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 1

Kharkiv
Kharkiv National
University of Radio Electronics
2022

UDC 621.3

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

Website: rt.nure.ua

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

The authors are responsible for the content of the article.

Editorial Team

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

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

Recommended by the Scientific and Technical Council of Kharkiv National University of Radio Electronics, protocol № 11 dated 08.12.2022.

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 1

Харків
Харківський національний
університет радіоелектроніки
2022

УДК 621.3

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

Сайт: rt.nure.ua

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

За зміст статті відповідальні автори.

Рекомендовано Науково-технічною радою Харківського національного університету радіоелектроніки, протокол № 11 від 08.12.2022.

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

CONTENT

SYSTEMS AND METHODS OF INFORMATION PROTECTION

Ye.V. Ostrianska, S.O. Kandiy, I.D. Gorbenko, M.V. Yesina Classification and analysis of vulnerabilities of modern information systems from classical and quantum attacks 7

S.O. Kandiy Analysis of DSTU 8961:2019 in random oracle model 22

V.I. Yesin, V.V. Vilihura The main categories of NewSQL databases and their features 37

D.V. Harmash Analysis of the Falcon signature compared to other signatures. GPV and Rabin frameworks 67

RADIO PHYSICS

O.V. Lazorenko, A.A. Onishchenko, L.F. Chernogor Multifractal analysis of model fractal and multifractal signals 72

RADIO LOCATION AND RADIO NAVIGATION

V.M. Kartashov, V.A. Pososhenko, A.I. Kapusta, M.V. Rybnykov, E.V. Pershin Features of the tasks of identifying and observing groups of unmanned letter vehicles 84

V.M. Kartashov, V.A. Pososhenko, K.V. Kolesnik, V.I. Kolesnik, R.I. Bobnev, A.I. Kapusta Algorithm for estimating the energy distribution of radar signals scattering on acoustic disturbances created by UAVs 93

I.V. Svyd, M.G. Tkach, I.I. Obod Comparative analysis of interference protection of "Friend-Foe" radar identification systems 101

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS

V.O Alieksieiev. D.V. Gretskih, D.S. Gavva, V.G. Lykhograi Wireless power transmission technologies 114

V.M. Borshchov, O.M. Listratenko, M.A. Protsenko, I.T. Tymchuk, O.V. Kravchenko, O.V. Suddia, I.V. Borshchov, M.I. Slipchenko Structural modeling and calculation of thermal conductivity of polyimide composite materials 133

V.V. Rapin Theoretical investigation of injection-locked differential oscillator 143

BIOMEDICAL RADIO ELECTRONICS

O.I. Dovnar, M.F. Babakov, V.I. Cherkis Using a fingerprint scanner to protect data in medical information systems 148

INFORMATION METHODS OF RADIO ENGINEERING, SIGNAL PROCESSING

I.V. Svyd, A.O. Serikov, I.I. Obod Applying MATLAB to Radar Systems Modeling 154

ABSTRACTS 159

ЗМІСТ

СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

Є.В. Острянська, С.О. Кандій, І.Д. Горбенко, М.В. Єсіна Класифікація та аналіз вразливостей сучасних інформаційних систем від класичних та квантових атак 7

С.О. Кандій Аналіз безпеки ДСТУ 8961:2019 у моделі випадкового оракула 22

В.І. Єсін, В.В. Вілігура Основні категорії NewSQL баз даних та їх особливості 37

Д.В. Гармаш Аналіз підпису FALCON в порівнянні з іншими підписами. Фреймворки GPV та Рабіна 67

РАДІОФІЗИКА

О.В. Лазоренко, А.А. Онищенко, Л.Ф. Чорногор Мультифрактальний аналіз модельних фрактальних і мультифрактальних сигналів 72

РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

В.М. Карташов, В.О. Посошенко, А.І. Капуста, М.В. Рибников, Є.В. Першин Особливості задач виявлення і спостереження груп безпілотних літальних апаратів 84

В.М. Карташов, В.О. Посошенко, К.В. Колісник, В.І. Колісник, Р.І. Бобнєв, А.І. Капуста Алгоритм оцінювання розподілу енергії радіолокаційних сигналів, які розсіюються на акустичних збуреннях, створених БПЛА 93

Свид І.В., Ткач М.Г., Обод І.І. Порівняльний аналіз завадозахищеності радіолокаційних систем ідентифікації за ознакою «свій-чужий» 101

ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

В.О. Алєксєєв, Д.В. Грецьких, Д.С. Гавва, В.Г. Лихограй Технології безпровідної передачі енергії 114

В.М. Борщов, О.М. Лістратенко, М.А. Проценко, І.Т. Тимчук, О.В. Кравченко, О.В. Суддя, І.В. Борщов, М.І. Сліпченко Структурне моделювання і розрахунок теплопровідності поліімідних композитних матеріалів 133

В.В. Рапін Теоретичне дослідження синхронізованого диференціального автогенератора (*англ.*) 143

БІОМЕДИЧНА РАДІОЕЛЕКТРОНІКА

О.Й. Довнар, М.Ф. Бабаков, В.І. Черкіс Використання сканеру відбитків пальців для захисту даних у медичних інформаційних системах 148

ІНФОРМАЦІЙНІ МЕТОДИ РАДІОТЕХНІКИ, ОБРОБКА СИГНАЛІВ

І.В. Свид, А.О. Сєріков, І.І. Обод Застосування MATLAB для моделювання радіолокаційних систем 154

РЕФЕРАТИ 159

УДК 004.056.5 **DOI:10.30837/rt.2022.4.211.01**

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

КЛАСИФІКАЦІЯ ТА АНАЛІЗ ВРАЗЛИВОСТЕЙ СУЧАСНИХ ІНФОРМАЦІЙНИХ СИСТЕМ ВІД КЛАСИЧНИХ ТА КВАНТОВИХ АТАК

Вступ

Завдяки останнім досягненням у квантових технологіях та потенціалу того, що практичні квантові комп'ютери можуть стати реальністю у майбутньому, відновився інтерес до розробки криптографічних технологій, захищених від звичайних та квантових атак. Наразі практично всім асиметричним криптографічним схемам, які сьогодні використовуються, загрожує потенційна розробка потужних квантових комп'ютерів. Постквантова криптографія є одним із способів боротьби із цією загрозою. Її безпека базується на складності математичних проблем, які наразі вважаються нерозв'язними ефективно – навіть за допомогою квантових комп'ютерів.

Безпека інформаційних систем забезпечується через захист від різноманітних загроз, що використовують вразливості системи. Загроза – це потенційне порушення безпеки, тоді як атака – це погроза, яка виконується. Процеси безпеки стосуються вибору та реалізації засобів контролю безпеки (так звані контрзаходи), які допомагають зменшити ризик, спричинений вразливостями.

Протоколи безпеки є будівельними блоками безпечного зв'язку. Вони реалізують механізми безпеки для надання послуг безпеки. Протоколи безпеки вважаються абстрактними під час аналізу, але вони можуть мати додаткові вразливості у реалізації. Ця робота містить цілісне дослідження протоколів безпеки. Розглядаються основи протоколів безпеки, таксономія атак на протоколи безпеки та їх впровадження, а також різні методи та моделі аналізу безпеки протоколів. Зокрема, уточнюються відмінності між інформаційно-теоретичною та обчислювальною безпекою, обчислювальними та символьними моделями. Крім того, надано огляд моделей обчислювальної безпеки для автентифікованого обміну ключами (AKE) і протоколів обміну ключами з автентифікацією пароля (PAKE).

Також було описано найважливіші моделі безпеки для протоколів AKE і PAKE. З появою нових технологій, які можуть мати інші вимоги до безпеки, а також завдяки збільшеним можливостям змагальності, завжди виникає потреба в розробці нових протоколів.

Для майбутнього використання постквантової криптографії недостатньо стандартизувати криптографічні алгоритми. Швидше, необхідно також адаптувати криптографічні протоколи до нових алгоритмів. Це пов'язано, наприклад, з тим, що в багатьох протоколах дозволена довжина відкритих ключів обмежена і більше недостатня для нових алгоритмів. Однак істотним моментом є те, що постквантові алгоритми, як правило, не слід використовувати окремо, а лише в гібридному режимі, тобто в поєднанні з класичною процедурою. Зміни в протоколах і стандартах повинні бути ініційовані та спільно розроблені галуззю. Ця робота вже триває для багатьох протоколів.

Метою статті є огляд, класифікація, аналіз та дослідження вразливостей інформаційних систем від класичних, квантових та спеціальних атак, виконані з урахуванням прогнозу щодо можливостей здійснення атак на постквантові криптографічні перетворення; вивчення моделей для оцінки безпеки для існуючих криптографічних протоколів, а також огляд та по-

рівняльний аналіз моделей безпеки та надання пропозицій щодо захисту від існуючих потенційних атак.

1. Попередні визначення квантово-безпечної криптографії

В останні роки спостерігається стійкий прогрес у створенні квантових комп'ютерів. У разі реалізації великомасштабних квантових комп'ютерів вони будуть загрожувати безпеці багатьох широко використовуваних криптосистем з відкритим ключем. Щоб протистояти загрозі сучасної асиметричної криптографії з боку квантових комп'ютерів, виникла нова галузь криптографічних досліджень – постквантова криптографія.

Постквантова криптографія займається розробкою та дослідженням асиметричних криптосистем, які, згідно із сучасними знаннями, не можуть бути зламані навіть потужними квантовими комп'ютерами. Тобто квантово-стійка криптографія – це криптографія, яка спрямована на надання криптографічних функцій і протоколів, які залишаються безпечними, навіть, якщо створено великомасштабні відмовостійкі квантові комп'ютери [1]. Ці методи базуються на математичних задачах, для розв'язання яких на сьогодні невідомі ані ефективні класичні алгоритми, ані ефективні квантові алгоритми. У сучасних дослідженнях застосовуються різні підходи до реалізації постквантової криптографії. До них належать, серед іншого:

- Криптографія на основі кодів: безпека схем на основі кодів ґрунтується на труднощах ефективного декодування загальних кодів з виправленням помилок.
- Криптографія на основі решітки: безпека схем на основі решітки базується на складності вирішення певних обчислювальних проблем на математичних решітках.
- Криптографія на основі гешування: безпека схем підпису на основі гешування базується на властивостях безпеки використаної геш-функції.
- Криптографія на основі ізогенії: схеми на основі ізогенії базують свою безпеку на тому факті, що важко знайти ізогенію між двома суперсингулярними еліптичними кривими, якщо така існує.
- Багатовимірна криптографія: безпека багатовимірної криптографії базується на припущенні, що багатовимірні поліноміальні системи рівнянь над скінченними полями важко вирішити.

Далі будуть розглянуті лише перші три класи, оскільки постквантові схеми, рекомендовані BSI, а також ті, що пройшли до фіналу конкурсу NIST та будуть стандартизовані, належать до цих класів. Багатоваріантні схеми мають довгу історію атак і виправлень. Криптографія, заснована на ізогеніях (відображення між еліптичними кривими зі спеціальними властивостями), є цікавою темою дослідження, яку, на думку BSI, слід вивчити далі, перш ніж розглядати рекомендацію.

Реалізація загроз, що спрямовані на програмні ресурси, може призводити до порушення вимог безпеки первинних інформаційних ресурсів, вплинути на інше ПЗ та, в окремих випадках, на функціонування апаратних ресурсів. А також порушити цілісність, справжність, доступність, неспростовність даних, що мають бути захищеними від несанкціонованих дій, які можуть привести до випадкової або умисної модифікації чи знищення. Саме тому у розд. 2 буде розглянуто основні атаки на постквантові криптографічні алгоритми.

2. Класифікація та аналіз основних атак на постквантові криптографічні перетворення

У цьому розділі коротко розглянуті деякі атаки на протоколи, схеми шифрування та їх реалізації.

Загрози в інформаційній безпеці можна розділити на чотири великі класи: розкриття або несанкціонований доступ до інформації; обман або прийняття неправдивих даних; порушення, переривання або запобігання коректній роботі; узурпація або несанкціонований контроль деякої частини системи [2]. Атаки також можна розділити на пасивні та активні [3].

Прослуховування є різновидом розкриття та несанкціонованого перехоплення інформації. Це пасивна атака, яка може бути або видаленням вмісту повідомлення, аналізом трафіку або переглядом файлів чи системної інформації. Маскування або підробка має місце, коли сутність видає себе за іншу сутність. Це можна вважати як вид обману та узурпації.

Несанкціонована зміна інформації є активною атакою, яка може бути обманом, або зривом і узурпацією. Затримка і відмова в обслуговуванні (DoS) є тимчасовими і довгостроковими гальмуваннями послуги відповідно. Хоча їх і можна вважати узурпацією, вони можуть грати допоміжну роль в обмані. Вони можуть бути результатом прямих атак або інших проблем, які не стосуються безпеки.

Під атаками впровадження маємо на увазі атаки, які використовують інформацію, яка отримана через витік через криптографічний примітив або його конкретне використання в протоколі безпеки. Наприклад, вимірювання споживання енергії або часу, необхідного для шифрування того самого повідомлення з різними секретними ключами може надати певну інформацію про секретні ключі.

2.1. Атаки на протоколи

Атака на протокол визначається та виконується відповідно до цілей безпеки або вимог безпеки протоколу, або моделі безпеки, в якій безпека протоколу доведена. Атака відбувається, коли порушується будь-яка властивість протоколу. Як зазначається в [42] атаки на протоколи безпеки можна в цілому розділити на пасивні та активні атаки. Також можна класифікувати атаки на протоколи безпеки на основі їх недоліків експлуатації [4].

У цьому розділі представлено неповний список стандартних атак на протоколи [5 – 8]. Це найпоширеніші типи атак, засновані на практичних сценаріях, завдяки чому зловмисник може спричинити збій протоколу. Список неповний, тому що в теорії є необмежені способи взаємодії криптоаналітика з одним або кількома (наприклад, паралельними) протоколами. Наведений нижче список не включає атаки, засновані на недоліках апаратного забезпечення або реалізації програмного забезпечення. Отже, найпоширенішими атаками на протоколи є:

- Атака з уособленням: це активна атака, спрямована на порушення автентичності. У цій атаці криптоаналітик намагається видати себе за одну або більше сутностей. Відповідно до змагальної моделі у відповідній моделі безпеки атака з уособленням може мати слабші варіанти обміну автентифікованими ключами (AKE) або протоколи обміну ключами з автентифікацією на основі пароля (PAKE), які будуть розглянуті у розд. 3. Ключова атака з уособленням компромісу (KCI) і атака з уособленням компрометації тимчасового ключа є слабшими варіантами атаки з уособленням у протоколах AKE, які потребують знання статичного секретного ключа та тимчасового секретного ключа (випадкове число) відповідно. Мета в таких варіантах – видати себе за іншу особу або іншу сутність скомпрометованій сутності [9].
- Атака «Людина посередині» (MITM): це варіант атаки з уособленням, де криптоаналітик знаходиться між двома сутностями та переконливо видає себе за обидві жертви. Практичні приклади включають атаку MITM на стільниковий зв'язок GSM мережі [10], протокол HTTPS [11] і EMV (Europay, MasterCard і Visa) протокол [12]. MITM можливий, коли протоколу бракує (взаємної) автентифікації.
- Атака зі спільним використанням невідомого ключа (UKS): це варіант атаки уособлення в протоколах AKE. Під час атаки UKS дві сутності мають спільний ключ сеансу, але вони мають різні представлення сеансу [13]. Атака UKS можлива коли протокол обміну ключами не може забезпечити автентифіковане зв’язування між сеансовим ключем та ідентифікаторами чесних об’єктів. Як правило, є два типи атак UKS [14, 15]: у першому типі, який називається Public UKS, атака замінює ключ, зловмисник реєструє відкритий ключ іншої сутності як власний відкритий ключ. У атаці UKS другого типу криптоаналітик має дійсний публічно-секретний ключ, сертифікований центром сертифікації, і намагається здійснити атаку UKS.

- Атака повтору: це активна атака, під час якої криптоаналітик втручається в запуск протоколу шляхом вставки деяких повідомлень із попередніх запусків протоколу або паралельно сесії. Це можна розглядати як комбінацію атак прослуховування та модифікації. Протокол вразливий до атаки відтворення, якщо він не забезпечує свіжість (freshness). Актуальність можна забезпечити за допомогою часових позначок, одноразових номерів або маркерів сеансу, а також лічильники [16].
- Атака Preplay: під час атаки Preplay супротивник готується до атаки шляхом імітації виконання протоколу та виконання набору операцій. Криптоаналітик виконує справжню атаку пізніше, коли є ймовірність здійснити той самий ряд операцій, що й у симуляції. Атака Preplay можлива, коли передбачуваний виклик у протоколах є виклик-відповідь [16, 17].
- Атака на відмову в обслуговуванні (DoS): DoS-атаки відносяться до широкого класу атак, які направлені на порушення доступності систем [18]. З точки зору протоколів, вони відносяться до атак, у яких зловмисник перешкоджає законним особам завершити протокол. На практиці вони можуть відбуватися проти серверів, які взаємодіють з багатьма клієнтами. Зловмисник може використати обчислювальні ресурси передбачуваного сервера (атака виснаження ресурсу) або перевищити кількість дозволених підключень до сервера (атака з розривом з'єднання). Неможливо повністю запобігти DoS-атакам, але можна зменшити їх вплив. Протоколи, які відкладають автентифікацію до кінця протоколу, набагато більш уразливі до атак DoS, ніж протоколи, які забезпечують автентифікацію на ранніх етапах.
- Атаки на дефекти: у атаках на дефекти зловмисник використовує відсутність належної перевірки типу повідомлення. Зловмисник надсилає повідомлення іншого типу, ніж очікується. Об'єкт-жертва не може виявити невідповідність типу та неправильно інтерпретує вміст повідомлення або поводиться неочікувано. Заходи протидії до атаки дефектів полягають у зміні порядку елементів повідомлення в наступному використанні одного й того самого повідомлення та гарантуванні того, що кожен ключ шифрування використовується один раз.
- Криптоаналіз: у протоколах безпеки криптографічні примітиви вважаються абстрактними та захищеними від атак. Однак є виняток, коли відомо, що ключ слабкий. Ці ситуації не повинні розкривати верифікатори або докази, які можуть бути використані для розкриття ключа. Важливо, щоб протоколи PAKE протистояли таким атакам:
  - офлайн-атака за словником: під час атаки за словником в автономному режимі, яка є пасивною, зловмисник підслуховує зв'язок між двома чесними об'єктами та отримує верифікатор, який можна використовувати для вилучення пароля за допомогою словника найбільш імовірних паролів. Зловмисник застосовує кожен пароль зі словника до отриманого верифікатора, поки не знайде правильний пароль, який задовольняє рівнянню верифікатора;
  - онлайн-атака за словником: під час онлайн-атаки за словником зловмисник використовує словник найбільш вірогідних паролів, але отримує верифікатор через онлайн-взаємодію з цільовою сутністю. Як контрзахід сервери зазвичай блокують обліковий запис користувача після кількох невдалих спроб.
- Атака за вибраним протоколом: під час атаки за вибраним протоколом новий протокол призначений для взаємодії з існуючим протоколом і створення вразливості. Ця атака заснована на сценарії взаємодії протоколу, де ключ використовується для кількох програм, наприклад, смарт-карти.
- Внутрішні недоліки дій: група атак заснована на відсутності деяких операцій, які є вирішальними для гарантування властивості безпеки. Прикладом є відсутність перевірки повідомлення, отриманого на третій фазі протоколу трьох проходів [19].

2.2. Атаки на алгоритми шифрування

Наївний спосіб атакувати схему шифрування полягає в атаці грубої сили або вичерпному пошуку ключа, коли зловмисник пробує всі можливі ключі в просторі ключів на парі

відкритий текст-шифртекст, доки не знайде ключ. Метою зловмисника є систематичне відновлення відкритого тексту із зашифрованого тексту або виведення ключа [5]. За класифікацією з [42] атаки на схеми шифрування можна розділити на наступні моделі атак:

- Під час атаки лише зашифрованим текстом криптоаналітик має лише зашифрований текст. Схема шифрування є абсолютно небезпечною, якщо вона вразлива до цієї атаки.
- Під час атаки з відомим відкритим текстом криптоаналітик також має певну кількість відкритого тексту та відповідного зашифрованого тексту.
- У атаці обраного відкритого тексту (CPA) криптоаналітик вибирає відкритий текст, а потім отримує відповідний зашифрований текст. Зловмисник використовує виведену інформацію, щоб відновити відповідний відкритий текст зашифрованого тексту, який раніше не бачив. Схеми шифрування з відкритим ключем є прикладом, коли зловмисник може зашифрувати будь-яке повідомлення за своїм вибором під відкритим ключем жертви. Адаптивна атака обраного відкритого тексту (CPA2) – це атака CPA, у якій вибір відкритого тексту зловмисником може залежати від зашифрованого тексту, створеного під час попередніх зашифрувань.
- Під час атаки за допомогою обраного зашифрованого тексту (CCA) зловмисник може розшифрувати довільні зашифровані тексти, наприклад за допомогою доступу до обладнання для розшифрування з надійно вбудованим ключем розшифрування. Мета полягає в тому, щоб вивести відкритий текст із раніше невидимого зашифрованого тексту. CCA має два спеціальні варіанти: у неадаптивній атаці зашифрованого тексту (CCA1) [20] зловмисник може мати доступ до системи лише протягом обмеженого часу або обмеженої кількості пар відкритий текст-зашифрований текст. Атаку називають неадаптивною, оскільки зловмисник не може адаптувати свої запити до оракула дешифрування відповідно до зашифрованого тексту виклику. В адаптивній атаці обраного зашифрованого тексту (CCA2) [21], яка є сильнішою, ніж CCA1, зловмисник має доступ до оракула розшифрування навіть після отримання виклику зашифрованого тексту.

Більшість із наведених атак можуть стосуватися до схем цифрового підпису та кодів автентифікації повідомлень (MAC), де метою зловмисника є підробка повідомлень або MAC. На основі наведених вище моделей атак у літературі представлено різні методи криптоаналізу. Найбільш широко використовуваними методами для криптоаналізу схем шифрування з симетричним ключем є диференціальний криптоаналіз [22, 23], лінійний криптоаналіз [24] і алгебраїчний криптоаналіз [25]. Інші методи включають комбіновані атаки [26], атаку «зустріч посередині» [27], інтегральний криптоаналіз [28], атаку з пов’язаним ключем [29] і атаку за визначенням [30].

Схеми шифрування з асиметричним ключем побудовані на нерозв'язності деяких складних проблем. Складні задачі, які використовуються в криптографії з відкритим ключем, включають факторизацію цілих чисел, задачу дискретного логарифмування (DLP) у відповідних групах, таких як мультиплікативні групи скінченних полів або адитивні групи еліптичних кривих над скінченними полями, задачі про рюкзак і задачі решітки, проблеми кодування та багатовимірні поліноміальні рівняння над малими кінцевими полями. Для схем шифрування було введено кілька понять безпеки, а саме:

- Нерозрізнюваність (IND), яка формалізує нездатність зловмисника дізнатися будь-яку інформацію про відкритий текст, що лежить в основі зашифрованого тексту виклику.
- Неподатливість (NM), яка формалізує нездатність супротивника перетворити даний зашифрований текст на інший зашифрований текст, щоб їхні відповідні відкриті тексти були «значуще пов’язані».
- Розпізнавання відкритого тексту [31, 32], яке формалізує нездатність криптоаналітика створити зашифрований текст, не знаючи базових повідомлень.

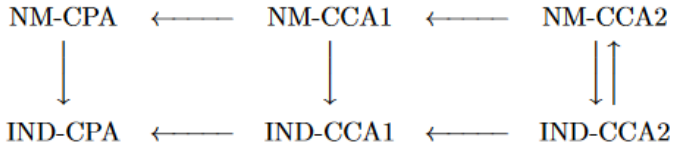


Рис. 1. Зв'язки між поняттями безпеки для схем шифрування з відкритим ключем [42]

Нерозрізнення є важливою властивістю для збереження конфіденційності. Однак у деяких випадках це може означати інші властивості безпеки, такі як цілісність, яка так чи інакше пов'язана з неподатливістю. На рис. 1 зображено співвідношення між поняттями нерозрізнення та неподатливості для схем шифрування з відкритим ключем під час атак CPA, CCA1 та CCA2 [31]. Стрілки позначають наслідки. Наприклад, якщо схема шифрування безпечна NM-CCA2, вона також безпечна NM-CCA1. Однак, якщо схема шифрування захищена NM-CCA1, вона може бути зламана в сенсі NM-CCA2. IND-CCA2 передбачає всі інші поняття. Поняття {IND-CPA, IND-CCA1, IND-CCA2} і {NM-CPA, NM-CCA1, NM-CCA2} можна визначити наступним чином (визначення 1 та визначення 2).

В и з н а ч е н н я 1. IND-CPA, IND-CCA1, IND-CCA2. Нехай $\Pi = (Gen, Enc, Dec)$ позначає схему шифрування з відкритим ключем, а $A = (A_1, A_2)$ позначає криптоаналітика з двома підалгоритмами. Для атаки $atk \in \{cpa, cca1, cca2\}$ і параметра безпеки $n \in \mathbb{N}$ ймовірність успіху криптоаналітика визначається як

$$Adv_{A,\Pi}^{ind-atk}(n) = \Pr\left[Exp_{A,\Pi}^{ind-atk-1}(n) = 1\right] - \Pr\left[Exp_{A,\Pi}^{ind-atk-0}(n) = 1\right], \quad (1)$$

для $b \in \{0,1\}$. Експеримент $Exp_{A,\Pi}^{ind-atk-b}(n) = b'$ визначається як

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

де

$$atk = cpa \Rightarrow O_1(\cdot) = \varepsilon, O_2(\cdot) = \varepsilon,$$
$$atk = cca1 \Rightarrow O_1(\cdot) = O_{Dec}(\cdot), O_2(\cdot) = \varepsilon,$$
$$atk = cca2 \Rightarrow O_1(\cdot) = O_{Dec}(\cdot), O_2(\cdot) = O_{Dec}(\cdot).$$

Схема шифрування безпечна в сенсі IND-ATK, якщо $Adv_{A,\Pi}^{ind-atk}(\cdot)$ є незначним у n [31].

В и з н а ч е н н я 2. NM-CPA, NM-CCA1, NM-CCA2. Нехай $\Pi = (Gen, Enc, Dec)$ позначає схему шифрування з відкритим ключем, а $A = (A_1, A_2)$ позначає криптоаналітика з двома підалгоритмами. Для атаки $akt \in \{cpa, cca1, cca2\}$ і параметра безпеки $n \in \mathbb{N}$ ймовірність успіху криптоаналітика визначається як

$$Adv_{A,\Pi}^{nm-atk}(n) = \Pr\left[Exp_{A,\Pi}^{nm-atk-1}(n) = 1\right] - \Pr\left[Exp_{A,\Pi}^{nm-atk-0}(n) = 1\right], \quad (2)$$

у якому для $b \in \{0,1\}$ експеримент $Exp_{A,\Pi}^{nm-atk-b}(n) = b'$ визначається як

$$
\begin{array}{ccc}
(pk, sk) & \leftarrow & Gen(1^n) \\
(m_0, m_1, s) & \leftarrow & A_1^{O_1}(pk) \\
c & \leftarrow & Enc_{pk}(m_1) \\
(R, c') & \leftarrow & A_2^{O_2}(m_0, m_1, s, c) \\
m' & \leftarrow & Dec_{sk}(c') \\
c \notin c' \wedge \perp \notin m' \wedge R(m_b, m') & \text{Тоді} & b' \leftarrow 1 \\
\text{Інакше} & & b' \leftarrow 1 \\
\text{Повернути} & b' &
\end{array},
$$

де m і c позначають вектори відкритих і зашифрованих текстів, $\perp$ позначає результат дешифрування, $R(\cdot)$ представляє поняття «значуще пов'язаних» і

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

Схема шифрування безпечна в сенсі NM-ATK, якщо $Adv_{A,\Pi}^{nm-atk}(\cdot)$ є незначним у n [31].

Уявлення про безпеку в схемах шифрування з симетричним ключем також можна моделювати як ігри з криптоаналітиком [33].

2.3. Атаки на реалізацію

Криптографічні примітиви та протоколи зазвичай вважаються абстрактними, коли вони розроблені. Ця математична абстракція є корисною для дослідження, але вона не охоплює всього сценарію, який може статися на практиці. Криптографічні алгоритми завжди реалізуються в програмному або апаратному забезпеченні фізичних пристроїв, на які впливає середовище. Зловмиснику може не знадобитися безпосередньо брати на себе обчислювальну складність зламу алгоритмів, щоб отримати відкритий текст або ключ. Інформація, отримана під час спостереження за обчисленнями або комунікацією конкретної реалізації, може дуже допомогти в криптоаналізі та може значно зменшити обчислювальну складність зламу криптосистеми. Далі коротко розглянемо деякі атаки на впровадження криптосистем.

Атаки з бічного каналу недорогі, реалістичні та зазвичай вважаються найнебезпечнішим типом фізичних атак. Обчислювальні пристрої пропускають інформацію не лише через взаємодію вводу-виводу, а й через фізичні характеристики обчислень, такі як енергоспоживання, час або електромагнітне випромінювання. Такий витік інформації може зламати багато криптосистем, які зазвичай використовуються. Атаки з бічного каналу можна розділити на пасивні та активні. У пасивних атаках по бічному каналу зловмисник не втручається в роботу цільової системи; в той час як під час активних атак бічним каналом (або атак через помилки) криптоаналітик має певний вплив на поведінку цільової системи. Атаки бічних каналів також можна класифікувати як інвазивні, неінвазивні та напівінвазивні. Інвазивні атаки вимагають розпакування, щоб отримати прямий доступ до внутрішніх компонентів пристрою. Неінвазивні атаки використовують лише зовнішню інформацію. Напівінвазивні атаки включають доступ до пристрою без пошкодження шару пасивації або електричного контакту з несанкціонованою поверхнею. Нижче наведено неповний список [42] атак бічними каналами. Таким чином, відомі атаки з бічних каналів включають такі:

- Атака за часом: під час атаки за часом деяка інформація про ключ або секретний параметр виводиться з часу роботи криптографічного алгоритму або пристрою. Тобто, атаки за часом використовують різницю в часі, необхідну пристрою для виконання конкретних операцій, наприклад, непостійний час для виконання двох різних інструкцій.

- Атака з аналізом потужності: в атаках з аналізом потужності цінну інформацію про операції або параметри отримують шляхом спостереження за енергоспоживанням криптографічного пристрою або модуля. Тобто, вони використовують той факт, що електронні пристрої споживають електроенергію під час роботи. Атаки з аналізом потужності можна розділити на простий аналіз потужності (SPA), де вимірювання енергоспоживання безпосередньо інтерпретуються, і диференціальний аналіз потужності (DPA), де статистичні функції застосовуються до вимірювань енергоспоживання.
- Атака електромагнітного аналізу: під час атаки електромагнітного аналізу певна інформація отримується шляхом вимірювання електромагнітних полів, що випромінюються пристроєм.
- Акустичний криптоаналіз: акустичні випромінювання можна розглядати як джерело інформації для атак бічними каналами. Приклади включають вилучення 4096-бітних ключів RSA з програмного забезпечення GnuPG за допомогою звуку, створеного комп'ютером під час розшифрування деяких вибраних зашифрованих текстів, або розпізнавання натиснутої клавіші за допомогою акустичного випромінювання клавіатури.
- Атаки на пам'ять: інформація бічного каналу з геш-значення процесора і DRAM може бути використано для криптоаналізу програмно реалізованих шифрів. Геш-значення центрального процесора знаходиться між процесором і основною пам'яттю, щоб прискорити час роботи. Атаки на основі геш-значення використовують вимірювання затримки, викликаної промахом гешу, який виникає, коли ЦП отримує доступ до даних, які не зберігалися в геш-значенні, і використовується для криптоаналізу шифрів, включаючи DES і AES [34].
- Атаки з ін'єкцією помилки: атаки з ін'єкцією помилки є активним аналогом атак із бічного каналу, коли зловмисник отримує інформацію про внутрішні стани алгоритму, провокуючи помилки в обчисленнях і порівнюючи правильний і помилковий результат. Збій може бути постійним, що незворотно пошкоджує криптографічний пристрій, або він може бути тимчасовим. Несправність може бути викликана зміною напруги, тактової частоти чи температури, або використанням світла, рентгенівського та мікрохвильового випромінювання, або вихрових струмів, викликаних магнітними полями тощо.

Пропоновані заходи протидії атакам із бічного каналу включають спеціальні рішення щодо впровадження, які пропонують лише часткове вирішення проблеми, і теоретичні рішення, які формально вирішують проблему. Стійка до витоків криптографія є активною областю досліджень, яка стосується обчислень за наявності витоку інформації та розглядає математичні рішення для вирішення атак бічними каналами. Стійка до витоків криптосистема залишається безпечною, навіть, якщо довільна, але обмежена інформація про секретний ключ і, можливо, інша інформація про внутрішній стан, потрапляє до зловмисника.

3. Моделі безпеки для криптографічних протоколів

Дійсні атаки на протоколи визначаються за допомогою моделей безпеки, а докази безпеки надають міркування щодо їх надання, часто шляхом зведення до передбачуваної складної проблеми. Важкі проблеми в більшості популярних криптографічних протоколів і примітивів відкритого ключа – це теоретико-числові проблеми, такі як проблема дискретного логарифмування та розкладання на множники, які неможливо здійснити цифровими комп'ютерами, але які можуть бути здійснені за допомогою квантового комп'ютера. При виборі параметрів безпеки для криптосистем необхідно розуміти та оцінювати вартість найвідоміших атак. Розробка захищених протоколів, які забезпечують певні бажані функції або служби безпеки для різних програм, є основною частиною сучасної криптографії.

Багато криптографічних протоколів можна сформулювати як багатостороннє обчислення (MPC). Наприклад, у протоколах обміну ключами кожен принципал має деякі секретні значення та хоче безпечно обчислити ключ сеансу без шкоди для секретних значень. У безпечному MPC набір із m сторін, кожна з яких має секретне значення x_i, хоче обчислити

спільну функцію $f(x_1,...,x_m)$, не розкриваючи жодної інформації про x_i. Правильність обчислень і конфіденційність вхідних даних є двома важливими цілями для безпечного МРС.

Поняття безпеки були визначені для протоколів МРС як в обчислювальній, так і в інформаційно-теоретичній безпеці [35].

Ідеальна модель для протоколів МРС складається з довіреної сторони, яка приватно отримує вхідні дані від учасників і обчислює для них функцію f. У реальній моделі учасники обчислюють f без довіреної сторони. Протокол вважається безпечним, якщо реальні параметри емулюють ідеальні параметри, тобто все, що криптоаналітик може отримати в реальних налаштуваннях, також можна отримати в ідеальних налаштуваннях. Хоча моделі безпеки, розроблені для загальних протоколів, будуть використовуватися для будь-якого криптографічного протоколу, простіше та ефективніше працювати з моделями безпеки, які спеціально розроблені для певного типу протоколу. Найбільш спеціалізовані та добре розроблені моделі захисту для аналізу криптографічних протоколів у обчислювальних налаштуваннях присвячені протоколам AKE та PAKE, які будуть розглянуті далі.

3.1. Моделі безпеки для AKE протоколів

Протоколи AKE є найбільш добре вивченим типом протоколів безпеки. У 1976 р. Діффі та Геллман [36] представили протокол узгодження ключів, який є вразливим до атак MITM через відсутність автентифікації.

Протоколи AKE повинні забезпечувати певні атрибути безпеки, і вони повинні протистояти добре відомим атакам, представленим у підрозд. 2.1. Актуальність є важливим атрибутом у протоколах обміну ключами. Встановлений ключ має бути новим, а не повторним з попередніх сеансів. Актуальність може бути забезпечена за допомогою часових позначок, nonces або лічильників [16]. Спираючись на роботу [42] очікується, що протокол узгодження ключів забезпечить пряму секретність, безпеку відомого ключа та спільний контроль ключів і буде стійким до атаки Деннінга – Сакко [6]. Попередня секретність зберігає безпеку ключів сеансу після розкриття матеріалу ключів, який використовується в протоколі для узгодження ключів сеансу. Безпека відомого ключа означає, що кожен запуск протоколу повинен створювати унікальний ключ сеансу. Злом сеансового ключа не повинен загрожувати іншим сеансовим ключам. Спільний контроль ключів гарантує, що всі передбачувані учасники залучені до генерації сеансового ключа, і гарантує, що жодна сутність не зможе змусити сеансовий ключ потрапити у заздалегідь визначений інтервал. Стійкість до атак Деннінга – Сакко не дозволяє зловмиснику відновити або вгадати секретні параметри, які використовуються в протоколі, після розкриття ключа сеансу.

Хоча протоколи AKE можна розглядати як особливі випадки МРС, багато моделей безпеки були спеціально розроблені для протоколів AKE в обчислювальних налаштуваннях. Першою ігровою моделлю безпеки для протоколів AKE була модель Bellare-Rogaway (BR93) [37], яка охоплювала взаємну автентифікацію та обмін ключами від попередньо спільних симетричних ключів. Пізніше Bellare і Rogaway представили модель BR95, яка є розширенням моделі BR93 і охоплює обмін ключами на основі сервера. У моделях BR безпека протоколу KE визначається в термінах розрізнення встановлених ключів сеансу від випадкових значень під час гри з криптоаналітиком PPT. Зловмисник має доступ до будь-яких публічних даних і контролює всі комунікації, взаємодіючи з набором оракулів, кожен з яких представляє екземпляр принципала в певному виконанні протоколу. Зловмисник взаємодіє з принципалами за допомогою запитів, які в основному є Send, Reveal, Corrupt і Test. Send дозволяє криптоаналітику змусити принципалів запускати протокол. Reveal моделює здатність зловмисника знаходити старі ключі сеансу. Corrupt моделює інсайдерські атаки криптоаналітика, повертає внутрішній стан оракула та встановлює довгостроковий ключ принципала на значення, вибране зловмисником. Потім зловмисник може контролювати поведінку пошкодженого принципала за допомогою надсилання запитів. Успіх зловмисника вимірюється з точки зору його переваги в розрізненні ключа сеансу від випадкового непов'язаного значення після

виконання тестового запиту. Моделі BR стали важливою віхою та призвели до появи інших моделей безпеки. Блейк-Вілсон і Менезес розширили моделі BR, щоб охопити відкриті ключі та угоду ключів. Bellare, Pointcheval і Rogaway розширили та модифікували модель BR95, а також представили модель BPR для протоколів на основі паролів. Далі наведено основні моделі безпеки для AKE протоколів:

- Модель CK01: Белларе, Канетті та Кравчик запровадили загальну структуру та модульний підхід для розробки та аналізу протоколів автентифікації та обміну ключами, який іноді називають моделлю BCK98. Згодом Канетті та Кравчик представили модель CK01 [38], яка вирішила проблеми з моделлю BCK98 і забезпечила прототип сучасної моделі безпеки шляхом поєднання моделей BR і BCK98. Основні ідеї в моделі CK01 подібні до ідей моделей BR, але модель CK01 дозволяє розкривати стани сеансу, що фіксує більше атрибутів безпеки. Після моделі BCK98 у моделі CK01 визначено дві змагальні моделі, а саме конкурентну модель неавтентифікованих посилань (UM) і змагальну модель автентифікованих посилань (AM). У моделі UM супротивник є активним і має повний контроль над комунікаційними лініями; у той час як у моделі AM супротивник не може фабрикувати повідомлення і може доставляти лише повідомлення, справді створені сторонами, без будь-яких змін чи доповнень.

Модель CK01 використовувалася для аналізу безпеки багатьох протоколів AKE. Однак вона отримала певну критику. Модель CK01 не надає конкретного визначення ідентифікаторів сесії. Крім того, визначення стану сеансу залежить від розробників протоколів, що може спричинити неоднозначність у доказах безпеки протоколів та їх реалізації. Будь-яка реалізація, в якій локальний стан (як показано супротивнику) містить більше інформації, ніж відповідне визначення в доказах, виходить за межі доказу. Інша проблема полягає в тому, що модель CK01 не фіксує деякі важливі атаки, такі як атака KCI та її варіанти.

- Модель HMQV: вирішує проблему того, що модель безпеки CK01 не фіксує атаки KCI. Докази конструкції та безпеки протоколу HMQV базувалися на новій формі підписів виклик-відповідь, яка була отримана зі схеми ідентифікації Шнорра. Для досягнення кращої продуктивності перевірки відкритого ключа, обов'язкові в протоколах MQV, були виключені з протоколів HMQV. Це робить протокол HMQV вразливим до атак малих підгруп, що дозволяє зловмиснику відновити статичний особистий ключ жертви. Хоча модель безпеки HMQV якимось чином ігнорувала модульний підхід у моделі безпеки CK01, але вона враховувала деякі інші поняття, такі як стійкість до атаки KCI та слабка ідеальна пряма секретність (wPFS).
- Модель eCK (extended-CK) була представлена LaMacchia та ін. [40] і є розширенням моделі CK01. Він усуває деякі недоліки в моделях BR і CK01. Зокрема, криптоаналітик може отримати тимчасові секрети, які належать до тестової сесії. Криптоаналітик може отримати довгостроковий ключ тестової сесії та свого партнера ще до завершення сесії. Модель eCK допускає різні комбінації для витоку довгострокових і тимчасових секретних ключів, але не обидва витоки відбуваються в одній сутності.

Порівняння між моделями CK01, HMQV і eCK наведено в [39], де зроблено висновок, що ці моделі непорівнянні, тобто безпека в кожній із цих трьох моделей не означає безпеку в двох інших моделях.

Кілька протоколів було запропоновано як вдосконалення протоколів MQV і HMQV, і для цих протоколів було представлено кілька моделей безпеки. Приклади включають протокол CMQV у моделі eCK, протокол UP у моделі Менезеса-Устаоглу (eCK+), протокол SMQV у моделі seCK, протокол FHMQV у моделі FHMQV і протокол UP+ у моделі vCK.

3.2. Моделі безпеки для протоколів PAKE

Протоколи AKE (PAKE) на основі паролів дозволяють двом або більше об'єктам автентифікувати один одного та спільно використовувати криптографічний ключ на основі попереднього спільного пароля, який запам'ятовує людина. Через низьку ентропію паролів такі

протоколи схильні до онлайн-атак підбору пароля, яким можна запобігти, обмеживши кількість невдалих спроб на стороні сервера. Мета протоколів PAKE полягає в тому, щоб єдиною реальною атакою була атака підбору пароля в режимі онлайн. Протоколи PAKE повинні бути стійкими до атак з підбіркою пароля в режимі офлайн і невиявлених онлайн [6].

Більшість існуючих протоколів PAKE мають докази або в моделі BPR, або в моделі BMP. Хоча ці моделі забезпечують певний рівень безпеки, вони мають обмеження щодо розповсюдження паролів. Тоді Канетті запропонував ідеальну функціональність для протоколів PAKE в універсально складеній (UC) структурі, де середовище емулює будь-який розподіл, неправильні паролі та пов'язані паролі. Однак він все ще не в змозі зафіксувати деякий витік інформації, який може статися в реальності.

Модель BPR була представлена як варіант моделі BR95 для протоколів PAKE. Він мав на меті боротися з вгадуванням пароля, секретністю пересилання, компрометацією сервера та втратою ключів сеансу. Подібно до моделі безпеки BR, перевага криптоаналітика в атаці визначається як подвоєна ймовірність того, що він виграє мінус один. Модель BPR забезпечує гарантії автентифікації. Зловмисник порушує автентифікацію клієнт-сервер, якщо якийсь серверний оракул завершує роботу, не маючи оракула-партнера. Зловмисник порушує автентифікацію між серверами, якщо якийсь клієнтський оракул завершує роботу, не маючи оракула-партнера. Криптоаналітик порушує взаємну автентифікацію, якщо якийсь оракул завершує роботу, не маючи оракула-партнера. Однак показано, що модель BPR є найслабшою моделлю безпеки серед моделей безпеки BR93, BR95 і CK01. Вона не фіксує деякі атаки, включаючи атаку UKS.

4. Формальна верифікація протоколів безпеки

Формальна перевірка коректності програмного забезпечення є важливою частиною практичної та теоретичної інформатики [42]. Оскільки протоколи безпеки можна розглядати як короткі програми або алгоритми, можна адаптувати методи коректності програмного забезпечення та інструменти для перевірки протоколів безпеки. Однак міркування про складність, важливість протоколів безпеки та той факт, що проблему безпеки за наявності зловмисника неможливо виявити за допомогою функціонального тестування програмного забезпечення, вказують на те, що існує потреба в спеціалізованих інструментах.

Перевірка протоколу безпеки означає перевірку того, що протокол правильний і працює відповідно до своїх цілей безпеки. Перевірка може вказувати на приклади збоїв або недоліків у аналізованому протоколі. Одночасне виконання протоколів, де об'єкт може мати різні ролі у різних виконаннях (наприклад, як ініціатор або відповідач), а також багатопротокольні атаки, згадані в підрозд. 2.1, роблять аналіз дуже складним, що не може бути охоплено евристичною перевіркою. Проблема верифікації є нерозв'язною в найзагальнішому вигляді. Для необмеженого розміру повідомлення за наявності активного супротивника або необмеженої кількості сеансів простір станів для дослідження є нескінченним, а проблема нерозв'язною. Однак збереження секретності є NP-складним для обмеженої кількості сеансів протоколу щодо моделі Долева – Яо [41] і вирішальним для необмеженої кількості сеансів за деяких додаткових обмежень.

Формальні методи, як визначено Медоузом, – це комбінація математичної або логічної моделі системи та її вимог разом із ефективною процедурою для визначення того, чи є доказ того, що система задовольняє вимоги, правильним. Формальна перевірка протоколів безпеки може розглядатися відповідно до специфікацій протоколу або реалізацій. Ідеальною метою є мати повністю автоматизований інструмент, який перевіряє безпеку реалізованого протоколу, але ця мета ще далека від досягнення.

Формальні методи перевірки протоколів безпеки відповідно до їх специфікацій можна загалом розділити на перевірку моделі та доведення теорем. У підході перевірки моделі будується кінцевий автомат, стани якого є всіма можливими проміжними станами виконання протоколу. Потім усі можливі виконання перевіряються на відповідність набору умов корек-

тності, щоб знайти атаку на протокол. Цей метод перевіряє, чи не досягнуто стану з небажаною властивістю, яка може вказувати на атаку. Правильність визначається просто через невдачу в пошуку атаки. Методи перевірки моделі, як правило, більше підходять для пошуку атак на протоколи, а не для підтвердження їх правильності. Через можливий паралельний сеанс протоколи безпеки загалом мають нескінченну кількість станів. Таким чином, відсутність атаки в кінцевій моделі не обов'язково означає відсутність атаки в нескінченному стані. Крім того, кількість станів у скінченній моделі може бути занадто великою і може сильно збільшуватися зі збільшенням кількості учасників і виконаних кроків. Методи перевірки моделі можуть забезпечити атаку, якщо виявлено, що протокол не задовольняє умові коректності. Однак вони не дають символічного доказу безпеки протоколу, якщо атаку не виявлено. У підході доведення теореми розглядаються та перевіряються всі можливі варіанти виконання протоколу на відповідність набору умов коректності. Ці методи, як правило, більше підходять для підтвердження правильності, а не для пошуку атаки на протоколи. Вони можуть використовувати аксіоматичний (дедуктивний) або індуктивний підхід.

Формальні методи перевірки протоколів безпеки іноді називають символьними моделями. У символьних моделях криптографічні примітиви розглядаються як ідеальні чорні скриньки. Однак вони мають ту перевагу, що спрощують створення інструментів автоматичної перевірки, а також існують численні ефективні інструменти для аналізу символьного протоколу.

На відміну від символьних моделей, криптоаналітик в обчислювальних моделях не виконує заздалегідь визначених дій для аналізу повідомлень, а моделюється як довільний алгоритм PPT. Обчислювальні моделі визначили сильнішу здатність змагатися, яка ближче до реального виконання протоколів. Вони дозволяють вибірково порушувати принципали під час виконання протоколу, наприклад, їх короткострокові чи довгострокові секрети або результати проміжних обчислень.

Таким чином, обчислювальні моделі забезпечують надійніші гарантії безпеки, наприклад, ідеальну пряму секретність або стійкість до атак із розкриттям стану. Однак вони здебільшого розроблені лише для протоколів ключових угод. Крім того, докази в обчислювальній моделі важче автоматизувати.

Висновки

1. Розробка потужних квантових комп'ютерів є загрозою для криптографії з відкритим ключем, яка використовується сьогодні. Постквантова криптографія пропонує квантово безпечну альтернативу криптосистемам із відкритим ключем, які зараз використовуються. Ці схеми можуть бути реалізовані на звичайних апаратних засобах.

2. Загрози в інформаційній безпеці можна розділити на чотири великі класи: розкриття або несанкціонований доступ до інформації; обман або прийняття неправдивих даних; порушення, переривання або запобігання коректній роботі; та узурпація або несанкціонований контроль деякої частини системи. Атаки також можна розділити на пасивні та активні.

3. Існує велика кількість можливих векторів атаки, які повинні бути враховані при реалізації криптографічної схеми. З часом ці атаки стали доступнішими для криптоаналітиків, оскільки вартість обладнання, для здійснення різних видів атак, продовжувала знижуватися. Обробка великих наборів вимірів також може потребувати значних обчислювальних ресурсів. З появою хмарних сервісів тепер стало набагато простіше застосовувати масштабований ресурс обробки на вимогу без вкладень у початкові витрати на інфраструктуру.

4. Захист від переліку атак, наведених у розд. 2, не гарантує безпеку протоколу, але можна очікувати, що новий протокол не успадковує помилки попередніх проєктів протоколів. Крім того, модель безпеки розглядає найважливіші атаки, які відбуваються в реальності. Це хороший вимір для оцінки моделей безпеки: якщо модель безпеки не дозволяє зловмиснику виконувати атаки, які можуть мати місце в реальності, докази безпеки будуть марними,

тому що будуть практичні сценарії порушення протоколу, які не зафіксовані у модель безпеки.

5. Потужність і складність атак покращилися завдяки вдосконаленню методів аналізу, які розвивалися від простих компараторів, таких як різниця середніх, за допомогою диференційного аналізу з використанням кореляції Пірсона, до новітніх математичних методів обробки та статистичних методів, таких як аналіз інформації. Оскільки вдосконалення атак продовжуються, як нові, так і існуючі контрзаходи потребуватимуть постійної оцінки, щоб гарантувати, що вони продовжують забезпечувати необхідний рівень захисту. Таким чином, у розд. 3 було розглянуто моделі безпеки, які були запропоновані для оцінки безпеки та боротьби з постійною загрозою від перерахованих в розд. 2 класів атак.

6. Розробка захищених протоколів, які забезпечують певні бажані функції або служби безпеки для різних програм, є основною частиною сучасної криптографії. З іншої точки зору, конструкцію будь-якої криптографічної схеми можна розглядати як проєкт безпечного протоколу для реалізації відповідної функціональності.

7. Протокол вважається безпечним, якщо реальні параметри емулюють ідеальні параметри, тобто все, що криптоаналітик може отримати в реальних налаштуваннях, також можна отримати в ідеальних налаштуваннях. Хоча моделі безпеки, розроблені для загальних протоколів, будуть використовуватися для будь-якого криптографічного протоколу, простіше та ефективніше працювати з моделями безпеки, які спеціально розроблені для певного типу протоколу.

8. Обчислювальні моделі забезпечують надійніші гарантії безпеки, наприклад, ідеальну пряму секретність або стійкість до атак із розкриттям стану. Однак вони здебільшого розроблені лише для протоколів ключових угод. Крім того, докази в обчислювальній моделі важче автоматизувати.

9. Найбільш спеціалізовані та добре розроблені моделі безпеки для аналізу криптографічних протоколів у обчислювальних налаштуваннях присвячені протоколам AKE та PAKE, були розглянуті в даній статті.

10. Також досі залишається багато відкритих питань щодо постквантової криптографії. З одного боку, стійкість до атак бічними каналами і безпека впровадження цих криптосистем ще недостатньо досліджені. З іншого боку, необхідні подальші дослідження можливих криптоаналітичних досягнень, як з класичними, так і з квантовими комп’ютерами. З огляду на ці питання у Європейському союзі було створено багато проєктів та ініціатив щодо дослідження багатьох питань, що стосуються постквантової криптографії та зокрема побудови великомасштабного квантового комп’ютера. Також досі залишається багато відкритих питань щодо постквантової криптографії. З одного боку, стійкість до атак бічними каналами і безпека впровадження цих криптосистем ще недостатньо досліджені. З іншого боку, необхідні подальші дослідження можливих криптоаналітичних досягнень, як з класичними, так і з квантовими комп’ютерами. З огляду на ці питання у Європейському союзі було створено багато проєктів та ініціатив щодо дослідження багатьох питань, що стосуються постквантової криптографії та зокрема побудови великомасштабного квантового комп’ютера.

Список літератури:

1. John Preuß Mattsson, Ben Smeets and Erik Thormarker Quantum-Resistant Cryptography. Ericsson Security Research. [Електронний ресурс]. Режим доступу: https://arxiv.org/ftp/arxiv/papers/2112/2112.00399.pdf.

2. M. Bishop, Introduction to computer security. Prentice Hall PTR, 2004.

3. W. Stallings. Cryptography and Network Security: Principles and Practice, 6th ed. Pearson Education, 2014.

4. S. Gritzalis, D. Spinellis Cryptographic protocols over open distributed systems: A taxonomy of flaws and related protocol analysis tools, in Safe Comp 97. Springer London, 1997, pp. 123–137. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/978-1-4471-0997-6.

5. A. J. Menezes, P. C. Van Oorschot, S. A. Vanstone, Handbook of applied cryptography. CRC press, 1996.

6. M. Toorani Cryptanalysis of a new protocol of wide use for email with perfect forward secrecy // Security and Communication Networks, vol. 8, no. 4, pp. 694-701, 2015.

7. R. Bird, I. Gopal, et al. Systematic design of a family of attack-resistant authentication protocols, Selected Areas in Communications // IEEE Journal on, vol. 11, no. 5, pp. 679–693, Jun 1993.

8. C. Boyd, A. Mathuria. Protocols for authentication and key establishment // Springer Science & Business Media, 2003.

9. M. Toorani. On vulnerabilities of the security association // IEEE 802.15.6 standard, in Financial Cryptography and Data Security, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2015, vol. 8976, pp. 245–260.

10. M. Toorani, A. Beheshti. Solutions to the GSM security weaknesses // Proceedings of the Second International Conference on Next Generation Mobile Applications, Services, and Technologies (NGMAST'08), Sept 2008, pp. 576–581.

11. H. Xia, J. C. Brustoloni. Hardening web browsers against man-in-the-middle and eavesdropping attacks // Proceedings of the 14th International Conference on World Wide Web. New York, 2005, pp. 489–498. [Електронний ресурс]. Режим доступу: http://doi.acm.org/10.1145/1060745.1060817.

12. S. Murdoch, S. Drimer, R. Anderson, M. Bond Chip and pin is broken // Security and Privacy (SP), 2010 IEEE Symposium on, May 2010, pp. 433–446.

13. M. Toorani. Cryptanalysis of a robust key agreement based on public key authentication // Security and Communication Networks, vol. 9, no. 1, pp. 19–26, 2016.

14. B. S. Kaliski. An unknown key-share attack on the MQV key agreement protocol // ACM Transactions on Information and System Security (TISSEC), vol. 4, no. 3, pp. 275–288, 2001.

15. S. Blake-Wilson, A. Menezes. Unknown key-share attacks on the station-to-station (sts) protocol // Public Key Cryptography, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1999, vol. 1560, pp. 154–170. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-49162-7 12.

16. L. Gong Variations on the themes of message freshness and replay // Proceedings of the Computer Security Foundations Workshop VI, vol. 6. Citeseer, 1993, pp. 131–126.

17. C. J. Mitchell, L. Chen. Comments on the s/key user authentication scheme // ACM SIGOPS Operating Systems Review, vol. 30, no. 4, pp. 12–16, Oct. 1996.

18. M. Eian, S. F. Mjølsnes. The modeling and comparison of wireless network denial of service attacks // Proceedings of the 3rd ACM SOSP Workshop on Networking, Systems, and Applications on Mobile Handhelds (MobiHeld'11), 2011, pp. 7:1–7:6. [Електронний ресурс]. Режим доступу: http://doi.acm.org/10.1145/2043106.2043113.

19. A. Shamir, R. Rivest, L. Adleman Mental poker // The Mathematical Gardner. Springer US, 1981, pp. 37–43. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/978-1-4684-6686-7.

20. M. Naor, M. Yung Public-key cryptosystems provably secure against chosen ciphertext attacks // Proceedings of the 22nd Annual ACM Symposium on Theory of Computing (STOC'90). New York, NY, USA: ACM, 1990, pp. 427–437.

21. C. Rackoff, D. R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack // Advances in Cryptology – CRYPTO'91, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1992, vol. 576, pp. 433–444.

22. E. Biham, A. Shamir. Differential cryptanalysis of des-like cryptosystems // Advances in Cryptology-CRYPTO'90, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1991, vol. 537, pp. 2–21. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-38424-3.

23. H. Wu, B. Preneel. Differential cryptanalysis of the stream ciphers Py, Py6 and Pypy // Advances in Cryptology – EUROCRYPT 2007, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, vol. 4515, pp. 276–290. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/978-3-540-72540-4.

24. M. Matsui, A. Yamagishi. A new method for known plaintext attack of feal cipher, in Advances in Cryptology – EUROCRYPT'92, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1993, vol. 658, pp. 81–91. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-47555-9.

25. G. Bard. Algebraic cryptanalysis // Springer Science & Business Media, 2009.

26. E. Biham, O. Dunkelman, N. Keller. New combined attacks on block ciphers // Fast Software Encryption, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2005, vol. 3557, pp. 126–144. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/11502760.

27. B. Zhu, G. Gong. Multidimensional meet-in-the-middle attack and its applications to katan32/48/64 // Cryptography and Communications, vol. 6, no. 4, pp. 313–333, 2014.

28. L. Knudsen, D. Wagner. Integral cryptanalysis // Fast Software Encryption, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2002, vol. 2365, pp. 112–127. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-45661-9.

29. E. Biham. New types of cryptanalytic attacks using related keys // Advances in Cryptology – EUROCRYPT'93, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1994, vol. 765, pp. 398–409. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-48285-7.

30. M. Hell, T. Johansson, L. Brynielsson An overview of distinguishing attacks on stream ciphers // Cryptography and Communications, vol. 1, no. 1, pp. 71–94, 2009.

31. M. Bellare, A. Desai, D. Pointcheval, P. Rogaway. Relations among notions of security for public-key encryption schemes // Advances in Cryptology – CRYPTO'98, ser. Lecture Notes in Computer Science. Springer Berlin

Heidelberg, 1998, vol. 1462, pp. 26–45. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/BFb0055718.

32. M. Bellare, P. Rogaway. Optimal asymmetric encryption – how to encrypt with RSA // Advances in Cryptology – EUROCRYPT'94, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1995, vol. 950, pp. 92–111.

33. J. Katz, Y. Lindell. Introduction to modern cryptography. Chapman & Hall / CRC, 2008.

34. D. A. Osvik, A. Shamir, E. Tromer. Cache attacks and countermeasures: The case of AES // Topics in Cryptology – CT-RSA 2006, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2006, vol. 3860, pp. 1–20.

35. R. Cramer, I. Damgard, J. B. Nielsen. Secure multiparty computation and secret sharing – an information theoretic approach. Book Draft, 2013.

36. W. Diffie, M. Hellman. New directions in cryptography // IEEE Transactions on Information Theory, vol. 22, no. 6, pp. 644–654, Nov 1976.

37. M. Bellare, P. Rogaway. Entity authentication and key distribution // Advances in Cryptology – CRYPTO'93, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 1994, vol. 773, pp. 232–249. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-48329-2.

38. R. Canetti, H. Krawczyk. Analysis of key-exchange protocols and their use for building secure channels // Advances in Cryptology – EUROCRYPT'01, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2001, vol. 2045, pp. 453–474. [Електронний ресурс]. Режим доступу: http://dx.doi.org/10.1007/3-540-44987-6.

39. C. Cremers. Examining indistinguishability-based security models for key exchange protocols: the case of CK, CK-HMQV, and eCK // Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security, ser. ASIACCS'11. New York, NY, USA: ACM, 2011, pp. 80–91.

40. B. LaMacchia, K. Lauter, A. Mityagin. Stronger security of authenticated key exchange // Provable Security, ser. Lecture Notes in Computer Science. Springer Berlin Heidelberg, 2007, vol. 4784, pp. 1–16.

41. D. Dolev, A. C. Yao. On the security of public key protocols // IEEE Transactions on Information Theory, vol. 29, no. 2, pp. 198–208, Mar 1983.

42. M. Toorani Security Protocols in a NutShell, Department of Informatics, University of Bergen, Norway, arXiv preprint arXiv:1605.09771, 2016. [Електронний ресурс]. – Режим доступу: https://arxiv.org/pdf/1605.09771.pdf.

Надійшла до редколегії 03.11.2022

Відомості про авторів:

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

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

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

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

УДК 004.056.5 **DOI:10.30837/rt.2022.4.211.02**

С.О. КАНДІЙ

АНАЛІЗ БЕЗПЕКИ ДСТУ 8961:2019 У МОДЕЛІ ВИПАДКОВОГО ОРАКУЛА

Вступ

ДСТУ 8961:2019 є діючим українським стандартом постквантового асиметричного шифрування та інкапсуляції ключів [1]. Стандарт грунтується на ANSI X9.98 [2], проте має ряд відмінностей. У літературі відомо не так багато робіт, що присвячені безпеці ДСТУ 8961:2019 та ANSI X9.98. Більшість досліджень присвячено або генерації загальносистемних параметрів [3], або атакам на основі редукції решіток [4].

У той же час, для механізмів інкапсуляції ключів, що є фіналістами конкурсу NIST PQC [5], у літературі можливо знайти всесторонній аналіз безпеки як для класичних атак, так і для атак з використанням квантових комп'ютерів [6, 7]. Безсумнівно, аналіз конкретних значень безпеки для наборів загальносистемних параметрів є важливим. Проте, такі оцінки мають сенс тільки у випадку, якщо криптоаналітик не може знайти "обхідні шляхи", якщо кожна атака зводиться до певного набору добре досліджених складних у теоретико-числовому сенсі проблем.

Зазвичай, для практичних застосувань механізми інкапсуляції ключів (KEM) вважаються безпечними [6, 7], якщо складність будь-якої атаки з адаптивно підібраними шифротекстами є занадто великою для практичної реалізації. Якщо KEM містить доказ того, що він є безпечним у зазначеному вище сенсі за умови складності певного невеликого набору теоретико-числових проблем при конкретних значеннях загальносистемних параметрів, то кажуть, що він є безпечним у стандартній моделі [8].

