах:

- поліімідна плівка DuPont™ Kapton® HN завтовшки 25 мкм з мінімальною теплопровідністю 0,12 Вт/(м·К) [3]**;**

- теплопровідна композитна поліімідна плівка KYPI – MT китайської компанії Suzhou (Сучжоу) Kying industrial materials Co. ltd завтовшки 25 мкм із теплопровідністю 0,36 Вт/(м·К) [4]**;**
- теплопровідна композитна поліімідна плівка DuPont™ Kapton® MT завтовшки 25 мкм із теплопровідністю 0,46 Вт/(м·К) [5]**;**
- теплопровідна композитна поліімідна плівка DuPont™ Kapton® MT + завтовшки 25 мкм із теплопровідністю 0,75 Вт/(м·К) [6].

1.2. Оцінка теплових опорів різних варіантів поліімідних діелектриків для комбінованих друкованих плат

У табл. 1 наведено оціночні розрахункові значення теплових опорів різних варіантів зразків ПІ плівок. Для розрахунку теплових опорів необхідне визначення їхньої залежності від фізичних і геометричних параметрів досліджуваних зразків поліімідних плівок. Вираз для теплових опорів з урахуванням геометричних і фізичних параметрів ділянки поширення тепла має такий вигляд для плоскої стінки [7, 8]:

$$R_\theta = \frac{h}{\lambda_T \times S}, \tag{1}$$

де h, мкм – товщина ділянки поширення тепла; λ_T, Вт /(м·К) – коефіцієнт теплопровідності; S, мм^2 – площа плоскої стінки; R_Θ, С/Вт – тепловий опір ділянок теплових ланцюгів ПІ плівок.

Таблиця 1

Значення теплових опорів ділянок теплових ланцюгів зразків ПІ плівок

Ділянка розповсюдження тепла	Значення теплового опору, °С/Вт
Діелектричний поліімідний шар з геометричними розмірами 30х30 мм, товщиною 25 мкм з теплопровідністю λ = 0,12Вт /(м·К)	0,23
Діелектричний поліімідний шар з геометричними розмірами 30х30 мм, товщиною 25 мкм з теплопровідністю λ = 0,36 Вт /(м·К)	0,077
Діелектричний поліімідний шар з геометричними розмірами 30х30 мм, товщиною 25 мкм з теплопровідністю λ = 0,46 Вт /(м·К)	0,06
Діелектричний поліімідний шар з геометричними розмірами 30х30 мм, товщиною 25 мкм з теплопровідністю λ = 0,75 Вт /(м·К)	0,037

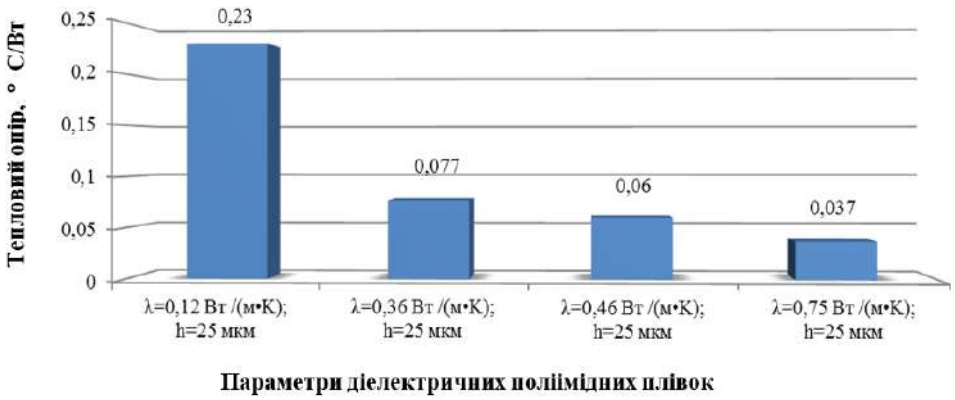


Рис. 1. Значення теплових опорів ділянок теплових ланцюгів ПІ плівок

Як випливає з табл. 1, для різних ділянок досліджених варіантів поліімідних плівок найменші значення теплових опорів можна отримати для зразків плівок завтовшки 25 мкм

із теплопровідністю $\lambda = 0{,}75$ Вт/(м·К). На рис. 1 наведено результати порівняння значень теплових опорів досліджуваних варіантів поліімідних плівок. Як випливає з рис. 1, обрані конструктивні параметри зразків поліімідних плівок мають хороші потенційні можливості для створення комбінованих друкованих плат на теплопровідній основі зі стандартними типорозмірами з істотно низькими значеннями теплових опорів поліімідних діелектриків (від ~ 0,23 до ~ 0,037 °С/Вт), особливо з теплопровідних поліімідних плівок.

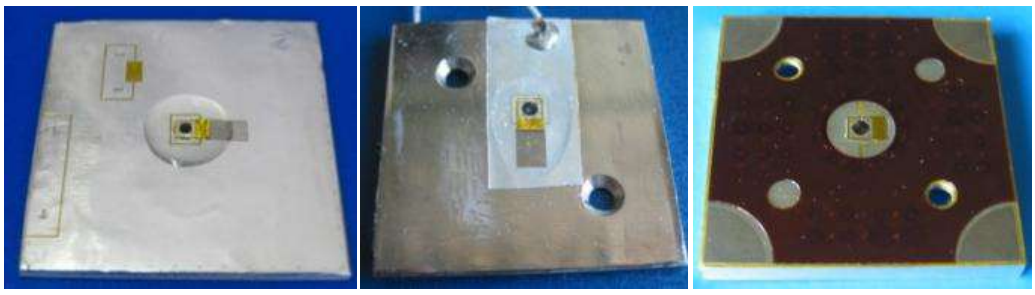
2. Результати та їх обговорення

2.1. Комбіновані теплопровідні плати для приймачів концентрованого сонячного випромінювання

Для відпрацювання технології виготовлення приймачів було розроблено та виготовлено різні типи тестових структур якості (ТСЯ) і один експериментальний зразок приймача концентрованого сонячного випромінювання на основі комбінованих плат на теплопровідних основах з діелектриками з поліміду. ТСЯ складалися з одного сонячного елемента (СЕ), комутаційної плати з комутаційним шаром з алюмінієвої фольги завтовшки 30 мкм, шаром ПІ діелектрика завтовшки 20 мкм і плоскої теплопровідної основи з алюмінію або міді. ТСЯ виготовляли шляхом приклеювання тильного контакту СЕ теплопровідним і електропровідним клеєм до плоского нікельованого мідного або алюмінієвого тепловідводу з розмірами 40х40 мм, приєднання комутаційної плати до фронтальних контактів СЕ методом УЗ-зварювання і герметизації СЕ оптично прозорою композицією Elastosil S690, Waker Silicones, Німеччина. Комутаційна плата на основі гнучкого безадгезивного алюміній-поліімідного носія приклеювалася до тепловідводу з боку поліміду діелектричним клеєм УП-10-14-2, Україна.

У ТСЯ і експериментальному зразку приймача було застосовано високоефективні триперехідні арсенід-галієві СЕ виробництва компаній ENE, Бельгія та Azur Space Solar Power GmbH, Німеччина, з ККД від 35 до 36 % відповідно. Для забезпечення електричного і теплового контакту між СЕ і тепловідводом використовували електропровідний теплопровідний клей Epo-Tek E4110, Німеччина, з товщиною клейового шару 10 – 20 мкм з теплопровідністю 1,5 Вт/(м·К). Комутаційна плата виготовлялася з безадгезивної алюміній-поліімідної плівки ФДІ-А-50. ТСЯ та експериментальний зразок приймача концентрованого сонячного випромінювання виготовлявся за допомогою Chip on flex (COF) – технології складання..

На рис. 2, *а, б, в* наведено зразки ТСЯ із розмірами 40х40 мм і різною конфігурацією комутаційних плат.



а *б* *в*

Рис. 2. Зразки ТСЯ: *а* – з комутаційною платою 40х40 мм; *б* – з комутаційною платою 30х10 мм; *в* – з комутаційною платою 40х40 мм покритої поліімідною плівкою для додаткового відведення тепла за допомогою випромінювання

На рис. 3 представлено експериментальний зразок багатоелементного приймача концентрованого сонячного випромінювання.

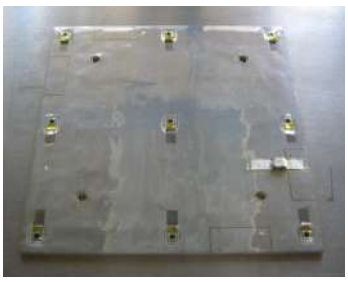


Рис. 3. Зразок приймача, виготовленого за COF-технологією складання
на плоскому алюмінієвому тепловідводі

До складу приймача концентрованого сонячного випромінювання входили:
– триперехідні GaInP / GaAs / Ge сонячні елементи, шт. – 9;
– алюміній-поліімідна комутаційна плата, шт. – 1;
– захисний діод, шт. – 1;
– алюмінієвий плоский тепловідвід, шт. – 1.

У табл. 2 наведено середні значення теплових опорів ділянок теплових ланцюгів клейових з'єднань і ПІ шарів у комбінованих платах на теплопровідних основах.

Таблиця 2

Середні значення теплових опорів ділянок теплових ланцюгів клейових шарів і ПІ шарів у комбінованих платах приймачів

Ділянка поширення тепла	Значення теплового опору, °С/Вт
Поліімідний шар	не більше 0,18
Шар діелектричного клею УП-10-14-2	не більше 0,25

Досліджені конструктивні параметри і теплові властивості розроблених комбінованих плат із застосуванням тонких поліімідних діелектриків і діелектричних адгезивів для приймачів концентрованого сонячного випромінювання підтвердили можливість забезпечення середніх значень повних теплових опорів ділянок теплових ланцюгів «ПІ шар – шар діелектричного адгезиву», що не перевищують 0,43 °С/Вт, і не поступаються за тепловою ефективністю комерційно прийнятним стандартним комбінованим структурам плат MC PCB [8, 9]:

2.2. Комбіновані теплопровідні плати для тривимірних конструкцій світлодіодних освітлювачів

Матеріали, що використовуються в корпусі світлодіодів, особливо підкладки з високою теплопровідністю, мають ключове значення для розробників у поліпшенні характеристик світлодіодів (СД), спрощенні процесу проєктування і зниженні загальних витрат. Поліімідні матеріали та їхні унікальні властивості, включно з надійністю, тонкістю, міцністю і гнучкістю ПІ плівок, широко використовуються розробниками в комбінованих платах у світлодіодному освітленні. Нові матеріали дають змогу створювати тривимірні конструкції СД модулів.

На рис. 4 показано зразок теплопровідної підкладки DuPont™ CooLam™ 3D, використаної в концепції заміни лампи A19.

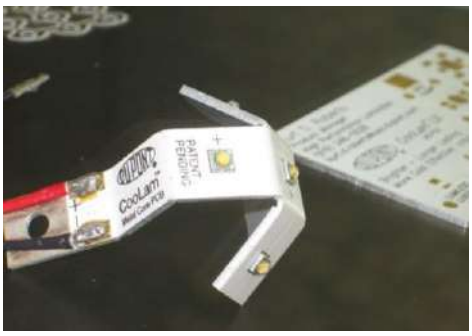


Рис. 4. Зразок теплопровідної підкладки DuPont™ CooLam™ 3D

Теплопровідні підкладки DuPont™ CooLam™ були розроблені для використання в друкованих платах з металевою основою. Ці термопластичні ламінати являють собою композит з металевої фольги і запатентованого теплопровідного поліімідного діелектрика, з'єднаного з металевою основою. Такі теплопровідні підкладки призначені для світлодіодних освітлювачів високої яскравості. Теплопровідні підкладки DuPont™ CooLam®, виготовлені із запатентованого поліімідного діелектрика, спеціально розробленого для забезпечення високої теплопровідності, розсіюють тепло швидше і надійніше, ніж звичайні плати на основі епоксидної смоли. Завдяки дуже низькому тепловому опору та максимальній робочій температурі 180 °C теплопровідні підкладки DuPont™ CooLam™ 3D забезпечують стабільну роботу в широкому діапазоні умов навколишнього середовища. Розроблена технологія дає змогу також виконувати одноразовий вигин комбінованої підкладки під кутом 90° або більше спільно з безперервною алюмінієвою основою завтовшки від 0,2 до 2 мм, що в ній є, ПІ діелектриком завтовшки приблизно 12 мкм та мідною комутуючою фольгою завтовшки приблизно 35 мкм. (рис. 4).

На рис. 5 представлено конструктивні варіанти об'ємних світлодіодних модулів на теплопровідних підкладках, що згинаються DuPont™ CooLam™ 3D.

Рис. 5. Приклади об'ємних світлодіодних модулів на теплопровідних підкладках, що згинаються DuPont™ CooLam™ 3D

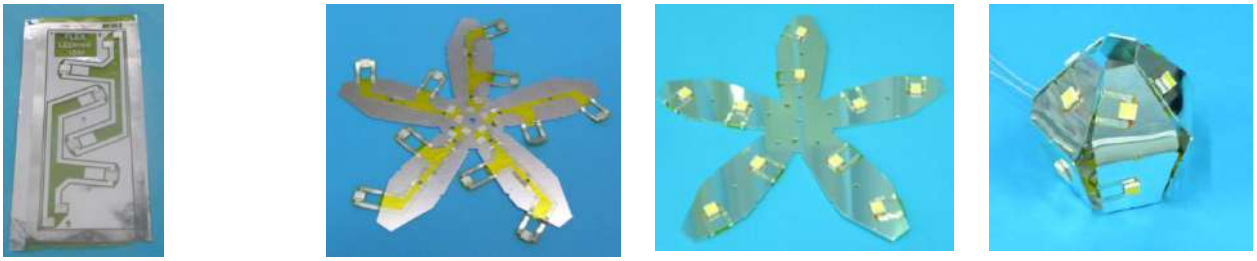
Згинальні підкладки з ПІ діелектриками дають змогу спростити конструкції об'ємних СД модулів, які коштують менше у виробництві та під час монтажу в освітлювачах, а також служать довше з плином часу. Теплопровідна плата, що згинається, дає змогу керувати світловіддачею світлодіодів, направляючи світло туди, куди потрібно [10, 11].

Авторами статті також були розроблені принципово нові дзеркалізовані об'ємні світлоефективні модулі (LEM Light Effective Module), що виготовляються на основі SMD, міні-COB і COB світлодіодів для інноваційних LED LEM LAMP світлодіодних світлоефективних ламп ретрофітів. Комбінована теплопровідна плата об'ємного світлодіодного модуля (СДМ) являє собою тримач-тепловідвід світлодіодного випромінювача, який виконаний у вигляді єдиного 3D теплопровідного світловідбиваючого дзеркалізованого елемента, що виготовляється з матеріалу MIRO-SILVER 4270AG компанії ALANOD GmbH & Co (Німеччина) різної товщини. Основа тримача-тепловідводу механічно утримує три або більше вигнутих відбивачів-радіаторів, які розташовані на відстані один від одного. На фронтальних

поверхнях теплопровідних відбивачів-радіаторів сформовано дзеркальні покриття із загальним коефіцієнтом відбиття до 98 %. Світловипромінювальні напівпровідникові прилади рівномірно встановлюються пайкою на контактні майданчики алюміній-поліімідних комутаційних плат. Плати зі світлодіодами приклеюють до поверхні відбивачів-радіаторів за допомогою теплопровідного клею TCOR компанії Electrolube (Великобританія) з теплопровідністю до 1,8 Вт/м·К.

Сутність запропонованих технічних рішень більш детально пояснюється конкретними прикладами їх виконання. На нижче наведених фотографіях представлено послідовність виготовлення різних типів світлодіодних модулів на SMD, міні-COB і COB світлодіодах, зовнішній вигляд модулів, їхній склад і компонування.

На рис. 6 представлено послідовність операцій складання зразка об'ємного світлодіодного міні-COB модуля потужністю 15 Вт на MCOB світлодіодах потужністю до 1,5 Вт.



а *б* *в* *г*

Рис. 6. Зразок дзеркалізованного об'ємного СДМ потужністю 15 Вт, його склад, компонування і послідовність складання: *а* – гнучка плата, *б* – гнучко-жорстка плата, *в* – гнучко-жорстка плата зі світлодіодами, *г* – об'ємний світлодіодний модуль

На рис. 7 представлено послідовність формування об'ємного світлодіодного SMD модуля.

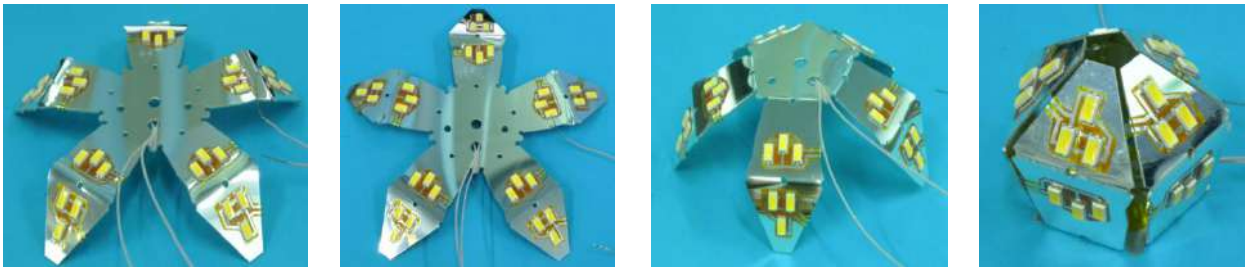


Рис. 7. Послідовність формування об'ємного світлодіодного SMD модуля

На рис. 8 представлено експериментальний зразок об'ємного дзеркалізованого світлодіодного модуля потужністю 15 Вт на COB світлодіодах.

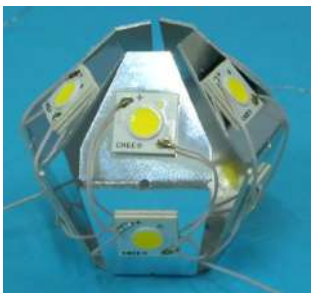


Рис. 8. Експериментальний зразок об'ємного дзеркалізованого СДМ потужністю 15 Вт на COB – світлодіодах компанії CREE, США

Запропоновані технічні рішення забезпечили підвищення енергетичної та оптичної ефективності об'ємних світлодіодних модулів при гарантуванні ефективного відведення тепла від світлодіодних приладів. Розроблена технологія дала змогу поліпшити теплові та

оптичні параметри світлодіодних модулів для нових прототипів потужних ламп із цоколем E27. Зокрема для ламп, що працюють у діапазоні потужностей від 15 до 40 Вт і більше в колбах із типорозмірами A95 і A105. Позитивний технічний результат було забезпечено завдяки збільшенню площі тримачів-тепловідводів для відводу тепла кондукцією (більш ніж у 2,5 – 3 рази порівняно з СДМ плоского типу з діаметром плоского радіатора до 80 мм) [12, 13].

2.3. Нові підходи для створення ефективних комбінованих друкованих плат на теплопровідних основах із діелектриками з поліміду

Авторами розглянуто варіанти технічних рішень комбінованих друкованих плат на теплопровідних основах з декількома типами ПІ плівок – основ з термозварювальними термопластичними адгезивними покриттями. Верхні комутувальні шари плат складаються з алюмінієвої або мідної фольги, теплопровідні основи плат виконані з міді або з алюмінієвого сплаву. Діелектричні шари плат виконуються з поліімідної плівки-основи завтовшки до 25 мкм із термопластичними покриттями з двох боків завтовшки від 3 до 15 мкм кожна, за допомогою яких ПІ діелектрик з'єднується з шаром металевої фольги та металевою теплопровідною основою.

На рис. 9 показано структуру комбінованої плати з діелектриком із поліімідної плівки з термозварювальними термопластичними покриттями з двох боків.

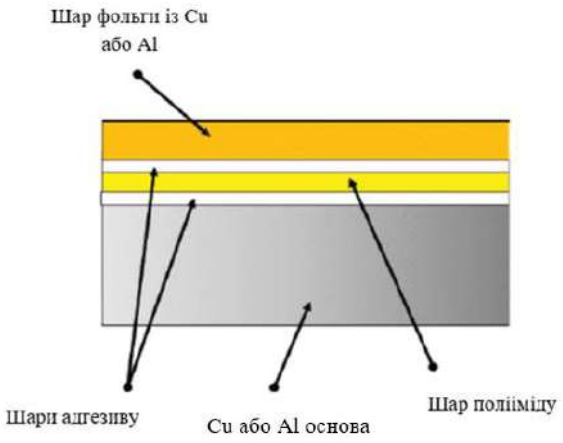


Рис. 9. Структура комбінованої плати з діелектриком із поліімідної плівки – основи з термозварювальними термопластичними покриттями

Наразі світове лідерство у створенні нових поліімідних матеріалів і в розширенні сфер їхнього застосування в промислових масштабах зберігає компанія DuPont (США). Відомі китайські компанії, такі як компанія Suzhou Kying Industrial Materials Co. Ltd, яка була заснована в 2003 р., і компанія SOLVER POLYIMIDE, яка заснована в листопаді 2010 р., також досягли істотних успіхів. Вони професійно займаються дослідженнями та новими розробками, серійним виробництвом поліімідних плівок, теплопровідних поліімідних плівок, ПІ плівок з термозварювальним термопластичним покриттям, а також експортом та імпортом різноманітних поліімідних матеріалів для двигунів та трансформаторів, матеріалів для 3D-друку, електротехнічних виробів і різноманітних пристроїв та обладнання.

2.3.1. Комбіновані друковані плати на теплопровідних основах із теплопровідними діелектриками з поліміду з термозварювальними термопластичними покриттями

Застосування в платах на теплопровідних основах промислових ПІ плівок з підвищеною теплопровідністю від 0,36 до 0,75 Вт/(м·К) з адгезивними покриттями порівняно зі стандартними поліімідними плівками з теплопровідністю 0,12 Вт/(м·К) дає змогу суттєво зменшити повні теплові опори таких плат та розширити сферу їхнього застосування.

Теплопровідні ПІ плівки з адгезивними покриттями являють собою тришарові структури, які безпосередньо впливають на характеристики повних теплових опорів комбінованих теплопровідних друкованих плат.

Компанія DuPont виробляє і поставляє на ринок теплопровідну термозварювальну поліімідно-фторопластову (ПМФ) плівку DuPont™ Kapton®120FMT616. Теплопровідна термозварювальна поліімідно-фторопластова плівка DuPont™ Kapton® 120 FMT616 володіє всіма перевагами теплопровідної поліімідної плівки Kapton® MT завтовшки 25 мкм, а також додатковими перевагами термозварювальним фторполімерним покриттям Teflon® FEP, що нанесене на обидва боки плівки Kapton® MT завтовшки 2,5 мкм кожне. Для використання в електронних пристроях плівки Kapton® MT і Kapton® FMT вдало поєднують свої високі електричні властивості, теплопровідність (0,46 Вт/(м•К)) і механічну міцність. Плівки Kapton® MT і Kapton® FMT мають вищий модуль пружності, ніж просто плівки Kapton® HN без покриття. Застосування DuPont™ Kapton®120FMT616 рекомендується в тих випадках, коли потрібна плівка, здатна до термозварювання, що має теплопровідність, вологостійкість і хімічну стійкість [14].

Компанія Suzhou Kying Industrial Materials Co. Ltd також виробляє і поставляє на ринок теплопровідну термозварювальну поліімідно-фторопластову плівку KYMIDE KYPIFMT 616. Плівка KYMIDE KYPIFMT 616 – це теплопровідна поліімідна плівка завтовшки 25 мкм, вкрита шарами фторполімерної смоли з двох боків завтовшки до 3 мкм кожен. Загальна товщина плівки становить 0,031 мм. Плівка KYMIDE KYPIFMT 616 має термоущільнювальні властивості і підвищену теплопровідність (0,36 Вт/(м•К)). Плівка має добру ізоляцію, хороший баланс електричних, хімічних і фізичних властивостей. Важливою перевагою ПМФ плівок є можливість фторопластових шарів зварюватися між собою і з поліімідом. Ця перевага дає змогу збільшити герметичність при з'єднанні шарів плівки, підвищити хімстійкість і гідростабільність багатошарових виробів [15].

Для отримання максимальних значень опору розшаровуванню (не менше 5 Н/см) багатошарових матеріалів великої площі процес пресування термозварювальних поліімідно-фторопластових плівок рекомендується проводити за таких технологічних параметрів: $T_{прес.}$ = 270 – 280°С; $P_{прес}$ = 5 – 10 МПа; $t_{прес.}$ = 10 мин. Процес монолітизації таких плівок відбувається за температури плавлення фторопластового шару (тобто за 270 – 280°С) [16].

2.3.2. Комбіновані друковані плати на теплопровідних основах з діелектриками з поліміду з термозварювальними термопластичними поліімідними покриттями

Компанія SOLVER POLYIMIDE виробляє і поставляє на ринок термопластичний поліімідний лак SolverPI-Liquid 1620. Термопластичний поліімідний лак SolverPI-Liquid 1620 має гарну ударну в'язкість при виготовленні у вигляді плівки. Лак призначений для радіаційностійкого покриття емальованих проводів і поліімідних плівок, а також для просочення вугілля- і склопластиків і для інших цілей. Температурний режим експлуатації виробів на основі лаку SolverPI-Liquid 1620 – від мінус 60 °С до плюс 220 °С, короткочасно до 300 °С. Лак також широко використовується у виробництві як поліімідний клей, що утворює низьков'язкі розплави при температурах 320 – 350 °С. Під час нагрівання відбувається перетворення клею з поліамідокислотної форми в поліімідну. При зазначеній температурі поліімідний клей переходить у в'язкотекучий стан і забезпечує щільне з'єднання поверхонь, що зшиваються і забезпечують хорошу якість склеювання (міцність розшаровування не менше 5 Н/см). За такого підходу склеювання поліімідних плівок малої площі (до декількох $дм^2$) можна здійснювати без попередньої активації поверхні плівок і не потрібне застосування спеціалізованого обладнання для щільного притиснення поверхонь, що склеюються, протягом усього часу термообробки. Застосовуваний для виготовлення ПІ покриттів рідкий термопластичний поліімідний лак SolverPI-Liquid 1620 має аналогічні властивості, що й типові рідкі поліімідні термореактивні лаки, і може бути використаний у технологічних процесах виготовлення ПІ матеріалів, що виготовляються з термореактивних ПІ лаків. Такі термопластичні покриття

мають гарну стійкість до високих температур, радіаційну стійкість, стійкість до вологи, корозійну стійкість, і мають малі значення коефіцієнта теплового розширення.

Термопластичні поліімідні покриття з лаку SolverPI-Adhesive 1620 можуть бути застосовані для виготовлення комбінованих друкованих плат на теплопровідній основі з тонкими діелектриками з будь-яких стандартних промислово випущених термореактивних ПІ плівок. А також з тонкими діелектриками з теплопровідних поліімідних плівок, що серійно випускаються і були розглянуті в цій статті раніше, такими як плівка DuPont™ Kapton® MT, плівка DuPont™ Kapton® MT + або плівка KYPI – MT. Термопластичні поліімідні покриття з лаку SolverPI-Adhesive 1620 також дають змогу виготовляти спеціалізовані двосторонні безадгезівні гнучкі ПІ плати з алюмінієвим покриттям [17, 18].

Плівки поліімідні з термопластичними поліімідними покриттями та поліімідні клеї з різних термопластичних ПІ лаків також можуть бути використані при створенні комбінованих плат типу MC PCB, що згинаються. Застосування поліімідних гнучких плівок з термопластичними ПІ покриттями дає змогу за необхідності згинати плоскі комбіновані плати з теплопровідною основою. При цьому можна виконати одноразовий вигин алюмінієвої основи, діелектрика і металевої комутуючої фольги спільно в композиті. Вигнута комбінована плата дає змогу створювати надійні складні 3 D конструкції виробів. Поліімідний матеріал (ПІ плівка – основа з термопластичними ПІ покриттями) діє як буфер механічних напружень і не передає ці механічні напруги через структуру плати, що призводить до збільшення терміну експлуатації електронних модулів на їхній основі.

На рис. 10 представлено варіанти вигнутих плат із металевою основою MC PCB компанії BAKNOR (Канада) [19].

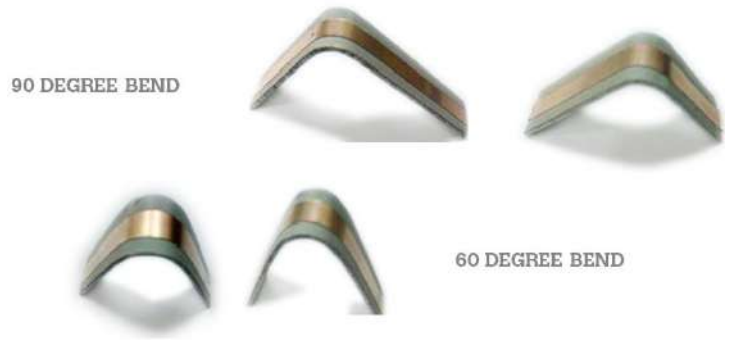


Рис. 10. варіанти вигнутих плат із металевою основою MC PCB компанії BAKNOR

Застосування сучасних промислових поліімідних плівок з термозварювальними фторопластовими покриттями та ПІ плівок з термопластичними поліімідними покриттями, що термозварюються, у комбінованих платах з теплопровідними основами, зокрема й таких, що згинаються, дають змогу суттєво розширити інноваційні можливості різних електронних модулів і друкованих вузлів, які розробляються.

Висновки

1. Проведено аналіз даних науково-технічних джерел щодо вибору сучасних матеріалів, призначених для використання при розробці та виготовленні комбінованих плат на теплопровідних основах з теплопровідними діелектриками з тонких плівок поліміду.

2. Розглянуто основні аспекти, що визначають проєктний вигляд друкованих плат на металевих теплопровідних основах, у яких застосовано як діелектрики поліімідні плівки, зокрема зі збільшеною теплопровідністю в інтервалі від 0,12 до 0,75 Вт/(м К).

3. Розглянуто конструктивні характеристики різних дослідницьких зразків тонких поліімідних плівок, зокрема теплопровідних, і виконано розрахунок їхніх теплових опорів. Результатами розрахунків підтверджено можливість створення комбінованих друкованих

плат на теплопровідних основах зі стандартними типорозмірами зі зменшеними значеннями теплових опорів поліімідних діелектриків від ~ 0,23 ° до ~ 0,037 °С/Вт.

4. Досліджено конструктивні параметри та теплові властивості розроблених комбінованих плат із застосуванням тонких поліімідних діелектриків та діелектричних адгезивів для приймачів концентрованого сонячного випромінювання, що підтвердили на практиці можливість забезпечення середніх значень повних теплових опорів ділянок теплових кіл «ПІ шар – шар діелектричного адгезиву», що не перевищують 0,43 С/Вт, та які не поступаються за тепловою ефективністю платам MC PCB зі стандартними розмірами.

5. Досліджено конструктивно-технологічні рішення об'ємних світлодіодних (СД) модулів при гарантуванні ефективного відводу тепла від світлодіодних джерел завдяки застосуванню комбінованих теплопровідних плат, що являють собою тримачі-тепловідводи світлодіодних випромінювачів, виконаних у вигляді єдиного 3D-тепловідбивного світловідбивального дзеркалізованого елемента, який є єдиним 3D-тепловідбивальним. Позитивний технічний результат було забезпечено завдяки збільшенню площі тримачів-тепловідводів більш ніж у 2,5 – 3 рази порівняно з СД модулями плоского типу. Розроблена технологія дала змогу поліпшити теплові та оптичні параметри світлодіодних модулів для прототипів потужних ламп із цоколем Е27. Зокрема для ламп, що працюють у діапазоні потужностей від 15 до 40 Вт і більше в колбах з типорозмірами А95 і А105.

6. Підтверджено потенційну можливість реалізації нових підходів для розроблення ефективних конструктивно-технологічних рішень комбінованих друкованих плат на теплопровідній основі, зокрема й тих, які можна вигинати, із застосуванням різних типів тонких теплопровідних ПІ плівок, які випускають у промисловості, із термозварювальними термопластичними покриттями, які термозварюються. Найкращий технічний результат забезпечують промислові теплопровідні ПІ плівки з термозварювальними термопластичними покриттями американських і китайських компаній, як-от DuPont, Suzhou Kying Industrial Materials Co.Ltd і Solver Polyimide. При теплопровідності використовуваних тонких ПІ плівок в інтервалі від 0,12 до 0,75 Вт/(м К) їхні тришарові структури з двосторонніми термопластичними адгезивними покриттями дають змогу забезпечити значення повних теплових опорів від 1,5 до 2,8 ^{o}С•см^{2}/Вт).

7. У компанії ТОВ "НВП "ЛТУ" (Україна) розроблено технологію та освоєно дрібносерійне виробництво поліімідних безадгезивних гнучких лакофольгових шаруватих матеріалів з потрібними геометричними розмірами та широкою номенклатурою застосовуваних алюмінієвих, мідних, нікелевих та інших типів фольги, які працюють у діапазоні температур від мінус 200 °С до плюс 250 °С для гнучких друкованих плат і які можна використати практично у всіх галузях спеціального приладобудування. Фахівці компанії успішно застосовують нові підходи для розроблення комерційно прийнятних конструкторських і технологічних рішень мікроелектронних модулів і друкованих вузлів на основі гнучких безадгезівних алюміній-поліімідних плат. Ці рішення дозволяють також ефективно використовувати Chip-on-board (COB) і Chip-on-flex (COF) технології для їх складання.

Список літератури:

1. Боднар Д. Металеві та композитні теплопровідні матеріали для потужних напівпровідникових корпусів // Компоненти та технології. 2014. № 12. С. 155 – 160.
2. Максимов А. Порівняльне дослідження теплопровідних властивостей матеріалів // Напівпровідникова світлотехніка. 2013. №4. С. 13 – 15.
3. Поліімідна плівка DuPont™ Kapton® HN, https://www.dupont.com/products/kapton-hn.html // офіційний сайт (дата звернення 05.02.2023).
4. Теплопровідна електроізолююча поліімідна плівка типу KYPI-MT (Китай), https://www.kying.com // офіційний сайт (дата звернення 05.02.2023).
5. Теплопровідна поліімідна плівка DuPont Kapton MT, https://www.dupont.com/products/kapton-mt.html // офіційний сайт (дата звернення 05.02.2023).
6. Теплопровідна поліімідна плівка DuPont™ Kapton® MT+, https://www.dupont.com/products/kapton-mt-plus.html // Офіційний сайт (дата звернення 05.02.2023).

7. Муравйов Ю. Особливості проектування та виробництва друкованих плат на металевій основі // Електронні компоненти. Україна. 2010. № 7/8. С. 83 – 86.
8. Борщов В.М. Дослідження теплових характеристик високоефективних приймачів концентрованого сонячного випромінювання нового покоління / В.М. Борщов, В.А. Антонова, О.М. Лістратенко, Я.Я. Костишин, Г.В. Буєров, І.Т. Тимчук, М.А. Проценко // Технологія приладобудування. 2012. №1. С. 3-9.
9. Борщов В.М., Лістратенко О.М., Проценко М.А. и др. Нові підходи до створення високоефективних приймачів випромінювання концентраторних сонячних модулів // Радіотехніка. 2019. Вип. 197. С. 123 – 136.
10. DuPont™ CooLam™ LA03525016 thermal substrate. Rich Wessel i Kurt Roberts, DuPont Circuit & Packaging Materials, Research Triangle Park, NC. How Substrate Materials Affect LED Reliability // Lighting Technology. July, 2012, https://www.techbriefs.com/component/content/article/tb/supplements/lt/features/articles/14414 // офіційний сайт (дата звернення 05.02.2023).
11. Patent US № US8707551B2. Bendable circuit board for LED mounting and interconnection // Daniel I. Amey et al., DuPont Electronics Inc., 2014.
12. Борщов В.М., Лістратенко О.М., Проценко М.А. и др. Нові конструктивно-технологічні рішення світлодіодних модулів для ламп-ретрофітів // Технологія та конструювання в електронній апаратурі. 2016. №6. С.3 – 10.
13. Борщов В.М., Лістратенко О.М., Проценко М.А. и др. Високоефективні об'ємні СДМ для надпотужних ламп побутового та промислового застосування // Оптоелектроніка та напівпровідникова техніка. 2017. Вип. 52. С.70 – 80.
14. Теплопровідна поліімідно-фторопластова плівка DuPont™ Kapton® 120FMT616, https://www.dupont.com/products/kapton-fmt.html // офіційний сайт (дата звернення 05.02.2023).
15. Теплопровідна поліімідно-фторопластова плівка, що термозварюється KYMIDE KYPIFMT 616 (Китай), https://www.kying.com // офіційний сайт (дата звернення 05.02.2023).
16. Власов І.С. Багатошарові полімерні матеріали та технологія отримання листів з орієнтованих поліімід-фторопластових плівок : автореф. дис. … канд. техн. наук. 2000. 22 с.
17. Термопластичний поліімідний лак SolverPI-Liquid 1620 (Китай), http://www.chinapolyimide.com/solverpi-liquid-1620 // офіційний сайт (дата звернення 05.02.2023).
18. Patent China № CN102408564B. Thermoplastic polyimide and preparation method of two-layer process adhesive-free double-side flexible copper clad plate using thermoplastic polyimide // Shengyi Technology Co Ltd., 2013.
19. Вигнуті плати з металевою основою MC PCB, http://www.baknor.com/heat-sinks/bendable-metal-core-pcbs // офіційний сайт (дата звернення 05.02.2023).

Надійшла до редколегії 02.03.2023

Відомості про авторів:
Борщов Вячеслав Миколайович – д-р техн. наук, професор, ТОВ «Науково-виробниче підприємство «ЛТУ», перший заступник директора – головний конструктор; Україна; e-mail: viatcheslav.borshchov@cern.ch; ORCID: https://orcid.org/0000-0002-5579-8932
Лістратенко Олександр Михайлович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», провідний науковий співробітник; Україна; e-mail: sasha.listratenko.12@gmail.com; ORCID: https://orcid.org/0000-0001-7643-5295
Проценко Максим Анатолійович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», начальник відділення – заступник головного конструктора; Україна; e-mail: max.protsenko.1978@gmail.com; ORCID: https://orcid.org/0000-0001-9313-1701
Тимчук Ігор Трохимович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», головний технолог; Україна; e-mail: ihortymchuk78@gmail.com; ORCID: https://orcid.org/0000-0002-6436-7253
Кравченко Олександр Вікторович – ТОВ «Науково-виробниче підприємство «ЛТУ», заступник начальника відділу; Україна; e-mail: kravcenkoaleksandr671@gmail.com; ORCID: https://orcid.org/0000-0002-7145-4304
Суддя Олександр Валерійович – ТОВ «Науково-виробниче підприємство «ЛТУ», науковий співробітник; Україна; e-mail: 4el1195@gmail.com; ORCID: https://orcid.org/0000-0002-2403-979X
Борщов Ілля Вячеславович – ТОВ «Науково-виробниче підприємство «ЛТУ», інженер; Україна, e-mail: illia.borshchov1@nure.ua**; ORCID:** https://orcid.org/0000-0002-6598-6988
Сліпченко Микола Іванович – д-р фіз.-мат. наук, професор, Інститут сцинтиляційних матеріалів НАНУ, провідний науковий співробітник; Україна; e-mail: naukovets.big@gmail.com; ORCID: https://orcid.org/0000-0002-4242-4800

УДК 621.383 **DOI:10.30837/rt.2023.1.212.11**

А.І. ЦОПА, д-р техн. наук, О.А. ЗАРУДНИЙ, канд. техн. наук

ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ ВИПРОМІНЮВАЧА ЛІДАРУ, ПОБУДОВАНОГО ЗА СХЕМОЮ ГЕНЕРАТОР-ПІДСИЛЮВАЧ

Вступ

Серед дистанційних методів контролю середньої та верхньої атмосфери одними з найбільш ефективних є лідарні, що забезпечується високою інформативністю процесів розсіювання оптичного випромінювання [1]. Сучасні лідарні дослідження в інтересах екології, метеорології та геофізики широко використовуються для моніторингу складу та динаміки атмосфери [1 – 4].

Для дослідження фотохімічних і динамічних процесів, що відбуваються у верхній мезосфері та нижній термосфері, активно застосовуються лідари, що використовують ефект резонансної флуоресценції на атомах металів та їх іонів, зокрема натрію [2, 4, 5]. У галузі атмосферних досліджень лідарні вимірювання натрієвого шару роблять важливий внесок у розуміння хімічних та динамічних процесів у районі мезопаузи [5, 6]. Крім дослідження хвильових процесів, на основі лідарних вимірювань форми спектральної лінії Na можна також відновлювати профіль температури в районі мезопаузи [7]. Як випромінювачі найчастіше використовуються рідинні лазери на органічних барвниках, що дозволяють у широких межах перестроювати довжину хвилі випромінювання.

Відомо, що застосування підсилювачів дозволяє, в загальному випадку, збільшити енергетичні параметри та ККД випромінювача лідару [1, 8], що забезпечує більший висотний діапазон лідарних вимірювань. При цьому вдається знизити вимоги до енергетичних характеристик генератора та покращити стабільність параметрів випромінювання, зменшивши навантаження на елементи резонатору. Оскільки для зондування атмосфери принципове значення має спектральна чистота випромінювання, переважно використовуються підсилювачі біжучої хвилі [8 – 10].

При побудові випромінювача лідара за схемою генератор-підсилювач в умовах постійної щільності накачування виникає проблема вибору співвідношення між протяжністю активного елемента генератора і протяжністю активного середовища підсилювача біжучої хвилі, яка б забезпечувала максимальний ККД всього випромінювача.

Основним завданням роботи була експериментальна перевірка результатів теоретичних досліджень, отриманих раніше [10 – 12], з метою визначення факторів, що впливають на вибір співвідношень довжин активних елементів генератора та підсилювача на основі органічного барвника родамін 6Ж з ламповим накачуванням при обмеженій їх сумарній протяжності.

Експериментальні дослідження ефективності випромінювача проводилися за умови загальної системи накачування для генератора і підсилювача, що забезпечувало однакові часові характеристики світлових імпульсів і значень ККД накачування.

Методика експерименту та конструктивні особливості випромінювачів

Експериментальні дослідження систем генератор-підсилювач проводилися раніше різними авторами [9, 13, 14]. Встановлено, що підсилювачі біжучої хвилі практично не змінюють просторові і спектральні характеристики пучка. Основною проблемою при дослідженні енергетичних параметрів було виділення корисної частини випромінювання із сумарної, що включає посилений шум. Так, у роботі [13] використовувалося просторове рознесення пучків за допомогою дифракційної решітки. Однак цей метод зручний тільки в тому випадку, коли спектри генерації та посиленого шуму перекриваються незначно.

Більш універсальним є використання спектрографів з позиційно-чутливими елементами [8], що дозволяють виділяти спектральні складові корисного сигналу. У цьому випадку вимі-

рювання коефіцієнта посилення проводиться порівнянням інтенсивності спектрів генератора і системи генератор-підсилювач. Для підвищення точності вимірювань енергії імпульсу до і після підсилювача лінії випромінювання інтегрувалися в межах їх ширини спектра. При цьому контроль фонової складової проводився оцінкою інтенсивності поблизу крил спектральних ліній корисного випромінювання.

У даній роботі авторами використовувалася фільтрація фонового випромінювання генератора за допомогою лінзового телескопу Ньютона (f=300мм) і розташованої у його фокальній площині польової діафрагми з кераміки діаметром близько 1,5 мм. Енергія на вході підсилювача змінювалася за допомогою каліброваних світлодільників та світлофільтрів. Діаметри підсилювача кювет близько 9 мм і генератора 7,5 мм підбиралися таким чином, щоб діаметр пучка на вході підсилювача приблизно відповідав поперечному розміру кювети підсилювача. Вимірювання енергії проводилися каліброваним вимірювачем ІКТ-1м калориметричного типу.

Відомо, що ефективність систем накачування з прямолінійними імпульсними лампами залежить не тільки від їх спектрально-енергетичних характеристик, але і від типу освітлювача, що застосовується. Найбільш ефективними освітлювачами, незважаючи на деяку неминучу неоднорідність накачування, є освітлювачі з циліндричними еліптичними дзеркалами [8, 12]. Безперечною перевагою дзеркальних освітлювачів є можливість зменшення теплового впливу ламп на активний елемент особливо в характерному для лідерів частотному режимі роботи. Ця теплова дія при досить високих потужностях накачування може бути такою, що при компактному компонуванні освітлювача для забезпечення високої якості пучка доведеться охолоджувати не тільки лампи, а й освітлювачі [8]. В експериментах використовувалася конструкція лазерного випромінювача на базі двоеліпсного освітлювача з суміщеним фокусом та дзеркальним покриттям. При цьому лампи розташовувалися у фокусах еліптичних циліндрів, а кювета з активною рідиною – у поєднаному фокусі, як показано на рис. 1.

Для збільшення протяжності активного середовища при постійній щільності накачування використовувався варіант конструкції лазерних випромінювачів, запропонований у [8, 12]. Суть способу полягає в послідовному поєднанні декількох двоеліпсних циліндричних освітлювачів (рис. 2), розташованих під кутом один до одного, щоб забезпечити можливість підключити лампи до розрядних ланцюгів і контуру рідинного охолодження. Рис. 2 ілюструє це на прикладі послідовного розташування двоеліпсних циліндрів, що використовуються для накачування загального АЕ. Такий спосіб дозволяє необмежено нарощувати довжину АЕ без збільшення протяжності ламп накачування. При цьому висока ефективність дзеркального освітлювача при раціональному виборі ексцентриситету поєднується з рівномірним накачуванням робочої речовини.

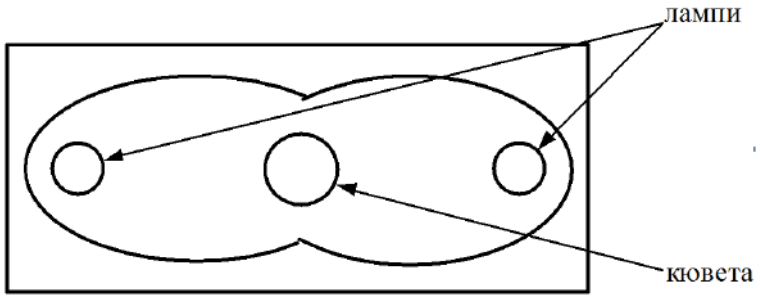


Рис. 1. Поперечний розріз двоеліпсного освітлювача кювети з активною рідиною

Для накачування активного елемента використовувалася імпульсна лампа типу ІСП-5000, яка має наступні характеристики: діаметр розрядного проміжку – 7 мм, довжина розрядного проміжку – 120мм. Гранична енергія лампи в мікросекундному діапазоні трива-

лостей залежно від тиску ксенону та тривалості імпульсу коливається від кількох сотень джоулів до одиниць кілоджоулів [8].

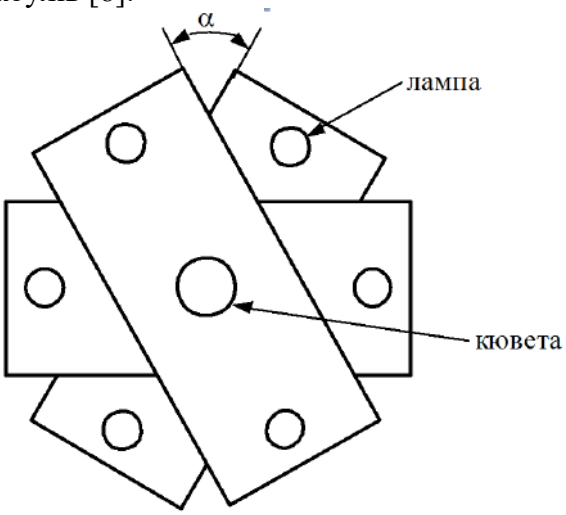


Рис. 2. Варіант послідовного поєднання двоеліпсних освітлювачів

Лазерна та підсилювальна головки мали ідентичне конструктивне виконання на базі двоеліпсних освітлювальних блоків. Довжини генератора та підсилювача змінювалися в ході досліджень з дискретністю, що визначається довжиною розрядного проміжку ламп ІСП-5000, яка дорівнює 12 см.

При виборі електричної схеми багатолампової системи накачування визначальною умовою є синхронність увімкнення всіх ламп. Несинхронність включення ламп, за часом порівнянна з часовими характеристиками розрядного імпульсу, викликає відповідну нестабільність параметрів лазерного випромінювання. У результаті аналізу літературних даних та експериментальних досліджень застосовувалася схема, представлена на рис. 3.

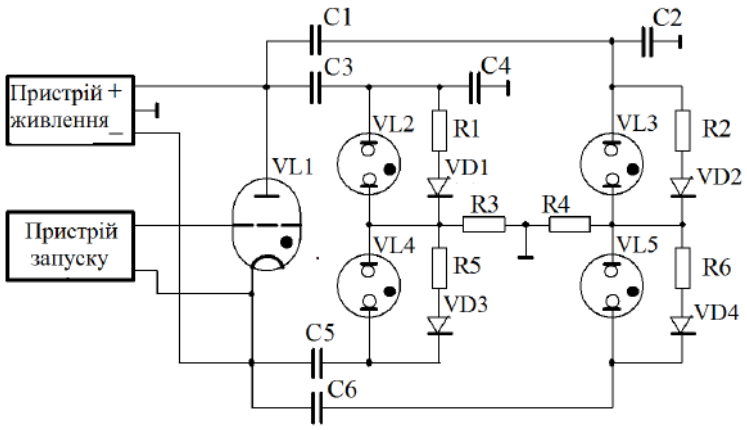


Рис. 3. Схема електрична багатолампової системи накачування

Наявність загального для усіх ламп накачування (VL2 – VL5) комутатора (VL1) забезпечує синхронність їх включення. Накопичувальні конденсатори (C1, C3, C5, C6) заряджаються від двополярного високовольтного джерела через резистори R1 – R6 і діоди VD1 – VD4. Розрядний контур утворений накопичувальними конденсаторами C3, C5 (C1, C6), розрядником VL1 і двома послідовно з'єднаними лампами-спалахами VL2, VL4 (VL3, VL5). У представленій на рис. 3 схемі до розряднику, як видно, підключені два незалежних розрядних контури. Однак схема дозволяє при необхідності збільшувати число пар ламп. Обмеженням може бути лише гранична імпульсна потужність комутатору. Послідовне з'єднання ламп забезпечує при вибраних розмірах розрядних проміжків ламп вигідніший режим узгодження і, що не менш важливо для багатолампових систем, приблизно вдвічі знижує навантаження

на розрядник за струмом комутації. Для полегшення пробою ламп необхідну асиметрію у початкові моменти розряду забезпечують конденсатори С2,С4 ємністю декілька сотен пікофарад.

Основними перевагами запропонованої схеми є такі:

- надійна синхронізація моменту включення ламп унаслідок застосування загального комутатора;
- можливість збільшення числа ламп за рахунок збільшення числа розрядних контурів, що підключаються до розрядника;
- висока напруга прикладається до ламп тільки на час дії імпульсу накачування, при цьому напруга середньої точки близька до нуля;
- напруга на кожному конденсаторі вдвічі менша за напругу, що прикладається через розрядник до ламп;
- надійний пробій розрядного проміжку лампи, оскільки через асиметрію розрядного ланцюга, що існує в початковий момент часу, до лампи прикладається подвійна напруга джерела живлення. Пробій однієї лампи, внаслідок різкого падіння напруги на ній, ініціює пробій у послідовно з'єднаній з нею лампі.

Для зменшення індуктивності розрядного кола застосовувалися малоіндуктивні елементи з використанням зворотних струмопроводів [8]. Так, зокрема, як комутатор застосовувалися малоіндуктивний розрядник типу РУ-70 і накопичувальні конденсатори типу К75-48, які з'єднувалися за допомогою коаксіальних кабелів з малим хвильовим опором. Для зменшення хвильового опору та погонної індуктивності як провідники коаксіальних кабелів використовувалися два обплетення, розділені декількома шарами фторопластової стрічки, що витримує робочу напругу розряду. Використання гнучких коаксіальних провідників дозволяє знизити механічне навантаження на електроди ламп та зменшити їх індуктивність. Багаторічний досвід експлуатації показав також їхню високу електричну надійність.

Експериментальні дослідження та обговорення результатів

В ході експериментів варіювалися протяжності активних елементів генератора і підсилювача з дискретним інтервалом, що визначається довжиною розрядного проміжку ламп накачування, рівного 120 мм. Коефіцієнт відбиття вихідного дзеркала резонатора у кожному випадку оптимізувався за максимальним значенням енергії генерації. Енергія на вході підсилювача регулювалася за допомогою комбінації світлоподілювальних пластин та нейтральних світлофільтрів. На рис. 4 зображена отримана експериментально залежність енергетичного коефіцієнта посилення підсилювача з довжиною активного елемента 360 мм від концентрації барвника активної рідини (σ_{01}^{max}nr) при фіксованій величині вхідної енергії, що дорівнювала 0,3 Дж.

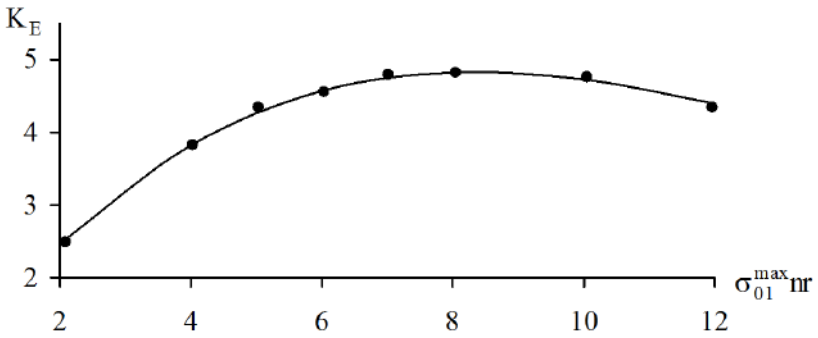


Рис. 4. Експериментально досліджена залежність енергетичного коефіцієнта посилення від концентрації активної рідини

З рис. 4 слідує, що зазначена залежність має плавний максимум і що максимальне значення енергетичного коефіцієнта посилення досягається при концентрації барвника, яка

відповідає умові $\sigma_{01}^{max}nr \approx 7 \div 8$, що якісно узгоджується з теоретичними оцінками [10]. При зазначених концентраціях барвника просторові характеристики випромінювання генератора помітно погіршуються порівняно з оптимальними при $\sigma_{01}^{max}rn \approx 5$ [10]. Тому в подальших експериментальних дослідженнях концентрація барвника вибиралася з умови $\sigma_{01}^{max}rn \approx 5$, що давало можливість використання як для генератора, так і для підсилювача загальної системи прокачування активної рідини [8]. При цьому, як видно з рис. 4, зниження коефіцієнта посилення порівняно з максимальним значенням було незначним і становило приблизно 10 %.

На рис. 5 представлено результати вимірювань максимального енергетичного коефіцієнта підсилення в режимі слабкого сигналу, тобто з вхідною інтенсивністю, значно меншою інтенсивності насичення ($I_{вх} << I_s$). Вхідна енергія світлового імпульсу тривалістю 2,2 мкс становила близько 0,1 мДж. Дослідження проводились при зміні довжини активного елемента підсилювача l_n від 12 до 60 см з інтервалом 12 см. Перевищення потужності накачування над порогом, що визначалось порівнянням осцилограм імпульсів накачування та генерації, дорівнювало 2,2. Тривалість світлового імпульсу накачування становила 3,5 мкс при енергії, що підводиться до кожної лампи, рівної 90 Дж. Отже, умови експерименту приблизно відповідали вихідним даним теоретичного розрахунку, представленого раніше [10].

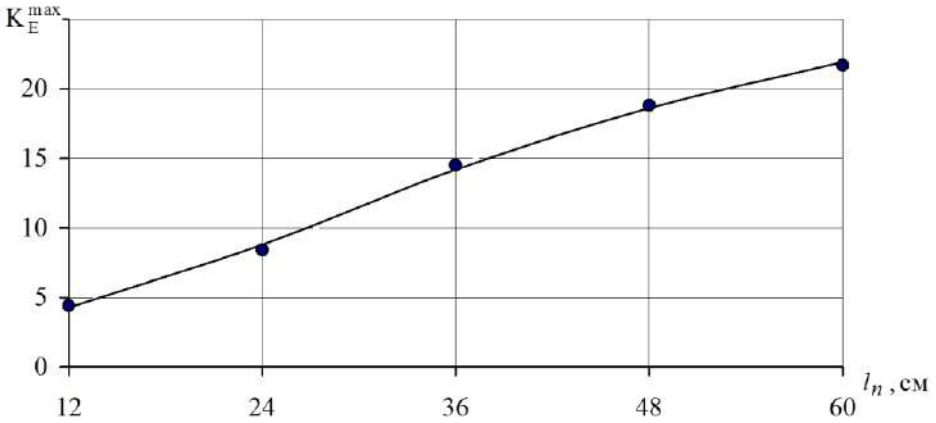


Рис. 5. Залежність максимального енергетичного коефіцієнта посилення від довжини активного елемента підсилювача у режимі слабкого сигналу

З рис. 5 видно, що енергетичний коефіцієнт посилення у даних умовах експерименту має досить велику величину і зростає зі збільшенням довжини підсилювача, що пояснюється малою величиною інтенсивності випромінювання на виході підсилювача порівняно з інтенсивністю насичення. Спочатку із зростанням довжини активного елемента підсилювача спостерігається майже експоненційне зростання коефіцієнта посилення, а в області великих довжин має місце насичення, що обумовлене зростанням інтенсивності уздовж активного елемента і збільшенням інтенсивності посиленого шуму. Результати експерименту показують потенційні можливості застосування підсилювача біжучої хвилі для посилення енергії випромінювання і якісно збігаються з результатами теоретичного аналізу [10, 16].

Режим слабкого сигналу не є оптимальним з точки зору загальної ефективності системи генератор-підсилювач. Тому доцільно було провести аналогічне експериментальне дослідження при величині вхідної інтенсивності одного порядку з інтенсивністю насичення активної рідини. На рис. 6 представлено результати вимірювань енергетичного коефіцієнта підсилення для вхідного сигналу з енергією 0,3 Дж, тривалістю 2,2 мкс при зміні довжини підсилювача. Тривалість світлового імпульсу накачування становила 3,5 мкс.

З рис. 6 видно, що коефіцієнт посилення має тенденцію до насичення при протяжності підсилювача більше 50 см, що відповідає результатам теоретичних розрахунків, отриманих раніше [10, 16]. Отже, можна зробити висновок про недоцільність подальшого збільшення

довжини підсилювача понад деяку величину, яка визначається енергією на вході підсилювача та інтенсивністю накачування, оскільки це призведе до помітного зниження загального ККД випромінювача, побудованого за схемою генератор-підсилювач.

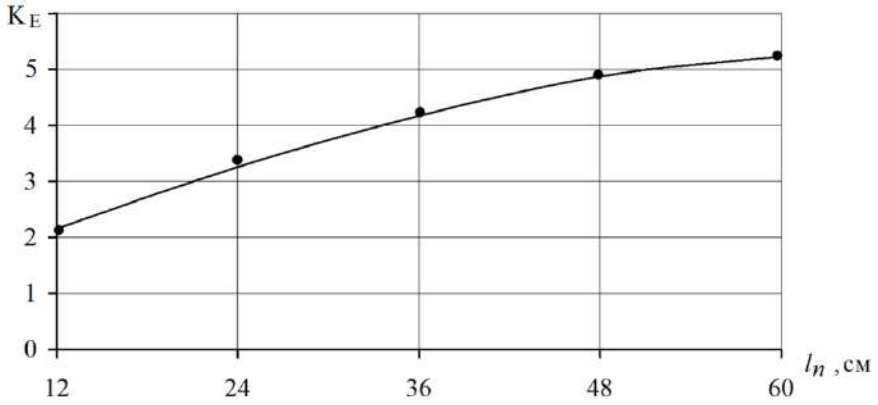


Рис. 6. Залежність енергетичного коефіцієнта посилення від довжини активного елемента підсилювача із порівняно великою вхідною інтенсивністю сигналу

З практичної точки зору важливо було також дослідити залежність енергії випромінювання при деяких фізичних обмеженнях, пов'язаних із сумарною енергією накачування системи генератор-підсилювач і, відповідно, довжиною області, що накачується. На рис. 7 представлено експериментальні результати досліджень енергії випромінювання при зміні довжини генератора при фіксованій сумарній довжині області накачування, що дорівнює 840 мм, яка обмежувалась граничним навантаженням загального комутатора (рис. 3).

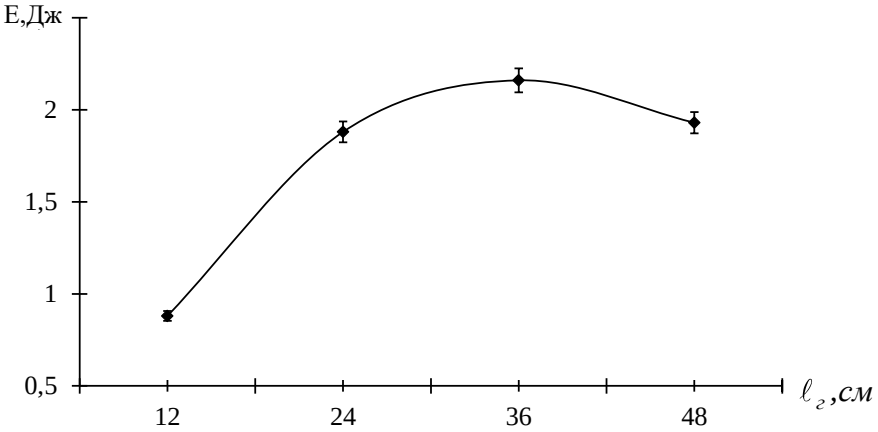


Рис. 7. Залежність вихідної енергії випромінювача, побудованого за схемою генератор-підсилювач від довжини генератора

Максимум енергії спостерігався за довжини активної області генератора 360 мм, тобто. приблизно половині загальної протяжності області, що накачується. При цьому енергія випромінювання системи генератор-підсилювач майже вдвічі перевищувала енергію генератора з тією ж довжиною, що свідчить про доцільність використання підсилювача у схемі випромінювача. Порівняння розрахункових [10] та експериментальних кривих показує, що найкраща ступінь узгодження має місце в області малих довжин підсилювача. Було встановлено, що однією з причин невідповідності коефіцієнтів посилення в області великих довжин було зниження ККД накачування зі збільшенням навантаження на загальний комутатор розрядного кола.

Висновки

1. Порівняння експериментальних результатів дослідження енергетичних характеристик системи генератор-підсилювач з теоретичними розрахунками показує їх якісну згоду і

дозволяє зробити висновок про те, що при однаковій протяжності області накачування перевага підсилювача починає позначатися лише починаючи з деякої довжини, яка залежить від щільності накачування і втрат резонатора. Якщо ж довжина області, що накачується, менше цієї величини, то вигідніше будувати випромінювач за схемою з одним тільки генератором.

2. Гранична довжина підсилювача та енергія випромінювача, побудованого за схемою генератор-підсилювач, обмежуються за рахунок зростання інтенсивності випромінювання, що посилюється уздовж активного елемента, а також збільшення інтенсивності посиленого шуму.

Список літератури:

1. Борейшо А. С., Ким А. А., Коняев М. А., Лугиня В. С., Морозов А. В., Орлов А. Е. Современные лидарные средства дистанционного зондирования атмосферы // Фотоника. 2019. Т. 13, № 7. С. 748 – 757.

2. X. Chu and G. Papen. Resonance fluorescence lidar for measurements of the middle and upper atmosphere // Laser Remote Sensing, T. Fujii, and T. Fukuchi, Eds., pp. 179 – 432, CRC Press (2005).

3. Mikhalev A.V., Tashchilin M.A. and Sakerin S.M. Effect of Atmospheric Aerosol on Ground-Based Airglow Observations // Atmospheric and Oceanic Optics. 2019. V.32. No.04. pp.410 – 415.

4. J. Wu, W. Feng, X. Xue, D. R. Marsh, J. M. C. Plane, X. Dou. The 27 Day Solar Rotational Cycle Response in the Mesospheric Metal Layers at Low Latitudes, Geophysical Research Letters, 10.1029/2019GL083888, 46, 13, (7199 – 7206), (2019).

5. Lidar observations of thermospheric Na layers up to 170 km with a descending tidal phase at Lijiang (26.7 N, 100.0 E), China / Q. Gao, X. Chu, X. Xue, X. Dou, T. Chen, J. Chen // Journal of Geophysical Research: Space Physics, 2015. Pp. 9213 – 9220.

6. Kylee Branning, Mark Conde, Miguel Larsen, Riley Troyer, Resolving Vertical Variations of Horizontal Neutral Winds in Earth's High Latitude Space-Atmosphere Interaction Region (SAIR) // Journal of Geophysical Research: Space Physics, 10.1029/2021JA029805, 127, 5, (2022)

7. Sharon L. Vadas, Erich Becker, Numerical Modeling of the Excitation, Propagation, and Dissipation of Primary and Secondary Gravity Waves during Wintertime at McMurdo Station in the Antarctic // Journal of Geophysical Research: Atmospheres, 10.1029/2017JD027974, 123, 17, (9326 – 9369), (2018).

8. Зарудный А.А., Плетенев В.Г., Верхоробин А.Л. Лазер повышенной спектральной яркости для исследования атмосферы // Радиотехника. 1998. Вып.102. С.170 – 175.

9. Шидловский В. Р., Шраменко М.В., Якубович С.Д. Перестраиваемый низкокогерентный источник света высокой спектральной яркости // Квантовая электроника. 2021. Т. 51:4. С. 287 – 292.

10. Зарудный А.А., Цопа А.И. Энергетические характеристики передатчика лидара, построенного по схеме генератор-усилитель // Радиотехника. 2018. Вып.192. С.56 – 60.

11. Петров В.В., Петров В. А., Купцов Г.В., Лаптев А.В., Кирпичников А. В., Пестряков Е.В. Моделирование процесса лазерного усиления с учётом зависимости теплофизических и лазерных характеристик среды от распределения температуры в активном элементе Yb:YAG // Квантовая электроника, 2020. Т. 50:4. С. 315 – 320.

12. Allain J.Y. High energy pulsed dye lasers for atmosferic sounding // Appl. Optics. 1989. V.18, №3. P.287 – 289.

13. Tunable dye laser amplifier chain for laser isotope separation / I. S. Grigoriev [et al.] // Quantum Electronics. 2004. Vol. 34, N.5. P. 447 – 450.

14. Vasnev N.A., Trigub M.V. and Evtushenko G.S. Features of Operation of a Brightness Amplifier on Copper Bromide Vapors in the Bistatic Scheme of a Laser Monitor // Atmospheric and Oceanic Optics. 2019. V. 32. No.04. P.483 – 489.

15. Звелто О. Принципы лазеров ; пер. под науч. ред. Т. А. Шмаонова. 4-е изд. СПб. : Лань, 2008. 720с.

16. Басецкий В.А., Зарудный А.А. Модель генерационных характеристик излучателя резонансного лидара // Радиотехника. 2010. Вып 160. С.124 – 129.

Надійшла до редколегії 02.02.2023

Відомості про авторів:

Цопа Олександр Іванович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри радіотехнологій інформаційно-комунікаційних систем. Україна; e-mail: oleksandr.tsopa@nure.ua; ORCID: https://orcid.org/0000-0002-4881-5343

Зарудний Олександр Андрійович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри радіотехнологій інформаційно-комунікаційних систем. Україна; e-mail: oleksandr.zarudnyi@nure.ua; ORCID: https://orcid.org/0000-0002-1612-0256

УДК 621.373.12 **DOI:10.30837/rt.2023.1.212.12**

В.Г. КРИЖАНОВСЬКИЙ, д-р техн. наук

СУЧАСНИЙ СТАН ТА ТЕНДЕНЦІЇ РОЗВИТКУ АВТОГЕНЕРАТОРІВ СІМЕЙСТВА КЛАСУ Е: ОГЛЯД

Вступ

Автогенератори (АГ) класу Е відносяться до пристроїв генерування високочастотної електромагнітної енергії з великим коефіцієнтом корисної дії (ККД), відрізняються добре розвиненою теорією роботи та використовуються в пристроях промислової електроніки, радіотехніки, системах передачі енергії та інформації, в тому числі і в кіберфізичних системах [1]. Після виходу монографії [1] пройшло п’ять років і за цей час відбувся відчутний прогрес як у галузі розробки цих пристроїв, які утворюють з різновидами цілу сім’ю пристроїв (класи: інверсний Е (Е-1), Е/F та інші [2, 3]), так і в розширенні сфери їх застосування. Простота конструкції АГ класу Е обумовлює їхню малу ціну, що особливо важливо для одноразових пристроїв спеціального призначення.

Мета роботи – розгляд стану та тенденцій розвитку АГ сімейства класу Е за останні п’ять років та формулювання питань класифікації цього типу.

Класифікація та принципи побудови автогенераторів класу Е

З’явившись як одна схема [4], АГ класу Е пройшли шлях розвитку конструкцій як високочастотних (ВЧ) автогенераторів до їх сучасного різноманіття, але при збереженні основних принципів роботи – усунення комутаційних втрат (виконання умов нульових напруги та її похідної в кінці інтервалу зачиненого стану транзистора) в ключовому режимі роботи. Завдяки цьому АГ класу Е демонструють високі значення ККД в діапазоні до 10 ГГц та в різноманітних схемних варіантах, від схем на дискретних компонентах до інтегральних схем з проєктними нормами 22 нм [1, 5]. В багатьох застосуваннях АГ класу Е можуть замінити каскадну схему «автогенератор-підсилювач потужності» і тому інтерес до цих пристроїв з боку розробників апаратури не знижується. Відносні недоліки АГ класу Е – низька стабільність частоти, складність налаштування на інший режим – продовжує залишатися в центрі уваги розробників, що і демонструють роботи останніх п’яти років, які оглянемо в цій роботі. Близькими питаннями, які мають загальну концептуальну базу з АГ класу Е, є АГ інших режимів роботи, які мають високий ККД – сімейства класів D, F, J та їх гібридних класів. Їх розгляд не є метою даного огляду, але цей матеріал буде корисним для розуміння їх сучасного стану.

Розгляд вже такої розгалуженої галузі, як АГ класу Е, краще проводити за допомогою зрозумілої класифікації. Питання класифікації пристроїв з високим ККД, до яких входять і АГ класу Е, не є простою справою з огляду на історію розвитку галузі, складність фізичних процесів, які покладені в основу класифікації, та інших чинників [6]. В літературі є різна систематизація АГ [1, 2], але корисною може бути і класифікація, яка представлена на рис. 1. Класифікація представлена у вигляді морфологічної таблиці і, відповідно, незаповнені клітини можуть бути вказівкою на створення нових конструкцій АГ.

Пропонується класифікувати АГ за типом кола зворотного зв’язку та структурою вихідної ланки, в таблиці більшість схем стосується схеми автогенератору на одному ключовому елементі, але для деяких рішень можливі і двотактні схеми. Також явно не представлені схеми генераторів зі зворотним зв’язком за рахунок імпедансу у колі стоку транзистора (генератор Колпіца), тому що така схема суперечить ідеї класу Е, і тому теоретично не може мати максимального ККД, хоча її використання можливе в реальних радіотехнічних пристроях, до того ж ця схема може бути краще придатною для реалізації у вигляді інтегральної схеми [1, 14]. Вказівка на конструкцію АГ міститься у розташуванні посилання на рис. 1. Тлумачення вимагають посилання, які відносяться до рядків «інші» – [22, 23, 25 – 29],

це конструкції, які поки існують у одиничних варіантах. У роботі [22] розглядається АГ класу EF2, в якому додано послідовний коливальний контур, що шунтує ключ. Це одне з рішень, яке дозволяє зменшити максимальну напругу на транзисторі та отримати більшу потужність при інтегральному виконанні АГ. Робота [23] описує двотактну схему АГ в інтегральному виконанні, яка має гарні шумові характеристики при високій енергетичній ефективності в діапазоні 4 ГГц. Тут виконання умов режиму класу Е дає гарну комбінацію параметрів АГ.

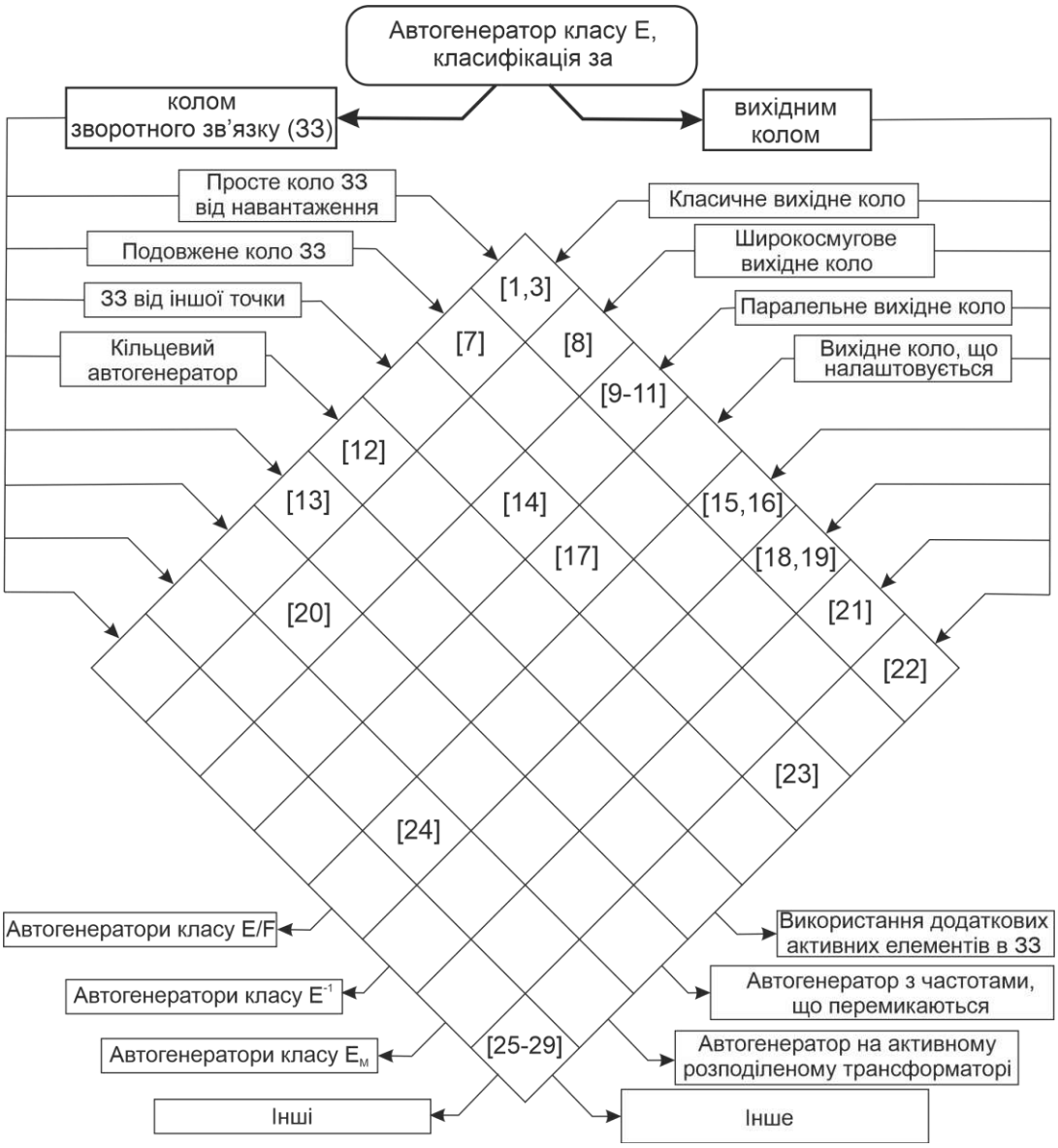


Рис. 1. Класифікація автогенераторів сімейства класу Е за структурою вихідного кола та кола зворотного зв'язку

Робота [25] присвячена максимізації діапазону синхронізації АГ класу Е, що дозволяє розширити можливості застосування синхронізованих АГ класу Е в радіотехнічних, комунікаційних та сенсорних системах. Цікавий різновид АГ класу Е розглянуто в [26] – це практично схема автогенератора з індуктивно зв'язаним додатковим резонатором, в який входить навантаження генератора. Вирішується нелінійна задача передати максимальну потужність з високим ККД і при цьому забезпечити виконання вимог до частоти генерації.

В роботі [27] розглядається підключення додаткової резонансної системи – відрізка довгої лінії до затвору транзистора, це дозволяє підвищити стабільність частоти генерації

при дії дестабілізуючих факторів на АГ класу Е. В [28] експериментально вивчено роботу двох взаємно синхронізованих АГ класу Е в режимах синфазних та протифазних коливань на їх виходах. Це допоможе здійснити просторове складання потужності АГ. У роботі [29] розглянуто автогенератор на безкорпусних GaN транзисторах на частоту 433 МГц для бездротової передачі енергії, цікавим цю роботу робить поєднання кількох граничних параметрів. Потім цей підхід був розвинений в розробці диференціального генератора на більшу потужність [30].

Аналіз діаграми рис. 1 свідчить про те, що увагу дослідників привертають гібридні режими роботи АГ сім'ї класу Е – режими класів Е/F, зокрема Е/Fmn, де індекси свідчать або про номери гармонік, для яких виконуються умови класу «інверсний F» (F-1), або вказують, що умови виконуються для всіх парних гармонік – E/Fodd. В роботах [15, 16] розглядаються однотактні генератори класу E/F3 які мають перевагу в зниженні максимальної напруги стоку транзистора, що важливо для НВЧ транзисторів та інтегральних мікросхем. В роботі [24] розглянуто нову конструкцію автогенератору з високим ККД і виконанням умов класу Е (точніше E/Fodd) з використанням активного розподіленого трансформатора – пристрою, який об'єднує потужність кількох транзисторів при роботі на спільне навантаження.

Далі розглянемо три конструкції АГ, які цікаві в плані подальшого розвитку цих пристроїв.

Автогенератор класу E/F_{odd} об'єднаний у активному розподіленому трансформаторі

Активний розподілений трансформатор в даному випадку має три пов'язані обмотки – відповідно одна пов'язана з навантаженням, друга складається з двох частин кожного з парціальних автогенераторів, а третя – з двох частин кіл зворотного зв'язку кожного автогенератора. Завдяки такій конструкції забезпечено потрібні фазові співвідношення кожного автогенератора та складання потужності у навантаженні. На рис. 2 показана спрощена схема АГ, за винятком обмотки, що пов'язана з навантаженням, та пасивних елементів кола зворотного зв'язку [24].

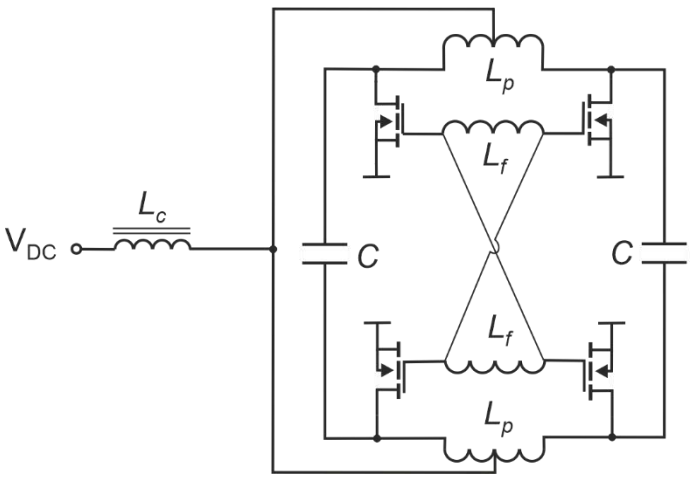


Рис. 2. Схема автогенератора класу E/F_{odd} [24]

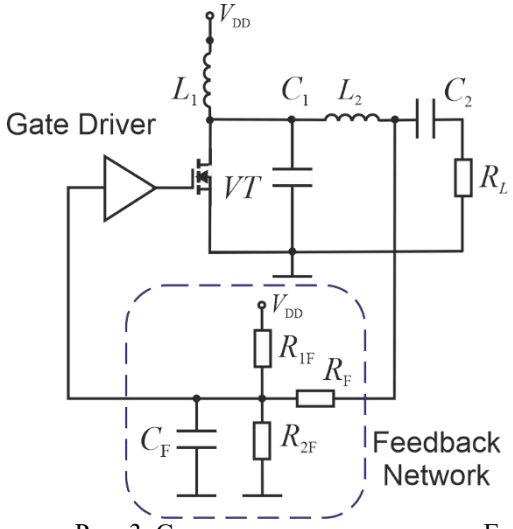


Рис. 3. Схема автогенератора класу Е з іншим колом зворотного зв'язку [17]

Перевагами такого АГ є симетрія схеми, – розподілений трансформатор це практично квадратна структура з трьох шарів, що утворюють обмотки, які індуктивним зв'язком передають енергію. Внаслідок цього форми напруги і струмів і транзисторах схожі між собою, що забезпечує високий ККД. Автогенератор на частоті 10,1 МГц має 324 Вт вихідної потужності при ККД 75 %. Принцип роботи та конструкція АГ можуть бути застосовані у мікрохвильовому діапазоні.

Автогенератор, який сам налаштовується на оптимальну роботу

В роботі [17] описано АГ класу Е, що відрізняється іншою принциповою схемою, в якій змінено коло зворотного зв'язку, внаслідок чого при зміні параметрів елементів, що може трапитися, наприклад, внаслідок старіння, майже не змінюється частота генерації і зберігається режим класу Е, а це забезпечує збереження вихідної потужності та ККД. Така схема може бути корисною в перетворювачах енергії та в системах бездротової передачі енергії. На рис. 3 показана принципова схема такого АГ.

Очікуваний ефект роботи досягається вибором параметрів елементів схеми, за рахунок чого фазовий зсув в колі зворотного зв'язку буде майже постійним і тому чутливість до зміни параметрів елементів істотно знижується, це дало підставу авторам роботи [17] назвати конструкцію самонастроюваним АГ класу Е.

Автогенератор, режим роботи якого не залежить від зміни навантаження

Імпеданс навантаження Z_L або його дійсна частина R_L є визначальним для характеристики режиму роботи підсилювачів та автогенераторів з високим ККД. Тому зміна його значення під час роботи, наприклад під впливом високочастотної потужності, призводить до зміни режиму роботи, і, як правило, до зниження ККД та зміни частоти. Існують різні рішення, як запобігти цим негативним наслідкам, – загальне рішення, це створити контур автоматичного регулювання, що потребує суттєвого ускладнення АГ. В роботі [19] надано схему АГ інверсного класу Е (Е-1), який за рахунок свого режиму роботи та деякої зміни кола зворотного зв'язку стає набагато менш чутливим до зміни навантаження. На рис. 4 показано схему такого АГ. Інверсний клас Е є дуальним до звичайного класу Е, тобто замість умови нульової напруги у мить перемикання діє умова нульового струму також при його нульовій похідній. Це відповідає взаємній заміні форм напруги та струму через активний пристрій (транзистор). В схемі на рис. 4 за рахунок паралельного контуру C_0L_2 на навантаженні діє синусоїдальна напруга, амплітуда якої значною мірою визначається напругою живлення V_{DD} і зберігає своє значення при зміні опору навантаження, відповідно не змінюються умови перемикання транзистора і автогенератор зберігає свою частоту. Цьому також сприяє наявність трансформатора L_2, L_3 в колі зворотного зв'язку, який фіксує зсув фаз, що дорівнює π, в колі зворотного зв'язку.

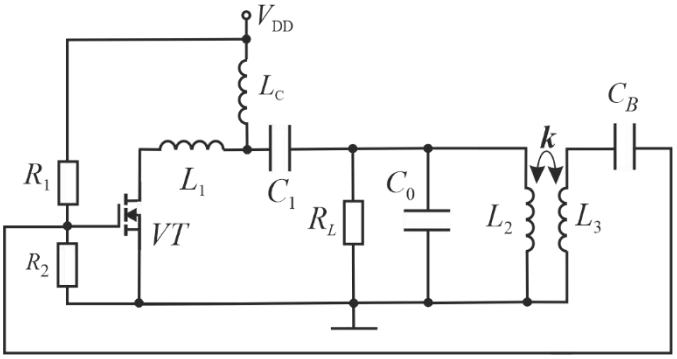


Рис. 4. Схема автогенератора класу E^{-1}, що не чутливий до навантаження [19]

Таким чином, дана схема [19] реалізує автономні коливання незалежно від змін навантаження та реактивних компонентів схеми. Ці властивості є перевагою в системах бездротової передачі енергії.

Тренди розвитку конструкцій автогенераторів класу Е та їх застосування

Цікавим є той факт, що майже половина джерел, які вказані на рис. 1, опубліковані за останні п'ять років. Це свідчить про потребу в таких пристроях і про появу нових застосу-

вань, де використовуються автогенератори з високим ККД. Зрозуміло, що АГ класу Е – це один з наявних класів пристроїв з високим ККД і конкуренцію їм складають окремі автогенератори та підсилювачі класу Е та інших високоефективних видів, які можуть мати кращу стабільність частоти та можуть бути більш гнучкими в комунікаційних застосуваннях. Властивості пристроїв класу Е – висока енергоефективність, розроблені методики проєктування, робота в широкому діапазоні параметрів – обумовлюють їх широке застосування і ця тенденція лише посилюється з розвитком аналогових інтегральних мікросхем, в яких застосовуються режими роботи класу Е [4, 31].

Однією з нових тенденцій є те, що АГ класу Е зараз все більше розробляються для виконання системних вимог та в тісному зв'язку з іншими елементами радіотехнічних [32], технологічних [33 – 35], біомедичних систем [36]. Це можуть бути підсхеми (елементи) для регулювання, частіше всього автоматичного, частоти, вихідної потужності та інше. Нові конструкції АГ класу Е можуть частково робити це без додаткових схем чи контролерів. Хоча, звісно, створення закінченої системи тільки з використанням самого автогенератору доволі складно зробити. Тому багато нових застосувань АГ класу Е саме пов'язані з їх інтеграцією у різні системи, і тут як раз важко провести межу між застосуванням автогенераторів чи підсилювачів класу Е, бо можна використовувати АГ класу Е в режимі синхронізації фактично як підсилювач, але з кращою надійністю або енергоефективністю [1]. Або використовувати контролер пристрою як генератор сигналу, а високочастотну потужність отримувати з виходу підсилювача класу Е.

Висновки

Протягом останніх п'яти років продовжували розвиватися автогенератори класу Е як і інші представники пристроїв з високим ККД, цей розвиток пов'язаний з розширенням їх використання в системах передачі енергії та інформації, в тому числі бездротової, в Інтернеті речей та біомедичних системах. Сучасний стан цих пристроїв дозволяє створювати джерела високочастотної енергії на частотах до мікрохвильового діапазону з потужністю від міліват до кіловат з різним виконанням. Тенденцією розвитку цих пристроїв є розробка конструкцій, які налаштовані під різні галузі застосування.

Запропоновано варіант класифікації автогенераторів класу Е, який вказує на можливі напрями розвитку цієї галузі техніки. Ці пристрої можуть бути застосовані для розробки енергоефективних пристроїв кіберфізичних систем та систем спеціального призначення.

Список літератури:

1. Крижановський В.Г., Макаров Д. Г., Чернов Д. В., Крижановський В. В. Автогенератори класу Е ; за ред. В. Г. Крижановського / ДонНУ ім. Василя Стуса. Вінниця : Нілан-ЛТД, 2017. 220 с.
2. Grebennikov A., Franco M. J. (2021) Switchmode RF and Microwave Power Amplifiers Third edition. Academic Press. 819 p.
3. Kazimierczuk M. (2014). RF Power Amplifiers. Second edition. Wiley. 687 p.
4. Ebert J., & Kazimierczuk M. Class E high-efficiency tuned power oscillator // IEEE Journal of Solid-State Circuits. 16(2). P. 62 – 66. https://doi.org/10.1109/jssc.1981.1051542
5. Seidel A., Wagner J., and Ellinger F. (2022). Polar Transmitter with Pseudo-Differential Inverse Class-E Output Stage in 22 nm FD-SOI // 14th German Microwave Conference (GeMiC). 01 – 04.
6. Крыжановский В.Г., Прилипская А.С. О классификации транзисторных усилителей мощности // Прикладная радиоэлектроника. 2010. Т. 9(4). С. 554 – 561.
7. Makarov D.G., Kryzhanovskyi V.V., Chernov D.V. (2016) Class E oscillator with electrically elongated feedback network // 2016 Intern. Conf. Radio Electronics & Info Communications (UkrMiCo). P.1 – 3. doi: 10.1109/UkrMiCo.2016.7739617.
8. Крыжановский В.Г. Автогенератор класса Е с расширенной полосой перестройки // Радиотехника. 2013. Вып. 175. С. 184 – 188.
9. Kurumizawa T. & Koizumi H. (2021) Voltage-Source Parallel Resonant Class E Oscillator // IEEE International Symposium on Circuits and Systems (ISCAS). 1 – 5. doi: 10.1109/ISCAS51556.2021.9401750.
10. Yamashita Y. & Wada K. (2017) Wireless power transmitter using parallel-tuned class-E power oscillator // International Symposium on Electronics and Smart Devices (ISESD). 287 – 290. doi: 10.1109/ISESD.2017.8253350.

11. Matsuhashi S., et al., (2020) Load-Independent Self-Tuned Parallel Resonant Power Oscillator // 2020 IEEE Energy Conversion Congress and Exposition (ECCE). 1571 – 1576. doi: 10.1109/ECCE44975.2020.9236069.

12. Крыжановский В.Г., Охрименко Ю.Г., Чернов Д.В. Анализ области устойчивой работы кольцевого автогенератора класса Е // Радиотехника. 2013. Вып. 175. С.189 – 194

13. Ahmadi M. M. & Pezeshkpour S. (2020). A Self-Starting Class-E Power Oscillator with an Inverting Gate Driver // IEEE Transactions on Industrial Electronics. 67(10). 8344 – 8354. https://doi.org/10.1109/tie.2019.2949533

14. Laskovski A. N. & Yuce M. (2010). Class-E oscillators as wireless power transmitters for biomedical implants. In E. Cianca (Ed.) // 3rd International Symposium on Applied Sciences in Biomedical and Communication Technologies 2010. Pp. 1 – 5. IEEE, Institute of Electrical and Electronics Engineers. DOI: 10.1109/ ISABEL. 2010. 5702913

15. Inaba T. & Koizumi H. (2018). Class E/F_3 Tuned Power Oscillator // IEEE Transactions on Power Electronics. 33(2). 1420–1427. https://doi.org/10.1109/tpel.2017.2686900

16. Krizhanovski V.G., Chernov D.V., Grebennikov Andrei. (2018) Low-Voltage Class E/F_3 High Frequency Oscillator // 14th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, Lviv-Slavske, Ukraine. 607 – 611. doi: 10.1109/TCSET.2018.8336275

17. Ahmadi M. M., & Salehi-Sirzar M. (2019). A Self-Tuned Class-E Power Oscillator // IEEE Transactions on Power Electronics. 34(5). 4434 – 4449. https://doi.org/10.1109/tpel.2018.2859387

18. Cantu H. I., Mury T., Fusco V.F. (2007) Inverse Class E amplifier and oscillator phase noise characteristics. European Microwave Conf., 9-12 Oct. 2007: proc. Munich, Germany. 740 – 742.

19. Komiyama Y., Matsuhashi S., Zhu W., Nguyen K., Uematsu T., Ito Y., Mishima T., & Sekiya H. (2022). Wireless power transfer system with load-independent inverse class-E oscillator // Nonlinear Theory and Its Applications, IEICE, 13(2). 465 – 470. https://doi.org/10.1587/nolta.13.465

20. Krizhanovski V., Kryzhanovskyi V., Grebennikov A., (2020) Class E oscillator with two switchable frequencies // IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET – 2020), IEEE. doi: 10.1109/TCSET49122.2020.235576

21. Miyahara R., Wei X., Nagashima T., Kousaka T. & Sekiya H. (2012). Design of Class-E_M Oscillator with Second Harmonic Injection // IEEE Transactions on Circuits and Systems I: Regular Papers. 59(10). 2456 – 2467. https://doi.org/10.1109/tcsi.2012.2188936

22. Madureira H., Deltimple N., Kerhervé E., Haddad S. (2013) Design of a class EF2 power oscillator for RF communication application // IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS). 763 – 766.

23. Barzgari M., Ghafari A., Nikpaik A. & Medi A. (2021). Even-Harmonic Class-E CMOS Oscillator // IEEE Journal of Solid-State Circuits. 1594 – 1609. https:// doi.org/10.1109/jssc.2021.3124971

24. Apperley T., Nielsen J. & Okoniewski M. (2020). A Class E/F_{odd} Power Oscillator Incorporating a Distributed Active Transformer // IEEE Transactions on Microwave Theory and Techniques. 68(6). P. 2409 – 2418. https://doi.org/10.1109/tmtt.2020.2977898

25. Yabe Y., Tanaka H.-A., Sekiya H., Nakagawa M., Mori F., Utsunomiya K., & Keida A. (2020). Locking Range Maximization in Injection-Locked Class-E Oscillator – A Case Study for Optimizing Synchronizability // IEEE Transactions on Circuits and Systems I: Regular Papers. 67(5). 1762 – 1774. https://doi.org/10.1109/tcsi.2019.2960847

26. Ardila V., Ramirez F. & Suarez A. (2021). Nonlinear Analysis of a High-Power Oscillator Inductively Coupled to an External Resonator // IEEE Microwave and Wireless Components Letters. 31(6). P. 737 – 740. https://doi.org/10.1109/lmwc.2021.3064246

27. Kryzhanovskyi V., Chernov D., Makarov D. & Krizhanovski V. (2022). A Simple Method to Increase the Stability of a Class E Power Oscillator // 2022 IEEE 16th Int. Conf. on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). pp. 785 – 788. doi: 10.1109/TCSET55632.2022.9766864.