На жаль, докази у стандартній моделі часто неможливо отримати для реальних KEM. Причиною цього є використання криптографічних геш-функцій [8]. Часто відсутні інструменти для того, щоб врахувати вплив алгебраїчної структури та властивостей геш-функції на безпеку перетворень. Тому, на практиці поширеною модифікацією є модель випадкового оракула [9], у якій усі геш-функції замінюються на випадкові оракули – ідеалізовані геш-функції, що не мають внутрішньої структури.

У роботі викладено перші результати спроби комплексного аналізу безпеки ДСТУ 8961:2019 у моделі випадкового оракула. Представлено докази IND-CCA2 безпеки схеми асиметричного шифрування та інкапсуляції ключів, що описані в ДСТУ 8961:2019.

1. Структура механізмів інкапсуляції ключів

З формальної точки зору, механізм інкапсуляції ключів є трійкою алгоритмів $(Gen, Encaps, Decaps)$ [7, 8], 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 та 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.

У класичному випадку [10] від механізмів інкапсуляції ключів вимагається властивість коректності:

$$\Pr[Decaps(sk, C) = k \mid (pk, sk) \leftarrow Gen(1^{\lambda}); (K, C) \leftarrow Encaps(pk)] = 1. \quad (1)$$

Згодом вимога коректності була узагальнена на випадок [11], коли алгоритм декапсуляції у незначній кількості випадків містить помилки декапсуляції:

$$\Pr[Decaps(sk,C)=k \mid (pk,sk)\leftarrow Gen(1^{\lambda});(K,C)\leftarrow Encaps(pk)]=negl(\lambda), \quad (2)$$

де $negl(\lambda)$ позначає незначну функцію, що залежить від параметра безпеки λ. Незначна функція – це функція, що зменшується швидше за будь-який поліном. Формальне визначення буде надано у розд. 2.

Сучасні механізми інкапсуляції ключів зазвичай не будуються *ad hoc* [5]. За основу береться деяка схема асиметричного шифрування. Схема асиметричного шифрування є трійкою алгоритмів (Gen, Enc, Dec), де:

- $Gen: 1^{\lambda}\rightarrow(pk,sk)$ – поліноміальний ймовірнісний алгоритм генерації ключової пари. Приймає параметр безпеки 1^{λ} та повертає ключову пару (pk,sk).
- $Enc:(pk,m)\rightarrow C$ – поліноміальний ймовірнісний алгоритм шифрування. Приймає публічний ключ pk та повідомлення m і повертає шифротекст C.
- $Dec:(sk,C)\rightarrow\{m,\perp\}$ – детермінований поліноміальний алгоритм розшифрування. Приймає секретний ключ sk та шифротекст C і повертає повідомлення m у разі вдалої декапсуляції та символ помилки $\perp$ – у разі виникнення помилок.

Від схеми шифрування, аналогічно, вимагається коректність:

$$\Pr[Dec(sk,Enc(pk,m))=m \mid (pk,sk)\leftarrow Gen(1^{\lambda})]=1. \quad (3)$$

Або її послаблений варіант

$$\Pr[Dec(sk,Enc(pk,m))=m \mid (pk,sk)\leftarrow Gen(1^{\lambda})]=negl(\lambda). \quad (4)$$

Для перетворення схеми асиметричного шифрування на механізм інкапсуляції ключів виконується дерандомізація схеми шифрування [12]. Ймовірнісний алгоритм $Enc(pk,m)$ перетворюється на детермінований алгоритм $Enc'(pk,r,m)$, який приймає випадкове значення r. Щоб детермінований алгоритм залишався безпечним, необхідно щоб для будь-якого m виконувалася нерівність

$$|r: Enc'(pk,r,m)=C|\leq\gamma(\lambda)=negl(\lambda). \quad (5)$$

Інакше кажучи, необхідно щоб кількість значень r, за яких при фіксованих pk,m можливо отримати шифротекст C, була незначною відносно параметра безпеки λ. Вимоги безпеки до асиметричної схеми при побудові механізму інкапсуляції ключів є дещо слабшими, ніж при загальному використанні схеми [6, 7, 11, 12]. Зазвичай вимагається лише безпека до атак з адаптивно підібраним повідомленням (зазвичай у моделі IND-CPA. Більш детально розглянуто у розд. 2). Далі до схеми застосовується перетворення типу CPA-to-CCA. Наприклад, перетворення Дента [10] або Фуджісакі – Окамото [11]. Узагальнена схема побудови механізму інкапсуляції ключів представлена на рис. 1. Зауважимо, що, звичайно, існують інші підходи до розробки механізмів інкапсуляції ключів, проте описаний є найбільш поширеним, і аналіз ДСТУ 8961:2019 зручно проводити саме при такій декомпозиції КЕМ.

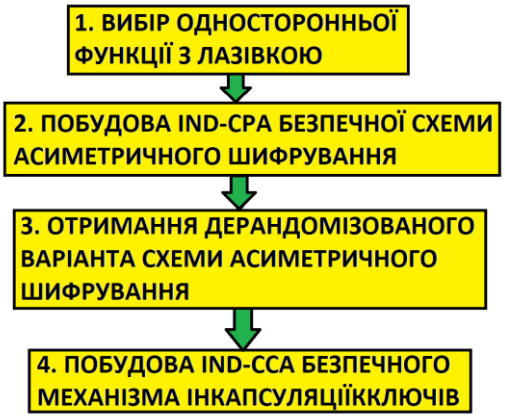


Рис. 1. Методологія побудови механізмів інкапсуляції ключів

2. Моделі безпеки на основі нерозрізнювальності

Найбільш поширеним підходом до оцінки безпеки криптографічних систем є моделі на основі нерозрізнювальності [13]. Такі моделі грунтуються на простій ідеї: якщо супротивник не зможе відрізнити шифротекст від випадкового набору бітів, то він не зможе витягти інформації з шифротексту про відкритий текст. Звичайно, на практиці завжди існує ймовірність, що криптоаналітик зможе витягти інформацію про відкритий текст. Він може хоча б просто вгадати його. Модель вимагає щоб ця ситуація мала незначну ймовірність, яку на практиці можна ігнорувати. Для формального визначення таких незначних ймовірностей використовується нотація незначних функцій [8]. Функція f є незначною, якщо для усіх поліномів p існує константа N_p, для якої виконується

$$f(x) \le \frac{1}{p(x)} \forall x \ge N_p . \tag{6}$$

Позначимо той факт, що функція $f(x)$ є незначною, як $f(x) = negl(x)$.

Звичайно, однієї нерозрізнювальності на практиці недостатньо. Криптоаналітик може методом грубої сили перебрати усі можливі варіанти, або використати певні алгебраїчні властивості для зменшення простору пошуку. Тому, усі оцінки у моделях на основі нерозрізнювальності передбачають, що існує деякий алгоритм $GenParams : \lambda \to params$, який для рівня безпеки λ генерує набір загальносистемних параметрів $params$, за яких складна задача, що лежить в основі криптографічної системи, є складною на практиці у сенсі рівня безпеки λ [11]. Зазвичай, докази у моделях на основі нерозрізнювальності носять асимптотичний характер, проте іноді доказ може використовувати конкретні властивості конкретних загальносистемних параметрів для доведення безпеки системи у обмеженій практичними потребами кількості випадків.

Найбільш поширеними моделями є [13]:

- IND-CPA (нерозрізнювальність для атак на основі підібраного відкритого текста);
- IND-CCA (нерозрізнювальність для атак на основі підібраного шифротекста).

Доказ безпеки у кожній моделі є доказом того, що криптоаналітик може отримати лише незначну перевагу при спробі відрізнити шифротекст від випадкових бітів при наявності доступу до деяких оракулів шифрування/розшифрування (інкапсуляції/декапсуляції). Структура оракулів і порядок доступу залежать від моделі. Для IND-CCA виділяють два випадки – IND-CCA1 та IND-CCA2. У IND-CCA1 доступ до оракулів можливий лише обмежений період часу.

Для деякої криптографічної системи $П$ позначимо перевагу криптоаналітика (супротивника) A у моделях IND-CPA, IND-CCA, IND-CCA2 як $Adv_{A,П}^{IND-ATK}(\lambda), ATK \in \{CPA, CCA1, CCA2\}$. Якщо виконується умова

$$Adv_{A,П}^{IND-ATK}(\lambda) = negl(\lambda), \tag{7}$$

то криптографічна система є безпечною у грі IND-ATK, де $ATK \in \{CPA, CCA1, CCA2\}$ відповідно. За визначенням криптоаналітик має відрізнити гру (експеримент) $Adv_{A,П}^{IND-ATK-0}(\lambda)$, у якій у якості віхідних даних дається справжній шифротекст, та гру $Adv_{A,П}^{IND-ATK-1}(\lambda)$, у якій у на вхід криптоаналітику даються випадкові дані. Відповідно для $ATK \in \{CPA, CCA1, CCA2\}$ маємо наступне визначення переваги:

$$Adv_{A,П}^{IND-ATK}(\lambda) = \left|\Pr\left[Exp_{A,П}^{IND-ATK-1}(\lambda) = 1\right] - \Pr\left[Exp_{A,П}^{IND-ATK-0}(\lambda) = 1\right]\right|. \tag{8}$$

Вміст ігор $Adv_{A,П}^{IND-ATK-0}(\lambda)$ та $Adv_{A,П}^{IND-ATK-1}(\lambda)$ залежить від криптографічної схеми. Для схеми асиметричного шифрування та механізмів інкапсуляції ключів він дещо відрізняється, проте сутність залишається та ж сама.

2.1. Моделі безпеки для асиметричних схем шифрування

Для асиметричної схеми шифрування $П = (Gen, Enc, Dec)$ ігри $Adv_{A,П}^{IND-ATK-0}(\lambda)$ та $Adv_{A,П}^{IND-ATK-1}(\lambda)$ для $ATK \in \{CPA, CCA1, CCA2\}$ зображені на рис. 2. Кожна гра відбувається між випробовувачем та супротивником (криптоаналітиком) [6, 13]. На початку гри випробовувач генерує випадкову ключову пару, два повідомлення m_0, m_1 та надсилає відкритий ключ і повідомлення супротивнику. Супротивник робить запити до оракула O_1 (під оракулом мається на увазі деякий алгоритм, який є "чорним ящиком" для супротивника). Після чого супротивник повідомляє випробовувачу, що готовий отримати завдання. У грі $Adv_{A,П}^{IND-ATK-0}(\lambda)$ випробовувач шифрує повідомлення m_0 та надсилає його шифротекст супротивнику, у грі $Adv_{A,П}^{IND-ATK-1}(\lambda)$ відповідні дії відбуваються для повідомлення m_1. Супротивник робить запити до оракула O_2 та генерує 0 якщо грає у гру $Adv_{A,П}^{IND-ATK-0}(\lambda)$ та 1, якщо грає у гру $Adv_{A,П}^{IND-ATK-1}(\lambda)$.

$$\begin{array}{ccc}
(pk, sk) & \leftarrow & Gen(1^{\lambda}) \\
(m_0, m_1, pk) & \rightarrow & A^{O_1}(m_0, m_1, pk) \\
b & \leftarrow & \{0,1\} \\
c & \leftarrow & Enc_{pk}(m_b) \\
b' & \leftarrow & A^{O_2}(m_0, m_1, pk, c) \\
Повернути & b' = b &
\end{array}$$

Рис. 2. Ігри $Adv_{A,П}^{IND-ATK-0}(\lambda)$ та $Adv_{A,П}^{IND-ATK-1}(\lambda)$ для асиметричної схеми шифрування

Значення оракулів O_1, O_2 для IND-CPA, IND-CCA1, IND-CCA2 зведені у табл. 1. Фактично використовується тільки оракул дешифрування.

Таблиця 1

Значення оракулів O_1, O_2 для схеми асиметричного шифрування

Модель	Оракул O_1	Оракул O_2
IND-CPA	-	-
IND-CCA1	Оракул дешифрування O_{Dec}	-
IND-CCA2	Оракул дешифрування O_{Dec}	Оракул дешифрування O_{Dec}

Оракул дешифрування може розшифрувати будь-який шифротекст, проте має деяку відмінність. Якщо шифротекст був виданий супротивнику як завдання, то оракул дешифрування незалежно від дій супротивника повертатиме $\perp$ на запит розшифрування завдання.

2.2. Моделі безпеки для механізмів інкапсуляції ключів

Для механізму інкапсуляції ключів $\Pi = (Gen, Encaps, Decaps)$ ігри $Adv_{A,\Pi}^{IND-ATK-0}(\lambda)$ та $Adv_{A,\Pi}^{IND-ATK-1}(\lambda)$ для $ATK \in \{CPA, CCA1, CCA2\}$ зображено на рис. 3. Кожна гра аналогічно відбувається між випробовувачем та супротивником (криптоаналітиком), проте вміст ігор відрізняється. Спочатку випробовувач генерує випадкову ключову пару та надсилає відкритий ключ супротивнику, який робить запити до оракула O_1. Після чого супротивник повідомляє випробовувачу, що готовий отримати завдання. У грі $Adv_{A,\Pi}^{IND-ATK-0}(\lambda)$ випробовувач генерує ключ K та його інкапсуляцію C, потім надсилає ключ і інкапсуляцію супротивнику. У грі $Adv_{A,\Pi}^{IND-ATK-1}(\lambda)$ випробовувач генерує ключ K та його інкапсуляцію C, але замість ключа надсилає супротивнику випадковий рядок бітів. Супротивник робить запити до оракула O_2 та генерує 0, якщо грає у гру $Adv_{A,\Pi}^{IND-ATK-0}(\lambda)$, та 1, якщо грає у гру $Adv_{A,\Pi}^{IND-ATK-1}(\lambda)$.

$$
\begin{array}{ccc}
(pk, sk) & \leftarrow & Gen(1^{\lambda}) \\
(pk) & \rightarrow & A^{O_1}(pk) \\
b & \leftarrow & \{0,1\} \\
(K_0, C) & \leftarrow & Encaps(pk) \\
K_1 & \leftarrow & \{0,1\}^{KeyLen} \\
b' & \leftarrow & A^{O_2}(pk, C, K_b) \\
\textit{Повернути} & b' == b &
\end{array}
$$

Рис. 3. Ігри $Adv_{A,\Pi}^{IND-ATK-0}(\lambda)$ та $Adv_{A,\Pi}^{IND-ATK-1}(\lambda)$ для механізму інкапсуляції ключів

Значення оракулів O_1, O_2 для IND-CPA, IND-CCA1, IND-CCA2 зведені у табл. 2. Фактично використовується тільки оракул декапсуляції.

Таблиця 2

Значення оракулів O_1, O_2 для механізму інкапсуляції ключів

Модель	Оракул O_1	Оракул O_2
IND-CPA	-	-
IND-CCA1	Оракул декапсуляції O_{Decaps}	-
IND-CCA2	Оракул декапсуляції O_{Decaps}	Оракул декапсуляції O_{Decaps}

Оракул декапсуляції діє аналогічно до оракула дешифрування та може декапсулювати ключ з інкапсуляції з обмеженнями на інкапсуляцію, що обрана у якості завдання.

Варто зазначити, що моделі $IND-CPA, IND-CCA1, IND-CCA2$ утворюють ланцюг включень:

$$IND-CPA \subseteq IND-CCA1 \subseteq IND-CCA2. \quad (9)$$

Тобто з безпеки у моделі IND-CCA2 випливає безпека у моделі IND-CCA1 і з безпеки у моделі IND-CCA1 випливає безпека у моделі IND-CPA.

2.3. Модель випадкового оракула

Докази у моделях на основі нерозрізнювальності мають, як правило, схожу структуру. В основі лежить техніка доказу "GAME HOPING" [9, 14]. Сутність техніки полягає у тому, що спочатку розглядається оригінальна гра, для якої складно зробити оцінку переваги супротивника. Далі гра дещо змінюється і вимірюється як зміниться при цьому перевага супротивника. Після серії таких змін можливо отримати гру, для якої легко оцінити перевагу, наприклад звести до гри, у якій супротивнику завжди даються виключно випадкові дані. Оскільки для останньої гри оцінка відома і відомо як змінюється перевага при переходах від однієї гри до іншої, то можливо отримати обмеження зверху для оригінальної гри.

Проблема такого підходу полягає у тому, що для деяких моментів важко оцінити як зміниться вплив на безпеку при їх заміні. В першу чергу це стосується геш-функцій. В їх алгебраїчній структурі можуть міститися вразливості або комбінація їх алгебраїчних властивостей з властивостями криптографічної системи може приводити до вразливостей.

На жаль, у загальному випадку рішень цієї проблеми невідомо. У сучасних розробках використовується так звана модель випадкового оракула, у якій геш-функції або елементи, що ведуть себе схожим чином, замінюються на випадкові оракули. Випадковий оракул – це ідеальна геш-функція, що не має структури. Супротивник замість викликів реальних функцій робить виклики випадкових оракулів для функцій, чий вплив важко оцінити. Існують приклади систем, які є безпечними у моделі випадкового оракула, проте небезпечні на практиці. Незважаючи на те, що таке послаблення дещо виключає з розгляду ряд атак, модель випадкового оракула де-факто є стандартом у сучасній криптографії. Безпека переважної більшості криптографічних систем оцінюється саме у моделі випадкового оракула. Більш того, доказ у моделі випадкового оракула дозволяє використовувати криптографічну схему як будівельний блок у більш складних моделях безпеки криптографічних протоколів на кшталт моделі Канетті – Кравчека [15] для протоколів узгодження ключів.

3. Відомості з теорії решіток

Введемо позначення: нехай $R_q = \mathbb{Z}_q[X]/(X^n - X - 1)$ – кільце поліномів над $\mathbb{Z}_q$ з твірним поліномом $X^n - X - 1$, а R_3 – множина поліномів кільця R_q, усі коефіцієнти яких належать до множини $\{-1, 0, 1\}$. Позначимо 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 грунтується на проблемі NTRU [16]. У загальному вигляді проблему NTRU можливо визначити наступним чином. Нехай $q > 2$ – ціле число, $\gamma > 2$ – дійсне чило, D – деякий розподіл ймовірностей над полем R_q. Проблема (D, γ, γ')-NTRU (проблема

пошуку) полягає у тому, щоб для $h \in R_q$ з розподілу D знайти пару $(f, g) \in R_q \times R_q \setminus \{0,0\}$, для яких виконується $g \cdot h = f \bmod q$ та $\| f \|, \| g \| \leq \sqrt{q} / \gamma$.

NTRU-припущенням є припущення, що для будь-якого рівня безпеки λ існують такі параметри для задачі NTRU, що перевага будь-якого алгоритма у вирішенні проблеми NTRU є незначною (у сенсі визначень в розд. 2).

На основі проблеми NTRU природнім чином будується одностороння функція з лазівкою. Функція $NTRU_h : R_q \times R_q \to R_q$ для деякого полінома $h \in R_q$ визначається так:

$$NTRU_h(m, r) = m + rh. \tag{10}$$

Поліном h задає базис-решітку з базисом $(1, h) \in R_q \times R_q$. Відображення $NTRU_h$ семплує деяку точку на цій решітці. Ісходні координати точки відновити доволі важко, якщо відомий тільки "поганий" базис $(1, h) \in R_q \times R_q$, проте, якщо відомий деякий гарний базис цієї ж решітки $(f, g) \in R_q \times R_q$, то задачу можливо вирішити за поліноміальний час. Припущення, що така функція (при відповідних параметрах) є односторонньою функцією з лазівкою, є необхідним припущенням для безпеки ДСТУ 8961:2019.

4. ДСТУ 8961:2019

Перед тим, як перейти до розгляду ДСТУ 8961:2019, зауважимо, що текст стандарту є доволі технічним і опис перетворення містить багато технічних деталей, які не впливають на аналіз у моделі випадкового оракула. Тому для полегшення аналізу введемо наступні геш-функції, які є еквівалентними до перетворень, що використовуються у ДСТУ 8961:2019:

$$\begin{gathered} BPGM : \{0,1\}^{8*\max MsgLenBytes + db} \times R_q \to R_q \\ MGF : R_q \to \{0,1\}_3 \\ H : R_q \to \{0,1\}^{\lambda} \\ KDF : R_q \to \{0,1\}^{K\_bytes} \end{gathered},$$

де λ – параметр безпеки, t – загальносистемний параметр, а maxMsgLenBytes, db, K_bytes – константи, що визначаються на основі загальносистемних параметрів. Також використовується бієктивне відображення:

$$\begin{gathered} Pad : \{0,1\}^{8*\max MsgLenBytes} \times \{0,1\}^{db} \to R_3 \\ Pad^{-1} : R_3 \to \{0,1\}^{8*\max MsgLenBytes} \times \{0,1\}^{db} \end{gathered},$$

яке кодує повідомлення (у вигляді строки бітів), довжину повідомлення та випадкову строку бітів у поліном з R_3.

Механізм інкапсуляції ключів ДСТУ 8961:2019 використовує у якості CPA-to-CCA перетворення гібридний варіант перетворень Дента власної розробки [10, 3] для отримання IND-CCA2 безпечного механізму інкапсуляції ключів з ймовірнісної схеми асиметричного шифрування. Схема асиметричного шифрування, що лежить в основі протоколу інкапсуляції ключів, наведена на рис. 4. Протокол інкапсуляції ключів наведено на рис. 5.

SkelyaPKE.Gen(1^λ):
1. $f \leftarrow_R R_3^{2t}$
2. $g \leftarrow_R R_3^{\left\lfloor \frac{2n}{3}+1 \right\rfloor}$
3. $if\ (3f+1)^{-1} = \perp$ goto 1
4. $h = (3f+1)^{-1} g \in R_q$
5. $return\ (pk = h, sk = f$

SkelyaPKE.Enc($msg, coins, h$):
1. $m = Pad(msg, coins)$
2. $r = BPGM(msg, coins, h)$
3. $R = rh$
4. $m' = m + MGF(R)$
5. $if\ m' \notin R_3^{2t,n-2t}$ return $\perp$
6. $c = R + m'$
7. $return(c)$

SkelyaPKE.Dec($c, (f, h)$):
1. $a = fc$
2. $m' = a\ mod\ 3$
3. $if\ m' \notin R_3^{2t,n-2t}$ return $\perp$
4. $R = c - m'$
5. $m = m' - MGF(R)$
6. $(msg, coins) = Pad^{-1}(m)$
7. $r' = BPGM(msg, coins, h)$
8. $R' = r'h$
9. $if\ R' = R\ return\ msg$
10. $return\ \perp$

Рис. 4. Ймовірністна асиметрична схема шифрування у ДСТУ 8961:2019

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. 5. Механізм інкапсуляції ключів ДСТУ 8961:2019

ДСТУ 8961:2019 підтримує три режими роботи для 256, 384 та 512 біт безпеки. Опис загальносистемних параметрів та їх значення для рівнів безпеки наведений у табл. 3.

Таблиця 3

Загальносистемні параметри для 256, 384, 512 біт безпеки

Параметр	Опис	256 біт	384 біта	512 біт
N	Ступінь поліномів	881	1201	1471
q	Великий модуль (модуль поля)	7673	9221	12251
p	Малий модуль	3	3	3
t	Визначає кількість ненульових коефіцієнтів в малих поліномах	159	192	255
d_g	Кількість ненульових коефіцієнтів в поліномі g. Визначається за формулою $d_g = \lfloor N/3 \rfloor$	293	400	490
d_f	Кількість ненульових коефіцієнтів в поліномі f. Визначається за формулою $d_f = 2t$	318	384	510
d_r	Кількість ненульових коефіцієнтів в поліномі r. Визначається за формулою $d_r = 2t$	318	384	510

5. Конкретні оцінки безпеки

Оскільки перетворення CPA-to-CCA гарантує безпеку механізму інкапсуляції ключа, то загальносистемні параметри, які забезпечують безпеку схеми асиметричного шифрування, будуть також безпечними і для механізму інкапсуляції ключів, то можливо сфокусуватися лише на схемі асиметричного шифрування.

На високому рівні, за умови безпеки загальної конструкції схеми, атаки для ДСТУ 8961:2019 можна поділити на три класи [17]:

- Атаки, що передбачають знання деякої кількості пар текст/шифротекст, які викликають помилки дешифрування.
- Атаки, направлені на односторонню функцію $NTRU_h$ у процедурі шифрування. Тобто, знайти прообраз функції для $e = NTRU_h(m', r)$.
- Атаки, направлені на односторонню функцію $NTRU_h$ у процедурі генерації ключової пари. Тобто, знайти прообраз функції для $NTRU_h(-f, g) = 0 \bmod q$.

Розглянемо перший клас атак. Інтуітивно зрозуміло, що чим складніше знайти помилку дешифрування, тим складніше реалізувати таку атаку. Помилка дешифрування відбувається у випадку, якщо $p(rg + m'F)$ має хоча б один коефіцієнт з абсолютним значенням більшим за $q/2$. Оскільки для будь-яких $u, v \in R_q$ має місце нерівності:

$$\begin{array}{l} \| uv \|_\infty \le 2 \| u \|_\infty \cdot \| v \|_1 \\ \| uv \|_\infty \le 2 \| u \|_2 \cdot \| v \|_2 \end{array}. \tag{11}$$

То, враховуючи, що $\| m' \|_\infty = \| g \|_\infty = 1$ та $\| F \|_1 = \| r \|_1 = 2t$, маємо умову відсутності помилок дешифрування:

$$\begin{array}{c} \| prg + m'F \|_\infty \le \| m \|_\infty + 3 \| m'F + rg \|_\infty \le \\ | 1 + 6(\| m' \|_\infty \| F \|_1 + \| g \|_\infty \| r \|_1) \le 1 + 24t < q/2 \end{array}. \tag{12}$$

Отже, якщо $q \ge 48t + 3$, то помилки дешифрування в ДСТУ 8961:2019 відсутні. Оскільки ця умова виконується для усіх наборів загальносистемних параметрів в стандарті, то помилки дешифрування відсутні і атаки такого класу відсутні для цієї криптосистеми.

У роботі [18] запропоновано техніку, яка показує як звести задачу вирішення рівняння $NTRU_h(-f, g)$ до знаходження прообраза $NTRU_h(m', r)$, що поєднує ці два вектори атак в один напрямок. Задача криптоаналізу ДСТУ 8961:2019 зводиться до криптоаналізу проблеми NTRU. Робіт з криптоаналізу NTRU існує чимало, зокрема для ДСТУ 8961:2019 авторами опубліковано методологію оцінки для гібридних атак та атак з використанням редукції решіток [3]. Детальний огляд результатів щодо редукції решіток та гібридним атакам виходить за межі цієї роботи.

Наостанок зауважимо, що ДСТУ 8961:2019 використовує поле $\mathbb{Z}_q[X]/(X^n - X - 1)$, яке не має нетривіальних підполів. Такий вибір дає захист від алгебраїчних атак на підполе [19] та S-Unit атак [20], теорія яких в останні роки дуже розвинулась для поля $\mathbb{Z}_q[X]/(X^n - X - 1)$. Алгебраїчні квантові атаки, що були викладені у серії робіт [21], також нерелевантні для ДСТУ 8961:2019 через використання нетипового поля та модульної структури NTRU.

6. Аналіз асиметричної схеми

Основний результат цього розділу викладено в теоремі 1.

Т е о р е м а 1. Якщо існує алгоритм A, що може перемогти у IND-CCA2 грі для SkeyaPKE за поліноміальний час з ймовірністю ε та робить $q_{BPGM}, q_{MGF}, q_{Dec}$ запитів до

випадкових оракулів BPGM, MGF та оракула дешифрування, то існує алгоритм B, що може знайти прообраз односторонньої функції NTRU з ймовірністю ε':

$$\varepsilon' \geq \varepsilon - \frac{q_{BPGM}}{|R_3^t|} - \frac{q_{Dec}\gamma}{|R_3^t|}. \tag{13}$$

Доказ.

Алгоритм B приймає у якості аргумента шифротекст c^*, відкритий ключ $pk = h$ та повертає m^*, r^* – відкритий текст та випадкове значення, що було використано для шифрування. Сутність алгоритму B полягає у симуляції IND-CCA2 гри для алгоритму A та працює наступним чином:

- Алгоритм B передає до A відкритий ключ pk.
- Алгоритм B отримує два випадкових повідомлення M_0, M_1 від IND-CCA2 іспитувача та передає їх до A.
- Алгоритм B обирає випадковий біт σ та передає c^* до B як шифротекст повідомлення $M^* = M_\sigma$ (тобто випадкові оракули відповідатимуть на запити від A таким чином, щоб повідомлення M_σ було результатом дешифрування c^*. Опис роботи оракулів нижче).
- Алгоритм A робить запити до оракулів.
- Алгоритм A повертає відповідь σ'. Якщо у запитах до оракулів є інформація, з якої можливо відновити m^*, r^*, то повернути m^*, r^*.

Повернути випадкові значення m^*, r^*.

Для симуляції випадкових оракулів використовуються два списки: $BPGM_{LIST}$ та MGF_{LIST}. Випадкові оракули виконують часткову симуляцію відповідних геш-функцій, як показано нижче.

Оракул для $BPGM(x)$ працює наступним чином:

- Якщо $(x, r) \in BPGM_{LIST}$, то повернути r.
- Якщо $x = M^* \| coins^* \| h$, то повернути r^*.
- Інакше обрати випадкове r, додати до $BPGM_{LIST}$ (x, r) та повернути r.

Оракул для $MGF(x)$ працює наступним чином:

- Якщо $(x, mask) \in MGF_{LIST}$, то повернути $mask$.
- Якщо $x = r^* \cdot h$, то повернути $m^* - MGF(r^* \cdot h)$.
- Інакше обрати випадкове $mask$, додати до MGF_{LIST} $(x, mask)$ та повернути $mask$.

Оракул дешифрування $Dec(c)$ працює наступним чином:

- Якщо $c = c^*$, то повернути $\perp$.
- Якщо у $BPGM_{LIST}$ містяться (x, r), для яких $Enc(x, r, pk) = c$, то повернути x.
- Інакше повернути помилку дешифрування.

Позначимо як W_{real}^{win}, та W_{oracle}^{win}, події, що відповідають перемозі A у IND-CCA2 грі з реальними геш-функціями та з оракулами відповідно, а $W_{real}^{bpgm}, W_{real}^{mgf}$ та $W_{oracle}^{bpgm}, W_{oracle}^{mgf}$ – події, що відповідають виклику оракулів для BPGM та MGF на даних з шифротексту у грі з реальними геш-функціями та з оракулами відповідно. Також введемо наступні позначення:

$$\begin{aligned} W_{real}^{query} &= W_{real}^{bpgm} \vee W_{real}^{mgf} \\ W_{oracle}^{query} &= W_{oracle}^{bpgm} \vee W_{oracle}^{mgf} \end{aligned}. \tag{14}$$

Розглянемо ймовірність перемоги у грі з реальними геш-функціями. За формулою повної ймовірності маємо:

$$\Pr[W_{real}^{win}] = \Pr[W_{real}^{win} \mid \neg W_{real}^{query}] \Pr[\neg W_{real}^{query}] + \Pr[W_{real}^{win} \mid W_{real}^{query}] \Pr[W_{real}^{query}]. \quad (15)$$

Якщо алгоритм A не робить запитів до геш-функцій, то він, відповідно, не має жодної інформації про біт σ, отже $\Pr[W_{real}^{win} \mid \neg W_{real}^{query}] = 1/2$. Підставивши у вираз для $\Pr[W_{real}^{win}]$, маємо:

$$\begin{aligned} \Pr[W_{real}^{win}] &= \frac{1}{2} \Pr[\neg W_{real}^{query}] + \Pr[W_{real}^{win} \mid W_{real}^{query}] \Pr[W_{real}^{query}] \\ &= \frac{1}{2}\left(1 - \Pr[\neg W_{real}^{query}]\right) + \Pr[W_{real}^{win} \mid W_{real}^{query}] \Pr[W_{real}^{query}] \\ &= \frac{1}{2} + \Pr[W_{real}^{query}]\left(\Pr[W_{real}^{win} \mid W_{real}^{query}] - \frac{1}{2}\right) \end{aligned} . \quad (16)$$

Оскільки $\Pr[W_{real}^{win} \mid W_{real}^{query}] - \frac{1}{2}$ завжди менше $\frac{1}{2}$, то маємо:

$$\Pr[W_{real}^{win}] \le \frac{1}{2}(1 + \Pr[W_{real}^{query}]). \quad (17)$$

Звідси випливає, що перевага у IND-CCA2 грі є меншою, ніж $\Pr[W_{real}^{query}]$, яку можливо розкласти як

$$\Pr[W_{real}^{query}] = \Pr[W_{real}^{bpgm} \wedge \neg W_{real}^{mgf}] + \Pr[W_{real}^{mgf}]. \quad (18)$$

Якщо подія W_{real}^{mgf} не відбулася, то A відповідно не має інформації про σ^*, отже можливо обмежити $\Pr[W_{real}^{bpgm} \wedge \neg W_{real}^{mgf}] \le \frac{q_{BPGM}}{|R_3^t|}$:

$$\Pr[W_{real}^{win}] \le \Pr[\neg W_{real}^{query}] \le \frac{q_{BPGM}}{|R_3^t|} + \Pr[W_{real}^{mgf}]. \quad (19)$$

Розглянемо як зміниться ймовірність перемоги, якщо замінити справжні геш-функції на випадкові оракули. Для алгоритму A різниця не буде помітна, якщо оракул дешифрування не поверне $\perp$ на валідний запит

Позначимо як W_{ind} подію, що зазначені вище випадки не стануться. У цьому випадку маємо:

$$\begin{aligned} \Pr[W_{oracle}^{mgf}] &\ge \Pr[W_{oracle}^{mgf} \mid W_{ind}] \Pr[W_{ind}] \\ &= \Pr[W_{real}^{mgf} \mid W_{ind}] \Pr[W_{ind}] \\ &\ge \left(\Pr[W_{real}^{win}] - \frac{q_{BPGM}}{|R_3^t|}\right) \Pr[W_{ind}] \end{aligned} . \quad (20)$$

Розглянемо ймовірність того, що оракул дешифрування не поверне $\perp$ на валідний запит. Нехай $c \ne c^*$ – деякий запит до оракула дешифрування, на який повернули $\perp$. Валідний шифротекст може бути відкинуто у грі з оракулами, якщо не було відповідного запиту до BPGM. Оскільки ймовірність звернення до випадкового оракула BPGM обмежена зверху як $\frac{\gamma}{|R_3^t|}$, де γ визначено формулою (5), то відповідно ймовірність того, що не відбудеться повернення $\perp$, складає $1 - \frac{q_{Dec}\gamma}{|R_3^t|}$.

Поєднуючи формули, маємо:

$$\varepsilon' \geq \varepsilon - \frac{q_{BPGM}}{|R_3^t|} - \frac{q_{Dec}\gamma}{|R_3^t|}. \quad (21)$$

7. Аналіз CPA-to-CCA перетворення

Доказ IND-CCA2 безпеки SkelyaPKE у моделі випадкового оракула грунтується на стандартних техніках. Використовується наступна технічна лема [10]:

Л е м м а 1. Позначимо як A, B, E деякі події в ймовірнісному просторі. Якщо $\Pr[A|\neg E] = \Pr[B|\neg E]$, то має місце нерівність $|\Pr[A] - \Pr[B]| \leq \Pr[E]$.

Т е о р е м а 2. Нехай $PPKE = (Gen, Enc, Dec)$ – деяка ймовірнісна схема асиметричного шифрування, а SkelyaKEM – механізм інкапсуляції ключей, що побудований за допомогою застосування перетворення на рис. 2 до PPKE. Якщо існує алгоритм A, що може перемогти у IND-CCA2 грі SkelyaKEM за поліноміальний час з ймовірністю ε та робить $q_{BPGM}, q_H, q_{KDF}, q_{Decaps}$ запитів до випадкових оракулів BPGM, H, KDF та оракула дешифрування, то існує алгоритм B, що може інвертувати PPKE з ймовірністю

$$\varepsilon' \geq \varepsilon - \frac{q_{Decaps}}{|R_3^{2t,N-2t}|} - \frac{\gamma q_{Decaps}}{2^{\lambda}}. \quad (22)$$

Доказ.

Алгоритм B приймає у якості аргумента шифротекст $PPKE - C_1^*$, відкритий ключ pk та повертає x. Сутність алгоритму B полягає у симуляції IND-CCA2 гри для алгоритму A та працює наступним чином:

- Алгоритм B генерує випадкові бітові строки C_2^*, K^*.
- Алгоритм B передає pk до A.
- Алгоритм B чекає доки A запросить завдання, на запит B посилає пару $\left(C^* = (C_1^*, C_2^*), K^*\right)$.
- Алгоритм B чекає доки A поверне біт σ'.
- Алгоритм B перевіряє чи існує $(x, r) \in BPGM_{LIST}$, для якого виконується $Enc(x, r, pk) = C_1^*$. Якщо так, то обчислює та повертає x.
- Алгоритм B перевіряє чи існує $(r, K) \in KDF_{LIST}$ або $(r, hash) \in H_{LIST}$, для якого виконується $x = Pad^{-1}(C_1^* - rh - MGF(rh)); Enc(x, r, pk) = C_1^*$. Якщо так, то повертає x.
- Алгоритм B повертає випадкове x.

Оракул для $BPGM(x)$ працює наступним чином:

- Якщо $(x, r) \in BPGM_{LIST}$, то повернути r.
- Якщо $Decaps(C_1^*, sk) = x$, то повернути $BPGM(x)$.
- Інакше обрати випадкове r, додати до $BPGM_{LIST}$ (x, r) та повернути r.

Оракул для $H(r)$ працює наступним чином:

- Якщо $(r, hash) \in H_{LIST}$, то повернути $hash$.
- Якщо $BPGM(Decaps(C_1^*, sk), pk) = r$, то повернути $H(r)$.
- Інакше обрати випадкове $hash$, додати до H_{LIST} $(r, hash)$ та повернути $hash$.

Оракул для $KDF(r)$ працює наступним чином:

- Якщо $(r, K) \in KDF_{LIST}$, то повернути K.
- Якщо $BPGM(Decaps(C_1^*, sk), pk) = r$, то повернути $KDF(r)$.
- Інакше обрати випадкове K, додати до KDF_{LIST} (r, K) та повернути K.

Оракул декапсуляції $Decaps((C_1, C_2))$ працює наступним чином:

- Якщо $C_1 = C_1^*$, то повернути $\perp$.
- Для кожного значення $(x, r) \in BPGM_{LIST}$ перевірити, чи виконується $Enc(x, r, pk)$ та $H(r) = C_2$. Якщо такої пари не знайдено, то повернути $\perp$.
- Обчислити $K = KDF(r)$ з використанням оракула для KDF.
- Повернути K.

Позначимо як W_{real}^{win} та W_{oracle}^{win} події, що відповідають перемозі A у IND-CCA2 грі з реальними геш-функціями та з оракулами відповідно. Розглянемо як зміниться ймовірність перемоги, якщо замінити справжні геш-функції на випадкові оракули. Для алгоритму A різниця не буде помітна, якщо:

- A робив запит на розшифрування завдання до того, як отримав завдання.
- A робив запит на декапсуляцію валідного шифротексту C_1, C_2, але A не робив відповідних запитів до H або BPGM.

Позначимо як W_{ind} подію, що зазначені вище випадки не стануться, отже, використовуючи Л е м м у 1, маємо:

$$\begin{gathered} \Pr[W_{real}^{win} \mid \neg W_{ind}] = \Pr[W_{oracle}^{win} \mid \neg W_{ind}] \Rightarrow \\ \mid \Pr[W_{real}^{win}] - \Pr[W_{oracle}^{win}] \mid \leq \Pr[W_{ind}] \end{gathered}. \tag{23}$$

Так як шифротекст був обраний випадково, то ймовірність першої події обмежена $\frac{q_{Decaps}}{\mid R_3^{2t,N-2t} \mid}$. Ймовірність того, що $H(BPGM(Decaps(C, sk), pk))$ обмежена $\frac{\gamma q_{Decaps}}{2^{\lambda}}$:

$$\frac{q_{Decaps}}{\mid R_3^{2t,N-2t} \mid} + \frac{\gamma q_{Decaps}}{2^{\lambda}}. \tag{24}$$

Оскільки перевага у грі з оракулом є нижньою оцінкою ймовірності того, що серед запитів до геш-функцій буде інформація, якої достатньо для дешифрування шифротексту C_1^*, то маємо

$$\varepsilon' \geq \varepsilon - \frac{q_{Decaps}}{\mid R_3^{2t,N-2t} \mid} - \frac{\gamma q_{Decaps}}{2^{\lambda}}. \tag{25}$$

8. Пропозиції щодо покращення

У розд. 6, 7 було показано, що схема асиметричного шифрування та механізм інкапсуляції ключів, що використовуються в ДСТУ 8961:2019, є безпечними в моделі випадкового оракула, проте також варто зауважити, що з аналізу також випливає ряд недоліків.

По-перше, стандарт визначає механізм інкапсуляції ключів, який фактично будується на основі IND-CCA2 безпечної схеми асиметричного шифрування. Для CPA-to-CCA перетворення, що використовується у ДСТУ 8961:2019, достатньо було б IND-CPA безпечної схеми. Механізм інкапсуляції ключів надзвичайно збитковий. Також у механізмі присутні деякі деталі, які не впливають на безпеку, проте при реалізації дещо уповільнюють схему. Наприклад, стандарт передбачає, що ключем інкапсуляції є поліном r, який виконує роль випадкового значення. Такий вибір потребує додаткових обчислень з кодування-декодування полінома в бітову строку, при тому, що не збільшує безпеку, а навпаки – ускладнює аналіз. В інших механізмах інкапсуляції ключів роль ключа виконує випадкове повідомлення m, що відповідає семантиці механізму інкапсуляції ключів як криптопримітива і не створює додаткових перешкод при аналізі.

Варто зауважити, що такий незвичний вибір може призвести до більш зручних доказів у моделі квантового оракула. Необхідні додаткові дослідження для цього випадку. Наведені докази у прямому вигляді доволі важко узагальнити на квантовий випадок через те, що квантовий оракул приймає усі запити в суперпозиції, і аргументи, що використовуються, неможк-

ливо використати. Втім, у роботі [12] запропоновано ряд технік, які могли б усунути цю перешкоду.

В реальному світі протоколи інкапсуляції ключів використовуються як складова частина більш складних протоколів. Зокрема, механізм інкапсуляції ключів можливо використовувати як основу для протоколів автентифікованого обміну ключами.

У роботах [22, 23] запропонований узагальнений протокол автентифікованого обміну ключів на основі довільного механізму інкапсуляції ключей, що є безпечним у моделі IND-CCA2. На рис. 6, 7 запропоновано протоколи SkelyaAKE1 з односторонньою автентифікацією та SkelyaAKE2 – з двохсторонньою автентифікацією.

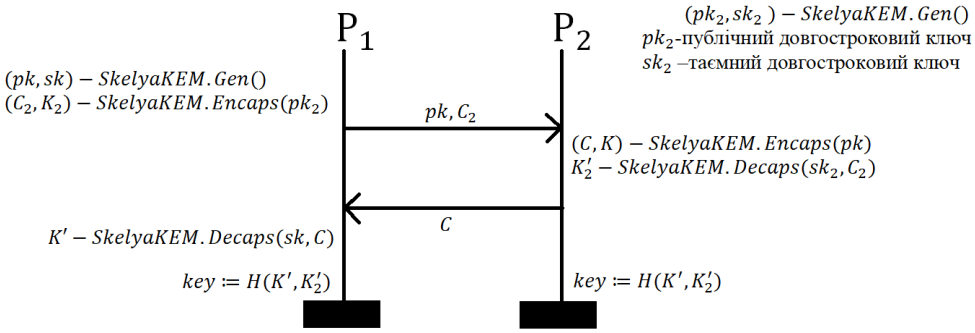


Рис. 6. Протокол SkelyaAKE1 з односторонньою автентифікацією

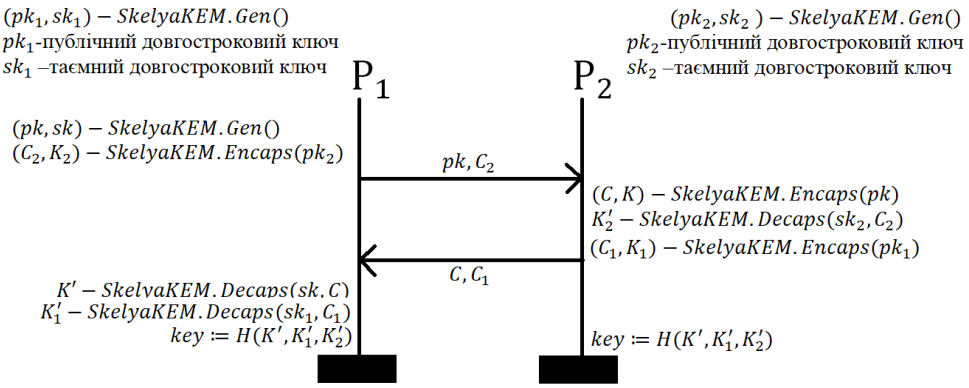


Рис. 7. Протокол SkelyaAKE2 з двохсторонньою автентифікацією

Сторони, що проводять автентифікацію, мають довгострокові ключі. Публічний довгостроковий ключ відомий іншій стороні. Цікавою задачею є створення протоколів на основі ДСТУ 8961:2019 з врахуванням його особливостей, а не з використанням узагальнених доказово безпечних конструкцій.

Висновки

1. В моделі випадкового оракула IND-CCA2 доведено безпеку схеми асиметричного шифрування, що описана в ДСТУ 8961:2019, та безпеку відповідного механізму інкапсуляції ключів. Оскільки стандарт містить лише технічний опис перетворень, у розд. 4 було введено формалізовану математичну модель без зайвих технічних деталей, що не впливають на оцінки безпеки.

2. Оскільки загальносистемні параметри в стандарті було обрано таким чином, щоб схема не мала помилок дешифрування, вдалося значно спростити доказ. У розд. 5 наведено схематичний огляд можливих векторів атак на ДСТУ 8961:2019, проте детальний аналіз є предметом подальших досліджень.

3. Важливо розуміти, що доказ у моделі випадкового оракула не означає, що взагалі атак немає, а лише доводить захист від широкого класу атак, що підпадають під модельні припущення. Наведений доказ може в подальшому бути вдосконалений з використанням більш сильних технік та з меншою кількістю модельних припущень.

4. Однією з важливих задач подальших досліджень є узагальнення отриманих результатів на квантовий випадок. Проте, для цього необхідні принципово інші техніки через те, що квантовий випадок приймає запити в суперпозиції, і побудувати оракулів дешифрування/декапсуляції аналогічно до того, як це приведено в доказах, стає майже неможливим.

5. Аналіз показав, що ДСТУ 8961:2019 має певну збитковість у сенсі безпеки. Конструкція може бути значно спрощена та пришвидшена без втрати безпеки. Безпеку, навпаки, можна значно підвищити.

6. Оцінка в моделі випадкового оракула сама по собі не забезпечує безпеку. Необхідно, щоб загальносистемні параметри відповідали необхідним рівням безпеки. Досліджуючи властивості загальних параметрів та доповнюючи існуючі докази, можна створити більш комплексні моделі безпеки, що враховують більшу кількість загроз і мають менше ризиків.

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. Горбенко І. Д., Горбенко Ю. І. Прикладна криптологія. Теорія. Практика. Застосування : монографія. Харків : Форт, 2012. 880 с.
3. Calculation of general parameters for NTRU Prime Ukraine of 6-7 levels of stability / I. D. Gorbenko and other // Telecommunications and Radio Engineering. Vol. 78. P. 327 – 340
4. Methods of building general parameters and keys for ntru prime Ukraine of 5th-7th levels of stability. product form / I.D. Gorbenko and other // Telecommunications and Radio Engineering. 2018. Vol 78. P. 579 – 594.
5. NIST Post-Quantum Cryptography Standartization Project [Electronic resource] // Online: https://csrc.nist.gov/Projects/post-quantum-cryptography/post-quantum-cryptography-standardization
6. 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
7. FrodoKEM. Learning With Errors Key Encapsulation Algorithm Specifications. And Supporting Documentation // [Electronic resource]. Online: https://frodokem.org/files/FrodoKEM-specification-20210604.pdf
8. Katz J., Lindell Y. Introduction to Modern Cryptography: Principles and Protocols // Chapman and Hall/CRC, 2007. 603 P.
9. Canetti R., Goldreich O., Halevi S. The random oracle methodology, revisited // 30th symposium on theory of computing. STOC, 1998. P. 209 – 218.
10. A. Dent. A Designer's Guide to KEMs // Lecture Notes in Computer Science. Vol. 2898. P. 28 – 44.
11. Howgrave-Graham N., Silverman J., Singer A. and William Whyte. NAEP: Provable security in the presence of decryption failures // Cryptology ePrint Archive, Report 2003/172. [Electronic resource]. Online: https://eprint.iacr.org/2003/172.
12. 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.
13. Goldreich O. Foundations of Cryptography/ Vol. 2. Cambridge University Press, 2000. 392 p.
14. Bellare M., Rogaway P. Random oracles are practical: A paradigm for designing efficient protocols // ACM Conference on Computer and Communications Security. 1993. Vol. 1. P. 62 – 73.
15. Canetti R., Krawczyk H. Analysis of key-exchange protocols and their use for building secure channels // EUROCRYPT. 2001. Vol. 2045. P. 453 – 474.
16. Hoffstein J., Pipher J., Silverman H. NTRU: a ring based public key cryptosystem // Algorithmic Nuber Theory. Third International Symposium. 1998. P. 267 – 288.
17. Hirschhorn P., Hoffstein J., Howgrave-Graham N. Choosing NTRUEncrypt Parameters in Light of Combined Lattice Reduction and MITM Approaches // Lecture Notes in Computer Science. Vol. 5536. P. 58 – 78.
18. Mol, P., Young M. Recovering NTRU Secret Key from Inversion Oracles // PKS. 2008. Vol. 4939. P. 18 – 36.
19. Micheli G., Heninger N., Shani B. Characterizing overstretched NTRU attacks // Journal of Mathematical Cryptology. 2020. Vol 14, Is 1. P. 110 – 119.
20. Bernstein D., Lange T. Non-randomness of S-unit lattices // [Electronic resource]. Online: https://s-unit.attacks.cr.yp.to/spherical.html
21. Eisenträger K., Hallgren S., Kitaev A. and Song F. A quantum algorithm for computing the unit group of an arbitrary degree number field // STOC. 2014. P. 293 – 302.
22. Fujioka A., Suzuki K., Xagawa K. and Yoneyama K. Strongly secure authenticated key exchange from factoring codes and lattices // PKC. 2012. Vol. 7293. P. 467 – 484.
23. Boyd C., Cliff Y., Gonzalez J. and Kenneth G. Efficient one-round key exchange in the standard model // ACISP. 2008. Vol. 5107. P. 69 – 83.

Надійшла до редколегії 04.10.2022

Відомості про автора:

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 004.065 **DOI:10.30837/rt.2022.4.211.03**

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

ОСНОВНІ КАТЕГОРІЇ NEWSQL БАЗ ДАНИХ ТА ЇХ ОСОБЛИВОСТІ

Вступ

У сучасному світі все гостріше постає проблема роботи з величезними обсягами даних та великими навантаженнями. Великі Web-застосунки, соціальні мережі, торгові платформи, портали новин, різні наукові дослідження, бізнес-аналітика, а також безліч інших областей, так чи інакше, стикаються з проблемами керування та аналізу даних великого обсягу. У таких проєктах велика кількість користувачів одночасно читають та записують інформацію, що вимагає від системи керування даними не тільки великої пропускної здатності та низьких затримок, а й масштабованості, надійності та узгодженості даних. Донедавна реляційні системи керування даними (СКБД) залишалися головним інструментом керування даними, яким потрібно щодня обробляти мільйони запитів. За деякими оцінками експертів, ринок реляційних баз даних (БД) приносить дохід понад 50 мільярдів доларів [1]. Таку популярність їм забезпечила декларативна мова запитів SQL, за допомогою якої користувачі можуть оперувати даними. Однак, незважаючи на велику популярність, досвід застосування, універсальність, забезпечення безпеки даних, що зберігаються, традиційні реляційні СКБД не завжди можуть задовольнити вимоги сучасних застосунків. Соціальні мережі, наприклад, вимагають виконання мільйонів операцій зчитування та мільярдів операцій запису в режимі, близькому до реального часу [2]. Для того щоб впоратися зі зростаючим обсягом даних і трафіку, були потрібні більші обчислювальні ресурси. Для вирішення цієї проблеми відомі два підходи – так звані вертикальне (scaling up або scaling vertically) та горизонтальне (scaling horizontally або scaling out) масштабування.

Багато хто пробував найбільш очевидний варіант вертикального масштабування своїх СКБД, переміщуючи базу даних на машину з більш досконалим та потужнішим (продуктивним) обладнанням. Однак це тільки підвищує продуктивність, але, як правило, не має такої ефективної віддачі (більші машини стають дедалі дорожчими), не кажучи вже про те, що існує фізична межа критичних характеристик відповідного обладнання. Крім того, переміщення бази даних з однієї машини на іншу – складний процес, що часто потребує значного часу простою, що є неприйнятним для певних застосунків. Альтернативою було використання великої кількості невеликих машин, об'єднаних у кластер (горизонтальне масштабування). Кластер невеликих машин може використовувати звичайне апаратне забезпечення і в результаті здешевити масштабування. Крім того, кластер надійніший – окремі машини можуть вийти з ладу, але весь кластер продовжить функціонування, незважаючи на ці збої, забезпечуючи високу надійність.

Коли відбулося зрушення у бік кластерів, виникла нова проблема – реляційні бази даних не були призначені до роботи на кластерах. Кластерні реляційні бази даних, такі як Oracle RAC або Microsoft SQL Server, ґрунтуються на концепції загальної дискової підсистеми. Вони використовують кластерну файлову систему, яка виконує запис даних у легко доступну дискову підсистему, але це означає, що дискова підсистема, як і раніше, є єдиним джерелом вразливості кластера. При цьому в роботі [3] зазначається, що все ж таки реляційні бази даних можуть працювати на різних серверах з різними наборами даних, ефективно виконуючи фрагментацію (sharding) бази даних. Хоча це дозволяє розділити робоче навантаження, вся процедура фрагментації, як правило, повинна контролюватись застосунком, який повинен стежити за тим, який сервер бази даних і за якою частиною даних звертається. Втім, сьогодні, наприклад, у СКБД Oracle Database використовується функція Oracle Sharding, яка дозволяє автоматично розподіляти та реплікувати дані в пулі баз даних Oracle, які не використовують спільно апаратне або програмне забезпечення [4].

Тому деякі компанії почали створювати спеціальне програмне забезпечення (ПЗ) проміжного рівня (або проміжне ПЗ) для поділу СКБД з одним вузлом – на кластер менш дорогих машин. Таке проміжне програмне забезпечення надає застосунку єдину логічну базу даних, що зберігається на декількох фізичних вузлах (серверах). Коли застосунок видає запити до такої бази даних, проміжне ПЗ перенаправляє та/або перезаписує їх, щоб розподілити їх виконання на одному або кількох вузлах у кластері. Вузли виконують ці запити та надсилають результати назад проміжному програмному забезпеченню, яке потім поєднує їх в одну відповідь застосунку. Як відомі приклади такого проміжного програмного забезпечення є кластер eBay на базі Oracle, кластер Google на базі MySQL, Facebook для їх власного кластера MySQL [5, 6]. Однак слід враховувати, що таке ПЗ проміжного рівня добре підходить для простих операцій, таких як читання або оновлення одного запису, але не для запитів, які оновлюють більше одного запису в транзакції або в таблицях з'єднання (їх виконати набагато складніше). Згодом деякі компанії відмовилися від проміжного програмного забезпечення та почали розробляти власні розподілені СКБД. Така їх поведінка була обумовлена кількома причинами. Насамперед, традиційні СКБД того часу були орієнтовані на узгодженість/несуперечність (consistency) та коректність (correctness) за рахунок доступності та продуктивності. Але цей компроміс вважався неприйнятним для застосунків, яким необхідно постійно перебувати в мережі та підтримувати велику кількість одночасних операцій. По-друге, вважалося, що використання повнофункціональної СКБД, такої, наприклад, як MySQL, як звичайне сховище даних вимагає надто великих накладних витрат. Крім того, так само вважалося, що реляційна модель є не найкращим способом представлення даних, а використання SQL – це надлишок для простих пошукових запитів. Тому, як відомо, до кінця 2000-х років з'явився різноманітний набір масштабованих і доступніших розподілених NoSQL СКБД. Основними причинами використання можливостей технології NoSQL можна вважати [3]: необхідність забезпечити доступ до даних, обсяг яких та вимоги до продуктивності змушують використовувати кластери; необхідність підвищення продуктивності розробки застосунків за рахунок використання зручнішого способу взаємодії з даними. Кожне рішення в рамках технології NoSQL використовує свою модель. Ці моделі можна розділити на кілька категорій [3, 7]: сімейство стовпців/сховище широких стовпців (column families / wide column store); сховища документів/документні бази даних (document store); ключ-значення (key value/tuple store); графові (graph); багатомодельні (multimodel); об'єктні/об'єктно-орієнтовані (object / object oriented); XML; багатозначні (multivalue); подієві (event, event stores, event sourcing); тимчасові ряди (time series) та деякі інші. Перевага використання системи NoSQL (принаймні така достатньо поширена думка існувала) полягала в тому, що розробники могли зосередитися на тих аспектах свого застосування, які були більш корисні для їхнього бізнесу, замість того, щоб турбуватися про те, як масштабувати СКБД. Однак, незважаючи на те, що всі вони спрямовані на вирішення проблеми масштабованості та мають певні переваги, кожному з них притаманні різні недоліки і кожне рішення може ефективно справлятися лише з певним класом завдань. Багато застосунків не можуть використовувати системи NoSQL, оскільки вони не можуть відмовитися від жорстких вимог, що висуваються до транзакцій та узгодженості. При цьому деякі організації, перш за все Google, виявили, що СКБД NoSQL змушують їх розробників витрачати занадто багато часу на написання коду для обробки суперечливих даних. Наприклад, решардинг (resharding) однією з систем зайняв понад два роки інтенсивних зусиль і включав координацію та тестування десятків команд для мінімізації ризиків, тоді як використання транзакцій робить їх більш продуктивними, оскільки вони являють собою більш зрозумілу для обговорення абстракцію [8]. Таким чином, єдиним доступним для цих організацій варіантом було або придбання потужнішої одновузлової машини і вертикальне масштабування СКБД, або розробка власного проміжного програмного забезпечення, що підтримує транзакції сегментованої БД. Обидва підходи дуже дорогі і тому для багатьох не підходили. Саме за цих умов і з'явилися системи NewSQL.

Сьогодні ландшафт в області NewSQL продовжує змінюватися, а стандартів поки що немає [9]. Цій темі присвячено чимало наукових статей, але на жаль, нерідко вони швидко старіють. Найчастіше актуальну інформацію можна знайти на сайтах відповідних систем та в блогах. Тому в цій статті намагатимемося усунути певною мірою цю прогалину, сконцентрувавши увагу на основних принципах, архітектурі та особливостях СКБД даного класу, а не на деталях реалізації, які з часом змінюються. При цьому для кращого розуміння викладеного матеріалу в роботі наводяться деякі теоретичні відомості, пов'язані з аспектами тематики, що розглядається, а також уточнюються деякі синонімічні терміни, в назві та перекладах яких у відомих релевантних джерелах присутня певна неоднозначність. Наприкінці наводяться деякі важливі узагальнені характеристики NewSQL СКБД для масштабованих рішень, що відрізняють їх від традиційних реляційних (RDBMS – Relational Database Management System) та NoSQL систем керування базами даних.

1. Основні категорії NewSQL баз даних

NewSQL – це клас сучасних систем керування реляційними базами даних, які прагнуть забезпечити високу продуктивність та масштабованість систем NoSQL, зберігаючи при цьому гарантії ACID (Atomicity, Consistency, Isolation, Durability) традиційної системи баз даних. Термін NewSQL був запропонований у 2011 р. аналітиком 451 Group (дослідницька компанія у сфері високих технологій, що базується в Нью-Йорку) Метью Аслетом [10]: «NewSQL – це скорочена назва різних нових постачальників масштабованих / високопродуктивних баз даних SQL. Раніше ми називали ці продукти «ScalableSQL», щоб відрізнити їх від існуючих продуктів реляційних баз даних. Оскільки це передбачає горизонтальну масштабованість, яка не обов'язково є функцією всіх продуктів, ми прийняли термін NewSQL у новому звіті».

Потреба в таких системах виникла в першу чергу у компаній (для їх корпоративних систем), які обробляють важливі дані (наприклад, фінансові системи, системи обробки замовлень і т. д.), яким потрібні рішення, що масштабуються, в той час як рішення NoSQL не могли забезпечити транзакційні механізми та не відповідали вимогам узгодженості даних. Серед застосувань систем NewSQL – розпізнавання шахрайства в реальному часі, електронна реклама, призначення розцінок та перехресні продажі в реальному часі, підтримка прийняття геопросторових рішень, proximity маркетинг, моніторинг ризиків, ціноутворення в реальному часі, Інтернет речей (IoT, internet of things) тощо [2, 11].

Системи баз даних NewSQL мають різні особливості та архітектури. Однак можна виділити їх такі спільні риси [11]: підтримка реляційної моделі даних та стандартного SQL; підтримка ACID транзакцій; масштабованість за рахунок поділу даних у кластерах без спільного використання ресурсів (shared nothing); та висока доступність за рахунок реплікації даних.

Розглянемо сучасні СКБД NewSQL з урахуванням їхнього групування на основі суттєвих аспектів їх реалізації. Для чого виділимо кілька категорій, які, на думку деяких дослідників, сьогодні найкраще являють собою системи NewSQL [6, 12]:

1. Системи, що створені з нуля з використанням нової архітектури.

2. Програмне забезпечення проміжного рівня (middleware), що дозволяє забезпечити можливість прозорого функціонування із сегментованою/фрагментованою на кількох вузлах (після так званого прозорого поділу – transparent sharding) бази даних.

3. Пропозиції «база даних як послуга» від постачальників хмарних обчислень, які також ґрунтуються на нових архітектурах.

1.1. Системи з використанням нової архітектури

У цю категорію включені системи NewSQL, розроблені на основі нової кодової бази, не спираючись на архітектуру застарілих систем, тобто створені з нуля. Усі СКБД цієї категорії засновані на розподілених (distributed) архітектурах, які працюють із ресурсами без поділу (спільного використання) та містять компоненти для підтримки управління паралельною роботою на кількох вузлах, відмовостійкості (за рахунок реплікації), керування потоками та

розподіленою обробкою запитів. Архітектура без поділу ресурсів (shared-nothing architecture) – це архітектура розподілених обчислень, в якій кожен запит на оновлення задовольняється одним вузлом у комп'ютерному кластері, що складається з кількох вузлів, які не використовують спільні ресурси (процесор, пам'ять, пристрій зберігання). Вперше цей термін був запроваджений Майклом Стоунбрейкером у роботі [13] при перерахуванні найпоширеніших архітектур для багатопроцесорних систем з високою швидкістю транзакцій. Перевага використання таких СКБД, створених для розподіленої обробки, полягає в тому, що всі частини системи можна оптимізувати для багатовузлових середовищ (multi-node environments). З цією метою, зокрема, використовуються оптимізатор запитів та протокол зв'язку між вузлами. Наприклад, більшість СКБД NewSQL можуть відправляти внутрішньозапитні дані безпосередньо між вузлами, а не відправляти в центральне розташування, як у деяких системах ПЗ проміжного рівня.