28. Krizhanovski V., Makarov D., Kryzhanovskyi V. & Grebennikov, A. (2021) Mutual synchronization of class E oscillators // IEEE 5th Int. Conf. on Information and Telecommunication Technologies and Radio Electronics (UkrMiCo) 254 – 257. doi: 10.1109/UkrMiCo52950.2021.9716687.

29. Rezk T. M., Fahmy G. A., Ibrahim S. A. & Ragai H. F. A 433 MHz e-GaN HEMT based Power Oscillator for Far Field Wireless Power Transfer // 2020 8th International Japan-Africa Conference on Electronics, Communications, and Computations (JAC-ECC). 2020. 80 – 83. doi: 10.1109/JAC-ECC51597.2020.9355883.

30. Rezk T. M., Fahmy G. A., Ibrahim S. A. & Ragai H. F. (2021). Design of a differential power oscillator for 433 MHz WPT using e-GaN HEMTs // Ain Shams Engineering Journal. https://doi.org/10.1016/j.asej.2021.09.008

31. Saheb Z. & El-Masry E. (2019). An energy-efficient and ultra-low-voltage power oscillator in CMOS 65 nm // Analog Integrated Circuits and Signal Processing. doi:10.1007/s10470-019-01431-z

32. Makhoul R., Zhuang J., Maynard X., Perichon P., Frey D., Jeannin, P.-O. & Lembeye, Y. (2019). A Very High Frequency Self-Oscillating Inverter Based on a Novel Free-Running Oscillator // IEEE Transactions on Power Electronics. 34(9). 8289 – 8292. https://doi.org/10.1109/tpel.2019.2904886

33. Jarndal A. & Petrovic T. (2018) GaN-Based Oscillators for Wireless Power Transfer Applications // 2018 International Conference on Advanced Computation and Telecommunication (ICACAT). 1-5. doi: 10.1109/ICACAT.2018.8933533

34. Jong-Ryul Yang, (2018) A Class E Power Oscillator for 6.78-MHz Wireless Power Transfer System // Electr Eng Technol. 13(1): 220 – 225.

35. Phaebua K., Lertwiriyaprapa T., Boonpoonga A., Rattanarungngam D. & Torrungrueng, D. (2022) An Experimental Study of Effect of Dielectric Materials on Wireless Power Transmission at 6.78 MHz // 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON). 1–4. doi: 10.1109/ECTI-CON54298.2022.9795554

36. Ahmadi, M. M., Pezeshkpour, S., & Kabirkhoo, Z. (2021). A High-Efficiency ASK-Modulated Class-E Power and Data Transmitter for Medical Implants // IEEE Transactions on Power Electronics, 1. https://doi.org/10.1109/tpel.2021.3092829

Надійшла до редколегії 05.03.2023

Відомості про авторів:

Крижановський Володимир Григорович – д-р техн. наук, професор, Донецький національний університет імені Василя Стуса (м. Вінниця), професор кафедри прикладної математики та кібербезпеки; Україна; email: v.krizhanovski @donnu.edu.ua; ORCID: https://orcid.org/0000-0002-2685-9740

УДК 624.323.64.012.8 **DOI:10.30837/rt.2023.1.212.13**

О.Д. МЕНЯЙЛО, канд. техн. наук, В.Г. МАХОНІН, М.С. СВІТЛИЧНИЙ

ДОСЛІДЖЕННЯ ПАРАМЕТРІВ ГЕНЕРАТОРА НА ЛАВИНО-ПРОЛІТНОМУ ДІОДІ

Вступ

Мікрохвильові коливання, що генеруються в *p-n*-переході, який включено в зворотному напрямку, були теоретично розглянуті ще в 1958 р. W.I. Read. Він запропонував таку структуру діода, який мав негативний динамічний опір. Робота генератора на основі такого діода базувалася на тому, що в *p-n*-переході в результаті зіткнень генеруються носії заряду, які взаємодіють з високочастотним полем зовнішньої коливальної системи, віддаючи йому свою енергію в певні проміжки часу.

Значно пізніше А.С. Тагер, А.І. Мельников, Г.П. Ковельов, А.М. Цебієв, а потім Т. Місава [1] займалися структурами *p-n* і *p-i-n* і, використовуючи принцип дії діода Read, варіюючи різні структури, вони створили так званий лавино-пролітний діод (IMPATT).

Сьогодні мікрохвильові генератори, побудовані на лавинних діодах, забезпечують досить високу ефективність у порівнянні з іншими твердотільними генераторами, що працюють у безперервному режимі. Генератори на лавино-пролітних діодах застосовуються в радіорелейних лініях, системах зв'язку, сліпої посадки літаків, електронних вимірювачах висоти, стаціонарних та переносних радіолокаційних станціях, і навіть в пристроях боротьби з шкідниками сільського господарствах [2, 3].

Певний інтерес представляє собою дослідження проблем технологічності, стабільності, регулювання вихідної частоти та потужності, синхронізації таких генераторів, їх електронного налаштування тощо [4]. Розгляду деяких з цих проблем присвячена ця стаття.

Особливості лавинного ефекту в лавинно-пролітному діоді

Сьогодні мікрохвильові генератори, побудовані на лавинних діодах, демонструють свою високу ефективність, яка обумовлюється в значній мірі їх принципом дії.

Проаналізуємо процеси, що протікають у *p-n*-переході, підключеному у зворотному напрямку (рис. 1, *а*). Можливий вид розподілу електричного поля в такому діоді представлено на (рис. 1). На цьому рисунку L_a представляє собою товщину напівпровідникового слою діода певної провідності. Розподіл кількості носіїв заряду, в цьому випадку, представлено на рис. 1. Природно, що концентрація носіїв збільшується від нуля всередині переходу до максимуму на його краях.

У частині переходу дякуючи зовнішній напрузі вільних носіїв заряду практично немає. Це спустошений шар (на рис. 1). Це шар напівпровідника, обмежений з обох боків шаром нейтрального напівпровідника. Напруженість електричного поля в цьому шарі прагне $E \to 0$, в той час як у площині, де просторовий заряд іонів змінює знак, вона максимальна. Із збільшенням зовнішньої напруги електричне поле також збільшується, а спустошений шар розширюється. Коли напруженість поля досягає деякого критичного значення $E = E_{krit}$, починається процес інтенсифікації ударної іонізації, що призводить до лавиноподібного зростання кількості носіїв заряду.

Імовірність іонізації сильно залежить від величини електричного поля, тому область, де утворюються носії заряду, є більш-менш вузьким шаром. Цей шар називається шаром множення. На практиці цей шар розташований по обидва боки технологічного переходу.

Завдяки зовнішньому електричному полю електрони, створені в шарі розмноження, дрейфують через відрізок збідненого шару до граничної області нейтрального напівпровідника, тоді як дірки проходять через *p*-шар, а електрони через *n*-шар.

Оскільки напруженість електричного поля велика, $E=E_{krit}>10^4 v/cm$, швидкість дрейфу носіїв заряду практично постійна і не залежить від електричного поля. Із збільшенням енергії носіїв заряду збільшується ймовірність того, що носії заряду більш часто будуть стикаються з іонами кристалічної решітки, що в кінцевому підсумку призводить до насичення швидкості носіїв.

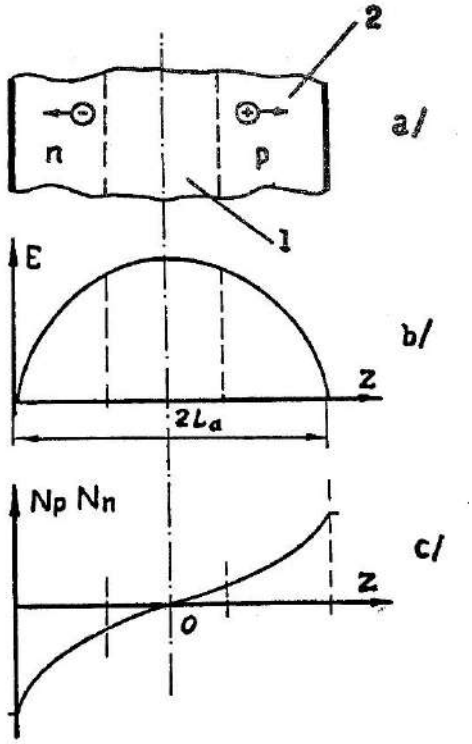


Рис. 1. Розподіл електричного поля та концентрації носіїв
в *p- n* напівпровідниковому переході

Таким чином, перехід *p-n*, підключений у зворотному напрямку, який знаходиться під дією зовнішнього електричного поля напруженістю $E=E_{krit}$, можна вважати еквівалентним діодові, в якому роль катода відіграє шар розмноження, а роль проміжку вільного пробігу відіграє решта шару напівпровідникової структури. Очевидно, струм, що надходить із шару розмноження, збільшується або зменшується в залежності від розміру силового поля, що переважає в цьому шарі.

Однак миттєве значення електричного поля визначає не величину самої лавини, а лише швидкість зміни струму через діод (для утворення лавини потрібен певний час), тому зміна струму не слідує відразу за зміною електричного поля, а затримується на фазовий кут π/2.

Зсув фаз між напруженістю електричного поля на діоді і струмом, що протікає через нього, призводять до негативного опору діода (рис. 2).

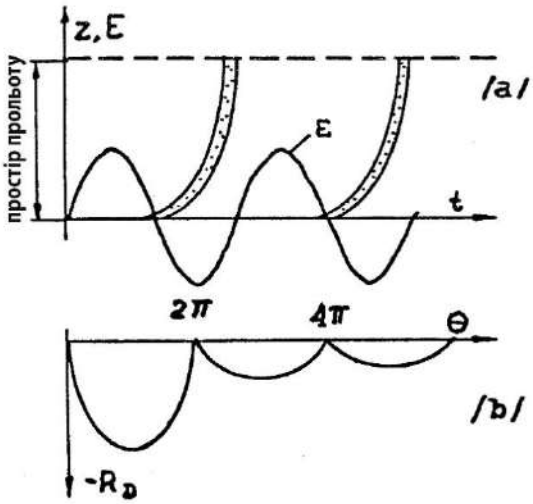


Рис. 2. Виникнення негативного опору в *p-n*-переході

В результаті наявності високочастотної напруги, підключеної до *p*-*n*-переходу, генеруються зарядові пакети шару множення, які негайно, якщо фазова затримка π /2, переходять в гальмівне високочастотне поле і пакети, що несуть заряд завдяки обміну енергією між високочастотним полем і пакетами носіїв заряду, опір діоду може вважатися негативним.

Звертає увагу, що для виникнення негативного опору, а отже і постійних незатухаючих високочастотних коливань необхідне певне фазове співвідношення між зовнішнім полем та процесами всередині діода.

Розподіл електричного поля для більш складної p^+-n-n^+ напівпровідникової структури представлено на рис. 3.

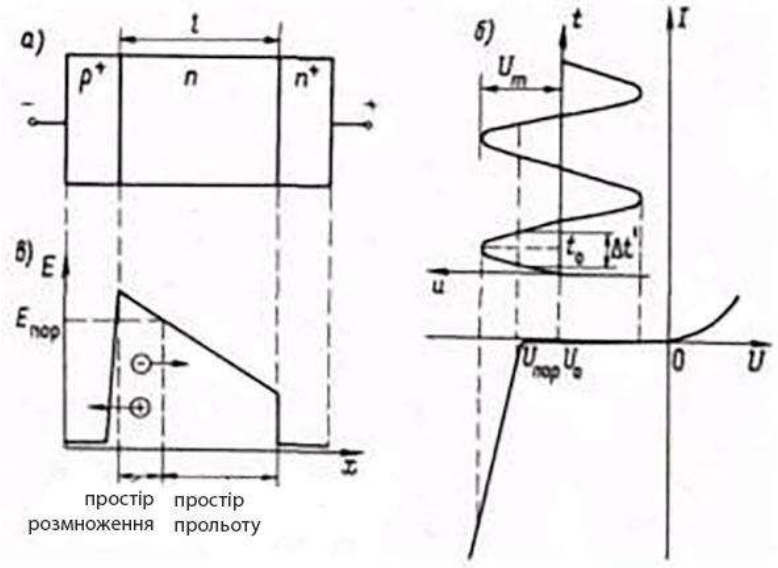


Рис. 3. Розподіл електричного поля ЛПД структури p^+-n-n^+

В результаті сумісної дії постійної та змінної напруги в області p^+-n-переходу, де сумарна напруженість електричного поля протягом певного часу перевищує порогове значення, виникає лавинний пробій, що приводить до лавиноподібного наростання числа носіїв заряду. Електрони, що дрейфують в пролітному просторі, утворюють імпульс конвекційного току, який в свою чергу взаємодіє з змінним електричним полем. Ця взаємодія супроводжується віддачею електронами своєї енергії змінному електричному полю, але тільки в тому випадку, якщо вони опиняються в гальмуючій фазі поля.

Для виготовлення ЛПД використовуються різні структури, такі, як чотиришарова структура Ріда, асиметричний різкий *p* -*n*-перехід , симетричний *p* -*n*-перехід, діод з двома дрейфовими областями ($p^+ - p - n - n+$), діод з двошаровою базою, діод з тришарової базою (модифікований діод Ріда). Кожна з наведених структур характеризується певною напругою пробою, розміром області лавинного множення і області дрейфу, ступенем впливу об'ємного заряду носіїв і температури, а також динамічними характеристиками асиметричного різкого *p*+ -*n*-переходу. Відзначимо, що в ЛПД розподіл концентрацій домішок в переходах має бути якомога ближче до ступінчатого. Найбільш важливим результатом фізичних процесів, що протікають в ЛПД, є наявність від'ємного опору. Комплексний опір ЛПД може бути досліджено експериментально. Цей опір, як показують результати експериментів, суттєво залежить від таких параметрів, як струм живлення, температура, частота.

На рис. 4 наведено одну з можливих залежностей комплексного опору лавино-пролітного діода з арсеніду галію від частоти та струму живлення при кімнатній температурі.

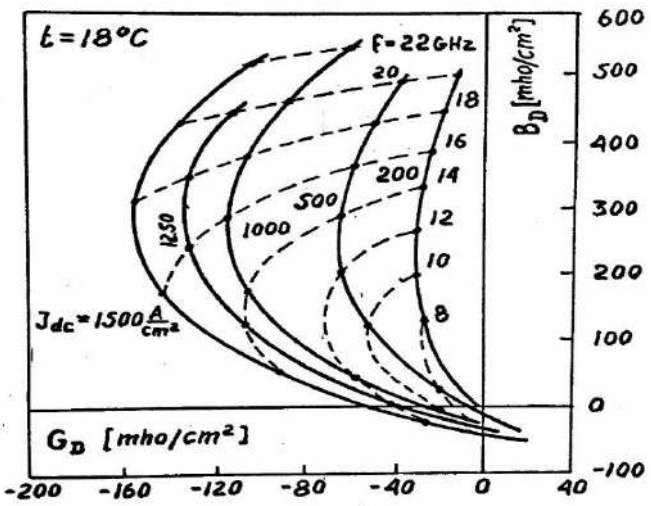


Рис. 4. Залежність комплексного опору ЛПД від частоти

Опис об'єкту дослідження

Для досліджень параметрів ГЛПД серед великого різноманіття різних конструктивних рішень була вибрана досить оригінальна конструкція генератора, що поєднує в собі як полозкові елементи, так і хвильоводний резонатор радіального типу. Конструктивною основою такого генератора є двосторонній фольгований діелектрик Polyguide.

Подальший експеримент показав, що використовувати відкриту конструкцію коливальної системи недоцільно, оскільки через низьку діелектричну проникність матеріалу Polyguide спостерігався високий рівень випромінювання в оточуюче середовище і, як наслідок, – низька добротність коливальної системи. Тому діод був по-суті розміщений в середині резонатора радіального типу, заповненого діелектриком.

Живлення на лавино-пролітний діод надходить через полозковий фільтр та відповідний отвір в радіальному резонаторі. Еквівалентна ємність, що присутня в резонаторі і роль якої виконує відповідний регулюючий гвинт, дозволяє проводити її налаштування в певному діапазоні частот. Для виведення мікрохвильової енергії використовується петля, розташована в резонаторі.

Регулюючи положення петлі, можна забезпечити узгодження між зовнішнім навантаженням і резонатором. Постійний струм живлення надходить на діод через фільтр низьких частот, виконаний за технологією друкованої плати. Для відводу теплової енергії передбачено відповідний радіатор з ребрами охолодження.

Розроблений генератор легко може бути поєднаний з іншими хвильоводними структурами, виконаними в полозковому варіанті (наприклад, циркуляторами).

Ескіз конструкції варіанту генератора на лавино-пролітному діоді, параметри якого були досліджені, наведено на рис. 5.

Як видно з рис. 5, основою генератора є двостороння металізована пластина Polyguide з змонтованим на ній алюмінієвим кільцем. За допомогою спеціального монтажного вузла лавино-пролітний діод розташовується в середині резонатора. З правого боку генератора знаходиться роз'єм постійного струму, а з лівого боку – роз'єм мікрохвильового виходу.

Механічне налаштування генератора відбувається за допомогою гвинта налаштування, розташованого навпроти лавинного діода. Виведення мікрохвильової енергії здійснюється за допомогою спеціальної петлі зв'язку. В такому ГДПД є можливість регулювати нахил петлі, тобто вихідну потужність генератора. В конструкції генератора, що розглядається, застосований лавино-пролітний діод типу Tesla VB 0234.

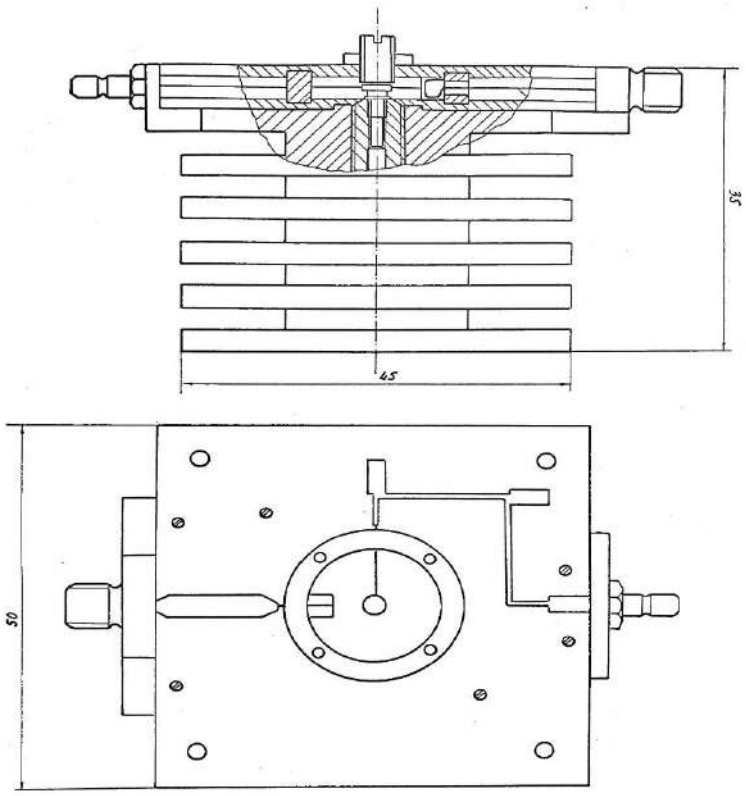


Рис. 5. Ескіз конструкції ГЛПД

Результати експериментальних досліджень

Основну увагу при дослідженнях ГЛПД приділено характеру залежності вихідної потужності від частоти налаштування при різних струмах живлення.

На рис. 6 наведено отриману експериментальним шляхом залежність вихідної потужності генератора від частоти при різних струмах живлення.

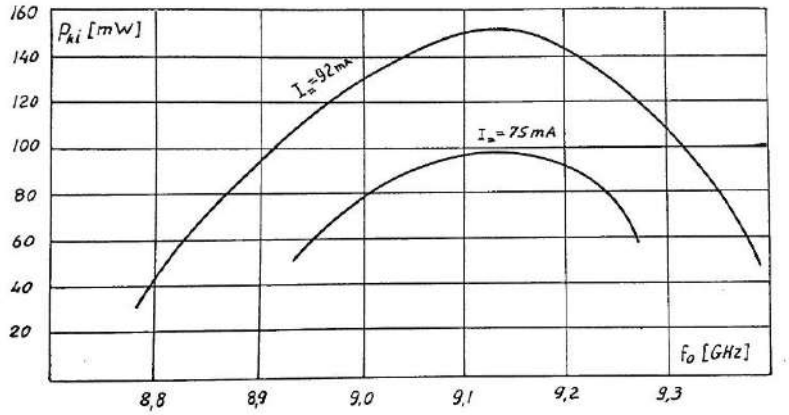


Рис. 6. Залежність вихідної потужності ГЛПД від частоти

Як видно з цього графіка, потужність струму значною мірою залежить як від постійного струму живлення, що протікає через лавинний діод, так і від частоти його налаштування. Звертає на себе увагу, що діапазон налаштування, а також вихідна потужність збільшується зі збільшенням постійного струму.

Як витікає з теорії лавинного ефекту, провідність лавинного діода є функцією постійного струму, що протікає через діод, тому природно, що зі зміною постійного струму змінюється як вихідна потужність, так і частота генератора. Результати дослідження вихідної частоти ГЛПД від струму живлення наведено на рис. 7.

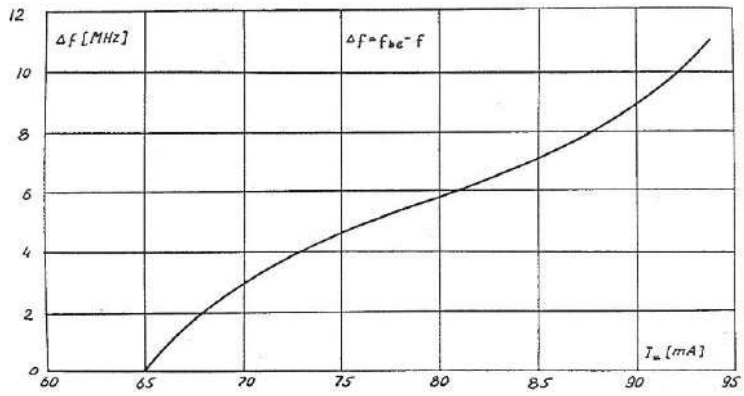


Рис. 7. Залежність вихідної частоти ГЛПД від струму живлення

Залежність, наведена на рис. 7, демонструє можливість електронного налаштування ГЛПД. Змінюючи постійний струм живлення, ми отримали можливість електронного налаштування частоти коливань в смузі близько 10 МГц. Додаткове налаштування генератора можливе також і механічним шляхом, за допомогою спеціального безлюфтового гвинта.

Цілком природно, що вихідні параметри ГЛПД залежать від добротності зовнішнього коливального контуру. Ця важлива характеристика генератора може бути визначена на основі вимірювання параметрів генератора як в холодному режимі, так і в режимі генерування. В цьому випадку коефіцієнт добротності генератора буде залежати від постійного струму, що протікає через діод. Результат таких досліджень наведено на рис. 8.

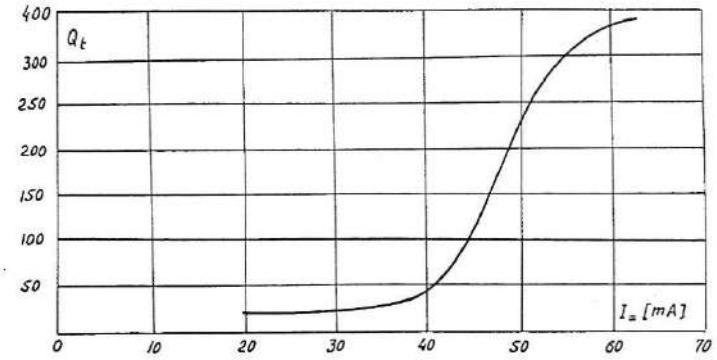


Рис. 8. Залежність коефіцієнта добротності від струму живлення

З цього графіка видно, що залежність навантаженої добротності Q_t від постійного струму слабшає поблизу струму збудження коливань (в нашому випадку струм збудження коливань становив I = 75 мА) і в подальшому залишається майже постійною.

Виходячи з цього, можна з високою достовірністю сказати, що добротність генератора, виміряна поблизу струму збудження, представляє загальну добротність генератора в робочому режимі.

Висновки

Таким чином, ГЛПД, що був досліджений, характеризується досить високими експлуатаційними параметрами, що дає можливість використовувати його у відповідних НВЧ пристроях. Отримані результати свідчать про його задовільну технологічність, високу працездатність та стабільну роботу.

Список литературы:

1. T. Misawa. Multiple uniform lagger approximationin analysis of negativ resistance in *p-n* junction in breakdown // IEEE Tran. ElectronDevices vol. ED-14 1967. Pp.795 – 808.

2. Сили И.И., Черенков А.Д. Параметри и стабильность частоти диодного генератрора с резонатором проходного типа // Енергосбережение Енергетика Енерграудит. 2015. №9. С. 53 – 59.

3. Сили И. И. Теоретический анализ процесса взаимодействия радиоимпульсов с колорадскими жуками в растительной среде картофеля // Технологический аудит и резервы производства. Харків, 2015. №4. С. 55 – 59.

4. Карушкин Н. Ф. Синхронизация генераторов на ЛПД импульсного и непрерывного действия в мм-диапазоне длин волн. Ч. 1. Конструкции генераторов и обобщенная модель их синхронизации внешним сигналом // Технология и конструирование в электронной аппаратуре. 2021 № 1–2 С. 10 – 20. http://dx.doi.org/10.15222/TKEA2021.1-2.10

5. Edgar Martinez: Next Generation of Terahertz Sources and Detectors. (PDF), S.4, abgerufen am 5. September 2021 (Vortragsfolie).

Надійшла до редколегії 11.02.2023

Відомості про авторів:

Меняйло Олександр Дмитрович – кандидат технічних наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри проєктування та експлуатації електронних апаратів, Україна; e-mail: oleksandr.meniailo@nure.ua; ORCID: https://orcid.org/0000-0002-3760-0523

Махонін Віктор Генадійович – Харківський національний університет радіоелектроніки, асистент кафедри проєктування та експлуатації електронних апаратів, Україна; e-mail: viktor.makhonin@nure.ua

Світличний Микита Сергійович – Харківський національний університет радіоелектроніки, студент ВСАм-22 кафедри проєктування та експлуатації електронних апаратів, Україна; e-mail: mykyta.svitlychnyi@nure.ua

RADAR AND RADIONAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

УДК 004.89: 621.396 **DOI:10.30837/rt.2023.1.212.14**

В.В. ЖИРНОВ, канд. техн. наук, С.В. СОЛОНСЬКА, канд. техн. наук

ІНТЕЛЕКТУАЛЬНА МОДЕЛЬ ЗОБРАЖЕНЬ ВІДМІТОК РАДІОЛОКАЦІЙНИХ ОБ'ЄКТІВ ДЛЯ ОГЛЯДОВИХ РЛС

Вступ

Наводяться результати розроблення інтелектуальної моделі зображень відміток радіолокаційних об'єктів для оглядових РЛС. Актуальність цих робіт полягає у створенні алгоритмів автоматичної обробки зображень радіолокаційних об'єктів для забезпечення ефективного автоматичного виявлення слабких корисних сигналів за рахунок накопичення сигнальної та логічної інформації в аналізованій комірці та в її околиці в складних завадових умовах. Удосконалення засобів забезпечення безпеки руху повітряного транспорту й автоматизація процесів управління його рухом вимагають ефективних процедур обробки сигнальної інформації. Актуальними є також питання більш повного використання та якісного підвищення інформаційних можливостей систем контролю, особливо в складних завадових умовах.

У відомих радіолокаційних інформаційних системах [1, 2] аналізуються моделі формування та динаміка змін зображень відміток об'єктів радіолокації, які формуються сукупністю відбитих від повітряних об'єктів радіолокаційних луна-сигналів. Ці зображення мають інформацію про повітряний об'єкт. Наприклад, якщо об'єкт точковий і рухомий, тоді формується зображення відмітки з пачки прийнятих елементарних відбитих сигналів у вигляді протяжної азимуту позначки, що має допплерівське зміщення несучої частоти РЛС. Відомі також операції в інтелектуальних інформаційних системах (ІІС), де використовуються алгоритми аналізу зображень відміток, в основі яких лежить модель дій людини-оператора та які пов'язані з можливістю паралельного сприйняття інформації з подальшим прийняттям рішення щодо аналізу відмінюючих ознак [2 – 4]. Важливу роль у розробці математичного забезпечення ІІС відіграють реляційні й логічні засоби представлення сигнальної інформації [4].

Існуючі підходи до моделей зображень відміток радіолокаційних об'єктів для оглядових РЛС не дозволяють створення на їх основі ефективних інформаційних систем для автоматичної їх обробки з метою виявлення та розпізнавання радіолокаційних об'єктів [1 – 4]. Перспективним вважається створення інтелектуальної моделі зображень відміток радіолокаційних об'єктів для оглядових РЛС, методів логічної обробки інтелектуальних зображень відміток і методу прийняття рішень, заснованого на локальному ознаковому описі [5 – 7]. В сучасних технологіях обробки сигналів та інформації недостатньо використовуються ті логічні операції, які подібні до семантики людини-експерта. Подібні операції широко використовують в технологіях інтелектуального аналізу процесів, таких, що, використовуючи дані про сигнальну обстановку, тобто, координати, форма, яскравість і передісторія, можуть отримувати корисну інформацію і передавати споживачеві.

Математичний опис інтелектуальної моделі зображень відміток радіолокаційних об'єктів для оглядових РЛС

Процес виявлення та розпізнавання повітряних об'єктів, що підлягає автоматизації, представлений як об'єднання двох основних завдань: розробка методики автоматичного конструювання інтелектуальних зображень радіолокаційних відміток та автоматичного конструювання текстів прийняття рішень про повітряні об'єкти, на основі аналізу смислових складових цих зображень. Алгоритм формування та логічної обробки інтелектуального зображення представлений у вигляді «вертикальної» структури у предикативній формі. Прискорення обчислень досягається тим, що групування доданків з однаковими множниками

у пірамідальному алгоритмі значно скорочує кількість операцій за рахунок виключення повторних процедур обчислень та операцій порівняння. При такому просторово-семантичному аналізі однакової обстановки радіолокації обсяг обробки зменшується в середньому у 2,6 рази зі збереженням необхідної достовірності. Радіолокаційна система шляхом огляду простору формує масив даних від кожної інформаційної комірки зони огляду, розміри якої визначаються роздільною здатністю або імпульсним об'ємом РЛС. З отриманих сигналів повітряних об'єктах формується карта чи матриця даних. Потім у результаті накопичення формується новий геометричний образ як віртуального зображення, тобто формується нова інтелектуальна модель зображення сигнальних позначок для точкових рухомих і малорухливих літальних апаратів типу літак, вертоліт, БПЛА (рис. 1).

Таким чином, база даних перетворюється на базу знань, в результаті аналізу якої визначаються предикатні ознаки, що містять інформацію про геометричний і смисловий образ символьних позначок. Зображення повітряних об'єктів на екрані індикатора РЛС формується як символи-образи. Формований масив даних є матрицею амплітуд $\|A\|$ розміром $M \times N$. Кожен елемент матриці i, j пов'язаний з відповідним інформаційним осередком зони огляду РЛС відповідно. Формування символьного масиву амплітуд $\|A\|$ здійснюється запам'ятовуванням величини амплітуди q_{ij} сигналу тривалість T огляду РЛС.

Ознаковий опис зображень радіолокаційних об'єктів

Процедури формалізації та аналізу геометричного та семантичного образу інтелектуальної моделі зображень відміток радіолокаційних об'єктів, що спостерігаються, здійснюються на основі алгебри предикатів [9 – 11]. При цьому визначається простір семантичних ознак, логічних зв'язків та залежностей на основі аналізу реальних процесів локації.

Ознакове виявлення та розпізнавання отриманих радіолокаційних зображень проводиться на логічному рівні обробки за допомогою алгебри кінцевих предикатів (АКП). АКП характеризується алфавітом подій А, що складається з k символів $\alpha_1, ..., \alpha_k$ і алфавітом змінних В, що складається з n символів $x_1, ..., x_n$, за допомогою яких може бути записаний будь-який n-місцевий k-й предикат, заданий над алфавітом подій А. У даній роботі для ознакового опису радіолокаційних зображень позначок повітряних об'єктів обрано n-місцевий двійковий предикат $f(x_1, ..., x_n)$. Інакше кажучи, символьна модель РЛ позначки точкового об'єкта – це пачка з n символів-предикатів $x_1, ..., x_n$ подій 0 чи 1 перевищення порога.

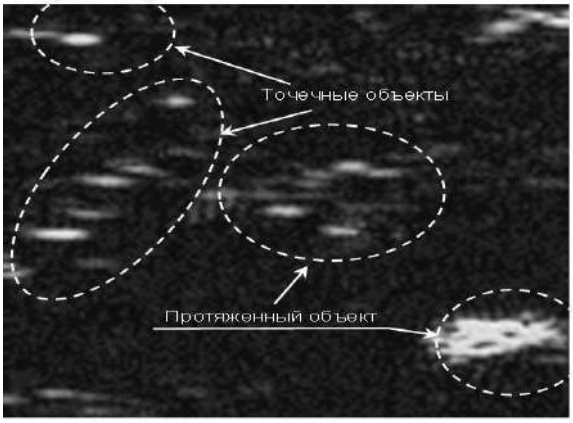
У ході досліджень використовувалися експериментальні дані (рис. 1), отримані під час запису відбитих сигналів оглядової РЛС сантиметрового діапазону (тривалість імпульсу 1 мкс, частота зондування 365 Гц). На рисунку наведено зображення розподілу амплітуд радіолокаційних позначок точкових та протяжних об'єктів у матриці розміром 129×129 реальних записів РЛС сигналів.

У розроблену інтелектуальну модель входять процедури формалізації та аналізу геометричного сигнального образу відміток радіолокаційних об'єктів на основі алгебри предикатів [7 – 10] та операції створення предикатної моделі семантичної складової для розпізнавання об'єктів локації, що спостерігаються.

Предикати $A(x)$ – це безліч подій сигналів з амплітудами q_{ij}, що перевищили поріг, з характеристикою $(t_{11}, t_{12}, \cdots, t_{ij}, \cdots, t_{mn})$, запишуться формулою

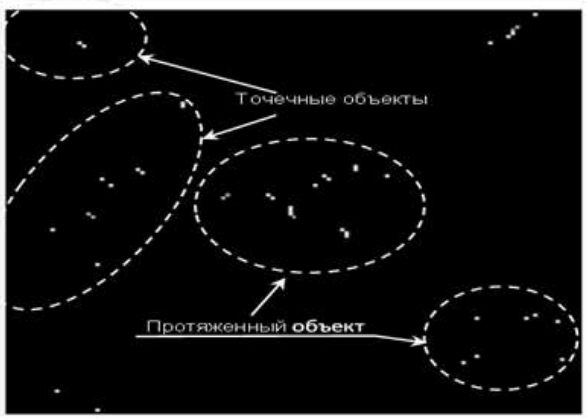
$$A(x) = t_{11} x^{q_{11}} \vee ... \vee t_{mn} x^{q_{mn}} = \bigvee_{i=1, j=1}^{mn} t_{ij} x^{q_{ij}}, \quad (1)$$

де $x^{q_{ij}}$ – форма впізнавання події, якщо $x = q_{ij}$, то $x^{q_{ij}} = 1$.



A

Рис. 1. Зображення реальних позначок об'єктів радіолокації



A5

Рис. 2. Інтелектуальні зображення радіолокаційних об'єктів

Пропонується система первинних семантичних ознак для опису смислової складової події в інформаційної комірці або між подіями у сусідніх інформаційних комірках та яка дозволяє формалізувати процеси формування інтелектуальної моделі зображення радіолокаційних позначок протягом кількох циклів зондувань РЛС. Для нашого випадку це:

– семантична ознака (унарний предикат) Z_{pij} наявності сигналу в a_{ij} інформаційному осередку (i, j – номери елементів зони огляду РЛС);

– семантичні ознаки (бінарні предикати) сусіднього осередку $Z_{d\,ij}$ та $Z_{a\,ij}$ переходу сигналу з поточного осередку a_{ij} до суміжного осередку за дальністю або азимутом.

Первинні семантичні ознаки формуються за таким правилом:

$$Z_{pij} = 1, \text{ at } A_{ij} > 0 \tag{2}$$

$$Z_{d\,ij} = 1, \text{ at } A_{i-1\,j} > 0 \wedge Z_{p\,ij} = 1 \tag{3}$$

$$Z_{a\,ij} = 1, \text{ at } Z_{p\,ij} = 1 \wedge A_{ij-1} > 0\,. \tag{4}$$

Запропонована система первинних семантичних ознак дозволяє описувати смислові залежності та зв'язки між інформаційними одиницями різного типу. Таким чином, маючи первинні семантичні ознаки, приступаємо до формування інтелектуальної моделі зображень сигнальних позначок для точкових рухомих та малорухомих літальних апаратів типу літак, вертоліт, БПЛА, а саме:

– формування простору семантичних ознак наявності сигналу (предикатів подій) $A = \{A_{11}, A_{12}, ..., A_{ij}, ..., A_{mn}\}$ у результаті аналізу амплітуд сигналів згідно з (1);

– формування простору первинних семантичних ознак Z_{pij}, $Z_{d\,ij}$, $Z_{a\,ij}$ згідно з (2) – (4), що дозволяє здійснювати ознаковий опис зображення позначки об'єкта $\left\{I\left(Z_{pij}, Z_{dij}, Z_{aij}\right)\right\}$ радіолокації на підставі семантичних зв'язків і відносин між подіями в інформаційних осередках в ході формування віртуального просторово-семантичного зображення радіолокаційних позначок;

– формування простору семантичних ознак зображень радіолокаційних позначок $\left\{I_k\left(Z_{pij}, Z_{dij}, Z_{aij}\right)\right\}$ для k об'єктів на основі аналізу залежностей первинних семантичних ознак;

– формування матриці ознакових описів інтелектуальних зображень відміток об'єктів радіолокації, створеної на безлічі $\left\{I_1, I_2, ..., I_k\right\}$ семантичних ознак зображень відміток на

основі геометричної $\left\{I_{g1}, I_{g2},...,I_{gk}\right\}$, смислової $\left\{I_{s1}, I_{s2},...,I_{sk}\right\}$ складових інтелектуальних зображень і семантичної ознаки флуктуації $\left\{I_{f1}, I_{f2},...,I_{fk}\right\}$ зображень відміток. Матриця розміру $k \times n$ (*k* рядків, *n* стовпців). Стовпці цієї матриці відповідають ознакам, наприклад, зображень відміток на основі геометричної $\left\{I_{g1}, I_{g2},...,I_{gk}\right\}$, смислової $\left\{I_{s1}, I_{s2},...,I_{sk}\right\}$ складових інтелектуальних зображень і семантичної ознаки флуктуації $\left\{I_{f1}, I_{f2},...,I_{fk}\right\}$ зображень відміток, а кожен рядок є ознаковим описом зображення відмітки одного об'єкта радіолокації. На рис. 3 наведено зразок матриці.

I_{g1}	I_{s1}	I_{f1}
I_{g2}	I_{s2}	I_{f2}
…	…	…
I_{gk}	I_{sk}	I_{fk}

Рис. 3. Матриця ознакових описів інтелектуальних зображень об'єктів радіолокації

– формування вектора прийнятих рішень про виявлення та розпізнавання зображень радіолокаційних позначок шляхом логічної обробки простору векторів семантичних ознак $W\left(I_g, I_s, I_f\right)$, яке задано на безлічі $\left\{I_1, I_2,...,I_k\right\}$ семантичних ознак зображень відміток на основі геометричної $\left\{I_{g1}, I_{g2},...,I_{gk}\right\}$, смислової $\left\{I_{s1}, I_{s2},...,I_{sk}\right\}$ складових інтелектуальних зображень та семантичної ознаки флуктуації зображень $\left\{I_{f1}, I_{f2},...,I_{fk}\right\}$.

Рис. 4. Пачка імпульсів, відбитих від літака

Розглянемо ці процеси докладніше. На рис. 4 наведена реальна, експериментально отримана пачка імпульсів, відбитих від літака. Тут дві інформаційні одиниці пов'язані ставленням «причина – наслідок»: ставленням появи сигналу в a_{ij} осередку (це семантична ознака Z_{pij} присутності сигналу); ставленням "сусідньої комірки" (це семантична ознака Z_{aij} переходу сигналу в суміжну за азимутом інформаційну комірку).

Наступним рівнем простору семантичних ознак радіолокаційних зображень об'єктів є семантичні ознаки зображень окремих відміток, що відображають відмітні ознаки процесу формування позначки I_{mk}, наприклад, точкового літального апарату як сукупності первинних семантичних ознак Z_{aij} сусідньої комірки за час *ln* наявності сигнальних імпульсів пачки. Тут семантична ознака радарної позначки дозволяє отримувати геометричну та смислову інформацію моделі зображення точкового рухомого об'єкта типу літак.

Формування ознакового простору інтелектуальних зображень об'єктів радіолокації

Складаємо семантичні рівняння для поряд розташованих елементів обробки (рис. 4) з умов отримання семантичних ознак Z_{dij}, Z_{aij} і шляхом їх вирішення для наступних дій щодо формування можливого типу інтелектуального зображення позначки, тобто номерів

$k = k_1$ і $l = l_1$ сусіднього елемента обробки за дальністю або азимутом. Визначаємо також, з якими із цих ознак працювати. Для цього при появі семантичного ознаку Z_{pij} наявності сигналу в a_{ij} інформаційному осередку складаємо семантичні рівняння для перевірки можливості формування семантичної ознаки $Z_{d\,ij}$ (приходу сигналу із сусідньої за дальністю $a_{i-1\,j}$ комірки) або семантичної ознаки $Z_{a\,ij}$ (переходу сигналу із суміжної за азимутом a_{ij-1} комірки) і отримаємо з умов (3) та (4):

$$\begin{matrix}(A_{i-1\,j} > 0 \wedge \mathrm{Z}_{p\,ij} = 1) = 1 \\ (\mathrm{Z}_{p\,ij} = 1 \wedge A_{ij-1} > 0) = 1\end{matrix}. \quad (5)$$

З аналізу варіантів розв'язків рівнянь (5) можна зробити такі висновки:

1. При виконанні 1-го рівняння формується семантична ознака $Z_{d\,ij}$, тобто сигнал переходить із сусідньої за дальністю $a_{i-1\,j}$ комірки, починає формуватися інтелектуальна модель зображення радарної позначки протяжного об'єкта або імпульсної перешкоди.

2. При виконанні 2-го рівняння формується семантична ознака Z_{aij}, тобто сигнал в досліджувану комірку переходить із сусідньої за азимутом комірки a_{ij-1} і починає формуватися інтелектуальна модель зображення радарної позначки літального апарата, тобто точкового об'єкта типу літак, вертоліт, БПЛА.

3. При виконанні одночасно 1-го та 2-го рівнянь формуються семантичні ознаки Z_{dij} і $Z_{a\,ij}$, тобто сигнал в досліджувану комірку переходить з сусідньої за дальністю $a_{i-1\,j}$ комірки, і з сусідньої за азимутом a_{ij-1} комірки. У цьому випадку продовжується формування інтелектуальної моделі зображення радіолокаційної позначки протяжного об'єкта.

У цій роботі детальніше вивчимо процес формування інтелектуальної моделі радіолокаційної позначки точкового літального апарату. Спочатку визначимо номер $l = l_1$ поруч розташованого елемента обробки з семантичною ознакою $Z_{a\,ij}$ сусіднього елемента обробки. Тут l_1 – номер початку пачки інтелектуальної моделі сигнальної позначки для точкового рухомого чи малорухливого об'єкта. Для першого кроку початку формування моделі приймемо l_1=0.

Виходячи з аналізу варіантів розв'язків рівнянь (6) та з урахуванням аналізу структурних елементів процесної моделі знань щодо виявлення та розпізнавання точкових рухомих об'єктів, на основі інтелектуального аналізу визначаємо черговість наступних процедур (кроків) обробки інтелектуальних зображень відміток об'єктів радіолокації.

На наступному кроці при складанні предикатних рівнянь для знаходження наступного номера $l = l_2$ елемента обробки з подібною семантичною ознакою враховуємо напрямок (a_{ij}, a_{ij+l_1}) формування інтелектуальної моделі пачечної структури сигнальної позначки точкового об'єкта, що позначився на першому кроці. При зміні номера координати l_1 напрямок пошуку збігається з напрямком азимутальної вісі координат j (вправо). Аналіз структурних елементів процесу міжперіодної обробки сигналів оглядових РЛС показує, що спочатку йде заповнення інформаційних комірок за дальністю i, а потім вже йде заповнення інформаційних комірок за азимутом j.

Таким чином, якщо є семантична ознака $Z_{a\,ij+l_1}$ сусідньої комірки за азимутом, то в наступному кроці обробки перевіряється наявність семантичної ознаки $Z_{a\,ij+l_2}$ в інформаційній комірці a_{ij+l_2}:

$$Z_{a\,ij+l_2} = (A_{ij} > 0 \wedge Z_{p\,i+l_2\,j} = 1) = 1 \tag{6}$$

Вирішуючи рівняння (6), знаходимо значення l_2. На виконання наступних операцій визначення номерів комірок уточнюємо напрям обчисленням градієнтів номерів за віссю i, тобто $\Delta l_2 = l_2 - l_1$.

При $\Delta l_2 = l_2 - l_1$ отримаємо

$$Z_{a\,ij+l_2+1} = (A_{ij+l_2} > 0 \wedge Z_{p\,ij+l_2+1} = 1) = 1. \tag{7}$$

На n-му кроці визначаємо $\Delta l_n = l_n - l_{n-1}$. Для цього кроку при $\Delta l_n = +1$:

$$Z_{a\,ij+l_n} = (A_{ij+l_{n-1}} > 0 \wedge Z_{p\,ij+l_n} = 1) = 1 \tag{8}$$

В результаті розв'язання системи *n* семантичних рівнянь (6) – (8) знаходимо всі значення номерів початку та кінця пачок $l_1 \cdots l_n$ та семантичні ознаки смислової I_S та геометричної I_g складових інтелектуального зображення позначки:

$$I_S = Z_{mij} = \bigwedge_{l_1}^{l_n} Z_{ai,j+l_n} = Z_{ai,j+l_1} \wedge Z_{ai,j+l_2} \wedge \cdots \wedge Z_{ai,j+l_{n-1}} \wedge Z_{ai,j+l_n} = 1. \tag{9}$$

З (9) отримуємо вигляд геометричної складової I_g семантичної ознаки інтелектуального зображення рухомого об'єкта на основі безлічі $\left\{Z_{ai,j+l_1}, Z_{ai,j+l_2}, \cdots, Z_{ai,j+l_{n-1}}, Z_{ai,j+l_n}\right\}$ первинних семантичних ознак відмітки з умови

$$I_g = 1 \text{ при } C1 \le (l_n - l_1) \le C2, \tag{10}$$

де $C1$ та $C2$ – мінімальна та максимальна тривалість пачок відміток.