Практично кожна з СКБД у цій категорії керує своїм власним первинним сховищем, що знаходиться або в пам'яті, або на диску. Це означає, що СКБД відповідає за розподіл бази даних за своїми ресурсами за допомогою спеціального механізму замість того, щоб покладатися на готову розподілену файлову систему (наприклад, HDFS – Hadoop Distributed File System) або структуру зберігання (наприклад, Apache Ignite). Це важливо, оскільки цей механізм дозволяє СКБД відправляти запит до даних, а не переносити дані в запит, що призводить до значно меншого мережного трафіку, оскільки передача запитів зазвичай вимагає менше мережного трафіку, ніж передача даних (а це не тільки кортежі, але також індекси та матеріалізовані уявлення) для обчислення [6, 14].

Управління власним сховищем також дозволяє СКБД використовувати складніші схеми реплікації, ніж це можливо з блоковою схемою реплікації, яка використовується в HDFS. Загалом це дозволяє цим СКБД досягати більш високої продуктивності, ніж у інших систем, які накладаються поверх інших існуючих технологій (наприклад, так звані системи «SQL на Hadoop», такі як Trafodion та Splice Machine, які забезпечують транзакції поверх Hbase). Але є й недоліки використання СКБД з урахуванням нової архітектури. Перш за все, багато організацій побоюються впроваджувати надто нові та неперевірені технології. При цьому слід також пам'ятати, що кількість людей, які мають досвід роботи із системою, набагато менша, ніж у постачальників популярніших СКБД. Також слід мати на увазі, що організація може втратити доступ до існуючих інструментів адміністрування та звітності. Деякі СКБД, такі як MariaDB Xpand (раніше Clustrix) і SingleStore (раніше MemSQL), щоб уникнути цієї проблеми підтримують сумісність із дротовим протоколом (wire protocol) MySQL.

Приклади систем з використанням нової архітектури: *MariaDB Xpand* [15] (раніше *Clustrix*; компанія Clustrix була придбана MariaDB у 2018 р. [16] і Clustrix був перейменований на MariaDB Xpand [17 – 19]; ClustrixDB більше недоступний як окремий об'єкт, тепер він входить до складу MariaDB Enterprise Server, розширюючи його можливості за рахунок розподіленої обробки даних та транзакцій, перетворюючи його на розподілену базу даних SQL, здатну масштабуватися до мільйонів транзакцій за секунду з архітектурою без загального доступу), *CockroachDB* [20], *Google Cloud Spanner* [8, 21, 22], *HyPer* [23], *SingleStore* (раніше *MemSQL* – з жовтня 2020 р. SingleStore) [24], *NuoDB* (у грудні 2020 р. придбана компанією Dassault Systemes) [25], *SAP HANA* [26, 27], *H-Store* (фінальна версія H-Store була анонсована в 2016 р., з тих пір ніяких нових розробок не було) [28]. Якщо ви хочете використовувати більш сучасну СКБД, засновану на архітектурі H-Store, вам слід використовувати *Volt Active Data* (раніше *VoltDB*; перейменована в 2022 р. [29, 30], *LeanXcale* [31].

1.2. Проміжне ПЗ, що забезпечує прозоре функціонування сегментованої на кількох вузлах БД

Термін проміжне програмне забезпечення (middleware – менеджер ресурсів, що пропонує свої застосунки для ефективного обміну та розгортання цих ресурсів у мережі [32]), було введено на початку 1980-х років. Сьогодні завдяки проміжному програмному забезпеченню

програміст має можливість реалізувати децентралізоване рішення замість того, щоб взаємодіяти і аналізувати різні компоненти [33]. Пропоновані рішення, що функціонують як прозорий шар для одновузлових систем, дозволяють застосункам та користувачам представляти базу даних як єдину централізовану одиницю, незважаючи на те, що остання розбита на кілька частин/сегментів і розподілена між кількома вузлами. При цьому фрагментація /сегментування відрізняється від технологій об'єднання/федералізації (federation) баз даних 1990-х років, оскільки кожен вузол [6]: працює з однією і тією ж СКБД; має лише частину загальної бази даних; не призначений для доступу та оновлення самостійно окремими застосунками.

Використання проміжного ПЗ дозволяє розподіляти запити, координувати транзакції, а також керувати розміщенням, реплікацією та секціонуванням даних між вузлами. Найчастіше на кожному вузлі СКБД встановлюється так званий шар прокладки (shim), який взаємодіє з проміжним програмним забезпеченням. Цей компонент відповідає за виконання запитів від імені проміжного програмного забезпечення у своєму локальному екземплярі СКБД та повернення результатів. В цілому, все це забезпечує застосунку можливість роботи з базою даних, представляючи її як єдину логічну БД без необхідності модифікації базової СКБД. Основною перевагою використання ПЗ проміжного рівня, що забезпечує прозоре функціонування сегментованої на кількох вузлах БД, є те, що воно найчастіше є простою заміною застосунку, який вже використовує існуючу СУБД з одним вузлом. Розробникам не потрібно вносити будь-які зміни до свого застосунку, щоб використовувати нову сегментовану (sharded) базу даних. Для сумісності з конкретною СКБД проміжне програмне забезпечення має підтримувати відповідний протокол обміну. Хоча проміжне ПЗ дозволяє компаніям легко масштабувати свою БД на кілька вузлів, у ранніх системах доводилося використовувати традиційну СКБД (наприклад, MySQL, Postgres, Oracle) на кожному вузлі. Але ці СКБД, як правило, засновані на диск-орієнтованій (disk-oriented) архітектурі, і тому вони, наприклад, не можуть використовувати диспетчер/менеджер зберігання (storage manager) або схему управління паралелізмом (concurrency), оптимізовану для зберігання даних, орієнтовану на пам'ять (memory-oriented), як у деяких системах NewSQL, побудованих на нових архітектурах. Дослідження показали, що успадковані компоненти диск-орієнтованих архітектур є значною перешкодою, яка не дозволяє традиційним СКБД масштабуватись, щоб використовувати переваги більшої кількості ядер центрального процесора та великих обсягів пам'яті [6, 34]. Підхід, пов'язаний з використанням проміжного програмного забезпечення, також може призвести до надмірного планування та оптимізації запитів до фрагментів даних, розміщених на окремих вузлах, для складних запитів, а саме один раз на проміжному ПЗ і один раз на окремих вузлах СКБД. Хоча, з іншого боку, це дозволяє кожному вузлу застосовувати свої власні локальні оптимізації для кожного запиту.

Приклади систем: *MariaDB MaxScale* [35], *ScaleArc* [36].

1.3. База даних як послуга

З розвитком індустрії хмарних обчислень та розвитком послуг, що надаються хмарою, база даних також може надаватися як відповідна послуга, яка називається «база даних як послуга» (database-as-a-service, DBaaS). Завдяки такій послузі організаціям не потрібно підтримувати СКБД ні на власному обладнанні, ні на віртуальній машині (virtual machine, VM), розміщеній у хмарі. Постачальник DBaaS відповідає за підтримку фізичної конфігурації бази даних, включаючи налаштування системи, реплікацію та резервне копіювання. Клієнту надається URL-адреса для підключення до СКБД, а також засоби для керування системою. Споживачі служби хмарних обчислень (клієнти DBaaS) платять за відповідну послугу залежно від обсягу спожитих ресурсів згідно з встановленим тарифом. При цьому, оскільки запити до баз даних дуже різняться по тому, як вони використовують обчислювальні ресурси, постачальники DBaaS зазвичай не вимірюють виклики запитів так само, як вони вимірюють операції в блокових службах зберігання (наприклад, Amazon S3, Google Cloud Storage). Замість цього

клієнти підписуються на цінову категорію, в якій вказано максимальний поріг використання ресурсів (зокрема, розмір сховища, обчислювальну потужність, виділену пам'ять), який гарантує їм постачальник [6].

Як і в більшості категорій служб хмарних обчислень, найбільші компанії є основними гравцями в області DBaaS (головним чином завдяки можливості забезпечити економію за рахунок масштабу). Але при цьому слід зазначити, що майже всі DBaaS просто надають керований екземпляр традиційної СКБД (наприклад, MySQL, Microsoft SQL Server) з одним вузлом. До таких постачальників хмарних обчислень можна віднести Google Cloud SQL (база даних як послуга на базі MySQL, PostgreSQL і Microsoft SQL Server; це повністю керована служба бази даних, яка допомагає настроювати, підтримувати, керувати та адмініструвати реляційні бази даних на Google Cloud Platform), Microsoft Azure SQL Database (це хмарний сервіс (як частина Microsoft Azure), що надає можливість зберігання та обробки реляційних даних, а також генерації звітності) та деякі інші. Однак такі системи, на думку деяких авторів [6], недоцільно відносити до систем NewSQL, оскільки вони використовують ті ж базові дискові СКБД, що базуються на архітектурах 1970-х років, навіть незважаючи на те, що деякі постачальники, такі як Microsoft, модернізували свої СКБД, щоб забезпечити найкращу підтримку мультиорендних (multitenant) розгортань [37]. Натомість, як NewSQL, вони пропонують розглядати тільки ті продукти DBaaS, які базуються на новій архітектурі. Яскравим представником таких продуктів є Aurora від Amazon для MySQL RDS (Amazon Aurora – це система управління реляційними базами даних, розроблена для хмари та сумісна з MySQL та PostgreSQL; Aurora доступна як частина служби реляційних баз даних (Relational Database Service, RDS) Amazon).

Використовуючи продукти цієї категорії, слід не забувати про можливі ризики їх застосування у компанії. Наприклад, клієнти таких компаній, як GenieDB, Xeround, які надавали «базу даних як послугу», були змушені шукати нового постачальника послуги з відповідним перенесенням даних через те, що постачальник DBaaS припинив свою діяльність.

Прикладом бази даних як послуга, яка доступна в цій категорії NewSQL станом на 2023 р., є *Amazon Aurora* [38].

2. Особливості СКБД NewSQL

Коли організація вирішує використовувати СКБД NewSQL, для ухвалення правильного рішення необхідно враховувати багато характеристик. Розглянемо основні характерні особливості NewSQL СУБД, щоб зрозуміти, що є новим і заслуговує на увагу в цих системах і як наявні можливості використовувати в подальшому при реалізації конкретних систем обробки даних.

2.1. Механізм реплікації

Основною властивістю технології NewSQL, яка викликає інтерес, є можливість ефективного функціонування баз даних на великому кластері. Залежно від вибраної моделі розподілу можна створити сховище даних, що забезпечує більшу доступність і надає можливість обробляти більший обсяг даних за більш інтенсивного трафіку операцій читання або запису, уникаючи навантаження та гальмування мережі. Існують два способи розподілу даних [3, 9, 39]: реплікація (replication) та сегментування/фрагментація (sharding або fragmentation).

Реплікація має на увазі копіювання одних і тих же даних на декількох вузлах, тобто при реплікації дані копіюються між кількома серверами, так, що кожен біт даних можна знайти в різних місцях. Фрагментація/сегментування означає розміщення різних даних на різних вузлах (поділ БД на секції/розділи, які називаються фрагментами), тобто різні дані розподіляються по кількох серверах, щоб кожен сервер діяв як окреме джерело підмножини даних. При цьому слід враховувати існування двох фундаментальних проблем проєктування [9]: як здійснювати фрагментацію і як оптимально розподілити/розмістити ці фрагменти. Який із цих способів або їх комбінація буде більш ефективною в конкретному випадку, залежить від специфікації проєкту БД. Наприклад, чи слід розділити конкретну таблицю на кілька розділів

або реплікувати її на кожному вузлі. Визначення оптимальної конфігурації для довільного застосунку – нетривіальне завдання, особливо для складного корпоративного застосунку з багатьма залежностями.

Перш ніж продовжити подальший виклад матеріалу, доцільно звернути увагу на той факт, що сьогодні в існуючих численних релевантних джерелах має місце певна неоднозначність у назвах деяких термінів та їх перекладів, пов'язаних із тематикою, що розглядається. Наприклад, те, що в одних джерелах називається секцією (partition), в MongoDB, Elasticsearch і SolrCloud називається «шард» (shard), в HBase – «регіон» (region), у BigTable – «сегмент» (tablet), в Cassandra і Riak – «віртуальний вузол» (vnode), в Couchbase – «віртуальна ділянка» (vBucket). Термін секціонування (partitioning) – являє собою спосіб умисного розбиття великого набору даних на менші, пов'язується з терміном шардинг (sharding) [40] і т. д. Щодо термінології, пов'язаної з реплікацією, то тут теж усе достатньо переплетено. У різних джерелах можна зустріти такі синонімічні назви, як реплікація типу «master-slave» («головний-підлеглий»/«ведучий-ведений»), також відомою під назвою «leader-based replication» – «реплікація з провідним вузлом» або реплікація типу active-passive – «активний-пасивний» [40]). При розподілі за схемою «master-slave» відбувається реплікація даних по багатьох вузлах. Один вузол призначається головним (master або primary). Цей головний вузол є надійним джерелом даних і зазвичай відповідає за виконання всіх модифікацій цих даних. Інші вузли є підлеглими/веденими (*slaves*) або вторинними (secondaries). Процес реплікації синхронізує підлеглі вузли із головними [3]. При цьому типі реплікації запит спочатку обробляється на одному вузлі, а потім СКБД передає результуючий стан іншим реплікам. Існуючий тип реплікації з кількома головними вузлами (multi-leader, multi-master) називають також реплікацією типу «головний-головний» (master-master) або реплікацією типу «активний-активний» (active-active) [40]. За такої схеми кожен з головних вузлів одночасно є підлеглим для інших головних. При цьому типі реплікації кожен вузол репліки одночасно обробляє один і той же запит. Наприклад, коли транзакція виконує запит, СКБД виконує цей запит паралельно на усіх репліках. Однорангову реплікацію (peer-to-peer) називають симетричною (symmetric) чи оновлення будь-якої копії (update-everywhere) [39]. При одноранговій реплікації всі репліки мають однакову вагу, усі можуть виконувати операції запису, і втрата будь-якої з них не призводить до втрати доступу до сховища даних [3]. Тому в цій роботі, як правило, вказуватиметься декілька термінів-синонімів, пов'язаних із конкретним контекстом.

Реплікація баз даних – давній предмет вивчення, її принципи не сильно змінилися з 1970-х років, коли їх тільки почали вивчати, оскільки фундаментальні обмеження мереж залишилися тими самими [40].

Найкращий спосіб, за допомогою якого організація може забезпечити високу доступність та надійність даних, наприклад, для свого OLTP-застосунку, – це реплікувати свою базу даних. Усі сучасні СКБД, включаючи системи NewSQL, підтримують той чи інший механізм реплікації. При цьому DBaaS мають явну перевагу в цій галузі, оскільки вони приховують від своїх клієнтів усі найдрібніші деталі налаштування реплікації. Вони спрощують розгортання реплікованої СКБД, при цьому адміністратору не потрібно турбуватися про передачу журналів та перевірку синхронізації вузлів.

Коли справа доходить до реплікації бази даних, є два конструктивні рішення, що залежать від того, як СКБД забезпечує узгодженість даних між вузлами, а саме, яку модель узгодженості вона підтримує [6]: модель сильної узгодженості (strong consistency) або модель кінцевої узгодженості (eventual consistency) – іноді званою слабко узгодженою (weakly consistent) моделлю. У СКБД, що підтримує модель сильної узгодженості, записи транзакції мають бути підтверджені та встановлені на всіх репліках, перш ніж ця транзакція вважатиметься зафіксованою (тобто довговічною – durable). Перевага цього підходу полягає в тому, що репліки можуть обслуговувати запити лише для читання та при цьому залишатись узгодженими. Тобто, якщо застосунок отримує підтвердження того, що транзакцію зафіксовано, то будь-які зміни, зроблені цією транзакцією, видно будь-якій подальшій транзакції в майбу-

тньому, незалежно від того, до якого вузла СКБД вони звертаються. Це також означає, що при збої репліки немає втрачених оновлень, оскільки всі інші вузли синхронізуються. Але для підтримки цієї синхронізації потрібно, щоб СКБД використовувала протокол атомарної фіксації (наприклад, двофазну фіксацію – two-phase commit, 2PC [41]), щоб гарантувати, що всі репліки узгоджуються з результатом транзакції, що має додаткові витрати і може призвести до зупинки у разі збою вузла. Також необхідно враховувати, що два різні оновлення на головному вузлі можуть бути виконані безпосередньо одне за одним, залишаючи так зване вікно неузгодженості (inconsistency window) на кілька мілісекунд. Однак затримка в роботі мережі означає, що на підлеглих вузлах вікно неузгодженості залишиться відкритим набагато довше [3]. Ось чому, наприклад, системи NoSQL вибирають слабо узгоджену модель, в якій не всі репліки повинні підтверджувати зміну, перш ніж СУБД повідомить про успішний запис.

Відомі системи NewSQL підтримують сильно узгоджену реплікацію. Наприклад, Spanner і CockroachDB забезпечують схему реплікації, оптимізовану для сильно узгоджених реплік глобальної мережі [8, 42, 43]. У Spanner [8] це досягається за рахунок синхронізації, що забезпечується за допомогою комбінації GPS [44] та атомного годинника. Щоб мінімізувати похибку годинника Google встановлює GPS-приймач або атомний годинник у кожному центрі обробки та зберігання даних (data center), завдяки чому годинник синхронізується з точністю в межах 7 мілісекунд [40]. У CockroachDB це досягається завдяки синхронізації, що забезпечується за рахунок особливого використання часових міток/позначок часу поточного часу (wall time, real-world time, wall-clock time) вузла, максимального зміщення часу для кластера та періодичного порівняння зміщення часу вузлів кластера між собою [45].

У принципі, у тому, як ці системи забезпечують таку узгодженість, немає нічого нового. Основи реплікації кінцевого автомата для СКБД вивчалися ще 1970-х роках [46]. Комерційна система управління реляційними базами даних, що забезпечує відмовостійкість та масштабованість NonStop SQL, була однією з перших розподілених СКБД, створених у 1980-х роках, що використовують реплікацію сильної узгодженості для забезпечення відмовостійкості [47]. Сьогодні традиційні реляційні бази даних намагаються забезпечити сильну узгодженість, уникаючи всіх можливих неузгодженостей [3].

Більшість СКБД NewSQL реалізують реплікацію типу master-slave (active-passive), оскільки використовують так звану недетерміновану схему управління паралелізмом. Це означає, що вони не можуть надсилати запити до реплік у міру їх надходження на головний вузол, тому що вони можуть виконуватися в іншому порядку на репліках, і стан баз даних буде різнитися на кожній репліці. Це пов'язано з тим, що їхній порядок виконання залежить від кількох факторів, у тому числі від мережевих затримок, «зависань» кешу та неузгодженості годинників. З іншого боку, СКБД, які підтримують детерміновану схему управління паралелізмом, наприклад, Volt Active Data (H-Store/VoltDB), не виконують цих додаткових кроків координації. Це пов'язано з тим, що СКБД гарантує, що операції транзакцій виконуються в тому самому порядку на кожній репліці, і, таким чином, стан бази даних гарантовано буде однаковим [48].

Одним із важливих аспектів систем NewSQL є також можливість реплікації по глобальній мережі (Wide Area Network, WAN). Будь-яку СКБД NewSQL можна налаштувати для забезпечення синхронного оновлення даних по глобальній мережі, але це спричинить значне уповільнення нормальної роботи. Таким чином, замість цього СКБД NewSQL підтримують асинхронні методи реплікації (оновлення однієї репліки поширюється на інші через деякий час, а не в тій же транзакції).

Реплікація та фрагментація є ортогональними методами. Використовувати можна будь-який із них або обидва одночасно. Якщо ви використовуєте реплікацію master-slave (active-passive) та фрагментацію, це означає, що у вас є кілька головних вузлів, але кожна одиниця даних має лише один головний вузол. Залежно від конфігурації, ви можете призначити вузол головним для одних даних і підлеглим для інших або поєднати обов'язки головного та підле-

глого на одному вузлі. Наприклад, використання однорангової реплікації та фрагментації – поширена стратегія у базах даних NoSQL типу «широко-колонкове сховище» (wide column store / column families) [3].

2.2. Секціонування (розбиття або шардинг)

Подібно до того, як зростаючий бізнес повинен розширюватися, щоб задовольнити потреби своїх клієнтів, масштабуватися повинні й застосунки, що сприяють на рівні інформаційних технологій розвитку цього бізнесу. Забезпечення масштабованості бази даних є одним із пріоритетних завдань для розробників сучасних застосунків. Про масштабованість починають замислюватися, коли у сервера виникають труднощі зі збільшенням навантаження [49]. Рано чи пізно збільшене навантаження на сервер призведе до появи вузьких місць і, як наслідок, до зниження продуктивності. Вертикальне масштабування, як зазначалося раніше, є надто важким та витратним. Найбільш привабливим рішенням є горизонтальне масштабування – розміщення бази даних на кластері серверів.

Майже всі розподілені системи керування базами даних NewSQL масштабуються шляхом розбиття бази даних на непересічні підмножини, відомі як розділи (partitions) або шарди (shards), причому кожен розділ розташовується на різних серверах. Існує два основних типи фрагментації/секціонування [9, 39, 50, 51]:

– вертикальне розбиття (vertical partitioning) – передбачає розподіл стовпців таблиці на дві чи більше таблиць, пов'язаних первинним ключем вихідної таблиці; отримані стовпці поділяються між безліччю вузлів (рис. 1);

горизонтальне розбиття (horizontal partitioning) – передбачає поділ рядків таблиці між двома чи більше таблицями з однаковою структурою (рис. 2).

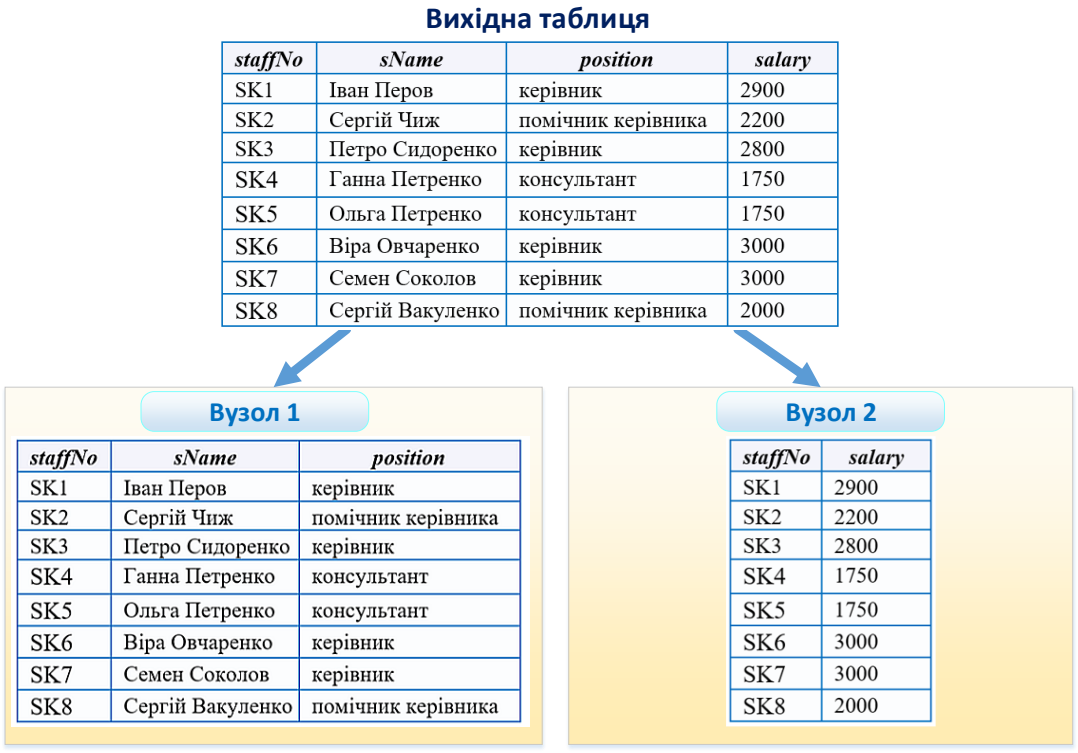


Рис. 1. Приклад вертикальної фрагментації

Основним способом, що знайшов більш широке застосування, масштабування баз даних для роботи на кількох вузлах є горизонтальна фрагментація [52]. Таблиці бази даних горизонтально поділяються на кілька фрагментів, межі яких ґрунтуються на значеннях одного (або кількох) стовпців таблиці (тобто атрибутів поділу).

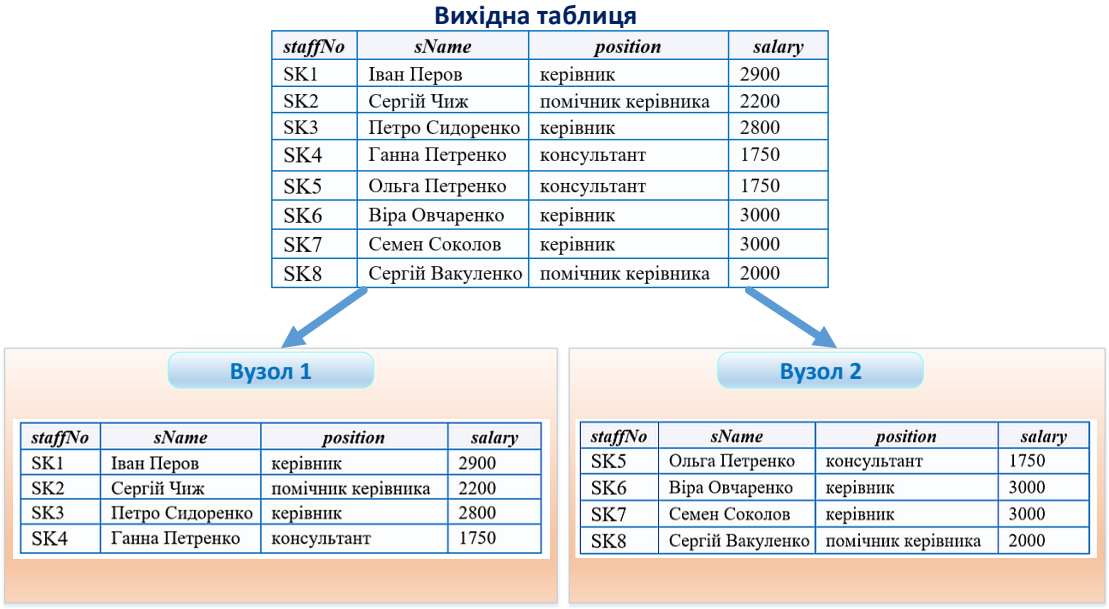


Рис. 2. Приклад горизонтальної фрагментації

СКБД надає кожному кортежу фрагмент (shard) на основі значень цих атрибутів, використовуючи один із методів. Пов'язані фрагменти кількох таблиць об'єднуються разом, щоб сформувати розділ, керований одним вузлом. Цей вузол відповідає за виконання будь-якого запиту, який потребує доступу до даних, що зберігаються в його розділі. Тільки системи DBaaS (зокрема Amazon Aurora) не підтримують цей тип розбиття [6].

Найбільш широко відомими та використовуваними підходами автоматичного горизонтального секціонування (розбиття) є [9, 52]:

– кругове розбиття (round-robin partitioning; стратегія, що забезпечує рівномірний розподіл даних; при n розділах/секціях, i-й кортеж у порядку вставки призначається розділу/секції (i mod n)),

– діапазонне розбиття (range partitioning; розділ/секція вибирається, виходячи з того, чи знаходиться ключ розбиття в межах певного діапазону; тобто розподіл кортежів здійснюється відповідно до набору предикатів),

– геш-розбиття (hash partitioning; застосовує геш-функцію до деякого атрибуту, в результаті застосування якої визначається номер розділу/секції).

Наприклад, геш-розбиття використовується в таких СКБД NewSQL, як Volt Active Data (VoltDB) та MariaDB Xpand (Clustrix). SAP HANA підтримує геш, діапазонне та кругове розбиття. LeanXcale підтримує діапазонне, геш розбиття та складове (composite) секціонування. Деякі СКБД NewSQL, такі як SingleStore (MemSQL), NuoDB і CockroachDB, мають власні моделі поділу, крім підтримки перерахованих вище методів секціонування.

Так, у СКБД NuoDB визначається один або кілька вузлів як менеджери/диспетчери зберігання (SM, storage managers), кожен з яких зберігає розділ бази даних. SM розбиває базу даних на блоки (звані мовою NuoDB «атомами» – atoms). Всі інші вузли в кластері позначаються як механізми транзакцій (TE, transaction engines – механізм транзакцій означає процес у базі даних, який надає клієнту застосунку доступ до бази даних [53]), які діють як кеш-пам'ять атомів. Для обробки запиту вузол TE витягує всі атоми, необхідні для цього запиту (або з відповідних SM, або з інших TE). TE встановлюють блокування запису на кортежі, а потім транслюють будь-які зміни до атомів іншим TE та SM. Щоб уникнути переміщення

атомів між вузлами туди та назад, NuoDB надає схеми балансування навантаження. NuoDB використовує ту ж схему розбиття, що й інші розподілені СКБД, але без необхідності попереднього секціонування бази даних або визначення зв'язків між таблицями.

SingleStore (MemSQL) також, як і NuoDB, використовує гетерогенну архітектуру (у якій не всі вузли однакові), що складається з вузлів-агрегаторів (aggregator nodes), призначених тільки для виконання, та вузлів-листів (leaf nodes), у яких зберігаються фактичні дані. Різниця між цими двома системами полягає в тому, як вони скорочують об'єм даних, що витягуються з вузлів зберігання (storage nodes) на вузли виконання (execution nodes). У NuoDB механізми транзакцій (TE) кешують атоми, щоб зменшити обсяг даних, які вони зчитують з диспетчерів зберігання (SM). Вузли-агрегатори SingleStore не кешують жодних даних, але листові вузли виконують частини запитів, щоб зменшити обсяг даних, що надсилаються на вузли-агрегатори. Це неможливо в NuoDB, тому що SM це тільки сховище даних. Ці дві системи можуть додавати додаткові ресурси виконання до кластеру СКБД (вузли TE NuoDB, вузли агрегатора SingleStore) без необхідності повторного розбиття бази даних.

На відміну від NuoDB та SingleStore, СКБД Volt Active Data використовує гомогенну архітектуру, де кожен вузол зберігає дані та виконує запити. І сьогодні ще належить з'ясувати, яка з архітектур гетерогенна або гомогенна є кращою з точки зору продуктивності або операційної складності [6].

Ще один новий аспект поділу в системах NewSQL полягає в тому, що деякі з них підтримують динамічну/живу міграцію (live migration). Це дозволяє СКБД переміщувати дані між фізичними ресурсами для повторного балансування та зменшувати кількість гарячих точок (hotspots) або збільшувати/зменшувати ємність СКБД без переривання обслуговування. Це схоже на повторне балансування в системах NoSQL, але складніше, оскільки СКБД NewSQL має підтримувати гарантії ACID для транзакцій під час міграції [14, 54]. Для цього у СКБД використовуються два підходи. Перший полягає в організації бази даних у вигляді безлічі крупнозернистих/гранулярних (coarse-grained) віртуальних/логічних розділів, поширених за фізичними вузлами [55].

Потім, коли СКБД необхідно виконати повторне балансування (re-balance – процес переміщення навантаження з одного вузла в кластері на інший; існують різні способи перебалансування [40]), вона переміщує ці віртуальні розділи між вузлами. Цей підхід використовується в MariaDB Xpand (Clustrix), а також у системах NoSQL, таких як Cassandra та DynamoDB. Інший підхід полягає в тому, щоб СКБД виконувала більш точне перебалансування шляхом перерозподілу окремих кортежів або кортежних груп за допомогою діапазонного розбиття. Цей підхід використовується в H-Store [54]. Рішення з іншою (власною) моделлю (підходом) поділу притаманне СКБД NewSQL Google Cloud Spanner. Сьогодні Spanner широко використовується як система керування базами даних OLTP для структурованих даних у Google. Spanner обслуговує десятки мільйонів запитів за секунду у всіх своїх базах даних, керуючи сотнями петабайт даних [22].

Таким чином, розміщуючи розділи на різних вузлах, що дозволяє масштабуватися до сотень або навіть тисяч вузлів [56], часто можна досягти майже лінійного прискорення, особливо для аналітичних запитів, коли кожен вузол може сканувати свої розділи паралельно [52]. Крім того, секціонування/фрагментація також може підвищити доступність, гарантуючи, що у разі збою одного розділу інші розділи зможуть відхилити деякі транзакції. В ідеалі СКБД також повинна мати можливість розподіляти виконання запиту на декілька розділів, а потім об'єднувати їх результати в один результат. Майже всі системи NewSQL (за винятком, напевно, ScaleArc), які підтримують вбудований поділ, надають цю функціональність.

Розподілена обробка транзакцій у розділених на секції (*partitioned*) базах даних – ідея не нова. Багато основ цих систем було взято з роботи авторів, які брали участь у кінці 1970-х у роботі над проєктом SDD-1 [57]. Крім того, на початку 1980-х років колективи авторів, які створили дві одновузлові СКБД – System R та INGRES, також створили розподілені версії своїх відповідних систем [58, 59]. Але через об'єктивні обставини ці СКБД так і не знайшли

широкого застосування. По-перше, на той час обчислювальні системи були дуже дорогими, тому більшість організацій не могли дозволити собі розгорнути свою базу даних на кластері машин. А по-друге, банально – не було потреби застосунків у високопродуктивній розподіленій СКБД. На той час очікувана пікова продуктивність СКБД зазвичай вимірювалася від десятків до сотень транзакцій на секунду. В даний час, обидва ці припущення вже без сумніву невірні. А створення великомасштабного застосунку з інтенсивним використанням даних сьогодні вже не є рідкістю.

Однак, незважаючи на певні досягнення в галузі розподіленої обробки транзакцій у сучасних розділених на секції базах даних, слід враховувати, що: *а*) час та ресурси, необхідні для створення та підтримки сегментованої архітектури БД, можуть переважити переваги її використання; *б*) на жаль, для робочих навантажень, що складаються з невеликих транзакцій, які стосуються кількох записів, жоден із наведених вище підходів горизонтального розбиття не є ідеальним. Якщо здійснюється доступ до більш ніж одного кортежу, то кругове і геш розбиття зазвичай вимагають доступу до кількох вузлів. Виконання розподілених транзакцій знижує продуктивність у порівнянні з виконанням транзакцій локально. Діапазонне розбиття може бути ефективнішим, але для цього потрібно ретельний вибір діапазонів, що може бути важко зробити вручну. Завдання секціонування/розбиття (partitioning problem) стає ще складнішим, коли транзакції стосуються кількох таблиць, які необхідно розділити за межами транзакцій. Наприклад, важко розділити дані для веб-сайтів соціальних мереж, де схеми часто характеризуються безліччю відносин «багато до багатьох» [52]. Крім того, підтримка цілісності даних може суттєво ускладнитися, оскільки функціонально залежні дані можуть виявитися фрагментованими та розміщуватись на різних вузлах. Тому сьогодні багато фахівців у світі проводять додаткові дослідження для вирішення цих проблем.

2.3. Пам'ять сховища

Із самого початку всі системи СКБД використовували архітектуру зберігання на основі дисків. Тобто в цих системах передбачалося, що основне місце зберігання бази даних знаходиться на запам'ятовуючому пристрої тривалого збереження з блоковою адресацією, такому як SSD (Solid-State Drive – твердотілий накопичувач) або HDD (Hard Disk Drive – жорсткий магнітний диск або накопичувач на магнітних дисках). І оскільки операції читання та запису на ці пристрої виконуються повільно порівняно зі швидкістю основної пам'яті, СКБД змушена використовувати основну/оперативну пам'ять (main memory) для кешування блоків, що зчитуються з диска, і для буферизації оновлень, отриманих в результаті транзакцій. Використання подібних довготривалих запам'ятовуючих пристроїв було викликано тим, що історично основна пам'ять була набагато дорожчою і мала обмежену ємність у порівнянні з дисками. Однак сьогодні сучасні технології досягли такого рівня розвитку, коли потужності та ціни такі, що можна зберігати всі бази даних OLTP (крім дуже великих) повністю у пам'яті (так звані in-memory database – (IMDB), або система баз даних в основній пам'яті – main memory database system (MMDB), або резидентна база даних – memory resident database). Перевага цього підходу полягає в тому, що він дозволяє проводити певну оптимізацію, оскільки СКБД більше не доводиться припускати, що транзакція може отримати доступ до даних у будь-який момент часу, і якщо дані не перебувають у пам'яті, СКБД доведеться перейти в стан очікування. Таким чином, ці системи можуть підвищити продуктивність, тому що багато компонентів, необхідних для обробки у відповідних ситуаціях, наприклад менеджер пула буферів або схеми управління паралелізмом з високим пріоритетом, просто стануть не потрібними.

Ідея зберігання бази даних повністю в оперативній пам'яті не нова [60, 61]. Так, IMS Fast Path [62], випущена в 1976 р., є однією з перших відомих систем, оптимізованих для даних, що знаходяться в пам'яті. В рамках проєкту MM-DBMS Університету Вісконсин-Медісон (University of Wisconsin-Madison) [63 – 65] були проведені основні дослідження, які заклали основу для багатьох аспектів СКБД з оперативною пам'яттю, включаючи індекси, обробку

запитів та алгоритми відновлення. У тому ж десятилітті було розроблено перші розподілені СКБД, засновані на архітектурі зберігання в основній пам'яті (проєкт PRISMA – великомасштабна дослідницька робота Philips Research Laboratory; однією з його цілей була розробка та реалізація розподіленої машини бази даних з оперативною пам'яттю [66]). Нова хвиля таких СКБД спостерігалася у 1990-х роках. У цей час з'явилися такі комерційні СКБД, як Altibase [67], Oracle TimesTen [68], які і сьогодні продовжують функціонувати та підтримуватись розробниками. Деякі NoSQL системи також підтримують можливість зберігання даних у пам'яті, наприклад, Redis (REmote DIctionary Server) – це мережеве сховище ключів та значень у пам'яті, яке можна використовувати як базу даних, кеш або брокер повідомлень. Починаючи з MongoDB Enterprise версії 3.2.6, MongoDB додала механізм зберігання пам'яті в механізми зберігання. Набори реплік MongoDB допускають гібридне розгортання бази даних у пам'яті та на диску.

Потенційною технічною перешкодою використання СКБД, заснованих на архітектурі зберігання в основній пам'яті, вважатиметься можливість втрати живлення. Незалежно від того, з якої причини відбувається втрата живлення, енергозалежна оперативна пам'ять втрачає всі дані. Однак з появою енергонезалежної пам'яті з довільним доступом або так званої технології енергонезалежної пам'яті (NVM, Non-Volatile Memory, Persistent Main Memory, Storage-Class Memory) бази даних в оперативній пам'яті можуть забезпечувати таку ж продуктивність і зберігати дані в разі втрати живлення. Крім того, застосування технології енергонезалежної пам'яті дозволяє використовувати однорівневу архітектуру зберігання даних без використання будь-яких зовнішніх запам'ятовуючих пристроїв, на відміну від традиційних транзакційних in-memory СКБД, змушених використовувати постійне зовнішнє сховище, щоб забезпечити важливу характеристику транзакції – довговічність.

Існує кілька СКБД NewSQL, заснованих на архітектурі зберігання в основній пам'яті, такі, наприклад, як HyPer, SingleStore, SAP HANA, Volt Active Data (H-Store/VoltDB). Ці СКБД за рахунок орієнтації на основну пам'ять працюють значно ефективніше при робочих навантаженнях OLTP систем. Робочі навантаження (workload – обсяг роботи, яку людина чи машина мають виконати за певний період часу [69]) OLTP характеризуються невеликими наборами даних, і великою кількістю простих запитів. Звичайною мірою робочих навантажень OLTP є кількість транзакцій за секунду [70].

Однією з особливостей систем NewSQL, побудованих на архітектурі зберігання в основній пам'яті, є можливість витіснення підмножини бази даних у постійне сховище, щоб зменшити пам'ять, яку вона займає. Це дозволяє СКБД підтримувати бази даних, розмір яких перевищує обсяг доступної пам'яті без необхідності зворотного перемикання на дискову архітектуру.

Загальний підхід, який отримав назву антикешування (anti-caching) [71, 40], полягає у використанні внутрішнього механізму відстеження всередині системи, щоб визначити, до яких кортежів більше немає доступу, а потім вибрати їх для виключення. Тобто за такого підходу за відсутності достатньої кількості доступної пам'яті «холодні» (що найдавніше використовувалися) дані з основної пам'яті переміщаються на диск безпечним для транзакцій способом і завантажуються у оперативну пам'ять при зверненні у майбутньому. Таким чином, «гарячі» дані знаходяться в основній пам'яті, а більш «холодні» – на диску в антикеш-частині системи. Антикешування дозволяє СКБД, побудованих на архітектурі зберігання в основній пам'яті, керувати базами даних, розмір яких перевищує обсяг колективної пам'яті на всіх вузлах [71]. Це нагадує те, що операційні системи роблять з віртуальною пам'яттю і файлами підкачування, але СКБД може керувати пам'яттю ефективніше, ніж операційна система, за рахунок маніпуляцій на рівні окремих записів, а не цілих сторінок пам'яті [40]. Подібний підхід з деякою модифікацією, що дозволяє більш ефективно здійснювати переміщення між основною пам'яттю та сховищем тривалого збереження, використовується у Volt Active Data (H-Store/VoltDB). Існує підтримка багаторівневого антикешування, що дозволяє визначити ієрархічну структуру сховища для перенесення кортежів із DRAM (dynamic random

access memory) у сховище тривалого збереження [28]. Також додано підтримку міграції (переміщення з основної пам'яті та повернення до основної пам'яті) даних на рівні кортежу при використанні енергонезалежної пам'яті. Замість використання так званих «tombstone» (віддалений запис у репліці розподіленого сховища даних) виключені кортежі переміщаються в окремий пул, що зберігається в NVM. Таким чином, можна безпосередньо звертатися до цих кортежів, не перериваючи і не перезапускаючи транзакцію.

Ймовірно, коли ширше використовуватимуться технології енергонезалежної пам'яті, знадобляться подальші зміни в конструкції підсистем зберігання [40, 72]. В даний час це ще відносно нова область досліджень, але вона заслуговує на пильну увагу. У цьому контексті можна відзначити SAP HANA – першу велику СКБД, яка активно використовує NVM [73, 74], а саме можливостей енергонезалежної пам'яті з більш високою щільністю (Intel Optane PMem), яка дозволяє зберігати дані як твердотілі накопичувачі та накопичувачі на жорстких магнітних дисках і забезпечує швидкість, аналогічну до основної пам'яті. Pmem (Persistent Memory) забезпечує унікальне поєднання доступної великої ємності та підтримки збереження даних. Завдяки інноваційній технології, що пропонує різні режими роботи, вона адаптується до потреб залежно від робочих навантажень [75].

SingleStore (MemSQL) використовує інший підхід до баз даних, обсяг яких перевищує обсяг пам'яті. У ній адміністратор може вручну вказати СКБД зберігати таблицю у форматі стовпців (columnar format). Ця СКБД не підтримує жодних метаданих відстеження в пам'яті для цих резидентних на диску кортежів. Вона організує ці дані у сховищі з журнальною структурою (log-structured storage), щоб зменшити накладні витрати на оновлення, які традиційно виконуються повільно у сховищах даних OLAP (online analytical processing).

2.4. Управління паралельною обробкою

Одночасне виконання транзакцій у розрахованих на багато користувачів базах даних може викликати певні проблеми, пов'язані із забезпеченням цілісності (integrity) та узгодженості (consistency) даних, наприклад, такі як: втрачене оновлення, «брудне» читання, неповторюване читання, фантомне читання. Тому керування паралельною обробкою (паралелізмом – concurrency control) є важливою концепцією, яку мають підтримувати бази даних, у тому числі NewSQL. Мета керування паралельною обробкою – запобігання непередбаченому впливу дій одного користувача на дії іншого. Іншими словами, щоб з одного боку в умовах паралельної обробки користувач отримав той же результат, як і у випадку, якби він був єдиним користувачем, а з іншого – щоб дії різних користувачів, якби й впливали один на одного, то тільки прогнозованим чином.

На цей час розроблено різні підходи керування паралельною обробкою, зокрема [9]:

– схеми двофазного блокування (2PL, two-phase locking). При цій стратегії транзакціям дозволяється накладати блокування в міру необхідності, але як тільки перше блокування знімається, транзакція вже не може накласти ніяких інших блокувань;

– упорядкування за позначками часу/часовими мітками (timestamp ordering – TO). СКБД передбачає, що транзакції виконуються в порядку їх позначок часу, при цьому транзакції не будуть виконувати операції, що чергуються, які порушать порядок, що серіалізується. Цей протокол вимагає, щоб всі розподілені вузли мали точно синхронізовані годинники;

– керування паралельним доступом через багатоверсійність/багатоверсійне керування конкурентністю (MVCC, multi-version concurrency control). MVCC – це метод керування паралелізмом, який зазвичай використовується СКБД для забезпечення одночасного доступу до бази даних та мовами програмування для реалізації транзакційної пам'яті. Що стосується конкретно забезпечення паралельного доступу до баз даних, суть цього методу полягає в наданні кожному користувачеві «знімку» бази. При цьому даний «знімок» має таку властивість, що зміни, що вносяться користувачем, невидимі іншим користувачам до моменту фіксації транзакції. Цей спосіб керування дозволяє домогтися того, що транзакції, що пишуть, не блокують ті транзакції, що читають, і транзакції, що читають, не блокують ті, що пишуть.

Для цього MVCC використовує позначки часу та інкрементуючі ідентифікатори (ID) транзакцій для досягнення узгодженості транзакцій. MVCC гарантує, що транзакції ніколи не доведеться чекати на читання об'єкта бази даних, підтримуючи кілька версій об'єкта. Кожна версія об'єкта (кортежу) має позначку часу для читання та позначку часу для запису, що дозволяє конкретній транзакції прочитати останню версію об'єкта, яка передує позначці часу читання. Проблема MVCC – наявність кількох версій елементів даних із різними значеннями. Щоб заощадити місце, старі версії даних повинні періодично видалятися. Але це можна робити тільки коли розподілена СКБД впевнена, що більше не з'явиться транзакція, якій потрібен доступ до видалених версій [9].

Однак, на жаль, сьогодні жоден метод чи механізм керування паралельною обробкою не є ідеальним для всіх випадків [76]. Усі вони передбачають певний компроміс. Наприклад, застосунок може дуже жорстко керувати паралельною обробкою, заблокувавши всю базу даних, але жодні інші програми не зможуть нічого робити під час її виконання. Такий захист надійний, але дорогий. Разом з тим існують заходи, які складніше запрограмувати або реалізувати, але які забезпечують більшу пропускну здатність. Доступні й інші заходи, які максимізують пропускну здатність, але мають низький рівень керування паралельною обробкою. Тому при розробці розрахованих на багато користувачів баз даних доводиться робити вибір серед цих заходів, йдучи на певні компроміси.

І якщо перші розподілені СКБД 1970-80-х років використовували схеми двофазного блокування, то майже всі системи NewSQL, побудовані на нових архітектурах, уникають 2PL через складність роботи із взаємоблокуванням. Замість цього поточна тенденція полягає у використанні варіантів керування паралельною обробкою з упорядкуванням за позначками часу. Найбільш широко використовуваний протокол у системах NewSQL – це децентралізоване керування паралельним доступом у вигляді багатоверсійності (MVCC). Цей механізм використовується практично у всіх системах NewSQL, заснованих на нових архітектурах, зокрема SingleStore, HyPer, SAP HANA, CockroachDB, MariaDB Xpand, LeanXcale. При цьому, незважаючи на інженерні оптимізації та налаштування, які ці системи використовують у своїх реалізаціях MVCC для підвищення продуктивності, основні концепції схеми не є новими. Подібні рішення використовувалися, наприклад, у СКБД VAX Rdb/ELN та InterBase на початку 1980-х років [6].

Інші системи NewSQL для керування паралельною обробкою використовують комбінацію 2PL та MVCC. При такому підході транзакції, як і раніше, повинні накладати блокування за схемою 2PL для модифікації бази даних. Коли транзакція змінює запис, СКБД створює нову версію цього запису так само, як у випадку з MVCC. Ця схема дозволяє запитам лише для читання уникнути необхідності накладання блокувань і, отже, не блокувати транзакції записи. Найбільш відомою реалізацією цього підходу є механізм зберігання (storage engine) InnoDB для СКБД MySQL. Також ця схема використовується в Spanner, NuoDB та MariaDB Xpand (Clustrix).

NuoDB покращує вихідний MVCC, використовуючи gossip/epidemic протокол для широкомовної передачі інформації про версії між вузлами. CockroachDB використовує для керування паралельною обробкою та забезпечення рівня ізоляції «серіалізований» (serializable) різновид MVCC [77]. Комерційна СКБД NewSQL Volt Active Data використовує керування паралельною обробкою на основі методу позначок часу, але замість чергування транзакцій, як і в MVCC, вона планує виконання транзакцій по одній за раз у кожному розділі. Вона також використовує гібридну архітектуру, в якій транзакції з одним розділом плануються децентралізованим чином, а транзакції з кількома розділами плануються за допомогою централізованого координатора (коли запит на транзакцію надходить на вузол, координатор надає запиту на унікальний ідентифікатор на основі позначки часу його надходження). Volt Active Data (H-Store/VoltDB) упорядковує транзакції на основі логічних позначок часу, а потім планує їх виконання у розділі, коли настає їхня черга. Коли транзакція виконується у розділі, вона має ексклюзивний доступ до всіх даних у цьому розділі, і, таким чином, системі

непотрібно встановлювати детальні блокування (fine-grained locks) та засувки (latches) для своїх структур даних. Це дозволяє транзакціям, яким потрібний доступ тільки до одного розділу, ефективно виконуватись, оскільки інші транзакції не конкурують одна з одною.

Недоліком підходу керування паралелізмом на основі розділів є те, що він не працює належним чином, якщо транзакції охоплюють кілька розділів, оскільки затримки мережного обміну даними змушують вузли не діяти в очікуванні повідомлень [6]. Слід зазначити, що керування паралельною обробкою на основі розділів також не є новою ідеєю. Вперше цей підхід було запропоновано у роботі [78].

Все проміжне програмне забезпечення та служби DBaaS успадковують схему керування паралелізмом архітектури СКБД, що лежить в їх основі. А оскільки більшість із них використовують MySQL, для керування паралельною обробкою вони використовують комбінацію 2PL та MVCC. В цілому, слід зазначити, що в основних схемах керування паралельною обробкою в системах NewSQL немає нічого принципово нового. Але це не зменшує їхньої значущості, оскільки реалізації в них відомих механізмів дозволяють ефективно функціонувати системі в контексті сучасного обладнання та розподілених операційних середовищ.

2.5. Вторинні індекси

Строго кажучи, індекси в цілому не є обов'язковим компонентом СКБД, але вони можуть істотно підвищити її продуктивність. Використання вторинного індексу – це спосіб ефективного доступу до записів у базі даних за допомогою деяких атрибутів таблиці, відмінних від атрибутів первинного ключа. Їх просто підтримувати у нерозділених на секції (non-partitioned) СКБД, оскільки вся база даних розташована на одиночному вузлі. Проблема з вторинними індексами виникає у розподіленій СКБД. Вона полягає в тому, що вторинні індекси не завжди можуть бути розділені так само, як і решта бази даних. Існуючі сьогодні конструктивні рішення для підтримки вторинних індексів у розподіленій СКБД залежать від того:

- де система їх зберігатиме;
- як вона підтримуватиме їх у контексті транзакцій.

У системі з централізованим координатором, як і в ПЗ проміжного рівня, що забезпечує прозоре функціонування з сегментованою на кількох вузлах БД, вторинні індекси можуть перебувати як на вузлі координатора, так і на вузлах сегментів. Перевага цього підходу полягає в тому, що у всій системі існує лише одна версія індексу, і тому його підтримувати легше. Усі системи NewSQL, засновані на нових архітектурах, децентралізовані та використовують секційовані вторинні індекси. Це означає, що кожний вузол зберігає частину індексу, а не повну копію. Компроміс між секційованими та реплікованими індексами полягає в тому, що з запитами, що мали місце, можливо, знадобиться охопити декілька вузлів, щоб знайти те, що вони шукають, але, якщо транзакція оновлює індекс, їй потрібно буде його змінити тільки на одному вузлі. У реплікованому індексі ролі міняються місцями. Пошуковий запит може бути задоволений лише одним вузлом у кластері, але щоразу, коли транзакція змінює атрибути, вказані у базовій таблиці вторинного індексу (ключ чи значення), СКБД має виконувати розподілену транзакцію, яка оновлює всі копії індексу.

Прикладом децентралізованого вторинного індексу, в якому поєднуються ці концепції, є MariaDB Xpand. Спочатку СКБД підтримує реплікований крупнозернистий (coarse-grained), заснований на діапазоні індекс на кожному вузлі, який відображає значення в розділи/частини (partitions). Це відображення дозволяє СКБД надсилати запити до відповідного вузла, використовуючи атрибут, який не є атрибутом розбиття таблиці. Потім ці запити будуть звертатися до другого секційованого індексу на цьому вузлі, який відображає точні значення в кортежі. Такий дворівневий підхід знижує об'єм координації, необхідної для синхронізації реплікованого індексу в кластері, оскільки він відображає лише діапазони, а не окремі значення. Додавання вторинного індексу до стовпців таблиць бази даних Google Cloud Spanner підвищує ефективність пошуку даних у цих стовпцях, у тому числі за рахунок ска-

нування індексу, а не повного сканування таблиці. Spanner зберігає такі дані у кожному вторинному індексі [79]: все ключові стовпці з базової таблиці; всі стовпці, включені до індексу; усі стовпці, зазначені в необов'язковому реченні STORING (діалект SQL для Google) або у реченні INCLUDE (діалект для PostgreSQL) визначення індексу. З часом Spanner аналізує таблиці, щоб упевнитися, що вторинні індекси використовуються для відповідних запитів. LeanXcale підтримує ефективні вторинні індекси, які розподіляються та зв'язуються з первинними даними. Запити до вторинних індексів можуть витягувати первинні дані локально і надавати повний результат з мінімальною кількістю циклів обробки. У SAP HANA підтримується довільна кількість унікальних та неунікальних вторинних індексів. Усередині вторинні індекси перетворюються на два різні варіанти, залежно від кількості задіяних стовпців [80]:

– індекси для окремих стовпців. Під час створення індексу для окремого стовпця сховище стовпців створює інвертований список (інвертований індекс – inverted index), який відображає ідентифікатори значень словника у відповідні записи у векторі індексу. Усередині створюються дві структури індексу: одна для дельта-таблиці та одна для основної таблиці (кожна таблиця сховища стовпців складається з двох окремих частин: основної таблиці та дельта-таблиці; у той час як основна таблиця призначена тільки для читання, сильно стиснута та оптимізована для читання, дельта-таблиця відповідає за відображення змін, виконаних операціями INSERT, UPDATE, DELETE). Коли цей індекс створюється для сховищ рядків, створюється лише один окремий індекс B+ дерева;

– індекси за кількома стовпцями (конкатеновані / складові індекси – concatenated indexes). Індекс з кількома стовпцями може бути корисним, якщо часто запитується певна комбінація атрибутів, або для прискорення обробки з'єднання, коли задіяно кілька атрибутів. Слід звернути увагу на те, що при створенні складеного / конкатенованого індексу окремі індекси для складових атрибутів не створюються (на відміну від первинного ключа, де також створюються індивідуальні індекси для кожного із складових атрибутів). Сховище стовпців підтримує інвертований індекс значень, інвертований геш-індекс та інвертований індивідуальний індекс для багатостолбцових індексів (multi-column indexes). При створенні складеного / конкатенованого індексу для сховища рядків створюється лише один окремий індекс B+ дерева.

Найбільш поширений спосіб створення вторинних індексів розробниками при використанні СКБД NewSQL, яка їх не підтримує – це застосування індексу з використанням розподіленого кеша в пам'яті, такого як Memcached [81]. Але використання зовнішньої системи вимагає, щоб застосунок підтримував кеш, оскільки СКБД не будуть автоматично скасовувати зовнішній кеш [6].

2.6. Механізм відновлення після збою

Важливою характеристикою СКБД NewSQL, пов'язаної із забезпеченням відмовостійкості (fault tolerance), є її механізм відновлення після збоїв (crash recovery). При цьому на відміну від традиційних СКБД, де основним завданням відмовостійкості є гарантія відсутності втрати оновлень, СКБД NewSQL також повинні мінімізувати час простою [6]. Так як передбачається, що сучасні веб-застосунки постійно будуть перебувати в мережі, а перебої в роботі сайту обходяться дорого.

Традиційний підхід до відновлення в системі з одним вузлом без реплік полягає в тому, що коли СКБД повертається в оперативний (online) режим після збою, вона завантажує останню контрольну точку (точку синхронізації між базою даних та журналом транзакцій), а потім відтворює свій журнал із записом наперед (WAL, write-ahead log) для повернення стану бази даних у той стан, у якому вона була на момент збою. Класичний метод цього підходу, відомий як ARIES (Algorithm for Recovery and Isolation Exploiting Semantics) [82], було винайдено дослідниками IBM у 1990-х роках. Усі основні СКБД реалізують той чи інший варіант ARIES, який підтримує часткові відкочування транзакцій, блокування з високим ступенем деталізації (наприклад, запис) та відновлення з використанням ведення журналу із

записом наперед. Однак у розподіленій СКБД з репліками традиційний підхід з одним вузлом не застосовується безпосередньо. Це пов'язано з тим, що під час збою головного вузла система активізує один із підлеглих вузлів в якості нового головного. Коли попередній головний вузол повертається до online режиму, він може просто завантажити свою останню контрольну точку і повторно запустити свій WAL, оскільки СКБД продовжує обробляти транзакції і, отже, стан бази даних просунувся вперед. Відновлювальному вузлу необхідно отримати оновлення від нового головного вузла (і, можливо, інших реплік – вузлів, на якому зберігається копія бази даних), які він пропустив, коли був у неробочому стані. Є два можливі способи зробити це. По-перше, відновлюючий вузол може завантажити свою останню контрольну точку та WAL зі свого локального сховища, а потім запросити зміни, які він пропустив із записів журналів інших вузлів. Поки вузол може обробляти журнал швидше, ніж до нього додаються нові оновлення, вузол врешті-решт прийде до того ж стану, що й інші реплік вузли. Це можливо, якщо СКБД використовує журнал із записом наперед, оскільки час застосування оновлень журналу безпосередньо до кортежів набагато менше, ніж час, необхідний для виконання вихідного оператора SQL. Інший варіант, що дозволяє скоротити час, необхідний для відновлення, передбачає відмову вузла, що відновлюється, від своєї контрольної точки на користь використання нової точки, з якої вузол буде відновлюватися. Тобто вузол, що відновлюється, може прийняти поточний стан нового головного вузла і продовжити роботу в якості звичайного вузла. Додаткова перевага цього підходу полягає в тому, що цей механізм можна використовувати і в СКБД для додавання нового вузла репліки.

LeanXcale надає механізм гарячого резервного копіювання (hot backup), що дозволяє виконувати повне резервне копіювання (full backup) розподіленої бази даних, гарантуючи послідовне відновлення на певний момент часу.

Ведення журналу в системах, що базуються на архітектурі зберігання в основній пам'яті, оптимізована для забезпечення високої пропускної здатності та низької затримки [61]. Оскільки введення-виведення журналу є основним вузьким місцем, ці системи намагаються максимально зменшити обсяг журналу навіть більше, ніж системи на основі дисків. Volt Active Data (H-Store/VoltDB) використовує спрощений варіант ведення журналу, званого журналом команд (command log) [48, 83], який реєструє лише запит на запуск транзакції (збереженої процедури). Тобто виклик кожної збереженої процедури являє собою транзакцію. Одна збережена процедура може включати безліч окремих операторів SQL, і кожен оператор SQL може змінювати сотні або тисячі рядків таблиці. Запис журналу для цієї схеми складається виключно з імені збереженої процедури, її вхідних параметрів та ідентифікатора транзакції. Процес протоколювання команд (ведення журналу команд) представлено на рис. 3. Частота (frequency), з якої транзакції записуються до журналу команд, налаштовується. Регулюючи частоту та тип ведення журналу (синхронний або асинхронний), ви можете збалансувати потреби вашого застосунку у продуктивності з бажаним рівнем надійності [83].

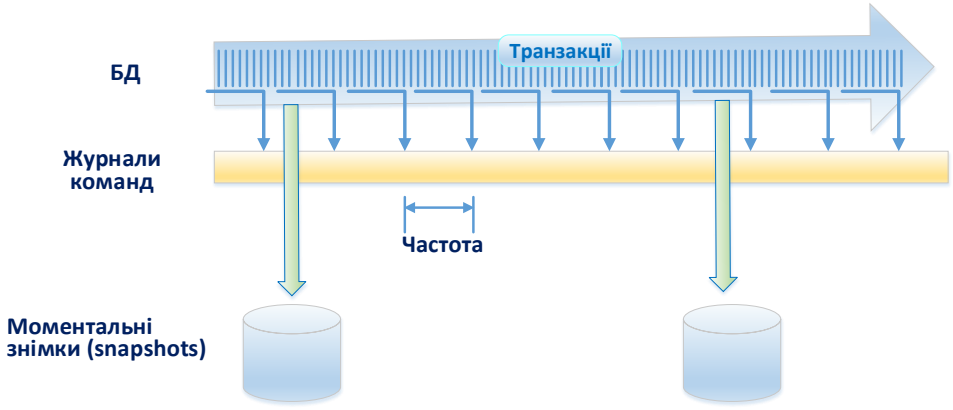


Рис. 3. Процес протоколювання команд (ведення журналу команд)

Щоб ще більше звести до мінімуму трафік журналу, деякі системи взагалі уникають реєстрації даних індексу. Реєструються лише оновлення базових даних. Після збою індекси будуються із нуля після відновлення базових записів. Крім того, СКБД, засновані на архітектурі зберігання в основній пам'яті, намагаються максимально розпаралелити введення-виведення журналів. На додаток, оскільки контрольні точки у цих системах зазвичай більше, ніж їх аналоги в дискових системах (це пов'язано з тим, що системи баз даних в основній пам'яті зазвичай записують усі (або більшість) рядків у сховище), були розроблені полегшені методи створення контрольних точок в основній пам'яті, наприклад, CALC (Checkpointing Asynchronously using Logical Consistency) [84], який переходить в режим копіювання при записі, записуючи тільки оновлення, що відбулися з моменту попередньої контрольної точки. Volt Active Data (H-Store/VoltDB) також використовує аналогічну схему копіювання під час запису для створення асинхронних моментальних знімків даних на кожному вузлі. Резервні копії та контрольні точки (РККТ) в in-memory базах даних, необхідні для забезпечення відмовостійкості БД, повинні зберігатися на постійних пристроях зберігання, таких як SSD або жорсткі диски. База даних повинна гарантувати актуальність резервних копій та мінімальний час відновлення [14].

У більшості СКБД, заснованих на архітектурі зберігання в основній пам'яті, дані відновлюються шляхом завантаження останньої дійсної (valid) контрольної точки з подальшим відтворенням кінця журналу, уникаючи повного скасування [61]. Це відноситься і до баз даних, встановлених на платформі СКБД Volt Active Data (H-Store/VoltDB), дані яких відновлюються шляхом завантаження останнього повного узгодженого моментального знімка (snapshot), після чого відтворюються (replay) транзакції зі свого журналу команд (починаючи з моменту останнього моментального знімка), щоб привести базу даних до узгодженого стану (рис. 4).

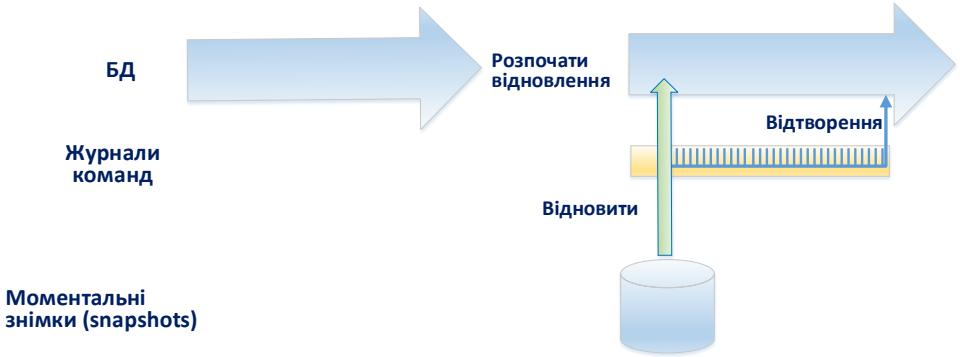


Рис. 4. Процес відновлення БД

Проміжне програмне забезпечення та системи DBaaS спираються на вбудовані механізми базових одновузлових СКБД, але додають додаткову інфраструктуру для вибору головного вузла та інші можливості керування. Системи NewSQL, засновані на новій архітектурі, використовують комбінацію готових компонентів (наприклад, ZooKeeper, Raft) та власних реалізацій існуючих алгоритмів (наприклад, Paxos) [61]. Все це стандартні процедури та технології, доступні у комерційних розподілених системах з 1990-х років.

2.7. Забезпечення безпеки

Безпека баз даних важлива для будь-якої організації користувача, які використовують великі набори даних як значущий актив. При цьому слід мати на увазі, що поняття безпеки відносяться не лише до даних, що зберігаються у базі даних. Порушення безпеки можуть торкнутися інших частин системи, що, своєю чергою, може вплинути на базу даних. Тобто безпека БД залежить від захищеності використовуваного обладнання, програмного забезпечення, власне даних від навмисних та/або випадкових загроз.

Розуміючи важливість питання безпеки, розробники СКБД, у тому числі NewSQL, намагаються вирішити ці питання в рамках своїх систем, не покладаючись на захист через застосунки. Засоби та методи захисту в різних NewSQL системах відрізняються один від одного, однак у тій чи іншій мірі в них, як правило, використовуються зрілі, перевірені часом механізми захисту, що застосовуються у традиційних реляційних СКБД. Так, наприклад, NuoDB подібно до класичних SQL БД підтримує механізми: обмежень цілісності (первинні, унікальний, зовнішні ключі, обмеження not null), надання прав доступу, ролей, ведення журналу бази даних. Крім того, у NuoDB підтримується прозоре шифрування даних (TDE – Transparent Data Encryption), аналогічно використовуваному в Oracle Database, Microsoft SQL Server. TDE гарантує, що дані користувача, що зберігаються в архіві, журналі, резервних копіях, spill-файлах (файли для збереження проміжних даних, коли в пам'яті недостатньо пам'яті для виконання запиту) будуть зашифровані перед записом на диск. Щоб захистити дані, збережені на диску, від несанкціонованого доступу на рівні операційної системи, СКБД SAP HANA також підтримує цей механізм шифрування на відповідному рівні (шифрування тому даних захищає область даних, шифрування журналу повторного виконання (журналу із записом наперед) захищає область журналу на диску). Дані зашифровуються за допомогою алгоритму AES-256-CBC. При цьому, якщо використовується енергонезалежна пам'ять, можна використовувати апаратне шифрування для захисту даних на диску. Коли законний користувач отримує доступ до бази даних SAP HANA за допомогою інтерфейсу клієнта (наприклад, ODBC, JDBC або HTTP), його можливість виконувати операції над об'єктами бази даних визначається наданими йому привілеями. Усі привілеї, надані користувачеві прямо чи опосередковано через ролі, об'єднуються. Це означає, що кожного разу, коли користувач намагається отримати доступ до об'єкта, система перевіряє авторизацію користувача, призначених йому ролей і безпосередньо наданих привілеїв. У SAP HANA використовуються кілька типів привілеїв: системні (system), об'єктні (object), аналітичні (analytic), пакетні (package) та прикладні (application). При цьому привілеї можуть зв'язуватися з роллю (роль може містити будь-яку кількість привілеїв). Як тільки всі запитані привілеї будуть знайдені, система надає доступ. У SAP HANA підтримується механізм маскування даних (data masking), що забезпечує так званий додатковий рівень керування доступом, який можна застосовувати до таблиць та уявлень. Методи анонімізації (anonymization), доступні в SAP HANA, дозволяють отримувати статистично достовірні відомості зі збережених даних, захищаючи при цьому конфіденційність окремих осіб.

Моделі авторизації в CockroachDB використовують особливу концепцію користувача та ролей. CockroachDB не має технічних відмінностей між роллю або користувачем. Так виконання операторів SQL CREATE USER і CREATE ROLE призведе до створення одного й того самого об'єкта з одним винятком: CREATE ROLE буде додано NOLOGIN опція за замовчуванням, що запобігає використанню користувача/ролі для входу в систему. Коли користувач підключається до бази даних або через вбудований клієнт SQL, або через клієнтський драйвер, CockroachDB перевіряє привілеї користувача та ролі для кожного оператора, що виконується. Якщо користувач не має достатніх привілеїв для виконання оператора, CockroachDB видає помилку. Крім привілеїв на: базу даних, функцію, схему, таблицю, послідовність, тип, у CockroachDB є так звані привілеї системного рівня – це особливий вид привілеїв, які застосовуються до всього кластера, що означає, що привілей не прив'язаний до будь-якого конкретного об'єкта у базі даних. Мережевий трафік у CockroachDB між вузлами, а також від клієнтів до вузлів шифрується за допомогою TLS (Transport Layer Security). Шифрування даних вузла на локальному диску забезпечується за допомогою прозорого механізму шифрування. Він дозволяє шифрувати всі файли на диску за допомогою AES у режимі лічильника з розміром ключа 128, 192 або 256 біт [85].

Безпека в LeanXcale досягається за рахунок:

- мережного шифрування, що забезпечує конфіденційність даних (дані зашифровані та доступні для читання тільки клієнту та LeanXcale), автентифікацію повідомлення

(підтверджує, що повідомлення надіслано з LeanXcale або допустимого клієнта), цілісність повідомлень (підтверджує, що жодне з повідомлень не було змінено з моменту відправлення), невідмовність (nonrepudiation – не дозволяє відправникам заперечувати надсилання зашифрованого повідомлення);

- шифрування даних (гарантує конфіденційність даних, що зберігаються на жорстких дисках, завдяки використанню сучасного алгоритму шифрування AES; рекомендується використання AES 256 у режимі GCM – Galois/Counter Mode);

- автентифікації користувача;

- авторизації, що управляє ролями та дозволами для кожної таблиці та користувача в базі даних.

Автентифікація в MariaDB Xpand здійснюється за допомогою облікових записів користувачів бази даних. Облікові записи користувачів бази даних визначаються ім'ям користувача, ім'ям хоста, з якого підключається обліковий запис, та методом автентифікації, налаштованим для облікового запису. MariaDB Xpand підтримує систему керування доступом, аналогічну до тієї, що використовується в MySQL. Привілеї можуть надаватися глобально, на рівні БД або на рівні таблиці. Також для керування доступом до об'єктів бази даних MariaDB Xpand використовує ролі. Тобто є можливість надання привілеїв ролям та відкликання привілеїв у ролей. Користувачеві може призначатися кілька ролей, і користувач може активувати кілька ролей одночасно. Після надання користувачеві ролі він може використовувати всі привілеї, якими володіє дана роль. При наданні ролі іншої ролі, ця роль успадковує всі привілеї, що належать наданої ролі, і користувачі з цією роллю можуть використовувати всі привілеї, що належать обом ролям.

Контроль доступу користувачів та груп до ресурсів Google Cloud Spanner на рівні проєкту, екземпляра Spanner та бази даних Spanner здійснюється за допомогою системи керування обліковими записами та доступом (IAM – Identity and Access Management). Використання IAM дозволяє надавати дозволи (permissions) користувачеві або групі без необхідності змінювати кожен екземпляр Spanner або дозвіл на базу даних окремо. Дозволи дають змогу користувачам виконувати певні дії з ресурсами Spanner. Дозволи не надаються користувачам безпосередньо. Користувачам надаються так звані попередньо визначені / зумовлені ролі (predefined roles) або ролі, що настроюються (custom roles), які мають один або кілька дозволів, пов'язаних з ними. Крім того, Cloud Spanner підтримує деталізований (fine-grained) контроль доступу, який поєднує переваги IAM з традиційним контролем доступу на основі ролей. Деталізований контроль доступу дозволяє налаштувати докладні політики доступу на рівні таблиць і стовпців. Для керування політиками на рівні таблиць та стовпців використовуються відповідні оператори DDL (Data Definition Language) SQL та функції IAM. Комплексна стратегія безпеки Google включає шифрування в стані спокою (зберігання), яке допомагає захистити вміст клієнтів від зловмисників. Зашифровується весь контент клієнтів Google, що зберігається, без будь-яких дій з боку останніх. Усі системи зберігання Google використовують однакову архітектуру шифрування, хоча деталі реалізації відрізняються від системи до системи. У Spanner є три рівні шифрування. Дані в стані спокою розбиваються на фрагменти підфайлів для зберігання і кожен фрагмент шифрується на рівні сховища за допомогою окремого ключа шифрування. Розмір кожного фрагмента може досягати кількох гігабайт. Ключ, який використовується для шифрування даних у блоці, називається ключем шифрування даних (DEK – data encryption key). Два фрагменти не будуть мати однаковий DEK, навіть якщо вони належать одному і тому ж клієнту або зберігаються на одному комп'ютері. Якщо фрагмент даних оновлюється, він шифрується за допомогою нового ключа, а не повторним використанням існуючого ключа. Такий поділ даних, у кожному з яких використовується свій ключ, обмежує ризик потенційної компрометації ключа шифрування даних лише цим блоком. Через великий обсяг ключів у Google та необхідності малої затримки та високої доступності ці ключі зберігаються поруч із даними, які вони шифрують. DEK зашифровуються («обгортаються» – wrapped) за допомогою ключа шифрування ключів (KEK – key en-

cryption key). Нарешті, кожен KEK шифрується ключем шифрування, керованим клієнтом (customer-managed encryption key). Google за допомогою алгоритму AES шифрує дані перед записом їх у систему зберігання БД або на апаратний диск. Шифрування вбудовано у всі системи зберігання. Кожен блок даних має унікальний ідентифікатор. Списки контролю доступу (ACL – access control lists) допомагають гарантувати, що кожен фрагмент може бути розшифрований лише службами Google, які працюють з авторизованими ролями, яким надається доступ лише в даний момент часу. Це обмеження доступу допомагає запобігти доступу до даних без авторизації, зміцнюючи безпеку та конфіденційність даних.

У SingleStore для розмежування доступу використовується метод керування доступом на основі ролей (RBAC – Role Based Access Control). Для цього визначається стандартний перелік таких ролей, можливі зв'язки між користувачами, ролями та групами: роль може мати декілька привілеїв; група може мати кілька ролей; у групі може бути декілька користувачів; користувач може мати кілька ролей; користувач може бути включений до декількох груп; користувачі успадковують дозволи, ролі груп, до яких вони належать. Крім того, у SingleStore підтримується детальний контроль доступу – захист на рівні рядків (Row Level Security). Безпека на рівні рядків у SingleStore досягається за допомогою створення уявлення таблиці зі спеціальним стовпцем ролей. Всі дії з базою даних записуються в журнал аудиту (існує 11 рівнів ведення журналу, які можна розділити на три категорії, кожна з яких має зростаючий рівень деталізації). Специфікація формату шифрування дисків SingleStoreDB сумісна з LUKS (Linux Unified Key Setup). Дані у стані спокою в SingleStoreDB можна захистити за допомогою IBM Guardium Data Encryption. Усю інформацію SingleStore, включаючи файли даних, резервні копії та журнали, можна захистити від несанкціонованого доступу, у тому числі з боку неавторизованих користувачів з правами адміністратора, за допомогою прозорого шифрування CipherTrust Transparent Encryption (CTE) від Thales. Цей процес також відомий як прозоре шифрування бази даних або TDE.

За замовчуванням Volt Active Data (VoltDB) не виконує жодних перевірок безпеки, коли клієнтський застосунок відкриває з'єднання з базою даних або викликає збережену процедуру. Це зручно при розробці та розповсюдженні застосунків у приватній мережі. Однак у загальнодоступних або напівприватних мережах важливо переконатися, що лише відомі клієнтські застосунки взаємодіють із базою даних. При відповідних налаштуваннях безпеки ім'я користувача та пароль, передані клієнтським застосунком, перевіряються сервером на відповідність їх користувачам, визначеним у файлі конфігурації. Якщо клієнтський застосунок передає допустиму пару імені користувача та пароля для облікового запису, термін дії якого ще не минув, з'єднання встановлюється. Коли застосунок викликає збережену процедуру, дозволи перевіряються знову. Якщо схема визначає, що користувачу призначено роль, що має доступ до цієї збереженої процедури, процедура виконується. В іншому випадку застосунку, що викликає, повертається помилка. Volt Active Data підтримує механізм ролей.

MariaDB MaxScale забезпечує корпоративну безпеку, перехоплюючи запити до бази даних до того, як вони досягнуть БД, що дозволяє запобігти розкриттю та пошкодженню даних, а також захистити базу даних від виведення з ладу через зловмисні або випадкові запити. У MariaDB MaxScale є механізм захисту від DoS-атак. Розширені функції безпеки у MariaDB MaxScale можуть відхиляти/блокувати небезпечні, непідтверджені (unapproved) або інші підозрілі запити, а також змінювати результати запитів для захисту даних [86]. Спеціальний фільтр обмеження результатів запобігає випадковим або зловмисним запитам, які роблять базу даних недоступною або розкривають великі обсяги даних, обмежуючи кількість результатів, які може повернути запит (наприклад, щоб запит не повертав усі десять мільйонів рядків у таблиці). Це не тільки зменшує трафік, а й запобігає розкриттю значних обсягів даних зловмисникові, який намагається різними способами їх отримати. Для захисту конфіденційної та/або особистої інформації від розкриття використовується механізм динамічного маскування. Для безпечних з'єднань із застосунками та базами даних MariaDB MaxScale підтримує SSL/TLS. Крім того, MariaDB MaxScale підтримує PAM (Pluggable Authentication

Modules) та GSSAPI (основною використовуваною реалізацією механізму GSSAPI (Generic Security Service Application Program Interface) є Kerberos) для автентифікації. MariaDB MaxScale може бути налаштований на використання проксі-протоколу передачі інформації про клієнта на сервер MariaDB для подальшого спрощення керування користувачами.

Amazon Aurora підтримує кілька способів автентифікації користувачів бази даних: автентифікацію за паролем (доступна за замовчанням для кластерів БД); IAM автентифікацію (для Aurora MySQL, Aurora PostgreSQL); Kerberos автентифікацію для одного кластера БД (для Aurora PostgreSQL). Конкретний користувач може увійти до бази даних, використовуючи лише один метод автентифікації. Amazon Aurora може шифрувати кластери БД. Для цього використовується алгоритм шифрування AES-256. Зашифровуватись можуть дані базового сховища кластерів БД, його автоматичні резервні копії, репліки читання та моментальні знімки. У процесі роботи Amazon Aurora виконує автентифікацію доступу і прозоро з мінімальним впливом на продуктивність розшифровує дані. Однак слід пам'ятати, що шифрування існуючого незашифрованого екземпляра Aurora на даний момент не підтримується. Тому, щоб використовувати шифрування Amazon Aurora для існуючої незашифрованої бази даних, потрібно спочатку створити новий екземпляр БД із «включеним» шифруванням, після чого перенести до нього відповідні дані. Aurora інтегрується з Amazon GuardDuty для виявлення потенційних загроз даним, що зберігаються в базах даних Aurora. GuardDuty RDS Protection створює профілі та відстежує дії по входу в систему та нові бази даних в обліковому записі користувача, крім того, він використовує спеціалізовані моделі машинного навчання для виявлення підозрілих входів до бази даних Aurora. Для управління доступом до даних в Amazon Aurora підтримується механізм так званих груп безпеки (контролюють доступ трафіку в кластер БД і з нього; в групі безпеки можна вказати до 20 правил), привілеїв та ролей.

У табл. 1 наведено зведені дані про деякі характеристики відомих систем NewSQL.

Якщо говорити в цілому про відмінності між традиційними реляційними, NoSQL та NewSQL базами даних для масштабованих рішень, то в табл. 2 наведено деякі важливі їх відмітні характеристики.

Узагальнюючи все сказане, а також деякі висновки, зроблені іншими авторами [87], слід зазначити, що основними перевагами баз даних NewSQL є:

– забезпечення більш високої узгодженості даних (підтримка ACID-транзакцій);

– можливість використання відомих, перевірених часом, операторів SQL та стандартного інструментарію;

– багатші можливості застосування методів аналізу з використанням SQL та розширень;

– кластеризація у стилі NoSQL з використанням більш традиційних моделей даних та запитів.

Крім того, можна також додати, що NewSQL системи дозволяють виконувати гібридну транзакційно-аналітичну обробку (HTAP), мета якої – виконання застосунків OLAP та OLTP на одних і тих самих даних. Це дозволяє проводити аналіз оперативних даних у реальному масштабі часу без традиційного поділу оперативної бази даних та сховища даних, уникаючи складнощів, пов'язаних з ETL (Extract, Transform, Load – витяг, перетворення, завантаження).

Основні недоліки NewSQL баз даних:

– архітектури в оперативній пам'яті можуть не підходити для обсягів, що перевищують кілька терабайт;

– підтримують лише частковий доступ до багатого інструментарію традиційних систем SQL.

ISSN 0485-8972 Radiotekhnika No. 211 (2022)
eISSN 2786-5525

Таблиця 1

Категорія	СКБД	Рік випуску	Поточний випуск/дата	Архітектура зберігання даних	Тип	Розбиття	Керування паралельною обробкою	Реплікація	Короткий опис
Системи з використанням нової архітектури	**MariaDB Xpand** (раніше **Clustrix)**	Clustrix, 2006	MariaDB Xpand: v6.1.0, 02.2023	Гібридна: у пам'яті та на SSD	OLTP	Підтримує	MVCC+2PL	active-active, active-passive	Розподілена БД SQL, яка ефективно масштабується для сучасних веб-застосунків з великим робочим навантаженням, які потребують суворої узгодженості та цілісності даних.
	CockroachDB	2014	v21.1.2, 06.2021	Дискове сховище	OLTP	Підтримує	MVCC	active-passive	Розподілена БД SQL з відкритим вихідним кодом, побудована на транзакційному та строго узгодженому сховищі ключ-значення. Задля узгодженості використовує алгоритм розподіленого консенсусу Raft.
	Google Cloud Spanner	2012 (Spanner) / 2017 (Cloud Spanner)	Немає даних	Дискове сховище	OLTP	Підтримує	MVCC+2PL	active-passive	Реплікована за глобальною мережею СКБД без спільного використання ресурсів, що використовує спеціальне обладнання (GPS, атомний годинник) для генерації позначок часу/часових міток (для високоточної синхронізації годинника). Широко використовується як СКБД OLTP для структурованих даних у Google та загальнодоступна у бета-версії як Cloud Spanner на хмарній платформі Google (GCP, Google Cloud Platform).
	Volt Active Data (раніше **H-Store/ VoltDB)**	2007 (H-Store) / 2008 (VoltDB)	v.12.1, 12.2022	У пам'яті (in-memory). РККТ – на SSD або HDD	OLTP	Підтримує	TO	active-active, active-passive	Розподілена база даних із сегментуванням та реплікацією даних. Є як комерційна версія, так і версія спільноти з відкритим вихідним кодом. Орієнтована на певний сегмент бізнес-обчислень, а саме на швидку обробку великих потоків даних.
	HyPer	2010	Немає даних	У пам'яті (in-memory)	HTAP: OLTP+ OLAP	Підтримує	MVCC	active-passive	База даних в оперативній пам'яті, метою якої є досягнення високої продуктивності як для робочого навантаження OLTP, так і для OLAP (HTAP – Hybrid transaction/analytical processing).
	SingleStore (раніше **MemSQL)**	2012	SingleStore: v8.0 12.2022	Rowstore – у пам'яті; Columnstore – на диску	HTAP	Підтримує	MVCC	active-passive	Розподілена БД у пам'яті, призначена як для транзакційних, так і для аналітичних робочих навантажень. Підтримує дротовий протокол MySQL. Дані для таблиць rowstore (зберігають інформацію у форматі рядків) зберігаються у пам'яті. Дані таблиць columnstore (сховище стовпців) зберігаються на диску. Миттєві знімки та журнали транзакцій зберігаються на диску.
	NuoDB	2013	v5.0.1, 2020	Гібридна: TE – у пам'яті, SM – на диску	OLTP	Підтримує	MVCC	active-passive	Розподілена система керування базами даних, що використовує архітектуру у стилі мікросервісів, яка поділяє обробку на три рівні: рівень управління, рівень транзакцій та рівень зберігання. Механізми транзакцій (TE) надають кеш, який є базою даних у пам'яті для швидкої обробки транзакцій. Диспетчери сховища (SM) забезпечують довговічність.

Продовження табл. 1

Категорія	СКБД	Рік випуску	Поточний випуск/дата	Архітектура зберігання даних	Тип	Розбиття	Керування паралельною обробкою	Реплікація	Короткий опис
	LeanXcale	2015	Немає даних	Дискове сховище	HTAP	Підтримує	MVCC LeanXcale	active-active	Швидка та масштабована БД, яка поєднує в собі характеристики SQL та NoSQL. Вона створена для прийому великих потоків даних у реальному масштабі часу, а також для можливості паралельного виконання різноманітних запитів для будь-якого використання.
	SAP HANA	2010	v.2.0 SPS06, 03.2022	У пам'яті (in-memory). РККТ – на дисках	HTAP	Підтримує	MVCC	active-passive	Система підтримує зберігання таблиць як по стовпцям, так і по рядках (у першому випадку для транзакційних навантажень, у другому – для аналітичних). SAP HANA активно використовує NVM.
ПЗ проміжного рівня	**MariaDB MaxScale**	2015	(v22.08.3, 12.2022); (v2.5.24, 01.2023); (v6.4.4, 11.2022)	-	HTAP	Підтримує	MVCC+2PL	active-active, active-passive	Проксі БД, маршрутизатор і балансувальник навантаження, який розширює можливості високої доступності, масштабованості та безпеки сервера MariaDB, відокремлюючи його від базової інфраструктури БД, та спрощує розробку застосунків. MariaDB MaxScale може надсилати різні запити до різних екземплярів БД.
	ScaleArc	2009	v.3.12 GA, 10.2017	-	HTAP	Підтримує	Mixed	active-passive	ScaleArc – балансувальник навантаження БД. Він дозволяє адміністраторам БД створювати високодоступні, масштабовані та прості в керуванні, обслуговуванні та міграції розгортання бази даних. Програмне забезпечення ScaleArc доступне для баз даних MySQL, SQL Server та Oracle. ScaleArc працює з Microsoft SQL Server та MySQL як локальне рішення, у хмарі для відповідних PaaS або як рішення DBaaS, включаючи Amazon RDS або AzureSQL.
БД як послуга	**Amazon Aurora**	2014	Aurora MySQL 3, 11.2021; Aurora PostgreSQL 14.3, 06.2022	Дискове сховище	OLTP	Не підтримує	MVCC	active-passive	Amazon Aurora – це реляційна база даних, що пропонується як сервіс в Amazon AWS. Заснована на версії MySQL з відкритим вихідним кодом, це комерційна база даних, заявлена як сумісна з MySQL та PostgreSQL та забезпечує високу пропускну спроможність.

Таблиця 2

Характеристика	RDBMS	NoSQL	NewSQL
Схема	Набір таблиць	Без схеми	Обидва підходи
SQL	Підтримується	Залежить від системи	Підтримується
ACID	Підтримується	Не підтримується	Підтримується
OLTP	Підтримується не повною мірою	Підтримується	Повна підтримка
Підтримка масштабування: вертикального / горизонтального	Так / Є реалізації для деяких СКБД (Oracle, SQL Server)	Так / Так	Так / Так
Можливість побудови розподілених систем	Так	Так	Так
Безпека	Висока	Низька	Висока
Архітектура зберігання даних	Дискове сховище / В пам'яті	Дискове сховище / В пам'яті	Дискове сховище / В пам'яті
Вплив збільшення розміру даних на продуктивність	Повільно	Швидко	Дуже швидко
Складність запитів	Низька	Висока	Висока
Робоче навантаження	OLTP / OLAP	OLTP	OLTP, OLAP, HTAP
Популярність (підтримка спільноти)	Величезна	Зростаюча	Зростаюча

Висновки

1. Поява NewSQL систем є відображенням об'єктивної потреби в масштабованих СКБД з підтримкою SQL та ACID-транзакцій, тобто таких систем, які прагнуть досягти продуктивності, порівнянної з NoSQL рішеннями, зберігаючи при цьому гарантії узгодженості даних. При цьому, безумовно, основною перевагою NewSQL систем є підтримка ACID-транзакцій, що уможливлює їх застосування там, де NoSQL рішення непридатні. Крім того, підтримка SQL дозволяє використовувати накопичений досвід та інструментарій для роботи з традиційними базами даних. Хоча при цьому слід враховувати, що створення схеми та оптимізація SQL запитів для виконання на кластері можуть ускладнити розробку застосунків.

2. NewSQL системи з новими архітектурами принципово відрізняються від традиційних SQL-орієнтованих СКБД початковою підтримкою розподіленої архітектури. Ці системи здатні добре масштабуватися горизонтально і забезпечувати високу продуктивність за певних типів транзакцій, що зачіпають невелику кількість вузлів. На таких транзакціях NewSQL системи наближаються за масштабованістю та продуктивністю до NoSQL рішень, але при цьому зберігаючи підтримку ACID та SQL. Крім того, NewSQL СКБД з новими архітектурами прагнуть підвищити продуктивність за рахунок оптимізації швидкості доступу до даних, у тому числі за рахунок зберігання бази даних повністю в оперативній пам'яті та ігнорування блокувань. При цьому, швидше за все, коли ширше використовуватимуться технології енергонезалежної пам'яті знадобляться подальші зміни в конструкції підсистем (в даний час це відносно нова галузь досліджень, що заслуговує на особливу увагу).

3. Використання програмного забезпечення проміжного рівня дозволяє прозоро розділяти дані між декількома екземплярами БД, знімаючи необхідність реалізації шардингу на стороні застосунку. Розробникам не потрібно вносити будь-які зміни до свого застосунку, щоб використовувати нову сегментовану базу даних. Для сумісності з конкретною СКБД проміжне програмне забезпечення має підтримувати відповідний протокол обміну. Продуктивність таких рішень загалом нижча, ніж, наприклад, систем з новими архітектурами, у тому числі через надмірність планування та оптимізації запитів до фрагментів даних, розміщених на окремих вузлах, для складних запитів. Хоча з іншого боку, це дозволяє кожному вузлу застосовувати власні локальні оптимізації для кожного запиту.

4. Певні NewSQL СКБД дозволяють проводити гібридну транзакційно-аналітичну обробку. Тобто дозволяють виконувати аналіз оперативних даних у реальному масштабі часу без традиційного поділу оперативної бази даних та сховища даних, уникаючи складнощів, пов'язаних із процесом ETL.

5. Через те, що найчастіше продуктивність є головним пріоритетом, вважається, що бази даних NewSQL мають більше вразливих місць у безпеці, ніж традиційні реляційні бази даних. Однак для більш обґрунтованого висновку все ж таки потрібні додаткові всебічні дослідження цієї проблеми.

6. Важливим висновком нашого аналізу вважатиметься те, що системи баз даних NewSQL не є принциповим відходом від існуючих принципів побудови БД. Вони швидше за все є наступним витком спіралі у розвитку технологій баз даних. Зокрема, існує низка рішень для масштабування класичних SQL-орієнтованих СКБД, що допомагають масштабувати існуючі застосунки без значних змін логіки роботи з даними. У більш довгостроковій перспективі можна вважати, що відбудеться конвергенція функцій у основних класах систем керування даними. Швидше за все, всі ключові системи будуть підтримувати ту або іншу форму реляційної моделі та SQL, а також операції OLTP та запити OLAP одночасно подібно до HTAP-систем.

Список літератури:

1. Abadi D., Ailamaki A., Andersen D., Bailis P., Balazinska M., Bernstein P., Boncz P., Chaudhuri S., et al // The Seattle Report on Database Research. ACM SIGMOD Record. 2020. 48. P. 44 – 53. https://doi.org/10.1145/3385658.3385668.
2. Gudivada V. N., Rao D., Raghavan V. V. Renaissance in database management: navigating the landscape of candidate systems // Computer. 2016. 49(4). P. 31–42. https://doi.org/10.1109/MC.2016.115.
3. Sadalage P. J., Fowler M. NoSQL Distilled A Brief Guide to the Emerging World of Polyglot Persistence. Addison-Wesley Professional, 2012. 188 p.
4. Using Oracle Sharding. Oracle Sharding Overview. URL: https://docs.oracle.com/en/database/oracle/oracle-database/19/shard/sharding-overview.html#GUID-0F39B1FB-DCF9-4C8A-A2EA-88705B90C5BF. (дата звернення: 17.02.2023).
5. Shute J., Vingralek R., Samwel B., Handy B., Whipkey C., Rollins E., Oancea M., Littlefield K., Menestrina D., Ellner S., Cieslewicz J., Rae I., Stancescu T., Apte H. F1: A distributed SQL database that scales // Proceedings of the 39th International Conference on Very Large Data Bases (VLDB) Endowment. 2013. 6(11). 1068 – 1079.
6. Pavlo A., Aslett M. What's really new with NewSQL? // ACM Sigmod Record. 2016. 45(2). P. 45 – 55. https://doi.org/10.1145/3003665.3003674.
7. NoSQL. URL: https://hostingdata.co.uk/nosql-database/ (дата звернення: 17.02.2023).
8. Corbett, J. C., Dean, J., Epstein, M., Fikes, A., Frost, C., Furman, J. J., Ghemawat S., Gubarev A., Heiser C., Hochschild P., Hsieh W., Kanthak S., Kogan E., Li H., Lloyd A., Melnik S., Mwaura D., Nagle D., Quinlan S., Rao R., Rolig L., Saito Y., Szymaniak M., Taylor C., Wang R., Woodford D. Spanner: Google's globally distributed database // ACM Transactions on Computer Systems (TOCS). 2013. 31(3). P. 1 – 22. https://doi.org/10.1145/2491245.
9. Özsu M. T., Valduriez P. Principles of Distributed Database Systems. Fourth Edition. Springer Cham, 2020. 674 p.
10. Aslett M. What we talk about when we talk about NewSQL. URL: http://blogs.the451group.com/information_management/2011/04/06/what-we-talk-about-when-we-talk-about-newsql/ (дата звернення: 17.02.2023).
11. Valduriez P., Jiménez-Peris R., Özsu M. T. Distributed database systems: The case for NewSQL // Hameurlain, A., Tjoa, A.M. (eds) Transactions on Large-Scale Data- and Knowledge-Centered Systems XLVIII. Lecture Notes in Computer Science. Vol 12670. Berlin, Heidelberg: Springer Berlin Heidelberg, 2021. P. 1 – 15. https://doi.org/10.1007/978-3-662-63519-3_1.
12. Moniruzzaman A. B. M. NewSQL: Towards next-generation scalable RDBMS for online transaction processing (OLTP) for Big Data management // International Journal of Database Theory and Application. 2014. 7(6) P. 121 – 130 http://dx.doi.org/10.14257/ijdta.2014.7.6.11.
13. Stonebraker M. The case for shared nothing // IEEE Database Eng. Bull. 1986. 9(1). P. 4 – 9.
14. Duggirala S. NewSQL databases and scalable in-memory analytics // Advances in Computers. Elsevier, 2018. 109. P. 49 – 76. https://doi.org/10.1016/bs.adcom.2018.01.004.
15. MariaDB Xpand. URL: https://mariadb.com/products/enterprise/xpand/ (дата звернення: 17.02.2023).
16. MariaDB. MariaDB Acquires Clustrix Adding Distributed Database Technology. URL: https://mariadb.com/newsroom/press-releases/mariadb-acquires-clustrix-adding-distributed-database-technology/ (дата звернення: 17.02.2023).

17. Namuag P. An Overview of MariaDB Xpand (formerly ClustrixDB). URL: https://severalnines.com/blog/overview-mariadb-xpand-formerly-clustrixdb/ (дата звернення: 17.02.2023).
18. Clustrix. URL: https://dbdb.io/db/clustrix (дата звернення: 17.02.2023).
19. MariaDB Xpand. URL: https://mariadb.com/docs/xpand/products/mariadb-xpand/ (дата звернення: 17.02.2023).
20. CockroachDB. URL: www.cockroachlabs.com (дата звернення: 17.02.2023).
21. Cloud Spanner. URL: https://cloud.google.com/spanner/ (дата звернення: 17.02.2023).
22. Bacon D. F., Bales N., Bruno N., Cooper B. F., Dickinson A., Fikes A., Fraser C., Gubarev A., Joshi M., Kogan E., Lloyd A., Melnik S., Rao R., Shue D., Taylor C., Holst M. H., Woodford D. Spanner: Becoming a SQL system // Proceedings of the 2017 ACM International Conference on Management of Data (SIGMOD '17). Association for Computing Machinery, New York, NY, USA, 2017. P. 331 – 343. https://doi.org/10.1145/3035918.3056103.
23. HyPer. URL: https://hyper-db.de/ (дата звернення: 17.02.2023).
24. SingleStore. URL: https://www.singlestore.com/ (дата звернення: 17.02.2023).
25. NuoDB. URL: https://www.nuodb.com. (дата звернення: 17.02.2023).
26. SAP HANA Cloud. URL: www.sap.com/products/hana.html (дата звернення: 17.02.2023).
27. Sikka V., Färber F., Lehner W., Cha S. K., Peh T., Bornhövd C. Efficient transaction processing in SAP HANA database: the end of a column store myth // Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data. 2012. P. 731 – 742).
28. H-Store. URL: https://hstore.cs.brown.edu/ (дата звернення: 17.02.2023).
29. Volt Active Data. URL: http://voltactivedata.com/ (дата звернення: 17.02.2023).
30. Simborg M. Introducing Volt Active Data. URL: https://www.voltactivedata.com/blog/2022/02/introducing-volt-active-data/ (дата звернення: 17.02.2023).
31. LeanXcale. URL: https://www.leanxcale.com/ (дата звернення: 17.02.2023).
32. Van Steen M., Tanenbaum A. S. Distributed systems. Third edition. Pearson Education, Inc. 2017. 596 p.
33. Gazis A., Katsiri E. Middleware 101: What to know now and for the future // Queue. 2022. 20(1). P. 10 – 23. https://doi.org/10.1145/3526211.
34. Harizopoulos S., Abadi D. J., Madden S., Stonebraker M. OLTP through the looking glass, and what we found there // Making Databases Work: the Pragmatic Wisdom of Michael Stonebraker. Association for Computing Machinery and Morgan & Claypool. 2018. P. 409–439. https://doi.org/10.1145/3226595.3226635.
35. MariaDB MaxScale. URL: https://mariadb.com/products/enterprise/components/#maxscale (дата звернення: 17.02.2023).
36. ScaleArc. URL: www.devgraph.com/scalearc (дата звернення: 17.02.2023).
37. Bernstein P. A., Cseri I., Dani N., Ellis N., Kalhan A., Kakivaya G., Lomet D. B., Manne R., Novik L., Talius T. Adapting Microsoft SQL server for cloud computing // 2011 IEEE 27th International Conference on Data Engineering, 2011. P. 1255 – 1263. https://doi.org/10.1109/ICDE.2011.5767935.
38. Amazon Aurora. https://aws.amazon.com/rds/aurora (дата звернення: 17.02.2023).
39. Connolly T. M., Begg C. E. Database systems: a practical approach to design, implementation, and management. Sixth edition. Harlow, Essex, England: Pearson Education Limited, 2015. 1329 p.
40. Kleppmann M. Designing data-intensive applications: The big ideas behind reliable, scalable, and maintainable systems. O'Reilly Media, Inc., 2017. 590 p.
41. Gray J., Reuter A. Transaction processing: concepts and techniques. Elsevier, 1992. 1070 p.
42. S. Kimball. Living without atomic clocks. URL: https://www.cockroachlabs.com/blog/living-without-atomic-clocks/ (дата звернення: 17.02.2023).
43. Spanner: TrueTime and external consistency. URL: https://cloud.google.com/spanner/docs/true-time-external-consistency (дата звернення: 17.02.2023).
44. TimeTools. What is the GPS Clock? URL: https://timetoolsltd.com/gps/what-is-the-gps-clock/ (дата звернення: 17.02.2023).
45. Kimball S., Sharif I. Living without atomic clocks. URL: https://www.cockroachlabs.com/blog/living-without-atomic-clocks/. (дата звернення: 17.02.2023).
46. Lamport L. The implementation of reliable distributed multiprocess systems. Computer Networks (1976). 1978. 2(2). P. 95–114. https://doi.org/10.1016/0376-5075(78)90045-4.
47. The Tandem Database Group. NonStop SQL: A distributed, high-performance, high-availability implementation of SQL // Gawlick, D., Haynie, M., Reuter, A. (eds) High Performance Transaction Systems. HPTS 1987. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg. 2005. Vol. 359. P. 60 – 104. https://doi.org/10.1007/3-540-51085-0_43.
48. Malviya N., Weisberg A., Madden S., Stonebraker M. Rethinking main memory OLTP recovery. In 2014 IEEE 30th International Conference on Data Engineering, Chicago, IL, USA, IEEE. 2014. P. 604 – 615, https://doi.org/10.1109/ICDE.2014.6816685.
49. Schwartz B., Zaitsev P., Tkachenko V. High performance MySQL: optimization, backups, and replication. Third Edition. O'Reilly Media, Inc., 2012. 826 p.
50. Harrington J. L. Relational database design and implementation. 4th edition. Morgan Kaufmann, 2016. 712 p.

51.Navathe S., Ceri S., Wiederhold G., Dou J. Vertical partitioning algorithms for database design // ACM Transactions on Database Systems (TODS). 1984. 9(4). P. 680–710. https://doi.org/10.1145/1994.2209.

52. Curino C., Jones E., Zhang Y., Madden S. Schism: a workload-driven approach to database replication and partitioning // Proceedings of the VLDB Endowment. 2010. 3(1-2). P. 48 – 57. https://doi.org/10.14778/1920841.1920853.

53. Law Insider. Legal Definitions Dictionary. URL: https://www.lawinsider.com/dictionary/transaction-engine (дата звернення: 17.02.2023).

54. Elmore A. J., Arora V., Taft R., Pavlo A., Agrawal D., Abbadi A. E. Squall: Fine-grained live reconfiguration for partitioned main memory databases // Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. 2015. P. 299–313. https://doi.org/10.1145/2723372.2723726.

55. Serafini M., Mansour E., Aboulnaga A., Salem K., Rafiq T., Minhas U. F. Accordion: Elastic scalability for database systems supporting distributed transactions // Proceedings of the VLDB Endowment. 2014. 7(12). P. 1035 –1046. https://doi.org/10.14778/2732977.2732979.

56. Ismail W., Muhammed A., Abdullah Z. H., Radman A., Hendradi R., Afandi R. R. A Survey of NewSQL DBMSs focusing on Taxonomy, Comparison and Open Issues. Journal of Computer Science & Computational Mathematics. 2021. P. 87-95. https://doi.org/10.20967/jcscm.2021.04.002.

57. Rothnie J. B., Bernstein P. A., Fox S., Goodman N., Hammer M., Landers T. A., Reeve C., Shipman D. W., Wong E. Introduction to a system for distributed databases (SDD-1) // ACM Transactions on Database Systems (TODS). 1980. 5(1), P. 1–17. https://doi.org/10.1145/320128.320129.

58. Williams R., Daniels D., Haas L., Lapis G., Lindsay B. G., Ng, P., Obermarck R., Selinger P., Walker A., Wilms P., Yost R. R*: An overview of the architecture // IBM Thomas J. Watson Research Division. 1981. P. 329 – 347.

59. Epstein R., Stonebraker M., Wong E. Distributed query processing in a relational data base system // Proceedings of the 1978 ACM SIGMOD international conference on management of data. (SIGMOD '78). Association for Computing Machinery, New York, NY, USA. 1978. P. 169 – 180. https://doi.org/10.1145/509252.509292.

60. DeWitt D. J., Katz R. H., Olken F., Shapiro L. D., Stonebraker M. R., Wood D. A. Implementation techniques for main memory database systems // Proceedings of the 1984 ACM SIGMOD international conference on management of data. 1984. P. 1 – 8. https://doi.org/10.1145/602259.602261.

61. Faerber F., Kemper A., Larson,P. A., Levandoski J., Neumann T., Pavlo A. Main memory database systems // Foundations and Trends in Databases. 2017. 8(1 – 2), P. 1 – 130.

62. Gawlick D., Kinkade D. Varieties of concurrency control in IMS/VS fast path // IEEE Database Eng. Bull. 1985. 8(2). P. 3 – 10.

63. Lehman T. J., Carey M. J. A Study of Index Structures for Main Memory Database Management Systems // Proceedings of the Twelfth International Conference on Very Large Data Bases, VLDB. 1985. P. 294 – 303.

64. Lehman T. J., Carey M. J. Query processing in main memory database management systems // Proceedings of the 1986 ACM SIGMOD international conference on Management of data. 1986. P. 239 – 250. https://doi.org/10.1145/16894.16878.

65. Lehman T. J., Carey M. J. A recovery algorithm for a high-performance memory-resident database system // ACM SIGMOD Record. 1987. 16(3). P. 104 – 117. https://doi.org/10.1145/38714.38730.

66. Kersten M.L., Apers P.M.G., Houtsma M.A.W., van KuykE.J.A., van de Weg R.L.W. A Distributed, Main-Memory Database Machine // Kitsuregawa M., Tanaka H. (eds) Database Machines and Knowledge Base Machines. The Kluwer International Series in Engineering and Computer Science. Springer, Boston, MA. 1988. Vol. 43. P. 353 – 369. https://doi.org/10.1007/978-1-4613-1679-4_26.

67. Altibase. URL: https://www.altibase.com (дата звернення: 17.02.2023).

68. TimesTen: Fastest OLTP Database, Ultra High Availability, Elastic Scalability. URL: https://www.oracle.com/database/technologies/related/timesten.html (дата звернення: 17.02.2023).

69. Cambridge Dictionary. URL: https://dictionary.cambridge.org/dictionary/english/workload. (дата звернення: 17.02.2023).

70. HPE Nimble Storage Deployment Considerations for Microsoft SQL Server. OLTP Workloads. URL: https://infosight.hpe.com/InfoSight/media/cms/active/public/tmg_HPE_Nimble_Storage_Deployment_Considerations_for_Microsoft_SQL_Server_doc_version_family.whz/xpm1491839725334.html (дата звернення: 17.02.2023).

71. DeBrabant J., Pavlo A., Tu S., Stonebraker M., Zdonik S. Anti-caching: A new approach to database management system architecture // Proceedings of the VLDB Endowment. 2013. 6(14). P. 1942 – 1953. https://doi.org/10.14778/2556549.2556575.

72. Arulraj J., Pavlo A., Dulloor S. R. Let's talk about storage & recovery methods for non-volatile memory database systems // Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15). Association for Computing Machinery, New York, NY, USA. 2015. P. 707 – 722. https://doi.org/10.1145/2723372.2749441.

73. Andrei M., Lemke C., Radestock G., Schulze R., Thiel C., Blanco R., Meghlan A., Sharique M., Seifert S., Vishnoi S., Booss D., Peh T., Schreter I., Thesing W., Wagle M., Willhalm T. SAP HANA adoption of non-volatile memory // Proceedings of the VLDB Endowment. 2017. 10(12). P. 1754 – 1765. https://doi.org/10.14778/3137765.3137780.

74. Intel Optane Persistent Memory and SAP HANA Platform Configuration. Technology overview and deployment guidelines for using Intel Optane persistent memory with SAP HANA. Configuration Guide. 2019. URL: https://cdrdv2-public.intel.com/753738/sap-hana-and-intel-optane-configuration-guide.pdf (дата звернення: 17.02.2023).

75. Intel Optane Persistent Memory. URL: https://www.intel.com/content/www/us/en/products/docs/memory-storage/optane-persistent-memory/overview.html (дата звернення: 17.02.2023).

76. Kroenke D. M., Auer D. J., Yoder R. C., Vandenberg S. L. Database processing fundamentals, design, and implementation. 15th edition. Pearson. 2018. 648 p.

77. Taft R., Sharif I., Matei A., VanBenschoten N., Lewis, J., Grieger, T., Niemi K., Woods A., Birzin A., Poss R., Bardea P., Ranade A., Darnell B., Gruneir B., Jaffray J., Zhang L., Mattis P. Cockroachdb: The resilient geo-distributed SQL database // Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data (SIGMOD '20). Association for Computing Machinery, New York, NY, USA, 2020. P. 1493 – 1509. https://doi.org/10.1145/3318464.3386134.

78. Garcia-Molina H., Salem K. Main memory database systems: An overview // IEEE Transactions on knowledge and data engineering. 1992. 4(6). P. 509–516. https://doi.org/10.1109/69.180602.

79. Google Cloud Spanner. Secondary indexes. URL: https://cloud.google.com/spanner/docs/secondary-indexes (дата звернення: 17.02.2023).

80. SAP HANA Performance Guide for Developers. Secondary Indexes. URL: https://help.sap.com/docs/SAP_HANA_PLATFORM/9de0171a6027400bb3b9bee385222eff/3441acf7dcf64e169ba94121acaf2350.html?version=2.0.04&locale=en-US (дата звернення: 17.02.2023).

81. Fitzpatrick B. Distributed caching with Memcached // Linux journal. 2004. 2004(124). P. 5.

82. Mohan C., Haderle D., Lindsay B., Pirahesh H., Schwarz P. ARIES: A transaction recovery method supporting fine-granularity locking and partial rollbacks using write-ahead logging // ACM Transactions on Database Systems (TODS). 1992. 17(1). P. 94 – 162. https://doi.org/10.1145/128765.128770.

83. Using VoltDB. V.12.1. URL: https://docs.voltdb.com/UsingVoltDB/ (дата звернення: 17.02.2023).

84. Ren K., Diamond T., Abadi D. J., Thomson A. Low-overhead asynchronous checkpointing in main-memory database systems // Proceedings of the 2016 International Conference on Management of Data (SIGMOD '16). Association for Computing Machinery, New York, NY, USA. 2016. P. 1539 – 1551. https://doi.org/10.1145/2882903.2915966.

85. Reid R. Practical CockroachDB: Building Fault-Tolerant Distributed SQL Databases. 1st ed. Apress. 2022. 254 p.

86. MariaDB MaxScale technical brief. Enterprise security. URL: https://mariadb.com/wp-content/uploads/2019/09/mariadb-maxscale-security_datasheet_1041.pdf (дата звернення: 17.02.2023).

87. Khasawneh T. N., AL-Sahlee M. H., Safia A. A. SQL, NewSQL, and NoSQL databases: A comparative survey // 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, IEEE. 2020. P. 013 – 021. https://doi.org/10.1109/ICICS49469.2020.239513.

Надійшла до редколегії 20.11.2022

Відомості про авторів:

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

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

УДК 003.026:004.056 **DOI:10.30837/rt.2023.4.211.04**

Д.В. ГАРМАШ

АНАЛІЗ ПІДПИСУ FALCON В ПОРІВНЯННІ З ІНШИМИ ПІДПИСАМИ. ФРЕЙМВОРКИ GPV ТА РАБІНА

Вступ

Квантовий комп'ютер може зруйнувати більшість, якщо не всі традиційні криптосистеми, що використовуються на практиці, а саме – всі системи на основі задачі факторизації цілих чисел (наприклад RSA) або завдання дискретного логарифмування (як традиційних, так і на еліптичних кривих Діффі – Хеллмана і DSA; а також всю криптографію, засновану на спаровуваннях). Деякі класичні криптосхеми, що базуються на обчислювально-складних завданнях, сильно відрізняються від зазначених вище і їх набагато складніше вирішити, вони залишаються незалежними від квантових обчислень. У даній роботі проведено огляд алгоритму Falcon.

1. Порівняння FALCON та схеми CRYSTALS-DILITHIUM

1.1. Crystals-Dilithium

Dilithium – схема підпису, яка наслідує концепцію створення схеми підпису зі схеми ідентифікації, використовуючи Fiat-Shamir з перериванням. Безпека даної схеми може бути скорочена до проблем безпеки Модульного-навчання з помилками (MLWE) і Модульного короткого цілочисельного рішення (MSIS). Вона створюється з метою дозволу швидким множенням використовувати NTT перетворення і уникати появи випадковості з дискретного поширення Gaussian замість вибору зразка з однорідного поширення.

Безпека Dilithium заснована на основоположних проблемах MLWE та MSIS. На даний момент не існує жодної ефективної атаки, що використовує модульну структуру і розглядається в якості безпеки, еквівалентної проблемам MLWE та MSIS.

На перевагу до інших пропозицій щодо підпису Dilithium підбирає з одноманітного поширення, уникаючи складний та неефективний відбір з дискретного поширення Gaussian. Модульна структура Dilithium забезпечує, що поліноміальне множення завжди виконується у тому ж самому кільці незалежно від рівня безпеки, який робить її легким для перемикання між рівнями. Множення може бути виконано ефективно через власні «дружні» параметри NTT. Використовуючи фокус для стискання відкритого ключа з фактором 2, Dilithium має найменший відкритий ключ плюс розмір підпису схеми на основі решіток, що використовують одноманітний відбір [3].

1.2. Falcon

Falcon – схема підпису, чий дизайн заснований на базі фреймворку Gentry-Peikert-Vaikuntanathan (GPV) для підписів на основі решіток, що використовують приховані функції. Він створює екземпляри даного тлумачення за допомогою решіток СКПН і ефективного зразку Gaussian, який створює схему, що є доказово безпечною на основі припущення, що SIS є складним, особливо у використаних решітках. Falcon був створений таким чином, що усі арифметичні операції можуть бути обчислені, використовуючи ефективні техніки Fourier-перетворення.

Він не вимагає (але може використовувати) одиницю з плаваючою крапкою і працювати ефективно на базі мікропроцесорів різного виду, включаючи Intel x86 і ARM cores. Постійний у часі шаблон Gaussian може бути використаний у Falcon.

Математична безпека Falcon покладається на твердість проблеми SIS над кільцями СКПН, яка виграє над довгою історією криптоаналізу для криптосистеми СКПН. Найбільш відомі атаки – загальні техніки решіток: не існує поширеного засобу для використання додаткової кільцевої структури, представленої у решітках СКПН. Для оцінювання безпеки проти

алгоритмів скорочення решіток, Falcon вживлює метод «Core-SVP», який також використовувався багатьма іншими представленнями NTT на основі решіток [2].

Коротко, Falcon – дуже компактний (найменший комбінований розмір відкритого ключа і підпису серед усіх кандидатів НІСТ) і ефективна схема пост-квантового підпису, чия безпека скорочується до добре оцінених припущень. Обрана кільцева структура і помилкове поширення дозволяються для ефективних реалізацій на основі FFT, які частково відміняють несприятливий вплив виконання помилкового поширення Gaussian і призводять до задовільного представлення на практиці. Насправді, можливо найбільшим недоліком Falcon здається складність розуміння усіх деталей тлумачення і реалізації схеми правильно.

Так само, як і їхні фізичні аналоги, цифрові підписи призначені для підтвердження того, що документ видано чи схвалено коректним чином. Разом зі схемами шифрування вони відіграють важливу роль у безпеці електронних комунікацій. Оскільки в цифровому світі все можна відтворити, цифрові підписи не можуть використовувати ті ж принципи, що й фізичні; натомість вони покладаються на складність математичних проблем.

Підписувач зберігає при собі закритий ключ, який він використовує кожного разу, коли обчислює підпис. З цим закритим ключем пов'язаний відкритий ключ, який він може публічно надіслати будь-кому. Кожного разу, коли підписувачу потрібно підписати повідомлення, він використовує свій особистий ключ, щоб вирішити якусь математичну складну задачу, яка залежить лише від повідомлення та відкритого ключа; рішенням буде підпис. З іншого боку, верифікатор генерує ту саму проблему (оскільки вона залежить лише від відкритих елементів) і використовує свій відкритий ключ, щоб перевірити, що підпис справді є вирішенням проблеми. Однак відкритий ключ не допомагає верифікатору самостійно вирішити проблему, якщо все це може здатися трохи абстрактним [3].

2. Схема Рабіна: приклад на основі факторизації

Відомо, що проблема розкладання на множники є складною: задано два дуже великі цілі числа p і q (скажімо, 1000 цифр кожне), комп'ютер може обчислити їхній результат $N = p \times q$ моментально, але відновлення (p, q) із заданим N недоступне для сучасних комп'ютерів. Є проблеми, які важко розв'язати, знаючи лише N, але вони вирішуються, враховуючи його розкладання на множники (p, q). Розглянемо, наприклад, задачу обчислення квадратного кореня: задане ціле число y, ми хочемо знайти таке ціле число $x^2 = y \bmod N$. Якщо ми знаємо лише N, це складна задача на класичному комп'ютері (принаймні, немає відомого ефективного методу її розв'язання), але перевірити, чи є x правильним рішенням, легко: просто перевірити $x^2 = y \bmod N$. Однак, якщо ми знаємо розкладання $N = p \times q$, то ця задача легко вирішується за допомогою цього алгоритму:

1) Обчислити квадратний корінь з y за модулем p і q. Є багато способів зробити це, наприклад алгоритм Тонеллі – Шенкса.

2) Скористатися китайською теоремою про залишки, щоб поєднати ці квадратні корені за модулями p і q у квадратний корінь за модулем N [4].

У цьому прикладі показано, які виникають проблеми (обчислення квадратного кореня):

1) Легко перевірити, важко вирішити за допомогою відкритого ключа N.

2) Легко вирішити за допомогою закритого ключа (p, q).

Криптографія з відкритим ключем використовує переваги асиметрії між тим, що можна досягти відкритим і закритим ключами. Наприклад, схема підпису Рабіна базується на конкретній проблемі, викладеній вище. Ця схема працює наступним чином:

1) Підписувач підписує повідомлення, спочатку надсилаючи його випадковій цілі y (використовуючи геш-функцію, тип функції, яка надсилає вхідні дані до випадкових на вигляд виходів). Потім він використовує свій особистий ключ (p, q), щоб обчислити квадратний корінь з $y \bmod N$: це рішення x слугуватиме підписом повідомлення.

2) Верифікатор використовує відкритий ключ N, щоб переконатися, що x є дійсним підписом повідомлення, перевіривши, що $x^2 = H(msg) \bmod N$.

Цікаво, що можна показати, що обчислення квадратного кореня та розкладання на множники є еквівалентними проблемами: якщо обчислення квадратного кореня за модулем N є складним, то складним є і розкладання N на множники. Схема Рабіна дотримується парадигми гешування, потім підписування: повідомлення спочатку гешується до цільового виклику, а рішенням для цього виклику є підпис. Falcon дотримується тієї ж парадигми, але замість цілочисельної факторизації використовує решітку [6].

2.1. Як влаштовані алгебраїчні решітки

Схема Рабіна не є постквантовою. Дійсно, її основну проблему, розкладання на множники, можна швидко вирішити за допомогою великомасштабного квантового комп'ютера. Однак його ідеї високого рівня можна адаптувати для роботи над проблемами решіток, які, як припускають, протистоять квантовим зловмисникам.

По суті, решітка – це нескінченна кількість точок, розташованих у вигляді сітки. Наприклад, на малюнку нижче зображена двовимірна решітка. Загалом, решітка може існувати в будь-якій додатній кількості вимірів.

Решітка має нескінченну кількість точок, що, звичайно, викликає питання практичності: чи потрібно зберігати нескінченну кількість точок? Звичайно, відповідь – ні, ми можемо бути ефективнішими. Першим кроком до практичності є робота лише з q-ірними ґратками; це решітки, координати точок яких є цілими і «обертають» деяке ціле число q, тобто, якщо ми зменшуємо за модулем q координати точки решітки, результатом все одно буде точка решітки [5].

2.2. Як відбувається створення цифрового підпису Falcone

Загальна інформація: фреймворк GPV.

Falcon слідує структурі, представленій у 2008 р. Гентрі, Пейкертом і Вайкунтатаном, яку скорочено називають фреймворком GPV. Деталі їх роботи можуть бути досить технічними, але ідея високого рівня полягає в наступному:

1) Відкритий ключ є довгою основою q-ї решітки.

2) Приватний ключ є (по суті) короткою основою тієї ж решітки.

3) Під час процедури підписання підписувач:

- генерує випадкове значення v;
- обчислює ціль $c = H(m||v)$, де H – геш-функція, яка надсилає вхідні дані до випадкової точки (на сітці), m – повідомлення;
- використовує свої знання про короткий базис для обчислення точки решітки v поблизу цілі c;
- на виході отримує (m, s), де $s=c-v$.

4) Верифікатор приймає підпис (m, s) в такому випадку, якщо вектор v короткий, та $H(m||v)-s$ є точкою на решітці, згенерованою його відкритим ключем [7].

2.3. NTRU решітки

Першим кроком для створення екземпляра GPV-структури є вибір класу криптографічно жорстких решіток: повинна бути можливість побудувати короткий і довгий базис для тієї самої решітки, щоб будь-кому з довгим базисом було важко знайти близькі вектори з такою ж точністю, як із коротким базисом. Для дидактичних цілей у прикладі використовується решітка розмірності $n = 2$. Однак на практиці ця розмірність недостатня для забезпечення безпеки: у низьких розмірностях алгоритми зменшення решітки, такі як LLL, можуть швидко відновити короткий базис з довгого базису. Подібно до того, як для захисту від класичних комп'ютерів RSA вимагає чисел у кілька тисяч біт, для захисту від класичних і квантових комп'ютерів криптосистеми на основі решітки зазвичай потребують розмірів у порядку величини $n = 1024$.

Зберігання баз таких великих розмірів може бути дорогим: кінцевий відкритий ключ легко може бути більшим за мегабайт. Щоб уникнути цієї проблеми, типово працювати зі структурованими решітками, де цілий базис можна отримати обертанням коефіцієнтів кількох початкових базисних векторів. Це значно зменшує розмір бази для зберігання. Falcon використовує решітки NTRU, які є класом таких структурованих решіток. Їх використання дозволяє зменшити розмір відкритого ключа до менш ніж 1,8 кілобайт. З моменту створення більше ніж 20 років решітки NTRU успішно витримали ретельну перевірку [8].

2.4. Швидка вибірка Фур'є

Другим кроком є вибір алгоритму для обчислення векторів тісної решітки на кроці 3.3 схеми підпису. Хоча алгоритм округлення Бабаї є дуже ефективним, відомо, що його не слід використовувати тут. Дійсно, результат алгоритму Бабаї завжди є паралелепіпедом, що має форму використовуваного базису, тому його використання призведе до повільного витоку закритого ключа. Натомість у своїй структурі Гентрі та його співавтори рекомендують працювати з модифікацією алгоритму Бабаї, де:

- кожен коефіцієнт округлюється випадковим чином. Це гарантує відсутність витоку інформації про закритий ключ;
- кожне округлення враховує попередні. Це дозволяє відбирати ближчі вектори, ніж із простим округленням.

Отриманий алгоритм, який часто називають семплером GPV, безпечніший і кращий, ніж алгоритм округлення Бабаї. Як додаткове вдосконалення, Falcon використовує структуру решіток NTRU, щоб зробити семплер GPV швидшим на два порядки. Falcon, який називається швидкою дискретизацією Фур'є, можна розглядати як гібрид між дискретизатором GPV і швидким перетворенням Фур'є, яке широко використовується в обробці сигналів. Нижче можна побачити результати роботи двох алгоритмів Falcone (512 та 1064 біт) [7]:

VARIANT	KEYGEN/S	SIGN/S	VERIFY/S	IPKI (BYTES)	ISIGI (BYTES)
Falcon-512	143	6081	37175	897	618
Falcon-1024	50	3075	17697	1793	1234

Висновки

1. Схеми цифрового підпису на решітках є основними претендентами на перемогу в конкурсі NIST PQC. Тому, їх подальший детальний аналіз та порівняння щодо основних характеристик стійкості є першочерговою задачею. Схема FALCON, як фіналіст другого етапу, потребує особливої уваги, оскільки має нетиповий дизайн, що використовує арифметику з плаваючою крапкою.

2. Одним із основних завдань конкурсу NIST США є розробка та прийняття постквантового чи постквантових стандартів ЕП. Зараз фаворіти – CRYSTALS-DILITHIUM та FALCON. Причому, подальше вирішення проблеми безпеки, тобто доведення криптографічної стійкості двох кандидатів-фіналістів, на стандарт ЕП FALCON, може ґрунтуватись на проблемах теорії та практики алгебраїчних решіток.

3. Можна дійти висновків, що Falcon використовує схему високого рівня (систему GPV) із двома компонентами (решітки NTRU та швидка вибірка Фур'є). Цілочисельний модуль завжди однаковий, q = 12289. Відкритий ключ – це один вектор із n цілих чисел від 0 до q–1, аналогічно для кожного підпису. Falcon доступний у двох варіантах, для n=512 або n=1024. Вони націлені на високу ефективність і високий рівень безпеки відповідно.

Список літератури:

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

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

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

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

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

6. McGrew D., Curcio M. Hash-Based Signatures draft-mcgrew-hash-sigs00[Електронний ресурс] / D. McGrew, M. Curcio. Режим доступу: https://tools.ietf.org/html/draftmcgrew-hash-sigs-00 .