Оскільки аналізований радіолокаційний об'єкт типу літальний апарат точковий і рухомий, то інтелектуальне зображення, що формується, має вигляд протяжної за азимутом позначки з амплітудними флуктуаціями пачки за рахунок доплерівського зміщення несучої частоти РЛС. З урахуванням амплітудних даних $q_{i,j}$ та використовуючи дані про форму згідно з (9), визначаємо ознаку флуктуації як допустиму різницю амплітуд відповідних комірок $\Delta g = q_{max} - q_{min} \ge G1$ амплітуд відповідних комірок:

$$I_f = 1 \text{ при } \Delta g = q_{max} - q_{min} \ge G1. \tag{11}$$

Для оцінки енергетичної ознаки I_e інтелектуальної моделі зображення пачки сигналів введено поняття накопиченої енергії [11] як суми амплітуд сигналів на їх предикати інформаційних комірок пачки у напрямку, що визначається вектором (l_n) розв'язків рівнянь (6) – (8). З урахуванням розподілу амплітуд $q_{i,j}$ у межах пачки визначаємо енергетичну ознаку пачки сигналів (позначок) рухомих повітряних об'єктів як сумарну амплітуду:

$$I_e = \sum_{l_1}^{l_n} q_{i,j+l_n} Z_{ai,j+l_n}. \tag{12}$$

Таким чином, шляхом аналізу рішень рівнянь (6) – (8), аналізу процедур отримання семантичних ознак смислової I_s (9), геометричної I_g (10), енергетичної I_e (12) складових інтелектуального зображення позначки та семантичної ознаки флуктуації I_f (11) можна скласти структуру та перелік процедурних та семантичних операцій формування та обробки інтелектуальних моделей зображень радіолокаційних об'єктів.

Види інтелектуальних зображень об'єктів радіолокації, отримані в результаті модельних експериментів, наведено на рис. 2.

Висновки

Отримано такі результати та їх новизна:

1. Нові інтелектуальні моделі зображень відміток, придатних для автоматичного прийняття рішення про виявлення та розпізнавання відміток радіолокаційних об’єктів.

2. Нові технології опису смислової залежності та зв'язки між інформаційними одиницями моделі зображень відміток для точкових рухомих та малорухомих літальних апаратів типу літак, вертоліт, БПЛА на основі нових первинних семантичних ознак.

3. Нові простори семантичних ознак на основі геометричної, процесної складових і флуктуацій відміток для опису зображень відміток об'єктів радіолокації.

Список литературы:

1. Radar Signal Processing and Its Applications / Jian Li, R. Hummel, P. Stoica, E. G. Zelnio. Springer, 2013. 279 p.
2. Skolnik M. I. (eds) (2021) Radar Handbook. McGraw-Hill, New York.
3. Russel S. Artificial intelligence. A modern approach, Second Edition / S. Russel, P. Norvig. Williams, 2006. 1410 p.
4. Бондаренко М. Ф. Теория интеллекта : учебник / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнаренко. Харьков : Изд-во СМИТ, 2007. 576 с.
5. Журавлев Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. 2005. Вып. 33. С. 5 – 68.
6. 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.
7. Жирнов В.В., Солонская С.В. Предикатная модель процессных знаний при обнаружении и распознавании пачечной структуры сигналов от летательных аппаратов в обзорных РЛС // Радиотехника. 2020. № 201. С. 137 – 144.
8. Jianping Ou, Jun Zhang and Ronghui Zhan. Processing Technology Based on Radar Signal Design and Classification // International Journal of Aerospace Engineering. Vol. 2020. Pp. 1 – 19. Article ID 4673763. https://doi.org/10.1155/2020/4673763.
9. Солонская С.В., Жирнов В.В. Предикатная модель процессных знаний при обнаружении и распознавании протяженных объектов типа облака, тучи, «ангел-эхо» в обзорных РЛС // Радиотехника. 2020. № 202. С 164 – 172.
10. Zhirnov V.V., Solonskaya S.V. Intelligent system for detection of low-visible air objects in surveillance radars // Telecommunications and Radio Engineering. 2020. Vol. 79, Iss. 17. P. 1513 – 1519. DOI: 10.1615/TelecomRadEng.v79.i17.20.
11. Advanced Methods and Deep Learning in Computer Vision.1st Edition / Ed.: E. R. Davies, Matthew Turk. Academic Press, 2021. Page Count: 586. ISBN: 9780128221099.

Надійшла до редколегії 07.02.2023

Відомості про авторів:

Жирнов Володимир Віталійович – канд. техн. наук, Харківський національний університет радіоелектроніки, п.н.с. НДЦ інтегрованих радіоелектронних систем і технологій, Україна; e-mail: nauka123@ukr.net; ORCID: http://orcid.org/0000-0002-2397-3126

Солонська Світлана Володимирівна – канд. техн. наук, НТУ "Харківський політехнічний інститут", Україна; e-mail: solonskaya@ukr.net, ORCID: https://orcid.org/0000-0002-8841-7825

УДК 621.396.96 **DOI: 10.30837/rt.2023.1.212.15**

І.В. СВИД, канд. техн. наук, С.В. СТАРОКОЖЕВ

РОЗПОДІЛЕНА ОБРОБКА РАДІОЛОКАЦІЙНОЇ ІНФОРМАЦІЇ СИСТЕМ СПОСТЕРЕЖЕННЯ ПОВІТРЯНОГО ПРОСТОРУ

Вступ

До основних джерел радіолокаційної інформації про повітряну обстановку у системі контролю повітряного простору відносяться первинні оглядові радіолокатори [1, 2], вторинні радіолокаційні системи [3 – 7] та системи ідентифікації за ознакою «свій-чужий» (Identification Friend or Foe (IFF)) [8 – 12]. В свою чергу первинні радіолокаційні системи в залежності від територіального розташування передавача та приймача діляться на однопозиційні та багатопозиційні [13 – 18]. При цьому слід зазначити, що аналіз інформаційної безпеки однопозиційних радіолокаторів [15, 16] показує їх вразливість як широкому спектру ненавмисних та навмисних завад, так і визначенні їх місця розташування. Це обумовлено простотою як виявлення випромінюючого передавача зондувального сигналу в однопозиційних радіолокаторах, так і оцінки його координат. Природно, це зумовило основний недолік однопозиційних радіолокаторів – низька стійкість до завад та живучість. Перехід до мережі радіолокаційних систем дозволяє значно послабити вплив навмисно спрямованих завад [17 – 19] на пункти прийому багатопозиційної мережі у зв'язку з неможливістю оцінити їх координати. Це дозволяє стверджувати, що мережі радіолокаційних систем мають більш високу стійкість до навмисних та внутрісистемних завад у порівнянні з однопозиційними радіолокаційними системами [20 – 22].

Для підвищення якості інформаційного забезпечення споживачів радіолокаційна інформація мережі радіолокаційних систем обробляється [23]. При цьому обробка радіолокаційної інформації може здійснюватися на сигнальному рівні [24 – 26] і на рівнях первинної [27] та вторинної [28] обробки радіолокаційної інформації.

Можна стверджувати, що у відомих роботах [29 – 32], зокрема, проведено систематичне введення в теорію, розробку та подано результати досліджень технології обробки інформації в радіолокаційних мережах систем спостереження повітряного простору. Розглянуто класичну теорію і методи обробки радіолокаційної інформації на наведених вище етапах обробки інформації радіолокаційних систем. Представлена технологія обробки радіолокаційної інформації доцільна як в управлінні повітряним рухом, так і в системі протиповітряної оборони. Названі системи описуються схожими алгоритмами обробки радіолокаційної інформації і, в цілому, мають загальні математичні основи.

Також у [33 – 36] розглянуто різні аспекти оптимальної обробки сигнальних даних та інформації. Показано, що підхід оптимізації обробки як сигналів, так і радіолокаційної інформації дозволяє істотно покращити характеристики в порівнянні з існуючим підходом до обробки радіолокаційних даних Слід зазначити, що деякі алгоритми оптимізації обробки радіолокаційних сигналів дозволяють прогнозувати продуктивність вже на етапі проєктування, а також служать для об'єднання інформації для багатоцільового відстеження з використанням розподіленої архітектури відстеження [37 – 41].

У роботах [24, 28] показано, що при виконуваних процедурах на етапах обробки інформації систем радіолокаційного спостереження оптимізація виявлення і вимірювання координат повітряних об'єктів можлива тільки при розподіленій обробці інформації у мережах спостереження. При цьому значення аналогового порогу для виявлення сигналу використовується і у якості параметру при спільній оптимізації обробки радіолокаційних інформації.

Метою роботи є аналіз якості об'єднання оцінок виявлення радіолокаційних сигналів та повітряних об’єктів при реалізації розподіленої обробки радіолокаційної інформації систем спостереження повітряного простору.

Оцінка якості виявлення повітряних об'єктів при розподіленій обробці радіолокаційної інформації

Як показано вище, основу інформаційного забезпечення споживачів становлять спільні або суміщені радіолокаційні системи спостереження. Однак така побудова радіолокаційної системи спостереження повітряного простору не завжди використовується для підвищення якості інформаційного забезпечення споживачів. Дійсно, виходячи з існуючих структур інформаційного забезпечення споживачів, вторинні радіолокаційні системи спостереження використовуються тільки для отримання бортової інформації ПО. Однак можна відзначити, що в окремих випадках, наприклад при вимірюванні дальності, до ПО використовують інформацію первинних РЛС, а азимута – інформацію запитальних радіолокаційних систем спостереження. Це підвищує якість інформаційного обслуговування споживачів.

В існуючих радіолокаційних систем спостереження повітряного простору споживачам видається оцінка вектору вимірювання $\hat{\vec{\alpha}}$, котра характеризується кореляційною матрицею помилок $\vec{C}^{-1}$, отриманою за результатами вимірювання координат ПО первинним радіолокатором. Інформація вторинних радіолокаційних систем спостереження використовується для отримання бортової інформації від ПО, яка також передається споживачеві. Слід зазначити, що для об'єднання оцінки вектору вимірювань ПО, отриманої первинними радіолокаційними системами спостереження, і польотної інформації, отриманої вторинними радіолокаційними системами спостереження, на запитальних радіолокаційних системах спостереження здійснюються всі ті процедури виявлення і вимірювання, що і на первинній радіолокаційний системі спостереження. Це може бути враховано при обробці радіолокаційної інформації, що призведе до підвищення якості радіолокаційної інформації, яка видається споживачам. Покажемо це.

Структуру первинної обробки радіолокаційної інформації в спільних та суміщених радіолокаційних системах спостереження можна представити у вигляді двоканальної структури, яка формує об'єднану радіолокаційну інформацію споживачам на основі вагового об'єднання результатів канального виявлення і оцінок векторів канальних вимірювань ПО з одночасним включенням до складу інформаційного блоку і польотної інформації, отриманої за каналами запитальних радіолокаційних систем спостереження.

Отримані в кожному з каналів оцінки векторів вимірювання спільно з кореляційними матрицями помилок вимірювання надходять на пристрій об'єднання оцінок. У пристрої об'єднання оцінок на основі аналізу надходять оцінки векторів вимірювання та кореляційних матриць помилок вимірювання, обчислюються оцінка результуючого вектору вимірювань і результуюча кореляційна матриця помилок. Результуючий вектор виміру $\hat{\vec{\alpha}}_p$ спільно з результуючою кореляційною матрицею похибок $\vec{C}_p^{-1}$ видаються споживачам.

Будемо враховувати, що одночасно виробляється оцінка вектору вимірювання $\vec{\alpha}$ в M каналах радіолокаційної обробки сигналів. Якщо задатися нормальним законом розподілу кожної з складових вектору $\vec{\alpha}$, а також врахувати некорельованість вимірювань в каналах обробки, то логарифм відношення правдоподібності з точністю до постійної величини можна записати як

$$\ln l = \sum_{k=1}^{M} \ln l_k = \sum_{k=1}^{M} \left[-\frac{1}{2} \left(\vec{\alpha}_k - \hat{\vec{\alpha}}_k \right)^T \vec{C}_k \left(\vec{\alpha}_k - \hat{\vec{\alpha}}_k \right) \right], \quad (1)$$

де $\hat{\vec{\alpha}}_k$ та $\vec{C}_k$ – оцінки та матриця точності оцінювання за сигналами кожної з радіолокаційних систем спостережень.

Легко показати, що вираз (1) з точністю до постійної величини можливо привести до вигляду

$$\ln l = -\frac{1}{2}\left(\vec{\alpha} - \hat{\vec{\alpha}}_p\right)^T \vec{C}_p \left(\vec{\alpha} - \hat{\vec{\alpha}}_p\right),$$

де $\hat{\vec{\alpha}}_p$ – результуюча оцінка, $\vec{C}_p$ – результуюча матриця точності, що визначаються з наступних виразів:

$$\hat{\vec{\alpha}}_p = \vec{C}_p^{-1} \sum_{k=1}^{M} \vec{C}_k \vec{\alpha}_k , \qquad \vec{C}_p = \sum_{k=1}^{M} \vec{C}_k . \tag{2}$$

Таким чином, на підставі виразу (2) можлива оцінка результуючого вектору вимірювання та результуючої матриці точності, а, отже, і результуючої кореляційної матриці похибок вимірювання при спільному використанні результатів вимірювання первинного та вторинного радіолокаторів.

В цьому випадку спостерігач має $\vec{R}$ матрицю сигнальної інформації $\vec{X} = \left\| x_{ij} \right\|$, де $\left\| x_{ij} \right\| = 1$, коли в елементі часового розділення $i = \overline{(1, M)}$, $j = \overline{(1, N)}$, яка відповідає просторовому дозволу, що розглядається, відбулося перевищення порога виявлення; коли ж не відбулося – то $\left\| x_{ij} \right\| = 0$.

Для рішення задачі виявлення необхідно отримати відношення правдоподібності та порівняти його з аналоговим порогом, обраним у відповідності до допустимої імовірності хибної тривоги виявлення повітряного об’єкта. Функції правдоподібності для гіпотез H_1 (наявності сигналу) та H_0 (відсутності сигналу) можна записати так:

$$L\left(x_i \middle| H_1\right) = \prod_{i=1}^{N} P_{sp}^{x_i}\left(x_i\right)\left[1 - P_{sp}\left(x_i\right)\right]^{1-x_i} , \tag{3}$$

$$L\left(x_i \middle| H_0\right) = \prod_{i=1}^{N} P_{p}^{x_i}\left(x_i\right)\left[1 - P_{p}\left(x_i\right)\right]^{1-x_i} , \tag{4}$$

де x_i – об’єднана послідовність нулів та одиниць з виходів приймальних пунктів мережі радіолокаційних систем.

Слід зазначити, що характерною особливістю вирішального пристрою виявлення ПО в спільній або суміщеній радіолокаційних системах спостереження є наявність двох порогів. Перший поріг встановлюється в порогових пристроях виявлювача сигналів кожного з каналів спільної радіолокаційної системи спостереження. Цей поріг аналоговий, і за допомогою тільки його можна змінювати умовну імовірність хибної тривоги на виході спільного виявлювача ПО. Другий поріг встановлюється в пороговому пристрої спільного виявлювача ПО і є порогом виявлення ПО. Він може бути тільки дискретним.

Проведемо оцінку характеристик виявлення ПО при спільному використанні сигналів первинного та вторинного каналів суміщеної радіолокаційної системи спостереження повітряного простору. Зазначимо, що об'єднання канальних рішень про виявлення повітряних об’єктів може здійснюватися на етапах:

- виявлення радіолокаційних сигналів;
- виявлення ПО.

Крім того, слід зауважити, що вибір вирішального правила при спільному виявленні сигналів суміщених радіолокаційних систем спостереження, як правило, повинен визначатися не тільки вимогами найкращого виявлення сигналів в таких системах. Дійсно, як зазначено вище, при виявленні ПО повинна бути проведена ідентифікація ПО. Це передбачає посилення вирішального правила, хоча при цьому результуючі характеристики виявлення можуть

погіршуватися. Жорсткість вирішального правила потрібна і при використанні сигналів спільних радіолокаційних систем спостереження для точного визначення координат ПО.

Проведемо порівняльний аналіз варіантів об'єднання рішень виявлення ПО.

Для першого варіанту бінарно-квантованої послідовності імпульсів з виходу детектора сигнали в кожному з каналів суміщених радіолокаційних систем спостереження надходять на виявлювач ПО. Результати канальних виявлень сигналів об'єднуються відповідно до правил «1 з 2» або «2 з 2» і далі об'єднана послідовність імпульсів надходить на виявлювач ПО. Завдання виявлювача ПО полягає в тому, щоб на основі аналізу надходження послідовності нулів та одиниць прийняти оптимальне рішення про наявність або відсутність ПО в прийнятій послідовності. Слід зазначити, що для вирішення завдання виявлення необхідно отримати відношення правдоподібності і порівняти його з порогом, обраним відповідно до допустимої імовірності хибної тривоги виявлення ПО.

Використовуючи вирази (3) та (4), відношення правдоподібності можна записати у вигляді

$$l(x_i)=\frac{L(x_i|H_1)}{L(x_i|H_0)}=\prod_{i=1}^{N}\left(\frac{P_{sp}(x_i)}{P_p(x_i)}\right)^{x_i}\left[\frac{1-P_{sp}(x_i)}{1-P_p(x_i)}\right]\geq l_0\,. \tag{5}$$

Здійснивши логарифмування виразу (5) та перетворивши отриманий вираз, отримуємо:

$$\sum_{i=1}^{N}x_i\eta_i\geq C\,, \tag{6}$$

де $$\eta_i=\ln\frac{p_{sp}(x_i)\left[1-p_{sp}(x_i)\right]}{p_p(x_i)\left[1-p_p(x_i)\right]},\quad C=\ln l_0-\sum_{i=1}^{N}\ln\frac{1-p_{sp}(x_i)}{1-p_{sp}(x_i)}\,.$$

Таким чином, алгоритм оптимального виявлення ПО (6) зводиться до підсумовування вагових коефіцієнтів η_i, що визначаються формами діаграм спрямованості антен відповідного каналу суміщених радіолокаційних систем спостереження, відповідних позиціям пачки, де $x_i=1$.

Як зазначено вище, у запитальних РСС у якості сигналів запиту і відповіді використовуються інтервально-часові коди (ІЧК). Оскільки для запитальних РСС характерно високе відношення сигнал-шум, то є можливість досягнення необхідних показників якості при обробці одиночних імпульсів ІЧК. Обробка прийнятих сигналів приймачем, при такій постановці розглянутого питання, полягає в декодуванні прийнятого сигналу й за його результатом – в прийнятті відповідного рішення.

Для підвищення імовірності прийнятого рішення при обробці кодованих сигналів, а також для захисту запитальних РСС від внутрішньо-системних завад у приймачі розглянутої системи можуть бути використані різні способи обробки сигналів, зокрема способи, що здійснюють міжперіодну обробку (МО) кодованих сигналів:

- декодування з попередньої МО сигналів;
- декодування з наступної МО сигналів.

У зв'язку із цим виникає інтерес до розгляду характеристик виявлення (ХВ) СВ при різних способах обробки, а також до оцінки впливу коефіцієнта готовності відповідача та імовірності подавлення сигналів у КВ оброблюваного пакета СВ на ХВ.

Отримаємо порівняльні ХВ для обох способів обробки СВ при дії в радіоканалі флуктуаційних завад. Розрахунки проведемо для критеріїв і особливостей побудови апаратури обробки ІЧК в існуючих радіолокаційних системах спостереження.

Припустимо, що коефіцієнт готовності літакового відповідача дорівнює одиниці і подавлення сигналів у радіоканалі відповіді відсутнє. На виході приймача сигналів здійснюється двійкове (бінарне) квантування сигналів, тобто при фіксованому співвідношенні

сигнал-завада q і обраному порозі обмеження знизу z_0 однозначно визначаються імовірності – P_{11} (імовірність виявлення одиночного імпульсу сигналу) і P_{01} (імовірність появи викиду шуму на заданій часовій позиції).

Припустимо також, що в дешифраторі здійснюється логіка n/n, а в пристрої МО сигналів застосовується логіка k/m, при якій фіксація виявлення сигналу відбувається за наявності будь-яких k імпульсів на m позиціях.

Порівняємо ХВ обох способів обробки з використанням критерію Неймана – Пірсона, тобто при фіксованому рівні хибних тривог знайдемо ХВ (імовірності виявлення кодованого сигналу) залежно від співвідношення сигнал-завада для моменту першого виявлення об'єкта (виконання критерію виявлення початку інформаційного пакета).

Для способу декодування з попередньої МО сигналів імовірність проходження кодованих сигналів і хибних тривог через пристрій МО сигналів можна записати як

$$D_1 = \sum_k^m C_m^k P_{11}^k (1-P_{11})^{m-k}; \;\; F_1 = \sum_k^m C_m^k P_{01}^k (1-P_{01})^{m-k} \; .$$

Імовірності P_{01} і P_{11} визначаються за заданими співвідношеннями:

$$P_{01} = e^{-z_0^2/2}; \;\; P_{11} = \int_{z_0}^{\infty} x e^{-(x^2+q^2)} I_0(qx) dx,$$

де $z_0 = z/\sigma$ – відношення поріг/шум; $I_0(qx)$ – модифікована функція Бесселя першого роду нульового порядку.

Імовірності проходження корисних та хибних сигналів через дешифратор можна визначити відповідно:

$$P_{1d} = \left[\sum_{i=k}^m C_k^m P_{11}^k (1-P_{11})^{m-k} \right]^n; \quad F_{1d} = \left[\sum_{i=k}^m C_k^m P_{01}^k (1-P_{01})^{m-k} \right]^n .$$

При декодуванні з наступною МО прийнятих сигналів імовірність проходження n імпульсних ІЧК і хибних сигналів через дешифратор

$$D_d = P_{11}^n; \quad F_d = P_{01}^n .$$

Імовірності виявлення корисних сигналів і хибних тривог на виході пристроїв МО обчислюються відповідно:

$$P_{2d} = \sum_{i=k}^m C_k^m D_d^k (1-D_d)^{m-k}; \quad F_{2d} = \sum_{i=k}^m C_k^m F_d^k (1-F_d)^{m-k} \; .$$

Наведені дослідження ХВ для двох способів обробки прийнятих сигналів при дії у каналі відповіді флуктуаційної завади не враховують імовірність подавлення СВ ХІЗ та реального коефіцієнта готовності відповідача.

Зазначимо, що наявність завад, як у каналі запиту, так і у каналі відповіді запитальних систем спостереження приводить до подавлення окремих імпульсів СВ, до утворення хибних СВ. У цьому випадку коефіцієнт готовності ЛВ, як правило, не дорівнює одиниці.

Отримаємо порівняльні характеристики виявлення для розглянутих способів обробки з урахуванням реального коефіцієнта готовності ЛВ й імовірності подавлення СВ. При цьому зробимо допущення, що імовірність подавлення СВ не впливає на утворення хибних тривог. У зв'язку з цим будемо визначати тільки імовірність виявлення кодованих сигналів.

Для першого способу декодування імовірність проходження кодованих сигналів через пристрій МО, з урахуванням коефіцієнта готовності ЛВ й імовірності подавлення СВ, можна визначити з співвідношення

$$D_1 = \sum_{i=0}^{m-k} C_m^i \left[P_0\left(1-P_p\right)\right]^{m-i}\left[1-P_0\left(1-P_p\right)\right]^i \sum_{j=0}^{m=k-i} C_{m-i}^j P_{11}^{m-j-i}\left(1-P_{11}\right)^j ,$$

де P_p – імовірність подавлення сигналів відповіді.

Імовірність виявлення кодованого сигналу на виході дешифратора в цьому випадку можна записати як

$$D_{1d} = \sum_{i=0}^{m-k} C_m^i \left[P_0\left(1-P_p\right)\right]^{m-i}\left[1-P_0\left(1-P_p\right)\right]^i \left[\sum_{j=0}^{m=k-i} C_{m-i}^j P_{11}^{m-j-i}\left(1-P_{11}\right)^j\right]^n . \qquad (7)$$

При декодуванні з наступною МО прийнятих сигналів імовірність проходження n-імпульсних ІЧК через дешифратор, з урахуванням впливу коефіцієнта готовності ЛВ та імовірності подавлення СВ, можна визначити як

$$D_d = P_0 P_p P_{11}^n .$$

Імовірність виявлення корисних сигналів на виході пристрою МО сигналів у цьому випадку

$$D_{2d} = \sum_{i=0}^{m-k} C_m^i \left(P_0 P_p P_{11}^n\right)^{m-i}\left[1-P_0 P_p P_{11}^n\right]^{m-i} . \qquad (8)$$

Вирази (7) і (8) отримано для загального випадку, коли P_p й P_{11} змінні. При $P_p = 1$ – це окремий випадок, коли враховуються тільки завади. При $P_{11} = 1$ – це інший окремий випадок, коли враховується тільки вплив імовірності подавлення СВ. У цьому випадку, як видно з (7) і 8), імовірність виявлення для обох способів обробки однакова.

Якщо припустити, що $P_{sp}(x_i)$ однакова в межах всієї ширини діаграми спрямованості антен суміщеної радіолокаційної системи спостереження (пачка прийнятих сигналів має прямокутну форму), то алгоритм (6) зводиться до виразу

$$\sum_{i=1}^{N} x_i \geq C_1 . \qquad (9)$$

Як випливає з (9), у разі прямокутної пачки процедура виявлення ПО зводиться до підрахунку одиниць в межах ширини пачки і порівняння числа накопичених імпульсів з пороговим числом C_1. Так як в цьому випадку схема виходить досить простою, а втрати в пороговому відношенні сигнал-шум незначні, то саме цей алгоритм широко використовується на практиці.

Аналіз ефективності алгоритмів інформаційного забезпечення проаналізуємо з урахуванням кінцевого результату, а саме – виявлення ПО. Використуємо правила виявлення ПО за пачкою двійково-квантованих сигналів, а також будемо розглядати випадок дешифрованих сигналів з виходів вторинних радіолокаційних систем спостереження. Будемо досліджувати два алгоритми об'єднання результатів виявлення:

- канальне накопичення і об'єднання результатів (НВ);
- об'єднання канальних рішень і накопичення (ВН).

Також проведемо порівняльний аналіз характеристик виявлення розглянутого і використовуваного на практиці виявлювачів ПО. Будемо розглядати випадок однакових значень

відносин сигнал-шум $q_i, i=\overline{(1,m)}$ для сигналів як первинного, так і вторинного каналів спільної радіолокаційних систем спостереження. У цьому випадку багатоканальне виявлення дає найбільший ефект. При такому розгляді питання у всіх каналах виявлення сигналів повинні бути однакові відносні пороги. Цим забезпечується однакова імовірність помилкової тривоги $F_i = F_0$. Однаковими будуть і імовірності виявлення $D_i = D_0, i=\overline{(1,m)}$. У цих умовах для незалежних флуктуацій амплітуд оптимальним вирішальним правилом для спільної обробки є правило «k з m». Вихідна імовірність хибної тривоги F визначається виявленням ПО. Переймаючись припустимою можливістю F, отримуємо для вибраного вирішального правила F_0 в кожному з каналів спільної радіолокаційної системи спостереження повітряного простору.

Таким чином, отримавши для заданої імовірності F і будь-якого вирішального правила імовірність F_0 в кожному приймальному каналі, можна обчислити вірогідність D_0 в кожному приймальному каналі, а потім обчислити вірогідність виявлення ПО D. При незалежних флуктуаціях сигналу в каналах прийому спільної радіолокаційної системи спостереження повітряного простору для кожного m існує оптимальне вирішальне правило. При малому числі каналів обробки, що нами розглядається, оптимальним є правило «1 з m». Однак в каналі ідентифікації повинне бути реалізовано тільки правило «m з m». Проробивши аналогічні операції з вихідними показниками якості виявлення сигналів, можемо отримати результуючі характеристики виявлення на виході виявлювача ПО. Другий поріг C_2 будемо вибирати, виходячи з половини пачки оброблюваних сигналів.

На рис. 1 – 3 наведено характеристики виявлення ПО для розглянутого та рекомендованого виявлювачів ПО при використанні двох каналів виявлення (первинний і вторинний).

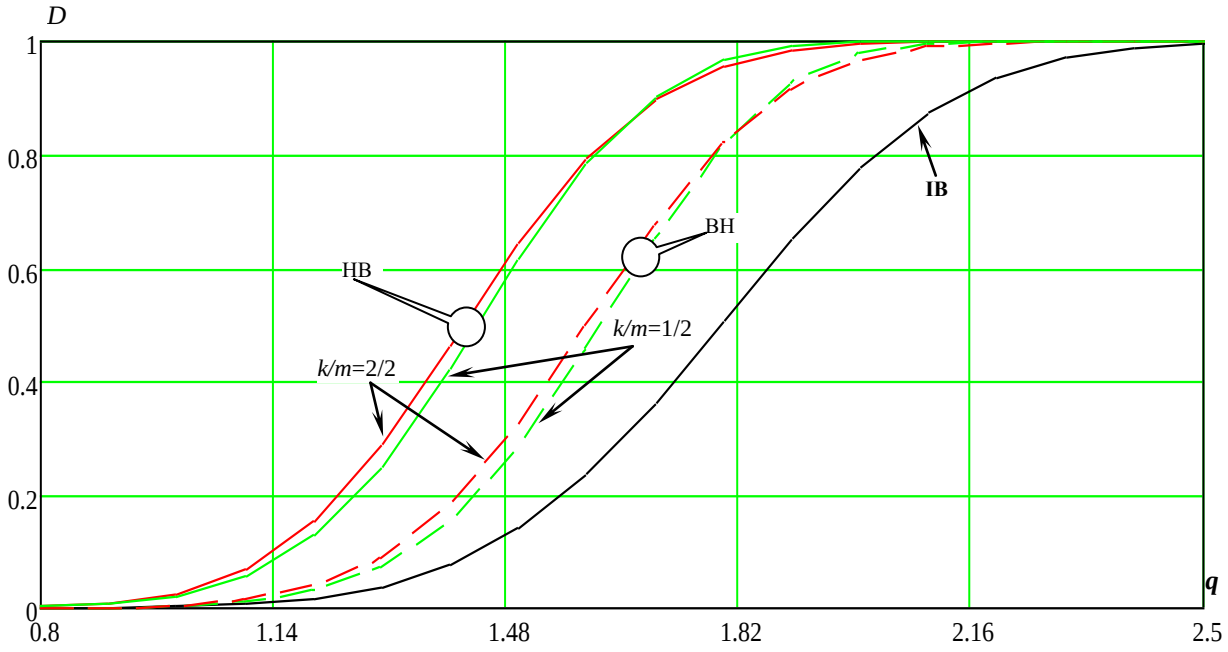


Рис. 1. Виявлення повітряних об'єктів при $P_0 = 1$

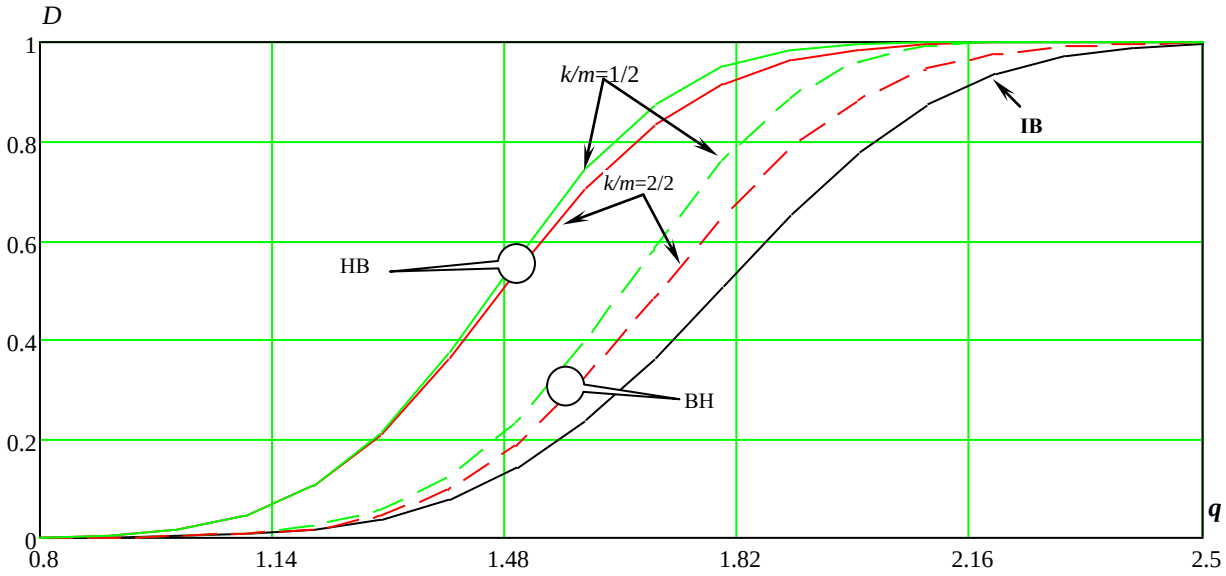


Рис. 2. Виявлення повітряних об’єктів при P_0 = 0,95

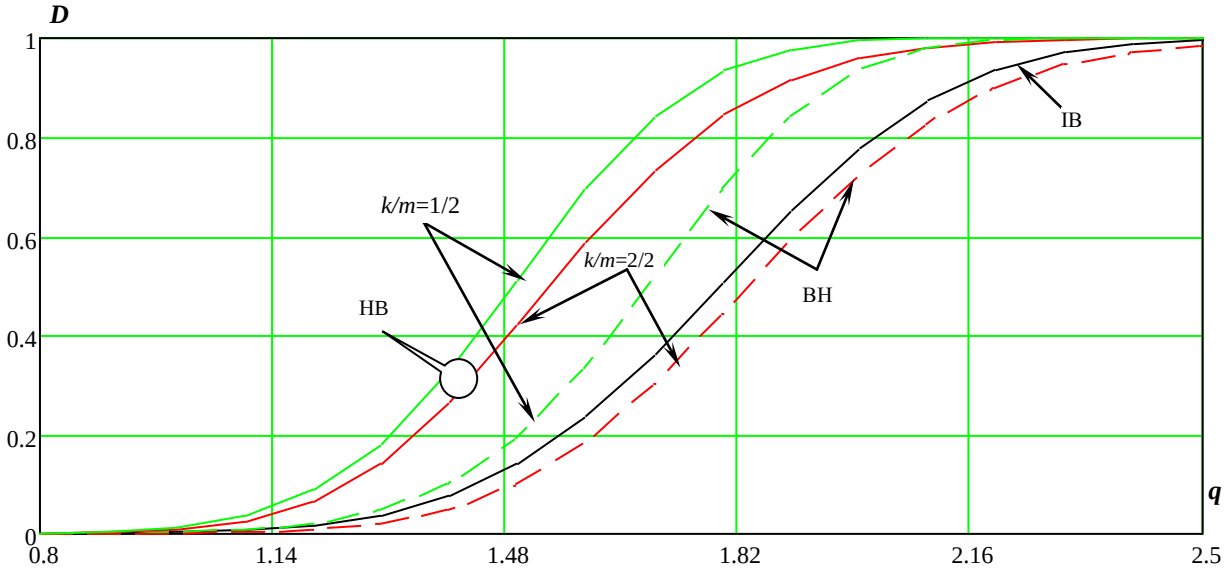


Рис. 3. Виявлення повітряних об’єктів при P_0 = 0,9

Як випливає з представлених залежностей, розглянута структура інформаційного забезпечення системи контролю повітряного простору має деякі переваги в порівнянні з тією, що використовується в даний час.

Так, для умови, що коефіцієнт готовності літакового відповідача дорівнює одиниці, що наведено на рис. 1, та відношення сигнал-шум, котре дорівнює 1,82, імовірність виявлення ПО для існуючого варіанту складає 0,53, а для варіантів виявлення-накопичення та накопичення-виявлення при вказаних критеріях виявлення складає 0,85 та 0,96 відповідно.

При зменшенні коефіцієнта готовності літакового відповідача до 0,95 (рис. 2) та відношенні сигнал-шум, що дорівнює 1,82, імовірність виявлення ПО для варіанта накопичення-виявлення та критерію k/m=1/2, складає 0,97, а для критерію k/m=2/2 – складає 0,95. Для варіанта виявлення-накопичення при критерію k/m=1/2 складає 0,8, а для критерію k/m=2/2 – складає 0,738.

При зменшенні коефіцієнта готовності літакового відповідача до 0,9 (рис. 3) та відношенні сигнал-шум, що дорівнює 1,82, імовірність виявлення ПО для варіанта накопичення-виявлення та критерію k/m=1/2 складає 0,95, а для критерію k/m=2/2 – складає 0,89. Для варіанта виявлення-накопичення при критерії k/m=1/2 складає 0,73, а для критерію k/m=2/2 – складає 0,49.

Порівняльний аналіз рис. 1 – 3 дозволяє зробити наступні висновки: якість інформаційного забезпечення споживачів на підставі запропонованої структури в порівнянні з використовуваною в даний час структурою обробки інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів, заснованого на накопиченні і з наступним об'єднанням; коефіцієнт готовності ЛВ істотним чином впливає на якість інформаційного забезпечення. Так, вже при $P_0 < 0{,}9$ використання цілочисельної логіки об'єднання інформації небажано.

Слід зазначити, що розглянуто випадок однакового відношення сигнал-шум в каналах обробки радіолокаційних систем спостереження. На практиці ж, відношення сигнал-шум вторинних каналів спільної радіолокаційної системи спостереження значно перевершує цей показник первинного каналу.

Висновки

Проведено аналіз ефективності алгоритмів інформаційного забезпечення на базі розподіленої обробки радіолокаційної інформації систем спостереження повітряного простору з урахуванням кінцевого результату, а саме – виявлення повітряних об’єктів за пачкою двійково-квантованих сигналів з урахуванням двох алгоритмів об'єднання результатів виявлення:

– канальне накопичення і об'єднання результатів;

– об'єднання канальних рішень і накопичення.

Це показує, що:

– якість інформаційного забезпечення споживачів на підставі запропонованої структури значно вище в порівнянні з використовуваною структурою обробки радіолокаційної інформації;

– якість інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів, заснованого на накопиченні сигналів з подальшим об'єднанням результатів виявлення;

– коефіцієнт готовності ЛВ істотним чином впливає на якість інформаційного забезпечення, вже при P0 < 0,9 використання цілочисельної логіки об'єднання інформації виявлення небажано.

Список літератури:

1. M. Skolnik. Improvements for air-surveillance radar // Proceedings of the 1999 IEEE Radar Conference. Radar into the Next Millennium (Cat. No.99CH36249), 1999, pp. 18 – 21. doi: 10.1109/NRC.1999.767195.

2. I. Svyd, I. Obod, O. Maltsev, V. Andrusevich, B. Bakumenko and O. Vorgul. Optimal Measurement of Signal Data Parameters of Requesting Radar Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2021, pp. 138 – 141. doi: 10.1109/UKRCON53503.2021.9575235.

3. F. L. Neindre, G. Ferre, D. Dallet, F. Letellier and K. Pitois. A Successive Interference Cancellation-based Receiver for Secondary Surveillance Radar // IEEE Transactions on Aerospace and Electronic Systems, 2022, doi: 10.1109/TAES.2022.3193649.

4. I. Obod, I. Svyd, O. Maltsev and S. Starokozhev. The Effect of Masking Interference on the Quality of Request Signal Detection in Aircraft Responders of the Identification Friend or Foe Systems // 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 2020, pp. 721 – 726. doi: 10.1109/PICST51311.2020.9467955.

5. M. Barbary, A. S. Hafez and T. Crew. An Industrial Design and Implementation Approach of Secondary Surveillance Radar System // 2021 International Telecommunications Conference (ITC-Egypt), 2021, pp. 1 – 9. doi: 10.1109/ITC-Egypt52936.2021.9513961.

6. I. Svyd, I. Obod, O. Maltsev and A. Hlushchenko. Secondary Surveillance Radar Response Channel Information Security Improvement Method // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), 2020, pp. 341 – 345. doi: 10.1109/DESSERT50317.2020.9125018.

7. M. Leonardi and D.D. Fausto. Secondary Surveillance Radar Transponders classification by RF fingerprinting // 2018 19th International Radar Symposium (IRS), 2018, pp. 1 – 10. doi: 10.23919/IRS.2018.8448244.

8. Свид І. В. Обробка радіолокаційної інформації систем спостереження повітряного простору : монографія. Дніпро : ЛІРА ЛТД, 2022. 224 с.

9. Y. Jiang, Z. Yang, C. Bo, and D. Zhang. Continuous IFF response signal recognition technology based on capsule network // Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 2021, pp. 455 – 468. doi: 10.1007/978-3-030-90196-7_39.

10. I. Svyd, I. Obod and O. Maltsev. Interference Immunity Assessment Identification Friend or Foe Systems // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 287 – 306, 2021. doi: 10.1007/978-3-030-71892-3_12.

11. T. M. Schuck, B. Shoemaker and J. Willey. Identification friend-or-foe (IFF) sensor uncertainties, ambiguities, deception and their application to the multi-source fusion process // Proceedings of the IEEE 2000 National Aerospace and Electronics Conference. NAECON 2000. Engineering Tomorrow (Cat. No.00CH37093), 2000, pp. 85 – 94. doi: 10.1109/NAECON.2000.894896.

12. V. Semenets, I. Svyd, I. Obod, O. Maltsev and M. Tkach. Quality Assessment of Measuring the Coordinates of Airborne Objects with a Secondary Surveillance Radar // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 105 – 125, 2021. doi: 10.1007/978-3-030-71892-3_5.

13. I. Ivashko, O. Krasnov and A. Yarovoy. Performance analysis of multisite radar systems // 2013 European Microwave Conference, 2013, pp. 1771-1774. doi: 10.23919/EuMC.2013.6687021.

14. Толюпа С.В., Дружинін В.А., Гордієвський О.Т. Розпізнавання групових об'єктів у багатопозиційних системах оперативного супроводження // Сучасний захист інформації. 2012. № 1. С. 66 – 70.

15. Обод І.І., Стрельницький О.О. Інформаційна безпека інформаційної мережі систем спостереження повітряного простору // Системи обробки інформації. 2015. № 9(134). С. 96 – 98.

16. Обод І.І., Стрельницький О.О. Захист інформації в мережі систем спостереження повітряного простору // Системи обробки інформації. 2016. № 2(139). С. 47 – 49.

17. J. Xu, X.-Z. Dai, X.-G. Xia, L.-B. Wang, J. Yu and Y.-N. Peng, Optimizations of Multisite Radar System with MIMO Radars for Target Detection // IEEE Transactions on Aerospace and Electronic Systems, vol. 47, no. 4, pp. 2329 – 2343, OCTOBER 2011. doi: 10.1109/TAES.2011.6034636.

18. Svyd I. Obod O. Maltsev O. Vorgul V. Chumak and B. Bakumenko. Estimation of the Spatial Coordinates of Air Objects in Synchronous Radar Networks for Airspace Observation // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), 2021, pp. 425 – 428. doi: 10.1109/PICST54195.2021.9772227.

19. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях радиолокации // Системи обробки інформації. 2006. № 9(58). С. 69 – 75.

20. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях // Системи обробки інформації. 2006. № 9(58). С. 69 – 71.

21. H. You, X. Jianjuan, G. Xin. Radar Data Processing with Applications. Publishing House of Electronics Industry, 2016. doi: 10.1002/9781118956878.

22. Chen Su, Chuanyun Zou, Liangyu Jiao, Qianglin Zhang. A MIMO Radar Signal Processing Algorithm for Identifying Chipless RFID Tags. Sensors (Basel). 2021 Dec 12;21(24):8314. doi: 10.3390/s21248314

23. Обод І.І., Стрельницький О.О., Андрусевич В.А. Методи підвищення якості інформаційного забезпечення системами спостереження повітряного простору // Системи обробки інформації. 2014. № 4(120). С. 53 – 55.

24. Обод І.І., Шевцова В.В. Порівняльний аналіз запитальних систем передачі інформації системи контролю повітряного простору // Зб. наук. пр. Харк. нац. ун-ту Повітряних Сил. 2013. № 1(34). С. 123 – 125.

25. І. Обод, І. Свид, О. Мальцев. Обробка даних радіолокаційних систем спостереження повітряного простору : навч. посібник. Харків : Друкарня Мадрид, 2021. 255 с.

26. J. Li, P. Stoica. MIMO Radar Signal Processing. Wiley-IEEE Press, 2008. 448 p.

27. S. M. Wu, G. A. Ybarra and W. E. Alexander. A complex optimal signal-processing algorithm for frequency-stepped CW data // IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, vol. 45, no. 6, pp. 754 – 757, June 1998. doi: 10.1109/82.686697.

28. Толюпа С.В., Дружинін В. А., Наконечний В.С., Цьопа Н.В., Батрак Є.О. Методи та алгоритми обробки радіолокаційної інформації у багатопозиційних системах зі змінною просторовою конфігурацією. Київ : Логос, 2014. 230 с.

29. Обод И.И. Обнаружение воздушных целей системой вторичной радиолокации // Радіоелектронні і комп'ютерні системи. 2005. № 3. С.25 – 28.

30. G. Lee, S. Lee, K. Kim and N. Kwak. Probabilistic Track Initiation Algorithm Using Radar Velocity Information in Heavy Clutter Environments // 2018 15th European Radar Conference (EuRAD), 2018, pp. 277 – 280. doi: 10.23919/EuRAD.2018.8546666.

31. Conte E., Daddio E., Farina A., and Longo M. Multistatic radar detection – Synthesis and comparison of optimum and suboptimum receivers // IEE Proceedings F: Communications Radar and Signal Processing, vol. 130, no. 6, pp. 484 – 494, 1983.

32. I. Prokopenko, V. Vovk and K. Prokopenko. Fast resource management algorithm for multi-position radar systems // 2015 16th International Radar Symposium (IRS), 2015, pp. 1045 – 1051. doi: 10.1109/IRS.2015.7226339.

33. V. Andrusevich and I. Obod. Assessment of the Quality of Information Support by Air Radar Surveillance Systems // Advanced Information Systems, vol. 5, no. 2, pp. 78 – 82, 2021. doi: 10.20998/2522-9052.2021.2.10.

35. I. Prokopenko, V. Vovk, S. Stavitsky and V. Medvedev. Optimization of use of resource in multi-position radar systems // 2014 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS), 2014, pp. 92 – 97. doi: 10.1109/MRRS.2014.6956673.

36. I. Obod, I. Svyd, O. Vorgul, O. Maltsev, O. Datsenko and N. Boiko Optimization of Data Processing Structure for Multi-Position Radar Surveillance Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON), 2021, pp. 133 – 137. doi: 10.1109/UKRCON53503.2021.9575286.

37. I. Svyd, I. Obod, O. Maltsev, O. Vorgul, I. Vorgul and I. Shevtsov Method for Increasing the Interference Immunity of the Channel for Measuring of the Short-Range Navigation Radio System // 2022 IEEE 16th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), Lviv-Slavske, Ukraine, 2022, pp. 802 – 807. doi: 10.1109/TCSET55632.2022.9767069.

38. I. Shevtsov et al. A Method for Increasing the Capacity of Radio Systems of Short-Range Navigation // 2022 IEEE 2nd Ukrainian Microwave Week (UkrMW), Ukraine, 2022, pp. 629 – 633. doi: 10.1109/UkrMW58013. 2022.10037138.

39. S. Starokozhev, M. Tkach, A. Hlushchenko, O. Datsenko, M. Chernyshov and V. Chumak. Frequency Efficiency Evaluation of Query Airspace Surveillance Systems // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 2021, pp. 501 – 505. doi: 10.1109/PICST 54195.2021.9772190.

40. S. Starokozhev, M. Tkach, A. Hlushchenko, O. Datsenko, M. Chernyshov and V. Chumak. Optimization of the Probability of Transmission of Flight Data in the Response Channel of Secondary Radar Systems // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 2021, pp. 511 – 515. doi: 10.1109/PICST54195.2021.9772199.

41. V. Semenets et al. Method of increasing the relative throughput of requesting radar systems // Przegląd Elektrotechniczny, vol. 1, no. 11, 2022, pp. 99 – 103. doi: 10.15199/48.2022.11.17.

Надійшла до редколегії 27.02.2023

Відомості про авторів:

Свид Ірина Вікторівна – кандидат технічних наук, доцент, завідувач кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: iryna.svyd@nure.ua; ORCID: http://orcid.org/0000-0002-4635-6542

Старокожев Святослав Валерійович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: sviatoslav.starokozhev@nure.ua; ORCID: https://orcid.org/0000-0002-1600-1337

УДК 621.396.96 **DOI:10.30837/rt.2023.1.212.16**

В.А. ТИХОНОВ, д-р физ.-мат. наук, О.В. КАРТАШОВ

СИНТЕЗ КОМПЛЕКСНОГО АЛГОРИТМУ ФУНКЦІОНУВАННЯ РАДІОАКУСТИЧНОГО ВИМІРЮВАЛЬНОГО КОМПЛЕКСУ

Вступ

Станції радіоакустичного зондування (РАЗ) атмосфери є перспективним засобом отримання інформації про висотний розподіл метеопараметрів в атмосфері Землі. Отримана інформація використовується в процесі вирішення актуальних науково-прикладних завдань: забезпечення зльоту, посадки та польотів літальних апаратів (як пілотованих, так і безпілотних), прогнозу погоди, прогнозування процесів поширення радіо-, акустичних та оптичних хвиль різних діапазонів [1, 2]. В даний час розвивається теорія радіоакустичного зондування атмосфери, створюються радіоакустичні станції різними науково-дослідними організаціями для виконання досліджень, станції РАЗ виробляються виробничими фірмами і пропонуються на ринок невеликими партіями [3].

Проте в цілому ефективність існуючих радіоакустичних засобів є недостатньою для вирішення актуальних прикладних завдань. Потреби практики формують необхідність суттєвого поліпшення основних тактико-технічних характеристик систем РАЗ. Тому існує необхідність у розвитку теорії систем радіоакустичного зондування атмосфери, методів проєктування радіоакустичних систем, а також у розробці відповідних перспективних підходів, структур та алгоритмів, що реалізуватимуться при побудові конкретних станцій, призначених для вирішення актуальних прикладних завдань [3, 4].

Синтез узагальненої структурної схеми комплексу

Проєктування складних інформаційно-керуючих систем поділяється на дві досить яскраво виражені стадії – системного проєктування, яке полягає в синтезі (розробці, виборі) та організації функцій структури системи в цілому, та технічного проєктування, що включає синтез, вибір та проєктування алгоритмів функціонування та технічної реалізації елементів системи [5]. На стадії системного проєктування основним предметом розгляду є структура (архітектура) майбутньої системи – фіксована сукупність елементів та зв'язків поміж них. Вивчення можливих варіантів структури дозволяє вирішити низку питань про вигляд проєктованої системи, абстрагуючись від конкретних елементів, на яких вона буде виконана [5 – 8].

При розробці систем радіоакустичного зондування атмосфери, як і розробки інших досить складних інформаційних систем, структурну схему доцільно отримувати евристичними методами, а проєктування елементів системи доцільно проводити за допомогою аналітичних методів синтезу і методів комп'ютерного моделювання. При синтезі структурної схеми радіоакустичного вимірювального комплексу (РАВК), що розробляється, далі використовується досвід проєктування подібних пристроїв, а також результати розгляду багатоканального радіоакустичного методу [5, 9, 10]. В результаті цього розгляду були визначені параметри сигналу (істотні параметри), що підлягають вимірюванню, розроблено методику відновлення характеристик атмосфери за отриманими значеннями параметрів сигналів, проаналізовано вплив різних факторів на точнісні показники, сформульовано найбільш загальні вимоги до вимірювального радіоакустичного комплексу. Все це дозволяє сформулювати вимоги до окремих елементів структури комплексу.

Розроблена узагальнена структурна схема радіоакустичного комплексу представлена на рис. 1. Вона містить такі елементи:

1. Передавальна радіо- і приймально-передавальна акустична антена.
2. Приймальна дискретна радіоантена.
3. Пристрій формування зондувальних акустичного та радіосигналів.

4. Пристрій перетворення і посилення радіосигналів, що приймаються.

5. Пристрій оптимального (квазіоптимального) виділення сигналів із перешкод, що виконує, наприклад, фільтрацію сигналів.

6. Пристрій виявлення та оцінки параметрів сигналів, виділених із перешкод.

7. Обчислювальний пристрій обробки інформації для розрахунку профілів метеовеличин за виміряними значеннями параметрів сигналів.

8. Пристрій керування РАВК, призначений для керування роботою всіх пристроїв (наприклад, з метою синхронізації), а також для адаптації до умов, що змінюються.

9. Приймач акустичного сигналу.

10. Пристрій обробки і визначення параметрів акустичного сигналу.

11. Пристрій збору апріорної інформації, яка служить для уточнення фізичних констант, що входять до розрахункових виразів, що використовуються при визначенні профілів метеопараметрів, та для визначення початкових значень параметрів зондувальних сигналів. Крім того, вимірювання вологості, температури, а також швидкості і напрямки вітру контактними метеодатчиками в приземному шарі атмосфери дозволяє здійснювати екстраполяцію значень швидкості звуку і швидкості вітру в першу висотну точку вимірювання методом РАЗ, забезпечуючи отримання при цьому досить вузького апріорного розподілу метеовеличин, що вимірюються, та сприяючи реалізації режиму захоплення у пристроях РАВК.

12. Пристрій відображення отриманої інформації та комунікаційні канали для зв'язку із споживачами.

Інформаційні зв'язки між елементами схеми на рис. 1 показані безперервною лінією, а зв'язки управління – пунктиром.

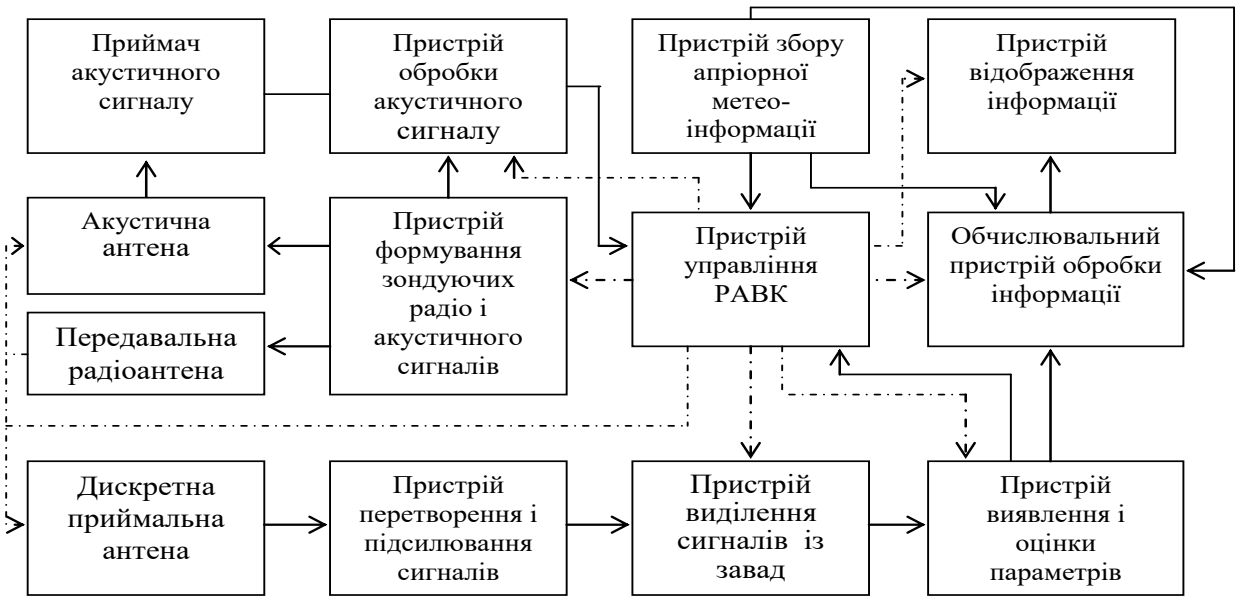


Рис. 1. Узагальнена структурна схема радіоакустичного вимірювального комплексу

Особливості запропонованої структурної схеми полягають, передусім, у тому, що вона представляє комплекс у досить узагальненому вигляді, відбиваючи ті основні функції, які мають виконуватися. Структурні схеми існуючих радарів, наприклад [3], є набором відомих блоків: вони розкривають структуру вже відомих, розроблених на інженерному рівні систем. Дана схема розкриває функціональну структуру перспективного комплексу, що синтезується, і є основою для його подальшого проєктування. Є й інші відмінності, але вони стануть очевиднішими в подальшому, при детальному розкритті елементів представленої структурної схеми.

Наступну розробку елементів структурної схеми РАВК та виконуваних системою функцій доцільно відповідно до системного підходу розбити на частини (системи), виконати

декомпозицію. Розбиття необхідно проводити таким чином, щоб кожна виділена частина мала власні показники якості, що однозначно пов'язані з показниками якості комплексу в цілому.

Процес розробки РАВК можна звести до розробки незалежних питань:

1. Синтез (вибір) типів зондувальних радіо- та акустичних сигналів, їх енергетичних параметрів, а також розробка відповідних пристроїв генерації (формування часової структури) та випромінювання (формування просторової структури) сигналів.

2. Синтез алгоритмів просторової та часової обробки сигналів для оптимального (квазіоптимального) виділення корисної інформації з сигналів, що приймаються на фоні шумів і перешкод, і розробка відповідних радіопристроїв та обчислювальних засобів для їх реалізації.

3. Розробка алгоритмів управління комплексом та адаптації комплексу до змінних зовнішніх умов, що формуються зовнішнім оточенням, а також розробка обчислювальних засобів, призначених для реалізації даних алгоритмів.

Зазначені частини можуть розроблятися окремо, але тільки в умовах їхньої постійної взаємодії та узгодження. Важливість, актуальність, взаємозв'язок і взаємозалежність, певна повнота переліку розв'язуваних завдань послужили основою декомпозиції їх із загального наукового завдання розвитку теорії та практики аналізованих систем.

Природно, що для створення ефективних радіоакустичних систем потрібна глибока теоретична розробка відповідних питань апаратурного характеру з використанням адекватних підходів, оскільки багато питань даного напряму не містяться в теорії радіолокаційних систем.

Теоретичне вивчення та дослідження властивостей радіоакустичних систем, а також синтез та розробка технічних рішень повинні мати комплексний системний характер, що враховує взаємний вплив різних підсистем, та виконуватися з використанням відповідних підходів та адекватного математичного апарату, що застосовуються в теорії систем, теорії стохастичного оптимального управління, теорії сигналів та ін.

У літературі сформувалася чітка думка [5, 13 – 18], що розробка сучасних ефективних радіосистем, що вирішують складні інформаційно-вимірювальні завдання та працюють в умовах різноманітної завадової обстановки, можлива лише на базі сучасних методів оптимізації. Причому системи певного, нехай досить вузького класу, що мають характерні особливості, повинні мати власну, адаптовану до наявної специфіки теорію, яка «обслуговує» цей напрямок.

Синтез (вибір) видів зондувальних сигналів

При проєктуванні комплексу РАЗ першим вирішується питання, яке пов'язане з вибором видів зондувальних сигналів та їх параметрів, а також методів їхньої генерації та випромінювання. До другого системного питання під час проєктування комплексу РАЗ можна приступати лише за наявності принципової визначеності у першому питанні.

Визначення типів зондувальних сигналів у загальному випадку має проводитися у межах завдання синтезу (оптимізації) сигналів. Питання, пов'язані з оптимізацією часової структури зондувальних сигналів інформаційних систем, розглянуті в [3 – 5], проте користуватися викладеною там методикою в практичному плані досить непросто, особливо стосовно радіоакустичного зондування, що має ряд істотних особливостей.

Вибір форми зондувальних радіо- та акустичного сигналів на практиці при побудові конкретної системи може здійснюватися шляхом аналізу відомих, що широко використовуються на практиці, поєднань сигналів. Такий аналіз, виконаний в [5], показує, що найбільш переважним є використання імпульсного акустичного сигналу з синусоїдальним заповненням і безперервного монохроматичного радіосигналу. Це пояснюється тим, що імпульсний акустичний сигнал створює локалізовану в просторі неоднорідність діелектричної проникності, що досить зручно при побудові системи та виконанні вимірювань, а безперервний

монохроматичний радіосигнал є найкращим при виконанні доплерівських вимірювань на поширених в радіолокації цілях. Однак застосування зазначеної комбінації сигналів у системах радіоакустичного зондування атмосфери не є оптимальним рішенням і призводить до появи певного роду специфічних похибок щодо характеристик атмосфери, природу яких вдалося розкрити тільки при виконанні теоретичних досліджень в даній галузі.

Таким чином, як випливає із загального розуміння принципів та особливостей функціонування радіоакустичних систем, основну увагу в процесі проєктування подібних систем слід приділяти вибору зондуючих сигналів, що використовуються. Процес вибору має бути заснований на обліку (аналізі) їх взаємодії із зовнішнім середовищем, взаємодії між собою (акустична – електромагнітна хвиля) та формуванням розсіяного сигналу. Саме на цьому етапі закладається корисна інформація в сигнал і зумовлюються багато характеристик станцій, і саме змістом цього етапу відрізняється в першу чергу теорія, що розробляється, від класичної теорії радіолокаційних сигналів, де етап формування розсіяного сигналу моделюється точковою метою.

У радіоакустичних системах використовуються зондувальні коливання різної фізичної природи – акустичні та електромагнітні, при цьому об'єкт, що розсіює, створюваний акустичним сигналом, не є точковим частотно-незалежним відбивачем і, отже, змінює при розсіюванні форму випромінюваних електромагнітних коливань. Відповідно до цього завдання аналізу зондувальних сигналів для радіоакустичних систем зондування атмосфери має полягати у спільному вивченні характеристик двох взаємозалежних видів сигналів – електромагнітного та акустичного.

Зондуючий сигнал радіоакустичних систем може бути представлений як векторний зондуючий сигнал, що складається з двох компонентів:

$$\vec{S} = \left| S_s,\ \mathrm{S}_e \right|,$$

де S_s – акустичний зондувальний сигнал; S_e – електромагнітний зондувальний сигнал.

Отже, методика аналізу та вибору зондувальних сигналів розглянутих систем повинна не тільки відповідати на питання, який сигнал є найкращим або найбільш прийнятним для певних фіксованих зовнішніх умов, але також і на питання, якими повинні бути сигнали за умов, що змінюються (в заданих межах), наприклад уздовж траси зондування.

Синтез алгоритмів обробки сигналів

Проєктування другої системи комплексу полягає у розробці алгоритмів, радіопристроїв та обчислювальних засобів для обробки отриманої з атмосфери інформації, починаючи від приймальних антен і закінчуючи розрахунком профілів метеовеличин. До основних питань даної частини відносяться синтез алгоритмів виявлення сигналів, що приймаються, а також алгоритм вимірювання координат центру плями і доплерівської частоти [4, 5].

Основними методами розробки алгоритмів виділення корисної інформації сигналів на фоні шумів є методи статистичного синтезу. Результати застосування цих методів залежать від типу сигналів, які на даному етапі проєктування визначені. Тому цю частину комплексу можна розробляти окремо.

Алгоритми і пристрої, що їх реалізують, призначені для обробки інформації, що отримується з атмосфери, в подальшому називатимемо системою обробки радіоакустичної інформації.

Причому слід зауважити, що методи часової та просторової обробки розсіяних сигналів систем радіоакустичного зондування атмосфери характеризуються рядом специфічних особливостей. Ці особливості зумовлені насамперед особливостями часової та просторової структури радіосигналу, розсіяного звуковою посилкою. Як випливає з результатів теоретичних досліджень і результатів численних експериментів, розсіяний на звуку радіосигнал має несиметричний спектр, а просторова структура розсіяного радіосигналу представляє локалізовану «пляму», яка до того ж переміщається поверхнею Землі.

Алгоритми управління комплексом

Управління комплексом полягає в такій зміні параметрів його пристроїв залежно від зміни зовнішніх умов, які забезпечують виконання розв'язуваної задачі з найкращими показниками. Основними показниками якості комплексу, як було показано раніше, є дальність, точність, просторове і часове розрізнення результатів вимірювань. У зв'язку з цим третє завдання, що полягає в проєктуванні системи управління, повинне включати синтез алгоритмів обробки інформації, що отримується з атмосфери, та алгоритмів формування керуючих впливів, які спрямовані на досягнення зазначених цілей управління, а також розробку відповідних технічних засобів [19 – 22].

Основні напрями впливу та прояви зовнішнього середовища, що впливають на роботу РАВК, – вплив атмосферних умов на звукову хвилю, яка поширюється в атмосфері, що полягає у зміні як часових, так і просторових параметрів акустичного зондувального сигналу, а також часової та просторової структури даного сигналу. Крім того, мають місце відбиття зондувальних сигналів від поверхні, що підстилає, місцевих предметів, пилу, гідрометеорів і т.д., значний вплив на роботу комплексу надають також природні і штучні перешкоди в акустичному і радіо каналах.

Таким чином, управління РАВК будемо розглядати як засіб досягнення цілей, що стоять перед РАВК, в системі вищого порядку, а з іншого боку як спосіб компенсації змін довкілля, які несприятливо впливають на комплекс і перешкоджають його ефективному (нормальному, оптимальному) функціонуванню.

Як відомо, система вважається керованою, якщо для неї виконуються необхідні умови спостережуваності та керованості [23, 24]. Перша умова полягає у забезпеченні можливості отримання інформації про поточний стан системи, а друга – у наявності каналів управління та керованих параметрів, що впливають на стан системи. Чим більше каналів управління, тим вище можливості з організації управління і тим вище його ефективність.

Під системою управління РАВК далі розумітимемо алгоритми обробки інформації та вироблення керуючих рішень, спрямованих на досягнення заданих цілей управління. Під терміном «управління» тут розуміється, з одного боку, позначення процесу досягнення поставленої мети, а з другого, – цілеспрямований вплив на структуру і параметри РАВК. На виконання цілей впливає значення вектору $\vec{X}(t)$ – вектору зовнішніх або вихідних параметрів РАВК, який визначає у загальному вигляді показник ефективності функціонування РАВК $E = E[\vec{X}(t)]$.

Об'єктом управління у розглянутій задачі, таким чином, є складові вектору внутрішніх параметрів РАВК $\vec{Y}(t)$, що визначають способи (алгоритми) використання його можливостей та витрачання його обмежених внутрішніх інформаційних, енергетичних та інших ресурсів. Для реалізації завдань управління комплексом необхідна наявність спеціальних інформаційних каналів управління, якими у межах комплексу передаються сформовані сигнали управління.

Показники якості інформації, що видається РАВК зовнішнім споживачам, залежать від вектору стану комплексу $\vec{Y}(t)$ (вектору внутрішніх параметрів), скоригованих системою управління, вектору стану зовнішнього середовища або вектору стану оточення комплексу $\vec{Z}(t)]$ та вироблених керуючих впливів $\vec{H}(t)]$:

$$\vec{X}(t) = F[t,\ \vec{\mathrm{Y}}(\mathrm{t}), \vec{\mathrm{Z}}(\mathrm{t}), \vec{\mathrm{H}}(\mathrm{t})],$$

де $F[\cdot]$ – оператор функціонування аналізованого комплексу як об'єкта управління.

Відповідно до розглянутого підходу мета управління полягає у досягненні таких значень вектору стану об'єкта $\vec{X}(t) = \vec{X}^*(t)$, які дозволяють забезпечувати максимум ефективності функціонування комплексу:

$$\max\ \mathrm{E}[\vec{X}(t)] = E[X^*(t)] = E^*(t).$$

Сигнали управління, вироблені системою управління та адаптації на підставі цілей управління $\vec{R}(t)$, інформації про стан зовнішнього середовища та стан вимірювального комплексу, можна подати таким чином:

$$\vec{U}(t) = f[t,\ \vec{\mathrm{Y}}_U(\mathrm{t}), \vec{\mathrm{Z}}(\mathrm{t}), \vec{\mathrm{R}}(\mathrm{t}), \vec{\mathrm{X}}(\mathrm{t})].$$

При вирішенні завдань управління в РАВК виникають певні труднощі, які обумовлені необхідними витратами обчислювального ресурсу для вирішення завдань у досить короткі проміжки часу, що визначаються динамікою зміни обстановки та функціонування РАВК. Швидкість динамічних процесів, що протікають у РАВК, або масштабування реального часу визначається, перш за все, швидкістю розповсюдження звуку в атмосфері. З іншого боку, труднощі процесів і завдань управління в РАВК пояснюються відсутністю в даний час методів формалізації та оптимізації завдань управління, які б отримували кількісні оцінки оптимальності управління в цілому.

З метою підвищення оперативності та зниження трудомісткості завдань управління комплексний алгоритм управління РАВК доцільно будувати за ієрархічним принципом. Ієрархічна структура управління будується шляхом виділення кількох підлеглих один одному рівнів управління. При цьому алгоритми управління вищих рівнів координують та впорядковують роботу підлеглих їм пристроїв. Інформація про стан атмосфери в систему управління надходить по приймальних акустичного та радіоканалів, а також через пристрій збору апріорної метаінформації. Ієрархічна система управління і адаптації РАВК представлена на рис. 2.

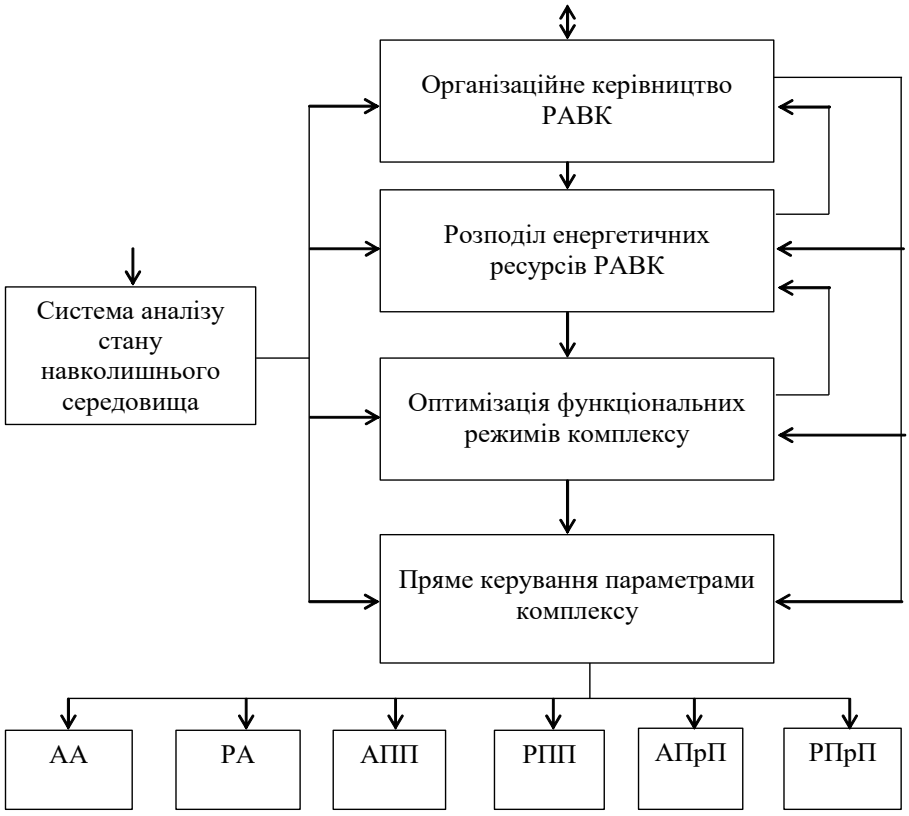


Рис. 2. Ієрархічна система управління і адаптації РАВК:
АА – акустична антена; РА – радіоантена; АПП – акустичний передавальний пристрій;
РПП – радіопередавальний пристрій; АПрП – акустичний приймальний пристрій;
РПрП – радіоприймальний пристрій

Відповідно до викладених загальних положень, РАВК як об'єкт управління може бути розділений на керовані елементи (пристрої), що відповідають основним пристроям радіоакустичного комплексу та системі обробки інформації. Основними об'єктами управління є такі елементи структури РАВК:

1. Передавальна радіо- та приймально-передавальна акустична антена: вибір використовуваної радіоантени (однієї з кількох просторово-рознесених антен), орієнтування антен у заданому напрямку.

2 Приймальна радіоантена: визначення просторового положення антени, вибір та підключення для аналізу відповідних приймальних елементів.

3. Пристрій адаптивного вибору видів зондуючих акустичного та радіосигналів та формування зондуючого векторного радіоакустичного сигналу:

- зміна несучої частоти акустичного сигналу;
- зміна несучої частоти радіосигналу.

4. Пристрій оптимального виділення інформативних сигналів на фоні перешкод:

- адаптивне формування опорних коливань у кореляційних системах обробки сигналів;
- управління центральною частотою налаштування вузькосмугових фільтрів;
- управління шириною смуги пропускання фільтрів.

Комплексний алгоритм функціонування радіоакустичного комплексу

При проєктуванні зазначених систем обробки інформації та управління користуватимемося широко застосовуваним в системотехніці методом декомпозиції – поділу систем на окремі, простіші підсистеми, що мають власні показники якості. Наприклад, завдання виявлення та оцінки параметрів вирішуватимемо окремо, використовуючи при цьому відповідні критерії, безпосередньо пов'язані з якістю інформації, що видається споживачеві [25, 26].

Алгоритми управління та обробки інформації з урахуванням зв'язків між ними утворюють комплексний алгоритм функціонування радіоакустичного комплексу.

Природно, що структура РАВК має певну стійкість до змін схеми побудови, цільового призначення, набору параметрів сигналу, що вимірюються, і т.д., тому ряд елементів структури може бути виконаний таким же чином, як і у відомих установках РАЗ, зокрема побудованих за основною схемою. До таких елементів відносяться передавальні радіо- та акустична антени, алгоритми і пристроїв оцінки доплерівської частоти, радіоприймач і т.д. При розробці елементів структури РАВК доцільно використати наявний досвід та застосувати апробовані технічні рішення.

Основну увагу в процесі проєктування слід приділяти тим елементам структури та алгоритмам їх роботи, які були відсутні у відомих установках РАЗ, або були, але по суті не виконували своїх функцій через низьку ефективність роботи.

Це, в першу чергу, алгоритми та пристрої обробки сигналу, що приймається, з метою оцінки координат центру плями та алгоритми, що забезпечують виконання умови Брегга при поширенні акустичного сигналу по трасі зондування. На етапі технічної реалізації розроблених алгоритмів постає питання про вибір відповідних технічних, зокрема обчислювальних засобів.

Питання прийому та обробки відбитого акустичного сигналу, необхідного для вимірювання поздовжньої складової швидкості вітру, у цій роботі не розглядається. Різні методичні питання та аспекти апаратурної реалізації подібних пристроїв описані у відповідній літературі.

Висновки

1. Розробку елементів структурної схеми РАВК та комплексного алгоритму його функціонування (виконуваних системою функцій) відповідно до системного підходу доцільно розбити на частини (системи), виконавши декомпозицію. Розбиття необхідно проводити таким чином, щоб кожна виділена частина мала власні показники якості, що однозначно пов'язані з показниками якості комплексу в цілому.

2. Процес розробки РАВК можна звести до розробки наступних незалежних питань:

- синтез (вибір) типів зондувальних радіо- та акустичних сигналів, їх енергетичних параметрів, а також розробка відповідних пристроїв генерації (формування часової структури) та випромінювання (формування просторової структури) сигналів;

- синтез алгоритмів просторової та часової обробки сигналів для оптимального (квазіоптимального) виділення корисної інформації з сигналів, що приймаються на фоні шумів і перешкод, і розробка відповідних радіопристроїв та обчислювальних засобів для їх реалізації;

- розробка алгоритмів управління комплексом і адаптації комплексу до зовнішніх умов, що змінюються, формуються зовнішнім оточенням, а також розробка обчислювальних засобів, призначених для реалізації даних алгоритмів.

3. Зондувальний сигнал радіоакустичних систем може бути представлений як векторний зондувальний сигнал, що складається з двох компонентів $\vec{S} = |S_s,\ S_e|$, де S_s – акустичний зондувальний сигнал; S_e – електромагнітний зондувальний сигнал. Відповідно до цього завдання синтезу (аналізу, вибору) зондувальних сигналів для радіоакустичних систем зондування атмосфери у процесі проєктування має полягати у спільному вивченні характеристик та виборі двох взаємозалежних видів сигналів – електромагнітного та акустичного.

4. Основними методами розробки алгоритмів виділення корисної інформації із сигналів на фоні шумів у комплексах РАЗ є методи статистичного синтезу. Результати застосування цих методів залежать від типів сигналів, які на даному етапі проєктування визначені. Тому цю частину комплексу можна розробляти окремо.

При цьому методи часової та просторової обробки розсіяних сигналів систем радіоакустичного зондування атмосфери характеризуються рядом специфічних особливостей, які обумовлені насамперед особливостями часової та просторової структури радіосигналу, розсіяного звуковою посилкою.

5. Третє завдання проєктування комплексу полягає в розробці системи управління, яке повинне включати синтез алгоритмів обробки інформації, що отримується з атмосфери, і алгоритмів формування керуючих впливів, які спрямовані на досягнення зазначених цілей управління. Далі розробляються відповідні технічні засоби, призначені для реалізації алгоритмів управління в реальному масштабі часу.

Список літератури:

1. Калистратова М.А., Кон А.И. Радиоакустическое зондирование атмосферы. Москва : Наука, 1985. 200 с.
2. Карташов В.М., Тихонов В.А., Олейников В.Н. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : ХНУРЭ, 2014. 312 с.
3. Карташов В.М. Модели и методы обработки сигналов систем радиоакустического и акустического зондирования атмосферы. Харьков : ХНУРЭ, 2011. 230 с.
4. 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.
5. Ситнік О.В., Карташов В.М. Радіотехнічні системи : навч. посібник. Харків : Сміт, 2009. 448 с.
6. Kartashov V.M., Oleynikov V.N, Sheyko S.A., Koryttsev I.V., Babkin S.I., Zubkov O.V. Peculiarities of small unmanned aerial vehicles detection and recognition // Telecommunications and Radio Engineering. 2019. Vol. 78, Iss. 9. pp. 771 – 781.
7. Oleynikov V. N., Zubkov O. V., Kartashov V. M., Korytsev I. V., Babkin S. I., Sheiko S. A. Investigation of detection and recognition efficiency of small unmanned aerial vehicles on their acoustic emission // Telecommunications and Radio Engineering. 2019. Vol. 78, Is. 9. pp. 759 – 770.
8. В.А. Тихонов, В.М. Карташов, В.М. Олейников, В.И. Леонидов, Л.П. Тимошенко, И.С. Селезнев, Н.В. Рыбников. Обнаружение-распознавание беспилотных летательных аппаратов с использованием составной модели авторегрессии их акустического излучения // Вісник НТУУ «КПІ». Радіотехніка. Радіоапаратобудування. 2020. Вип. №81. С. 38 – 46.
9. V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 – 25, 2020. Volume 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). pp. 35 – 39.

10. O. Sotnikov, V. Kartashov, O. Tymochko, O. Sergiyenko, V. 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. pp. 537 – 578.
11. Developing and Applying Optoelectronics in Machine Vision/ O. Sergiyenko, J.C. Rodriguez-Quiñonez. IGI Global, 2016. 341p.
12. Koryttsev S., Sheiko V., Kartashov O., 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.
13. 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.
14. 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. pp.1885 – 1892.
15. 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.
16. 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.
17. Vasilchenko A., Kartashov V. Analysis of influence exerted by longitudinal Doppler effect upon output signal of sodar antenna array // Telecommunications and Radio Engineering. Vol. 66, Iss.9. pp. 841 – 847. DOI: 10.1615/TelecomRadEng.v66.i9.50.
18. 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. pp. 461 – 468. DOI: 10.1615/TelecomRadEng.v77.i5.90.
19. Карташов В.М., Тихонов В.А., Воронин В.В. Особенности построения и применения комплексных систем дистанционного зондирования атмосферы // Радиотехника. 2016. Вып. 186. С. 184 – 185.
20. 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.
21. 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.
22. Красненко Н.П. Акустическое зондирование атмосферы. Новосибирск : Наука, 1986. 167 с.
23. 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.
24. 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.
25. Карташов В.М. Функции рассеяния сигналов систем зондирования атмосферы // Радиотехника. 2001. №118. С.61 – 65.
26. Карташов В.М., Пащенко С.В. Алгоритм формирования оценок максимального правдоподобия параметров радиосигнала, рассеянного акустическим волновым пакетом // Радиотехника. 2011. №164. С. 35 – 40.

Надійшла до редколегії 21.02.2023

Відомості про авторів:

Тихонов В'ячеслав Анатолійович – д-р ф.-м. наук, професор, Харківський національний університет радіоелектроніки, професор кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; email: vyacheslav.tykhonov@nure.ua, ORCID: https://orcid.org/0000-0002-4618-4787

Карташов Олександр Володимирович – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем; Україна; email: oleksandr.kartashov@nure.ua

УДК 621.396.96 **DOI: 10.30837/rt.2023.1.212.17**

І.В. СВИД, канд. техн. наук, М.Г. ТКАЧ

СИНТЕЗ І АНАЛІЗ ВИЯВЛЮВАЧА ТРАС ПОВІТРЯНИХ ОБ'ЄКТІВ ЗАПИТАЛЬНОЇ РАДІОЛОКАЦІЙНОЇ СИСТЕМИ

Вступ

Створення єдиного радіолокаційного простору неможливе без впровадження однієї інформаційної мережі на основі існуючих та перспективних радіолокаційних систем спостереження повітряного простору.

В існуючих інформаційних мережах радіолокаційних систем спостереження супровід повітряних об'єктів, як правило, здійснюється за інформацією первинних радіолокаційних систем спостереження [1 – 4], а вторинні радіолокаційних систем спостереження використовуються як джерела додаткової радіолокаційної інформації [5 – 8]. У той самий час, перехід на автоматичне залежне спостереження [9 – 12] передбачає обов'язкову наявність лише вторинних радіолокаційних систем спостереження. У зв'язку з цим актуальними є питання розробки методів та алгоритмів супроводу повітряних об'єктів за інформацією вторинних радіолокаційних систем спостереження. При цьому специфіка побудови та функціонування вторинних радіолокаційних систем спостереження суттєво відрізняється від первинних радіолокаційних систем спостереження.

Вказана специфіка вторинних радіолокаційних систем спостереження повітряного простору обумовлена [13 – 16]:

- реалізацією літакового відповідача за принципом відкритої системи масового обслуговування з відмовами;
- одноканальним принципом обслуговування сигналів запиту;
- використання специфічних сигналів (інтервально-часових кодів, позиційних кодів) у якості сигналів запиту та відповіді;
- несинхронним принципом побудови мережі запитальних радіолокаційних систем спостереження повітряного простору.

Наведені особливості реалізації запитальних радіолокаційних систем спостереження та наявність навмисних/ненавмисних (внутрісистемних) завад зумовили кінцеве значення коефіцієнта готовності відповідачів P_0 запитальних радіолокаційних систем спостереження. Таким чином, ця обставина повинна бути враховуватися при реалізації пристроїв супроводу повітряних об'єктів за інформацією запитальних радіолокаційних систем спостереження. У роботах [17 – 20] розглянуто синтез оптимального виявлювача трас повітряних об'єктів в єдиній постановці питання виявлення в інформаційних системах: виявлювач сигналів запиту, виявлювач сигналів відповіді, виявлювач повітряного об'єкта і власне виявлювач траєкторії.

Однак послідовність виконуваних процедур виявлення дозволяє реалізувати виявлювачі трас повітряних об'єктів з проміжними прийняттями рішень про виявлення сигналів відповіді, виявлення повітряних об'єктів та виявлення траси повітряних об'єктів.

При цьому слід зазначити, що зміна структури інформаційного забезпечення системи контролю повітряного простору, яка обумовлена переходом до автоматичного залежного спостереження [21 – 24] дещо змінює підхід до інформаційного забезпечення споживачів. Так, перехід вторинних радіолокаційних систем спостереження до основних джерел радіолокаційної інформації ставить задачу реалізації супроводу повітряних об'єктів за інформацією цих радіолокаційних систем спостереження. Необхідно також відзначити, що раніше функція супроводу повітряних об'єктів вирішувалася тільки на основі радіолокаційної інформації первинних радіолокаційних систем спостереження [25 – 27]. Якщо теорія і практика побудови фільтрів супроводу повітряних об'єктів за інформацією первинних радіолокаційних систем спостереження досить докладно розглянута в існуючій технічній літературі

[1, 28 – 31], то розгляд указаних питань для вторинних радіолокаційних систем спостереження має деякі прогалини [32 – 35]. Дійсно, специфіка побудови вторинних радіолокаційних систем спостереження зумовила наявність деяких параметрів, які відсутні в первинних радіолокаційних систем спостереження [36 – 39].

При цьому слід зазначити, що вторинна та третина обробка радіолокаційної інформації в системі контролю використання повітряного простору, як правило, завжди виконувалася з використанням інформаційних технологій [1 – 4]. Переваги широкого використання цифрової обробки сигналів дозволили застосувати інформаційні технології обробки радіолокаційної інформації з етапу первинної обробки радіолокаційної інформації. Це дозволило підвищити ефективність інформаційного забезпечення споживачів та здійснити сумісну оптимізацію обробки сигналів первинної і вторинної обробки радіолокаційної інформації.

Побудова існуючих вторинних радіолокаційних систем спостереження за принципом несинхронної мережі, обслуговування першого вірно прийнятого сигналу запиту та відкритої одноканальної системи масового обслуговування з відмовами [4, 40 – 42] негативно впливає на якість обробки інформації вторинних запитальних систем. Така побудова останніх відкриває широкі можливості з несанкціонованого використання відповідачів інформаційних систем, що розглядаються, а також для повної паралізації літакових відповідачів шляхом постановки корельованих завад необхідної інтенсивності. Так, робота літакового відповідача у полі дії багатьох вторинних радіолокаційних систем спостереження, що створюють внутрісистемні завади, призводить до того, що коефіцієнт готовності відповідача P_0 завжди менше одиниці. Коефіцієнт готовності відповідача залежить від інтенсивності потоку сигналів запиту, утвореного потоком сигналів запиту від запитальних радіолокаційних систем спостереження, потоком навмисних корельованих завад, а також потоком сигналів запиту, що утворився з потоку навмисних і ненавмисних некорельованих завад [43 – 46].

Синтезуємо структуру виявлювача траси повітряних об’єктів за інформацією запитальних радіолокаційних систем спостереження єдиної інформаційної мережі з проміжними прийняттями рішень щодо виявлення сигналів, повітряних об’єктів і трас.

Синтез виявлювача траси повітряних об’єктів запитальних радіолокаційних систем

Синтезуємо виявлювач, в якому рішення щодо виявлення траси повітряних об’єктів приймається на основі перевищення суми одиничних рішень про виявлення сигналів відповіді запитальних радіолокаційних систем спостереження x_{ijk}, де $i=\overline{1...n}$, $j=\overline{1...N}$, $k=\overline{1...K}$, де n – значність інтервально-часового коду сигналів відповіді, N – число сигналів відповіді у пачці сигналів, що приймаються, K – число відміток повітряних об’єктів, за якими приймається рішення про виявлення траси з вагами, що визначаються показниками якості виявленого імпульсів запиту, порогового рівня, обчисленого на основі певного критерію. При цьому слід зазначити, що наявність єдиного порогового пристрою прийняття рішення про виявлення траси повітряних об’єктів дещо ускладнює реалізацію синтезованої структури виявлювача траси повітряних об’єктів.

Легко побачити, що викладена процедура виявлення траси повітряних об’єктів складається з трьох процедур виявлення, кожна з яких ідентична за підходом до синтезу оптимального виявлювача трас повітряних об’єктів. Розглянемо її.

Після прийняття рішення про виявлення сигналів відповіді на подальшу обробку надходить реалізація $x_l=1$, якщо в елементі часового дозволу $l=\overline{1,L}, L=n(N)(K)$, що відповідає аналізованому просторовому дозволу, відбулося перевищення порога; якщо ж не сталося – $x_l=0$.

Для прийняття рішення про наявність або відсутність сигналу при сумісній обробці піддається обробці сукупність нулів і одиниць x_l. Вочевидь, що x_l – випадкова величина, яка підпорядковується розподілу Бернуллі:

$$P(x_l) = P_l^{x_l} (1 - P_l)^{1-x_l}, \tag{1}$$

де P_l – імовірність перевищення порога в l-м каналі обробки. При відсутності сигналу $P_l = F_l$ – імовірність хибної тривоги, а при дії сигналу $P_l = P_0 D_l$ – імовірність виявлення. У подальшому будемо вважати, що коефіцієнт готовності літакового відповідача входить у імовірність виявлення сигналів запиту.

Завдання виявлення сигналів відповіді можна розглядати у різних постановках. Дійсно, в аналізованому виявлювачі можливе керування напругою порога спрацювання вихідного порогового пристрою, а також напругою порога вхідного порогового пристрою. Розглянемо показники виявлювача під час управління величиною порога лише у вихідному пороговому пристрою, тобто у значно вужчій постановці питання.

Будемо вважати, що на вхід пристрою спільної обробки сигналів надходить сукупність випадкових величин x_l. Імовірності всіх можливих комбінацій x_l, як при відсутності, так і за наявності сигналу (гіпотези H_0 і H_1), тобто, $P(x_{ij}|H_0)$ та $P(x_{ij}|H_1)$, довільні, однак відомі. Для кожної конкретної сукупності x_l можливо сформувати відношення правдоподібності:

$$\Lambda = P(x_l|H_1) / P(x_l|H_0). \tag{2}$$

Порівняння відношення правдоподібності Λ з порогом, визначеним за допустимою імовірністю хибної тривоги, забезпечує оптимальне за критерієм Неймана – Пірсона рішення про наявність або відсутність сигналу. Так як шуми у каналах обробки незалежні, то можна записати співвідношення

$$P(x_1,..., x_L|H_0) = \prod_{l=1}^{L} P(x_l|H_0) = \prod_{l=1}^{L} F_l^{x_l} (1 - F_l)^{1-x_l}. \tag{3}$$

Можна стверджувати, що при впливі сигналу перевищення порогів у каналах обробки сигналів – незалежні події. В цьому разі

$$P(x_1,..., x_L|H_0) = \prod_{l=1}^{L} P(x_l|H_1) = \prod_{l=1}^{L} D_l^{x_l} (1 - D_l)^{1-x_l}. \tag{4}$$

З урахуванням виразів (3) та (4) загальний вираз відношення правдоподібності можна записати у наступному вигляді:

$$\Lambda = \prod_{l=1}^{L} D_l^{x_l} (1 - D_l)^{1-x_l} \Big/ \prod_{l=1}^{L} F_l^{x_l} (1 - F_l)^{1-x_l}. \tag{5}$$

Здійснимо логарифмування виразу (5), отримаємо вираз

$$\ln \Lambda = \sum_{l=1}^{L} x_l (\ln D_l - \ln F_l) + (1 - x_l)[\ln(1 - D_l) - \ln(1 - F_l)].$$

Позначимо множники при випадкових величинах x_l як

$$Q_l = \ln D_l - \ln F_l - \ln(1 - D_l) + \ln(1 - F_l) = \ln\left[\frac{D_l(1 - F_l)}{(1 - D_l)F_l}\right] \tag{6}$$

та відкинемо доданки, які не залежать від x_l, тоді отримаємо оптимальний за критерієм Неймана – Пірсона алгоритм виявлення при поєднанні попередніх рішень всіх каналів обробки:

$$\ln \Lambda = \sum_{l=1}^{L} Q_l x_l \underset{<}{>} z_0, \tag{7}$$

де z_0 – поріг, який визначається імовірністю хибної тривоги.

Отже, спільна обробка сигналів зводиться до вагового підсумовування одиниць і нулів x_l, що відображають прийняті у пасивному і в активному каналах обробки попередні рішення. Вагові коефіцієнти (6) підвищують роль каналу обробки, де вище імовірність D_l і нижче імовірність F.

Оскільки x_l рівні 0 або 1, то ліва частина (7) є сумою $k < Ln < MN$ вагових коефіцієнтів Q_l, а значить, може приймати лише певні дискретні значення. Значення порога z_0 в цьому випадку може лежати в межах $0 < z_0 < \sum_{l=1}^{L} Q_l$, щоб, з одного боку, не приймалося завжди тривіальне рішення про виявлення, а з іншого – тривіальне рішення про невиявлення. Якщо всі Q_l різні і сума будь-якої групи Q_l не збігається з сумою будь-якої іншої групи, то при різних комбінаціях значень x_l для розглянутого випадку можливі $2^L - 1$ різних значень $\ln \Lambda > 0$. Обираючи поріг виявлення в інтервалах між значеннями Q_l та їх різних сум, можна сформувати $2^L - 1$ різних правил виявлення.

При фіксованих імовірностях попередніх рішень у каналах обробки F_l та D_l різні вирішальні правила дають різні значення імовірностей F і D.