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

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

Надійшла до редколегії 03.12.2022

Відомості про автора:

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

RADIO PHYSICS
РАДІОФІЗИКА

УДК 621.372(075.8) **DOI:10.30837/rt.2022.4.211.05**

О.В. ЛАЗОРЕНКО, д-р фіз.-мат. наук, А.А. ОНИЩЕНКО,
Л.Ф. ЧОРНОГОР, д-р фіз.-мат. наук

МУЛЬТИФРАКТАЛЬНИЙ АНАЛІЗ МОДЕЛЬНИХ ФРАКТАЛЬНИХ І МУЛЬТИФРАКТАЛЬНИХ СИГНАЛІВ

Вступ

Одним із актуальних напрямків сучасної фрактальної радіофізики [1] є мультифрактальний аналіз сигналів і процесів різного походження. На сьогодні створено багато різноманітних методів мультифрактального аналізу, що дозволяють досить ефективно отримувати інформацію про фрактальні та мультифрактальні властивості досліджуваних об'єктів. У переважній більшості робіт (див., наприклад, [2 – 4]) основний наголос, на жаль, робиться на особливості алгоритму самого методу, а не на трактовку отриманих числових мультифрактальних характеристик. Але для фахівця-практика вкрай потрібно отримати саме прості та зрозумілі поради щодо коректної інтерпретації отриманих результатів проведених досліджень. Саме тому тематика даної роботи виглядає актуальною.

Метою роботи є дослідження особливостей мультифрактальних характеристик модельних фрактальних і мультифрактальних сигналів, отриманих із використанням двох найбільш відомих методів мультифрактального аналізу – методу Wavelet Transform Modulus Maxima (WTMM) і методу Multi-Fractal Detrended Fluctuation Analysis (MF DFA).

Методи мультифрактального аналізу та відповідні числові характеристики

Метод WTMM як метод мультифрактального аналізу сигналів і процесів був запропонований у 1988 р. А. Арнеодо (A. Arneodo), Г. Грассо (G. Grasseau) та М. Холлшнайдером (M. Hollschneider) в роботі [5]. Він отримав подальший розвиток у роботах С. Малла [6] та А. Арнеодо, Е. Бакрі та Дж. Мьюзі [7 – 9]. Даний метод ґрунтується на використанні двох інтегральних перетворень з апарату вейвлет-аналізу – безперервного вейвлет-перетворення (БВП) та аналітичного вейвлет-перетворення (АВП) для отримання функції мультифрактального спектру $f(\alpha)$.

На нашу думку, найбільш вдало основи метода WTMM викладено в роботах [4, 6].

У межах даної роботи метод WTMM використовується для всього аналізованого сигналу в цілому. До мультифрактальних характеристик, що традиційно дозволяє отримати даний метод, належать мінімальне ($\alpha_{\min}$) та максимальне ($\alpha_{\max}$) значення показника Гьольдера α для функції мультифрактального спектру $f(\alpha)$, її ширина $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$ та значення узагальненого показника Херста α^*, що відповідає положенню єдиного максимуму функції $f(\alpha)$ [10].

Будучи узагальненням відомого метода монофрактального аналізу (метода DFA – Detrended Fluctuation Analysis), метод MF DFA (multi-fractal DFA) з'явився в 2002 р. в роботі Я. Кантельхардта (J. Kantelhardt) та ін. [2].

Для практиків, що використовують системи комп'ютерної математики (СКМ) MATLAB/SciLab, є корисною відмінна оглядова стаття [11]. На сьогодні метод MF DFA популярний серед дослідників (див., наприклад, [11 – 14]).

На відміну від методу WTMM, який застосовується до всього сигналу в цілому, а тому дає глобальні значення мультифрактальних характеристик, у алгоритмі методу MF DFA використовується ковзаюча віконна функція у часовій області скінченної ширини, через що з'являється можливість отримати локальні (у межах вікна) значення мультифрактальних

характеристик. Зазвичай отримані величини асоціюються з положенням середини вікна на часовій осі. Отже, всі наведені вище мультифрактальні характеристики стають функціями часу: $\alpha_{\min}(t)$, $\alpha_{\max}(t)$, $\Delta\alpha(t)$ і $\alpha^*(t)$ [11].

Останнє виявляється важливим у випадку дослідження нестаціонарних (у сенсі мультифрактальних властивостей) сигналів і процесів. Більш того, порівняння локальних і глобальних значень згаданих характеристик теж є цікавим і корисним.

Досить великий практичний досвід роботи авторів з такими сигналами та процесами викликав необхідність розширити перелік мультифрактальних характеристик, додавши декілька нових.

Часто вважають, що в ідеальному випадку функція мультифрактального спектру $f(\alpha)$ добре апроксимується параболою, і ця парабола є симетричною відносно вертикальної прямої, що проходить через її точку максимуму. Але практичний досвід підказує, що реальні мультифрактальні спектри часто виявляються несиметричними. Тому для урахування цього факту слід ввести додаткову числову характеристику.

Цю характеристику будемо називати *коефіцієнтом асиметрії функції мультифрактального спектру* та визначимо його через інші параметри:

$$K_f = \ln\frac{\alpha_{\max} - \alpha^*}{\alpha^* - \alpha_{\min}}. \qquad (1)$$

Аналізуючи співвідношення (1), легко побачити, що для симетричної функції мультифрактального спектру $f(\alpha)$, коли $\alpha^* = (\alpha_{\min} + \alpha_{\max})/2$, маємо $K_f = 0$. Якщо максимум асиметрично зсунуто праворуч від симетричного положення ($\alpha^* > (\alpha_{\min} + \alpha_{\max})/2$), спостерігається $K_f < 0$, а якщо ліворуч ($\alpha^* < (\alpha_{\min} + \alpha_{\max})/2$), то $K_f > 0$. Важливо також, що коли мультифрактал у граничному випадку перетворюється на монофрактал з показником Гьольдера α_0, для якого $\alpha_{\min} = \alpha_0 + \varepsilon$, $\alpha_{\max} = \alpha_0 - \varepsilon$, $\alpha^* = \alpha_0$, $\varepsilon \to +0$, маємо

$$\lim_{\varepsilon\to+0} K_f(\varepsilon) = \lim_{\varepsilon\to+0} \ln\frac{\alpha_{\max} - \alpha^*}{\alpha^* - \alpha_{\min}} = \lim_{\varepsilon\to+0} \ln\frac{\alpha_0 - \varepsilon - \alpha_0}{\alpha_0 - \alpha_0 - \varepsilon} = \ln 1 = 0.$$

Друга корисна числова характеристика, що ми пропонуємо використовувати для мультифрактального аналізу сигналів і процесів, є аналогом показника широкосмуговості, який застосовується для надширокосмугових (НШС) сигналів [15]. Вона задається співвідношенням

$$\mu_\alpha = \frac{\Delta\alpha}{\alpha^*}$$

та має назву *показник відносної ширини мультифрактального спектру*.

Показник відносної ширини мультифрактального спектру сигналу μ_α для строго монофрактального сигнала перетворюється на 0. Для інших видів сигналів, у тому числі й мультифрактальних, він є додатним. Така числова характеристика дозволить оцінювати, наскільки досліджуваний сигнал або процес є близьким до монофрактального.

Ще одне запропоноване нами доповнення набору традиційних числових характеристик, що використовуються під час проведення мультифрактального аналізу сигналів і процесів, стосується *розмірності* так званого *носія мультифракталу* (multi-fractal support).

Добре відомо (див., наприклад, [6]) що у випадку сигналу або процесу носієм мультифрактала є гладка крива, фрактальна розмірність D_F якої завжди дорівнює одиниці. Саме тому точка максимуму функції мультифрактального спектра $\left(\alpha^*, f(\alpha^*)\right)$ майже завжди має

$f(\alpha^*)=1$. Але у випадку монофракталу з показником Гьольдера α_0, як було вже вказано вище, увесь мультифрактальний спектр колапсує у точку (α_0,α_0). І оскільки для монофрактала $0<\alpha_0<1$, то й ордината точки положення максимуму зколапсованого спектра виявляється меншою за одиницю. У переважній більшості випадків це не відіграє суттєвої ролі, але при дослідженні нестаціонарних сигналів і процесів це зовсім не так.

Якщо сигнал або процес $s(t)$ є нестаціонарним у сенсі зміни у часі його фрактальних властивостей, то аналіз числових характеристик мультифрактального спектра треба проводити у часовому вікні скінченної ширини T (див., наприклад, [11]). Положення цього вікна на часовій осі відносно досліджуваного сигналу під час обчислення мультифрактальних характеристик перетворює ці характеристики на функції часу. Так саме стається і з ординатою точки максимуму функції мультифрактального спектра, тобто вона стає функцією часу $f(\alpha^*)\equiv f_\alpha = f_\alpha(t)$.

Важливо, що на графіку $f_\alpha(t)$ провали нижче рівня $f(\alpha^*)=1$ можуть сигналізувати про те, що мультифрактальний процес переходить у монофрактальний режим. Побачити це одночасно на трьох залежностях $\alpha_{\min}(t)$, $\alpha_{\max}(t)$ і $\Delta\alpha(t)$ значно складніше та не зовсім зручно.

Таким чином, для кожного сигналу, що досліджуватиметься нижче із застосуванням методу MF DFA, ми будемо використовувати часові залежності $\alpha_{\min}(t)$, $\alpha_{\max}(t)$, $\Delta\alpha(t)$, $K_f(t)$, $\mu_\alpha(t)$ і $f_\alpha(t)$, наводячи результати аналізу у спеціально створеному зручному форматі.

Модельні фрактальні та мультифрактальні сигнали

Для проведення досліджень, про які йдеться нижче, нами було створено великий набір з двох десятків модельних фрактальних (точніше кажучи, монофрактальних) сигналів (ФС) і мультифрактальних сигналів (МФС), до того ж як детермінованих, так і стохастичних. Але через обмеженість обсягу статті наведемо тільки п’ять моделей детермінованих ФС і МФС.

Модель 1. Дана модель є моделлю косинусної функції Вейєрштраса – Мандельброта [16]

$$s_1(t)=\sum_{n=-\infty}^{+\infty}\frac{1-\cos(\lambda^n t)}{\lambda^{(2-D)n}},$$

з фрактальною розмірністю $D=1.5$, $\lambda>1$. Вона є строго однорідною, а її графік є самоафінним [17].

Модель 2. Модель зі стрибком фрактальної розмірності. Цей модельний сигнал створено на основі з двох послідовно розташованих у часовій області модельних фрактальних НШС (ФНШС) сигналів $s_0(t)$ однакової довжини, перший з яких має фрактальну розмірність $D=1.8$, а другий – $D=1.2$. Сама же модель $s_0(t)$ є моделлю детермінованого фрактального НШС (ФНШС) сигналу, що базується на узагальненій функції Вейєрштраса [3], в якій всі випадкові фази дорівнюють нулю:

$$s_1(t)=\left[1-b^{2D-4}\right]\frac{\sum_{n=0}^{M} b^{(D-2)n}\cos\left(2\pi s b^n t\right)}{1-b^{(2D-4)(M+1)}},$$

де t – часова змінна, b – параметр масштабування за часом, D – фрактальна розмірність сигналу, $1<D<2$, M – кількість гармонік, які використовуються для побудови фізичного фракталу (якщо $M\to\infty$, то ми отримуємо математичний фрактал).

Модель 3. Модель із лінійно зменшуваною фрактальною розмірністю. Це модель ФС на основі узагальненої функції Вейєрштраса, у якої фрактальна розмірність $D(t)$ із часом зменшується за лінійним законом від 1.8 до 1.2.

Модель 4. Модель з наявністю нефрактальної частини. Ця модель складена з двох частин, перша з яких є косинусною функцією Вейєрштраса – Мандельброта з фрактальною розмірністю $D = 1.5$, а друга – НШС сигналом (показник широкосмуговості $\mu = 0.2$), що не має фрактальних властивостей. Крім того, також потрібна модифікація моделі 4, яку назвемо моделлю 4а, де НШС сигнал замінено на $s(t) = \sqrt{(t-5)/5}$, $t \in [5,10]$.

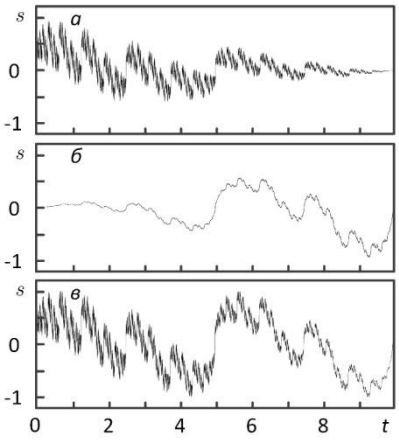


Рис. 1. Модельний детермінований ФС (модель 5), створений як адитивна сума двох ФНШС сигналів (модель $s_0(t)$) з різними значеннями фрактальної розмірності D: *а* – ФНШС сигнал з $D = 1.8$ та лінійно зменшуваною у часі амплітудою, *б* – ФНШС сигнал з $D = 1.2$ та лінійно збільшуваною у часі амплітудою, *в* – власне сама модель 5

Модель 5. Складна модель (рис. 1, в), що утворена адитивною сумою двох ФС, які базуються на модельному ФНШС сигналі $s_0(t)$. Перший з них має фрактальну розмірність $D = 1.8$, а його амплітуда зменшується за лінійним законом (рис. 1, *а*). Другий має $D = 1.2$, натомість його амплітуда зростає також за лінійним законом (рис. 1, *б*).

Формально кажучи, лише модель 1 є моделлю монофрактального сигналу. Моделі 2 і 3 належать до МФС з точки зору глобальних характеристик, але залишаються монофрактальними у локальному сенсі. У розд. 3 вони також будуть розглядатися та порівнюватися із моделями, що є мультифрактальними як у глобальному, так і у локальному сенсах. Моделі 4 і 4а взагалі є фрактальними тільки на першій своїй половині. Але вони вкрай потрібні, щоб побачити, як поводяться досліджувані методи фрактального аналізу, коли їх застосовують до нефрактального сигналу. Модель 5 є типовою моделлю мультифрактального сигналу, оскільки є мультифрактальною як у глобальному, так і у локальному сенсі. Загальна кількість відліків кожного модельного сигналу $N = 8192$.

Результати мультифрактального аналізу модельних сигналів

Спочатку розглянемо результати використання методу WTMM. Він застосовувався для аналізу одразу всієї реалізації модельного ФС або МФС сигналу.

Для моделі 1 маємо $\alpha_{\min} = 0.55$, $\alpha_{\max} = 0.61$, $\Delta\alpha = 0.06$, $\alpha^* = 0.58$, $K_f = 0.0$, $f_\alpha = 1.0$, $\mu_\alpha = 0.10$. Оскільки це модель монофрактального сигналу ($D = 1.5$), згадана раніше теорія стверджує, що його функція мультифрактального спектру $f(\alpha)$ повинна колапсувати у точку (α_0, α_0), де $\alpha_0 = 2 - D = 0.5$. Наразі видно, що це не так. Мультифрактальний спектр став дуже вузьким ($\mu_\alpha = 0.10$), але не точковим ($\mu_\alpha = 0$). Цей практичний факт відомий у літературі [2] і зазвичай пояснюється тим, що теорія створена для математичних фракталів, а ФС у вигляді скінченного дискретного вектора даних є фізичним фракталом. Також слід зазначити, що отримана величина узагальненого показника Херста ($\alpha^* = 0.58$) є дещо завищеною відносно істинного відомого значення ($\alpha_0 = 0.50$).

Для моделі 2 маємо $\alpha_{\min} = 0.30$, $\alpha_{\max} = 0.87$, $\Delta\alpha = 0.57$, $\alpha^* = 0.58$, $K_f = 0.0$, $f_\alpha = 1.0$, $\mu_\alpha = 0.98$. Це типова модель МФС, тому її функція мультифрактального спектра є доволі широка ($\mu_\alpha = 0.98$). Обидві монофрактальні компоненти, показники Гьольдера яких дорівнюють відповідно $\alpha_1 = 0.2$ і $\alpha_2 = 0.8$, утворюють цю ширину ($\Delta\alpha \approx \alpha_2 - \alpha_1$), хоча

увесь спектр виявляється дещо зсунутим праворуч ($\alpha^* > (\alpha_1 + \alpha_2)/2$) на величину, аналогічну тій, яка отримана для моделі 1. Сам спектр є симетричним ($K_f = 0.0$).

Для моделі 3 маємо $\alpha_{\min} = 0.28$, $\alpha_{\max} = 0.86$, $\Delta\alpha = 0.58$, $\alpha^* = 0.57$, $K_f = 0.0$, $f_\alpha = 1.0$, $\mu_\alpha = 1.02$. Будучи моделлю МФС, модель 3 істотно відрізняється від моделі 2 тим, що у першій є всі складові в інтервалі $\alpha \in [0.2, 0.8]$, а у другій – тільки дві кінцеві ($\alpha_1 = 0.2$ і $\alpha_2 = 0.8$). Результати ми отримали фактично однакові. Звідси можна зробити важливий висновок: розрізнити такі моделі виключно за виглядом функції мультифрактального спектру всієї реалізації сигналу неможливо. Тобто одному виглядові мультифрактального спектра можуть відповідати абсолютно різні сигнали. Це не виглядає особливо дивним, якщо згадати, що й однакові значення фрактальних розмірностей двох фракталів нічого не кажуть про подібність їх структур.

Для моделі 4 маємо $\alpha_{\min} = 0.34$, $\alpha_{\max} = 3.99$, $\Delta\alpha = 3.65$, $\alpha^* = 1.83$, $K_f = 0.37$, $f_\alpha = 1.0$, $\mu_\alpha = 1.99$. Це важливий результат, оскільки він показує, яким негативним чином наявність нефрактального (у даному разі, гармонічного) сигналу впливає на результати мультифрактального аналізу. За логікою попередніх двох результатів і зважаючи на те, що будь-яка гладка крива має $\alpha_1 = 1.0$, можна було б очікувати, що функція мультифрактального спектра розташується праворуч від монофрактальної складової ($\alpha_2 = 2 - D = 0.5$), але не сильно відходячи від $\alpha_1 = 1.0$. Але вийшло зовсім не так. Утворився широкий ($\mu_\alpha = 1.99$) асиметричний ($K_f = 0.37$) спектр, до того ж отримані значення $\alpha_{\max} = 3.99$ і $\alpha^* = 1.83$ порушують умови фрактальності ($0 < \alpha^* < 1$). Аналогічні результати можна отримати і у випадку, якщо в даній моделі замінити гармонічну функцію на якусь іншу гладку (не обов'язково, періодичну) функцію (модель 4а). Цей результат добре пояснює, чому в літературі є поради щодо усунення трендів (наприклад, з використанням поліномів, фур'є- та вейвлет-фільтрації і т. і.) з експериментальних даних. Треба також зазначити, що у випадку адитивної суми фрактального та нефрактального сигналів негативний вплив останнього є істотно меншим.

Для моделі 5 маємо $\alpha_{\min} = 0.31$, $\alpha_{\max} = 0.87$, $\Delta\alpha = 0.56$, $\alpha^* = 0.44$, $K_f = 1.20$, $f_\alpha = 1.0$, $\mu_\alpha = 1.23$. Особливість даної моделі у порівнянні з моделями 2 і 3 полягає у тому, що подібно до другої моделі вона має тільки дві монофрактальні компоненти з ($\alpha_1 = 0.2$ і $\alpha_2 = 0.8$), але вони не змінюють одна одну, а задаються із змінним у часі співвідношенням амплітуд. Важливо відзначити, що величини $\alpha_{\min}$, $\alpha_{\max}$, $\Delta\alpha$, f_α отримано майже такі самі, натомість через істотну асиметрію спектра ($K_f = 1.20$) значення узагальненого показника Херста виявилося суттєво іншим, ніж для моделей 2 і 3. Відповідно збільшилося і значення показника відносної ширини функції мультифрактального спектра ($\mu_\alpha = 1.23$). До речі, саме для характеристики подібних сигналів і було запропоновано нові числові характеристики, такі як K_f і μ_α.

Тепер звернемося до методу MF DFA, який дозволяє отримати інформацію про нестаціонарну фрактальну структуру досліджуваних модельних сигналів. Наведені нижче результати отримані для ковзаючого прямокутного вікна шириною 1/10 від загальної довжини дискретного вектору даних аналізованого сигналу.

Декілька слів про спеціальний формат представлення результатів застосування методу MF DFA (рис. 2). На рис. 2, *а* розташовано аналізований сигнал у часовій області. Нижче послідовно зображено часові залежності числових характеристики мультифрактального спект-

ру: $\alpha_{\min}(t)$ (рис. 2, б), $\alpha_{\max}(t)$ (рис. 2, в), $\Delta\alpha(t)$ (рис. 2, г), $\alpha^*(t)$ (рис. 2, д), $K_f(t)$ (рис. 2, е) і $f_\alpha(t)$ (рис. 2, є). У самому низу перебуває графік функції спектральної густини (ФСГ) БВП сигналу (рис. 2, ж). Для всіх модельних сигналів у цьому місці використовувався вейвлет

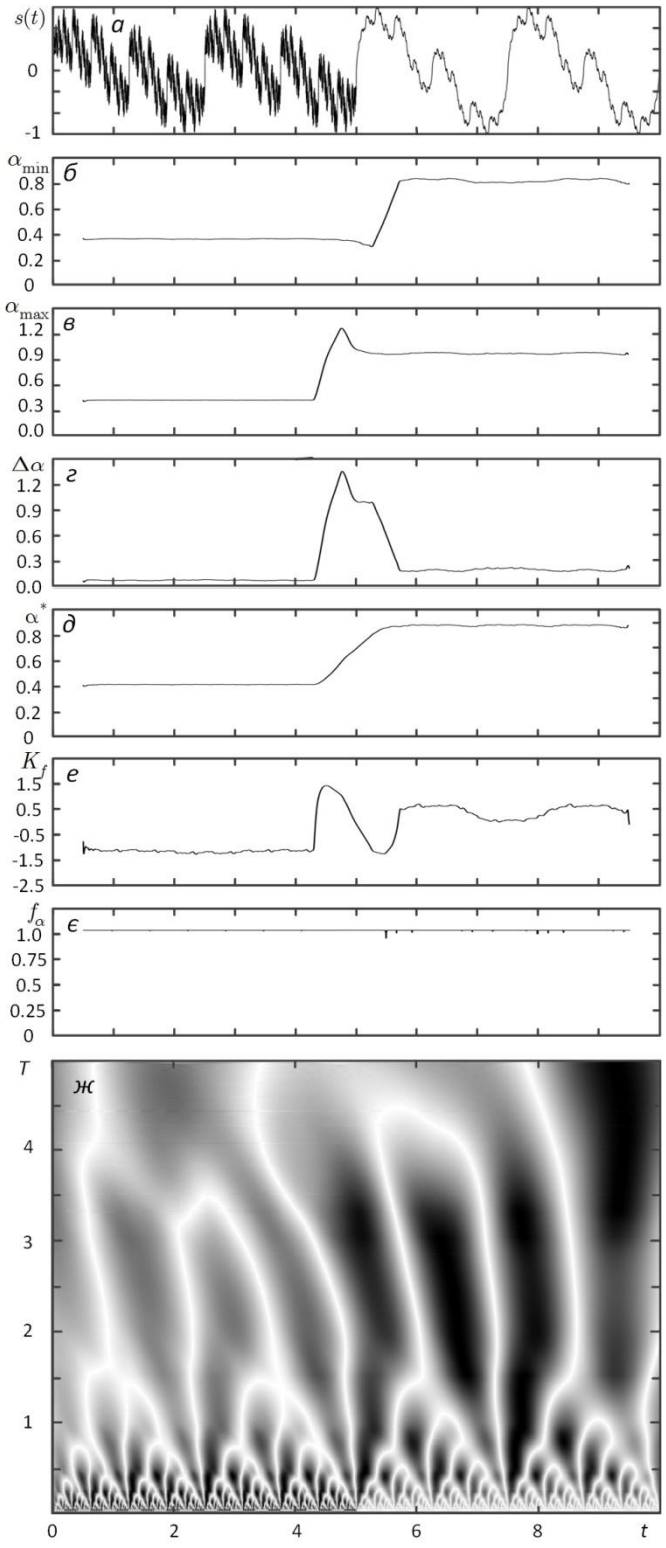


Рис. 2. Результати мультифрактального аналізу модельного ФС (модель 2): *а* – сигнал у часовій області, *б* – $\alpha_{\min}(t)$, в – $\alpha_{\max}(t)$, *г* – $\Delta\alpha(t)$, *д* – $\alpha^*(t)$, *е* – $K_f(t)$, *є* – $f_\alpha(t)$, *ж* – ФСГ БВП сигналу

Добеші четвертого порядку (db4 за поширеною класифікацією СКМ MATLAB/SciLab). ФСГ БВП добре відображає часо-частотний склад аналізованого сигналу, а тому є вкрай корисною для розуміння часових змін числових характеристик мультифрактального спектру.

Розглянемо особливості отриманих результатів для кожної моделі.

Модель 1, як і слід було очікувати, має майже незмінні у часі залежності $\alpha_{\min}(t), \alpha_{\max}(t)$, $\Delta\alpha(t)$, $\alpha^*(t)$, $K_f(t)$ і $f_\alpha(t)$. Слово «майже» означає, що певні флуктуації існують, але вони малі. Величини $\alpha_{\min}(t)$, $\alpha^*(t)$ і $f_\alpha(t)$ майже не відрізняються від аналогічних результатів методу WTMM ($\alpha_{\min} = 0.55$, $\alpha^* = 0.58$, $f_\alpha = 1.0$), середні значення $\Delta\alpha(t)$ (приблизно 0.15) і $\alpha_{\max}(t)$ (приблизно 0.7) дещо перевищують дані методу WTMM ($\Delta\alpha = 0.06$, $\alpha_{\max} = 0.61$), до того ж спектр виявляється трохи асиметричним ($K_f(t) \approx -0.5$ на відміну від $K_f = 0.0$).

Модель 2 (рис. 2), що містить два послідовно розташованих у часі монофрактальних сигнали з $D = 1.8$ і $D = 1.2$, чітко демонструє переваги метода MF DFA відносно вивчення часових змін фрактальних характеристик досліджуваного сигналу. Оскільки метод WTMM дає, у певному сенсі, якесь середнє значення мультифрактальних характеристик ($\alpha_{\min} = 0.30$, $\alpha_{\max} = 0.87$, $\Delta\alpha = 0.57$, $\alpha^* = 0.58$, $K_f = 0.0$, $f_\alpha = 1.0$, $\mu_\alpha = 0.98$) для обох присутніх монофрактальних компонент, то метод MF DFA завдяки вікну у часовій області демонструє положення у часі кожної з компонент, а також зону переходу від однієї до іншої, коли до вікна потрапляє різне співвідношення для кількості енер-

гій обох компонент (рис. 2). Для першої компоненти ($D=1.8$) при $t\in[0.5,4.2]$ маємо $\alpha_{\min}(t)=0.35$, $\alpha_{\max}(t)=0.40$, $\Delta\alpha(t)=0.05$, $\alpha^*(t)=0.40$, $K_f(t)=-1.1$, $f_\alpha(t)=1.0$. Для другої компоненти ($D=1.2$) при $t\in[5.8,9.5]$ маємо $\alpha_{\min}(t)=0.80$, $\alpha_{\max}(t)=1.00$, $\Delta\alpha(t)=0.20$, $\alpha^*(t)=0.90$, $K_f(t)=0.0$, $f_\alpha(t)=1.0$. У середній зоні ($t\in[4.2,5.8]$) спостерігається перехід, де важливо, що узагальнений показник Херста $\alpha^*(t)$ зростає приблизно за лінійним законом. Ширина цієї середньої зони, зрозуміло, залежить від ширини вікна, що використовується, і дорівнює приблизно двом таким ширинам.

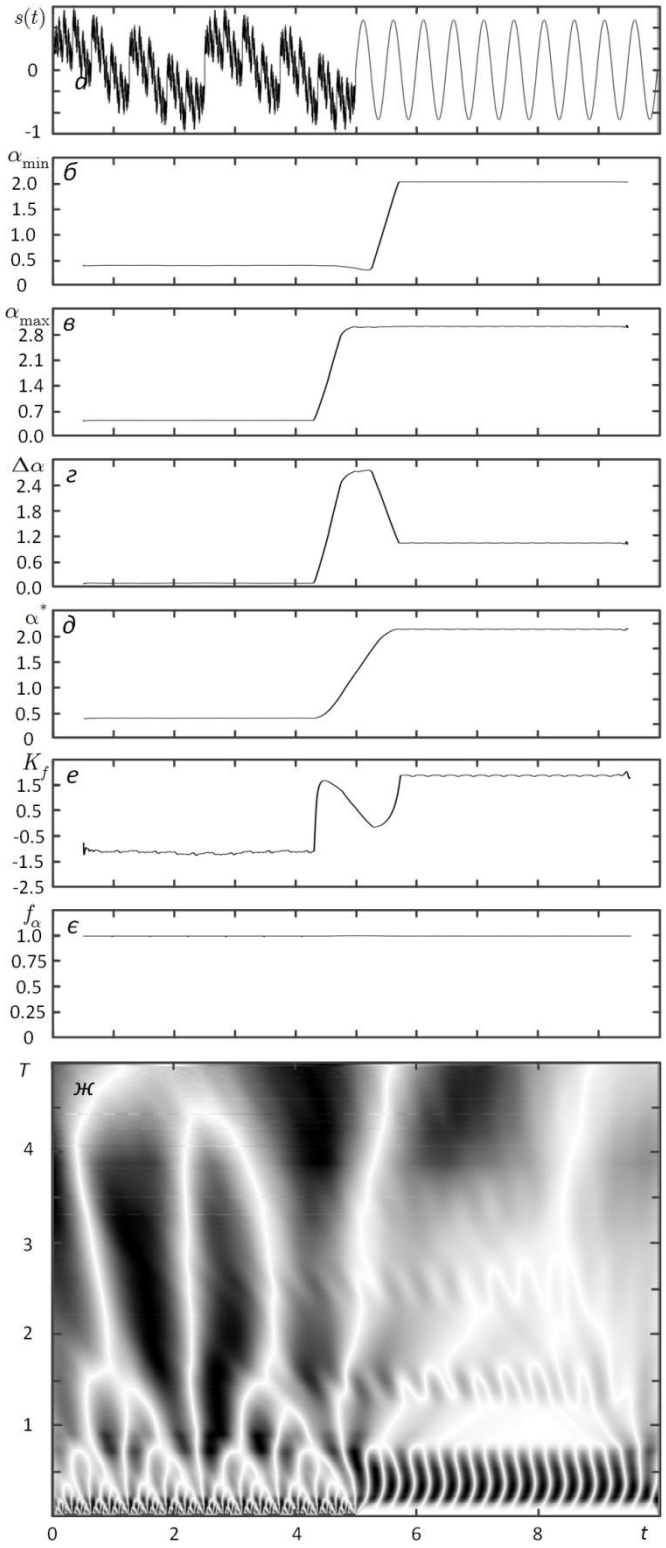


Рис. 3. Результати мультифрактального аналізу модельного ФС (модель 4): *а* – сигнал у часовій області, *б* – $\alpha_{\min}(t)$, *в* – $\alpha_{\max}(t)$, *г* – $\Delta\alpha(t)$, *д* – $\alpha^*(t)$, *е* – $K_f(t)$, *є* – $f_\alpha(t)$, *ж* – ФСГ БВП сигналу

Модель 3, у якої фрактальна розмірність $D(t)$ із часом зменшується за лінійним законом від 1.8 до 1.2, також у методі MF DFA показує зовсім інші результати, ніж у методі WTMM ($\alpha_{\min}=0.28$, $\alpha_{\max}=0.86$, $\Delta\alpha=0.58$, $\alpha^*=0.57$, $K_f=0.0$, $f_\alpha=1.0$, $\mu_\alpha=1.02$). Зрозуміло, що в цілому модель є МФС, але в кожен окремий момент її мультифрактальність залежить від ширини віконної функції. Маємо лінійне зростання $\alpha_{\min}(t)$, $\alpha_{\max}(t)$, $\alpha^*(t)$ і майже незмінність $\Delta\alpha(t)$, $K_f(t)$ і $f_\alpha(t)$. Зазначимо, що $\alpha^*(t)$ лінійно зростає приблизно від 0.4 до 0.8, що певною мірою відрізняється від того, що мало бути (від 0.2 до 0.8).

Модель 4 (рис. 3, *а*), що містить монофрактальну та нефрактальну гармонічну частини, є досить важливою для практичного розуміння результатів мультифрактального аналізу. Метод WTMM для всієї моделі дає $\alpha_{\min}=0.34$, $\alpha_{\max}=3.99$, $\Delta\alpha=3.65$, $\alpha^*=1.83$, $K_f=0.37$, $f_\alpha=1.0$, $\mu_\alpha=1.99$. Метод MF DFA показує (рис. 3), що для монофрактальної частини ($D=1.5$) при $t\in[0.5,4.2]$ маємо $\alpha_{\min}(t)=0.35$, $\alpha_{\max}(t)=0.41$, $\Delta\alpha(t)=0.06$, $\alpha^*(t)=0.40$, $K_f(t)=-1.0$, $f_\alpha(t)=1.0$, а для нефрактальної частини при

$t \in [5.8, 9.5]$ – $\alpha_{\min}(t) = 2.00$, $\alpha_{\max}(t) = 3.00$, $\Delta\alpha(t) = 1.00$, $\alpha^*(t) = 2.15$, $K_f(t) = 1.9$, $f_\alpha(t) = 1.0$. У середній частині ($t \in [4.2, 5.8]$) $\alpha_{\min}(t)$, $\alpha_{\max}(t)$, $\alpha^*(t)$ зростають приблизно за лінійним законом. Важливо, що формально мультифрактальний спектр нефрактальної частини має форму, схожу з тією, яка спостерігається для МФС, але його числові характеристики далеко виходять за межі умов фрактальності ($0 < \alpha^*(t) < 1$). Нефрактальність на часочастотній площині (рис. 3, *ж*) відображається у зникненні деревоподібної структури, характерної саме для фракталів. Аналогічні результати отримано також для моделі 4а.

Модель 5, яка демонструє МФС, складений з двох монофрактальних компонент, для котрих відношення амплітуд у часі змінюється від 0 до ∞, є особливою, оскільки у кожен момент часу присутні обидві компоненти. Метод WTMM, що не враховує часового розташування компонент, дає $\alpha_{\min} = 0.31$, $\alpha_{\max} = 0.87$, $\Delta\alpha = 0.56$, $\alpha^* = 0.44$, $K_f = 1.20$, $f_\alpha = 1.0$, $\mu_\alpha = 1.23$. Як вже було сказано, більшість із цих параметрів не відрізняється від того, що метод WTMM дає для моделей 2 і 3. Метод MF DFA виявляє, що функції $\alpha_{\min}(t)$, $\alpha_{\max}(t)$, $\alpha^*(t)$ і $\Delta\alpha(t)$ зростають нелінійно, причому швидкість їх зростання збільшується із часом.

Тепер на основі результатів мультифрактального аналізу модельних сигналів, отриманих з використанням методів WTMM і MF DFA, необхідно зробити певні висновки для різних видів сигналів, які, сподіваємося, будуть цікавими та корисними для практиків. Такими видами наразі є монофрактальні, мультифрактальні та нефрактальні сигнали.

Мультифрактальний аналіз монофрактальних сигналів

1. На відміну від теоретичних положень, отриманих для математичних фракталів, під час мультифрактального аналізу модельних монофрактальних сигналів, що належать до фізичних фракталів, мультифрактальний спектр не колапсує у точку. Це пов'язано, насамперед, з тим, що модельний сигнал має вектор даних скінченої довжини, а тому як і у будь-якого фізичного фрактала, фрактальні властивості існують лише у обмеженому з обох боків діапазоні масштабів. Отримані мультифрактальні спектри є досить вузькими, добрим індикатором чого виступає показник відносної ширини мультифрактального спектру μ_α. Для монофрактальних сигналів він зазвичай задовольняє умові μ_α □ 1. Для математичного монофрактала, зрозуміло, ми отримаємо $\mu_\alpha = 0$. Сказане є справедливим як у глобальному (метод WTMM), так і у локальному (метод MF DFA) сенсі.

2. Як глобальні, так і локальні отримані оцінки узагальненого показника Херста виявляються дещо завищеними для всіх моделей. Отже, відповідні оцінки фрактальної розмірності, у відповідності до цього, стають заниженими.

3. Параметр f_α не завжди своїм зменшенням відносно одиниці вказує на монофрактальність сигналу. Скоріш за все, причина – та сама, що описана у першому висновку.

4. Стохастичні моделі у порівнянні з детермінованою (модель 1) на окремих реалізаціях у цілому показують гірший результат у сенсі стабільності мультифрактальних характеристик. Однак це можна покращити створенням великого набору цих випадкових реалізацій із наступним усередненням по ансамблю.

5. Шукаючи на часових графіках мультифрактальних характеристик перехід досліджуваного сигналу у монофрактальний режим, у першу чергу слід співставляти графіки функцій $f_\alpha(t)$, $\Delta\alpha(t)$ і $\mu_\alpha(t)$. Те, що перша від них не відхилилася від одиниці, ще не заперечує монофрактальності. На практиці виконання умов $\Delta\alpha \le 0.2$ і $\mu_\alpha \le 0.25$ є вагомішими.

Мультифрактальний аналіз мультифрактальних сигналів

1. Моделі мультифрактальних сигналів можуть істотно відрізнятися одна від одної у сенсі поведінки мультифрактальних характеристик у часі. Все залежить від того, як саме

монофрактальні складові розташовано у сигналі. Вони можуть змінювати одна одну стрибком (модель 2), поступово (модель 3), або взагалі існувати одночасно із різним відношенням амплітуд (модель 5). Але всі вони у глобальному сенсі є мультифрактальними сигналами.

2. Метод WTMM, що досліджує глобальні мультифрактальні характеристики МФС, у випадку, коли різні за часовою фрактальною структурою моделі мають однаковий діапазон використаних значень показника Гьольдера, дає для них фактично однакові результати та не дозволяє впевнено відрізнити одну від одної. Отже, обмежувати мультифрактальний аналіз виключно методом WTMM є недоцільним.

3. Локальні мультифрактальні характеристики МФС у межах часового вікна скінченної довжини, що використовується у методі MF DFA, можуть бути як майже монофрактальними (модель 2 у зонах, коли до вікна потрапляє тільки одна монофрактальна складова), близькими до монофрактальних (модель 3, частина моделі 5), так і повністю мультифрактальними (середня зона моделі 2, друга частина моделі 5). Виявлення таких інтервалів є корисним під час аналізу експериментальних даних.

3. Узагальнений показник Херста $\alpha^*(t)$ в методі MF DFA є досить добрим індикатором часової поведінки фрактальних особливостей досліджуваного сигналу. Результати аналізу показують, що він добре віддзеркалює часову поведінку показника Гьольдера $\alpha(t)$, що застосовувався під час створення самих модельних сигналів. Більш того, він добре корелює з часовими змінами показника Херста $H(t)$, що застосовується у монофрактальному аналізі.

4. Коефіцієнт асиметрії функції мультифрактального спектру, що відображає відхилення досліджуваного зразка спектру від симетрії, також часто є корисним. Наприклад, метод WTMM, саме він дозволяє відрізнити модель 5 від моделей 2 і 3, оскільки решта мультифрактальних характеристик у цьому випадку приймають майже однакові значення.

Мультифрактальний аналіз нефрактальних сигналів

1. Метод WTMM, реагуючи на наявність нефрактальної гармонічної компоненти в моделі 4 (рис. 3), створює аномально широкий ($\mu_\alpha \approx 2$) мультифрактальний спектр, положення максимуму ($\alpha^* = 1.83$) якого істотно порушує умови фрактальності ($0 < \alpha^* < 1$). Тобто гармонічна компонента веде себе як складова з $\alpha_0 > 2$. В цілому це здається дещо дивним, оскільки добре відомо, що гладка функція повинна мати $\alpha_0 = 1$. Джерелом проблем не є періодичність функції, як про це було сказано у роботі [2], де обґрунтовувалася необхідність видалення гармонічних трендів для проведення мультифрактального аналізу. Ми встановили, що заміна періодичної функції на неперіодичну (модель 4а) зовсім не впливає на результат. До речі, це пояснює, чому в інших роботах (див., наприклад, [18]) є поради взагалі усувати будь-який тренд з використанням поліноміальної апроксимації.

2. Метод MF DFA демонструє, що мультифрактальний спектр чисто гармонічного сигналу без будь-яких інших домішок (рис. 3) (модель 4 при $t \in [5.8, 9.5]$) сам по собі є досить широким ($\Delta\alpha(t) = 1.00$, $\mu_\alpha \approx 0.5$) і асиметричним ($K_f(t) = 1.9$), до того ж максимум зсунуто ліворуч. Узагальнений показник Херста виявляється екстремально великим ($\alpha^*(t) = 2.15$). Заміна періодичної функції на неперіодичну (модель 4а) призводить фактично лише до зміни напрямку нахилу мультифрактального спектру.

3. У разі наявності в аналізованому сигналі адитивної суміші фрактальної та нефрактальної компонент, негативний вплив нефрактальної компоненти істотно зменшується, особливо коли її амплітуда є меншою порівняно з амплітудою фрактальної компоненти.

4. Найбільшу ширину мультифрактального спектру в методі MF DFA для обох моделей (моделі 4 і 4а) має середня зона ($t \in [4.2, 5.8]$), до того ж ширина зростає по мірі того, як все більша кількість відліків нефрактального сигналу потрапляє до аналізуючого вікна у часовій області.

5. На нашу думку, існуючі поради усувати з аналізованого сигналу всі тренди (як періодичні, так і неперіодичні) є дещо категоричними. Зробити це, дійсно, можна, наприклад, з використанням поліномів або фур'є- чи вейвлет-фільтрації. Різноманітних засобів для цього сьогодні вистачає. Але дослідник має бути впевнений, що усунувши тренд, він не спотворить сам досліджуваний процес. Наведемо відповідний приклад. Розглянемо модель ФНШС сигналу $s_1(t)$, що базується на функції Вейєрштраса (будь-яка половина сигналу на рис. 2, *а*). Її можна апроксимувати, наприклад, чотирипелюстковим нефрактальним НШС сигналом. А після цього видалити такий тренд. Чим буде те, що ми отримаємо у результаті, сказати важко, але точно не ФНШС сигналом. Звідси висновок: відповідальність за те, що у кожному випадку вважати трендом, а що – корисним сигналом, лежить на дослідникові. Тут можуть сказати, що це збільшує суб'єктивність результатів аналізу. Це насправді так. Але фрактальний і мультифрактальний аналізи самі по собі, подібно до будь-якого виду спектрального аналізу, дійсно є суб'єктивними по своїй природі.

Метод коригуючої функції для мультифрактального аналізу

Вище декілька разів зазначалося, що узагальнений показник Херста α^* майже завжди видає явно завищену оцінку для кожного модельного сигналу. Добре відомо [19], що й фактично кожен метод монофрактального аналізу дає зсунуту оцінку фрактальної розмірності D модельного сигналу. Автори у роботі [19] вже запропонували вихід із такого становища, створивши метод коригуючої функції (КФ) для монофрактального аналізу.

Аналогічну ідею можна застосувати й тепер, модифікувавши метод КФ для мультифрактального аналізу. Отже, основна ідея методу КФ для мультифрактального аналізу полягає у зниженні відхилення отримуваної оцінки узагальненого показника Херста від величини показника Гьольдера аналізованого сигналу за рахунок створення спеціальної КФ на основі модельних сигналів із заздалегідь відомими значеннями показника Гьольдера.

Зважаючи на обмежений обсяг даної статті, зазначимо, що модифікація методу КФ полягає у формальній заміні у наведених у роботі [19] співвідношеннях оцінки фрактальної розмірності D^* на оцінку узагальненого показника Херста α^*, яку позначатимемо як $\hat{\alpha}$, а відомого істинного значення фрактальної розмірності D модельного сигналу – на значення його показника Гьольдера α.

Отже, в результаті коригування положення максимуму мультифрактального спектра α^* відбувається його паралельний перенос. Тому відповідних змін зазнають $\alpha_{\min}$, $\alpha_{\max}$ і μ_α. Величини ж $\Delta\alpha$ і K_f залишаться незмінними.

Залишається додати, що у наших дослідженнях для побудови програмних реалізацій КФ методів WTMM і MF DFA ми використовували детерміновану модель монофрактального сигналу – модель 1, оскільки вона має найкращу часову стабільність мультифрактальних характеристик із усіх розглянутих нами моделей.

Висновки

1. Запропоновано нові числові характеристики, що є корисними для проведення мультифрактального аналізу сигналів і процесів, обґрунтовано доцільність їх створення, продемонстрована корисність та ефективність. До них належать коефіцієнт асиметрії функції мультифрактального спектру, показник відносної ширини мультифрактального спектру та розмірність носія мультифрактала.

2. Із використанням методів WTMM і MF DFA проведено мультифрактальний аналіз набору модельних ФС і МФС. Виявлено особливості мультифрактального аналізу -, мультифрактальних і нефрактальних сигналів і процесів, сформульовано відповідні рекомендації для практиків. Розроблено зручні формати представлення результатів аналізу.

3. Встановлено, що під час переходу МФС до монофрактального режиму функція мультифрактального спектру фізичного фрактала не колапсує у точку, як це має відбуватися у

теорії для математичного фрактала. Ознаками появи такого переходу є зменшення $\Delta\alpha$ і μ_α, а також виконання умови $f_\alpha < 1$. Результати моделювання свідчать, що на практиці режим вже можна вважати монофрактальним, якщо $\Delta\alpha \le 0.2$ і $\mu_\alpha \le 0.25$ навіть коли $f_\alpha = 1$.

4. Продемонстровано, що мультифрактальний аналіз нефрактальних сигналів призводить до появи мультифрактальних спектрів із аномальними значеннями мультифрактальних характеристик. По-перше, такі спектри суттєво порушують умови фрактальності ($0 < \alpha^* < 1$) для узагальненого показника Херста, який завжди для них $\alpha^* > 1$. По-друге, вони також виявляються дуже широкими ($\Delta\alpha \ge 1$), а сам нефрактальний сигнал у мультифрактальному спектрі формально поводить себе як монофрактальна складова з $\alpha_0 > 1$.

5. Доведено на практичних прикладах, що існуючі в літературі поради усувати з аналізованого сигналу всі тренди (як періодичні, так і неперіодичні) є дещо категоричними. Дослідник має бути впевнений, що усунувши тренд, він не спотворить сам досліджуваний процес, оскільки тренд може виявитись невід'ємною частиною досліджуваного процесу. Прикладом цього є ФНШС процес. Відповідальність за те, що саме у кожному конкретному випадку вважати трендом, а що – корисним сигналом, лежить на самому дослідникові. Збільшення суб'єктивності результатів аналізу наразі не є загрозою, оскільки фрактальний і мультифрактальний аналізи самі по собі, подібно до будь-якого виду спектрального аналізу, дійсно є суб'єктивними за своєю природою.

6. Встановлено, що оцінки $\hat{\alpha}$ однієї з найважливіших мультифрактальних характеристик – узагальненого показника Херста α^*, що отримуються як методом WTMM, так і методом MF DFA, є завищеними (у діапазоні $\alpha = 0.2 - 0.8$ відхилення $\hat{\alpha}$ від істинного значення α складає $90 \div 5$ %), що негативним чином впливає на результати мультифрактального аналізу реальних сигналів і процесів. Продемонстровано, що функціональна залежність $\hat{\alpha} = F(\alpha)$ є нелінійною для обох аналізованих методів. Більш того, виявлена також залежність величини $\hat{\alpha}$ від кількості точок дискретного вектору даних аналізованого сигналу N, тобто у результаті маємо $\hat{\alpha} = F(\alpha, N)$.

7. Із метою покращення точності оцінки мультифрактальних характеристик створено метод КФ для мультифрактального аналізу. Його застосування дозволило істотно знизити відхилення отримуваної оцінки узагальненого показника Херста від істинної відомої величини показника Гьольдера аналізованого сигналу. Так, за мінімальної дозволеної кількості дискретних відліків даних аналізованого сигналу ($N_{\min} = 32$) у діапазоні $\alpha = 0.2 - 0.8$ відхилення скоригованого значення узагальненого показника Херста від істинного значення α складає лише $8 \div 3$ %, а саме істинне значення α стабільно потрапляє у середину довірчого інтервалу (рівень надійності складає 0.9).

Список літератури:

1. Лазоренко О. В., Черногор Л. Ф. Фрактальная радиофизика. 1. Теоретические основы // Радиофизика и радиоастрономия. 2020. Т. 25, № 1. С. 3 – 77.

2. Kantelhardt J. W., Zschiegner S. A., Koscielny-Bunde E., Havlin S., Bunde A., Stanley H. E. Multifractal detrended fluctuation analysis of nonstationary time series // Physica A: Statistical Mechanics and Its Applications. 2002. Vol. 316, No. 1 – 4. P. 87 – 114.

3. Jaffard S. Multifractal Formalism for Functions. Part E Results Valid for All Functions // SIAM J. Math. Anal. 1997. Vol. 28, No. 4. P. 944-970.

4. Arneodo A., Audit B., Kestener P. and Roux S. Multifractal Formalism based on the Continuous Wavelet Transform // Scholarpedia. 2007, Vol. 3, P. 1-20.

5. Arneodo A., Grasseau G., and Holschneider M. Wavelet transform of multifractals // Phys. Rev. Lett. 1988. Vol. 61. P. 2281 – 2284.

6. Mallat S. A Wavelet Tour of Signal Processing. San Diego, CA: Academic Press, 1998.

7. Muzy J. F., Bacry E., Arneodo A. Wavelets and multifractal formalism for singular signals: Application to turbulence data // Physical Review Letters. 1991. Vol. 67, No. 25. P. 3515–3518.

8. Arneodo A., Bacry E., and Muzy J. F. The thermodynamics of fractals revisited with wavelets // Physica A. 1995. Vol. 213. P. 232–275.

9 Muzy J.-F., Bacry E., Arnéodo A. Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus-maxima method // Physical Review E, American Physical Society (APS). 1993. Vol. 47, No. 2. P. 875 – 884.

10. Weiss B., Clemens Z., Bódizs R., Vágó Z., Halász P. Spatio-temporal analysis of mon-ofractal and multifractal properties of the human sleep EEG // Journal of Neuroscience Methods. 2009. Vol. 185. P. 116–124.

11. Ihlen E. A. F. Introduction to multifractal detrended fluctuation analysis in Matlab // Frontiers in Physiology. June 2012, Vol. 3, Article 141.

12. Telesca L., Lapenna V., Macchiato M. Mono- and multi-fractal investigation of scaling properties in temporal patterns of seismic sequences // Chaos, Solitons and Fractals. 2004. Vol. 19. P. 1–15.

13. Ge E., Leung Y. Detection of crossover time scales in multifractal detrended fluctuation analysis // Journal of Geographical Systems. 2012. Vol. 15, No. 2. P. 115 – 147.

14. Sarlis N. V., Skordas E. S., Mintzelas A., Papadopoulou K. A. Micro-scale, mid-scale, and macroscale in global seismicity identified by empirical mode decomposition and their multifractal characteristics // Scientific Reports. 2018. Vol. 8. P. 9206.

15. Astanin LY, Kostylev A A. Ultrawideband Radar Measurements: Analysis and Processing. London : The Institute of Electrical Engineers, 1997.

16. Feder J. Fractals. New York and London : Springer, 1988. 284 p.

17. Turcotte D. L. Fractals and Chaos in Geology and Geophysics. Cambridge: Cambridge University Press, 1997. 398 p.

18. Jaffard S. Multifractal Formalism for Functions. Part I: Results Valid for All Functions // SIAM J. Math. Anal. 1997. Vol. 28, No. 4. P. 944-970.

19. Лазоренко О. В., Онищенко А. А., Чорногор Л. Ф. Метод коригуючої функції для фрактального аналізу // Радіотехніка. 2022. Вип. 210. С. 177 – 187.

Надійшла до редколегії 07.10.2022

Відомості про авторів:

Лазоренко Олег Валерійович – д-р фіз.-мат. наук, доцент, Харківський національний університет імені В. Н. Каразіна, завідувач кафедри загальної фізики, Україна; e-mail: Oleg.V.Lazorenko@karazin.ua; ORCID: https://orcid.org/0000-0002-0250-8671

Онищенко Андрій Анатолійович – Харківський національний університет радіоелектроніки, старший викладач кафедри фізики, Україна; e-mail: andrey.onishchenko@nure.ua, ORCID: https://orcid.org/0000-0002-2118-9119

Чорногор Леонід Феоктистович – д-р фіз.-мат. наук, професор, Харківський національний університет імені В.Н. Каразіна, завідувач кафедри космічної радіофізики, Україна; e-mail: Leonid.F.Chernogor@karazin.ua; ORCID: https://orcid.org/0000-0001-5777-2392

RADIO LOCATION AND RADIO NAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

УДК 621.396.96, 621.397.48:004.932.2 **DOI:10.30837/rt.2022.4.211.06**

В.М. КАРТАШОВ, д-р техн. наук, В.О. ПОСОШЕНКО, канд. техн. наук,
М.В.РИБНИКОВ, А.І. КАПУСТА, Є.В. ПЕРШИН

ОСОБЛИВОСТІ ЗАДАЧ ВИЯВЛЕННЯ І СПОСТЕРЕЖЕННЯ ГРУП БЕЗПІЛОТНИХ ЛІТАЛЬНИХ АПАРАТІВ

Вступ

Поява та значне розширення функціональних можливостей безпілотних літальних апаратів (БПЛА) дозволило застосовувати їх при вирішенні найрізноманітніших завдань у різних галузях людської діяльності [1 – 4]: при спостереженні за різними об'єктами та територіями, для доставки вантажів, ретрансляції радіосигналів, у сільськогосподарській діяльності, військовій справі тощо. Однак поряд з виконанням корисних функцій БПЛА несуть і значну потенційну загрозу, розширюючи можливості протиправних дій у різних сферах діяльності [5 – 8].

При вирішенні завдань запобігання несанкціонованим діям з використанням БПЛА в даний час найчастіше використовуються радіолокаційні, оптичні, інфрачервоні та акустичні методи та засоби для їх спостереження [9 – 18]. Питанням виявлення БПЛА присвячено значну кількість публікацій у періодичній пресі та матеріалах міжнародних наукових конференцій, і кількість їх стрімко зростає. Як випливає з літератури, до теперішнього часу досягнуто певних наукових і технічних результатів, що забезпечують виявлення та ідентифікацію типів безпілотних літальних апаратів з різним ступенем достовірності в різних умовах застосування. Проте загалом ситуація така, що потреби практики виявлення БПЛА задовольняються нині далеко ще неповною мірою [1 – 3].

Сучасна тенденція підвищення ефективності використання БПЛА полягає у переході від одиночного до групового застосування, що реалізується в рамках реалізації стратегії мережецентричного управління [2 – 4]. Розробка наукових та технічних основ групового застосування є розвитком ідей щодо підвищення ефективності їх використання та досягнення необхідних результатів при малих витратах сил та засобів. Матеріальною основою групового використання БПЛА є вдосконалення технічних характеристик літальних апаратів та розвиток спеціалізованих БПЛА, призначених на вирішення деяких конкретних завдань.

Основні задачі використання угруповань БПЛА:

- підвищення ймовірності виконання поставленого завдання та ефективності використання наявних засобів шляхом багаторазового дублювання та комплексування функціональних можливостей літальних апаратів, а також шляхом спеціалізації окремих апаратів у групі;
- маскування основного напрямку та цілей групи БПЛА, завантаження та дезорганізація систем виявлення, спостереження, управління, цілерозподілу та впливу шляхом відволікання наявних засобів системи виявлення та впливу на безліч цілей, що входять до групи БПЛА;
- перевищення можливостей засобів протиповітряної оборони (ППО) шляхом використання значної кількості об'єктів у групі та виснаження наявних ресурсів;
- формування «віртуальної насиченої повітряної обстановки» з метою імітації масованого застосування засобів нападу;
- зменшення психологічної стійкості противника та його деморалізація.

В даний час розробляються математичні методи побудови груп БПЛА, технології їх застосування при вирішенні різних завдань як в умовах моделювання ситуацій, так і в натурних експериментах і реальних умовах.

Вочевидь, чим складнішим є алгоритм функціонування групи БПЛА, чим більш вона неоднорідна і автономна, тим складніші завдання здатна виконувати. При цьому бортовий комплекс функціонування та управління кожного конкретного БПЛА також повинен відповідати завданням групи.

Як випливає з викладеного, завдання спостереження за групою безпілотних літальних апаратів є значно складнішим порівняно із завданням спостереження одиночних БПЛА [19 – 24].

Основним завданням статті є розгляд особливостей вирішення сукупності завдань, пов'язаних з виявленням та спостереженням групи БПЛА, комплексно інтегрованою системою, що включає різні інформаційні канали.

Просторове ешелонування окремих каналів комплексних систем

Розглянемо інформаційні, енергетичні та пошукові можливості окремих засобів виявлення, що входять до складу інтегрованої системи спостереження БПЛА, з метою побудови алгоритму ефективної спільної обробки вхідних сигналів, що надходять, з урахуванням різних можливостей окремих каналів (за дальністю, розпізнаванням тощо).

Можливості різних радіо-, оптичних та акустичних засобів виявлення та супроводу малих БПЛА наведено в таблиці [2, 26]:

Можливості різних методів із супроводу та спостереження малих БПЛА

Характеристика	Радіо		Оптичні			Акустичні
	Засоби РЛР (РЛС)	Засоби РРТР	Засоби ОЕР у видимому діапазоні	Засоби ОЕР в ІЧ діапазоні	Лазерні засоби	Засоби АР
Виявлення у денний час	+	+	+	–	+	+
Виявлення у нічний час	+	+	–	+	+	+
Виявлення в умовах природних перешкод	+	+	+	+	+	+
Виявлення БПЛА серед природних об'єктів (насамперед – птахів)	–	+	–	–	–	±
Виявлення у складних погодних умовах	±	+	–	–	–	–
Ідентифікація БПЛА	–	+	±	±	–	+
Селекція одиночних та групових цілей	+	+ (по різним каналам)	+	+	+	+ (для БПЛА різних типів)
Супровід та формування траєкторії	+	+ (для багатопозиційної системи)	+	+	+	+ (для багатопозиційної системи)
Дальність дії	висока	висока	середня	середня	середня	низька

Радіолокаційна характеристика БПЛА – ефективна площа розсіювання, що визначає потужність розсіяного радіосигналу та можливості його енергетичного виявлення, визначається формою та розмірами об'єкта, матеріалами, з якого він виготовлений, довжиною хвилі та поляризацією радіосигналу. При вирішенні завдань розпізнавання та класифікації БПЛА використовується його сигнатура, що є, по суті, радіолокаційним портретом об'єкта. Радіолокаційна сигнатура (мікродоплерівська сигнатура) БПЛА визначається кінематичними властивостями цілі, а також модуляцією розсіяного сигналу елементами літального апарату, що рухаються, – гвинтами, лопатками турбореактивного двигуна і т.д., фізичними і геометричними особливостями цілі.

Відомо, що зменшення масогабаритних характеристик БПЛА супроводжується суттєвим зменшенням дальності виявлення радіолокаційними засобами. При використанні в конструкції літальних апаратів композитних, радіопрозорих матеріалів процес їх виявлення з використанням радіолокаційних станцій також ускладнюється [27, 28].

Методи та технічні засоби оптико-електронного спостереження, що працюють у видимому діапазоні електромагнітних хвиль, забезпечують непогані характеристики при спостереженні БПЛА, у тому числі малорозмірних та малошвидкісних. У той же час є залежність результатів оптичного спостереження літальних апаратів від стану атмосфери, погодних умов і часу доби.

Результати натурних випробувань [26] показують, що середня дальність оптичного спостереження БПЛА оптико-електронними засобами розвідки при його спостереженні з бокових ракурсів становить 150 – 700 м, а спереду – 100 – 400 м [27].

Технічні засоби оптико-електронного спостереження БПЛА у видимому діапазоні спектра мають недостатні пошукові можливості та потребують зовнішнього цілевказання при початковому виявленні об'єкта.

Використовуються засоби оптико-електронного виявлення БПЛА в ІЧ-діапазоні, які забезпечують найбільшу ефективність у нічний час.

У літературі немає достовірних даних про дальність виявлення БПЛА з використанням тепловізійних камер, однак зазначається, що дальність у цьому випадку не перевищуватиме дальності спостереження БПЛА у видимому діапазоні спектра.

Застосування акустичних засобів спостереження дозволяє виявляти БПЛА, визначати його пеленг, клас літального апарату. Однак є й недоліки акустичних систем, що обмежують можливості їх застосування для спостереження БПЛА [30]: порівняно невеликі відстані виявлення БПЛА – до 0,8 км за висотою та до 1,2 км за дальністю, невисока точність оцінки координат БПЛА, спричинена, насамперед, вітряною рефракцією акустичних хвиль в атмосфері.

Таким чином, найкращі можливості пошуку БПЛА має радіолокаційний метод. Пошукові можливості оптичного, інфрачервоного та акустичного методів значно слабші. Зони виявлення різних методів і засобів, що входять до складу комплексної системи, можуть бути представлені графічно (рис. 1).

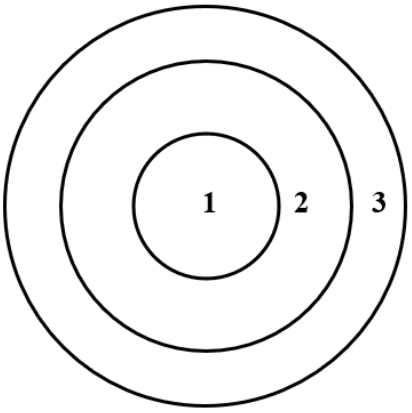


Рис. 1. Графічне подання зон виявлення БПЛА каналами інтегрованої системи: 1 – акустичний канал; 2 – оптичний та інфрачервоний канали; 3 – радіолокаційний канал

Комплексування інформаційних каналів

З позицій статистичної теорії інформаційних радіосистем є два підходи до комплексування засобів спостереження [31 – 33]. З використанням першого підходу завдання комплексування вирішується на етапі первинної обробки інформації (обробки сигналів), під час використання другого – на етапі вторинної обробки інформації (на етапі винесених рішень).

При першому підході до комплексування у межах статистичної теорії радіосистем за результатами спостереження векторного процесу, складовими якого є вхідні сигнали різних каналів, виробляється синтез алгоритмів первинної обробки сигналів у кожному каналі, і навіть відбувається об'єднання інформації, одержуваної кожному з каналів [34, 35]. Такий підхід дозволяє синтезувати оптимальну (відповідно до обраного критерію якості) інтегровану систему обробки інформації (ІСОІ), яка дозволяє отримати

максимальну кількість корисної інформації з сигналів, що спостерігаються на входах інформаційних каналів.