Отриманий алгоритм виявлення аналізованого випадку синтезу повторюється при різному числі сумованих ваг тричі, тобто здійснюються три процедури виявлення: 1) виявлення сигналів відповіді; 2) виявлення повітряного об'єкту; 3) виявлення траси повітряного об'єкту. У цьому випадку загальний вираз для вирішального правила виявлення трас повітряних об'єктів за інформацією запитальних радіолокаційних систем спостереження можна записати в наступному вигляді:

$$R = \sum_{k=1}^{K} Q_k \left(x_k = \left[\sum_{j=1}^{N} Q_j \left\{ x_j = \sum_{i=1}^{n} Q_i x_i \underset{<}{>} z_{01} \right\} \right] z_{02} \right) z_{03}, \tag{8}$$

де $z_{0i}\left(i = \overline{1-3}\right)$ – пороги прийняття рішення кожною із зазначених процедур виявлення.

Коли виконується умова $F_i = F_0$, $P_{0i} D_i = D_0$, то $Q_1 = ... = Q_k = Q$, $k = n(N)K$, тоді у виразі (8) можемо Q винести за знак суми та розділити обидві частини на постійну величину Q. За таких припущень вираз (8) можна записати як

$$R = \sum_{k=1}^{K} \left(x_k = \left[\sum_{j=1}^{N} \left\{ x_j = \sum_{i=1}^{n} x_i \underset{<}{>} z_{01} \right\} \right] z_{02} \right) z_{03}. \tag{9}$$

Як видно з (9), при всіх можливих x_i в кожному з попередніх виявлювачів, що розглядаються, величина $R > 0$ може приймати тільки m різних значень. У цьому випадку отримуємо відоме правило виявлення «k з m», згідно з яким сигнал вважається виявленим, якщо попереднє виявлення сталося хоча б у k з m каналів часової обробки. Легко бачити, що

$2^m - 1$ вирішальних правил, що витікають з (8), при незначних відмінностях у вагових коефіцієнтах включатимуть всі m вирішальних правил типу «k з m», що одержуються з (9).

Таким чином, характерною особливістю виявлювача трас повітряних об'єктів у цьому випадку є наявність чотирьох порогів. Перший поріг z_0 встановлюється в пороговий пристрій виявлювача сигналів коду у відповідь. Цей поріг аналоговий і його можна змінювати, тобто змінювати умовну імовірність хибної тривоги на виході виявлювача трас повітряних об'єктів. Тобто відбувається зміна кінцевої мети розглянутого виявлювача. Другий поріг z_{01} встановлюється в пороговий пристрій при виявленні сигналів у відповідь, третій z_{02} – при виявленні повітряних об'єктів. І четвертий z_{03} – при виявленні трас повітряних об'єктів. Ці пороги можуть бути аналоговими при реалізації алгоритму (8) або дискретними при реалізації алгоритму (9). У першому випадку вдається реалізувати $\left(2^n - 1\right)\left(2^N - 1\right)\left(2^K - 1\right)$ вирішальних правил, тоді як у другому – nNK.

Характерною особливістю такої реалізації (послідовного ухвалення рішення про попереднє виявлення) є те, що можна змінювати послідовність процедур виявлення. Зокрема можна реалізувати виявлювач трас у такій послідовності: виявлювач складових сигналу-відповіді; виявлювач сигналу у відповідь; виявлювач траси; виявлювач повітряних об'єктів. Вирішальне правило такої структури виявлювача трас повітряних об'єктів запитальними радіолокаційними системами спостереження можливо подати в наступному вигляді:

$$R = \sum_{k=1}^{K}\left(x_k = \left[\sum_{j=1}^{N}\left\{ x_j = \sum_{i=1}^{n} x_i \begin{matrix} > \\ < \end{matrix} z_{01} \right\} \right] z_{03} \right) z_{02}. \qquad (10)$$

Відмінність алгоритмів (9) та (10) позначається на схемній побудові виявлювачів трас повітряних об'єктів за інформацією запитальних радіолокаційних систем спостереження.

Таким чином, структура синтезованого квазіоптимального виявлювача трас повітряних об'єктів запитальними радіолокаційних систем спостереження більш уніфікована в схемотехнічній побудові.

Представляє інтерес проведення аналізу синтезованого алгоритму виявлення трас повітряних об'єктів запитальними радіолокаційними системами спостереження повітряного простору з метою порівняння якісних показників синтезованих структур виявлювачів.

Аналіз виявлювача траси повітряних об'єктів запитальних радіолокаційних систем

Будемо вважати, що на вхід запитальних радіолокаційних систем спостереження можуть надходити флуктуаційні та імпульсні (хаотичні, внутрісистемні тощо) завади. Проведемо порівняльний аналіз імовірності виявлення трас повітряних об'єктів за інформацією запитальних радіолокаційних систем спостереження. Модульність побудови виявлювача трас повітряних об'єктів, як показано вище, дозволяє розглядати цю структуру в наступних послідовностях попередніх виявлень:

- виявлювач повітряних об'єктів – виявлювач сигналів відповіді – виявлювач траси повітряних об'єктів (І варіант);
- виявлювач сигналів відповіді – виявлювач повітряних об'єктів – виявлювач траси повітряних об'єктів (ІІ варіант);
- виявлювач повітряних об'єктів – виявлювач траси повітряних об'єктів – виявлювач сигналів відповіді – (ІІІ варіант).

Отримаємо математичні вирази для виявлення трас повітряних об'єктів запитальними радіолокаційними системами при використанні першого варіанту.

Будемо враховувати, що у виявлювачі повітряних об'єктів використовується логіка K/N, для виконання якої необхідна наявність імпульсів сигналів відповіді на одних і тих же ділянках дальності K із N запитів (тобто K виступає в якості цифрового порога виявлення

повітряного об'єкту). У пристрої виявлення сигналів відповіді застосовується логіка n/n, для виконання якої потрібна наявність усіх імпульсів в кожній повторній посилці. При цьому у пристрої виявлення траєкторій використовуються критерій виявлення траєкторій l/m.

При такій постановці питання імовірність виявлення повітряних об'єктів D_1 за результатами виявлення одиночних сигналів відповіді запитальних радіолокаційних систем для зазначеної логіки визначається наступним чином:

$$D_1 = \sum_{i=0}^{N-K} C_N^i P_0^{N-1}(1-P_0)\left[\sum_{i=0}^{N-K-i} C_{N-1}^l P_1^{N-l-i}(1-P_1)^l\right],$$

де P_1 – імовірність виявлення одиночних імпульсів сигналів відповіді.

Імовірність виявлення сигналів відповіді можна визначити з виразу

$$D_{11} = \sum_{i=0}^{N-K} C_N^i P_0^{N-1}(1-P_0)\left[\sum_{i=0}^{N-K-i} C_{N-1}^l P_1^{N-l-i}(1-P_1)^l\right]^n.$$

Імовірність виявлення траси повітряних об'єктів вторинною радіолокаційною системою спостереження для аналізованого варіанта побудови визначимо з виразу

$$D_{111} = \sum_{i=0}^{m} C_m^i D_{11}^i (1-D_{11})^{m-i}.$$

Отримаємо вирази для виявлення трас повітряних об'єктів для другого варіанта побудови виявлювача трас. Імовірність виявлення n-імпульсних сигналів відповіді

$$D_2 = P_1^n P_0.$$

Імовірність виявлення повітряних об'єктів на виході виявлювача повітряних об'єктів

$$D_{22} = \sum_{i=0}^{N-K} C_N^i \left(P_1^n P_0\right)^i.$$

В цьому разі імовірність виявлення трас повітряних об'єктів на виході виявлювача траси визначається як

$$D_{222} = \sum_{i=1}^{m} C_m^i D_{22}^i (1-D_{22})^{m-i}.$$

Отримаємо вираз для виявлення траси повітряних об'єктів для третього варіанта побудови виявлювача. Імовірність виявлення повітряних об'єктів на підставі результатів виявлення одиночних сигналів відповіді можливо записати в наступному вигляді:

$$D_3 = \sum_{i=0}^{N-K} C_N^i P_0^{N-1}(1-P_0)^i \sum_{l=0}^{N-K-i} C_{N-1}^l P_1^{N-l-i}(1-P_1)^l.$$

Імовірність виявлення траси повітряних об'єктів за одиночним сигналом відповіді

$$D_{33} = \sum_{i=0}^{m} C_m^i D_3^i (1-D_3)^{m-i}.$$

Імовірність виявлення траси повітряних об'єктів на виході виявлювача трас можна визначити з виразу

$$D_{333} = D_{33}^n.$$

Оцінимо вплив флуктуаційних завад у каналі сигналів відповіді, коефіцієнта готовності літакового відповідача вторинних радіолокаційних систем спостереження та критерію виявлення траси на значення цифрового порога виявлення повітряних об’єктів запитальних радіолокаційних систем спостереження для пачки сигналів відповіді, що становить 25.

Представлені на рис. 1 – 3 залежності дозволяють проводити порівняльний аналіз існуючих і перспективних запитальних вторинних радіолокаційних систем спостереження за якістю виявлення трас повітряних об’єктів $D = f(K, P_0, k/m)$ при дії у каналі відповіді флуктуаційних та імпульсних завад за логікою обробки k/m, що дорівнює 2/2 (рис. 1), 2/3 (рис. 2), 3/3 (рис. 3). При цьому червоним кольором позначено розрахунки для першого варіанта реалізації виявлювача траси повітряних об’єктів, зеленим – другого варіанту реалізації виявлювача траси повітряних об’єктів, синім – третього варіанту реалізації виявлювача траси повітряних об’єктів.

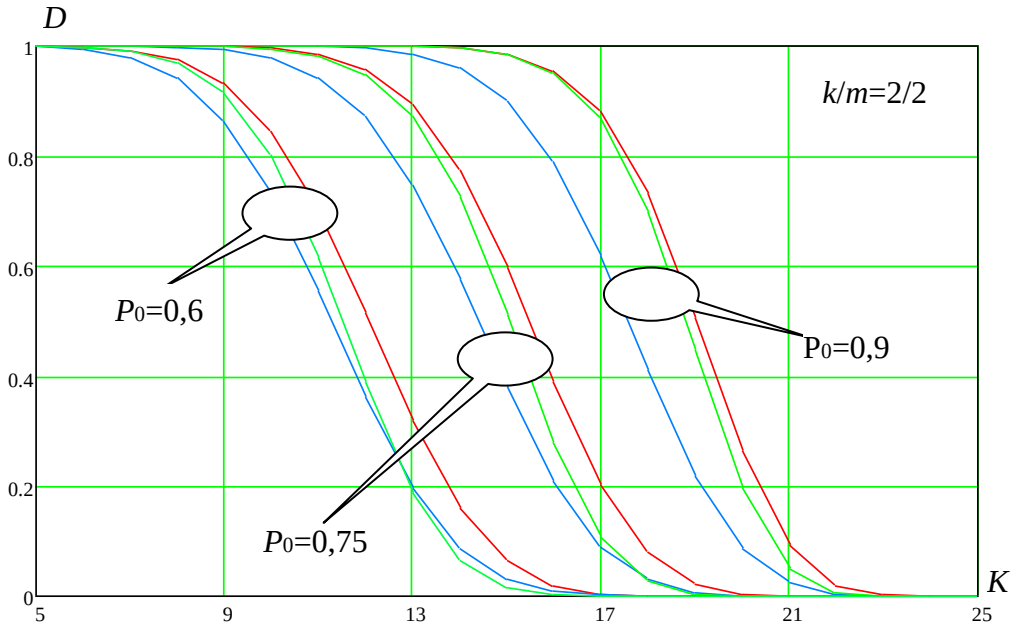


Рис. 1. Імовірність виявлення траси повітряного об’єкта при *k*/*m*=2/2

Представлені порівняльні характеристики якості виявлення трас повітряних об’єктів запитальними вторинними радіолокаційними системами спостереження показали, що перший варіант виявлювача кращий в порівнянні з іншими, розглянутими в роботі. Дійсно, така структура виявлювача трас повітряних об’єктів найменш чутлива до негативної дії коефіцієнта готовності відповідача запитальних радіолокаційних систем спостереження.

Слід зазначити, що у роботі розглянуто випадок однакового відношення сигнал-шум в каналах обробки радіолокаційних систем спостереження. На практиці ж, відношення сигнал-шум вторинних каналів спільної радіолокаційних систем спостереження значно перевищує цей показник відносно до первинного каналу.

Порівняльний аналіз рис. 1 – 3 дозволяє зробити наступні висновки:

- якість інформаційного забезпечення споживачів на підставі запропонованої структури вище в порівнянні зі структурою обробки інформації запитальних радіолокаційних систем, що використовується в даний час;

- якість інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів, заснованого на накопиченні і з наступним об'єднанням інформації запитальних радіолокаційних систем;

- коефіцієнт готовності літакових відповідачів істотно впливає на якість інформаційного забезпечення споживачів системи контролю повітряного простору.

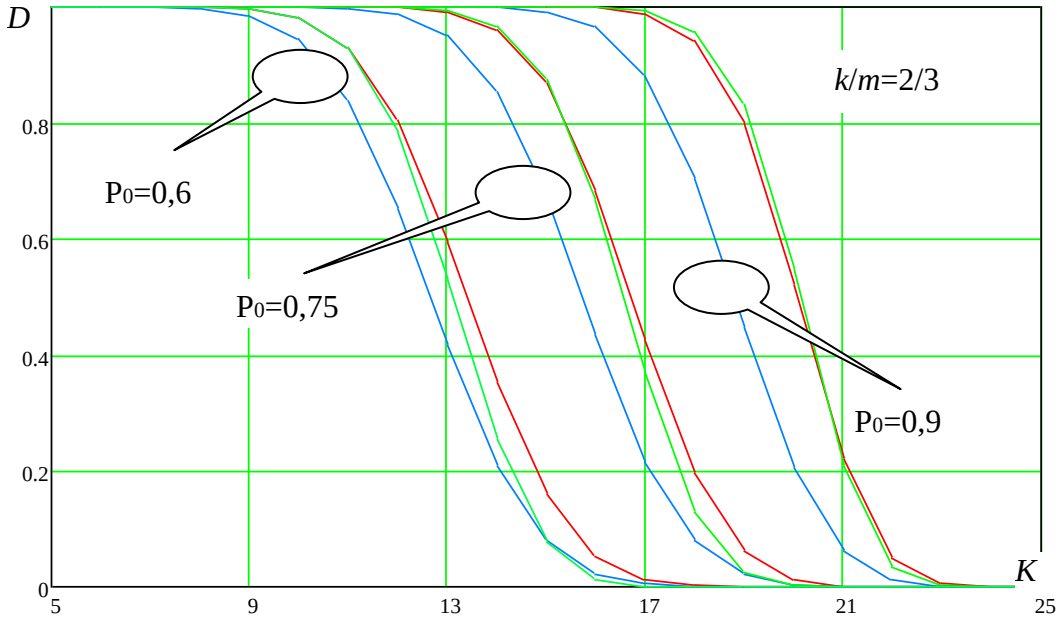


Рис. 2. Імовірність виявлення траси повітряного об’єкта при *k*/*m*=2/3

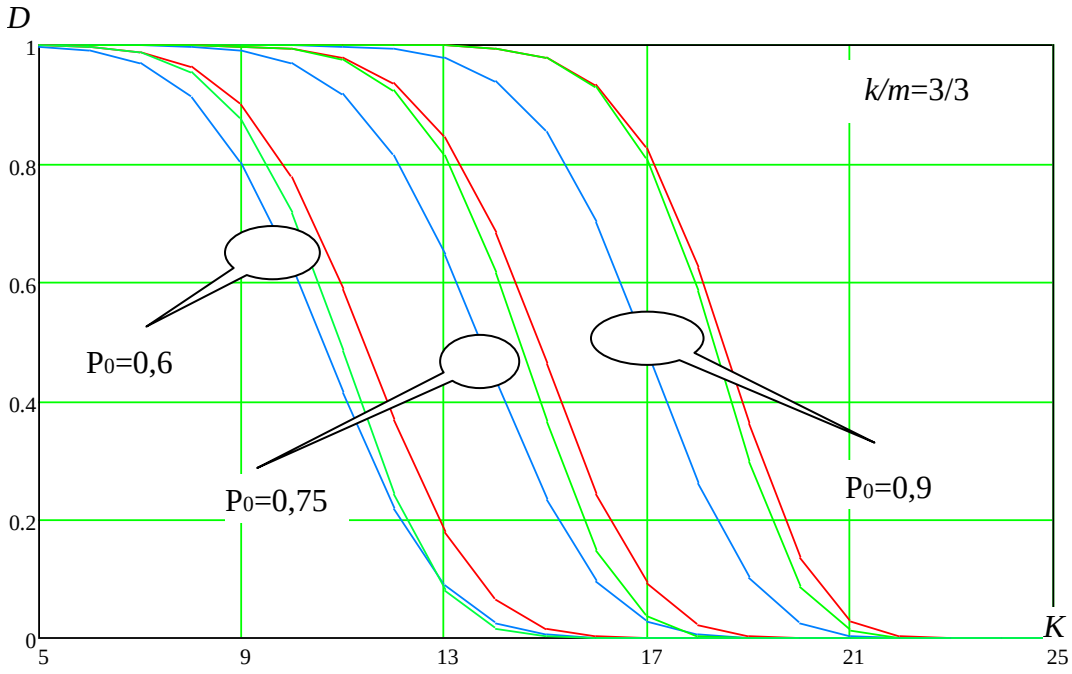


Рис. 3. Імовірність виявлення траси повітряного об’єкта при *k*/*m*=3/3

Висновки

Проведено синтез та аналіз структури виявлювача траси повітряних об’єктів запитальними радіолокаційними системами спостереження повітряного простору, що дозволило:

- провести порівняльний аналіз якості виявлення трас повітряних об’єктів розглянутими конфігураціями структури виявлювача трас повітряних об’єктів при різних послідовностях проведення операцій: виявлення сигналів відповіді, виявлення повітряного об’єкта та виявлення траси повітряного об’єкта;

- підвищити якість інформаційного забезпечення споживачів системи контролю повітряного простору на підставі запропонованої структури в порівнянні зі структурою обробки інформації запитальних радіолокаційних систем, що використовується в даний час;

- показати, що якість інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів, заснованого на накопиченні і з наступним об'єднанням інформації запитальних радіолокаційних систем;

- оцінити вплив коефіцієнта готовності літакових відповідачів запитальних радіолокаційних систем на якість інформаційного забезпечення споживачів системи контролю повітряного простору.

Список літератури:

1. І. Свид, І. Обод. Завадостійкість радіолокаційних систем ідентифікації за ознакою свій-чужий. Харків : Друкарня Мадрид, 2021. 253 с. doi: 10/30837/978-617-7988-76-1.
2. І. Обод, І. Свид, О. Мальцев. Обробка даних радіолокаційних систем спостереження повітряного простору : навч. посібник. Харків : Друкарня Мадрид, 2021. 255 с.
3. J. Li, P. Stoica. MIMO Radar Signal Processing. Wiley-IEEE Press, 2008. 448 p.
4. Свид І. В. Обробка радіолокаційної інформації систем спостереження повітряного простору: монографія. Дніпро : ЛІРА ЛТД, 2022. 224 с.
5. M. Barbary, A. S. Hafez and T. Crew. An Industrial Design and Implementation Approach of Secondary Surveillance Radar System // 2021 International Telecommunications Conference (ITC-Egypt). 2021. pp. 1 – 9. doi: 10.1109/ITC-Egypt52936.2021.9513961.
6. I. Svyd, I. Obod, O. Maltsev, V. Andrusevich, B. Bakumenko and O. Vorgul. Optimal Measurement of Signal Data Parameters of Requesting Radar Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON). 2021. pp. 138 – 141. doi: 10.1109/UKRCON53503.2021.9575235.
7. F.L. Neindre, G. Ferre, D. Dallet, F. Letellier and K. Pitois. A Successive Interference Cancellation-based Receiver for Secondary Surveillance Radar // IEEE Transactions on Aerospace and Electronic Systems, 2022. doi: 10.1109/TAES.2022.3193649.
8. I. Obod, I. Svyd, O. Maltsev and S. Starokozhev. The Effect of Masking Interference on the Quality of Request Signal Detection in Aircraft Responders of the Identification Friend or Foe Systems // 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T). 2020. pp. 721 – 726. doi: 10.1109/PICST51311.2020.9467955.
9. V. Semenets et al. Method of increasing the relative throughput of requesting radar systems // Przegląd Elektrotechniczny. 2022. Vol. 1, no. 11. pp. 99 – 103. doi: 10.15199/48.2022.11.17.
10. I. Svyd, I. Obod, O. Maltsev and A. Hlushchenko. Secondary Surveillance Radar Response Channel Information Security Improvement Method // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2020. pp. 341 – 345, doi: 10.1109/DESSERT50317.2020.9125018.
11. M. Leonardi and D.D. Fausto. Secondary Surveillance Radar Transponders classification by RF fingerprinting // 2018 19th International Radar Symposium (IRS). 2018. pp. 1 – 10. doi: 10.23919/IRS.2018.8448244.
12. I. Obod, I. Svyd, O. Vorgul, O. Maltsev, O. Datsenko and N. Boiko. Optimization of Data Processing Structure for Multi-Position Radar Surveillance Systems // 2021 IEEE 3rd Ukraine Conference on Electrical and Computer Engineering (UKRCON). 2021. pp. 133 -ю137. doi: 10.1109/UKRCON53503.2021.9575286.
13. P. Švábeník, D. Zeman, R. Balada and Z. Fedra. Separation of secondary surveillance radar signals // 2011 34th International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 2011. pp. 487 – 490. doi: 10.1109/TSP.2011.6043683.
14. I. Svyd, I. Obod and O. Maltsev. Interference Immunity Assessment Identification Friend or Foe Systems // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 287 – 306, 2021. doi: 10.1007/978-3-030-71892-3_12.
15. W. Konle. Separate processing of primary and secondary radar data in multi radar tracking // 2013 14th International Radar Symposium (IRS), Dresden, Germany, 2013. pp. 361.-.366.
16. V. Semenets, I. Svyd, I. Obod, O. Maltsev and M. Tkach. Quality Assessment of Measuring the Coordinates of Airborne Objects with a Secondary Surveillance Radar // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, 2021. pp. 105 – 125. doi: 10.1007/978-3-030-71892-3_5.
17. I. Ivashko, O. Krasnov and A. Yarovoy. Performance analysis of multisite radar systems // 2013 European Microwave Conference, 2013. pp. 1771 – 1774. doi: 10.23919/EuMC.2013.6687021.
18. Толюпа С.В., Дружинін В.А., Гордієвський О.Т. Розпізнавання групових об'єктів у багатопозиційних системах оперативного супроводження // Сучасний захист радіолокаційної інформації. 2012. No 1. С. 66 – 70.
19. Обод І.І., Стрельницький О.О. Інформаційна безпека інформаційної мережі систем спостереження повітряного простору // Системи обробки радіолокаційної інформації. 2015. № 9(134). С. 96 – 98.
20. Обод І.І., Стрельницький О.О. Захист радіолокаційної інформації в мережі систем спостереження повітряного простору // Системи обробки радіолокаційної інформації. 2016. № 2(139). С. 47 – 49.

21. 12. J. Xu, X. -Z. Dai, X. -G. Xia, L. -B. Wang, J. Yu and Y. -N. Peng, Optimizations of Multisite Radar System with MIMO Radars for Target Detection // IEEE Transactions on Aerospace and Electronic Systems. Vol. 47, no. 4. pp. 2329 – 2343, OCTOBER 2011. doi: 10.1109/TAES.2011.6034636.

22. 18. I. Svyd, I. Obod, O. Maltsev, O. Vorgul, V. Chumak and B. Bakumenko, Estimation of the Spatial Coordinates of Air Objects in Synchronous Radar Networks for Airspace Observation // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T). 2021. pp. 425-428. doi: 10.1109/PICST54195.2021.9772227.

23. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях радиолокации // Системи обробки радіолокаційної інформації. 2006. № 9(58). С. 69 – 75.

24. Обод И.И., Булай А.Н., Луценко Ю.А. Оценка точности определения местоположения воздушных объектов в синхронных информационных сетях // Системи обробки радіолокаційної інформації. 2006. № 9(58). С. 69 – 71.

25. H. You, X. Jianjuan, G. Xin. Radar Data Processing with Applications // Publishing House of Electronics Industry, 2016. doi: 10.1002/9781118956878.

26. Chen Su, Chuanyun Zou, Liangyu Jiao, Qianglin Zhang. A MIMO Radar Signal Processing Algorithm for Identifying Chipless RFID Tags. Sensors (Basel). 2021 Dec 12;21(24):8314. doi: 10.3390/s21248314

27. Обод І.І., Стрельницький О.О., Андрусевич В.А. Методи підвищення якості інформаційного забезпечення системами спостереження повітряного простору // Системи обробки радіолокаційної інформації. 2014. № 4(120). С. 53 – 55.

28. Обод І.І., Шевцова В.В. Порівняльний аналіз запитальних систем передачі радіолокаційної інформації системи контролю повітряного простору // Зб. наук. пр. Харк. нац. ун-ту Повітряних Сил. 2013. № 1(34). С. 123 – 125.

29. Обод И.И. Обнаружение воздушных целей системой вторичной радиолокации // Радіоелектронні і комп'ютерні системи. 2005. № 3. С.25 – 28.

30. І. Свид, В. Семенець, О. Мальцев, М. Ткач, С. Старокожев, О. Даценко, І. Шевцов. Порівняльний аналіз методів визначення координат повітряних об'єктів системами широкозонової мультилатерації // Радіотехніка. 2022. Вип. 209. С. 162 – 177. doi: 10.30837/rt.2022.2.209.16.

31. S. M. Wu, G. A. Ybarra and W. E. Alexander. A complex optimal signal-processing algorithm for frequency-stepped CW data // IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing. June 1998. Vol. 45, no. 6. pp. 754 – 757.. doi: 10.1109/82.686697.

32. Толюпа С.В., Дружинін В. А., Наконечний В.С., Цьопа Н.В., Батрак Є.О. Методи та алгоритми обробки радіолокаційної інформації у багатопозиційних системах зі змінною просторовою конфігурацією. Київ : Логос, 2014. 230 с.

33. М. Ткач, І. Свид, О. Воргуль, С. Старокожев, О. Мальцев, А. Глущенко. Оцінка відносної пропускної здатності запитальних систем спостереження повітряного простору // Радіотехніка. 2022. Вип. 208. С. 28 – 37. doi: 10.30837/rt.2022.1.208.03.

34. G. Lee, S. Lee, K. Kim and N. Kwak. Probabilistic Track Initiation Algorithm Using Radar Velocity Information in Heavy Clutter Environments // 2018 15th European Radar Conference (EuRAD), 2018. pp. 277 – 280. doi: 10.23919/EuRAD.2018.8546666.

35. Conte E., Daddio E., Farina A., and Longo M. Multistatic radar detection – Synthesis and comparison of optimum and suboptimum receivers // IEE Proceedings F: Communications Radar and Signal Processing. 1983. Vol. 130, no. 6. pp. 484 – 494.

36. M. K. Abdul-Hussein, O. Strelnytskyi, I. Obod, I. Svyd and H. Alrikabi, Evaluation of the Interference's Impact of Cooperative Surveillance Systems Signals Processing for Healthcare // International Journal of Online and Biomedical Engineering (iJOE). 2022. Vol. 18, no 03. pp. 43 – 59. doi: 10.3991/ijoe.v18i03.28015.

37. I. Prokopenko, V. Vovk and K. Prokopenko. Fast resource management algorithm for multi-position radar systems // 2015 16th International Radar Symposium (IRS). 2015. pp. 1045 – 1051. doi: 10.1109/IRS.2015.7226339.

38. V. Andrusevich and I. Obod. Assessment of the Quality of Information Support by Air Radar Surveillance Systems, Advanced Information Systems. 2021. Vol. 5, no. 2. pp. 78 – 82. doi: 10.20998/2522-9052.2021.2.10.

39. I. Prokopenko, V. Vovk, S. Stavitsky and V. Medvedev. Optimization of use of resource in multi-position radar systems // 2014 IEEE Microwaves, Radar and Remote Sensing Symposium (MRRS), 2014. pp. 92 – 97. doi: 10.1109/MRRS.2014.6956673.

40. I. Shevtsov et al. A Method for Increasing the Capacity of Radio Systems of Short-Range Navigation // 2022 IEEE 2nd Ukrainian Microwave Week (UkrMW). Ukraine, 2022. pp. 629 – 633. doi: 10.1109/UkrMW58013.2022.10037138.

41. S. Starokozhev, M. Tkach, A. Hlushchenko, O. Datsenko, M. Chernyshov and V. Chumak. Frequency Efficiency Evaluation of Query Airspace Surveillance Systems // 2021 IEEE 8th International Conference on Problems of Infocommunications, Science and Technology (PIC S&T), Kharkiv, Ukraine, 2021. pp. 501 – 505. doi: 10.1109/PICST54195.2021.9772190.

42. S. Starokozhev, M. Tkach, A. Hlushchenko, O. Datsenko, M. Chernyshov and V. Chumak. Optimization of the Probability of Transmission of Flight Data in the Response Channel of Secondary Radar Systems // 2021 IEEE 8th

International Conference on Problems of Infocommunications, Science and Technology (PIC S&T). Kharkiv, Ukraine, 2021. pp. 511 – 515. doi: 10.1109/PICST54195.2021.9772199.

43. Свид І.В., Ткач М.Г., Обод І.І. Порівняльний аналіз завадозахищеності радіолокаційних систем ідентифікації свій-чужий // Радіотехніка. 2022. Вип. 211. С. 101 – 113. doi: 10.30837/rt.2022.4.211.08.

44. І. Свид, М.Ткач, А.Сєріков, О. Коротіч, С. Дацько, Д. Сухоруков, Т. Мачоніс. Обробка інформації мереж радіолокаційних систем спостереження повітряного простору // Радіотехніка. 2022. Вип. 210. С. 137 – 145. doi: 10.30837/rt.2022.3.210.11.

45. Ткач М.Г. Оцінка відносної пропускної здатності літакових відповідачів вторинних радіолокаційних систем спостереження повітряного простору // Радіотехніка. 2021. Вип. 207. С. 123 – 131. doi: 10.30837/rt.2022.3.210.11.

46. І. Свид, І. Воргуль, С. Старокожев, М. Ткач, О. Мальцев, І. Шевцов. Порівняльний аналіз завадостійкості каналу передачі інформації вторинних радіолокаційних систем // Радіотехніка. 2022. Вип. 208. С. 44 – 54. doi: 10.30837/rt.2022.1.208.05.

Надійшла до редколегії 28.02.2023

Відомості про авторів:

Свид Ірина Вікторівна – кандидат технічних наук, доцент, завідувач кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: iryna.svyd@nure.ua; ORCID: http://orcid.org/0000-0002-4635-6542

Ткач Марія Геннадіївна – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: mariia.zavorotna@nure.ua; ORCID: http://orcid.org/0000-0002-4248-7633

MEANS OF TELECOMMUNICATIONS
ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ

УДК 621.396.004 **DOI:10.30837/rt.2023.1.212.18**

Л.О. ТОКАР, канд. техн. наук, О.А. КОЛТАКОВ, В.Є. ЦИЛЮРИК

СТВОРЕННЯ ТЕСТОВОГО СТЕНДУ CALL-ЦЕНТРУ ДЛЯ БАЛАНСУВАННЯ НАВАНТАЖЕННЯ СЕРВЕРІВ ASTERISK У КЛАСТЕРІ

Вступ

У зв’язку з зростанням попиту на телекомунікаційні послуги сервери голосових викликів та викликів даних повинні забезпечувати більшу пропускну здатність обробки викликів. Інтернет та мобільні пристрої сприяли збільшенню потреб в комутаційних потужностях телекомунікаційних серверів. Через вимогливі додатки можливості існуючих серверів обробки викликів доведено до краю.

Виходячи з цього, одним із рішень є кластеризація серверів викликів. Кластер серверів викликів є групою з декількох окремих серверів викликів, що охоплюють певну географічну область. Щоб подолати зростаючий попит на пропускну здатність, необхідні дослідження параметрів продуктивності кластерів серверів викликів для отримання оцінки про якість та збалансованість роботи центру обробки викликів. Кластеризація серверів у call-центрі забезпечує переваги як масштабованості, так і надійності. Тема роботи є актуальною, що обумовлено необхідністю ефективної обробки викликів серверами у кластері для забезпечення потрібної відмовостійкості та мінімізації експлуатаційних витрат для телефонних компаній.

Call-центр являє собою центр обробки телефонних викликів, принцип побудови якого ґрунтується на маршрутизації викликів агентів за певними правилами, які розробляються в компанії та дозволяють якісно та ефективно обслуговувати клієнтів, а також підтримувати черги дзвінків. Він здатний не тільки приймати та обробляти запити, але й використовувати для контактів з клієнтами звичайну пошту, факсимільний та мобільний зв'язок, Інтернет, SMS (short message service) тощо.

Процес моделювання call-центру являє собою віртуалізацію всієї системи, або деяких її частин. Це необхідно для більш детального розгляду деяких аспектів функціонування мережі, подальшого уникнення проблем при побудові реальної телефонної мережі, виявлення слабких ланок мережі, аналізу трафіку тощо [1]. Особливістю процесу віртуалізації є налаштування АТС для обслуговування клієнтських пристроїв. Для цього доцільно створити віртуальну машину з встановленою на ній віртуальною АТС, наприклад Asterisk, FreePBX або Elastic, налаштувати її в одну мережу з клієнтськими пристроями та маршрут до неї. Варто зазначити, що встановлений SIP-сервер сам не обробляє медіапотоки – це робить окремий медіа-сервер, використовуючи протокол RTP. У реалізаціях IP-АТС завжди SIP-сервер й медіа-сервер перебувають у одному фізичному сервері. Однак системи з великим навантаженням (наприклад, у великих VoIP-операторів) можуть використовувати медіа-сервер, встановлений на іншій фізичній машині, щоб краще справлятися з обробкою сесій. Також можливе розподілення навантаження на кілька медіа-серверів [2].

Основна частина

Основними вимогами щодо стабільного функціонування call-центру можна визначити: відмовостійкість системи; якісний та швидкий доступ до мережі Інтернет; використання системи, яка здатна інтегрувати в собі декілька технологій (передача даних, голосу, відео); безпека дзвінків шляхом шифрування даних; масштабованість; здійснення дзвінків в загальну абонентську мережу PSTN (public switched telephone network) [3].

Балансування навантаження є важливим аспектом кластерів у call-центрі. Один із серверів обробки викликів у кластері може бути перевантажений, у той час як інші сервери обробки викликів можуть працювати нижче за призначений рівень навантаження. У разі такої незбалансованої ситуації можна розподілити додатковий трафік перевантаженого серверу у кластері на слабозавантажений сервер. Такий рівень розподілу навантаження не дозволяє одному серверу викликів стати вузьким місцем у продуктивності.

Аналіз балансування навантаження показав, що методи динамічного балансування навантаження дозволяють використовувати оперативну інформацію про стан системи прийняття рішень для коригування балансування навантаження під час роботи, що впливає на продуктивність системи. Прийняття рішень можливо використовувати на різних рівнях: рівні завантаження ЦП (центрального процесора), доступної пам'яті тощо [4].

У літературі [5] запропоновано використовувати двокаскадний комутатор з балансуванням навантаження для масштабування до швидкості оптоволокна. Такий підхід вважається більш ефективним, ніж інші підходи, такі як i-SLIP, що називається алгоритмом вирішення конфліктів. Його засновано на ітеративному циклічному зіставленні з ковзанням. У літературі [6] показано, що для трафіку з багатопротокольною комутацією за мітками MPLS (multiprotocol label switching) алгоритми динамічного розподілу навантаження кращі за продуктивністю, ніж алгоритм найкоротшого шляху. Зазначено, що алгоритми динамічного балансування навантаження ефективні як з легких, так і з важких умов навантаження.

Алгоритм перенаправлення запитів для кластерів веб-серверів досліджено в [7]. Проведено класифікацію кожного нового запиту на основі очікуваного впливу на ресурси сервера. При цьому кожному запиту надається вага, яка використовується для розрахунку навантаження на сервер. Щоб скоротити час збору інформації про навантаження, вибираються лише два сервери з пулів серверів, і новий запит перенаправляється на найменш завантажений сервер.

У [8] запропоновано використовувати стратегії розподілу навантаження, що засновано не тільки на ресурсах ЦП, але й на врахуванні ресурсів пам'яті. Метою запропонованих стратегій розподілу навантаження є мінімізація як часу простою ЦП, так і кількості відмов сторінок в гетерогенних розподілених системах.

У [9] використано мобільні агенти для забезпечення розподілу навантаження у глобальному мережному середовищі, такому як Інтернет. Мобільні агенти діють як координатори від імені серверів і представляють себе на віддалених майданчиках для координації дій щодо переміщення робочих місць.

У роботі створено схему організації call-центру компанії. Основним елементом виступає кластер з декількох SIP-серверів (якщо компанія здійснює досить велику кількість дзвінків). Сервер може бути як фізичний, так і віртуальний (орендована віртуальна машина (ВМ). SIP-сервери зв’язуються з CRM системою, яка призначена для автоматизації стратегій взаємодії із замовниками (клієнтами). Разом вони формують ядро системи, яке має швидкісне з’єднання з мережею Інтернет, де серверам виділяються номери, що виділені SIP-провайдером. Офіс call-центру, з’єднаний з ядром, складається з IP-телефонів та комп’ютерів. Він може територіально знаходитись як в одному будинку з ядром мережі, так і в іншому місці.

В якості програмного забезпечення запропоновано обрати Asterisk. Модульна архітектура Asterisk дозволяє легко підключати в комутаційне поле будь-яку бізнес-логіку, написану практично будь-якою мовою програмування, або реалізовану власною мовою діалплану Asterisk. Серед основних функцій Asterisk слід зазначити: підтримку як протоколів IP-телефонії, так й традиційних ліній зв'язку; підтримку відеозв'язку; підтримку шифрування розмов; наявність простих й добре документованих інтерфейсів для інтеграції з іншими системами; підтримку всіх базових та розширених функцій АТС; наявність готових дистрибутивів [10]. Asterisk може працювати практично на будь-якій платформі Linux та на деяких

інших ОС (операційних системах), таких як Solaris, BSD, MacOS X. На рис. 1 представлено схему організації call-центру компанії.

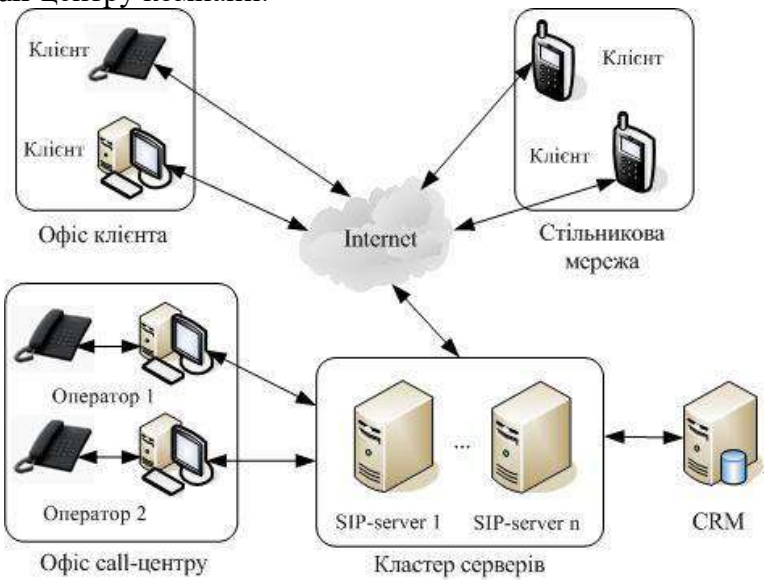


Рис. 1. Схема організації call-центру компанії

В роботі створено мережну модель call-центру, яка визначає загальну структуру, що моделюється: об'єкти в системі, а також їх фізичне розташування, взаємозв'язки та конфігурації. Мережну модель call-центру показано на рис. 2.

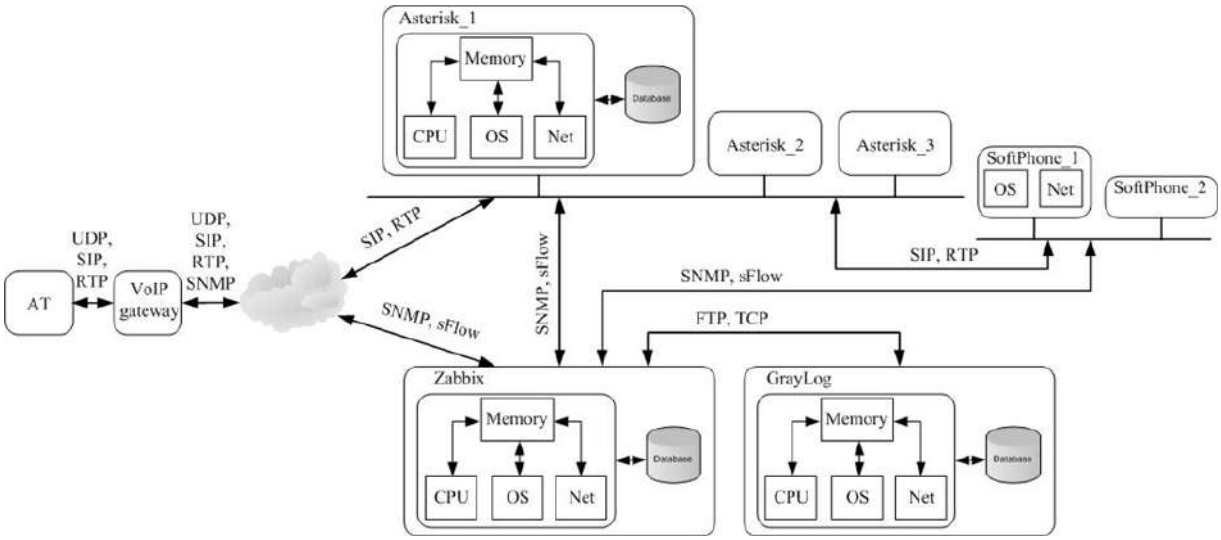


Рис. 2. Мережна модель call-центру

Основними структурними блоками домену мережі є підмережі, вузли та канали зв'язку. У цьому розумінні один сервер виклику становить об'єкт підмережі в домені мережі. Кластер створюється в мережному домені шляхом з'єднання кількох серверів викликів через канали зв'язку.

Абонентський термінал являє собою пристрій, який формує виклики до АТС на виділений її зовнішній номер. На даному сегменті мережі визачено тип трафіку – це UDP, SIP, RTP, SNMP. VoIP шлюз призначено для підключення телефонних апаратів до АТС через IP-мережу та передачі й обробки голосового трафіку.

Asterisk являє собою сервер АТС, що керує голосовим трафіком та викликами в мережі, у складі якого слід відзначити блок пам’яті, центральний процесор, мережний адаптер,

операційну систему та базу даних. Характер трафіку визначається переважно протоколами SIP та RTP. В моделі показано три сервери, які об'єднано в кластер.

У якості серверу моніторингу, що збирає необхідні метрики з хостів мережі та записує всі події до log серверу GrayLog, обрано open source продукт Zabbix. Такий вибір обумовлено простотою підключення для моніторингу будь-яких параметрів інформаційних систем у IT-інфраструктурах. Запропоновано застосування Zabbix`а для дослідження кластеру серверів викликів. Тип трафіку визначено протоколами SNMP, sFlow та SIP. Сервер логів GrayLog з'єднано з сервером моніторингу Zabbix. Логи зберігаються в його виділеній базі даних, до якої можна звернутися у будь який момент. Основний тип трафіку: TCP, FTP.

Налаштування тестового стенду починається з налаштування віртуальної машини. У якості середовища обрано VMWare.

У роботі використано технологію віртуалізації, яка дасть змогу побудувати мережу call-центру та провести необхідне тестування. У якості платформи для налаштування мережі використано гіпервізор VMWare ESXI 6.7 та кліент vCenter. Гіпервізор VMware ESXi – це потужний інструмент, який емулює апаратні ресурси, дозволяє безпечно виконувати машинні інструкції, ізолює та поділяє ресурси віртуальних машин [11]. В табл. 1 наведено обрані параметри віртуальних машин для налаштування.

Таблиця 1

Name	CPU	Memory	OS	IP-address	Hard Disk
1	2	3	4	5	6
Asterisk_1	4	6 Gb	Ubuntu 20.04	192.168.200.112	60 Gb
Asterisk_2	4	6 Gb	Ubuntu 20.04	192.168.200.104	60 Gb
Asterisk_3	4	6 Gb	Ubuntu 20.04	192.168.200.109	60 Gb
Zabbix	6	6 Gb	Ubuntu 20.04	192.168.200.107	60 Gb
Gray_Log	4	6 Gb	Ubuntu 20.04	192.168.200.110	60 Gb
SoftPhone_1	2	4 Gb	Ubuntu 20.04	192.168.200.108	60 Gb
SoftPhone_2	2	4 Gb	Ubuntu 20.04	192.168.200.111	60 Gb
Mikrotik	1	130 Mb	RouterOS	192.168.200.254	500 Mb
SiPp	4	6 Gb	Ubuntu 20.04	192.168.100.10	60 Gb

Вікно налаштування параметрів ВМ (віртуальної машини) наведено на рис. 3.

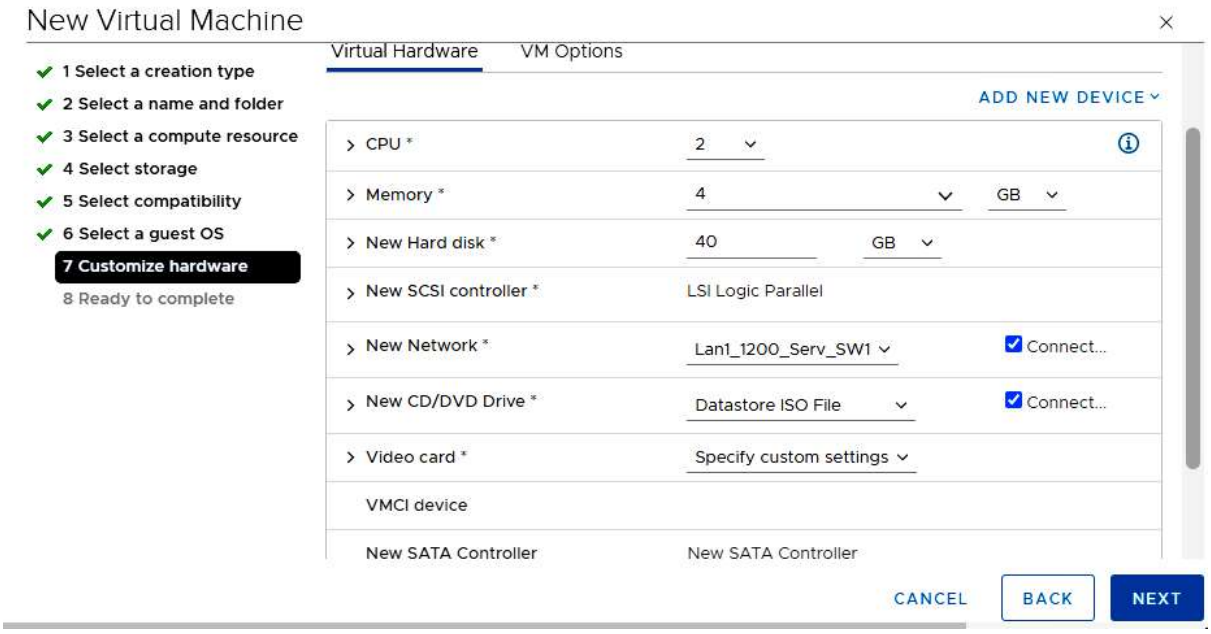


Рис. 3. Налаштування параметрів ВМ

При створенні ВМ використано наступні параметри: CPU (central processing unit), memory, місце на жорсткому диску, мережний адаптер та ISO образ ОС (операційна система). Перевагу надано ОС Ubuntu Server 20.04 за таких ознак: безкоштовність, досить велике коло користувачів, легкість для розгортання будь яких мережних або програмних сервісів, постійне оновлення, стабільність роботи [12]. В якості ядра мережі call-центру обрано образ RouterOS від компанії Mikrotik, основні функції якого наступні: маршрутизація трафіку в мережі, видавання адрес хостам, виконання ролі firewall для безпеки локальної мережі.

Наступним етапом є налаштування АТС Asterisk та створення кластеру серверів.

Встановлення АТС Asterisk на сервер Ubuntu здійснювалося з готових репозиторіїв. Готовий пакет asterisk встановлено з офіційних джерел [13, 14]. Надалі йде перевірка стану служби sudo systemctl status asterisk та проводиться конфігурація сервера. Продукт налаштовується дуже гнучко й має масу можливостей. Основними файлами конфігурації виступають sip.conf extensions.conf. Файл sip.conf відповідає за налаштування внутрішніх й зовнішніх каналів SIP в Asterisk. При цьому використано об'єкти конфігурації – піри. В налаштуваннях діє принцип успадкування, як і в більшості конфігів Asterisk. Подальші налаштування зводяться до глобальних налаштувань драйвера SIP Asterisk, які поширюються на всі об'єкти, але можуть бути перевизначені для окремих бенкетів у їх категоріях.

Слід зазначити один із найважливіших конфігураційних файлів Asterisk – extensions.conf, у якому визначається обробка та маршрутизація вхідних та вихідних викликів. Цей файл керує поведінкою всіх з'єднань, що проходять через АТС. Зміст файлу extensions.conf розбито на секції, в яких можуть бути або визначені статичні налаштування та визначення або команди плану набору, що виконуються; в цьому випадку вони називаються контекстами. Секції, призначені для статичних налаштувань, називаються general і globals, а імена контекстів визначаються системним адміністратором системи.

В роботі було прийнято рішення обмежитись лише двома віртуальними телефонами для перевірки викликів в мережу. На наступному кроці налаштовано Sip-транки для об’єднання серверів Asterisk в кластер. Для цього на кожному з серверів введено новий пір, який буде використовувати сервер для підключення до наступного. Тобто налаштовується реєстрація серверів меж собою, але без підтвердження. На рис. 4 показано процес перевірки реєстрації серверів між собою на прикладі Asterisk_1.

```
aster1*CLI> sip show peers
Name/username              Host                                  Dyn Forcerport Comedia    ACL Port     Status
ion
201/201                    192.168.200.113                        D  Auto (No)  No             57085    Unmonitored
202/202                    192.168.200.100                        D  Auto (No)  No             50981    Unmonitored
asterisk2                  192.168.200.104                           Auto (No)  No             5060     OK (1 ms)
asterisk3                  192.168.200.109                           Auto (No)  No             5060     OK (1 ms)
4 sip peers [Monitored: 2 online, 0 offline Unmonitored: 2 online, 0 offline]
```

Рис. 4. Інформація про реєстрацію в системі на прикладі Astersik_1

Подальшим кроком є налаштування SoftPhone. В якості IP телефонів в роботі використано програмне забезпечення Zoiper5. На відміну від більшості тестових стендів, на ВМ встановлено графічну оболонку XFCE для взаємодії з програмою.

Після цього проводилося створення Zabbix серверу та додавання host моніторингу. Основна задача сервера Zabbix – це моніторинг всієї мережі call-центру, зняття необхідних метрик, наприклад стану каналів, стану SIP-транків, навантаження ВМ тощо. Доменне ім'я серверу в роботі не виділяється. Підключення виконується за його IP-адресою та портом. Для обробки коду та створення динамічного контенту для веб-сервера встановлюється та налаштовується веб-сервер apache в середовищі Ubuntu, база даних MySql та PHP (personal home page) [15].

На наступному етапі розгорнуто платформу за допомогою репозиторіїв. Zabbix має чотири основні елементи, за допомогою яких можна моніторити певне робоче середовище й

збирати про нього повний пакет даних для оптимізації роботи: сервер, проксі, агент, веб-інтерфейс є частиною сервера системи й вимагає для роботи веб-сервер [16]. Для тестування та моніторингу обладнання в мережі потрібно додати відповідні параметри. На рис. 5 показано процес додавання нового host до системи.

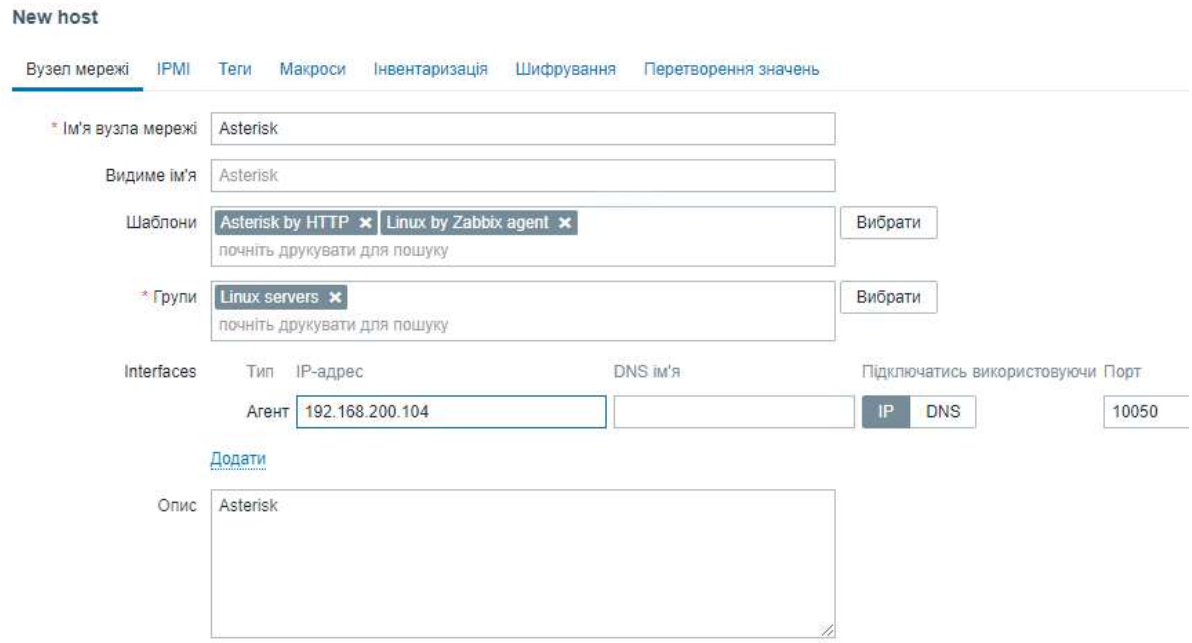


Рис. 5. Процес додавання host моніторингу

Крім інструмента моніторингу, який збирає необхідні метрики в системі та надає їх адміністратору, Zabbix може бути певним чином сконфігуровано для оперативного реагування на певні прояви системи. Робиться це за допомогою скриптів та тригерів. Тому в подальших налаштуваннях слід звернути увагу на тригери в Zabbix, які є логічними висловлюваннями, що відображають стан системи. Висловлювання, що використовуються в тригерах, є дуже гнучкими. Адміністратор може використовувати їх для створення складних логічних тестів з огляду на статистику з моніторингу [17]. Функції тригерів дозволяють посилатися на зібрані значення, поточний час та інші фактори. Стан (вираз) тригера перераховується щоразу, коли Zabbix сервер отримує нове значення даних, якщо це значення даних є частиною висловлювання. Zabbix підтримує такі важливі висловлювання тригерів: невідома важливість; в інформаційних цілях; попереджувальний; середня проблема; сталося щось важливе; надзвичайний [18]. На рис. 6 показано приклад тригерів, які використовуються для Asterisk, їх можна вибрати з шаблону при налаштуванні моніторингу host.

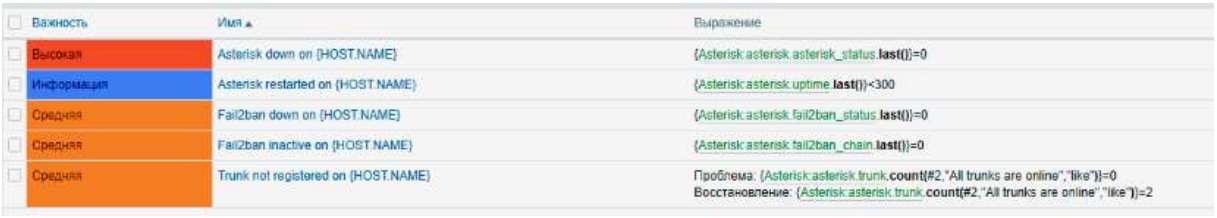


Рис. 6. Тригери для серверів Asterisk

У шаблоні присутні шість елементів даних, які визначаються в агенті, п’ять тригерів та один графік.

Віджети, графіки та карти мереж в Zabbix потрібні для візуалізації даних. Це, насамперед, полегшує сприйняття інформації адміністратором системи.

Після налаштувань для візуалізації даних в Zabbix отримано графіки та карту мережі. Карта мережі являє собою змодельовану структуру мережі call-центру. На рис. 7 показано графік з трафіком серверу для Asterisk_1.

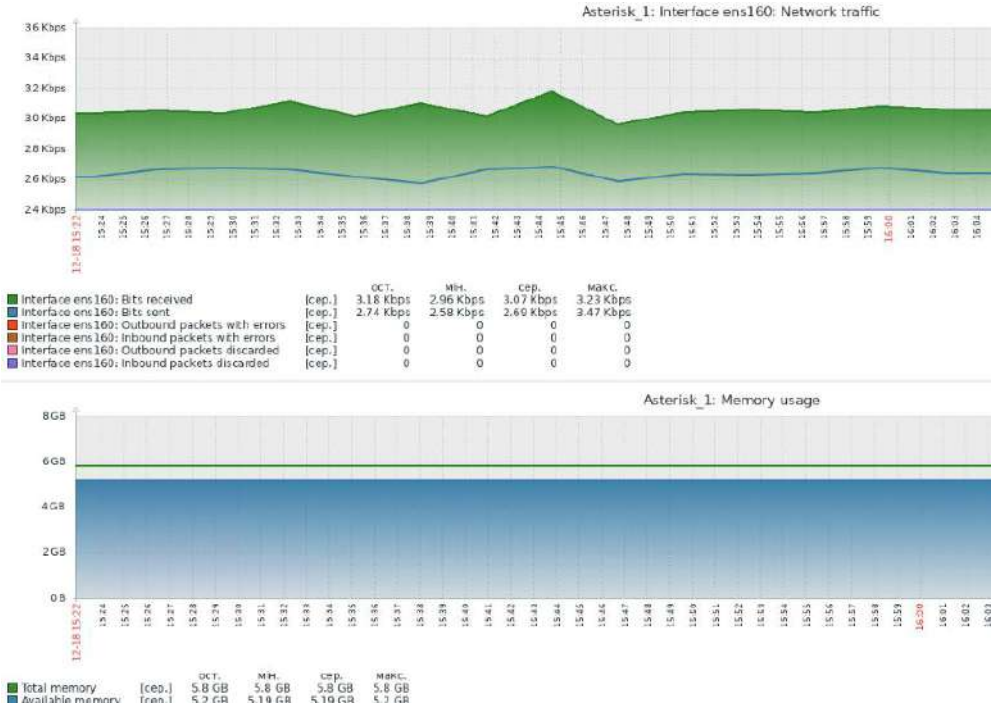


Рис. 7. Графіки пропускної здатності інтерфейсу ens160 та пам’яті сервера Astersik_1

На рис. 8 показано топологію мережі call-центру (карту мережі call-центру), яку отримано в результаті налаштувань.

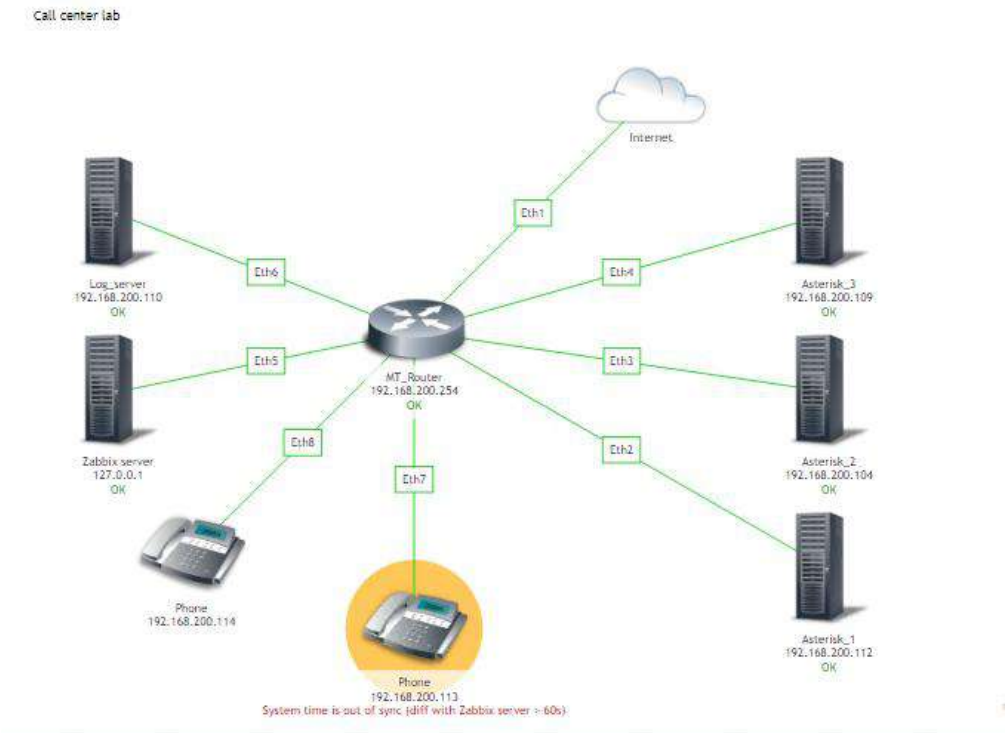


Рис. 8. Карта мережі call-центру

Для прикладу телефон з адресою 192.168.200.113 підсвічено жовтим кольором (в його системі заздалегідь був налаштований інший час). Видно, що даний тригер відображається на мапі як попереджувальний та відображає коротку інформацію адміністратору.

У роботі проведено аналіз та керування балансуванням системою Zabbix. Для цього показано процес тестування навантаження викликами на три сервери Asterisk та реалізація можливостей Zabbix як інструмента балансування навантаження. Розраховано кількість викликів, з якими здатний працювати call-центр.

Тестування навантаження відбувається за допомогою утиліти SIPp, яка є потужною утилітою для створення навантаження на SIP обладнання. Зазвичай SIPp використовується для перевірки відмови системи IP-телефонії, виявлення максимально допустимого навантаження або DDoS-атак (distributed denial-of-service). Сценарій сесії в SIPp описується в XML (extensible markup language) файлі. Можливо скористатися одним із безлічі сценаріїв, що розповсюджуються в комплекті з SIPp, або створити свій сценарій сесії в SIPp.

У роботі створено спеціальний SIP-peer з ім'ям sipp для прийняття викликів від SIPp. В Створено власний сценарій сесії в SIPp, приклад налаштування параметрів якого наведено на рис. 9.

```
sipp@phone:/etc$ cd /etc/sipp-3.3
sipp@phone:/etc/sipp-3.3$ sudo sipp 192.168.200.112 -s 201 -i 192.168.100.10 -d 2h -l 60 -aa -mi 10.10.10.2 -rtp_echo -nd -r
10
```

Рис. 9. Команда для старту відправки викликів SIPp з заданими параметрами

Особливістю сценарію сесії в SIPp є можливість Asterisk авторизувати SIPp не за паролем, а за IP-адресою, яку вказано в полі host. На рис. 10 показано фрагмент роботи утиліти SIPp при заданих параметрах навантаження.

```
1 calls (limit 2)                    Peak was 2 calls, after 30 s
1 Running, 1 Paused, 3 Woken up
0 dead call msg (discarded)          390 out-of-call msg (discarded)
3 open sockets
1309 Total RTP pckts sent            0.000 last period RTP rate (kB/s)

                                 Messages  Retrans   Timeout   Unexpected-Msg
    INVITE ---------->             7         0         0
       100 <----------             7         0         0         0
       180 <----------             7         0         0         0
       200 <----------  E-RTD1      6         0         0         0

       ACK ---------->             6         0
           [ NOP ]
     Pause [   8000ms]             6                             2
           [ NOP ]
     Pause [   1000ms]             4                             0
       BYE ---------->             4         0         0
       200 <----------             4         0         0         0

------ [+|-|*|/]: Adjust rate ---- [q]: Soft exit ---- [p]: Pause traffic -----
```

Рис. 10. Фрагмент роботи утиліти SIPp

В процесі навантаження кількість викликів поступово збільшувалась на 50. Виявилось, що один сервер Asterisk з поточними його параметрами здатний обробити максимум 915 одночасних викликів. Результати тестування показано на графіку (рис. 11).

Виходячи з результатів, отриманих на графіку, для сервера Asterisk було встановлено порогове значення завантаженості CPU в 90 % (йому відповідає 850 викликів). Саме при ньому для уникнення повного перевантаження сервера потрібно запускати в роботу додаткові сервери.

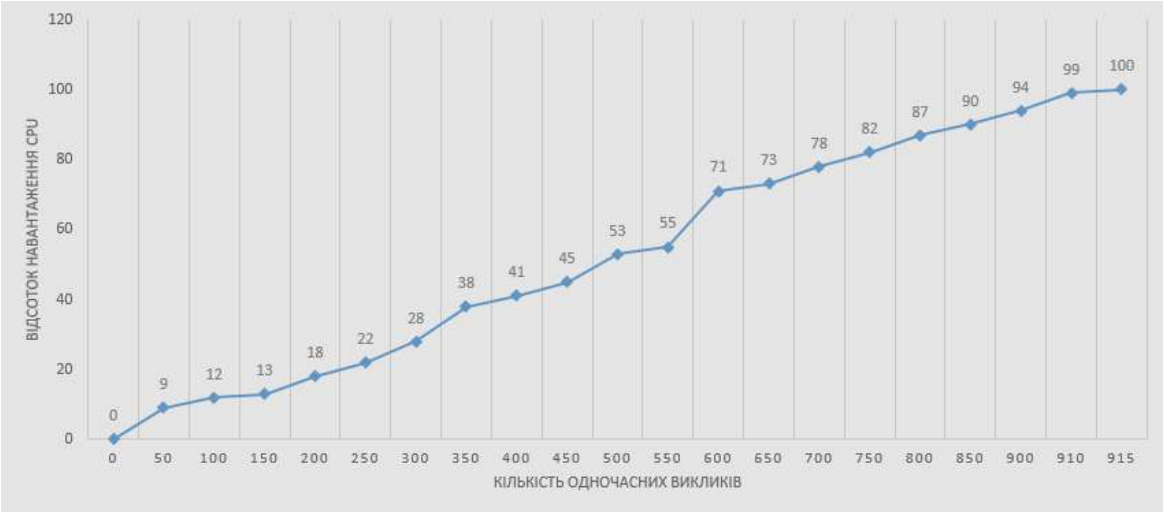


Рис. 11. Графік залежності навантаження CPU від кількості викликів

За допомогою Zabbix запущено процес балансування навантаження на кластер SIP серверів, виходячи з метрик стану CPU та стану SIP-транку. Для того, щоб Zabbix мав змогу перевіряти стан CPU та стан транку, розроблено необхідний скріпт та додано його в існуючі параметри моніторингу host.

В мережі call-центру ВМ Asterisk_1 виступає як в ролі АТС, так і в ролі сервера балансування навантаження [19]. В якості програмного забезпечення, яке відповідає за балансування, на сервері встановлено та налаштовано модуль Kamailio. За замовчуванням його вимкнено.

Для реалізації заданих умов перевірки використано можливості скриптів та тригерів в Zabbix. При виявленні заданого відсотка завантаженості CPU (в даному випадку 90 %) налаштовано скрипт, за яким Zabbix відправляє команду до сервера Asterisk на увімкнення даного модуля. На рис. 12 показано процес створення тригеру, котрий реагує на завантаження CPU host Asterisk.

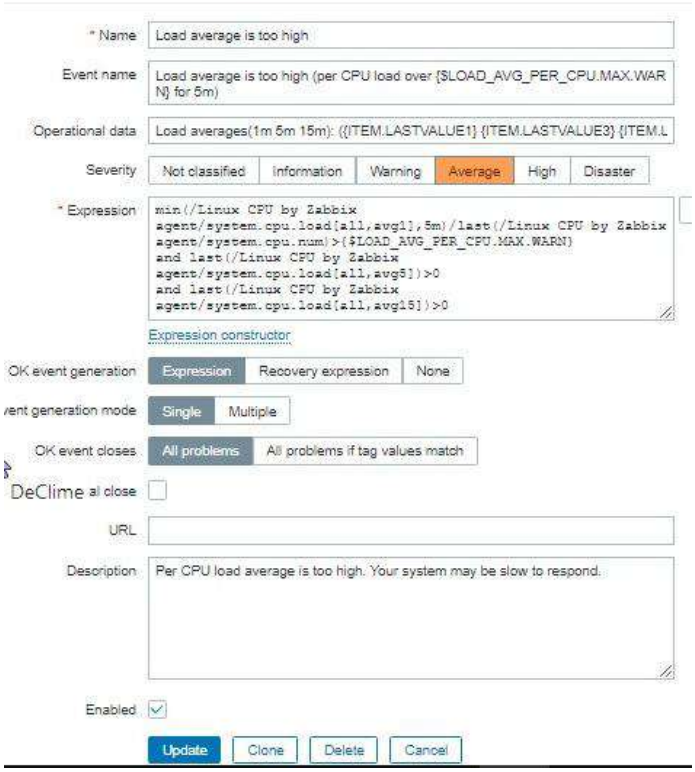


Рис. 12. Процес створення тригеру навантаження CPU

Таким чином, в мережі call-центру налаштовано можливості тестового стенду з використанням Zabbix для балансування навантаження в кластері серверів Asterisk. Було виявлено, що кластер серверів Asterisk здатний обслуговувати 2550 викликів одночасно.

Висновки

Проаналізовано проблеми телекомунікаційних серверів у центрах обробки викликів, що пов`язані з питаннями забезпечення більшої пропускної здатності. Одним з рішень є кластерізація серверів викликів.

Розглянуто особливості процесу моделювання call-центру. Доведено, що для огляду деяких аспектів функціонування мережі необхідна віртуалізація всієї системи або деяких її частин. Це безсумнівно корисно для огляду функціонування мережі, виявлення проблем з трафіком та аналізом параметрів слабких ланок мережі. Проаналізовано особливості діяльності та реалізації мережі call-центру. У якості ядра даної структури розглянуто кластер серверів з використанням віртуальної IP-АТС Asterisk. Такий вибір обумовлено модульною архітектурою Asterisk та можливістю роботи на багатьох ОС. Показано, що одним із основних аспектів якісного функціонування call-центру є балансування навантаження серверів.

Проведено аналіз балансування навантаження з використанням різних алгоритмів та стратегій.

Для розробки та налаштування тестового стенду в роботі створено схему організації call-центру компанії, основним елементом якої виступає кластер SIP-серверів, та мережну модель call-центру. Мережна модель визначає загальну структуру об'єктів, а також їх фізичне розташування, взаємозв'язки та конфігурації. Для створення мережі call-центру та необхідного тестування використано технологію віртуалізації. У якості платформи для налаштування мережі використано гіпервізор VMWare ESXI 6.7 та кліент vCenter. Проведено налаштування АТС Asterisk та створення кластеру серверів. В мережі call-центру ВМ Asterisk_1 виступає як в ролі АТС, так і в ролі сервера балансування навантаження.

Для тестування кластеру серверів викликів, які змодельовано як географічно розподілені сегменти телекомунікаційної мережі, запропоновано використати open source продукт Zabbix. Для візуалізації даних в Zabbix отримано характеристику пропускної здатності для сервера Asterisk_1 та карту мережі, яка фактично являє собою змодельовану структуру мережі call-центру.

Показано процес тестування навантаження викликами на три сервери Asterisk та реалізація можливостей Zabbix для балансування навантаженням. Тестування навантаження відбувається за допомогою утиліти SIPp, яка є потужною утилітою для створення навантаження на SIP обладнання. Створено власний сценарій сесії в SIPp для прийняття викликів. При цьому розрахована кількість викликів, з якими здатний працювати call-центр. Виявлено, що один сервер Asterisk з поточними його параметрами здатний обробити максимум 915 одночасних викликів.

Запущено процес балансування навантаження на кластер SIP серверів, виходячи з метрик стану CPU та стану SIP-транку. Виявлено, що кластер серверів Asterisk здатний обслуговувати 2550 викликів одночасно.

Список літератури:

1. Токар Л.О. Особливості побудови віртуальних АТС // Радіотехніка. 2022. Вип. 208. С. 55 – 64. doi:10.30837/rt.2022.1.208.06.

2. Voxlink. URL: https://voxlink.com/ (дата звернення 14.12.2022).

3. X. Xiao, J. Sun, J. Yang. Operation and maintenance(O&M) for data center: An intelligent anomaly detection approach // Computer Communications. 2021. Vol. 178. pp. 141 – 152. doi:10.1016/j.comcom.2021.06.03.

4. K. Gardner, J. Abdul Jaleel, A. Wickham, S. Doroudi. Scalable load balancing in the presence of heterogeneous servers // Performance Evaluation Review. 2020. Vol. 48, no. 3. pp. 37 – 38. doi:10.1145/3453953.3453961.

5. A. Siokis, K. Christodoulopoulos, N. Pleros, E. Varvarigos. Electro-optic switches based on space switching of multiplexed WDM signals: Blocking vs non-blocking design trade-offs // Optical Switching and Networking. 2017. Vol. 25. pp. 40 – 56. doi:10.1016/j.osn.2017.

6. D. Medhi, K. Ramasamy. Routing and Traffic Engineering using MPLS", in Network Routing. 2018. chapter 23. pp. 766 – 785. doi: 10.1016/B978-0-12-800737-2.00027-2.

7. Ataie Reza Ehsan, Sayed Entezari-Maleki, Etesami Ehsan, Egger Bernhard, Sousa Leonel, Movagharg Ali. Modeling and evaluation of dispatching policies in IaaS cloud data centers using SANs // Sustainable Computing, Informatics and Systems. 2022. Vol. 33. pp. 88 – 102. doi:10.1016/j.suscom.2021.

8. C. Li, Q. Cai, Y. Lou. Optimal data placement strategy considering capacity limitation and load balancing in geographically distributed cloud // Future Generation Computer Systems. 2022. Vol. 127. pp. 142 – 159. doi:10.1016/j.future.2021.08.014.

9. M. Ali, S. Bagchi. Probabilistic normed load monitoring in large scale distributed systems using mobile agents // Future Generation Computer Systems. 2019. Vol. 96. pp. 148 – 167. doi:10.1016/j.future.2019.01.053.

10. Баскаков І. В., Пролетарський А. В., Мельников С. А., Федотов Р. А. IP-телефонія у комп'ютерних мережах // Інтернет-університет інформаційних технологій, 2020. 227 с.

11. О.А. Колтаков, Л.О. Токар. Віртуалізація ресурсів підприємства // Матеріали IV Міжнар. студ. наук. конф. Наука сьогодення: від досліджень до стратегічних рішень. 17 черв. 2022. С.178 – 180.

12. Граннеман Скотт., Linux. Карманный справ очник. Sams Publishing, 2019. 464 с.

13. Pelayo Nuno, Carla Suárez, Eva Suárez. A Diagnosis and Hardening Platform for an Asterisk VoIP PBX // Security and Communication Networks. 2020. pp. 1 – 14. doi:10.1155/2020/8853625.

14. Linux Open Source Software Technologies. URL: https://losst.pro/ (дата звернення 19.01.2023).

15. Uytterhoeven Patrik, Olups Rihards. Zabbix 4 Network Monitoring. Third Edition Packt, 2019. 798 p.

16. Van Baekel Brian, Liefting Nathan. Zabbix 6 IT Infrastructure Monitoring. Packt Publishing, 2022. 506 p.

17. ZABBIX 6.2 Improve your monitoring performance. URL: https://www.zabbix.com /(дата звернення 22.12.2022).

18. A. Pradana, I. Widiasari, R.Efendi. Implementasi Sistem Monitoring Jaringan Menggunakan Zabbix Berbasis SNMP // Security and Communication Networks. 2022. Vol. 19(2). pp. 248 – 262. doi:10.24246/aiti.v19i2.248-262.

19. Andrea Clementia, Emanuele Nataleb, Isabella Ziccardi. Parallel Load Balancing on constrained client-server topologies // Theoretical Computer Science. 2021. Vol. 8952021. pp. 16-33. doi:10.1145/3350755.3400232.

Надійшла до редколегії 15.02.2023

Відомості про авторів:

Токар Любов Олександрівна – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри інфокомунікаційної інженерії ім. В.В. Поповського (ІКІ); Україна; e-mail: liubov.tokar@nure.ua; ORCID: https://orcid.org/0000-0002-7780-1928

Колтаков Олександр Анатолійович – компанія IT-Lance, системний адміністратор, Україна, email: oleksandr.koltakov@nure.ua

Цилюрик Вадим Євгенович – Харківський національний університет радіоелектроніки, магістр кафедри інфокомунікаційної інженерії ім. В.В. Поповського, Україна, email: vadym.tsyliuryk@nure.ua

ABSTRACTS РЕФЕРАТИ

SYSTEMS AND METHODS OF INFORMATION PROTECTION
СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

UDC 004.056.5

Methods and means of static and dynamic code analysis / *A.O. Gapon, V.M. Fedorchenko, O.V. Sievierinov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 7 – 13.

The purpose of the article is to explore the methods and tools used to analyze software code in order to identify errors and potential problems. Static and dynamic code analysis are key processes in software development, as they allow you to detect errors in the early stages of development, reduce the risk of problems later and ensure high quality of the software product. The article discusses various methods and means of static and dynamic code analysis. For each method and tool, examples of their use and advantages and disadvantages are described.

The article will be useful for software developers who want to improve the quality of their products and reduce the risk of problems. It will provide readers with an in-depth understanding of code analysis techniques and tools and help them choose the most appropriate tool for their needs.

Key words: SAST; DAST; HAST; data–flow analysis; Symbolic execution.

1 fig. Ref: 4 items.

УДК 004.056.5

Методи та засоби статичного та динамічного аналізу коду / *А.О. Гапон, В.М. Федорченко, О.В. Сєвєрінов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 7 – 13.

Метою статті є дослідження методів та інструментів, які використовуються для аналізу програмного коду з метою виявлення помилок та потенційних проблем. Статичний та динамічний аналіз коду є ключовими процесами в розробці програмного забезпечення, оскільки дозволяють виявити помилки на ранніх етапах розробки, зменшити ризик виникнення проблем у пізніший час та забезпечити високу якість програмного продукту. Розглядаються різні методи та засоби статичного та динамічного аналізу коду. Для кожного методу та інструменту наводяться приклади їх використання та описуються переваги та недоліки.

Стаття буде корисна розробникам програмного забезпечення, які хочуть покращити якість своїх продуктів та зменшити ризик виникнення проблем. Вона надасть читачам глибоке уявлення про методи та інструменти аналізу коду та допоможе їм вибрати найбільш підходящий засіб для їхніх потреб.

Ключові слова: SAST; DAST; HAST; аналіз потоку даних; символьне виконання.

Іл. 1. Бібліогр.: 4 назв.

UDC 004.056.5

Side-channel attacks on CRYSTALS-KYBER, countermeasures and comparison with SKELYA (DSTU 8961-2019) / *Ya.A. Derevianko, I.D. Gorbenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 14 – 29.

Although the mathematical problems used in post-quantum cryptography algorithms appear to be mathematically secure, a class of attacks known as side-channel attacks may prove to be a threat to the security of such algorithms. Side-channel attacks affect the hardware on which the cryptographic algorithm runs, they are not attacks on the algorithm itself.

The good news is that side-channel analysis on new post-quantum cryptographic algorithms started early, even before the algorithms were standardized, given that older algorithms still face side-channel problems.

Kyber is a lattice-based post-quantum algorithm based on the complexity of the M-LWE problem. Kyber offers a secure public key encryption (PKE) scheme against a chosen plaintext attack (CPA) and a secure key encapsulation mechanism against a chosen ciphertext attack (CCA).

This paper provides a study of side-channel and fault-injection attacks on lattice-based schemes, with focus on the Kyber (KEM).

Considering the wide range of known attacks, the protection of the algorithm requires the implementation of individual countermeasures. The paper presents and tests a number of countermeasures capable of providing/improving protection against existing SCA/FIA for Kyber KEM.

The obtained results show that the presented countermeasures incur a reasonable performance cost. Therefore, the use of special countermeasures in real implementations of lattice-based schemes, either alone or as an augmentation of general countermeasures, is necessary.

Key words: post-quantum algorithm; side channels; Skelya algorithm; Crystals-Kyber; NIST.

3 tabl. 9 fig. Ref: 28 items.

УДК 004.056.5

Атаки бічними каналами на CRYSTALS-KYBER, контрзаходи та порівняння з алгоритмом СКЕЛЯ (ДСТУ 8961-2019) / *Я.А. Дерев'янко, І.Д. Горбенко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 14 – 29.

Незважаючи на те, що математичні задачі, що застосовуються у алгоритмах постквантової криптографії, здаються математично безпечними, клас атак, відомий як атаки бічними каналами, може виявитися загрозою безпеці таких алгоритмів. Атаки бічними каналами стосуються апаратного забезпечення, на якому працює алгоритм криптографії, вони не є атаками на сам алгоритм.

Хорошою новиною є те, що аналіз бічних каналів на нових алгоритмах постквантової криптографії почався завчасно, ще до стандартизації алгоритмів, враховуючи, що старіші алгоритми все ще стикаються з проблемами бічних каналів.

Kyber – це постквантовий алгоритм на основі решітки, заснований на складності задачі M-LWE. Kyber пропонує безпечну схему шифрування з відкритим ключем (PKE) проти атаки вибраного відкритого тексту (CPA) і захищений механізм інкапсуляції ключів проти атаки вибраного шифротексту (CCA).

У роботі надається дослідження атак бічними каналами і атак із впровадженням помилок на схеми на основі решітки, з основним акцентом на Kyber (KEM).

Враховуючи широкий спектр відомих атак, захист алгоритму потребує впровадження індивідуальних контрзаходів. Представлено і протестовано ряд контрзаходів, здатних забезпечити/покращити захист відносно існуючих SCA/FIA для Kyber KEM.

Отримані результати показують, що представлені контрзаходи спричиняють розумні витрати на продуктивність. Тому використання спеціальних контрзаходів у реальних реалізаціях схем на основі решітки, або окремо, або як підсилення загальних контрзаходів, є необхідним.

Ключові слова: постквантовий алгоритм; бічні каналі; алгоритом Скеля; Crystals-Kyber; NIST

Табл. 3. Іл. 9. Бібліогр.: 28 назв.

UDC 004.056.5

An overview of threats to data security and integrity in cloud computing / *M.V. Yesina, A.A. Kravchenko, S.O. Kravchenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 30 – 35.

Cloud computing has become an integral part of our lives, and today it is used almost everywhere. In general, cloud computing is a concept of providing IT resources in the form of services. There are two cloud computing models: deployment models, which differ in the type of cloud management and access to the cloud and the level of security, and service models, which differ in the level of service provision, which affects, among other things, the level of responsibility of the service provider and the consumer. Cloud services began to gain popularity in 2009, and the demand for them has grown exponentially every year. They became especially popular during the pandemic in 2019, when people had to stay at home without interrupting their work processes, and now, in post-covid times, they also remain popular due to their convenience, high availability, easy scalability and cost savings. Due to the widespread use of cloud computing services, a high level of security is required. Unfortunately, the popularity of cloud computing has its drawbacks – in addition to the fact that it is more difficult to monitor the security of a remote environment than the security of a local computer, there are many other threats. In today's reality, people use cloud computing technologies in large volumes, for example, at work, for personal purposes, etc., as they have great trust in these technologies. This is the reason why it is necessary to maintain a high level of security and constantly improve it. Cloud computing security threats are usually divided into confidentiality, integrity, and availability threats. To prevent the loss of confidential information, service providers must ensure its integrity. Users want to be sure that their data will not fall into the hands of an intruder or third-party services. Therefore, this article discusses the most common threats to data security and integrity in cloud computing and the existing methods that prevent these vulnerabilities and possible problems at different levels and with the help of different tools.

Key words: cloud computing; cloud computing security; integrity; integrity threats; security threats; integrity methods.

Ref: 29 items.

УДК 004.056.5

Огляд загроз безпеці та цілісності даних у хмарних обчисленнях / *М.В. Єсіна, А.А. Кравченко, С.О. Кравченко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 30 – 35.

Хмарні обчислення стали невід'ємною частиною нашого життя, і сьогодні їх використовують майже усюди. Взагалі, хмарні обчислення являють собою концепцію надання IT ресурсів у вигляді послуг. Розрізняють дві моделі хмарних обчислень – розгортання, що відрізняються за типом управління хмарою та доступу до неї та рівнем безпеки, і моделі обслуговування, які відрізняють за рівнем надання послуг – це впливає, окрім іншого, на рівень відповідальності постачальника і споживача послуг. Хмарні сервіси почали набувати своєї популярності у 2009 р., і з кожним роком попит на них лише зростав у геометричній прогресії. Особливо великої популярності вони набули під час пандемії у 2019 р., коли люди повинні були залишатись вдома, не зупиняючи при цьому робочих процесів, і зараз, в постковідні часи, вони теж залишаються популярними через їх зручність, високу доступність, легку масштабованість та економію коштів. Через розповсюджену експлуатацію послуг хмарних обчислень виникає необхідність високого рівня безпеки. На жаль, у великій популярності є свої мінуси – окрім того, що слідкувати за безпекою віддаленого середовища складніше, ніж за збереженістю локального комп'ютера, існує ще безліч загроз. Люди використовують технології хмарних обчислень у великих обсягах, наприклад, на роботі, в особистих цілях та інше, так як вони мають велику довіру до цих технологій. Саме це є

причиною необхідності підтримувати безпеку на високому рівні і весь час її вдосконалювати. Загрози безпеки хмарних обчислень зазвичай поділяють на загрози конфіденційності, цілісності та доступності. Щоб запобігти втраті довіреної інформації провайдери послуг мають забезпечити її цілісність. Користувачі хочуть бути впевнені в тому, що їх дані не попадуть до рук зловмисника або на сторонні сервіси. Тож, дана стаття розглядає найбільш поширені загрози безпеки і цілісності даних в хмарних обчисленнях та існуючі методи, що на різних рівнях та за допомогою різних інструментів запобігають цим вразливостям та можливим проблемам.

Ключові слова: хмарні обчислення; безпека хмарних обчислень; цілісність; загрози цілісності; загрози безпеки; методи забезпечення цілісності.

Бібліогр.: 29 назв.

UDC 004.056.5

Models of threats to cloud services / *M.V. Yesina, V.V. Onoprienko, A.V. Tolok* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 36 – 41.

Cloud services have become popular due to their advantages over traditional computing. The cloud provides remote access to software, hardware, and other services. This has allowed companies to be more productive and enabled remote work. Cloud services have fewer hardware and infrastructure requirements, which reduces the cost of maintaining and supporting information technology. The future success of organizations will depend, not least, on the extent to which they implement cloud computing in their operations. According to forecasts, spending on cloud IT technologies will continue to grow and in 2025 will exceed spending on traditional IT technologies. Security of cloud services is becoming a critical issue as more and more companies complete their digital transformation. Despite the many benefits, cloud services also face their own security threats and challenges. Since cloud services store and process a significant amount of sensitive information, a cloud breach can lead to data leaks that can hinder business development and cause significant damage to a company's reputation. There are risks associated with the unavailability of cloud services in case of technical problems and dependence on external providers. Therefore, companies should carefully assess potential threats and take appropriate measures to protect their data and business in general when using cloud services. There are many methods to help determine how prepared your organization is for the growing number of threats. Threat modeling is one of the methods for predicting and preparing for possible threats. Using modeling frameworks allows you to allocate resources and plan possible actions during an attack. There are many modeling frameworks available, but it is important to remember that these frameworks have their advantages and disadvantages, so the choice depends on the context and needs of a particular system. Analyzing, evaluating, and comparing existing methods for modeling and protecting against threats in cloud services is the main objective of this article.

Key words: cloud computing; threats of cloud services; modeling systems.

3 tabl. 1 fig. Ref: 11 items.

УДК 004.056.5

Моделі загроз хмарних послуг / *М.В. Єсіна, В.В. Онопрієнко, А.В. Толок* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 36 – 41.

Хмарні послуги стали популярними завдяки своїм перевагам над традиційними обчисленнями. Хмара дає можливість отримувати віддалений доступ до програмного забезпечення, обладнання та інших послуг. Це дозволило компаніям бути більш продуктивними і зробило можливою віддалену роботу. Хмарні послуги мають менше вимог до обладнання та інфраструктури, що знижує витрати на утримання та підтримку інформаційних технологій. Майбутній успіх організацій буде залежати від обсягу впровадження хмарних обчислень у свою роботу. За прогнозами витрати на хмарні ІТ-технології будуть зростати та у 2025 р. будуть перевищувати витрати на традиційні ІТ-технології. Безпека хмарних послуг стає критичною проблемою, оскільки все більше компаній завершують свою цифрову трансформацію. Незважаючи на велику кількість переваг, хмарні послуги також стикаються зі своїми власними загрозами та викликами безпеки. Оскільки хмарні послуги зберігають та обробляють значну кількість конфіденційної інформації, злам хмари може призвести до витоку даних, що може стати на шляху розвитку бізнесу та завдати значної шкоди репутації компанії. Існують ризики, зв'язані з недоступністю хмарних послуг у випадку технічних проблем та залежності від зовнішніх провайдерів. Тому, підприємства повинні ретельно оцінювати потенційні загрози та приймати відповідні заходи для захисту своїх даних та бізнесу в цілому при використанні хмарних послуг. Існує багато методів, які допоможуть визначити, наскільки ваша організація готова до зростаючої кількості загроз. Моделювання загроз один з методів прогнозування та підготовки до можливих загроз. Використання фреймворків моделювання дозволяє розподілити ресурси та спланувати можливі дії під час атаки. Існує багато фреймворків моделювання, але важливо пам’ятати, що ці фреймворки мають свої переваги та недоліки, тому вибір залежить від контексту та потреб конкретної системи. Аналіз, оцінка та порівняння існуючих методів моделювання та захисту від загроз у хмарних послугах є основною метою цієї статті.

Ключові слова: хмарні обчислення; загрози хмарних послуг; системи моделювання.

Табл. 3. Іл. 1. Бібліогр.: 11 назв.

UDC 004.056.5

Scientific and methodological bases of analysis, evaluation and results of comparison of existing and promising (post-quantum) asymmetric cryptographic primitives of electronic signature, protocols of asymmetric encryption and key encapsulation protocols / *Yu.I. Gorbenko, M.V. Yesina, V.A. Ponomar, I.D. Gorbenko, E.Yu. Kapt'ol* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 42 – 65.

Currently, world civilization is taking significant steps in science and practice related to quantum calculations. Significant steps are being taken to achieve the competitive advantage of countries in the field of quantum information science and the practice of introducing quantum technologies. Scientific and practical research is first aimed at reducing the risks related to quantum computers on cybersecurity, economic and national security. Although the full range of quantum computers is still unknown, it is obvious that further technological and scientific leadership of states will at least partially depend on the country's ability to maintain a competitive advantage in quantum computing and quantum information science. However, along with the potential advantages, quantum calculations are likely to cause significant risks about economic and national security. Specific actions are determined that technologically developed states that begin a long-term process of transferring vulnerable computer systems to quantum-resistant cryptography. An important problem in cryptology is to analyze ways of reducing risks for vulnerable cryptographic systems and the state of their development, adoption and implementation at the international and national levels of post-quantum standards of asymmetric cryptotransformations of electronic signatures (ES), asymmetric ciphers (AC) and key encapsulation protocols (KEP). Therefore, the processes of reducing risks for vulnerable existing standardized cryptographic systems and determining the directions of development of mathematical methods and the study of the prospects for their application in the creation of standardized AC, KEP and ES are significantly significant. They are reduced to justification and definition of mathematical methods and mechanisms that will create promising (post-quantum) standardized AC, KEP and ES. The course of analysis, evaluation and results of comparison of existing and post-quantum asymmetric cryptotransformations of AC, KEP and ES, and standardization at the international and national levels, including for transitional and post-quantum periods, are the main objective of this article.

Key words: post-quantum cryptography; quantum computer; electronic signature; asymmetric encryption protocol; key encapsulation protocol.

9 tabl. 10 fig. Ref: 23 items.

УДК 004.056.5

Науково-методичні основи аналізу, оцінки та результати порівняння існуючих та перспективних (постквантових) асиметричних криптографічних примітивів електронного підпису, протоколів асиметричного шифрування та протоколів інкапсуляції ключів / *Ю.І. Горбенко, М.В. Єсіна, В.А. Пономар, І.Д. Горбенко, Є.Ю. Каптьол* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 42 – 65.

Наразі робляться суттєві кроки щодо досягнення конкурентної переваги країн у галузі квантової інформаційної науки та практики впровадження квантових технологій. Наукові та практичні дослідження направлені на зниження ризиків, що пов’язані з квантовими комп’ютерами щодо кібербезпеки, економічної та національної безпеки. Хоча повний спектр застосування квантових комп’ютерів ще невідомий, проте очевидно, що подальше технологічне та наукове лідерство держав, принаймні частково, буде залежати від здатності країни підтримувати конкурентну перевагу в галузі квантових обчислень та квантової інформаційної науки. Але, поряд з потенційними перевагами, квантові обчислення ймовірно будуть викликати щодо економічної та національної безпеки значні ризики. Визначаються конкретні дії, які мають зробити технологічно розвинені держави, що розпочинають багаторічний процес переведення вразливих комп’ютерних систем на квантово-стійку криптографію. Важливою проблемою в криптології є аналіз шляхів зниження ризиків для вразливих криптографічних систем та стану їх розроблення, прийняття та впровадження на міжнародному та національному рівнях постквантових стандартів асиметричних криптоперетворень електронних підписів (ЕП), асиметричних шифрів (АСШ) та протоколів інкапсуляції ключів (ПІК). Тому процеси зниження ризиків для вразливих існуючих стандартизованих криптографічних систем та визначення напрямків розвитку математичних методів та дослідження перспектив їх застосування в ході створення стандартизованих АСШ, ПІК та ЕП є суттєво значимими. Вони зводяться до обґрунтування та визначення математичних методів та механізмів, які дозволять створити перспективні (постквантові) стандартизовані АСШ, ПІК та ЕП. Хід аналізу, оцінки та результати порівняння існуючих та постквантових асиметричних криптопримітивів АСШ, ПІК та ЕП та стандартизації на міжнародному та національному рівнях, в тому числі для перехідного та постквантового періодів, є основною метою статті.