При другому підході компонентами векторного процесу, що спостерігається, будуть вихідні дані пристроїв первинної обробки сигналів. Вони є рішенням, прийнятим на етапі виявлення, оцінки координат об'єкта тощо. Таким чином, здійснюється синтез інтегрованої системи вторинної обробки інформації (ІСВОІ). Оскільки синтез ІСВОІ здійснюється за наявних обмежень на структуру та параметри пристроїв первинної обробки, які фізично реалізовані, то якість інформації на виході ІСВОІ може виявитися нижчою порівняно з якістю вихідних результатів ІСОІ. Зниження якості інформації обумовлено існуючими обмеженнями на структуру системи.

Незважаючи на деякий можливий програш ІСВОІ порівняно з ІСОІ виконання оптимізації на етапі вторинної обробки (етапі рішень) може виявитися дуже ефективним на практиці, оскільки спирається на використання тих пристроїв первинної обробки сигналів, які є та використовуються для побудови відповідних інформаційних каналів.

Використання математичних методів статистичної теорії радіосистем дозволяє досить гнучко виконувати синтез оптимальних структур комплексних систем виявлення, дозволу та вимірювання параметрів груп БПЛА під час використання різних технічних засобів інформаційних каналів.

Комплексування алгоритмів виявлення

Синтезована оптимальна комплексна система обробки – ІСОІ може виявитися складною, особливо в тому випадку, коли використовуються вимірювачі – рознесені в просторі. В цьому випадку при реалізації ІСОІ необхідно використовувати канали зв'язку з досить високою пропускною здатністю. Система виходить набагато простішою для реалізації на практиці, коли здійснюється комплексування алгоритмів виявлення на етапі вторинної обробки [33].

У цьому випадку у кожному з каналів завдання виявлення груп БПЛА вирішується незалежно один від одного. Подальше комплексування здійснюється внаслідок спільної обробки вихідних даних виявлювачів каналів, тобто їх рішень про наявність чи відсутність об'єктів. На етапі вторинної обробки оптимізація комплексної обробки заснована на використанні критерію максимуму відношення правдоподібності, для формування якого використовуються отримані раніше приватні рішення канальних виявлювачів.

Розглянемо синтез оптимального виявлювача на етапі вторинної обробки сигналів (рішень) n інформаційних каналів комплексної системи виявлення БПЛА.

Виявлювач *i*-го каналу реалізує певну вирішальну функцію $\vartheta_i(\cdot)$ в результаті аналізу на проміжку часу [0, T] $y_i(t)$ (чи зображення) та виносить рішення $\gamma_i(y_i(t))=1$ про наявність корисного сигналу чи рішення $\gamma_i(y_i(t))=0$ про його відсутність, з ймовірностями відповідно до правильного виявлення D_i чи хибної тривоги F_i.

На виходах канальних виявлювачів формується випадковий вектор рішень $\gamma_1,...,\gamma_n$, компоненти якого набувають значення 0 або 1 з ймовірностями

$$P(\gamma_i=1\,|\,\theta=0)=F_i,\;\; P(\gamma_i=0\,|\,\theta=0)=1-F_i$$
$$P(\gamma_i=1\,|\,\theta=1)=D_i,\;\; P(\gamma_i=0\,|\,\theta=1)=1-D_i.$$

Відповідно до критерію відношення правдоподібності під час використання вхідних сигналів $\gamma_1,...,\gamma_n$ формується підсумкове рішення R_1 про наявність корисного сигналу або R_0 про його відсутність

$$\Lambda_n=\frac{P(\gamma_1,...,\gamma_l\,|\,\theta=1)}{P(\gamma_1,...,\gamma_l\,|\,\theta=0)}\lessgtr H_n. \qquad (1)$$

Беручи до уваги статистичну незалежність канальних спостережень γ_i, співвідношення (1) запишемо у вигляді

$$\sum_{i=1}^{n} \gamma_i \ln\left[\frac{D_i(1-F_i)}{F_i(1-D_i)}\right] <> H_n . \qquad (2)$$

Отриманий вираз визначає алгоритм оптимального комплексування канальних виявлювачів груп БПЛА на етапі вторинної обробки. Відповідно до (2), обробка полягає у підсумовуванні рішень виявлювачів $\gamma_i = 1$, винесених у каналах, з канальними вагами:

$$\eta_i = \ln\left[D_i(1-F_i) / F_i(1-D_i)\right].$$

У тому випадку, якщо значення ймовірностей правильного виявлення та хибної тривоги канальних виявлювачів груп БПЛА дорівнюватимуть – $D_i = D,\ F_i = F,\ i = 1,...,n$, то вагові коефіцієнти каналів набувають однакових значень $\eta_i = \eta$. Поріг виявлення H_n слід вибирати відповідно до критерію Неймана – Пірсона, виходячи з ймовірності хибної тривоги F_n для комплексної системи у цілому.

При розрахунку характеристик виявлення ICBOI (значень D_n и F_n) слід приймати до уваги співвідношення

$$D_n = \sum_{m=h}^{l} C_l^m D^m (1-D)^{l-m},\ \ F_n \le \sum_{m=h}^{l} C_l^m F^m (1-F)^{l-m} . \qquad (3)$$

Структурну схему об'єднання рішень канальних виявлювачів у багатоканальній системі виявлення груп БПЛА наведено на рис. 2.

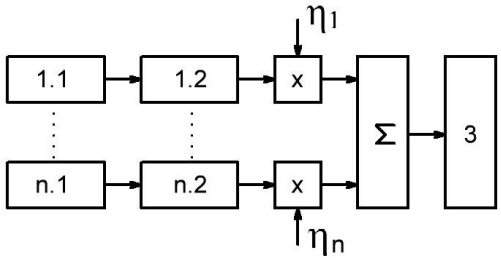


Рис. 2. Структурна схема об'єднання рішень комплексної системи виявлення груп БПЛА:
1.1, n.1 – формувачі сигналів (зображень) інформаційних каналів, 1.2, n.2 – виявлювачі сигналів у каналах,
3 – вирішувальний пристрій

Таким чином, математичні методи статистичної теорії радіосистем дозволяють здійснити оптимальний синтез комплексних систем виявлення та вимірювання параметрів груп БПЛА, оптимізацію алгоритмів обробки багатомодальних сигналів при використанні різних інформаційних каналів та технічних засобів як на етапі первинної, так і на етапі вторинної обробки інформації.

Методи обробки сигналів в інтегрованій системі з використанням цілевказівки

Одне із завдань РЛС в комплексній інтегрованій системі спостереження БПЛА полягає у видачі просторових координат виявлених цілей на певному рубежі для цілевказівки оптико-електронним засобам, що дозволить їм зробити енергетичне виявлення групової цілі без додаткового пошуку (допошуку), або обмежити зону допошуку до прийнятних розмірів. Таким чином відбувається «зав'язування» процесу обробки інформації в комплексній системі за кожною груповою ціллю, виявленою спочатку РЛС, який далі включає все більші ресурси (апаратні, обчислювальні, інтелектуальні). У міру наближення групи БПЛА до об'єкта, що

охороняється, відбувається все більш різноманітна обробка вхідних сигналів і зображень, що надходять на вхід комплексної системи, і витягується з них все більша кількість інформації. У свою чергу групова ціль, у міру наближення до об'єкта, що охороняється, забезпечує можливість для отримання все більшої кількості різноманітної інформації.

Послідовність вирішення сукупності завдань у комплексній інтегрованій системі спостереження БПЛА у міру наближення групової цілі до об'єкта, що охороняється, представляється наступною:

- виявлення групової цілі (енергетичне виявлення);
- оцінка координат групи об'єктів;
- просторове розрізнення та визначення кількості апаратів у групі;
- розпізнавання (визначення типу) кожного окремого апарату;
- оцінка координат кожного літального апарату окремо;
- визначення складу групи (однорідна, неоднорідна);
- визначення спеціалізації групи та розтин характеру її завдань.

Найкращі пошукові можливості та найбільшу дальність має радіолокаційний метод. Саме з використанням методу радіолокації проводиться первинне енергетичне виявлення групи літальних апаратів і оцінка просторових координат групи. Отримані у процесі розв'язання даних завдань результати є основою виконання цілевказання – вказання попередніх просторових координат групи іншим засобам комплексної інтегрованої системи для наступного узгодженого виконання сукупності завдань спостереження.

Обробка та об'єднання багатомодальних сигналів у комплексній системі спостереження за групами БПЛА зменшує наявну невизначеність та сприяє зменшенню похибок, за якими ознаки оцінюються системою [36 – 38]. У цьому випадку використовується присутність в сигналах окремих каналів інформації, що взаємно доповнюється. Наявна надмірність інформації також сприяє підвищенню надійності комплексної системи за наявності аномальних помилок або збоїв у каналах.

Істотним є те, що додаткова інформація з кількох модальностей дозволяє використовувати ознаки сигналів, які неможливо однозначно інтерпретувати за наявності інформації лише від кожної модальності окремо [39, 40]. Наявність можливості проводити паралельну обробку даних у каналах декількох модальностей, що використовуються, дозволяє також більш оперативно отримувати інформацію про групу БПЛА. Рішення (виявлення, ідентифікація) із заданими показниками якості можна отримати і при використанні тільки одного або меншої кількості інформаційних каналів, але це потребує більшого часу для накопичення інформації.

Об'єднання інформації окремих каналів у комплексній системі спостереження груп БПЛА, можливе лише на рівні сигналів, на рівні ознак і рівні рішень [41, 42]. При цьому можуть бути реалізовані такі стратегії об'єднання даних:

- раннього об'єднання, реалізовані лише на рівні сигналів, одержуваних від БПЛА;
- раннього об'єднання, реалізовані лише на рівні ознак опису груп БПЛА;
- пізнього об'єднання, реалізовані на семантичному рівні прийняття рішення;
- гібридного об'єднання.

Використання різних видів стратегій об'єднання даних, зокрема стратегії гібридного об'єднання багатомодальних сигналів наявних каналів інтегрованої системи, дозволяє проводити обробку та об'єднання інформації з урахуванням наявної специфіки завдань, що вирішуються даною системою, та можливостей технічних засобів у кожному каналі.

Значні можливості для об'єднання канальної інформації в інтегрованих системах відкриваються з розвитком нейронних мережевих технологій. Об'єднання інформаційних каналів у цьому випадку, зокрема, здійснюється не на рівні ознак, що формуються в окремих концептах, а шляхом об'єднання наявної в каналах інформації в єдине семантичне мультимодальне подання (мультимодальну функцію) [43, 44].

Сигнали, зображення та результати аналізу в каналах, отримані за групою БПЛА, при використанні нейронних мереж та раннього об'єднання зливаються ще до того, як відповідні

канальні уявлення детально вивчені та сформовані відповідні ознаки [45]. У разі пізнього злиття спочатку здійснюється вивчення канальної інформації з допомогою нейронних мереж. У цьому випадку отримані оцінки канальних функцій формують вектор оцінок мультимодальної функції, які є вхідними даними для системи машинного навчання і подальшої інтерпретації отриманої багатоканальної інформації.

При використанні просторового ешелонування каналів і цілевказівок в інтегрованій системі реалізується послідовне накопичення інформації з каналів системи, що послідовно підключаються, проводиться її обробка з використанням нейромережевих або традиційних інформаційних технологій інтерпретації та прийняття рішень за групами БПЛА.

Висновки

1. Сучасна тенденція підвищення ефективності застосування БПЛА полягає в переході від одиночного, до їхнього групового застосування, що реалізується в рамках реалізації стратегії мережецентричного управління.

Відповідно до цього при побудові комплексної інтегрованої системи виявлення та спостереження за БПЛА, що включає різні канали, доцільно враховувати особливості функціонування системи, пов'язані з виявленням і спостереженням груп БПЛА.

2. Розглянуто інформаційні, енергетичні та пошукові можливості окремих засобів виявлення, що входять до складу інтегрованої системи спостереження БПЛА, з метою побудови ефективного алгоритму спільної обробки вхідних сигналів, що надходять, з урахуванням різних можливостей окремих каналів (за дальністю, розпізнаванням тощо). Показано, що найкращими пошуковими можливостями і найбільшою дальністю володіє метод радіолокації виявлення груп БПЛА, далі ідуть за спадною оптичний, інфрачервоний і акустичний методи.

Синтезовано оптимальний алгоритм виявлення груп БПЛА у комплексній інтегрованій системі, що поєднує рішення про виявлення, винесені у приватних каналах. Відповідно до синтезованого алгоритму комплексна обробка полягає у підсумовуванні рішень окремих виявлювачів з деякими вагами, що визначаються якістю рішень, прийнятих у каналах. Якість рішень, у свою чергу, залежить від технічних засобів каналів, що використовуються, та умов спостереження.

3. Запропоновано послідовність вирішення сукупності взаємопов'язаних завдань у комплексній інтегрованій системі спостереження БПЛА у міру наближення групової цілі до об'єкта, що охороняється. Послідовність включає наступні операції: виявлення групової цілі (енергетичне виявлення); оцінка координат групи об'єктів; просторовий розрізнення та визначення кількості апаратів у групі; розпізнавання (визначення типу) кожного окремого апарату; оцінка координат кожного літального апарату окремо; визначення складу групи (однорідна, неоднорідна); визначення спеціалізації групи та розтин характеру її завдань.

Список літератури:

1. Кошкин Р.П. Беспилотные авиационные системы. Москва : Стратегические приоритеты, 2016. 676 с.
2. Макаренко С. И., Тимошенко А. В., Васильченко А. С. Анализ средств и способов противодействия беспилотным летательным аппаратам. Ч. 1. Беспилотный летательный аппарат как объект обнаружения и поражения // Системы управления, связи и безопасности. 2020. № 1. С. 109-146. DOI: 10.24411/2410-9916-2020-10105.
3. Kartashov V.M., Oleynikov V.N, Sheyko S.A., Koryttsev I.V., Babkin S.I., Zubkov O.V. Peculiarities of small unmanned aerial vehicles detection and recognition // Telecommunications and Radio Engineering, 2019. Vol. 78, Iss. 9. P. 771 – 781.
4. Карташов В.М., Олейников В.Н., Шейко С.А. и др. Особенности обнаружения и распознавания малых беспилотных летательных аппаратов // Радиотехника. 2018. Вып. 195. С. 235 – 243.
5. 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, №5. P.461 – 468.
6. Kartashov V. M., Oleynikov V. N., Sheyko S. A., Babkin S. I., Koryttsev I. V., Zubkov O. V., Anokhin M. A. Information characteristics of sound radiation of small unmanned aerial vehicles // Telecommunications and Radio Engineering. 2018. Vol.77, Iss. 10. P. 915 – 924.

7. Карташов В.М., Олейников В.Н., Шейко С.А. и др. Информационные характеристики звукового излучения малых беспилотных летательных аппаратов // Радиотехника. 2017. Вып. 191. С. 181 – 187.

8. Kartashov V. M., Oleynikov V. N., Zubkov O. V, Sheyko S. A. Optical detection of unmanned air vehicles on a video stream in a real-time // The Fourth International Conferenceon Information and Telecommunication Technologies and Radio Electronics (UkrMiCo'2019), 9-13 September 2019, Odessa, Ukraine. 4 p.

9. Oleksandr Sotnikov, Vladimir Kartashov, Oleksandr Tymochko, Oleg Sergiyenko, Vera Tyrsa, Paolo Mercorelli, Wendy Flores-Fuentes. Methods for Ensuring the Accuracy of Radiometric and Optoelectronic Navigation Systems of Flying Robots in a Developed Infrastructure. Chapter 16 // Machine Vision and Navigation; Editors: Sergiyenko**,** Oleg, Flores-Fuentes**.** Wendy**,** Mercorelli**,** Paolo. P.537 – 578.

10. 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, Iss. 9. P 759 – 770.

11. Kartashov V., Oleynikov V., Koryttsev I., Zubkov O., Babkin S., Sheiko S. Processing and Recognition of Small Unmanned Vehicles Sound Signals // 2018 International Scientific-Practical Conference on Problems of Infocommunications. Science and Technology (PIC S and T 2018). Proceedings, 31 January 2019. P. 392 – 396.

12. Kartashov V., Oleynikov V., Koryttsev I., Sheyko S., Zubkov O., Babkin S., Selieznov I. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles // 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 25-29 Feb. 2020. P. 1-4.

13. Kartashov V.M., Oleynikov V.N, Zubkov O.V., Korytsev I.V., Babkin S. I., Sheiko S.A., Kolendovskaya M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles // Telecommunications and Radio Engineering. 2020. Vol. 79, Iss. 9. P. 769 – 780.

14. Oleynikov V., Zubkov O., Kartashov V., Koryttsev I., Sheyko S., Babkin S. Experimental estimation of direction finding to unmanned air vehicles algorithms efficiency by their acoustic emission // 2019 International Scientific-Practical Conference: Problems of Infocommunications. Science and Technology (PIC S and T 2019). Proceeding, 2019. P. 175 – 178.

15. Semenets V.V., Kartashov V.M., Leonidov V.I. Features of Acoustic Noise of Small Unmanned Aerial Vehicles // Telecommunications and Radio Engineering. 2020. Vol. 79, Iss. 11. P. 985 – 995. DOI: 10.1615/TelecomRadEng.v79.i11.80.

16. Тихонов В.А., Карташов В.М., Олейніков В.М. и др. Обнаружение-распознавание беспилотных летательных аппаратов с использованием составной модели авторегрессии их акустического излучения // Вісник НТУУ «КПІ». Радіотехніка. Радіоапаратобудування. 2020. №81. С. 38 – 46.

17. Kartashov V. M., Tikhonov V. A., Voronin V. V. Features of Construction and Application of Complex Systems for the Atmosphere Remote Sounding // Telecommunications and Radio Engineering. 2017. Vol. 78, Iss.8. P.743-749.

18. Карташов В.М., Олейников В.Н., Колендовская М.М. и др. Комплексирование изображений при обнаружении беспилотных летательных аппаратов // Радиотехника. 2020. Вып. 201. С.120 – 129.

19. Kartashov V.M., Tikhonov V.A., Voronin V.V., Tymoshenko L.P. Complex model of random signal in problems of acoustic sounding of atmosphere // Telecommunications and Radio Engineering. 2016. Vol. 75, Iss. 20. P. 1885 – 1892.

20. Developing and Applying Optoelectronics in Machine Vision**.** Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez; 2016, IGI Global, 341 p.

21. Sytnik O., KartashovV. Methods and Algorithms for Technical Visionin Radar Introscopy. Chapter 13 // Optoelectronics in Machine Vision-Based Theories and Applications. IGI Global, 2019. P. 373 – 391.

22. 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.

23. Карташов В.М. Модели и методы обработки сигналов систем радиоакустического и акустического зондирования атмосферы; Харьков : ХНУРЭ, 2011. 234 с.

24. Карташов В.М., Олейников В.Н., Воронин В.В. и др. Методы комплексной обработки и интерпретации радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов // Радиотехника. 2020. Вып. 202.

25. Countering rogue drones. FICCI Committee on Drones, EY, 2018. 31 p.

26. Ростопчин В. В. Ударные беспилотные летательные аппараты и противовоздушная оборона – проблемы и перспективы противостояния // Беспилотная авиация [Электронный ресурс]. 2020. URL: https://www.researchgate.net/publication/331772628_Udarnye_ bespilotnye_letatelny e_apparaty _i_protivovozdusnaa_oborona_-problemy_i_perspektivy_ protivostoania (дата обращения 18.10.2020).

27. Еремин Г. В., Гаврилов А. Д., Назарчук И. И. Малоразмерные беспилотники – новая проблема для ПВО // Отвага [Электронный ресурс]. 29.01.2015. № 6 (14). URL: http://otvaga2004.ru/armiya-i-vpk/armiya-i-vpkvzglyad/malorazmernye-bespilotniki/ (дата доступа 18.10.2020).

28. Ананенков А. Е., Марин Д. В., Нуждин В. М. и др. К вопросу о наблюдении малоразмерных беспилотных летательных аппаратов // Труды МАИ. 2016. № 91. С. 19.

29. Изделия и комплексы противодействия беспилотным летательным аппаратам [Доклад]. СПб.: АО «НИИ «Вектор», 2018. 51 с.

30. Гейстер С. Р., Джеки А. М. Решение задачи обнаружения маловысотных легкомоторных летательных аппаратов путем использования акустических и сейсмических полей // Наука и военная безопасность. 2008. № 1. С. 42 – 46. URL: http://militaryarticle.ru/nauka-i-voennayabezopasnost/2008/12105-reshenie-zadachi-obnaruzhenija-malovysotnyh (дата обращения 18.09.2022).
31. Сосулин Ю.Г. Теоретические основы радиолокации и радионавигации : учеб. пособие для вузов. Москва : Радио и связь, 1992. 304 с.
32. Карташов В.М. и др. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы. Харьков : ХНУРЭ, 2014. 312 с.
33. Ситнік О.В., Карташов В.М. Радіотехнічні системи : навч. посібник. Харків : Сміт, 2009. 448 с.
34. Shirman Y.D., Manzhos V.N. The theory and technique of processing radar information against the background of interference. Москва : Radio and communications, 1981. 416 p.
35. Koch W., Koller J., Ulmke M. Ground target tracking and road map extraction // ISPRS J. Photogramm. Remote Sens. 2006; 61:197–208. doi: 10.1016/j.isprsjprs.2006.09.013.
36. Hengy S., Laurenzis M., Schertzer S., Hommes A., Kloeppel F., Shoykhetbrod A., Geibig T., Johannes W., Rassy O., Christnacher F. Multimodal UAV detection: Study of various intrusion scenarios // Proceedings of the Electro-Optical Remote Sensing XI International Society for Optics and Photonics. Warsaw, Poland. 11–14 September 2017. p. 104340P.
37. Laurenzis M., Hengy S., Hammer M., Hommes A., Johannes W., Giovanneschi F., Rassy O., Bacher E., Schertzer S., Poyet J.M. An adaptive sensing approach for the detection of small UAV: First investigation of static sensor network and moving sensor platform // Proceedings of the Signal Processing, Sensor/Information Fusion, and Target Recognition XXVII International Society for Optics and Photonics; Orlando, FL, USA. 16–19 April 2018. p. 106460S.
38. Park S., Shin S., Kim Y., Matson E.T., Lee K., Kolodzy P.J., Slater J.C., Scherreik M., Sam M., Gallagher J.C., et al. Combination of radar and audio sensors for identification of rotor-type unmanned aerial vehicles // Proceedings of the 2015 IEEE SENSORS. Busan, Korea. 1–4 November 2015. P. 1 – 4.
39. Charvat G.L., Fenn A.J., Perry B.T. The MIT IAP radar course: Build a small radar system capable of sensing range, Doppler, and synthetic aperture (SAR) imaging // Proceedings of the 2012 IEEE Radar Conference. Atlanta, GA, USA. 7–11 May 2012; pp. 0138–0144.
40. Liu H., Wei Z., Chen Y., Pan J., Lin L., Ren Y. Drone detection based on an audio-assisted camera array // Proceedings of the 2017 IEEE Third International Conference on Multimedia Big Data (BigMM); Laguna Hills, CA, USA. 19–21 April 2017; pp. 402–406.
41. Басов О.О., Карпов А.А. Анализ стратегий и методов объединения многомодальной информации // Обработка информации и управления. 2015. №2. С.7-14.
42. 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.
43. Atrey P. K., Hossain M. A., Kankanhalli M. S. Multimoda Fusion for Multimedia Analysis: a Survey // Multimedia Systems. 2010. Vol. 16. Iss. 6. P. 345 – 379.
44. Годунов А. И., Шишков С. В., Бикеев Р. Р. Взаимосвязь машинного (технического) зрения с компьютерным зрением при идентификации малогабаритного беспилотного летательного аппарата // Труды междунар. симпозиума «Надежность и качество». 2015. Т. 1. С. 213 – 217.
45. Зайцев А. В., Назарчук И. И., Красавцев О. О., Кичулкин Д. А. Особенности борьбы с тактическими беспилотными летательными аппаратами // Военная мысль. 2013. № 5. С. 37 – 43.

Надійшла до редколегії 04.11.2022

Відомості про авторів:
Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідуючий кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; e-mail: volodymyr.kartashov@nure.ua, ORCID: https://orcid.org/0000-0001-8335-5373
Посошенко Віталій Олександрович – канд. техн. наук, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vitalii.pososhenko@nure.ua; ORCID: https://orcid.org/0000-0003-0867-9161
Рибников Микола Володимирович – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, email: mykola.rybnykov@nure.ua, ORCID: https://orcid.org/0000-0003-1340-8788
Капуста Анастасія Ігорівна – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, e-mail: anastasiia.kapusta@nure.ua, ORCID: https://orcid.org/0000-0003-2206-1552
Першин Євгеній Васильович – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, e-mail: yevhenii.pershyn@nure.ua, ORCID: https://orcid.org/0000-0002-4573-9381

УДК 621.396.96, 621.397.48 **DOI:10.30837/rt.2022.4.211.07**

В.М. КАРТАШОВ, д-р техн. наук, В.О. ПОСОШЕНКО, канд. техн. наук,
К.В. КОЛІСНИК, канд. техн. наук, В.И. КОЛІСНИК, Р.И. БОБНЄВ, А.И. КАПУСТА

АЛГОРИТМ ОЦІНЮВАННЯ РОЗПОДІЛУ ЕНЕРГІЇ РАДІОЛОКАЦІЙНИХ СИГНАЛІВ, ЯКІ РОЗСІЮЮТЬСЯ НА АКУСТИЧНИХ ЗБУРЕННЯХ, СТВОРЕНИХ БПЛА

Вступ

Ряд проблем, пов'язаних з радіолокацією об'єктів з динамічною ефективною поверхнею розсіювання (ЕПР), вимагають пошуку оцінок розподілу енергії корисного сигналу на інтервалі спостереження після розв'язання задачі виявлення конкретного об'єкта в умовах малого співвідношення сигнал/шум і апріорної невизначеності щодо комплексної огинаючої пакету радіовідбиттів.

До таких об'єктів можна віднести атмосферні неоднорідності, які виникають в результаті функціонування безпілотного літального апарату (БПЛА). Звукові хвилі, що породжують такі неоднорідності, при поширенні в атмосфері поступово згасають. При цьому вони постійно взаємодіють з коливаннями, які постійно утворюються. Тому конфігурація звукового поля поблизу БПЛА постійно змінюється, що і зумовлює динамічний характер розсіювання радіохвиль на атмосферних збуреннях, а отже їх ЕПР [1 – 10].

Отримання оцінок розподілу нешумової енергії на інтервалі спостереження дозволяє сформувати додаткові інформаційні ознаки для розпізнання класів виявлених БПЛА, а також для просторового розрізнення окремих БПЛА, які знаходяться на одній дальності у разі їхнього групового застосування.

Алгоритм оцінювання, який розглядається, може знайти застосування при радіолокації протяжних розподілених цілей (наприклад: метеорних утворень у вищих прошарках атмосфери Землі), а також для дослідження атмосферних вітрових рухів.

Постановка задачі

Розглядається адитивна модель суміші пакету розсіяних радіосигналів та шуму виду

$$\dot{Y}(t) = \dot{S}(t) + \dot{N}(t), \text{ где } \dot{S}(t) = \sum_{i=1}^{k} \dot{S}_i(t) = \sum_{i=1}^{k} \dot{b}_i \cdot \dot{X}_i(t), \quad (1)$$

де $\dot{S}_i(t) = \dot{b}_i \dot{X}_i(t)$ – i-й відносно початку пакету імпульсний сигнал у пачці, що містить k імпульсів; $\dot{b}_i$ – довільний амплітудний множник i-го імпульсу, який потрібно оцінити; $\dot{X}_i(t)$ – відомий з точністю до фази опорний сигнал, який відповідає зондуючому радіоімпульсу: $x_i(t) = \mathrm{Re}\{\dot{X}_i(t) e^{j\omega_o t}\}$ [11].

Для відносно великих співвідношень сигнал/шум $q \gg 1$ оцінка розподілу енергії сигналу на інтервалі спостереження з прийнятною похибкою здійснюється шляхом визначення модулів коефіцієнтів $\dot{b}_i$. Для малих значень $q \approx 1$ подібний підхід є неприйнятним, оскільки похибка оцінювання стає співрозмірною з модулем максимального значення амплітуди пакету відбитих сигналів $\dot{b}_{imax}$ і навіть перевищує його. У цьому випадку розподіл нешумової енергії слід шукати у вигляді розподілу блоків незмінного рівня її оцінки E_A на однакових інтервалах часу $T_{\text{бл.}} = l \cdot T_{\text{з}}$, де $T_{\text{з}}$ – період зондуючих радіосигналів, використовуючи принцип статистичного накопичення енергії всередині кожного блоку. Причому величина коефіцієнту l – зворотно пропорційна значенню співвідношення q.

Таким чином, потрібно сформулювати алгоритм обробки радіолокаційних сигналів, розсіяних на акустичних збуреннях, створених БПЛА, який повинен працювати у реальному масштабі часу і надавати оцінки розподілу корисної енергії на інтервалі спостереження з деталізацією, яка залежить від поточного співвідношення сигнал/шум.

Алгоритм оцінювання

Одна з важливих вимог до алгоритму оцінювання полягає в тому, щоб він був реалізований у єдиному методологічному плані з алгоритмом виявлення. Такий підхід дозволяє отримати оптимальне поєднання апаратного та програмного забезпечення функціонування цих алгоритмів та створює передумови для організації їх роботи в реальному масштабі часу.

Алгоритм виявлення БПЛА, описаний у [11], дозволяє реалізувати цей підхід. Робота цього алгоритму заснована на тій обставині, що адитивна суміш корисного сигналу з шумом виду (1) являє собою випадковий вузькосмуговий процес, для якого квадратурні складові $Y_c(t)$ і $Y_s(t)$ повного вектору $\dot{Y}$(t) мають нормальний розподіл для різних типів випадкової величини $\dot{Y}$(t) [11]. У цьому випадку оцінка енергії на інтервалі спостереження у разі приведення її до дисперсії шуму являє собою реалізацію випадкової величини $\hat{\xi}$, яка розподілена за центральним або нецентральним законом χ^2_N з N ступенями свободи і чисельною оцінкою параметра нецентральності λ [12].

При цьому, якщо всі вибірки огинаючої вузькосмугового випадкового процесу, які є нормованими за дисперсією шуму, розбити на k блоків, розташованих один до одного в хронологічному порядку, то можна записати: $\hat{\xi} = \sum_{j=1}^{k}(\xi_j)$, де ξ_j – реалізація випадкової величини, яка має розподіл χ^2_L з L ступенями свободи, де L=N/k та певними параметрами нецентральності λ_i такими, що $\sum_j^k(\lambda_j)$=λ.

Таким чином, після прийняття рішення про виявлення БПЛА на конкретній дальності ми фіксуємо набір з k випадкових величин, що мають розподіл χ^2_M з однаковим числом ступенів свободи M, але різними в загальному випадку параметрами нецентральності λ_i. У цьому зв’язку привабливою є ідея безеталонного оцінювання параметра, що характеризує сукупність об'єктів у припущенні, що закон розподілу цього параметра є відомий (заданий) [13]. Ця ідея базується на методах теорії порядкових статистик [14, 15], які вимагають замість процедури порівняння шуканого параметра об'єкта з певним зразком упорядкування вибірки з кількох об'єктів.

Розглянемо деякий гіпотетичний випадковий процес $\dot{Y}_{\text{г}}(t)$ виду (1), який являє собою адитивну суміш шуму $\dot{N}(t)$ та гіпотетичного пакету корисних сигналів $\dot{S}_{\text{г}}(t)$. Нехай $\dot{S}_{\text{г}}(t) = \dot{b}_{\text{г}} \cdot \dot{X}(t)$. Таким чином, гіпотетичний сигнал $\dot{S}_{\text{г}}(t)$ є еквівалентним з точністю до постійного множника $\dot{b}_{\text{г}}$ опорному (зондуючому) сигналу $\dot{X}(t)$. Причому значення параметра $\dot{b}_{\text{г}}$ виберемо так, щоб оцінка $\xi_{\text{г}} = \xi$, тобто повинна виконуватися рівність: $\sum_{j=1}^{k}(\xi_{j\text{г}}) = \sum_{j=1}^{k}(\xi_j)$. Сенс розгляду пакету $\dot{S}_{\text{г}}(t)$ полягає у моделюванні умови безперервного та рівномірного надходження нешумової енергії на інтервалі спостереження (0; $T_{\text{н}}$). Потім результати, отримані за допомогою моделі $\dot{S}_{\text{г}}(t)$, поширимо на алгоритм оцінювання реального пакету корисних сигналів $\dot{S}(t)$.

Оцінимо величину математичного очікування $M[\xi]$ випадкової величини ξ, як результат операції усереднення суми k реалізацій $\xi_{j\text{г}}$, $j = \overline{1,k}$:

$$\widehat{M}[\xi] = \frac{1}{k} \cdot \sum_{j=1}^{k}(\xi_{j\text{г}}).$$

З іншої сторони:

$$\widehat{M}[\xi] = \frac{1}{k} \cdot \sum_{j=1}^{k}(\xi_j) = \hat{\xi}.$$

Оскільки згідно з [12] для випадкової величини ξ, що має розподіл χ^2_N з параметром нецентральності λ, перші два моменти мають вигляд: $M[\xi] = N + \lambda$, $D[\xi] = 2 \cdot N + 4 \cdot \lambda$, то $\hat{\xi}$/k= $L + \lambda_{\text{г}}$. Звідки оцінка $\lambda_{\text{г}}$: $\lambda_{\text{г}} = (\hat{\xi} - N)/k$, а оцінка загального параметра нецентральності $\hat{\lambda} = k \cdot \lambda_{\text{г}}$.

Утворимо з гіпотетичних локальних сум $\xi_{j\text{г}}$ варіаційний ряд:

$$\hat{\xi}_{\text{г}}(j), j = \overline{1,k},\ \text{де}\ \hat{\xi}_{\text{г}}(1) \le \hat{\xi}_{\text{г}}(2) \le \ldots . \hat{\xi}_{\text{г}}(k-1) \le \hat{\xi}_{\text{г}}(k). \quad (2)$$

Знаючи параметр нецентральності $\lambda_{\text{г}}$, число ступенів свободи L (а отже, знаючи аналітичний вираз для щільності ймовірності $P_{\text{н}}(\xi)$ [16] нецентрального розподілу χ^2_L, число k членів варіаційного ряду), можна знайти значення математичного очікування M_j та дисперсії D_j для кожної j-ї порядкової статистики з варіаційного ряду (2) відповідно до таких виразів [14]:

$$M_j = \int_{-\infty}^{\infty} \xi \cdot \psi_j(\xi) \cdot d\xi,\ D_j = \int_{-\infty}^{\infty} \xi^2 \cdot \psi_j(\xi) \cdot d\xi\ , \quad (3)$$

$$\psi_j(\xi) = \frac{k!}{(j-1)!\cdot(k-j)!} \cdot [F(\xi)]^{j-1} \cdot [1 - F(\xi)]^{k-j} \cdot f(\xi), \quad (4)$$

де $\psi_j(\xi)$ – густина ймовірності j-ї порядкової статистики для варіаційного ряду (2), що містить k елементів; $f(\xi) = P_{\text{н}}(\xi)$ – щільність ймовірності нецентрального розподілу χ^2_L з параметром нецентральності $\lambda_{\text{г}}$; $F(\xi)$ – функція розподілу випадкової величини ξ.

Для широкого кола практично важливих розподілів, серед яких – нормальне та всі усічені (у тому числі χ^2_L при досить великому значенні L) виконується важливе співвідношення:

$$D_k[\,\hat{\xi}_j\,] \ll D[\xi], \text{ де } j = \overline{1,k}\ ,$$

$D_k[\hat{\xi}_j]$ – дисперсія j-ї порядкової статистики; $D[\xi]$ – дисперсія випадкової величини ξ; k – кількість елементів варіаційного ряду.

В табл. 1 – 5 в якості прикладу зведено чисельні дані розрахункових величин M_j та D_j для фіксованих значень параметрів k, $\lambda_{\text{г}}$, $N_{\text{св.}}$ Оцінки M_j та D_j отримано відповідно до виразів (3) з застосуванням аналітичних виразів для щільності ймовірності центрального (у разі $\lambda_{\text{г}} = 0$) та нецентрального розподілу (у разі, коли $\lambda_{\text{г}} \neq 0$) χ^2_n з n ступенями свободи:

$$p_{ц}(\xi) = (1/(2^{n/2}\Gamma(n/2))\xi^{n/2-1}e^{-\xi/2},$$

де Г(х) – гамма-функция Эйлера; $P_{ц}(\xi)$ =0 для $\xi \le 0$;

$$p_{нц}(\xi) = (e^{-0.5(\xi+\lambda)}\xi^{(n/2)-1}\sum_{j=0}^{\infty}(\lambda\xi/4)^j)/(j!\Gamma(j+n/2)))/2^{n/2},$$

для $\xi \ge 0$, где λ – параметр нецентральності.

Таблиця 1

Розрахункові значення математичного очікування M_jта дисперсії D_j порядкових статистик (за виразами (3), (4)) при λ=0, k = 64, $N_{\text{св.}} = 8$

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M_j	1.701	2.222	2.581	2.870	3.120	3.343	3.549	3.740	3.921	4.093	4.259	4.420	4.575	4.728	4.877	5.023
D_j	0.559	0.520	0.501	0.491	0.485	0.482	0.481	0.481	0.481	0.483	0.484	0.487	0.489	0.432	0.496	0.499

j	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
M_j	5.168	5.310	5.452	5.592	5.732	5.871	6.010	6.149	6.288	6.428	6.568	6.709	6.851	6.995	7.140	7.286
D_j	0.503	0.508	0.512	0.516	0.521	0.526	0.531	0.536	0.542	0.548	0.554	0.561	0.568	0.574	0.582	0.590

j	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
M_j	7.435	7.586	7.739	7.895	8.053	8.215	8.381	8.551	8.725	8.903	9.088	9.278	9.475	9.679	9.891	10.113
D_j	0.598	0.606	0.615	0.625	0.635	0.645	0.656	0.667	0.680	0.693	0.707	0.722	0.738	0.755	0774	0.793

j	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
M_j	10.345	10.589	10.847	11.120	11.412	11.726	12.065	12.436	12.847	13.309	13.837	14.458	15.218	16.207	17.639	19.870
D_j	0.815	0.839	0.865	0.894	0.928	0.965	1.008	1.058	1.116	1.188	1.279	1.396	1.556	1.800	2.246	3.985

Таблиця 2

Розрахункові значення математичного очікування M_jта дисперсії D_j порядкових статистик (за виразами (3), (4)) при λ=8, $k = 64$, $N_{св.} = 8$

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M_j	4.093	5.234	5.997	6.600	7.113	7.567	7.979	8.361	8.713	9.057	9.380	9.691	9.991	10.283	10.567	10.845
D_j	1.239	1.113	1.070	1.013	0.909	0.973	0.960	0.952	0.947	0.943	0.940	0.939	9390	0.940	0.942	0.945

j	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
M_j	11.118	11.386	11.651	11.913	12.172	12.430	12.686	12.941	13.196	13.451	13.705	13.981	14.218	14.476	14.735	14.997
D_j	0.948	0.951	0.956	0.959	0.965	0.971	0.976	0.982	0.089	0.997	1.005	1.013	1.022	1.030	1.040	1.051

j	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48
M_j	15.262	15.529	15.800	16.074	16.558	16.637	16.926	17.222	17.524	17.833	18.151	18.477	18.815	19.163	19.524	19.900
D_j	1.062	1.074	1.086	1.100	1.113	1.127	1.144	1.159	1.177	1.197	1.217	1.358	1.260	1.287	1.314	1.343

j	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64
M_j	20.392	20.703	21.135	21.592	22.077	22.597	23.257	23.766	24.437	25.187	26.040	27.037	28.248	28.808	32.033	34.991
D_j	1.374	1.409	1.447	1.490	1.540	1.594	1.657	1.732	1.819	1.924	2.061	2.232	2.468	2.828	3.506	7..130

Таблиця 3

Розрахункові значення математичного очікування M_jта дисперсії D_j порядкових статистик (за виразами (3), (4)) при λ=3, $k = 32$, $N_{св.} = 8$

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M_j	3.012	3.972	4.651	5.211	5.705	6.157	6.582	6.987	7.379	7.762	8.139	8.514	8.889	9.266	10.647	10.035
D_j	1.015	0.960	0.940	0.933	0.934	0.939	0.947	0.958	0.970	0.985	1.001	1.019	1.039	1.060	1.084	1.109

j	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
M_j	10.431	10.839	11.261	11.701	12.162	12.648	13.167	13.725	14.334	15.008	15.768	16.649	17.711	19.069	21.006	24.218
D_j	1.138	1.169	1.203	1.242	1.285	1.333	1.389	1.454	1.531	1.623	1.739	1.889	2.095	2.406	2.061	4.768

Таблиця 4

Розрахункові значення математичного очікування M_jта дисперсії D_j порядкових статистик (за виразами (3), (4)) при λ=3, $k = 16$, $N_{св.} = 6$

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M_j	2.540	3.607	4.434	5.165	5.853	6.525	7.198	7.890	8.614	9.390	10.242	11.207	12.344	13.772	15.775	19.262
D_j	1.076	1.105	1.140	1.181	1.229	1.282	1.343	1.414	1.496	1.594	1.715	1.871	2.082	2.398	2.956	4.427

Таблиця 5

Розрахункові значення математичного очікування M_jта дисперсії D_j порядкових статистик (за виразами (3),(4)) при λ=5, $k = 16$, $N_{св.} = 6$

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M_j	3.324	4.658	5.671	6.557	7.381	8.180	8.976	9.787	10.633	11.534	12.517	13.625	14.923	16.544	18.80	22.668
D_j	1.353	1.361	1.385	1.421	1.465	1.518	1.581	1.654	1.740	1.845	1.974	2.142	2.371	2.713	3.317	4.956

Тому у якості справжніх оцінок $\hat{B}_j$ j-х локальних сум $\hat{\xi}_j$ можна прийняти значення математичних очікувань M_j j-х порядкових статистик $j = \overline{1,k}$ [13]:

$$\hat{B}_j = M_j; \hat{B}_{jmax} = M_{jmax}. \quad (5)$$

При цьому утворюється помилка, значення якої $\delta_j = \xi_j - M_j$, а дисперсія оцінки $\widehat{B}_j$ дорівнює дисперсії j-ї порядкової статистики: $\sigma_{j\text{г}}^2 = D_j$.

Таким чином, упорядковуючи вибірку локальних сум $\xi_{j\text{г}}$ досить великого обсягу, можна, користуючись оцінкою (5), отримати значення $\hat{B}_j$, найкращі в сенсі мінімуму дисперсії помилки для обраного закону розподілу випадкової величини ξ при рівномірному розподілі корисної енергії на інтервалі спостереження $(0; T_{\text{н}})$.

Розглянемо реальний варіаційний ряд $\hat{\xi}_j, j = \overline{1,k}$, складений із експериментально отриманих локальних сум ξ_j. Головна його відмінність від варіаційного ряду (2) полягає в тому, що локальні значення енергії корисного сигналу E_j у загальному випадку заздалегідь нерівномірно розподілені в кожному j-му інтервалі аналізу, який входить до загального інтервалу спостереження. Ця обставина робить неприйнятним алгоритм оцінювання $\hat{B}_j$ (5), оскільки цей алгоритм інваріантний до розподілу сумарних нешумових енергій E_{s1} і E_{s2} пакетів корисних сигналів $\dot{S}_1(t)$ і $\dot{S}_2(t)$ на інтервалі спостереження $(0;\ T_{\text{н}})$ при формальній рівності параметрів нецентральності $(\hat{\lambda}_1 \approx \hat{\lambda}_2)$. Ускладнимо алгоритм (5).

Нехай $d_j = \sqrt{D_j}$ – середньоквадратичне відхилення j-ї порядкової статистики (2), а τ – деякий стандартизований коефіцієнт, який враховує конкретний розподіл випадкової величини ξ. У першому наближенні $\tau \approx 1$.

Тоді, якщо для усіх значень реального варіаційного ряду $(\hat{\xi}_j, j = \overline{1,k})$ виконується умова

$$M_j - d_j \cdot \tau \le \xi_j \le M_j + d_j \cdot \tau\ , \tag{6}$$

можна стверджувати, що всі проміжні інтервали аналізу (на яких отримані оцінки ξ_j) містять приблизно однакову енергію E_A нешумового сигналу $\dot{S}(t)$:

$$E_{Aj} = E_A = const, j = \overline{1,k}. \tag{7}$$

Значення E_A є пропорційним деякому середньому параметру нецентральності λ_j також однакового для кожного з k проміжних інтервалів аналізу:

$$E_A{\sim}\lambda_j;\ \lambda_j = \lambda_A = const, j = \overline{1,k}; \tag{8}$$

$$E_\Sigma = \sum_{j=1}^{k} E_{Aj} = k \cdot E_{A;}\ E_\Sigma{\sim}\xi'. \tag{9}$$

Якщо для довільної порядкової статистики $\hat{\xi}_j$ виконується співвідношення

$$\xi_j < \left(M_j - d_j \cdot \tau\right), \tag{10}$$

то j-й інтервал аналізу містить менше енергії нешумового сигналу, ніж середня оцінка E_A, тобто $E_{Aj} < E_A$. Причому різниця $\Delta E_j = E_A - E_{Aj}$ пропорційна різниці $(M_j - \hat{\xi}_j)$:

$$|\Delta E_j|{\sim}|M_j - \xi_j|. \tag{11}$$

Аналогічно, якщо для i-ї порядкової статистики виконується умова

$$\xi_i > (M_i + d_i \cdot \tau), \tag{12}$$

то i-й інтервал варіаційного ряду містить більше енергії корисного сигналу на величину ΔE_i, ніж усереднена по k інтервалам аналізу оцінка E_A. Тобто

$$|\Delta E_i|{\sim}\left|M_i - \hat{\xi}_i\right| \tag{13}$$

Отже, проміжні інтервали аналізу, котрим виконується умова (10), формують своєрідний «енергетичний фонд» $E_{\text{э}}$. Цей фонд має бути перерозподілений між тими елементами аналізу, для яких виконується нерівність (12) відповідно до різниць (11) та (13). Тому оцінкою енергії $\hat{E}_j$ на інтервалах аналізу, що відповідають елементам варіаційного ряду $(\hat{\xi}_j, j = 1, k)$, можуть бути визначені так:

$$E_j = E_A + (\Delta E_+) \cdot 1\left(\hat{\xi}_j - M_j - d_j \cdot \tau\right) - (\Delta E_-) \cdot 1(M_j - \hat{\xi}_j - d_j \cdot \tau), \quad (14)$$

де

$$\Delta E_+ = \frac{\hat{\xi}_j - M_j}{G} \cdot E_Э;$$

$$\Delta E_- = \frac{M_i - \hat{\xi}_i}{M_i} \cdot E_A;$$

$$G = \sum_{j=1}^{k} [(\hat{\xi}_j - M_j) \cdot 1(\hat{\xi}_j - M_j) \cdot 1(\hat{\xi}_j - M_j - d_j \cdot \tau)];$$

$$E_Э = \sum_{i=1}^{k} [\frac{M_i - \hat{\xi}_i}{M_i} \cdot E_A \cdot 1\left(M_i - \hat{\xi}_i - d_i \cdot \tau\right)];$$

1(X) – функція включення.

Оцінка E_A отримується як функція від параметра $\hat{\lambda}$, який обчислюється за допомогою виразів: $\lambda_г = \frac{\xi'}{k} - L = \frac{\xi' - N}{k}$; $\hat{\lambda} = k \cdot \lambda_г$.

Остаточно

$$E_\Sigma = \varphi(\hat{\lambda}); \quad E_A = \frac{E_\Sigma}{k} = \frac{\varphi(\hat{\lambda})}{k}.$$

Дисперсія D_{Ej} оцінок локальних значень енергій $\hat{E}_j$ за умови великої кількості (k) локальних інтервалів аналізу визначається сумою дисперсій оцінки сумарного параметра нецентральності $D_{\hat{\lambda}}$ та оцінок математичного очікування порядкових статистикD_j:

$$D_{Ej} = D_{\hat{\lambda}} + D_j.. \quad (15)$$

Відповідно до [13 – 15] величина $D_j \ll D_\xi$, де $D_\xi = \sigma_{bi}^2$ – дисперсія оцінки амплітуди одиночного відбитого сигналу (1) при малому співвідношенні сигнал/шум (q). Величина $D_{\hat{\lambda}}$ так само як і σ_{bi}^2 – обернено пропорційна значенню q, але не для одиночного парціального імпульсу $\dot{b}_i \cdot \dot{x}_i(t)$, що входить до пакету радіовідбиттів $\dot{S}(t)$ (1), а для енергії $Э_s$ всього пакету сигналів:

$$D_{\hat{\lambda}} \sim \frac{1}{q_{\hat{\lambda}}} = \frac{1}{\frac{Э_s}{N_0}} = \frac{N_0}{\sum_i Э_i},$$

де $Э_i$ – енергії всіх парціальних сигналів $\dot{b}_i \cdot \dot{x}_i(t)$, що становлять пачковий сигнал $\dot{S}(t)$ згідно з моделлю (1). Тому $D_{\hat{\lambda}} \ll D_\xi$ при виконанні умови $\sum_i Э_i \gg N_0$. Таким чином, у загальному випадку $D_{Ei} \ll D_\xi$.

Використовуючи однозначну відповідність між елементами варіаційного ряду $\hat{\xi}_j$ та елементами хронологічно правильної послідовності ξ_j, легко отримати оцінку впорядкованого за часом розподілу сумарної енергії E_Σ на інтервалі спостереження $T_Н$ за проміжними інтервалами аналізу.

Слід зауважити, що використовувати алгоритм оцінювання так, як його описано вище, недоцільно через великий обсяг чисельних розрахунків, що ускладнює практичну реалізацію даного алгоритму у реальному масштабі часу. Щоб вирішити цю проблему пропонується наступне. Попередньо розраховуються чисельні значення математичного очікування та дисперсії за формулами (3), (4) для різних співвідношень параметрів k, $\lambda_г$, $N_{св.}$ Ці значення заносяться до довготривалої пам'яті, яка сформована у вигляді трьохвимірної матриці. При цьому значення параметру $\lambda_г$ обираються в діапазоні прийнятних величин від $\lambda_{гmin}$ до $\lambda_{гmax}$ з невеликим кроком зміни $\Delta\lambda$, які визначаються експериментально. Значення параметрів k та $N_{св.}$ доцільно обирати у форматі 2^P, що є зручним з точки зору апаратного та програмного втілення запропонованих алгоритмів.

Висновки

Розробка та дослідження методів оцінювання розподілу нешумової енергії на інтервалі спостереження є логічним продовженням зусиль щодо створення алгоритмів виявлення радіосигналів, розсіяних на атмосферних неоднорідностях, які виникають внаслідок функціонування БПЛА. Подібні методи потрібні для уточнення алгоритмів виявлення, класифікації виявлених БПЛА за додатковими інформаційними ознаками, підвищення роздільної здатності при виявленні декількох апаратів, розташованих на одній дальності при груповому застосуванні БПЛА, з'ясування часових параметрів еволюції руху БПЛА у часі та у просторі тощо.

Бажання отримати впевнене виявлення БПЛА та оцінювання його характеристик на максимально можливій дальності призводить до необхідності обробляти корисні радіолокаційні сигнали при малому співвідношенні сигнал/шум. А це, в свою чергу, робить неможливим через велику похибку отриманих оцінок здійснювати процедури оцінювання методом порівняння з еталонами фізичних величин.

В цьому зв'язку перспективним, на наш погляд, є алгоритм оцінювання, заснований на методах теорії порядкових статистик, які передбачають замість порівняння чисельних реалізацій з еталоном формування з них варіаційного ряду при умові апріорного знання функції розподілу цих реалізацій. При цьому використовується той факт, що для певних розподілів випадкової величини, серед яких є нормальний та всі обмежені, дисперсія оцінки у вигляді математичного очікування певної порядкової статистики суттєво менше дисперсії прямого вимірювання при малому співвідношенню сигнал/шум. У запропонованому алгоритмі використовується розподіл χ^2_N, який використано при вирішенні задачі виявлення. Тобто задачі виявлення БПЛА та оцінювання розподілу корисної енергії на інтервалі спостереження вирішено в єдиному методологічному плані.

Для заощадження часу та обчислювального ресурсу при обробці у реальному масштабі часу сигналів, що приймаються, запропоновано використовувати попередньо розраховані масиви чисельних значень математичного очікування та дисперсії порядкових статистик у вигляді матриць, елементи яких розташовано відповідно до значень номеру порядкової статистики j, розмірності варіаційного ряду k, параметру нецентральності λ.

Список літератури:

1. Карташов В.М., Олейников В.Н., Шейко С.А. и др. Особенности обнаружения и распознавания малых беспилотных летательных аппаратов // Радиотехника. 2018. Вып. 195. С.235 – 243.
2. Карташов В.М., Олейников В.Н., Воронин В.В. и др. Методы комплексной обработки и интерпретации радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов // Радиотехника. 2020. Вып. 202. С. 173 – 182.
3. Карташов В.М., Олейников В.Н., Шейко С.А. и др. Информационные характеристики звукового излучения малых беспилотных летательных аппаратов // Радиотехника. 2017. Вып. 191. С. 181 – 187.
4. Красненко М.П. Акустическое зондирование атмосферы. Новосибирск : Наука СО, 1986. 167с.
5. Карташов В.М., Харченко О.И., Посошенко В.А. и др. Обнаружение беспилотных летательных аппаратов с ипользованием рассеяния радиоволн на акустических возмущениях среды, создаваемых летательным аппаратом // Радіотехніка. 2021. Вип. 206.
6. Moses A. Radar-based detection and identification for minia ture air vehicles / A. Moses, M.J.Rutherford, K.P. Valavanis // IEEE International Conference on Control Applications.
7. Даник Ю.Г., Пулеко І.В., Бугайов М.В. Виявлення безпілотник літальних апаратів на основі аналізу акустичних та радіолокаційних сигналів // Вісник ЖДТУ. 2014. № 4 (71). С.71 – 80.
8. Карташов В.М., Посошенко В.О., Воронин В.В. и др. Методы обнаружения-распознавания радиолокационных, акустических, оптических и инфракрасных сигналов беспилотных летательных аппаратов // Радиотехника. 2021. Вып. 205. С.138 – 153.
9. Карташов В.М. Модели и методы обработки сигналов систем радиоакустического и акустического зондирования атмосферы. Харьков : ХНУРЭ, 2011. 234 с.
10. Карташов В.М. Функции рассеяния сигналов систем зондирования атмосферы // Радиотехника. Харьков. 2001. Вып. 118. С.61 – 65.
11. Карташов В.М., Посошенко В.А., Колесник В.И. и др. Обнаружение радиолокационных сигналов, рассеянных на акустических возмущениях, создаваемых БПЛА // Радиотехника. 2021. Вып. 207. С. 113 – 122.

12. Урковиц. Обнаружение неизвестных детерминированных сигналов по энергии // ТИИЭР. 1967. Т.55 №4. С. 50 – 59.
13. Ефимов А.Н., Кутеев В.М. Безэталонные измерения и идентификация методами теории порядковых статистик // Автоматика и телемеханика. 1978. №12. С.30 – 35.
14. Дейвид Г. Порядковые статистики. Москва : Наука, 1973. 335с.
15. Боярский Э.А. Порядковые статистики. Москва : Статистика, 1972. 281с.
16. Королюк В.С., Портенко Н.И., Скороход А.В., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. Москва : Наука. Гл. ред. физ.-мат. лит., 1985. 640 с.

Надійшла до редколегії 12.10.2022

Відомості про авторів:
Карташов Володимир Михайлович – д-р техн. наук, професор, Харківський національний університет радіоелектроніки, завідувач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: volodymyr.kartashov@nure.ua; ORCID: https://orcid.org/0000-0001-8335-5373
Посошенко Віталій Олександрович – канд. техн. наук, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: vitalii.pososhenko@nure.ua; ORCID: https://orcid.org/0000-0003-0867-9161
Колісник Костянтин Васильович – канд. техн. наук, Харківський національний університет радіоелектроніки, доцент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: kolesniknet@ukr.net
Колісник Вікторія Іванівна – Харківський національний університет радіоелектроніки, асистент кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: viktoriia.kolisnyk@nure.ua; ORCID: https://orcid.org/0000-0002-2382-9124
Бобнєв Роман Олександрович – Харківський національний університет радіоелектроніки, старший викладач кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна; email: roman.bobniev@nure.ua; ORCID:https://orcid.org/0000-0002-9322-9722
Капуста Анастасія Ігорівна – Харківський національний університет радіоелектроніки, аспірант кафедри медіаінженерії та інформаційних радіоелектронних систем, Україна, e-mail: anastasiia.kapusta@nure.ua, ORCID: https://orcid.org/0000-0003-2206-1552

УДК 621.396.967.2 **DOI:10.30837/rt.2022.4.211.08**

І.В. СВИД, канд. техн. наук, М.Г. ТКАЧ, І.І. ОБОД, д-р техн. наук

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЗАВАДОЗАХИЩЕНОСТІ РАДІОЛОКАЦІЙНИХ СИСТЕМ ІДЕНТИФІКАЦІЇ «СВІЙ-ЧУЖИЙ»

Вступ

Інформаційне забезпечення системи контролю повітряного простору [1] вирішується головним чином радіолокаційними системами спостереження, до яких відносяться первинні [2 – 4], вторинні [5 – 8] радіолокатори, а також системи радіолокаційної ідентифікації за ознакою «свій-чужий» (IFF) [9 – 16]. В теперішній час існують дві системи радіолокаційної ідентифікації за ознакою «свій-чужий» «Пароль» [17] та MkXIIA [18 – 22]. При цьому слід зазначити, що перша з вказаних систем радіолокаційної ідентифікації за ознакою «свій-чужий» працює в частотному діапазоні, який відрізняється від частотного діапазону роботи систем вторинної радіолокації, то друга працює в частотному діапазоні вторинної радіолокації. Однак, як показано в роботах [23, 24], один з головних інформаційних ресурсів системи контролю повітряного простору побудовано на принципах одноканальної чи двоканальної системи передачі інформації. Це дозволяє зацікавленій стороні несанкціоноване використовувати цей інформаційний ресурс для подальшого визначення координат повітряних об’єктів, з одного боку, та перекручувати інформацію цих інформаційних ресурсів, з другого боку, що призводить до непередбачуваних результатів.

Мета роботи – оцінка завадозахищеності існуючих систем радіолокаційної ідентифікації за ознакою «свій-чужий».

Тенденції розвитку та коротка характеристика систем радіолокаційної ідентифікації за ознакою «свій-чужий»

Система IFF MkXIIA є розвитком раніше впроваджених систем IFF MkXA та IFF MkXII, забезпечуючи розширені можливості ідентифікації «свій-чужий» (IFF) при збереженні сумісності з цивільними системами вторинної радіолокації.

Так, у 1980-х роках було додано новий цивільний режим Mode S, який дозволяв кодувати у сигналі відповіді значно більшу кількість інформаційних даних. Це використовувалося для кодування розташування повітряного об’єкта (ПО), використовуючи дані з навігаційної системи. Це основна частина системи запобігання зіткнень, яка дозволяє комерційним ПО дізнаватися про місцезнаходження інших ПО у цьому районі та уникати їх, не вдаючись до наземних операторів. Розвиток режиму S призвів до створення технології ADS-B [25, 26].

Основні концепції режиму ADS-B було воєнізовано як режим 5, який є просто криптографічною закодованою версією координатних даних режиму S. Новостворена система IFF XIIA є розвитком створених систем IFF MkXA та IFF MkXII, забезпечує розширені можливості процесу ідентифікації «свій-чужий» зі збереження сумісності з цивільними вторинними радіолокаторами спостереження (SSR). STANAG 4193 (Частина III) визначає встановлені характеристики системи IFF MkXIIA. Вимоги, застосовні до будь-якої конкретної категорії обладнання, вказані у стандарті STANAG 4193 (частина I), який описує систему IFF MkXIIA та визначає загальні характеристики системи. Класифіковані характеристики системи IFF MkXIIA визначені STANAG 4193 (частина II). У додатку до STANAG 4193 (Частина III) визначаються відповідні функціональні вимоги та вимоги до характеристик, які застосовуються до обладнання запитувача і відповідача. Встановлені вимоги вважаються мінімально необхідними для забезпечення належних загальних характеристик системи IFF [18].

В даний час авіація країн НАТО оснащена системою IFF MkXIIA, в якій додатково використовується імітостійкий режим Mode 5. Особливостями сигналів режиму Mode 5 є наявність опорної групи імпульсів з часово-імпульсною модуляцією, що визначає режим ідентифікації, та частотна модуляція з безперервною фазою імпульсів преамбули і подавлен-

ня бічних пелюсток 16-розрядною послідовністю коду Уолша, котра використовується при декодуванні інформаційної групи.

Сигнал режиму Mod5 аналітично може бути представлено так:

$$S_{m5}(t) = A\sum_{n=0}^{3} s_1(t-t_n,k)p(t-t_n) + A_p s_1(t-t_p,k)p(t-t_p) + A\sum_{m=0}^{M-1} s_1(t-t_n,k)p(t-t_m),$$

де $p(t-t_n) = \{1 | t_i \in [t_0 + iNT, t_0 + (i+1)NT]; 0 | \notin [t_0 + iNT, t_0 + (i+1)NT]\}$; A та A_p – амплітуди імпульсів преамбули та подавлення бічних пелюсток; t_n та t_p – моменти часу появи імпульсів преамбули та подавлення бічних пелюсток; t_m – моменти часу появи m-го імпульсу з M імпульсів інформаційної групи; $s_1(t,k) = \sum_{i=0}^{N-1} U_m \cos\left(2\pi f_s t + \frac{\pi b_i(k)}{2T} + \varphi_i\right), t_0 \le t \le t_0 + NT$ – частотне представлення імпульсу сигналу запиту (СЗ) частотно-маніпульованою 16-розрядною послідовністю Уолша; U_m – амплітуда сигналу; t_0 – час початку сигналу; $N = 16$ – кількість елементів в сигналі; T – тривалість елемента (посилки) сигналу; f_s – несійна частота сигналу; $b_i(k) = \mathrm{sgn}(wal(k,\theta_i))$ – знак i-го елемента посилки на інтервалі часу $[t_0 + iT,\ t_0 + (i+1)T]$; $wal(k,\theta)$ – функція Уолша за номером k, яка визначена на інтервалі $[0, NT]$; $\theta_i = t_i / NT = i / N, \theta_i \in [0,1]$ – безрозмірний час, в моменти якого формується значення $b_i(k) \in \{-1;1\}$; $\varphi_i = \frac{\pi}{2} - \sum_{j=1}^{i-1} b_j - \frac{(i-1)\pi}{2} b_i + \varphi_s$; φ_s – початкова фаза сигналу; $\beta = \Delta f T = 0.5$ індекс модуляції; Δf – рознос частот.

Імпульси преамбули можна представити в вигляді

$$g_k(t) = \sum_{n=1}^{L-1} a_k(n)p(t-nT),$$

де $p(t)$ – імпульс тривалості nT; $a_k(n), 0 \le n \le L$ – кодова послідовність, яка визначає розстановку імпульсів преамбули за часом.