Ключові слова: постквантова криптографія; квантовий комп’ютер; електронний підпис; протокол асиметричного шифрування; протокол інкапсуляції ключів.

Табл. 9. Іл. 10. Бібліогр.: 23 назв.

UDC 004.056.5

Security analysis of promising key encapsulation mechanisms in the core-SVP model / *S.O. Kandiy* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 66 – 84.

The study of key encapsulation mechanisms on structured lattices is one of the important directions in modern post-quantum cryptography, as many mechanisms are either already standardized (DSTU 8961:2019 "Skelya") or are promising candidates for standardization (CRYSTALS-Kyber). Estimating the complexity of lattice reduction for cryp-

tographic schemes is an old problem. Asymptotic estimates differ greatly from experimental values, therefore, a number of heuristic methods were developed to solve practical problems. The coreSVP model is a standard means of assessing the security of cryptographic schemes on lattices. The purpose of the work is to analyze the encapsulation mechanisms of DSTU 8961:2019 "Skelya" and CRYSTALS-Kyber keys in the coreSVP model. The analysis was performed using two popular heuristics – GSA (Geometric Series Assumption) and the Chen-Nguyen simulator. The analysis showed that the Chen-Nguyen simulator gives slightly lower estimates than the GSA heuristic. As a result of the analysis, it was found that 8961:2019 The "Skelya" and CRYSTALS-Kyber in the coreSVP model for classical computers have slightly lower than declared security values, but for quantum computers the key encapsulation mechanisms provide the declared security levels. Note that during the analysis, the accuracy of the GSA heuristics and the Chen-Nguyen simulator were analyzed separately. Examples of parameters for which heuristics do not give sufficiently accurate results are given. The performed analysis does not take into account the algebraic structure of lattices used in 8961:2019 "Skelya" and CRYSTALS-Kyber. The inclusion of an algebraic structure in the analysis is a further direction of work. The use of simulators is a promising direction, however, more accurate simulators that take into account the structuring of LWE and NTRU arrays are needed.

Key words: post-quantum cryptography; algebraic lattices; DSTU 8961:2019 "Skelya", CRYSTALS-Kyber; BKZ; SVP; core-SVP.

12 tab. 13 fig. Ref: 26 items.

УДК 004.056.5

Аналіз безпеки перспективних механізмів інкапсуляції ключів у моделі core-SVP / *С.О. Кандій* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 66 – 84.

Дослідження механізмів інкапсуляції ключів на структурованих решітках є одним з важливих напрямів у сучасній постквантовій криптографії, оскільки багато механізмів або вже стандартизовані (ДСТУ 8961:2019 "Скеля"), або є перспективними кандидатами на стандартизацію (CRYSTALS-Kyber). Оцінка складності редукції решіток для криптографічних схем є давньою проблемою. Асимптотичні оцінки сильно відрізняються від експериментальних значень, тому для вирішення практичних задач був розроблений ряд евристичних методів. Модель coreSVP є стандартним засобом оцінки безпеки криптографічних схем на решітках. Мета роботи – аналіз механізмів інкапсуляції ключів ДСТУ 8961:2019 "Скеля" та CRYSTALS-Kyber у моделі coreSVP. Аналіз виконаний з використанням двох популярних евристик – GSA (Geometric Series Assumption) та симулятора Чена – Нгуєна. У результаті аналізу показано, що симулятор Чена – Нгуєна дає нижчі оцінки, ніж евристика GSA. В результаті аналізу виявлено, що 8961:2019 "Скеля" та CRYSTALS-Kyber у моделі coreSVP для класичних комп'ютерів мають значення безпеки дещо нижчі, ніж заявлені, проте для квантових комп'ютерів механізми інкапсуляції ключів забезпечують заявлені рівні безпеки. Зауважимо, що під час аналізу окремо проаналізовано точність евристик GSA та симулятора Чена – Нгуєна. Наведено приклади параметрів, за яких евристики не дають достатньо точного результату. Аналіз не враховує алгебраїчну структуру решіток, що використовуються у 8961:2019 "Скеля" та CRYSTALS-Kyber. Включення алгебраїчної структури до аналізу є подальшим напрямком роботи. Використання симуляторів є перспективним напрямком, проте необхідні більш точні симулятори, що враховують структурованість LWE та NTRU решіток.

Ключові слова: постквантова криптографія; алгебраїчні решітки; ДСТУ 8961:2019 "Скеля"; CRYSTALS-Kyber; BKZ; SVP; core-SVP.

Табл. 12. Іл. 13. Бібліогр.: 26 назв.

UDC 621.396:004.056

Ways to improve the efficiency of methods and means of counteracting unauthorized speech recording and their comparative analysis / *A.N. Oleynikov, V.A. Pulavskyi, O.H. Bilotserkivets* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 85 – 89.

The features of using ultrasonic, electromagnetic and acoustic methods to counteract unauthorized recording of speech on sound recording devices are considered. The advantages and unused possibilities of the methods are noted. Ways to increase the effectiveness of the ultrasonic method of counteraction consist in using the two-frequency ultrasonic method, in positioning the device for emitting ultrasonic vibrations no more than two meters from the target, in using high-intensity ultrasonic vibrations, in using the ultrasonic method together with the acoustic method, and others. Ways to increase the effectiveness of the electromagnetic method consist in using amplitude-pulse modulation of the interference signal, narrowly directed antennas, a device with a maximum allowable power, and others; Ways to increase the effectiveness of the acoustic method consist in creating speech-like interference from the speech of the interlocutor (synchronized with speech pauses), reducing the distance between the source of interference radiation and the source of speech in relation to the distance between the interlocutors due to organizational protection measures. (adapted acoustic method).

The most promising is the adapted acoustic method having the greatest potential for guaranteed countermeasures, subject to the requirements for its optimal use.

Analyzing the results of the experiment, we conclude that the adapted acoustic method is the most effective. The effectiveness of the method confirms the suppression range from 1.8 m to 3.5 m depending on the device, for comparison, in the ultrasonic experiment, one of the indicators was 0.9 m, and in the electromagnetic experiment, even

0 m. The formation of interference through the acoustic channel is higher the described method ensures the universality of the proposed method to the type of device for suppressing unauthorized speech recording, regardless of the underlying method – electromagnetic, ultrasonic or acoustic.

Key words: speech; record; acoustic; electromagnetic; ultrasonic; method; means; suppression; efficiency.

6 fig. Ref: 2 items.

УДК 621.396:004.056

Шляхи підвищення ефективності методів та засобів протидії несанкціонованому запису мови та їх порівняльний аналіз / *А.М. Олейніков, В.А. Пулавський, О.Г. Білоцерківець* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 85 – 89.

Розглядаються особливості застосування ультразвукового, електромагнітного та акустичного методів протидії несанкціонованому запису мовлення на звукозаписні пристрої. Відзначаються переваги та невикористані можливості методів. Запропоновано шляхи підвищення ефективності ультразвукового методу протидії шляхом: використання двочастотного ультразвукового методу, розташування пристрою випромінювання ультразвукових коливань не більше двох метрів до цілі, застосування ультразвукових коливань великої інтенсивності, використання ультразвукового методу разом з акустичним та інші. Підвищення ефективності електромагнітного метода шляхом передбачає використання амплітудно-імпульсної модуляції завадового сигналу, використання вузькоспрямованих антен, використання пристрою з максимально допустимою потужністю та інше. Підвищення ефективності акустичного методу шляхом передбачає створення мовоподібних завад з мови співрозмовника (синхронізованих з мовними паузами), зменшення відстані між джерелом завадового випромінювання і джерелом мови по відношенню до відстані між співрозмовниками за рахунок організаційних заходів захисту.(адаптований акустичний метод).

Найбільш перспективним є адаптований акустичний метод, що має найбільші можливості гарантованої протидії, при дотриманні вимог для оптимального його застосування.

Аналізуючи результати експерименту, отримуємо висновок, що адаптований акустичний метод є найбільш ефективним. Ефективність методу підтверджує дальність подавлення, яка склала від 1,8 до 3,5 м в залежності від пристрою, для порівняння в ультразвуковому експерименті один із показників склав 0,9 м, а в електромагнітному навіть 0 м. Формування перешкоди по акустичному каналу описаним способом забезпечує універсальність запропонованого методу до будь-якого типу пристрою придушення несанкціонованого запису мови незалежно від методу, що лежить в його основі, – електромагнітного, ультразвукового чи акустичного.

Ключові слова: мова; запис; акустичний; електромагнітний; ультразвуковий; метод; засіб; подавлення; ефективність.

Іл. 6. Бібліогр.: 2 назв.

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS
ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

UDC 621.391:537.86:519

Applying factorization to increase the resolving ability of the parametric estimation of the power spectral density / *V.A. Tikhonov, V.M. Bezruk* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. № 212. P. 90 – 101.

We consider a possibility of the factorization of parametric spectral power density (PSM) estimation of a random process based on autoregressive linear prediction model to increase the spectrum resolution. Factorization refers to the decomposition of the multimode PSM into simpler single-mode components. Factorization makes it possible not only to decompose a complex multimode PSM into simple single-mode components, but also to analyze more accurately the low-, medium- and high-frequency components of the SPM of a random process. The main attention is paid to the study of the problem of increasing the resolving power of SPM estimation by its factorization by the Yule-Walker and Berg method.

Key words: random process; spectral power density; spectral resolution; factorization; autoregressive model.

15 fig. Ref: 18 items.

УДК 621.391:537.86:519

Застосування факторизації для підвищення роздільної здатності параметричної оцінки спектральної щільності потужності / *В.А. Тихонов, В.М. Безрук* / Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 90 – 101.

Розглянуто можливості факторизації параметричної оцінки спектральної щільності потужності (СЩП) випадкового процесу на основі моделі авторегресії лінійного передбачення для підвищення роздільної здатності спектра. Під факторизацією розуміється розкладання багатомодової СЩП на простіші одномодові складові. Факторизація дає змогу не тільки розкласти складну багатомодову СЩП на простіші одномодові складові, а й точніше аналізувати низькочастотні, середньочастотні та високочастотні складові СПМ випадкового процесу. Основну увагу приділено дослідженню задачі підвищення роздільної здатності оцінки СЩП шляхом її факторизації методом Юла – Уокера і Берга.

Ключові слова: випадковий процес; спектральна густина потужності; роздільна здатність спектра; факторизація; модель авторегресії.

Іл. 15. Бібліогр.: 18 назв.

UDC 537.868

Influence of ferrimagnetic resonance on conversion of electromagnetic energy by a system consisting of two cylinders into a mechanical one / *G.L. Komarova* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 102 – 114.

This work presents the analysis of the integral equation of macroscopic electrodynamics, the solution of the problem of diffraction of a plane polarized electromagnetic wave on a system consisting of two ferrite cylinders of radii corresponding to spatial resonance ($R \leq 0.1 \cdot \lambda o$, λo is the wavelength in free space). The electromagnetic fields inside the first (second) cylinder are presented as the sum of the fields of the solitary first (second) cylinder, a plane-parallel wave falls on it and is scattered by the second (first) solitary cylinder. The expressions for the fields satisfy Maxwell's equations, boundary conditions for two cylinders, and integral equations. The influence of the distance between the centers of the cylinders on the strength of the electromagnetic field in the middle of the ferrite cylinders has been studied. It has been established that in a system consisting of two cylinders, a group resonance arises due to their mutual arrangement in space. The transformation of microwave energy on a system consisting of two ferrite cylinders depending on the value of their resonant radii at ferrimagnetic resonance has been studied. An inhomogeneous electromagnetic wave created by propagating in free space with a power flux density of 622 kW/m^2 and a length of 3.2 cm reflected from a metal screen acts on a system of ferrite cylinders, the total length of which is 1.28 m, and the resonant radius is 3.863 mm with a force equal to 4 N. The results of studying the phenomenon of diffraction on a system consisting of two ferrite cylinders show that the total force with which the inhomogeneity of a standing electromagnetic wave acts on two cylinders is 2.8 times greater than the force acting on a solitary cylinder.

Key words: cylinder system; ferrimagnetic resonance; electromagnetic energy; conversion; mechanical energy.

1 tabl. 4 fig. Ref: 15 items.

УДК 537.868

Вплив феримагнітного резонансу на перетворення мікрохвильової енергії системою, що складається з двох циліндрів в механічну / *Г.Л. Комарова* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 102 – 114.

Аналізується інтегральне рівняння макроскопічної електродинаміки, вирішується задача дифракції плоскої поляризованої електромагнітної хвилі на системі, що складається з двох феритових циліндрів, величини радіусів яких відповідають просторовому резонансу ($R \leq 0{,}1 \cdot \lambda o$, λo _ довжина хвилі у вільному просторі). Електромагнітні поля всередині першого (другого) циліндра представлені у вигляді суми полів відокремленого першого (другого) циліндра, на який падає плоско паралельна хвиля і розсіяна другим (першим) відокремленим циліндром. Вирази для полів задовольняють рівнянням Максвелла, граничним умовам для двох циліндрів та інтегральним рівнянням. Досліджено вплив відстані між центрами циліндрів на величину напруженості електромагнітного поля в середині феритових циліндрів. Встановлено, що в системі, що складається з двох циліндрів, виникає груповий резонанс, обумовлений взаємним розташуванням їх у просторі. Досліджено перетворення мікрохвильової енергії в механічну системою, що складається з двох феритових циліндрів в залежності від величини їх резонансних радіусів при феримагнітному резонансі. Неоднорідна електромагнітна хвиля, створена у вільному просторі з щільністю потоку потужності кВт/м^2 і довжиною 3,2 см, та відбита від металевого екрану, діє на систему феритових циліндрів, загальна довжина яких дорівнює 1,28 м, а резонансний радіус дорівнює 3,863 мм, з силою, що дорівнює 4 Н. Результати дослідження дифракційного явища на системі, що складається з двох феритових циліндрів, показують, що загальна сила, з якою неоднорідність стоячої електромагнітної хвилі діє на два циліндри, в 2,8 рази більше сили, що діє на відокремлений циліндр.

Ключові слова: система циліндрів; феримагнітний резонанс; електромагнітна енергія; перетворення; механічна енергія.

Табл. 1. Іл. 4. Бібліогр.: 15 назв.

UDC 621.357

Combined heat conductive boards with polyimide dielectrics / *V.M. Borshchov, O.M. Listratenko, M.A. Protsenko, I.T. Tymchuk, O.V. Kravchenko, O.V. Syddia, I.V. Borshchov, M.I. Slipchenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 115 – 126.

Heat-conductive properties of thin heat-conductive polyimide dielectrics have been studied and their thermal resistances have been calculated. Possibility of creating combined printed circuit boards on heat-conductive bases with reduced thermal resistances of polyimide dielectrics from ~ 0.2 to ~ 0.04 °C /W is confirmed.

Design parameters and thermal properties of the combined boards with thin polyimide (PI) dielectrics for receivers of concentrated solar radiation are studied. Possibility of providing thermal resistances of PI dielectrics not exceeding 0.43 °C/W has been confirmed.

Technical solutions of volumetric light-emitting diode (LED) modules on combined heat-conductive boards, which are 3D-holders-heat sinks, made in the form of single heat-conductive light-reflecting mirrored element, are studied. High thermal characteristics of the modules were achieved due to increase in the area of heat sink holders by more than 2.5 – 3 times compared to flat-type LED modules.

Scientific and technical sources were analyzed for selection of modern polyimide materials intended for development and manufacture of combined boards on heat-conductive bases with dielectrics made of polyimide films with increased thermal conductivity up to 0.36 – 0.75 W/(m•K).

Potential possibility of creating effective combined printed circuit boards on heat-conductive basis, including those that can be bent, is confirmed using modern industrially manufactured thin heat-conductive PI films with heat-sealable thermoplastic coatings that provide the value of total thermal resistance of boards from 1.5 up to 2.8 °C•cm^2/W.

Key words: combined printed circuit boards; heat-conductive polyimide composites; solar concentrator receivers; LED modules.

2 tab. 10 fig. Ref: 19 items.

УДК 621.357

Комбіновані теплопровідні плати з діелектриками з поліміду / *В.М. Борщов, О.М. Лістратенко, М.А. Проценко, І.Т. Тимчук, О.В. Кравченко, О.В. Суддя, І.В. Борщов, М.І. Сліпченко* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 115 – 126.

Досліджено теплопровідні властивості тонких теплопровідних поліімідних діелектриків і виконано розрахунок їхніх теплових опорів. Підтверджено можливість створення комбінованих друкованих плат на теплопровідних основах зі зменшеними тепловими опорами поліімідних діелектриків від ~ 0,2 до ~ 0,04 °С/Вт.

Досліджено конструктивні параметри та теплові властивості комбінованих плат з тонкими поліімідними (ПІ) діелектриками для приймачів концентрованого сонячного випромінювання. Підтверджено можливість забезпечення теплових опорів ПІ діелектриків, що не перевищують 0,43 С/Вт.

Досліджено технічні рішення об'ємних світлодіодних (СД) модулів на комбінованих теплопровідних платах, що являють собою 3D-тримачі-тепловідводи, які виконані у вигляді єдиного теплопровідного світловідбивного дзеркалізованого елементу. Досягнуто високі теплові характеристики модулів завдяки збільшенню площі тримачів-тепловідводів більш ніж у 2,5 – 3 рази порівняно з СД модулями плоского типу.

Проаналізовано науково-технічні джерела для вибору сучасних поліімідних матеріалів, які призначені для розроблення та виготовлення комбінованих плат на теплопровідних основах із діелектриками з поліімідних плівок зі збільшеною теплопровідністю до 0,36 – 0,75 Вт/(м•К).

Підтверджено потенційну можливість створення ефективних комбінованих друкованих плат на теплопровідній основі, зокрема й тих, які можна згинати, із застосуванням сучасних промислово виготовлених тонких теплопровідних ПІ плівок з термозварювальними термопластичними покриттями, що мають змогу забезпечити значення повних теплових опорів плат від 1,5 до 2,8 °С•см^2/Вт.

Ключові слова: комбіновані друковані плати; теплопровідні поліімідні композити; концентраторні сонячні приймачі; світлодіодні модулі.

Табл. 2. Іл. 10. Бібліогр.: 19 назв.

UDC 621.383

Experimental studies of a lidar emitter built according to the oscillator-amplifier scheme / *A.I. Tsopa, A.A. Zarudny* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 127 – 133.

The results of experimental studies of the energy characteristics of the leader transmitter built according to the lamp-pumped organic-dye oscillator-amplifier scheme are presented. When constructing the lidar emitter according to the oscillator-amplifier scheme under conditions of constant pump density, the problem arises of choosing the ratio between the length of the active element of the generator and the length of the active medium of the traveling wave amplifier, which ensures the maximum efficiency of the entire emitter. The main objective of the work was experimental verification of the results of theoretical studies in order to determine the factors influencing the choice of the ratio of the lengths of the active elements of the generator and amplifier based on the organic dye rhodamine 6G with lamp pumping with their limited total length.

The results of the experiments confirm the theoretical conclusions that there are optimal ratios of the lengths of the generator and amplifier, at which the radiation energy is maximum. The limiting length of the amplifier and the energy of the emitter, built according to the scheme, the oscillator-amplifier are limited due to an increase in the intensity of the radiation that is amplified along the active element, as well as an increase in the intensity of the amplified noise.

Key words: lidar; oscillator; amplifier; flashlamp; resonator; pump.

7 fig. Ref: 16 items.

УДК 621.383

Експериментальні дослідження випромінювача лідару, побудованого за схемою генератор-підсилювач / *О.І. Цопа, О.А. Зарудний* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 127 – 133.

Наведено результати експериментальних досліджень енергетичних характеристик передавача лідару, побудованого за схемою генератор-підсилювач на органічному барвнику з ламповим накачуванням. При побудові випромінювача лідара за схемою генератор-підсилювач в умовах постійної щільності накачування виникає проблема вибору співвідношення між протяжністю активного елемента генератора і протяжністю активного середовища підсилювача біжучої хвилі, яка б забезпечувала максимальний ККД всього випромінювача. Основ-

ним завданням роботи була експериментальна перевірка результатів теоретичних досліджень з метою визначення факторів, що впливають на вибір співвідношень довжин активних елементів генератора та підсилювача на основі органічного барвника родамін 6Ж з ламповим накачуванням при обмеженій їх сумарній протяжності.

Результати експериментів підтверджують теоретичні висновки про те, що існують оптимальні співвідношення довжин генератора і підсилювача, при яких енергія випромінювання є максимальною. Гранична довжина підсилювача та енергія випромінювача, побудованого за схемою генератор-підсилювач обмежуються за рахунок зростання інтенсивності випромінювання, що посилюється уздовж активного елементу, а також збільшення інтенсивності посиленого шуму.

Ключові слова: лідар; генератор; підсилювач; лампа-спалах; резонатор; накачка.

Іл. 7. Бібліогр.: 16 назв.

UDC 621.373.12

Current state and development trends of class E oscillators: an overview / *V.G. Krizhanovski* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 134 – 140.

An analysis of the current state of power generators of the class E family was carried out. They include classes: inverse E (E^{-1}), with an injection of harmonics – class EM, hybrid classes E/F_n and EF_n, including those built based on an active distributed transformer scheme. New developments of such oscillators over the past five years are considered. Significant progress has been made in the development of new feedback schemes. The study of oscillator schemes with power summation and the use of synchronized oscillators is considered. In connection with the appearance of new active devices, circuits using additional active elements – drivers of powerful transistors and the use of two-stroke circuits for the construction of a class E oscillator key are spreading. The proposed classification of oscillators is based on the principle of building an output circuit and a feedback circuit, the morphological table of types is constructed class E oscillators. The principles of operation and characteristics of some new schemes of class E oscillators and their parameters are considered. The trends in the development of class E generators are determined, the main of which is the design of generators for operation in radio engineering systems and matching their parameters with the needs of such systems, as well as in systems of power (industrial) electronics, wireless energy transmission, biomedical and information systems. This is done in accordance with the trend of increasing the oscillator signal parameters while maintaining high efficiency. The variety of power levels, frequencies, and type of performance of class E oscillators is preserved and expanded, in the form of integrated circuits, which allows increasing the parameters of transmitters, sensors, and systems of compatible energy and information transmission.

Key words: generators of class E; circuits of oscillators; feedback links; classification of generators of the class E family.

4 fig. Ref: 36 items.

УДК 621.373.12

Сучасний стан та тенденції розвитку автогенераторів сімейства класу е: огляд / *В.Г. Крижановський* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 134 – 140.

Проведено аналіз сучасного стану потужних автогенераторів сімейства класу Е. До них відносяться класи: інверсний Е (E^{-1}), з інжекцією гармонік – клас E_M, гібридні класи E/F_n та EF_n, в тому числі побудовані на основі схеми активного розподіленого трансформатору. Розглянуто нові розробки таких осциляторів за останні п’ять років. Визначено значне просування у розробки нових схем зворотного зв’язку. Вказано на вивчення схем автогенераторів з підсумовуванням потужності та використанням синхронізованих автогенераторів. У зв’язку з появою нових активних пристроїв, поширюються схеми з використанням додаткових активних елементів – драйверів потужних транзисторів та використання двотактних схем побудови ключа автогенератору класу Е. Запропоновано класифікацію автогенераторів за принципом побудови вихідного кола та кола зворотного зв’язку, побудовано морфологічну таблицю типів автогенераторів класу Е. Розглянуто принцип роботи та характеристики деяких нових схем автогенераторів класу Е та їх параметри. Визначено тренди розвитку автогенераторів класу Е, головним з яких є конструювання автогенераторів для роботи в радіотехнічних системах та узгодженні їх параметрів з потребами таких систем, а також в системах силової (промислової) електроніки, бездротової передачі енергії, біомедичних та інформаційних системах. Це робиться відповідно до тенденції підвищення параметрів сигналу автогенераторів одночасно зі збереженням високого коефіцієнту корисної дії. Зберігається та розширюється різноманіття рівнів потужності, частот та виду виконання автогенераторів класу Е, зокрема у вигляді інтегральних схем, що дозволяє підвищити показники трансмітерів, сенсорів та систем суміс-ного передавання енергії і інформації.

Ключові слова: автогенератори класу Е; схеми автогенераторів; ланки зворотного зв’язку; класифікація автогенераторів сімейства класу Е.

Іл. 4. Бібліогр.: 36 назв.

UDC 624.323.64.012.8

Study of parameters of the avalanche diode generator / *O.D. Menyailo, V.G. Mahonin, M.S. Svitlichnyi* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. № 212. P. 141 – 147.

The article is devoted to the development and analysis of an avalanche diode generator. The equivalent circuit of the generator is considered and its simplified analysis is carried out. Using the elements of graphic analysis, the conditions of stability of generator oscillations were obtained. The original design of the avalanche diode generator is proposed and considered. A feature of the generator design is the use of a combined oscillating system, which is a three-dimensional resonator in the form of a metal ring made of aluminum alloy placed on a double-sided foil dielectric. Thanks to the top cover made of the same foil dielectric, the oscillating system has a closed nature. The electromagnetic energy supply and output system is made using printed technologies based on a double-sided foil dielectric. The avalanche diode is located in the center of this structure and has a thermal connection with the external radiator. On the one hand, such a decision allowed us to increase significantly the Q-factor of the oscillating system compared to the oscillating system made by the printed method and, at the same time, gave an opportunity to combine the developed auto-generator with other elements of the waveguide path made in the printed version. The developed generator has the possibility of both mechanical, by changing the volume of the resonator with the help of special backlash-free elements, and electronic adjustment, by changing the supply current. The article contains the results of experimental studies of the proposed design of the auto-generator, in particular the dependence of the output power on the frequency of oscillations and on the supply current, as well as the Q factor on the supply current. The research results indicate a fairly high Q-factor of such an oscillating system and, as a result, increased stability of oscillations. In addition, this design to a certain extent improves the overall manufacturability of the design and its material capacity in comparison with the waveguide version.

Key words: avalanche-flying diode; experimental research; electromagnetic energy; transformation; mechanical energy; foil dielectric.

8 fig. Ref: 5 items.

УДК 624.323.64.012.8

Дослідження параметрів генератора на лавино-пролітному діоді / *О.Д. Меняйло, В.Г. Махонін, М.С. Світличний* / Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 141 – 147.

Стаття присвячена розробці та аналізу генератора на лавинно-пролітному діоді. Розглянуто еквівалентну схему генератора та проведено її спрощений аналіз. З використанням елементів графічного аналізу одержано умови стабільності коливань генератора. Запропоновано та розглянуто оригінальну конструкцію генератора на лавинно-пролітному діоді. Особливістю конструкції генератора є застосування комбінованої коливальної системи, яка представляє собою об'ємний резонатор у вигляді металевого кільця з алюмінієвого сплаву, розміщеного на двосторонньому фольгованому діелектрику. Завдяки верхній кришці, виконаній з такого ж фольгованого діелектрика, коливальна система має закритий характер. Систему живлення та виводу електромагнітної енергії виконано з використанням друкованих технологій на базі двостороннього фольгованого діелектрика. Лавинно-пролітний діод розміщено в центрі цієї конструкції та має тепловий зв'язок з зовнішнім радіатором. Таке рішення дозволило, з одного боку, суттєво підвищити добротність коливальної системи у порівнянні з коливальною системою, виконаною друкованим способом і, в той же час, надає можливість поєднувати розроблений автогенератор з іншими елементами хвильоводного тракту, виконаними в друкованому варіанті. Розроблений генератор має можливість як механічного, шляхом зміни об'єму резонатора за допомогою спеціальних без люфтових елементів, так і електронного налаштування, шляхом зміни струму живлення. Стаття містить результати експериментальних досліджень запропонованої конструкції автогенератора, зокрема залежності вихідної потужності від частоти коливань та від струму живлення а також добротності від струму живлення. Результати досліджень свідчать про досить високу добротність такої коливальної системи і як результат підвищену стабільність коливань. Крім того ця конструкція в певній мірі покращує загальну технологічність конструкції та її матеріалоємність у порівнянні з хвильоводним варіантом.

Ключові слова: лавинно-пролітний діод; експериментальні дослідження; електромагнітна енергія; перетворення; механічна енергія; фольгований діелекрик.

Іл. 8. Бібліогр.: 5 назв.

RADAR AND RADIONAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

UDC 004.89: 621.396

Intelligent model of radar object images for surveillance radars / *V. Zhyrnov, S. Solonska* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. № 212. P. 148 – 154.

The results of developing an intelligent model of radar object images for surveillance radars are presented. The relevance of this work deals with the development of algorithm for automatic processing images of radar objects that provide effective detection of weak true signals due to the accumulation of signal and logical information in the analyzed cell and in its surroundings under interferences. The improvement of air safety tools and the automation of air traffic management processes require effective procedures to process signal information. The issues of more complete use and qualitative improvement of the information-processing capabilities of control systems are also topical, especially in difficult conditions of interfering signals. The basis of this study is the idea of using an intellectual model of radar object images for automatic decision-making on detection and recognition of radar objects, built on the space of semantic features. The main result is optical object recognition, similar to how an expert can easily recognize aerial objects

and their types when viewing radar object images. Based on semantic features intelligent model of radar object images has been developed, which makes it possible to effectively detect and classify aerial objects. It is worth noting that the characteristic description of intelligent model of radar object images for point, extended, moving and stationary radar objects is the mathematical description of procedures and relationships at perception and analysis of signals in the form of distinguishing features or properties. As a result, various virtual images of radar object are generated in the form of spatial-semantic and spectral-semantic models. The main features and structural elements of the model are given. It is shown that the advantages of this model are related to the possibility of characteristic description of the radar object images using the algebra of finite predicates.

Key words: semantic analysis; radar signal; identification; aerial object.

4 fig. Ref: 11 items.

УДК 004.89: 621.396

Інтелектуальна модель зображень відміток радіолокаційних об'єктів для оглядових РЛС / *В.В. Жирнов, С.В. Солонська* / Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 148 – 154.

Наведено результати розроблення інтелектуальної моделі зображень відміток радіолокаційних об'єктів для оглядових РЛС. Актуальність роботи полягає у створенні алгоритмів автоматичної обробки зображень радіолокаційних об'єктів для забезпечення ефективного автоматичного виявлення слабких корисних сигналів за рахунок накопичення сигнальної та логічної інформації в аналізованій комірці та в її оточенні в складних умовах завад. Удосконалення засобів забезпечення безпеки руху повітряного транспорту й автоматизація процесів управління вимагають ефективних процедур обробки сигнальної інформації. Актуальним є також питання більш повного використання і якісного поліпшення інформаційних можливостей систем управління, особливо в складних умовах заважаючих сигналів. В основу розробки покладена ідея використання інтелектуальної моделі зображень радіолокаційних об'єктів для автоматичного прийняття рішень з виявлення та розпізнавання об'єктів РЛС, побудованих на просторі семантичних ознак. Головним результатом є оптичне розпізнавання об'єктів, схоже на те, як експерт може легко розпізнавати аерооб'єкти та їх типи при перегляді зображень радіолокаційних об'єктів. На основі семантичних властивостей розроблено інтелектуальну модель зображення радіолокаційних об'єктів, що дозволяє ефективно виявляти й класифікувати повітряні об'єкти. Варто відзначити, що ознаковий опис інтелектуальної моделі зображень радіолокаційних об'єктів для точкових, протяжних, рухомих та нерухомих радіолокаційних об'єктів є математичним описом процедур і взаємозв'язків при сприйнятті та аналізі сигналів, представлених як відмінюючі ознаки або властивості. В результаті створюються різні віртуальні зображення позначок радіолокаційних об'єктів у вигляді просторово-семантичних та спектрально-семантичних моделей. Наведено основні особливості та структурні елементи моделі. Показано, що переваги даної моделі пов'язані з можливістю ознакового опису зображення радіолокаційного об'єкту з використанням алгебри кінцевих предикатів.

Ключові слова: семантичний аналіз; радіолокаційний сигнал; ідентифікація; повітряний об'єкт.

Іл. 4. Бібліогр.: 11 назв.

UDC 621.396.96

Distributed processing of radar information in airspace surveillance systems / *I.V. Svyd, S.V. Starokozhev* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. № 212. P. 155 – 165.

The work is dedicated to the analysis of the quality of combining assessments of the radar signals and airborne objects detections in the implementation of distributed processing of radar information of airspace surveillance systems. The main sources of radar information about the air situation in the airspace control system are primary surveillance radars, secondary radar systems and identification systems on the basis of "friend or foe". It should be noted that the analysis of the information security of single-position radars shows their vulnerability in a wide range of unintentional and intentional interference, as well as determining their location. This is due to the ease of detection of the emitting transmitter of the probing signal in single-position radars. It led to the main disadvantage of single-position radars – low noise immunity and low survivability. The transition to a network of radar systems can significantly reduce the impact of deliberately directed interference. It also allows the use of methods for distributed processing of radar information in airspace surveillance systems.

Analysis of the effectiveness of information support algorithms based on distributed processing of radar information of airspace surveillance systems, taking into account the final result, makes it possible to detect airborne objects using a packet of binary-quantized signals, taking into account two algorithms for combining detection results: channel accumulation and combining results; association of channel solutions and accumulation. It shows following: – the quality of consumer information support based on the proposed structure is much higher compared to the used radar information processing structure; the quality of information support for consumers has the best performance when using the signal processing method based on the accumulation of signals with the subsequent combination of detection results; the availability factor of the aircraft transponder significantly affects the quality of information support, already at P_0<0.9 the use of integer logic for combining detection information is undesirable.

Key words: radar system; airspace; surveillance system; radar information processing; distributed processing; evaluation.

3 fig. Ref: 41 items.

УДК 621.396.96

Розподілена обробка радіолокаційної інформації систем спостереження повітряного простору / *І.В. Свид, С.В. Старокожев* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 155 – 165.

Роботу присвячено аналізу якості об'єднання оцінок виявлення радіолокаційних сигналів та повітряних об’єктів при реалізації розподіленої обробки радіолокаційної інформації систем спостереження повітряного простору. Основними джерелами радіолокаційної інформації про повітряну обстановку в системі контролю повітряного простору є первинні оглядові радіолокатори, вторинні радіолокаційні системи та системи ідентифікації за ознакою «свій-чужий». Слід зазначити, що аналіз інформаційної безпеки однопозиційних радіолокаторів показує їх вразливість у широкому спектрі ненавмисних та навмисних завад, так і визначенні їх місця розташування. Це обумовлено простотою виявлення випромінюючого передавача зондувального сигналу в однопозиційних радіолокаторах. Що і зумовило основний недолік однопозиційних радіолокаторів – низька стійкість до завад та живучість. Перехід до мережі радіолокаційних систем дозволяє значно послабити вплив навмисно спрямованих завад. Також дозволяє використовувати методи розподіленої обробки радіолокаційної інформації у системах спостереження повітряного простору.

Аналіз ефективності алгоритмів інформаційного забезпечення на базі розподіленої обробки радіолокаційної інформації систем спостереження повітряного простору з урахуванням кінцевого результату дозволяє виявлення повітряних об’єктів за пачкою двійково-квантованих сигналів з урахуванням двох алгоритмів об'єднання результатів виявлення: канальне накопичення і об'єднання результатів; об'єднання канальних рішень і накопичення. Це показує, що: якість інформаційного забезпечення споживачів на підставі запропонованої структури значно вище в порівнянні з використовуваною структурою обробки радіолокаційної інформації; якість інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів, заснованого на накопиченні сигналів з подальшим об'єднанням результатів виявлення; коефіцієнт готовності літакового відповідача істотним чином впливає на якість інформаційного забезпечення, вже при $P_0 < 0,9$ використання цілочисельної логіки об'єднання інформації виявлення небажано.

Ключові слова: радіолокаційна система; повітряний простір; система спостереження; обробка радіолокаційної інформації; розподілена обробка; оцінка.

Іл. 3. Бібліогр.: 41 назв.

UDC 621.396.96

Synthesis of a complex algorithm for the operation of a radio-acoustic measuring complex / *V.A. Tikhonov, A.V. Kartashov* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 166 – 174.

Stations of radio-acoustic sounding of the atmosphere are a promising means of obtaining information about the altitudinal distribution of meteorological parameters in the Earth's atmosphere used in the process of solving current scientific and applied tasks to ensure aircraft flights, weather forecasting, etc. However, the effectiveness of the existing radio-acoustic means is insufficient, and there are practical needs for development of appropriate prospective structures and algorithms, which will be implemented when constructing specific stations designed to solve actual applied tasks.

The article presents a synthesis of the structural diagram of a promising radio-acoustic measuring complex (RAVC) and a complex algorithm of its functioning. It is shown that the process of the RAVC development can be reduced to the development of independent issues. They are as follows: synthesis (selection) of types of sounding radio and acoustic signals, their energy parameters, synthesis of spatial and temporal signal processing algorithms for optimal selection of useful information from signals received against a background of noise, development of complex management algorithms and complex adaptation to external conditions formed by the external environment, as well as development of technical means intended for the implementation of the specified algorithms. According to this task of synthesis (selection) of sounding signals for radio-acoustic systems of sounding the atmosphere in the design process, it should be a joint study of the characteristics and selection of two interdependent types of signals – electromagnetic and acoustic.

Key words: remote sensing of the atmosphere; complex; algorithm; estimation of parameters; management; temperature; signal; synthesis, interference.

2 fig. Ref: 26 items.

УДК 621.396.96

Синтез комплексного алгоритму функціонування радіоакустичного вимірювального комплексу / *В.А. Тихонов, О.В. Карташов* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 166 – 174.

Станції радіоакустичного зондування атмосфери є перспективним засобом отримання інформації про висотний розподіл метеопараметрів в атмосфері Землі, яка використовується в процесі вирішення актуальних науково-прикладних завдань з забезпечення польотів літальних апаратів, прогнозу погоди та ін. Проте ефективність існуючих радіоакустичних засобів є недостатньою і існують потреби практики з розробці відповідних перспективних структур та алгоритмів, що реалізуватимуться при побудові конкретних станцій, призначених для вирішення актуальних прикладних завдань.

Виконано синтез структурної схеми перспективного радіоакустичного вимірювального комплексу (РАВК) і комплексного алгоритму його функціонування. Показано, що процес розробки РАВК можна звести до розробки наступних незалежних питань: синтез (вибір) типів зондувальних радіо та акустичних сигналів, їх

енергетичних параметрів, синтез алгоритмів просторової та часової обробки сигналів для оптимального виділення корисної інформації з сигналів, що приймаються на фоні шумів, розробка алгоритмів управління комплексом і адаптації комплексу до зовнішніх умов, що формуються зовнішнім оточенням, а також розробка технічних засобів, призначених для реалізації вказаних алгоритмів. Відповідно до цього завдання синтезу (вибору) зондувальних сигналів для радіоакустичних систем зондування атмосфери у процесі проектування має полягати у спільному вивченні характеристик та виборі двох взаємозалежних видів сигналів – електромагнітного та акустичного.

Ключові слова: дистанційне зондування атмосфери; комплекс; алгоритм; оцінка параметрів; управляння; температура; сигнал; синтез, завада.

Іл. 2. Бібліогр.: 26 назв.

UDC 621.396.96

Synthesis and analysis of the trace detector of air objects of an interrogating radar system / *I.V. Svyd, M.G. Tkach* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. № 212. P. 175 – 185.

The paper considers the features of tracking air objects in existing information networks of radar systems for monitoring airspace. It is shown that the tracking is carried out according to the information of the primary radar surveillance systems, and the secondary radar surveillance systems are used as sources of additional radar information. But the transition to automatic dependent surveillance implies the mandatory use of only request radar surveillance systems. Therefore, the problems of formulating methods and algorithms for tracking air objects based on information from secondary radar surveillance systems are relevant. The specifics of the construction and operation of secondary radar surveillance systems differ significantly from primary radar surveillance systems. The work carried out the synthesis and analysis of the structure of the tracks detector of air objects by interrogating radar systems for monitoring the airspace, namely: a comparative analysis of the quality of identifying the tracks of air objects was carried out; the quality of information support to consumers of the airspace control system with the proposed structure was improved in comparison with the used information processing structure; it is shown that the quality of information support for consumers has more preferable indicators when using the signal processing method during acquisition with subsequent information merging; the influence of the readiness factor of aircraft transponders of interrogative radar systems on the quality of information support for consumers of the airspace control system was evaluated.

Key words: radar system; interrogation radar system; synthesis; analysis; track; signal; decision; detector; method; an air object; surveillance system; information network.

3 fig. Ref: 46 items.

УДК 621.396.96

Синтез і аналіз виявлювача трас повітряних об’єктів запитальної радіолокаційної системи / *І.В. Свид, М.Г. Ткач* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 175 – 185.

Розглянуто особливості супроводу повітряних об'єктів в існуючих інформаційних мережах радіолокаційних систем спостереження повітряного простору. Показано, що супровід здійснюється за інформацією первинних радіолокаційних систем спостереження, а вторинні радіолокаційних систем спостереження використовуються як джерела додаткової радіолокаційної інформації. Але перехід на автоматичне залежне спостереження передбачає обов'язкове використання лише запитальних радіолокаційних систем спостереження. Тому актуальними є питання розробки методів та алгоритмів супроводу повітряних об’єктів за інформацією вторинних радіолокаційних систем спостереження. Продемонстровано, що специфіка побудови та функціонування вторинних радіолокаційних систем спостереження суттєво відрізняється від первинних радіолокаційних систем спостереження. Проведено синтез та аналіз структури виявлювача траси повітряних об’єктів запитальними радіолокаційними системами спостереження повітряного простору. Проведено порівняльний аналіз якості виявлення трас повітряних об’єктів; забезпечено підвищення якості інформаційного забезпечення споживачів системи контролю повітряного простору на підставі запропонованої структури в порівнянні зі структурою обробки інформації, що використовується; показано, що якість інформаційного забезпечення споживачів має кращі показники при використанні методу обробки сигналів при накопиченні з наступним об'єднанням інформації; оцінено вплив коефіцієнта готовності літакових відповідачів запитальних радіолокаційних систем на якість інформаційного забезпечення споживачів системи контролю повітряного простору.

Ключові слова: радіолокаційна система; запитальна радіолокаційна система; синтез; аналіз; траса; сигнал; рішення; виявлювач; метод; повітряний об’єкт; система спостереження; інформаційна мережа.

Іл. 3. Бібліогр.: 46 назв.

MEANS OF TELECOMMUNICATIONS
ЗАСОБИ ТЕЛЕКОМУНІКАЦІЙ

UDC 621.396.004

Creating a call center test bench for load balancing Asterisk servers in a cluster / *L.O. Tokar, O.A. Koltakov, V.Y. Tsyliuryk* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2023. №212. P. 186 – 196.

The article deals with the issues of increasing throughput in call centers. The current solution is to cluster call servers and evaluate their characteristics to ensure efficient operation and the necessary fault tolerance. It is shown that one of the main aspects of the quality functioning of the call center is load balancing of servers in the cluster.

The features of the call-center modeling process are considered. The organization scheme of the call center of the company and the network model of the call center have been created. Virtualization technology was used to create a network model of a call center. The VMWare ESXI 6.7 hypervisor and the vCenter client were used as a network configuration platform.

An analysis of load balancing was carried out using different algorithms and strategies.

Asterisk PBX was configured and a server cluster was created. A test bench was developed and configured using the Zabbix open source product to explore a cluster of call servers. A bandwidth characteristic for the Asterisk _1 server and a network map were obtained, which actually represents a simulated structure of the call center network.

The process of load testing on three Asterisk servers and the implementation of Zabbix load balancing capabilities are shown. A custom SIPp session script has been created for accepting calls and load testing. The number of calls that the call center is able to handle is calculated. It has been determined that a single Asterisk server with its current settings can handle a maximum of 915 concurrent calls. The process of load balancing on a cluster of SIP servers has been launched. A cluster of Asterisk servers has been found to be capable of handling 2550 simultaneous calls.

Key words: virtualization; balancing; server; cluster; load; testing; calls; Asterisk.

1 tab. 12 fig. Ref: 19 items.

УДК 621.396.004

Створення тестового стенду call-центру для балансування навантаження серверів Asterisk у кластері / *Л.О.Токар, О.А. Колтаков, В.Є. Цилюрик* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2023. Вип. 212. С. 186 – 196.

Розглянуто питання підвищення пропускної здатності у центрах обробки викликів. Актуальним рішенням визначено кластерізацію серверів викликів та оцінку їх характеристик для забезпечення ефективної роботи та необхідної відмовостійкості. Показано, що одним із основних аспектів якісного функціонування call-центру є балансування навантаження серверів у кластері.

Розглянуто особливості процесу моделювання call-центру. Створено схему організації call-центру компанії та мережну модель call-центру. Для створення мережної моделі call-центру використано технологію віртуалізації. У якості платформи для налаштування мережі використано гіпервізор VMWare ESXI 6.7 та кліент vCenter.

Проведено аналіз балансування навантаження з використанням різних алгоритмів та стратегій.

Проведено налаштування АТС Asterisk та створення кластеру серверів. Розроблено та налаштовано тестовий стенд з використанням open source продукту Zabbix для дослідження кластеру серверів викликів. Отримано характеристику пропускної здатності для сервера Asterisk_1 та карту мережі, яка фактично являє собою змодельовану структуру мережі call-центру.

Показано процес тестування навантаження викликами на три сервери Asterisk та реалізація можливостей Zabbix для балансування навантаженням. Для прийняття викликів та тестування навантаження створено власний сценарій сесії в SIPp. Розраховано кількість викликів, з якими здатний працювати call-центр. Виявлено, що один сервер Asterisk з поточними його параметрами здатний обробити максимум 915 одночасних викликів. Запущено процес балансування навантаження на кластер SIP-серверів. Виявлено, що кластер серверів Asterisk здатний обслуговувати 2550 викликів одночасно.

Ключові слова: віртуалізація; балансування; сервер; кластер; навантаження; тестування; виклики; Asterisk.

Табл. 1. Іл.12. Бібліогр.: 19 назв.

COLLECTION OF SCIENTIFIC PAPERS
RADIOTEKHNIKA
Issue 212
In English and Ukrainian Russian

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 212
Англійською та українською мовами

Коректор Л.І. Сащенко

Підп. до друку 30.03.2023. Формат 60x90/8. Папір офсет. Гарнітура Таймс. Друк. ризограф.
Ум. друк. арк**.** 10,9. Обл.-вид. арк. 10,1. Тираж 300 прим. Зам. № 32. 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.