Слід зазначити, що в режимі Mode 5 головними є підрежими M5L1 (імітостійка ідентифікація) та підрежим M5L2 (підтримка функцій ADS-B). Різниця підрежимів ідентифікації проводиться аналізом розстановки імпульсів преамбули в часі. Для підрежиму M5L1 потрібна наявність імпульсу P_2 та імпульсу P_1, а для M5L2 – щонайменше три імпульси з чотирьох. При цьому імпульс P_1 є «опорним», щодо якого визначається часове розташування інших. Якщо позначити наявність/відсутність імпульсу логічної змінної P_i, то логіку прийняття рішення про підрежим ідентифікації можна представити у виді

$$\begin{cases} P_{\text{M5L1}} \leftrightarrow P_1 P_2 \\ P_{\text{M5L2}} \leftrightarrow (P_1 P_2 P_3 P_4) \vee (\bar{P}_1 P_2 P_3 P_4) \vee (P_1 \bar{P}_2 P_3 P_4) \vee (P_1 P_2 \bar{P}_3 P_4) \vee (P_1 P_2 P_3 \bar{P}_4) \end{cases},$$

де наявність імпульсів на очікуваних часових позиціях визначається підрежимом ідентифікації.

Антенна система відповідача ПО утворена значною кількістю слабко спрямованих антен, що розширює можливості щодо оптимізації обробки даних відповідачів ПО за часовими та просторовими параметрами.

В даний час широко використовуються дві системи ідентифікації: MkXIIA і «Пароль» [17], які використають різні частотні діапазони. Система IFF «Пароль» реалізована за прин-

ципом запитальних асинхронних безадресних систем [17], а радіолокаційна система IFF MkXIIA відноситься до беззапитальних систем [20].

Розглянемо порівняльну завадозахищеність систем IFF на прикладі існуючих систем ідентифікації ПО за ознакою «свій-чужий».

Під завадостійкістю будемо розуміти здатність інформаційного засобу виконувати свої функції з необхідними показниками якості за наявності ненавмисних і навмисних завад.

Під прихованістю системи будемо розуміти неможливість радіорозвідки визначити місце розташування інформаційного засобу за межами зони його видимості.

Завадостійкість та прихованість сумісно визначають завадозахищеність.

Під можливістю несанкціонованого використання будемо розуміти можливість безпосереднього або шляхом запиту використання сигналів чи інформації розглянутого засобу.

Принцип побудови існуючих запитальних систем IFF зумовив широкі можливості зацікавленої сторони, як у подавлені запитальних IFF в системному плані, так і в отриманні інформації шляхом несанкціонованого включення відповідача ПО [27 – 29].

Використання в якості сигналів запиту і відповіді інтервально-часового та позиційного кодів [30 – 32], тобто використання сигналів без внутрішньоімпульсної модуляції та випромінювання сигналу відповіді слабко спрямованою антенною системою зумовили широке використання відповідача ПО системами радіорозвідки зацікавленої сторони як одного з основних інформаційних джерел.

Нижче покажемо, що такі принципи побудови радіолокаційних систем IFF обумовлюють низьку завадозахищеність існуючих радіолокаційних систем IFF [33 – 35], тобто невисоку завадостійкість та практично відсутність енергетичної прихованості роботи відповідача ПО [36 – 38].

Оцінка завадостійкості запитальних систем радіолокаційної ідентифікації за ознакою «свій-чужий»

Мережа запитальних радіолокаційних систем IFF за принципом побудови відноситься до несинхронних, що означає, по-перше, відсутність синхронізації за часом випромінювання СЗ окремих запитувачів та, по-друге, відсутність будь-якої часової синхронізації між різними системами ідентифікації.

Відповідачі запитальних радіолокаційних систем IFF за принципом обслуговування сигналів запиту відносяться до систем масового обслуговування (СМО) з обслуговуванням першого правильно прийнятого СЗ.

За принципом побудови відповідачі відносяться до одноканальних СМО з відмовами. Суть цих систем полягає в тому, що при обслуговуванні заявки (СЗ) відповідач радіолокаційних систем IFF закривається на певний час, який називають часом паралізації. Величина часу паралізації залежить від режиму роботи. Наявність часу паралізації відповідача ПО обмежує завадостійкість, як самого відповідача, так і радіолокаційних систем IFF у цілому. Оскільки відповідачі обслуговують будь-який правильно прийнятий запит (навіть імітований зацікавленою стороною), то вони належать до відкритих СМО. Наявність тільки одного каналу обслуговування СЗ та паралізація відповідача ПО при обслуговувані СЗ дозволяє віднести відповідач ПО до одноканальної відкритої СМО з відмовами. Будемо враховувати, що завадостійкість запитальної системи IFF характеризує імовірність виявлення повітряного об’єкту радіолокаційною системою IFF P_s, яку визначають як імовірність отримання потрібного числа сигналів відповіді на запити запитальної радіолокаційної системи IFF, яка розглядається. Розглянемо залежність цього показника P_s від інтенсивності завад.

Наземний радіозапитувач прийме сигнал відповіді від відповідача ПО тоді і тільки тоді, якщо одночасно відбудуться дві події:

- відповідач ПО прийме, правильно декодує сигнал запиту і сформує сигнал відповіді (імовірність цієї події дорівнює завадостійкості відповідача ПО P_0);
- сигнал відповіді відповідача ПО прийме та виявить наземний радіозапитувач.

Розглянемо імовірності цих двох подій при наявності завад та проаналізуємо імовірність одночасного їх виконання.

Будемо вважати, що сумарний потік завад утворюється потоком СЗ сусідніх систем IFF, потоком навмисних корельованих завад від зацікавленої сторони та потоком хаотичних імпульсних завад (навмисних та ненавмисних некорельованих завад). Розрахунки проведемо для сумарного потоку сигналів неімітостійких та імітостійких режимів роботи радіолокаційних систем IFF за ознакою «свій-чужий» і для існуючих алгоритмів виявлення ПО. Цей алгоритм полягає в проведенні ідентифікації ПО на першому етапі в неімітостійкому режимі, і за позитивним результатом здійснюється перехід до другого етапу – ідентифікації за ознакою «свій-чужий» в імітостійкому режимі. Крім того, при оцінці завадостійкості будемо оцінювати вклад різних небажаних ситуацій у сумарну оцінку завадостійкості систем IFF.

Проведемо оцінку завадостійкості відповідачів запитальних систем IFF при спільній дії потоків сигналів запиту та хаотичної імпульсної завади.

Вплив потоку СЗ призводить, як показано вище, до паралізації літакового відповідача на час паралізації, який визначається режимом запиту. Зазначимо, що при прийманні СЗ за основною пелюсткою діаграми спрямованості антени запитувача відповідач ПО повністю паралізується на час обслуговування, а при прийманні СЗ за бічними пелюстками діаграми спрямованості антени запитувача відповідач ПО паралізується на час між імпульсом СЗ, амплітуда якого запам’ятовується, і імпульсом подавлення бокових пелюсток. Хаотична імпульсна завада (ХІЗ) (навмисна або ненавмисна) впливає на роботу відповідача ПО:

- по-перше, подавлює окремі імпульси СЗ, що робить неможливим обслуговування даного СЗ;
- по-друге, паралізує роботу відповідача ПО через утворення хибних СЗ (хибна тривога першого і другого роду).

Оцінимо завадостійкість відповідача ПО при впливі вказаних завад. При надходженні на вхід відповідача ПО потоків СЗ і ХІЗ відповідач не сформує сигнал відповіді, якщо станеться хоча б одна з таких несприятливих ситуацій:

– СЗ даного запитувача подавиться через утворення з ХІЗ випереджаючих хибних СЗ (хибна тривога першого роду), які призводять до випромінювання сигнал відповіді або спрацьовування схеми подавлення бокових пелюсток (імовірність P_1);

– сигнал запиту даного запитувача подавиться через випереджаючі СЗ сусідніх запитувачів або запитувачів зацікавленої сторони (імовірність P_2);

– окремі імпульси коду запиту даного запитувача подавляться на високій частоті через збіг за часом імпульсів різних СЗ при несприятливих фазових співвідношеннях (імовірність P_3);

– СЗ даного запитувача подавиться через випереджаючі хибні СЗ, що утворюються при взаємодії першого імпульсу СЗ даного запитувача з випереджаючими (на базу коду) імпульсами ХІЗ або СЗ (хибна тривога другого роду) і призводять до випромінювання сигнал відповіді або спрацьовування схеми подавлення бокових пелюсток (імовірність P_4);

– СЗ даного запитувача подавиться через появу на позиції сигналу хибного імпульсу подавлення, який утворився з завад (імовірність P_5);

– сигнал запиту подавиться через спрацьовування схеми часової селекції відповідачів (імовірність P_6);

– СЗ подавиться через інерційність схем вхідних формувачів дешифратора і обмеження завантаження відповідача (імовірність P_7).

Визначимо, що імовірність цих подій в припущенні, що потоки СЗ і ХІЗ впливають на коди запитів даного запитувача незалежно один від одного і кількість джерел, що формують загальний потік СЗ, достатньо велика для характеристики потоку як пуасонівського.

Нехай на вхід відповідача поступають:

– потік ХІЗ інтенсивністю λ_0;

– потік СЗ інтенсивністю λ_1, який включає потоку СЗ сусідніх запитувачів і потоку СЗ, імітованих зацікавленою стороною;

– потік СЗ, які викликають спрацьовування схеми подавлення бічних пелюсток, інтенсивністю λ_2.

Припустимо, що тривалість імпульсів потоку СЗ однакова, незмінна за часом і збігається з тривалістю імпульсів корисного сигналу. Припустимо також, що загальні потоки СЗ складаються з k частин неімітостійкого режиму та $(1-k)$ частин імітостійкого режиму.

Сумісна дія ХІЗ і потоку СЗ призводить до високочастотного подавлення окремих імпульсів потоку СЗ при несприятливих фазових співвідношеннях, внаслідок чого інтенсивність потоку СЗ зменшується.

Імовірність того, що хоча б один імпульс ХІЗ збіжиться за часом з імпульсом потоку СЗ і подавить його, становить:

$$P_p = \gamma\left(1-e^{-\lambda_0\tau_0}\right),$$

де γ – коефіцієнт інтерференційного подавлення, який визначає імовірність інтерференційного подавлення імпульсу прийнятого сигналу запиту при його збіжності за часом з імпульсом завади.

Через високочастотне подавлення зменшується інтенсивність потоку СЗ, які викликають випромінювання сигнал відповіді $\lambda_1^1 = \lambda_1\left(1-P_p\right)^n$ та інтенсивність потоку СЗ, які викликають спрацьовування схеми подавлення бокових пелюсток $\lambda_2 = \lambda_2\left(1-P_p\right)^n$, де n – значність коду сигналів запиту.

Імовірність того, що хоча б один СЗ потрапить в випереджаючий інтервал і подавить СЗ даної радіолокаційної системи IFF за рахунок часу паралізації t_1 відповідача ПО в неімітостійкому режимі при випромінюванні сигнал відповіді, визначається відповідно:

від ХІЗ: $P_1^1 = 1-e^{-\lambda_x t_1}$, від потоку СЗ: $P_1^2 = 1-e^{-k\lambda_1 t_1}$,

де λ_x – середня кількість хибних n-імпульсних кодів, що призводять до випромінювання сигналу відповіді; $k = \lambda_н / \lambda_1$ – відносна частка неімітостійкого режиму в загальній інтенсивності потоку СЗ; $\lambda_н$ – інтенсивність потоку СЗ неімітостійкого режиму.

Середню кількість хибних n-імпульсних кодів, які призводять до випромінювання сигналу відповіді, можна визначити за формулою

$$\lambda_x = n\tau_0^n\lambda_0^{n-1}\left(1-\tau_s/\tau_0\right),$$

де τ_S – задана тривалість селекції імпульсів за часом.

Імовірність того, що хоча б один СЗ потрапить в випереджаючий інтервал і подавить СЗ даної запитальні системою IFF за рахунок часу паралізації t_2 відповідача ПО в імітостійкому режимі при випромінюванні сигнал відповіді, визначається відповідно:

від ХІЗ: $P_1^3 = 1-\exp\left(-\lambda_x t_2\right)$, від потоку СЗ: $P_1^4 = 1-\exp\left[-\left(1-k\right)\lambda_1 t_2\right]$.

Результуюча імовірність подавлення СЗ даного запитувача системи через паралізацію відповідача при випромінюванні сигнал відповіді складає

$$P_1 = 1-\prod_{i=1}^{4}\left(1-P_1^i\right).$$

Тут і далі розрахунки проведено за умови, що інтенсивність λ_1 потоку СЗ, випромінюваних за бічними пелюстками діаграми спрямованості антени запитувача, в три рази переви-

щує інтенсивність λ_0 потоку СЗ, випромінюваних за основним пелюсткою діаграми спрямованості антени запитувача.

Імовірність P_2 того, що хоча б один СЗ попаде в випереджаючий інтервал і подавить СЗ даної радіолокаційної системи IFF за рахунок часу паралізації t_3 відповідача ПО при спрацьовуванні схеми подавлення бокових пелюсток в неімітостійкому режимі, визначається відповідно:

від ХІЗ: $P_2^1 = 1 - e^{-\lambda_x t_3}$, від потоку СЗ: $P_2^2 = 1 - e^{-k\lambda_2 t_3}$.

Імовірність того, що хоча б один СЗ попаде в випереджаючий інтервал і подавить СЗ даної радіолокаційної системи IFF за рахунок часу паралізації t_4 відповідача ПО при спрацьовуванні схеми подавлення бокових пелюсток в імітостійкому режимі, визначається відповідно:

від ХІЗ: $P_2^3 = 1 - e^{-\lambda_x t_4}$, від потоку СЗ: $P_2^4 = 1 - e^{-(1-k)\lambda_2 t_4}$.

Результуюча імовірність подавлення СЗ даного запитувача радіолокаційної системи IFF через паралізацію відповідача при прийманні СЗ по бічних пелюстках діаграми спрямованості антени запитувача становить:

$$P_2 = 1 - \prod_{i=1}^{4} \left(1 - P_2^i\right).$$

Імовірність подавлення одного будь-якого імпульсу СЗ даного запитувача через збіжність з імпульсами потоків ХІЗ і СЗ становить

$$P_{10} = \gamma \left[1 - e^{-\lambda_s \tau_0}\right],$$

де $\lambda_s = \lambda_0 + \lambda_1^1 + \lambda_2^1$ – інтенсивність сумарного потоку завад та СЗ.

З урахуванням n імпульсів СЗ імовірність подавлення сигналу запиту складає

$$P_3 = 1 - \left(1 - P_{10}\right)^n.$$

Імовірність P_4 подавлення СЗ даного наземного радіозапитувача через появу випереджаючих хибних кодів запиту, що утворюються в результаті взаємодії першого імпульсу коду запиту з випереджаючими імпульсами потоку СЗ і призводять до випромінювання сигналу відповіді або спрацьовування схеми подавлення бічних пелюсток, визначається співвідношенням

$$P_4 = \left(1 - P_{01}\right)^n \left[1 - \left(1 - P_{10}\right)^{n+1}\right].$$

Другий співмножник враховує можливі ситуації утворення хибних випереджаючих кодів запиту: n кодів запиту, що призводять до випромінювання коду відповіді, і одного коду сигналу подавлення, який призводить до спрацьовування схеми подавлення бокових пелюсток.

Імовірність хибної тривоги другого роду P_{01} визначається за формулою

$$P_{01} = 1 - e^{-\lambda_s \tau_0}.$$

Імовірність P_5 подавлення запиту даного запитувача через появу на позиції сигналу хибного імпульсу подавлення, який утворився з завад, визначається за формулою

$$P_5 = \left(1 - P_{10}\right)^n P_{01}^{n-1}.$$

Імовірність P_6 подавлення СЗ в результаті спрацьовування схем часової селекції відповідачів визначається співвідношенням

$$P_6 = 1 - e^{-2\lambda_s \tau_0}.$$

Імовірність P_7 подавлення кодів запиту через інерційність вхідних формувачів відповідача ПО визначається за формулою

$$P_7 = 1 - \left(1 - P_f\right)^n,$$

де $P_f = 1 - e^{-\lambda_s \tau_f}$ – імовірність подавлення одного імпульсу коду через інерційність формувача.

З аналізу ймовірностей несприятливих ситуацій, що призводять до подавлення СЗ наземного запитувача, видно, що максимальною є імовірність P_1 подавлення СЗ даного запитувача радіолокаційної системи IFF через утворення з ХІЗ випереджаючих хибних СЗ. Приблизно втричі меншою є імовірність P_2 подавлення СЗ запитувача радіолокаційної системи IFF через випереджаючі СЗ сусідніх запитувачів або запитувачів зацікавленої сторони. Імовірності всіх інших несприятливих ситуацій є незначними порівняно з імовірністю P_1.

Якщо середня кількість СЗ перевищує припустиму величину завантаження відповідача λ_i, то імовірність відповіді при роботі схеми обмеження завантаження відповідача ПО зменшується і становить: $P_{lv} = \lambda_i / \lambda_3$, де $\lambda_3 = \lambda_1 + \lambda_2$.

Імовірність випромінювання відповіді відповідачем ПО на запит даної запитальної системою IFF становить:

$$\text{при } \lambda_3 < \lambda_M, \quad P_0 = \prod_{i=1}^{7} \left(1 - P_i\right), \qquad \text{при } \lambda_3 > \lambda_M \quad P_0 = P_{lv} \prod_{i=1}^{7} \left(1 - P_i\right).$$

Розрахунки за наведеними виразами подано на рис. 1. При цьому вважали, що інтенсивність потоку ХІЗ $\lambda_0 = 0; 2 \cdot 10^4; 4 \cdot 10^4$, а інтенсивність λ_1 потоку СЗ, які призводять до випромінювання сигналу відповіді, в п’ять разів менше інтенсивності λ_2 потоку СЗ, які викликають спрацьовування схеми подавлення бічних пелюсток.

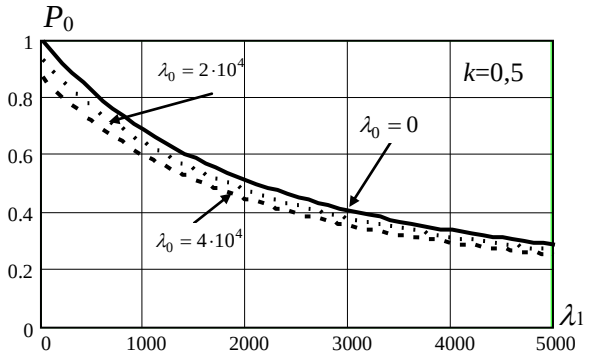


Рис. 1. Оцінка завадостійкості відповідача ПО запитальної системи IFF

З наведених результатів можна зробити такі висновки:

1. Збільшення інтенсивності потоку СЗ призводить до різкого зниження завадостійкості відповідача ПО, що вказує на низьку завадостійкість відповідача ПО (а також систем IFF в цілому). Некорельовані завади (ХІЗ) порівняно слабко впливають на КГ відповідача ПО. Так, при $k = 0,5$ і наявності потоку СЗ $\lambda_1 = 5000$ вплив інтенсивних некорельованих завад

($\lambda_0 = 40000$) призводить до порівняно незначного зниження КГ відповідача ПО з 0,3 до 0,27. Це означає, що найбільш небезпечною для радіолокаційної системи IFF є навмисна корельована завада. Ця обставина дозволяє стверджувати, що основним видом завад при подавлені систем IFF у системному плані є постановка навмисних корельованих завад. З наведеного розрахунку виходить, що інтенсивність потоку СЗ, яка дорівнює 5000, що потребує вилученню 10000 імпульсів, більш ніж на порядок ефективніше за вилучення 40000 імпульсів некорельованої завади.

2. Збільшення частки імітостійкого режиму в загальному потоку СЗ призводить до суттєвого зниження завадостійкості відповідача ПО. Так, при відсутності ХІЗ ($\lambda_0 = 0$) та при інтенсивності потоку СЗ $\lambda_1 = 4000$ збільшення частки імітостійкого режиму з 0,5 до 0,9 призводить до зменшення КГ відповідача ПО майже втричі – з 0,35 до 0,12.

Наведена оцінка завадостійкості показує, що відповідач ПО, як правило, не досягає максимального завантаження. Дійсно, у відповідача ПО для підвищення завадостійкості обмежують максимальну кількість відповідей: в радіолокаційних систем IFF "Пароль" вона збільшена до 3500. Але наведені розрахунки показують, що кількість відповідей відповідача ПО ніколи не сягає такого великого значення. Це вказує на неправильне визначення коефіцієнта завантаження, через що відповідачі ПО існуючих систем IFF не будуть відсіювати СЗ малої потужності. Неправильне визначення максимального завантаження відповідача ПО призводить до зниження завадостійкості відповідача ПО і запитальної системою IFF в цілому. Зацікавлена сторона може несанкціоновано запитувати відповідач ПО і отримувати від нього інформацію або паралізувати його застосуванням завад потрібної інтенсивності за допомогою одного запитувача, розташованого на значній відстані.

Таким чином, наведена оцінка завадостійкості відповідача ПО вказує на низьку стійкість відповідача ПО до впливу навмисних корельованих завад і дозволяє оцінити пропускну здатність існуючих систем IFF.

Оцінка завадостійкості запитальної радіолокаційної системи IFF при сумісній дії потоків сигналів запиту та хаотичної імпульсної завади у каналі запиту та відповіді

Будемо вважати, що апаратура оброблення запитувача реалізує двоетапний алгоритм квазіоптимального виявлення пачки сигналу відповіді, а коефіцієнт готовності відповідача ПО постійний в межах всієї пачки сигналу відповіді. На вхід приймача наземного запитувача надходять незалежні потоки ХІЗ в каналі відповіді інтенсивністю λ_0 та потік сигналу відповіді.

При відсутності завад імовірність виявлення пачки сигналу відповіді запитувачем, тобто повітряного об’єкту радіолокаційними системами IFF, при застосуванні логіки "*k* із *M*" визначається за формулою

$$P_s = \sum_{i=k}^{M} C_M^i P_0^i \left(1 - P_0\right)^{M-i},$$

де $C_M^i = \left(M!/i!(M-i)!\right)$ – біноміальні коефіцієнти.

Наявність ХІЗ в радіоканалі відповіді призводить до подавлення окремих імпульсів сигналу відповіді. Сигнал відповіді не буде прийнятий наземним запитувачем, якщо станеться хоча б одна з таких подій:

– сигнал відповіді відповідача ПО буде подавлений на високій частоті через збіжність за часом імпульсів ХІЗ і сигнал відповіді при несприятливих фазових співвідношеннях (імовірність P_8);

– сигнал відповіді подавиться в результаті інерційності схем оброблення прийнятих сигналів (імовірність P_9).

Імовірність P_8 високочастотного подавлення сигналу відповіді через збіжність за часом імпульсів ХІЗ і сигналу відповіді при несприятливих фазових співвідношеннях, за аналогією з імовірністю P_3, визначається за формулою

$$P_8 = 1-\left(1-P_{10}\right)^n,$$

де $P_{10} = \gamma\left[1-e^{-\lambda_s^{pr}\tau_0}\right]$ – імовірність того, що хоча б один імпульс з потоку сигналу відповіді збіжиться за часом з імпульсом ХІЗ і через несприятливі фазові співвідношення буде подавлений; $\lambda_s^{pr} = \lambda_0 + \lambda_1 P_0$ – інтенсивність сумарного потоку ХІЗ та сигналів відповіді у каналі відповіді.

Імовірність P_9 подавлення сигналу відповіді в результаті інерційності схем оброблення прийнятих сигналів визначається за наступним виразом:

$$P_9 = 1-\left(1-P_f^{pr}\right)^n,$$

де $P_f^{pr} = 1-e^{-\lambda_s^{pr}\tau_f}$ – імовірність подавлення одного імпульсу коду через інерційність схеми оброблення прийнятого сигналу.

Тоді імовірність P_{pr} подавлення сигналу відповіді можливо обчислити як

$$P_{pr} = 1-\prod_{i=8}^{9}\left(1-P_i\right),$$

а імовірність виявлення ПО за пачкою сигнал відповіді запитувачем з урахуванням впливу ХІЗ визначається за формулою

$$P_s = \sum_{i=k}^{M} C_M^i \left(\left[\left(1-P_{pr}\right)P_0\right]^i \left[1-\left(1-P_{pr}\right)P_{pr}\right]P_0\right)^{M-i}.$$

Залежність імовірності виявлення ПО P_s радіолокаційних систем IFF від інтенсивності λ_1 потоку СЗ, яка призводить до випромінювання сигналу відповіді, наведено на рис. 2.

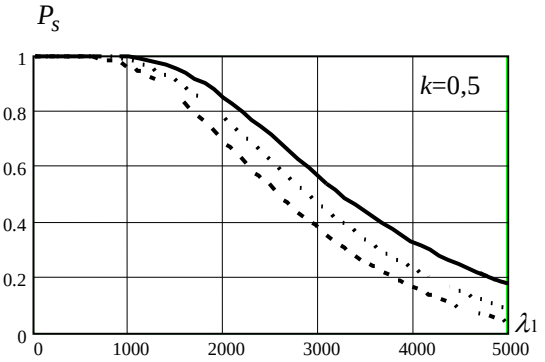


Рис. 2. Залежність $P_s = f\left(\lambda_1, \lambda_0, k\right)$

При розрахунках вважали, що інтенсивність λ_1 потоку СЗ, яка призводить до випромінювання сигналу відповіді, в п’ять разів менше інтенсивності λ_2 потоку СЗ, яка викликає спрацьовування схеми подавлення бокових пелюсток, інтенсивність потоку ХІЗ в каналі відповіді $\lambda_0 = 10^4$. Залежності наведено для різних коефіцієнтів неімітостійкості k та інтенсивності потоку ХІЗ в каналі запиту $\lambda_0 = 0;\ 2\cdot10^4;\ 4\cdot10^4$.

З порівняння рис. 2 видно, що збільшення в сумарному потоку СЗ частки імітостійкого режиму призводить до істотного зниження імовірності виявлення ПО запитальною системою IFF. Так, при $\lambda_1 = 3000$ та $\lambda_0 = 2 \cdot 10^4$ збільшення в сумарному потоці СЗ частки імітостійкого режиму з 0,5 ($k = 0,5$) до 0.9 ($k = 0,1$) призводить до зниження P_s з 0,48 до 0,03. Це означає, що введення імітостійкого режиму дійсно дозволило виключити імітацію сигналу «Я свій» зацікавленою стороною. Але застосування зацікавленою стороною навмисної корельованої завади призведе до паралізації відповідача ПО, що робить неможливим визначення державної приналежності виявлених повітряних об'єктів.

Наведені розрахунки завадостійкості систем IFF показали суттєву залежність імовірності виявлення ПО від наявності та інтенсивності навмисних корельованих завад, імітованих зацікавленою стороною. Невірний вибір коефіцієнта завантаження відповідача ПО радіолокаційної системи IFF «Пароль» дозволяє зацікавленій стороні паралізувати роботу радіолокаційної системи IFF імітованими сигналами запиту малої потужності, тобто на повній відстані, визначеній прямою видимістю.

Оцінка енергетичної прихованості відповідачів ПО запитальних систем радіолокаційної ідентифікації за ознакою «свій-чужий»

Оцінимо енергетичну прихованість відповідача ПО за критерієм дальності виявлення сигналів відповіді. В якості системи радіотехнічної розвідки розглянемо різницево-далекомірну систему, яка включає три приймальних пункти.

Система радіотехнічної розвідки здатна вирішити координатну задачу при виявлені на всіх приймальних пунктах одного імпульсу сигналу відповіді ($n = 1$) або всього сигнал відповіді $n = 2$ чи $n = 3$. Ця можливість закладена у структурі сигналу відповіді, в якості яких, як правило, використовують систему сигналів незначного об'єму. Дійсно, кількість сигналів відповіді існуючих систем IFF відома, тривалість імпульсів, які входять до сигнал відповіді, також відома. Ці обставини дозволяють зацікавленій стороні здійснювати виявлення не тільки окремних імпульсів сигналу відповіді систем IFF, а також і сигнал відповіді у цілому, при використанні, наприклад, багатоканальних виявлювачів, за рахунок апріорно відомих сигналів відповіді.

Позначимо довжину електромагнітних хвиль (ЕМХ) λ, потужність передавача відповідача ПО P, коефіцієнт підсилення антени відповідача ПО G, ефективну площу приймальної антени A, порогову чутливість приймача $P_{pr\,min}$. Відповідач ПО збуджує в розкриві приймальної антени системи радіотехнічній розвідки, розташованої на відстані r від відповідача ПО, густину потоку потужності ЕМХ $S_{pr} = PG / 4\pi r^2$. Потужність сигналу на вході приймача становить (без врахування поляризаційних втрат):

$$P_{pr} = S_{pr} A = PGA / 4\pi r^2 . \tag{1}$$

Для виявлення сигналу відповіді необхідно, щоб відношення сигнал/шум перевищувало пороговий рівень. Відношення сигнал-шум оцінимо за формулою

$$q = \sqrt{P_{pr} / N_0}, \tag{2}$$

де $N_0 = kT\left(K_{ch} - 1\right)$ – спектральна густина потужності шумів, $k = 1,38 \cdot 10^{23}$ Вт/(Гц·К) – стала Больцмана, K_{ch} – коефіцієнт шуму приймача, $T = 290$ – температура.

Дальність виявлення сигналу відповіді (за сигналами Мод 5) типового відповідача ПО системою радіотехнічної розвідки, розрахована за формулами (1), (2), наведена на рис. 3 (для $F = 10^{-6}$). Дальність виявлення розраховувалась для одного імпульсу сигналу відповіді та для сигналу відповіді цілком.

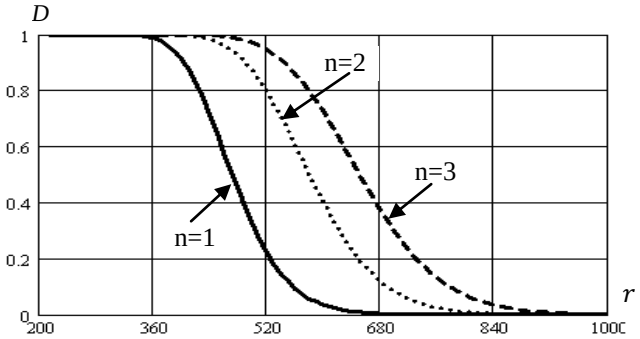


Рис. 3. Дальність виявлення сигналів систем IFF

Наведені розрахунки показують, що системи радіотехнічної розвідки здатні виявити сигнал відповіді сучасних відповідача ПО на великій дальності. Наприклад, дальність виявлення відповідача ПО всього за одним імпульсом сигналу відповіді з імовірністю $D = 0{,}5$ становить 470 км, що відповідає дальності прямої видимості на висоті 12,4 км (при висоті підйому антени 10 м). Таким чином, зона виявлення сигналів відповіді відповідача ПО систем радіотехнічної розвідки, як правило, обмежується відстанню прямої видимості та значно перевищує зони виявлення первинних систем.

Представлені розрахунки дальності виявлення сигналів опорної групи імпульсів з часово-імпульсною модуляцією Mod 5 показують, що наявність у випромінюваних сигналах ідентифікації радіоімпульсів без внутрішньої модуляції доводить, що існуючі системи IFF MkXIIA позбавлені енергетичної прихованості і, як наслідок, не відносяться до завадозахищених. Дійсно, використання сигналів опорної групи, яка включає чотири імпульси преамбули з часово-імпульсною модуляцією, яка визначає підрежим ідентифікації, а також імпульсу подавлення бокових пелюсток дозволяє здійснювати обчислення координат повітряних об’єктів, які використають зазначену систему IFF MkXIIA за рахунок використання технології WAM [39, 40].

Висновки

Розглянуто існуючі системи радіолокаційної ідентифікації за ознакою «свій-чужий» з точки зору оцінки завадозахищеності. Наведений аналіз завадозахищеності існуючих систем радіолокаційної ідентифікації об’єктів за ознакою «свій-чужий», побудованих на принципах запитальних та беззапитальних інформаційних систем, показав, що використання прямокутних радіосигналів з часово-імпульсною модуляцією у якості сигналів запиту та відповіді, які випромінюють повітряні об’єкти, має низьку завадозахищеність та виключає енергетичну прихованість відповідачів повітряних об’єктів і, як наслідок, надає можливість здійснювати несанкціоноване обчислення координат повітряних об’єктів зацікавленою стороною на основі випромінених сигналів ідентифікації за ознакою «свій-чужий».

Список літератури:

1. N. Ntombela and P. Umenne. Access Control with Automated on Duty Notification Tool in air traffic Services // Artificial Intelligence Big Data Computing and Data Communication Systems (icABCD) 2020 International Conference on, pp. 1-5, 2020. doi: 10.1109/icABCD49160.2020.9183828.

2. D. Cohen and Y. C. Eldar. Sub-nyquist radar systems: Temporal, spectral, and spatial compression // IEEE Signal Processing Magazine, vol. 35, no. 6, pp. 35 – 58, 2018. doi:10.1109/MSP.2018.2868137.

3. H. You, X. Jianjuan, and G. Xin. Radar Data Processing with Applications. 2016. doi: 10.1002/9781118956878.

4. X. Li and J. Du. Performance optimization algorithm of Radar Signal Processing System // Cluster Computing, vol. 20, no. 1, pp. 359 –370, 2016. doi: 10.1007/s10586-016-0710-6.

5. Обод И.И. Помехоустойчивые системы вторичной радиолокации. Москва : ЦИНТ, 1998. 118 с.

6. G. Jiang, Y. Fan and H. Yuan. Assessing the Capacity of Air Traffic Control Secondary Surveillance Radar System // 2019 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC), Taiyuan, China, 2019, pp. 1-3, doi: 10.1109/CSQRWC.2019.8799146.

7. 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.

8. I. Svyd, I. Obod, O. Maltsev, I. Shtykh, G. Maistrenko, and G. Zavolodko. Comparative quality analysis of the air objects detection by the Secondary Surveillance Radar // 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO), 2019. doi: 10.1109/ELNANO.2019.8783539.

9. X. Du, K. Liao and X. Shen. Secondary Radar Signal Processing Based on Deep Residual Separable Neural Network // 2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), 2020, pp. 12-16, doi: 10.1109/ICPICS50287.2020.9202372.

10. Свид І. В., Обод І. І. Завадостійкість радіолокаційних систем ідентифікації за ознакою «свій-чужий» : монографія. Харків : Друкарня Мадрид, 2021. 254 с.

11. Obod I., Svyd I., Maltsev O. and Bakumenko B. Comparative Analysis of Noise Immunity Systems Identification Friend or Foe // 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 2020, pp. 751-756, doi: 10.1109/ELNANO50318.2020.9088856.

12. Sharifi-Tehrani O., Sadeghi A. and Razavi S. M. J. Design and Simulation of IFF/ATC Antenna for Unmanned Aerial Vehicle // Majlesi Journal of Mechatronic Systems, vol. 6, no. 1, Jun. 2017.

13. Svyd I., Obod I., Maltsev O., Tkachova T. and Zavolodko G. Improving Noise Immunity in Identification Friend or Foe Systems // 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON), Lviv, Ukraine, 2019, pp. 73-77, doi: 10.1109/UKRCON.2019.8879812.

14. Strelnytskyi O., Svyd I., Obod I., Maltsev O., Voloshchuk O. and Zavolodko G. Assessment Reliability of Data in the Identification Friend or Foe Systems. // 2019 IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO), Kyiv, Ukraine, 2019, pp. 728-731, doi: 10.1109/ELNANO.2019.8783397.

15. 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), 2021, pp. 501-505, doi: 10.1109/PICST54195.2021.9772190.

16. Порівняльний аналіз завадостійкості каналу передачі інформації вторинних радіолокаційних систем / І.В. Свид, І.Ю. Воргуль, С.В. Старокожев и др. // Радіотехніка. 2022. Вип. 208. С. 44 – 54. doi: 10.30837/rt.2022.1.208.05.

17. Маляренко А.С. Системы вторичной радиолокации для управления воздушным движением и государственного радиолокационного опознавания : справочник. Харьков : ХУПС, 2007. 78 с.

18. Technical Characteristics of the IFF Mk XIIA System Part III: Installed System Characteristics. NATO – STANAG 4193 PT III. 2016.

19. H. Duan, Y. Cheng, B. Shen, K. He, and G. Bai. LFM interference cancellation algorithm based on MDPT-WC for mark XIIA mode 5 // 2020 IEEE 20th International Conference on Communication Technology (ICCT), 2020. doi: 10.1109/ICCT50939.2020.9295892.

20. L. Huan, Z. Feng, L. Y. Bai, and W. Jian. One joint demodulation and Despreading algorithm for MOD5 // The Open Automation and Control Systems Journal, vol. 7, no. 1, pp. 386–397, 2015. doi: 10.2174/1874444301507010386.

21. Li Sheng-qiang. Analysis on Data Format of Mode 5 in Western Mark XIIA // Journal of the University of Electronic Science and Technology of China. Vol.40 No.4. 2011.

22. W.-H. Kim, S.-Y. Jung, Y.-S. Lee, and S.-M. Chang. Mark XIIA(Mode 5) IFF system integration and certification test for surface to air missile system // Journal of the Korea Institute of Military Science and Technology, vol. 25, no. 2, pp. 160 – 168, 2022. doi:10.9766/kimst.2022.25.2.160.

23. Обод І.І., Шевцова В.В. Порівняльний аналіз запитальних систем передачі інформації системи контролю повітряного простору // Зб. наук. праць Харків. нац. ун-ту Повітряних Сил. 2013. № 1(34). С. 123-125.

24. Обод І.І., Шевцова В.В. Відносна пропускна спроможність запитальних систем передачі інформації системи контроля повітряного простору // Системи обробки інформації, 2013. № 2(109). С. 74 – 76.

25. H. Li, F. Chen, and J. Wang. A preamble detecting algorithm of MOD-5 interrogating signal // Applied Mechanics and Materials, vol. 543-547, pp. 2733–2737, 2014. doi:10.4028/www.scientific.net/AMM.543-547.2733.

26. J. Sun. The 1090 Megahertz Riddle: A Guide to Decoding Mode S and ADS-B Signals, 2nd ed. TU Delft OPEN Publishing, 2021. doi:10.34641/mg.11.

27. Обод І.І., Свид І.В., Штих І.А. Завадозахищеність запитальних систем спостереження повітряного простору. Харків : ХНУРЕ, 2014. 310 с.

28. Обод І.І., Стрельницький О.О., Андрусевич В.А. Інформаційна мережа систем спостереження повітряного простору. Харків : ХНУРЕ, 2015. 270 с.

29. Бакуменко Б.В., Обод І.І. Завадозахищеність запитувальних радіотехнічних систем // Системи озброєння і військова техніка. 2006. № 2(6). С. 26 – 28.

30. Бакуменко Б.В., Обод І.І. Методи підвищення завадозахищеності запитувальних радіотехнічних систем // Системи обробки інформації. 2006. № 9(58). С. 10 – 12.

31. Порівняльний аналіз методів визначення координат повітряних об’єктів системами широкозонової мультилатерації / І. В. Свид, В. В. Семенець, О. С. Мальцев и др. // Радіотехніка. 2022. Вип. 209. С. 162 – 177. doi: 10.30837/rt.2022.2.209.16.

32. Обод И.И., Абрамов А.Д., Крупка А.В. Пространственная избирательность ответчиков как метод повышения помехоустойчивости запросных радиотехнических систем // Моделювання та інформаційні технології : зб. наук. праць НАНУ. 2005. № 33. С. 103 – 107.

33. Обробка інформації мереж радіолокаційних систем спостереження повітряного простору / І.В. Свид, М.Г. Ткач, А.О. Сєріков и др. // Радіотехніка. 2022. Вип. 210. С. 137-145. doi: 10.30837/rt.2022.3.210.11.

34. Обод І.І., Свид І.В., Черних О.П. Оцінка якості передачі інформації у запитальних каналах передачі систем спостереження повітряного просторум // Восточно-Европейский журнал передовых технологий. Метрология, стандартизация, сертификация. Харків, 2011. № 3/11(51). С. 52 – 54.

35. Обод І.І.,Свид І.В., Штих І.А. Методи підвищення завадозахищеності літакових відповідачів запитальних систем спостереження повітряного простору // Системи обробки інформації. 2015. № 1(126). С. 41 – 43.

36. 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.

37. Обод І.І.,Свид І.В, Мальцев О.С. Обробка даних радіолокаційних систем спостереження повітряного простору: навчальний посібник. Харків : Друкарня Мадрид, 2021. 255 с.

38. I. Svyd, O. Maltsev, I. Obod, and G. Zavolodko. Fusion method of primary surveillance radar data and IFF systems data // 2020 IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT), 2020. doi: 10.1109/DESSERT50317.2020.9125040.

39. D. He, X. Lu, W. Wang and J. Su. Analysis of Wide Area Multilateration Localization Accuracy Under Different Stations Layout ond Aircraft Height // DEStech Transactions on Engineering and Technology Research, 2017, doi: 10.12783/dtetr/iceta2016/7068.

40. I. Obod, I. Svyd, O. Maltsev, G. Zavolodko, and S. Leonov, WAM systems: Comparative Analysis of Information Support Quality // 2020 IEEE International Conference on Problems of Infocommunications. Science and Technology (PIC S&T), 2020. doi: 10.1109/PICST51311.2020.9468085.

Надійшла до редколегії 10.11.2022

Відомості про авторів:

Свид Ірина Вікторівна – канд. техн. 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

Обод Іван Іванович – д-р техн. наук, професор, професор кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: ivan.obod@nure.ua; ORCID: https://orcid.org/0000-0002-9898-0937

УДК 662.396.67: 621.314.6 **DOI:10.30837/rt.2022.4.211.09**

В.О. АЛЄКСЄЄВ, Д.В. ГРЕЦЬКИХ, д-р техн. наук,
Д.С. ГАВВА, канд. техн. наук, В.Г. ЛИХОГРАЙ, канд. фіз.-мат. наук

ТЕХНОЛОГІЇ БЕЗПРОВІДНОЇ ПЕРЕДАЧІ ЕНЕРГІЇ

Вступ

Безпровідна передача енергії (БПЕ) має вже свою порівняно довгу історію і стала одним з актуальних напрямків в науці і техніці, що стрімко розвивається. Початком цьому послужив винахід В. Брауном ректени, яка знайшла успішне застосування в різних галузях БПЕ. До основних з них відносяться: передача енергії сфокусованим мікрохвильовим променем на Землю зі сонячних космічних електростанцій, або на стратосферні ретранслятори, а потім на Землю та наступне її перетворення в постійний струм ректенами; енергопостачання БПЛА (від малорозмірних до стратегічних, у тому числі й висотних ретрансляційних платформ) сфокусованим мікрохвильовим променем з поверхні Землі, де ректена є енергоустановкою цих БПЛА; передача енергії сфокусованим мікрохвильовим променем до важкодоступних об'єктів (об'єкти можуть розміщатися високо в горах, на островах морів і великих озер, а також в інших місцях, до яких з технічних, економічних або інших причин створення повітряних, надводних, підземних і підводних ліній передачі електроенергії неможливе або недоцільне). Освоєння більш високочастотних діапазонів радіо- та оптичного випромінювань, впровадження нових технологій в області мікро- і наноелектроніки, розвиток безпровідного зв'язку стимулювало появу нових напрямків БПЕ і дозволило по-новому поглянути на існуючі. В даний час можна вже виділити ряд технологій БПЕ, що стрімко розвиваються і відрізняються між собою за технічною реалізацією систем БПЕ, вирішуваними ними задачами, діапазоном робочих частот та режимом роботи їх передавальних і приймальних підсистем.

Умовно стаття складається з трьох частин. У першій частині наведено літературний огляд різних технологій БПЕ, а саме – їх особливостей, галузей застосування та тенденцій подальшого розвитку. У другій частині наведено результати досліджень, проведених у Харківському національному університеті радіоелектроніки (ХНУРЕ) в напрямку створення математичної моделі (ММ), що дозволяє проводити з єдиних позицій аналіз та оптимізацію систем БПЕ, в яких використовуються різні технології передачі енергії. У третій частині наведено нові результати, пов'язані з перевіркою адекватності розробленої колективом ХНУРЕ нелінійної ММ електродинамічного рівня системи БПЕ шляхом порівняння результатів розрахунків щодо розробленої моделі з відомими експериментальними даними.

1. Види технологій безпровідної передачі енергії

Безпровідна передача енергії – це загальний термін для ряду різних технологій передачі енергії за допомогою електромагнітних полів (ЕМП) [1, 2]. Під технологією передачі енергії будемо розуміти сукупність різних методів, що використовують ЕМП для передачі енергії [1, 3, 4]. Одними з чинників, що визначають вибір тієї або іншої технологій передачі енергії (рис. 1), є відстань, на яку передається енергія (передача енергії в ближній зоні, зонах Френеля та Фраунгофера [5]) та вид використовуваної електромагнітної енергії (ЕМЕ).

У ближньому полі (ближня зона) енергія передається на короткі відстані за допомогою змінних магнітних полів з використанням індуктивного зв'язку між витками проводу, або змінними електричними полями (рис. 2) з використанням ємнісного зв'язку між металевими електродами. Порядок робочих частот таких систем БПЕ – Гц – МГц.

У проміжному та дальньому полі (зона Френеля та зона Фраунгофера) енергія передається електромагнітними хвилями (ЕМХ) радіо- або оптичного діапазону (у даній статті системи БПЕ і їхні компоненти оптичного діапазону [6 – 13] не розглядатимуться).

Рис. 1. Технології безпровідної передачі енергії

На рис. 3 схематично пояснюється суть технології БПЕ сфокусованим мікрохвильовим променем у зоні Френеля (енергія за допомогою спеціального радіопередавального пристрою передається на одній частоті) і технології збору ЕМЕ з навколишнього середовища (відбувається збір ЕМЕ з навколишнього середовища, яка створюється радіопередавальними пристроями радіоелектронних систем (РЕС) того або іншого призначення та подальше перетворення її в постійний струм).

Загальним для всіх технологій БПЕ є те, що вони засновані на перетворенні високочастотних ЕМП у постійний струм або на частоту електромережі (50/60 Гц). Різницею між ними є діапазон робочих частот, природа первинних джерел енергії, відстань передачі енергії та пристрої для передачі, приймання й перетворення електромагнітної енергії, що використовуються в системах БПЕ. Таким чином, узагальнену схему системи БПЕ можна подати як на рис. 4.

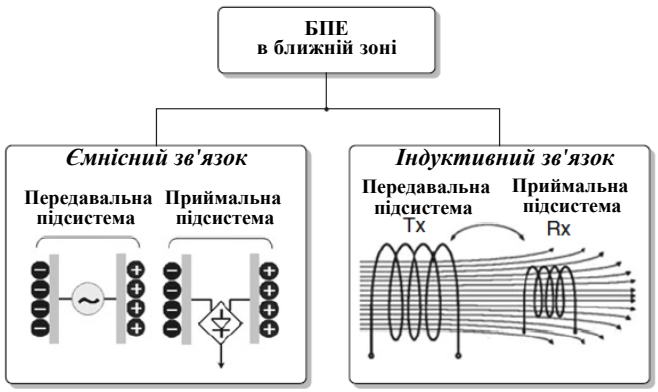
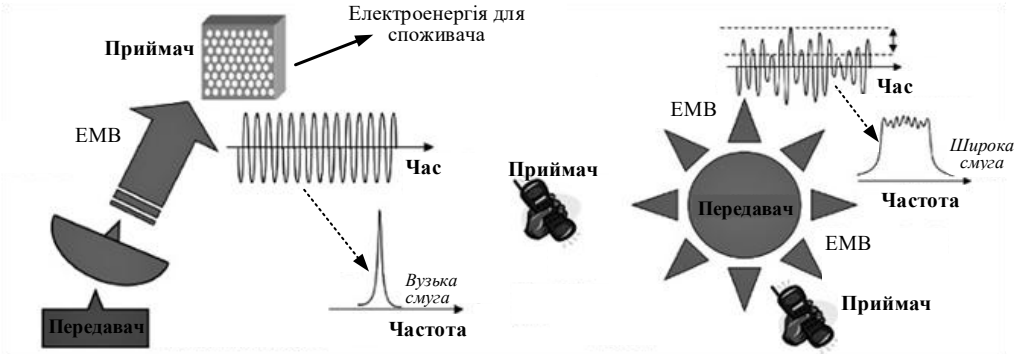


Рис. 2. Способи БПЕ в ближній зоні



а *б*

Рис. 3. Технологія БПЕ сфокусованим мікрохвильовим променем у зоні Френеля – *а* та збір енергії з навколишнього ЕМП у зоні Фраунгофера – *б*

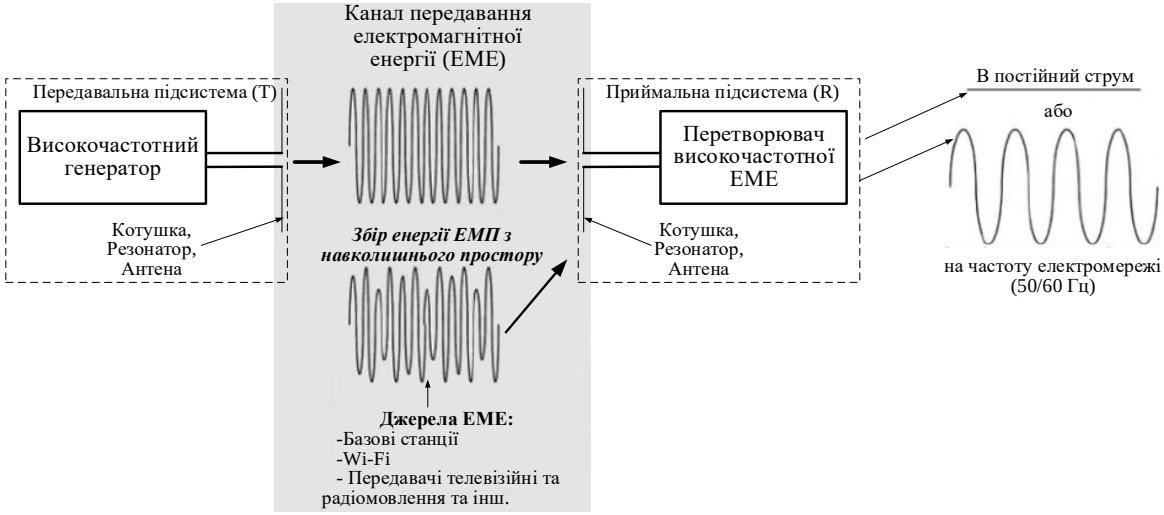


Рис. 4. Узагальнена структурна схема системи БПЕ

2. Галузі застосування технологій БПЕ

2.1. Безпровідна передача енергії в ближній зоні

За останнє десятиліття дослідницький інтерес до передачі енергії в ближній зоні сильно зріс. Розробляються, удосконалюються та серійно випускаються безпровідні зарядні пристрої для мобільних телефонів, планшетів, смарт-годинників та інших пристроїв побутової електроніки [14]. Застосування безпровідної передачі енергії має багато переваг при зарядці акумуляторів електромобілів [15 – 19]. Велику увагу з боку як наукового середовища, так і промисловості приваблюють системи підводного безпровідного енергопостачання глибоководних безпілотних апаратів (рис. 5) [20]. Інтелектуальні системи моніторингу здоров'я, які можна носити або приєднувати до тіла людини (рис. 6), вважаються наступним поколінням персональних портативних пристроїв для дистанційної медичної практики [21]. Щоб забезпечити довгострокову роботу передових біоелектронних пристроїв, які імплантуються у людське тіло, актуальним є рішення завдань щодо організації їхнього безпровідного живлення [22 – 27].

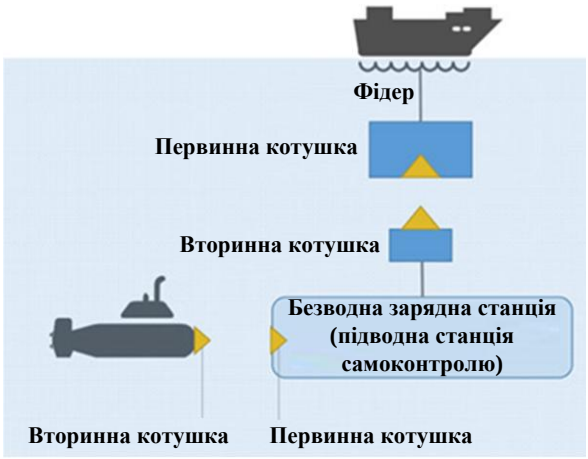


Рис. 5. Система БПЕ для живлення безпілотних підводних апаратів

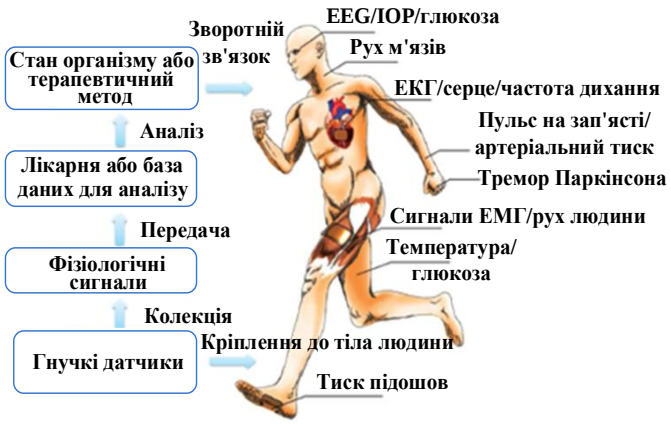


Рис. 6. Медичні датчики на тілі людини

2.2. Безпровідна передача енергії в зоні Френеля

Ідея безпровідної передачі енергії мікрохвильовим променем отримала практичне втілення у 60-і роки XX ст. у зв'язку з розвитком радіолокації, освоєнням мікрохвильового діапазону хвиль та у зв'язку з винаходом В. Брауном антени-випрямляча (ректени) [28, 29]. Результатом його робіт стало створення нового класу енергетичних систем – систем БПЕ мікрохвильовим променем. Такі системи складаються з передавальної підсистеми, завданням якої є перетворення енергії первинного джерела в енергію сфокусованого ЕМВ, та приймальної підсистеми у вигляді ректени, що розташована у зоні Френеля (рис. 3, *а*). Завданням ректени є приймання і перетворення сфокусованого ЕМВ в постійний струм, що надходить до споживача енергії.

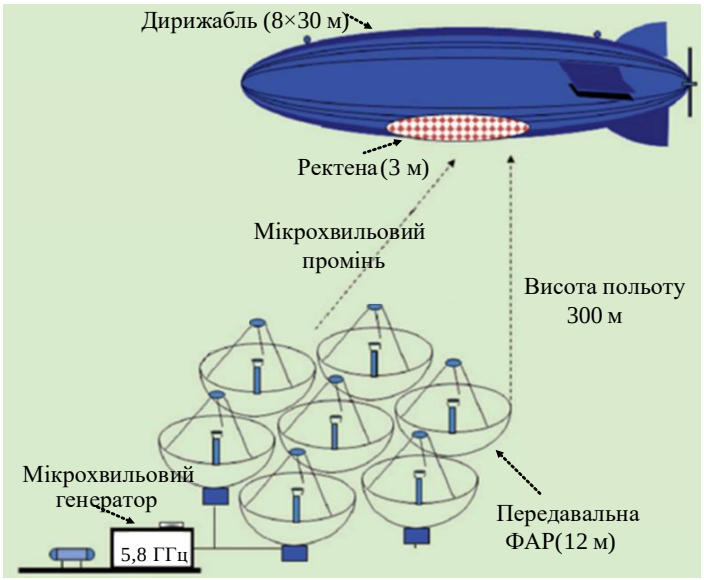


Рис. 7. Демонстрація польоту дирижабля, що живиться мікрохвильовим променем [47]

Галузі застосування систем БПЕ мікрохвильовим променем різноманітні. По-перше це створення альтернативних джерел енергії – сонячних космічних електростанцій (СКЕС), енергія з яких повинна передаватися сфокусованим мікрохвильовим променем і прийматися наземними ректенними системами [1, 30 – 34]. В [35] зазначена гостра необхідність у найближчій перспективі почати практичну реалізацію проектів СКЕС, зокрема, відзначені серйозні наміри США та Японії створити вже незабаром (до 2025 р.) потужні СКЕС. По-друге це безпровідна передача енергії до важкодоступних наземних об'єктів [36 – 41]. По-третє, дистанційне енергопостачання об'єктів, які знаходяться тривалий час в повітрі (рис. 7) [42 – 51] та ін.

2.3. Безпровідна передача енергії в зоні Фраунгофера (збір ЕМЕ з навколишнього простору)

В останні роки активно провадяться дослідження технології збору енергії ЕМП в навколишньому середовищі [52] та перетворення її у постійний струм. Збір енергії ЕМП можна розділити на два типи: збір зовнішньої радіочастотної (РЧ) енергії та спеціальний збір РЧ енергії. Збір зовнішньої РЧ енергії, пов'язаний з джерелами, наявними в навколишньому середовищі (базові станції GSM, Wi-Fi, цифрове телебачення, радіомовлення, тощо (рис. 8)). Якщо навколишні джерела ЕМП не забезпечують достатньої кількості енергії для задоволення вимог додатків, то використовується спеціальний збір необхідної кількості енергії за допомогою створення додаткових джерел (передавач радіочастотної енергії (рис. 8)). Як приклад застосування останнього підходу можна навести запропоновану у [53, 54] систему контролю підвіски автомобіля на основі технологій БПЕ. Однак, як зазначено у [52], цей вид спеціальної системи збору енергії може призвести до великих витрат, особливо у випадку великомасштабних мереж, де може знадобитися встановлення великої кількості таких джерел.

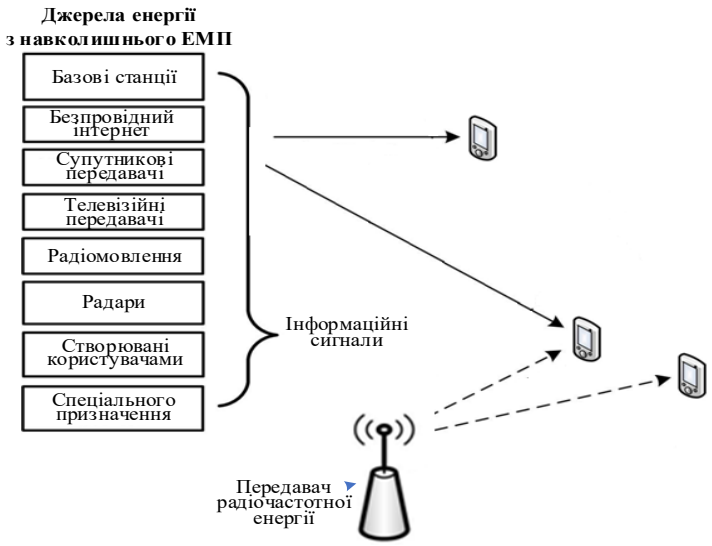


Рис. 8. Енергопостачання малопотужних пристроїв за рахунок видобування енергії з навколишнього ЕМП

На рис. 9 наведено приклад побудови типового вузла мережі датчиків з безпровідним живленням. Енергія, що вилучена з навколишнього простору, може використовуватися безпосередньо або накопичуватися і зберігатися для подальшого використання, що дозволяє реалізувати альтернативні джерела енергії для тих місць, де немає енергетичних систем, або виникають труднощі енергоживлення різного характеру.

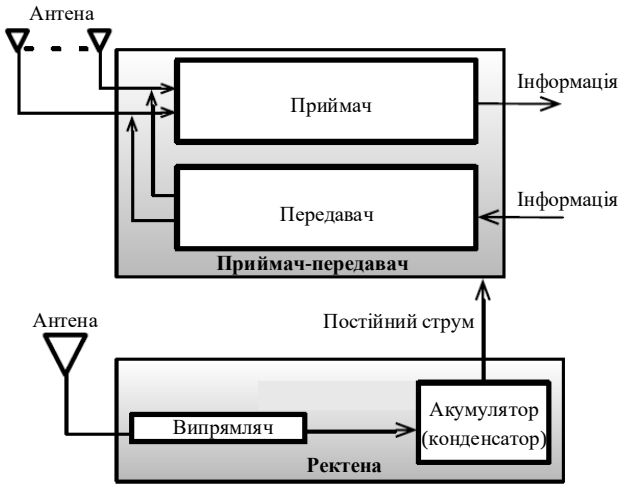


Рис. 9. Архітектура вузла з безпровідним живленням

Актуальність вирішення завдань

збору енергії пов'язана з тим, що більшість електронних пристроїв, таких як датчики в промислових, комерційних і медичних додатках (наприклад, моніторингу забруднення повітря, лісових пожеж, контролю стану різних механізмів, устаткування та будівельних споруд і т.п.), безпровідні пристрої та інші схеми з низьким енергоспоживанням, живляться від батарей. Однак навіть дуже якісні батареї мають обмежений термін використання. Заміна батарей стає "дорогим задоволенням", коли у віддалених місцях знаходяться сотні датчиків. Крім цього виникають нагальні екологічні проблеми, пов'язані з утилізацією відпрацьованих батарей. Застосування ж технологій збору ЕМЕ, що створюється РЕС різного класу та призначення, практично не потребує технічного обслуговування цього устаткування є економічно ефективним (застосовуючи технології збору енергії, пристрої та обладнання можуть стати самоокупними по відношенню до енергії, необхідної для роботи, тим самим забезпечується практично необмежений термін експлуатації) і привабливим з точки зору екологічної безпеки.

Ряд виробників вже пропонують для ринку комерційні рішення систем збору енергії [55 – 58]. Експериментальні оцінки енергетичних параметрів системи збору РЧ енергії, які зроблені деякими дослідниками, узагальнені в табл. 1. Бачимо, що дана технологія має найменшу енергоємність у порівнянні з іншими. Оскільки схеми збору енергії призначені для роботи з відносно невеликими напругами та струмами, вони покладаються на сучасну елементну базу для досягнення високої ефективності перетворення ЕМП у постійний струм.

Таблиця 1

Потужність передавача	Частота	Відстань	Потужність у навантаженні ректени
4 В [59]	902–928 МГц	15 м	5.5 мкВт
1,78 В [60]	868 МГц	25 м	2.3 мкВт
1,78 В [55]	868 МГц	27 м	2 мкВт
3 В [61]	915 МГц	5 м	189 мкВт
3 В [61]	915 МГц	11 м	1 мкВт
960 кВ (ТВ-вишка) [62]	674–680 МГц	4,1 км	60 мкВт

Технологія збору РЧ енергії має безліч практичних застосувань. Коротко розглянемо деякі з них.

Інтернет речей (IoT) (англ. Internet of Things) [63] прокладає шлях до повсюдних послуг у різних сферах життєдіяльності, однак при цьому виникають проблеми в досягненні енергоефективної роботи пристроїв IoT [64]. В [65] підкреслюється актуальність застосування технології збору енергії в додатках IoT, розглядаються різні джерела енергії та методи її збору. Робота [66] присвячена розрахунку вузлів IoT на надзвичайно низькі споживання потужності за рахунок використання джерел енергії з навколишнього простору. Автори описали сучасний стан справ у даній галузі й дали чітке уявлення про методи енергозбереження та подальшої стратегії розвитку «зеленого» IoT. В [67] автори виступають за те, що збір енергії з навколишнього середовища може бути єдиним життєздатним варіантом для продовження терміну служби великомасштабних взаємозалежних мереж IoT.

Автоматизація промисловості. Енергія є ключовим чинником розвитку Індустрії 4.0 шляхом інтеграції обчислювальних процесів з фізичними процесами в інтелектуальному виробничому середовищі. У такому середовищі використовуються різні інтелектуальні пристрої для моніторингу фізичних виробничих процесів, які постійно задіяні в енергоємних операціях [68]. Енергоємна операція для цих пристроїв, які в основному живляться від батарей [69], є критичним місцем їх енергоефективної роботи. Автори роботи [69] пропонують усувати цю критичну ланку за рахунок застосування технології збору енергії з навколишнього простору. В [70] запропонована концепція оптимізації виробничих операцій у рамках розумної промисловості шляхом збору РЧ енергії в декількох місцях виробництва. Зібрана енергія подається у мікроелектромеханічні системи (МЕМС) і датчики, які встановлені на виробничій лінії. Таким чином, система збору енергії може зменшити трудовитрати на технічне обслуговування та експлуатацію, пов'язані із частим

перезарядженням акумуляторів або заміною батарей. У рамках такої ж концепції автори в [71] описують ряд технологій збору енергії, що підходять для автоматизації промисловості, досліджують енергоспоживання невеликих сенсорних пристроїв, розгорнутих у виробничому цеху, а потім оцінюють потенціал збору енергії, який можна використати в промислових процесах.

Практична схема збору та керування радіочастотною енергією для безпровідних сенсорних мереж (БСМ) на основі поліпшеного алгоритму маршрутизації з високою енергоефективністю запропонована в [72]. Автори брали до уваги необхідне енергоспоживання сенсорних вузлів, доступність РЧ енергії, оцінювали статистику зміни рівня густини потоку потужності, запас зібраної енергії в контексті побудови БСМ. Подібні дослідження проводилися й у [73], де були викладені основні вимоги та принципи енергоживлення пристроїв БСМ шляхом збору РЧ енергії. В [74] наведений огляд різних джерел енергії з навколишнього середовища для живлення пристроїв БСМ. Автори роботи стверджують, що рівень потужності 10 – 100 мкВт хоч і малий, але цілком достатній для реалізації ряду додатків на основі БСМ.

Медична інформатика. Веб-рішення для охорони здоров'я в сполученні з інтелектуальними системами на основі IoT забезпечують універсальний доступ до даних для лікарів та лабораторій по всьому світу [75]. У цій прикладній галузі були численні реалізації вузлів натільних комп'ютерних мереж (англ. body area network, безпровідна натільна комп'ютерна мережа WBAN) [76 – 79] і медичних пристроїв, що носять на тілі людини [80 – 82], включаючи системи для моніторингу стану здоров'я пацієнтів і безпровідних інтелектуальних механізмів ін'єкцій [76 – 79]. Завдяки інтеграції можливості збору РЧ енергії, малопотужні медичні пристрої можуть надавати дані по запиту в режимі реального часу. В [76] розроблена конструкція системи енергопостачання датчика WBAN, який складається з невеликої тридіапазонної антени та тридіапазонного випрямляча з ефективністю перетворення 59 % при вхідній потужності -10 дБм. Запропонований датчик компактний і підходить для самоконтролю тіла людини. В [80] наведено всебічний огляд як наукової літератури, так і наявних у продажі пристроїв збору РЧ енергії для живлення медичних пристроїв, що носять на тілі людини.

Радіочастотна ідентифікація (RFID) (англ. Radio Frequency IDentification) – є однією зі зрілих безпровідних технологій малого радіуса дії, у якій дані кодуються в цифровому виді в невеликих радіотранспондерах (також відомих як мітки або смарт-мітки). RFID належить до технологічного сімейства автоматичної ідентифікації та збору даних, де мітки не повинні перебувати в межах прямої видимості для зчитування даних.

Виявлено безліч робіт, у яких використовується збір РЧ енергії в різних додатках RFID [83 – 90]. В [83] докладно описується ряд різних міток RFID і відзначається необхідність збору РЧ енергії для живлення активних міток RFID. В [84] збір РЧ енергії визначений як джерело «зеленої» енергії, що підходить для ряду сенсорних додатків у суворих умовах. Всебічний огляд досягнень щодо реалізації RFID-датчиків з особливим акцентом на збір РЧ енергії наведений в [85]. Авторами роботи [83] описана трифазна система збору енергії, у якій використовується дводіапазонна антена (900 МГц, 2,45 ГГц). Ефективність перетворення РЧ енергії в постійну напругу 30 % на частоті 2,4 ГГц. У роботі [87] обговорюють конструкцію ректени для RFID. Ректена перетворює у постійний струм електромагнітну енергію, створювану джерелами, що працюють на частоті 2,45 ГГц. Автори у [89] описують систему збору РЧ енергії для RFID-міток з ефективністю перетворення 50 %.

Розумні будинки та моніторинг стану конструкцій. Розумний будинок зазвичай ставиться до структури, що використовує автоматизовані процеси для керування його роботою (наприклад, опаленням, вентиляцією, освітленням, кондиціонуванням повітря, безпекою й т.п.). Це дозволяє забезпечити зв'язок між будинком і мешканцями для поліпшення умов їхнього життя, підвищення енергоефективності будинку, виявити потенційні ризики, пов'язані з конструкцією будинку [91]. У роботах [92 – 96] обговорюється

перевага збору РЧ енергії для живлення датчиків, розгорнутих у будівельних конструкціях, у порівнянні з високими витратами на обслуговування акумуляторів. Оскільки датчики часто встановлюються в недоступних місцях усередині будинку або конструкції [94]. Автори у [93, 94] демонструють впровадження датчиків як частини кіберфізичних систем [68] для реалізації моніторингу стану конструкції розумних будинків. Мережі датчиків складаються із прототипів LoRaWAN без батареї, які живляться від системи збору РЧ енергії. В [95] розроблена конструкція компактної ректени площею 66 см^2, яка працює на частоті 868 МГц для додатків моніторингу стану конструкцій будинків. В роботі [96] викладені досягнення в галузі збору енергії з навколишнього середовища, яка використовується для живлення датчиків у додатках моніторингу стану конструкції будівель.

3. Параметри систем збору радіочастотної енергії

Розглянемо основні параметри, що характеризують систему збору РЧ енергії (рис.10).

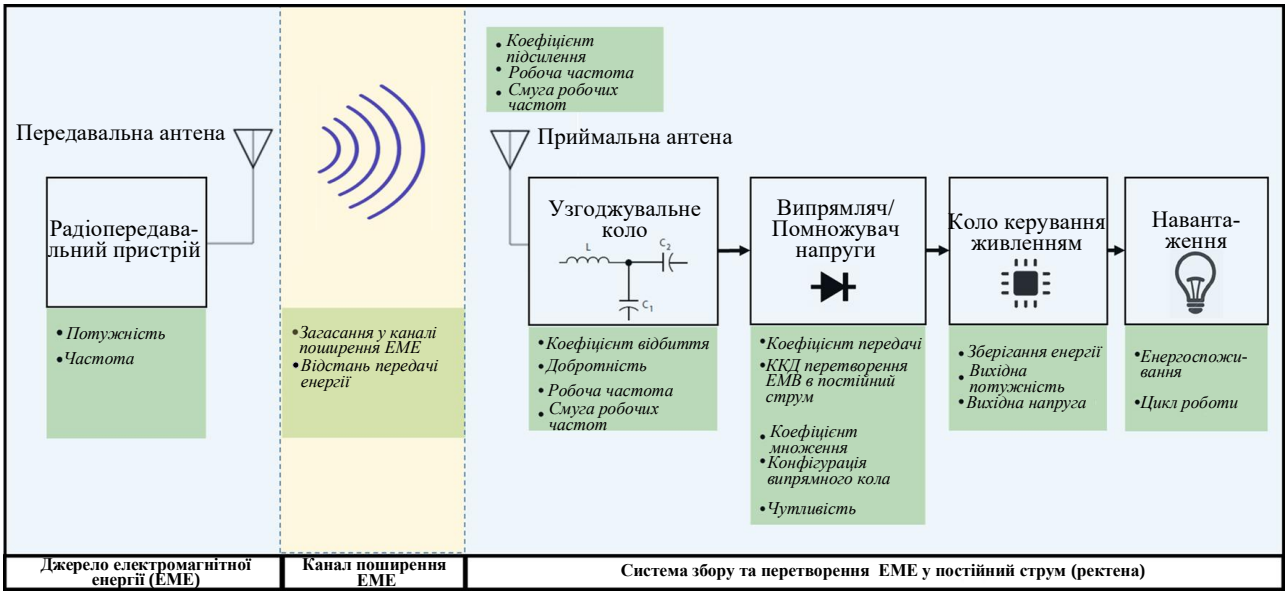


Рис. 10. Структурна схема системи збору електромагнітної енергії

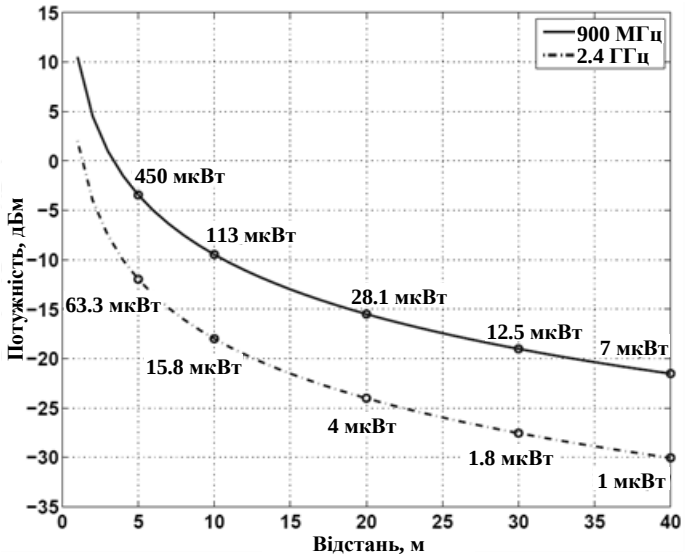


Рис. 11. Вплив відстані на потужність, що приймається при зборі РЧ енергії

Відстань передачі ЕМЕ здебільшого залежить від робочої частоти. Відомо, що загасання при передачі ЕМЕ у атмосфері на високих частотах більше, ніж на низьких частотах (рис. 11 [52]). ЕМХ на низьких частотах проникають глибше у матерію, що необхідно враховувати, якщо технології збору енергії застосовуються для імплантованих пристроїв у тіло людини.

ККД перетворення ЕМЕ в постійний струм (або ККД випрямлення) визначається таким чином

$$\eta_{\text{в}} = \frac{P_0}{P_{\text{вх}}}, \qquad (1)$$

де P_0 – потужність постійного струму в навантаженні ректени; $P_{\text{вх}} = P_{\text{макс}} = \dfrac{e_a^2}{8R_a(f_0)}$ – максимальна потужність, яку може витягти випромінювач ректени з поля падаючої ЕМХ; e_a – амплітуда напруги холостого ходу, що наводиться падаючою хвилею на клемах випромінювача; $R_a(f_0)$ – активна частина вхідного опору випромінювача на робочій частоті f_0.

Рівень побічного випромінювання ректени

$$\xi(nf_0) = \frac{P_\Sigma(nf_0)}{P_{\text{вх}}}, \quad n = 2,3..., \tag{2}$$

де $P_\Sigma(nf_0)$ – потужність, яку випромінює ректена на частоті *n*-ї гармоніки.

Потужність постійного струму в навантаженні ректени відрізняється від $P_{\text{вх}}$ на величину потужності втрат, яка, в свою чергу, складається з потужності втрат в випрямному елементі, потужності втрат в елементах вхідного фільтру (узгоджувальному колі), потужності втрат в елементах вихідного фільтра, потужності, що втрачається за рахунок випромінювання ректенного елемента на частотах вищих гармонік і за рахунок проходження цих гармонік в навантаження, а також потужності, що перевипромінена ректенним елементом на основній частоті. У зв'язку із цим визначальними принципами при розробці ректен є досягнення високого ККД ректени, мінімізація випромінювання на частотах гармонік, простота конструкції, прийнятні масогабаритні показники, низька вартість, надійність і придатність для серійного виробництва. ККД випрямлення залежить від рівня падаючої потужності, опору навантаження, способу включення випрямного діода в схему і т.д. [97, 98]. У [98] відзначено, що на енергетичні характеристики ректен впливають численні фактори і є три взаємопов'язані рівня вирішення проблеми створення якісних ректен: перший – поліпшення параметрів окремих випрямних елементів, другий – оптимізація параметрів і характеристик окремих ректенних елементів, третій – оптимізація характеристик усієї ректени в цілому.

За наявності втрат в резонансних узгоджувальних колах (рис. 10) з часом амплітуда коливань зменшується за експоненційним законом. Швидкість зміни накопиченої енергії в резонансному колі характеризують власною добротністю, яку знаходять за формулами [99]:

$$Q_0 = \omega_0 \frac{W_{\text{нак}}}{P_{\text{втр}}} = 2\pi \frac{W_{\text{нак}}}{\Delta W} \text{ або } Q_0 = \frac{f_0}{\Delta f}, \tag{3}$$

де $W_{\text{нак}}$ – енергія, що накопичена у резонансному колі; $P_{\text{втр}}$ – потужність втрат енергії в резонансному колі; $\Delta W = TP_{\text{втр}}$ – зміна енергії ЕМП в резонансному колі за один період коливань T; f_0 та Δf – частота настройки резонансного (узгоджувального) кола (рис. 10) та його смуга пропускання відповідно.

Для потужних систем БПЕ сфокусованим мікрохвильовим променем вводиться такий параметр як гранично допустимий рівень густини потоку падаючої потужності $П_д$ на ректену, що визначається можливостями випрямного діода Шотткі (допустимими значеннями вхідної потужності $P_д$, прямого струму $I_д$ і зворотної напруги $U_д$, при перевищенні яких діод пробивається). Системи БПЕ малої потужності (системи збору РЧ енергії) потрібно характеризувати таким параметром як чутливість [52]. Чутливість визначається як мінімальна падаюча потужність, яка необхідна для запуску роботи системи збору енергії. Тобто це здатність збирати енергію та працювати з низькою падаючою

на ректену густиною потоку потужності. Чим вища чутливість системи збирання, тим краща ефективність перетворення потужності падаючого на ректену ЕМВ в постійний струм. Чутливість кількісно визначається виразом [52]

$$S_{[\text{дБм}]} = 10\lg\frac{P_{\min}}{1\text{мВт}}, \tag{4}$$

де $P_{\min}$ – мінімальна потужність, необхідна системі для виконання завдання щодо перетворення ЕМВ в постійний струм.

Результатом роботи системи збору ЕМЕ є живлення кінцевих пристроїв постійним струмом. Тому вихідна потужність постійного струму є ще одним показником для оцінки ефективності системи збору ЕМЕ.

4. Математична модель системи БПЕ

У [100 – 106] розроблена нелінійна математична модель (ММ) електродинамічного рівня системи БПЕ, яка використовує поєднання електродинамічного і схемотехнічного підходів та використовує поняття змінних стану. Автори [100 – 106] зазначають, що нелінійна ММ електродинамічного рівня системи БПЕ, на відміну від існуючих, відкриває нові широкі можливості щодо розвитку методів аналізу та оптимізації систем БПЕ з наступних причин:

– є універсальною, тому що дозволяє проводити аналіз і оптимізацію систем БПЕ, в яких використовуються різні технології передачі енергії;

– дозволяє врахувати всю сукупність нелінійних ефектів, які виникають в системах БПЕ через наявність в їх складі антен та трактів живлення з нелінійними характеристиками (передавальні активні фазовані антенні решітки, паразитні нелінійності в передавальних антенах, ректени);

– враховує взаємні зв'язки в самій системі БПЕ (внутрішньо системні процеси [102], а також її електродинамічну взаємодію з іншими РЕС і взаємодію інших РЕС з нею (міжсистемна взаємодія [101], що дозволяє коректно вирішувати задачі електромагнітної сумісності ще на етапі моделювання систем БПЕ, а також вирішувати задачі щодо проєктування РЕС нових класів з одночасною передачею інформації та енергії;

– є гнучкою, тому що при аналізі передбачає можливість зміни конфігурації системи БПЕ в залежності від її призначення (технології передачі енергії) і можливість проведення аналізу окремих її підсистем, пристроїв, вузлів.

Дана ММ ґрунтується на моделі антен з нелінійними елементами (АНЕ), що складається з лінійних (ЛБ) та нелінійних (НБ) багатополюсників. У [100] сформульовано етапи побудови нелінійної ММ електродинамічного рівня системи БПЕ. Розглядалася система БПЕ, в якій передавальна та приймальна підсистеми мають довільні конфігурації (рис. 4) і до складу яких входять антени та тракти їх живлення з НЕ. Зазначено, що в передавальній (T) та приймальній (R) підсистемах системи БПЕ можна виділити нелінійні підсхеми (НПС), лінійні підсхеми (ЛПС), випромінювальні структури (ВС), генератори та навантаження. Показано, що систему БПЕ довільної конфігурації можна уявити як АНЕ. Для цього нелінійні підсхеми передавальної підсистеми НПС_T і приймальної підсистеми НПС_R було об'єднано в нелінійну підсхему НПС_TR, яку названо нелінійною підсхемою системи БПЕ (рис. 12). Лінійні підсхеми передавальної підсистеми ЛПС_T і приймальної підсистеми ЛПС_R об'єднано в лінійну підсхему ЛПС_TR системи БПЕ. Випромінювальну структуру передавальної підсистеми ВС_T і ректени ВС_R об'єднано в підсхему ВС_TR – випромінювальну структуру системи БПЕ, а джерела (генератори) і навантаження в підсхему, яку названо зовнішні пристрої (ЗП_TR) системи БПЕ. Кожній підсхемі у відповідність були поставлені ЛБ та НБ. Для більш повного опису системи БПЕ з погляду її електродинамічної взаємодії з іншими РЕС (міжсистемна взаємодія) на рис. 12 виділені додаткові групи входів (перетини $\delta'_T - \delta'_T$ й $\delta'_R - \delta'_R$), по яких і враховується ця міжсистемна взаємодія.

Символом $\wedge\vee$ відзначена та обставина, що при об'єднанні в один багатополюсник $\text{ВС}_\text{Т}$ і ВС_R та розсіювачів взаємний зв'язок між ними враховується в параметрах (компонентних рівняннях) об'єднаного багатополюсника ВС_TR.

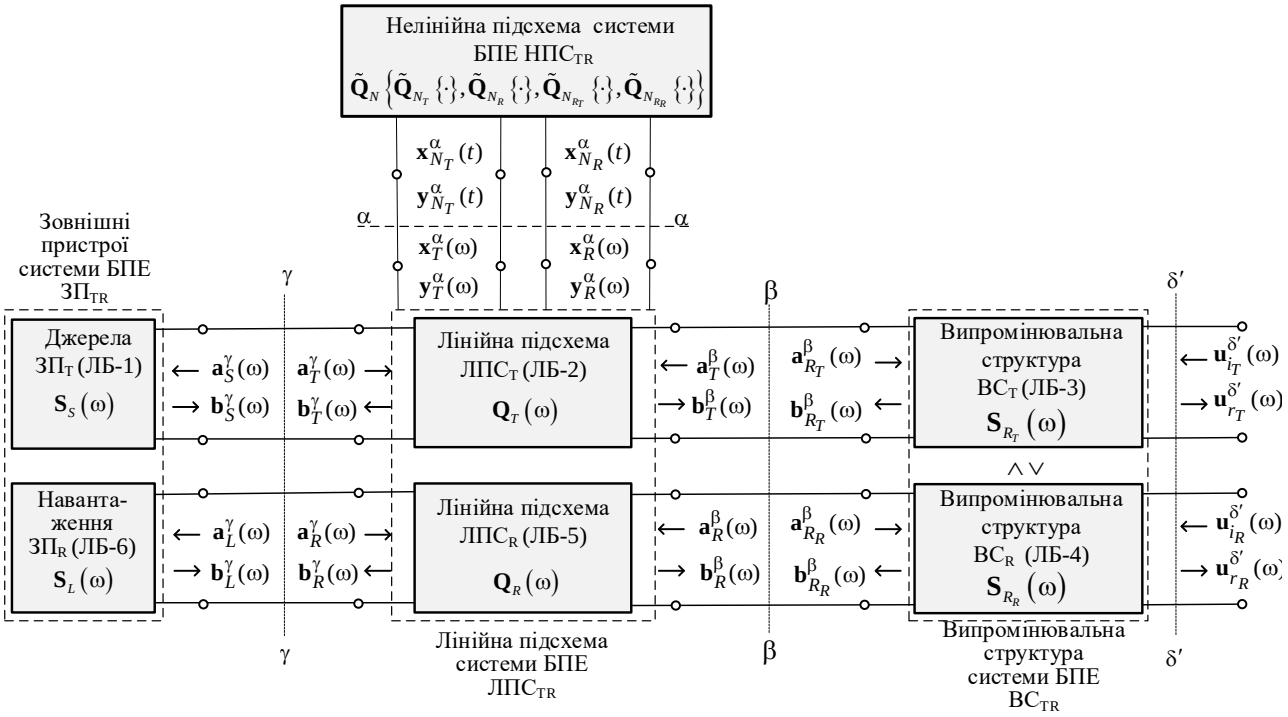


Рис. 12. Структурна схема системи БПЕ

У [100] обґрунтовано, що з метою створення узагальненої моделі системи БПЕ режими входів НБ доцільно описувати в часовій області як векторами струмів $\mathbf{i}^{\alpha}_{N_{**}}(t)$ і напруг $\mathbf{u}^{\alpha}_{N_{**}}(t)$, так і векторами падаючих $\mathbf{a}^{\alpha}_{N_{**}}(t)$ та відбитих $\mathbf{b}^{\alpha}_{N_{**}}(t)$ хвиль. Тому до розгляду введені узагальнені вектори (рис. 12 перетин $\alpha-\alpha$):

$$\mathbf{x}^{\alpha}_{N_{**}}(t)=\left[\mathbf{u}^{\alpha 1}_{N_{**}}(t)\quad \mathbf{i}^{\alpha 2}_{N_{**}}(t)\quad \mathbf{a}^{\alpha 3}_{N_{**}}(t)\right]^T,\ \mathbf{x}^{\alpha}_{N_{**}}(t)=\left[\mathbf{u}^{\alpha 1}_{N_{**}}(t)\quad \mathbf{i}^{\alpha 2}_{N_{**}}(t)\quad \mathbf{a}^{\alpha 3}_{N_{**}}(t)\right]^T,$$

де верхній індекс T – операція транспонування; $\mathbf{u}^{\alpha 1}_{N_{**}}(t),\mathbf{i}^{\alpha 2}_{N_{**}}(t),\ \mathbf{a}^{\alpha 3}_{N_{**}}(t)$ – вектори вхідних впливів на відповідних входах НБ; $\mathbf{i}^{\alpha 1}_{N_{**}}(t),\mathbf{u}^{\alpha 2}_{N_{**}}(t),\ \mathbf{b}^{\alpha 3}_{N_{**}}(t)$ – вектори відгуків на відповідних входах НБ; $**$ відповідає передавальній підсистемі T або приймальній R.

Враховуючи сказане та умови з'єднання в перетинах $\beta-\beta$ і $\gamma-\gamma$ систему БПЕ було подано у вигляді багатовходової АНЕ (рис. 13).

Нелінійна підсхема НПС_TR системи БПЕ описується в часовій області нелінійним діагональним матричним оператором

$$\tilde{\mathbf{Q}}_N\{\cdot\}=diag\left\{\tilde{\mathbf{Q}}_{N_T}\{\cdot\},\tilde{\mathbf{Q}}_{N_R}\{\cdot\}\right\}.$$

Блоки матриці $\tilde{\mathbf{Q}}_N\{\cdot\}$ мають наступну структуру:

$$\tilde{\mathbf{Q}}_{N_{**}}\{\cdot\}=diag\left\{\tilde{\mathbf{G}}_{N_{**}}\{\cdot\},\tilde{\mathbf{R}}_{N_{**}}\{\cdot\},\tilde{\mathbf{S}}_{N_{**}}\{\cdot\}\right\},$$

де $**$ відповідає T або R; $diag\{\cdot\}$ – діагональна матриця, $\tilde{\mathbf{G}}_N\{\cdot\},\tilde{\mathbf{R}}_N\{\cdot\},\tilde{\mathbf{S}}_N\{\cdot\}$ – нелінійні оператори, що задають зв'язок між вхідними впливами (відповідно напругами, струмами,

падаючими хвилями) і відгуками (струмами, напругами, відбитими хвилями, відповідно) на входах НБ ($НПС_{TR}$)

Зовнішні параметри $НПС_{TR}$ пов'язані системою компонентних рівнянь:

$$\mathbf{y}_N^{\alpha}(t) = \tilde{\mathbf{Q}}_N\left\{\mathbf{x}_N^{\alpha}(t)\right\}. \tag{5}$$

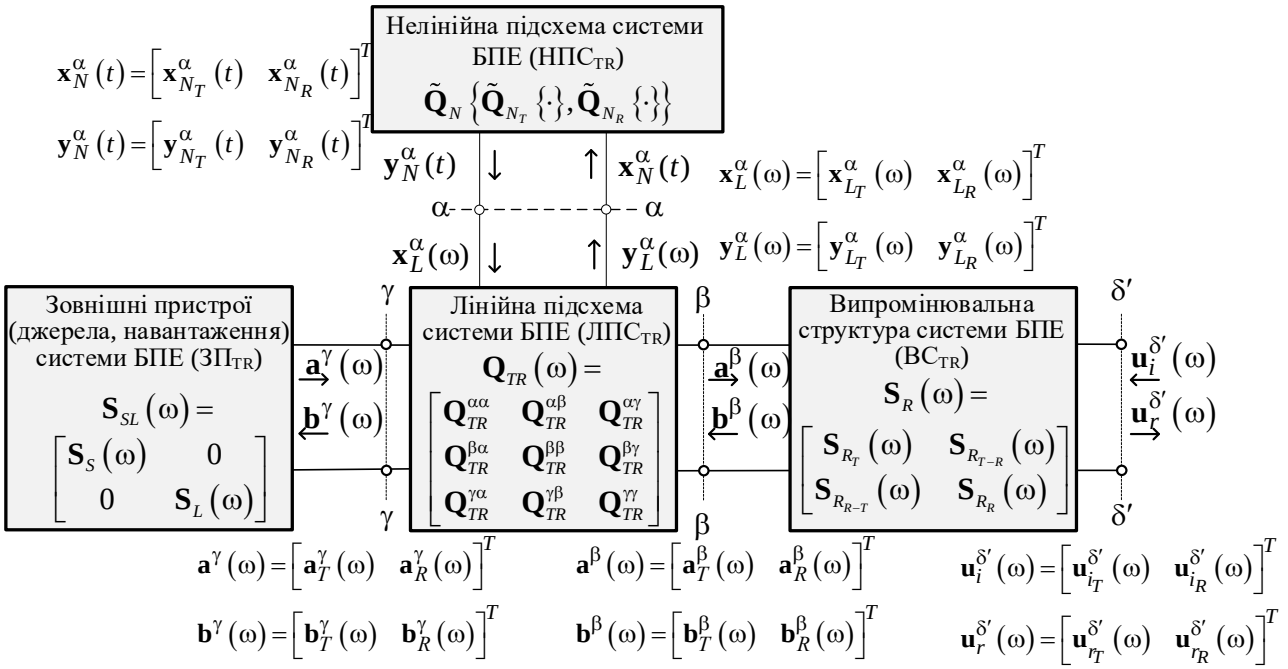


Рис. 13. Подання системи БПЕ у вигляді багатовходової АНЕ

Багатополюсник зовнішніх пристроїв системи БПЕ $ЗП_{TR}$ характеризується блоковою матрицею розсіяння $\mathbf{S}_{SL}(\omega)$ (рис. 13). Багатополюсник $ЛПС_{TR}$ зручно описувати блоковою змішаною матрицею $\mathbf{Q}_{TR}(\omega)$, тому що режими входів багатополюсника $ЛПС_{TR}$, який з'єднується, відповідно, із багатополюсниками $ЗП_{TR}$, $ВС_{TR}$ і нелінійним багатополюсником $НПС_{TR}$ (рис. 13), доцільно характеризувати векторами, аналогічними тим, які характеризують режими відповідних входів приєднуваних до нього багатополюсників. Багатополюсник $ВС_{TR}$ описується блоковою матрицею розсіяння $\mathbf{S}_R(\omega)$ (рис. 13), яка пов'язує падаючі та відбиті хвилі в перерізі $\beta-\beta$ (внутрішньосистемна взаємодія) й амплітуди збіжних та розбіжних хвиль у каналах вільного простору в перерізі $\delta'-\delta'$. Вектори $\mathbf{u}_i^{\delta'}(\omega)$ та $\mathbf{u}_r^{\delta'}(\omega)^T$ є комплексними амплітудами сферичних взаємоортогональних типів хвиль, за допомогою яких досліджувана система БПЕ взаємодіє з іншими системами (міжсистемна взаємодія).

У [100] розглянуто виведення рівнянь стану та вихідних рівнянь системи БПЕ. Зазначається, що найдоцільніше вектором змінних стану вибрати один з векторів, які описують режими входів $ЛПС_{TR}$, з'єднуваних з $НПС_{TR}$, тобто або вектор $\mathbf{x}_L^{\alpha}(\omega)$, або вектор $\mathbf{y}_L^{\alpha}(\omega)$. З точки зору аналізу системи БПЕ вибір $\mathbf{x}_L^{\alpha}(\omega)$ або $\mathbf{y}_L^{\alpha}(\omega)$, як вектора змінних стану, є цілком рівноправним, тому що і той, і другий вектори, по-перше, однозначно визначають режим входів у перерізі $\alpha-\alpha$, і, по-друге, знаючи один з цих векторів, можна визначити режим усієї системи БПЕ, тобто режими у перерізах $\beta-\beta$, $\gamma-\gamma$, $\delta-\delta$ (рис. 13). При такому виборі змінних стану на етапі складання рівнянь стану всю лінійну підсхему системи БПЕ досить описати лише відносно перерізів $\alpha-\alpha$.

Далі з врахуванням (5) та умов з'єднання у перетині $\alpha-\alpha$ отримано рівняння стану системи БПЕ в часовій області:

$$\mathbf{y}_L^{\alpha}(t)=\tilde{\tilde{\mathbf{Q}}}_N\left\{\mathbf{x}_L^{\alpha}(t)\right\} \text{ або } \mathbf{x}_L^{\alpha}(t)=\tilde{\tilde{\mathbf{Q}}}_N^{-1}\left\{\mathbf{y}_L^{\alpha}(t)\right\}, \tag{6}$$

де

$$\tilde{\tilde{\mathbf{Q}}}_N\{\cdot\}=diag\left\{\tilde{\tilde{\mathbf{Q}}}_{N_T}\{\cdot\},\tilde{\tilde{\mathbf{Q}}}_{N_R}\{\cdot\}\right\},\ \tilde{\tilde{\mathbf{Q}}}_{N_T}\{\cdot\}=diag\left\{-\tilde{\mathbf{G}}_{N_T}\{\cdot\},\tilde{\mathbf{R}}_{N_T}\{-\mathbf{E}\cdot\},\tilde{\mathbf{S}}_{N_T}\{\cdot\}\right\},$$

$$\tilde{\tilde{\mathbf{Q}}}_{N_R}\{\cdot\}=diag\left\{-\tilde{\mathbf{G}}_{N_R}\{\cdot\},\tilde{\mathbf{R}}_{N_R}\{-\mathbf{E}\cdot\},\tilde{\mathbf{S}}_{N_R}\{\cdot\}\right\}.$$

З розв'язку рівняння стану визначається вектор $\mathbf{x}_L^{\alpha}(t)$. Надалі розглядається тільки періодичний або квазіперіодичний сталий режим роботи систем БПЕ. Відповідно встановлюється зв'язок між векторами $\mathbf{x}_L^{\alpha}(t)$ та $\mathbf{x}_L^{\alpha}(\omega)$, $\mathbf{y}_L^{\alpha}(t)$ та $\mathbf{y}_L^{\alpha}(\omega)$.

В ході одержання рівнянь стану (6) жодних обмежень (окрім режиму збудження) не вводилося. Тому отримані рівняння стану дають можливість описати притаманні системам БПЕ нелінійні ефекти, зв'язані з утворенням у їх відгуку нових спектральних складових із частотами $\nu_n \neq \omega_k \ \ (k=\overline{0,q})$, відмінними від частот вхідних впливів ω_k, а також нелінійну залежність вектора змінних стану $\mathbf{x}(\omega)$ від рівня вхідних впливів.

Система вихідних рівнянь системи БПЕ [100]:

$$\begin{bmatrix}\mathbf{b}^{\gamma}(\nu_n)\\ \mathbf{u}_r^{\delta'}(\nu_n)\end{bmatrix}=\begin{bmatrix}\tilde{\mathbf{Q}}_{TR}^{\gamma\alpha}(\nu_n) & \tilde{\mathbf{Q}}_{TR}^{\gamma\gamma}(\nu_n) & \tilde{\mathbf{Q}}_{TR}^{\gamma\delta'}(\nu_n)\\ \tilde{\mathbf{Q}}_{TR}^{\delta'\alpha}(\nu_n) & \tilde{\mathbf{Q}}_{TR}^{\delta'\gamma}(\nu_n) & \tilde{\mathbf{Q}}_{TR}^{\delta'\delta'}(\nu_n)\end{bmatrix}\begin{bmatrix}\mathbf{x}_L^{\alpha}(\nu_n)\\ \mathbf{a}^{\gamma}(\omega_k)\\ \mathbf{u}_i^{\delta'}(\omega_k)\end{bmatrix}. \tag{7}$$

Знаючи вектор вхідних параметрів системи БПЕ, можна визначити всі її зовнішні параметри, що описують міжсистемну взаємодію та внутрішньосистемні процеси [103, 104]. Фізичний зміст блоків матриці розсіяння випромінювальної структури системи БПЕ та співвідношення, що дозволяють розрахувати всі блоки матриці розсіяння, які необхідні при складанні і розв'язанні рівнянь стану системи БПЕ, а також визначені її зовнішніх характеристик, наведено у [101, 102].

5. Експериментальна перевірка адекватності запропонованої моделі

Для перевірки адекватності запропонованої моделі у [100 – 104] та методики аналізу систем БПЕ було проведено розрахунок за розробленою моделлю і порівняння з експериментальними результатами, які отримані різними авторами для простих систем БПЕ. Розглядається тільки передача енергії в ближній зоні, так як даний випадок дозволяє, з одного боку, підтвердити достовірність ММ [100 – 104] і, з іншого – показати її універсальність.

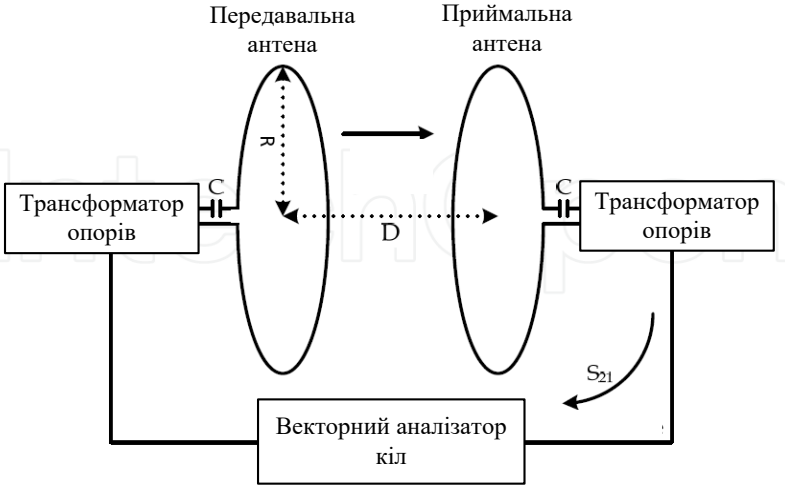


Рис. 14. Схема експерименту з вимірювання коефіцієнта передачі системи БПЕ [107]

Як перший приклад розглянуто передачу енергії між двома малими кільцевими антенами та проведено порівняння з експериментальними результатами [107]. Схема експерименту показана на рис. 14. Автори цієї роботи досліджували коефіцієнт передачі за потужністю

$k_P = \frac{P_L}{P_{in}} = |S_{21}|^2$ між передавальною та приймальною антенами, розташованими в ближній зоні. Передавальна та приймальна антени налаштовувалися в резонанс на одній і тій же частоті. Особливістю даної моделі системи БПЕ є те, що автори використовували два випромінювачі з узгоджувальними пристроями замість узгодження за допомогою додаткових розсіювачів, що входять до складу системи. Перевагою використовуваної системи є можливість виключення індуктивностей з магнітним зв'язком, що призводить до спрощення системи. Кожна з цих антен має один виток радіусом 150 мм і підключені послідовно ємності для настроювання на необхідну резонансну частоту. Експеримент було проведено для відстаней між двома антенами відповідно 49, 80, 170 й 357 мм. Передбачалося, що середня робоча частота системи $f_0 \approx 13.5$ МГц ($\lambda_0 \approx 22.2$ м). Таким чином, антени мали радіус $0.0067\lambda_0$, а відстані між ними становили – $0.002\lambda_0$, $0.0036\lambda_0$, $0.0077\lambda_0$ й $0.016\lambda_0$. На рис. 14 наведено отримані за запропонованою моделлю частотні залежності коефіцієнта передачі за потужністю даної системи БПЕ при різних відстанях між приймальною і передавальною антенами та результати експериментальних вимірювань з [107].

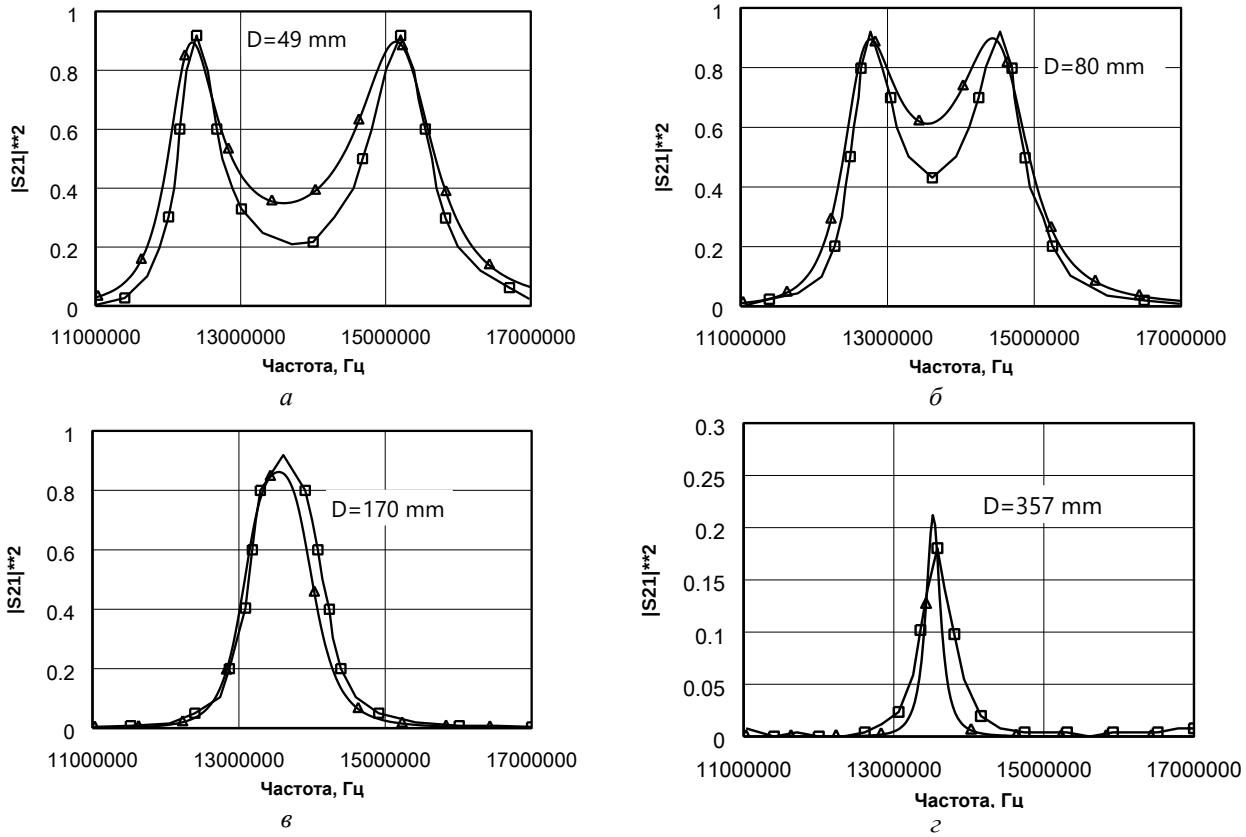


Рис. 15. Частотні залежності коефіцієнта передачі за потужністю системи БПЕ при різних відстанях між приймальною і передавальною антенами (-Δ- –результати моделювання, -□- – експериментальні значення [107])

Як видно з рис. 15, при відстані між антенами 49 мм система має найбільший коефіцієнт передачі за потужністю на двох частотах, приблизно на 12.4 і 15.2 МГц. При зменшенні відстані а, отже, зв'язку між антенами, різниця резонансних частот зменшується і вони приблизно рівні 12.7 та 14.6 МГц (рис. 15, *б*), відповідно. При відстанях більше 170 мм поділ резонансних частот зникає (рис. 15, *в, г*), і система має одну резонансну частоту приблизно 13.6 МГц, а ефективність системи значно знижується зі збільшенням відстані. Порівняння графіків на рис. 15, *а – г* показує досить гарний збіг результатів моделювання та експерименту.

Як другий приклад було використано результати експериментальних вимірювань параметрів розсіювання системи БПЕ, показаної на рис. 16 [108]. При виготовленні експериментального макета використано мідні провідники і керамічні конденсатори ємністю 100 пФ для настроювання випромінювачів в резонанс. Вимірювання S-параметрів проводилося за схемою, показаної на рис. 14, однак без використання трансформаторів опорів. Експериментальні результати з [108] показано на рис. 17. На цих рисунках показані також S-параметри, отримані в результаті моделювання. За даними авторів провідність матеріалу випромінювачів становила $6.8 \cdot 10^7$ См/м. Таке значення використано при моделюванні випромінювачів, що мають кінцеву провідність. Як видно з рис. 17,*а* наявність втрат в провідниках випромінювачів істотно знижує коефіцієнт передачі системи, а рис. 17,*б* демонструє досить гарний збіг теоретичних та експериментальних результатів.

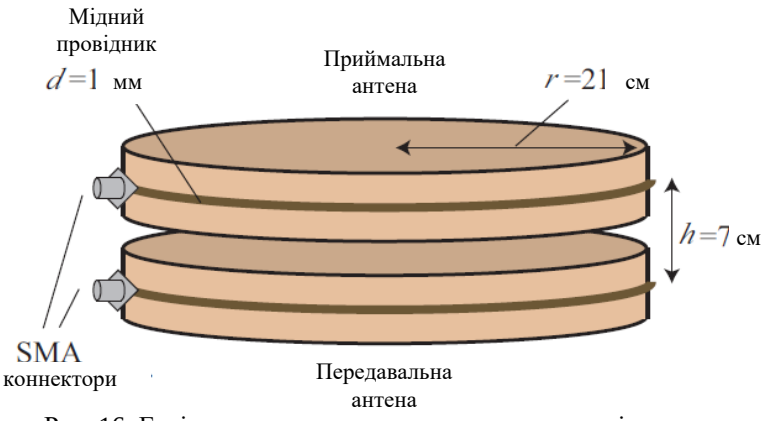
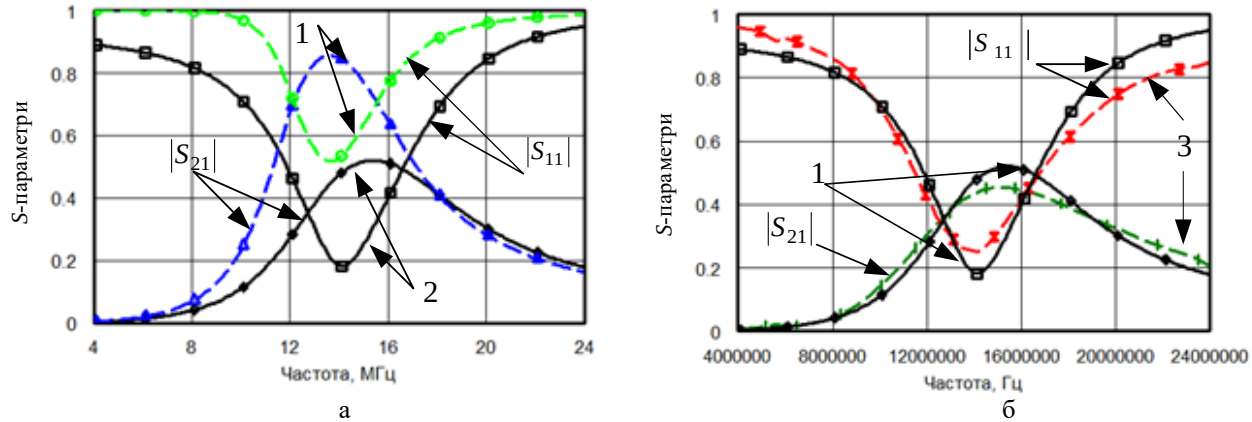


Рис. 16. Ескіз експериментального макета для вимірювання S-параметрів системи БПЕ [108]



а б

Рис. 17. Параметри розсіювання системи БПЕ, показаної на рис. 16: *а* – результати моделювання для випромінювачів, виконаних з ідеальних провідників (криві 1) та матеріалу з кінцевою провідністю (криві 2);
б – порівняння результатів моделювання (криві 1) з результатами експерименту (криві 3) [108]

Ще одним із прикладів для порівняння було обрано систему БПЕ, фотографія якої запозичена з [109] та показана на рис. 18. Ця система БПЕ складається з двох випромінювачів (передавального і приймального) та трьох розсіювачів. Випромінювачі та розсіювачі за допомогою ємностей налаштовувалися в резонанс на частоту близько 70 МГц. Радіус випромінювачів і розсіювачів становив 37 мм, значення резонансних ємностей 33 пФ, відстань між елементами 70 мм, а опори джерела і навантаження 50 Ом. На рис. 19 наведено частотні залежності коефіцієнта передачі системи, що отримані за пропонованою моделлю (крива 1), експериментальні результати з роботи [109] (крива 2) і результати розрахунку по моделі, запропонованої в цій же роботі (крива 3).

Як видно з порівняння, результати розрахунку за запропонованою моделлю набагато краще збігаються з експериментальними результатами, ніж результати моделювання, отримані авторами [109]. Таким чином, наведені в цьому підрозділі результати свідчать про достовірність моделі, запропонованої в даній роботі, адекватність отриманих з її допомогою результатів і її універсальність.

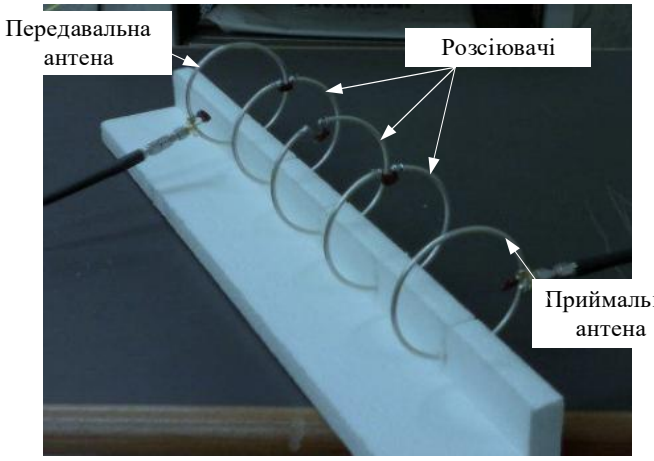


Рис. 18. Фотографія експериментального макета системи БПЕ [109]

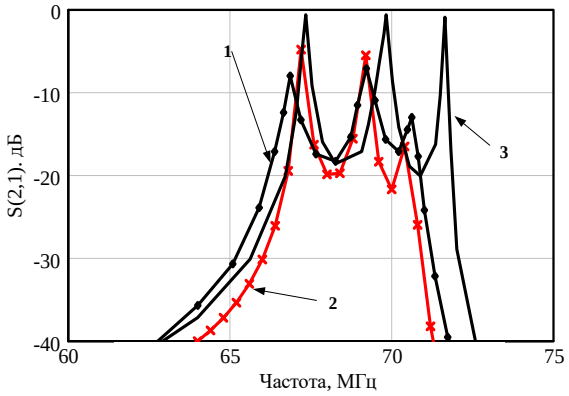


Рис. 19. Частотні залежності коефіцієнта передачі системи БПЕ, що наведена на рис. 18

Висновки

Останнім часом все більша увага приділяється напрямку науки і техніки, пов'язаному з розробкою інноваційних технологій створення нових високоефективних систем БПЕ. Видно, що до теперішнього часу в галузі БПЕ зроблено чимало, підсилюється її вплив на процес розвитку світової енергетики та технічний рівень розробок у цій галузіі досить високий. Проте, незважаючи на значні досягнення в галузі БПЕ, все ще багато питань перебувають у стадії досліджень, причому це стосується не тільки окремих підсистем, але й усієї системи БПЕ в цілому, що функціонує в умовах реальної електромагнітної обстановки. Одні з цих питань були вже вирішені колективом лабораторії антен ХНУРЕ. Було показано, що для аналізу та оптимізації систем БПЕ, в яких використовуються принципово різні технології безпровідної передачі енергії, можна застосувати підхід, в основі якого лежить єдине уявлення, на електродинамічному рівні, про функціонування систем БПЕ широкого класу й призначення. У даній роботі були продовжені дослідження, започатковані колективом лабораторії антен ХНУРЕ щодо створення ММ системи БПЕ. А саме – доведено адекватність розробленої нелінійної ММ електродинамічного рівня системи БПЕ шляхом порівняння результатів розрахунків щодо розробленої моделі з відомими експериментальними даними. Показано, що розроблена колективом ХНУРЕ нелінійна ММ електродинамічного рівня системи БПЕ є адекватна та дозволяє проводити суворий аналіз та оптимізацію як систем БПЕ широкого класу й призначення в цілому, так і окремих їх підсистем, функціональних блоків та вузлів.

Список літератури:

1. Shinohara N. Wireless Power Transfer via Radiowaves. John Wiley & Sons. 2014. 238 p.

2. Nikoletseas S., Yang Y., Georgiadis A. Wireless Power Transfer Algorithms, Technologies and Applications in Ad Hoc Communication Networks. Springer Nature Switzerland AG. 2016. 745 p.

3. Lu X., Wang P., Niyato D., Kim D.I., Han Z. Wireless charging technologies: Fundamentals, standards, and network applications // IEEE Communications Surveys & Tutorials. 2015. vol. 18, no. 2. P. 1413 – 1452.

4. Sun T., Xie X., Wang Z. Wireless power transfer for medical microsystems. Springer. 2013.

5. Сазонов Д.М. Антенны и устройства СВЧ. Москва : Высш. шк., 1988. 432 с.

6. Грецких Д.В., Гомозов А.В., Цикаловский Н.М., Аль-Самарай Ш.Ф.А. Области применения и современные тенденции развития наноректенн // Технология приборостроения. 2012. №2. С. 36 – 42.

7. Дьячков П.Н. Электронные свойства и применения нанотрубок. Москва : БИНОМ. Лаборатория знаний, 2011. 488 с.

8. Слепян Г.Я., Максименко С.А., Кужир П.П. Современные тенденции развития наноэлектромагнетизма: аналитический обзор [Электронний ресурс] // НИУ «Ин-т ядерных проблем» БГУ. 2012. Режим доступу: http://elib.bsu.by/bitstream/123456789 /18999/1/fanem_2012.pdf.

9. Novotny L., N. van Hulst. Antennas for Light // Nat. Photon. 2011. N 5. P. 83 – 90.

10. Zhu Z. Optical rectenna solar cells using graphene geometric diodes / Z. Zhu, S. Grover, K. Krueger, G. Moddel // 37th IEEE Photovoltaic Specialists Conference. 2011. P. 20 – 22.

11. Joshi S. Infrared Optical Response of Geometric Diode Rectenna Solar Cells / S.Joshi, Z. Zhu, S. Grover, G. Moddel // 38th IEEE Photovoltaic Specialists Conference. 2012. P. 2976 – 2978.

12. Kotter D.K., Novack S.D., Slafer W.D., Pinhero P.J. Theory and manufacturing processes of solar nanoantenna electromagnetic collectors // Journal of Solar Energy Engineering-transactions of The Asme. 2010. Vol. 132, N 1. P. 1 – 10.

13. Pan Y., Rosamond M.C., McDonald A. at al. Design and performance of micro-rectenna arrays for thermal energy harvesting // 40th International Conference on Infrared, Millimeter, and Terahertz waves (IRMMW-THz). 2015. P. 1 – 2.

14. Qassim Abdullahi S., Rahil Joshi, Symon K. Podilchak, Sadeque R. Khan, at al. Design of a wireless power transfer system for assisted living applications // Desmulliez and Apostolos Georgiadis Wireless Power Transfer. 2019. Vol. 6, Is. 1. P. 41 – 56.

15. Mickel Budhia, Grant A. Covic, John T. Boys. Design and Optimization of Circular Magnetic Structures for Lumped Inductive Power Transfer Systems // IEEE Transactions on Power Electronics. 2011. Vol. 26, Is. 11. P. 3096 – 3108.

16. Taylor M. Fisher, Kathleen Blair Farley, Yabiao Gao, Hua Bai and Zion Tsz Ho Tse. Electric vehicle wireless charging technology: a state-of-the-art review of magnetic coupling systems // Wireless Power Transfer. 2014. Vol. 1, Is. 02. P 87 – 96.

17. Hassler M., Atasoy O., Twelker K., Kesler M., Birkendahl J. and Krammer J. A comparison on simulated, analytic, and measured impedance values for an inductive power transfer system // Wireless Power Transfer. 2020. Vol. 7, Is. 1. P. 51 – 59.

18. Poguntke T., Schumann P., Ochs K. Radar-based living object protection for inductive charging of electric vehicles using two-dimensional signal processing // Wireless Power Transfer. 2017. Vol. 4, Spec. Is. 2: Contactless Charging for Electric Vehicles. P. 88–97.

19. Zhang Z., Zhang B., Deng B., Wei X. and Wang J. Opportunities and challenges of metamaterial-based wireless power transfer for electric vehicles // Wireless Power Transfer. 2018. Vol. 5(1). P. 9 – 19.

20. Jing Zhou, Kan Guo, Zhonghua Chen, Hui Sun and Sideng Hu. Design considerations for contact-less underwater power delivery: a systematic review and critical analysis // Wireless Power Transfer. 2020. Vol. 7, Is. 1. P. 76 – 85.

21. Wang X., Liu Z., Zhang T. Flexible sensing electronics for wearable/attachable health monitoring // Small. 2017. Vol. 13(25). P. 1 – 19.

22. Curry E.J., Ke K., Chorsi M.T., Wrobel K.S., at al. Biodegradable piezoelectric force sensor // Proc. Natl. Acad. Sci. USA. 2018, Vol. 115. P. 909 – 914.

23. Devansh R. Agrawal, Yuji Tanabe, Desen Weng, Andrew Ma, at al. Conformal phased surfaces for wireless powering of bioelectronic microdevices // Nature Biomedical Engineering. 2017. Vol. 1, Article number: 0043. P. 1 – 9.

24. Mustafa F. Mahmood, Saleem Lateef Mohammed, Sadik Kamel Gharghan, Ali Al-Naji, and Javaan Chahl. Hybrid Coils-Based Wireless Power Transfer for Intelligent Sensors // Sensors. 2020. Vol. 20(9). P. 1 – 24.

25. Monti G., Arcuti P., Tarricone L. Resonant Inductive Link for Remote Powering of Pacemakers // IEEE Transactions On Microwave Theory And Techniques. 2015. Vol. 63, No. 11. P. 3814 – 3822.

26. Moore J., Castellanos S., Xu S., Wood B., at al. Applications of Wireless Power Transfer in Medicine: State-of-the-Art Reviews // Annals of Biomedical Engineering. 2019. Vol. 47, No. 1. P. 22 – 38.

27. Xiong Q. Wireless Charging Device for Artificial Cardiac Pacemaker / Quan Xiong // International Conference on Information Technology and Management Innovation (ICITMI). 2015. P. 765–768.

28. Brown W.C. Experimental involving a microwave beam to power and position a helicopter // IEEE Transactions on Aerospace and Electronic Systems. 1969. Vol. AES-5, Is. 5. P. 692 – 702.

29. Brown W.C. Adapting Microwave Techniques to Help Solve Future Energy Problems //G MTT International Microwave Symposium Digest of Technical Papers. 1973. Vol. 73.1. P. 189 – 191.

30. Brown W.C., Eves E.E. Microwave power transmission and its application to space // IEEE Transactions on Microwave Theory and Techniques. 1992. Vol. 40, No 8. P. 1239 – 1250.

31. Yang Y., Zhang Y., Duan B., at al. A novel design project for space solar power station (SSPS-OMEGA) // Acta Astronautica. 2016. Vol. 121. P. 51 – 58.

32. Glaser P.E. An overview of the solar power satellite option // IEEE Transactions on Microwave Theory and Techniques. 1992. Vol. 40, Is. 6. P. 1230 – 1238.

33. Shinohara N. Power without wires // IEEE Microwave Magazine. 2011. V.12, No 7. P. 64 – 73.

34. Shoki H. Issues and Initiatives for Practical Use of Wireless Power Transmission Technologies in Japan // Microwave Workshop Series on Innovative Wireless Power Transmission: Technologies, Systems, and Applications (IMWS), IEEE MTT-S International. 2011. P. 87 – 90.

35. Возобновляемая энергетика. Пути повышения энергетической и экономической эффективности // Труды Междунар. форума «Возобновляемая энергетика. Пути повышения энергетической и экономической эффективности REENFOR – 2014» ; под. ред. О.С. Попеля, Д.О. Дуникова. Москва : ОИВТ РАН, 2014. 478 с.

36. Гомозов А.В., Гомозов В.И., Шокало В.М., Грецких Д.В., Аль-Самарай Ш.Ф.А. Передающая подсистема беспроводной передачи энергии к труднодоступным объектам на основе многопозиционной системы излучателей с фокусировкой излучения. Ч. 1 // Радиотехника. 2011. №165. С. 112 – 118.

37. Гомозов А.В., Гомозов В.И., Шокало В.М., Грецких Д.В., Аль-Самарай Ш.Ф.А. Передающая подсистема беспроводной передачи энергии к труднодоступным объектам на основе многопозиционной системы излучателей с фокусировкой излучения. Ч. 2 // Радиотехника. 2011. №167. С. 18 – 24.

38. Gomozov A.V., Shokalo V.M., Gretskih D.V., Al-Sammarrai Sh.F.A. Principles of construction and application of microwave systems for wireless energy transmission of ground and space basing // Computational problems of electrical engineering. 2012. Vol. 2, № 1. P. 15 – 23.

39. Shimokura N., Kaya N., Shinohara M., Matsumo H. Point-to-point microwave power transmission experiment // Scripta Technica, Inc. Electr Eng Jpn. 1997. No 120(1). P. 33 – 39.

40. Applications of wireless power transmission via radio frequency beam / Report ITU-R SM.2392-0. 2016. 33 p.

41. MHI Successfully Completes Ground Demonstration Testing of Wireless Power Transmission Technology for SSPS – Expanding the Potential for New Industrial Applications [Электронний ресурс] // Press information. 2015. Режим доступу до ресурсу: http://www.mhi-global.com/news/story/1503121879.html.

42. East T. A self-steering array for the SHARP microwave-powered aircraft // IEEE Transactions on Antennas and Propagation. 1992. Vol. 40, No 12. P. 1565 – 1567.

43. Sohlesak J.J., Alden A., Ohno T. SHARP (Stationary high altitude platform): rectenna and low altitude tests // Globecom 85: IEEE Glob. Telecommun. conf. New Orleans. 1985. Vol. 2. P. 960 – 964.

44. Fujino Y., Fujita M., Kaya N., et al. A Dual Polarization Microwave Power Transmission System for Microwave Propelled Airship Experiment // ISAP'96 Proceedings, Chiba, Japan. 1996. P. 393 – 396.

45. Fujino Y., Fujita M. Development of a High-Efficiency Rectenna for Wireless Power Transmission – Application to Microwave-Powered Airship Experiment // J. Commun. Res. Lab. 1999. Vol. 43, No 3. P. 367 – 37.

46. Gavan J., Tapuchi S. Microwave wireless-power transmission to high-altitude-platform systems // URSI Radio Science Bulletin. 2010. Vol. 2010, No 334. P. 25 – 42.

47. Dickinson R.M. Power in the sky: Requirements for microwave wireless power beamers for powering high-altitude platforms // IEEE Microwave Magazine. 2013. Vol. 14, Is. 2. P. 36 – 47.

48. Yuichiro O., Naohiro T. Study of Electric Aircraft Charged by Beamed Microwave Power // IHI Engineering Review. 2015. Vol. 48, No 2. P. 29 – 32.

49. Shimamura K., Sawahara H., Oda A., et al. Feasibility study of microwave wireless powered flight for micro air vehicles // Wireless Power Transfer. 2017. Vol. 4, No 2. P. 146 – 159.

50. Takabayashi N., Shinohara N., Mitani T., Furukawa M., Fujiwara T. Rectification Improvement With Flat-Topped Beams on 2.45-GHz Rectenna Arrays // IEEE Transactions on Microwave Theory and Techniques. 2020. Vol. 68, Is. 3. P. 1151 – 1163.

51. Гомозов А.В., Грецких Д.В., Цикаловский Н.М., Шарапова Е.В. Радиотехническая система беспроводного энергоснабжения беспилотных летательных аппаратов // Космическая техника. Ракетное вооружение. Сб. науч.-техн. ст. ГП" КБ Южное". 2015. №1 (108). С. 36 – 41.

52. Sherazi H.H.R., Zorbas D., O'Flynn B. A Comprehensive Survey on RF Energy Harvesting: Applications and Performance Determinants // Sensors. 2022, Vol. 22, No. 2990. P. 1 – 36.

53. Грецких Д.В., Лихограй В.Г., Щербина А.А., Сакало С.Н., Ткачева Т.С. Система контроля подвески автомобиля на основе технологий беспроводной передачи энергии // Радиотехника. 2020. Вып. 201. С. 52 – 63.

54. Gretskih D., Luchaninov A., Lykhograi V., Shcherbina A., Sakalo S. Researching the possibility of wireless energy transmission for the power supply condition monitoring system of a car's suspension // IEEE Ukrainian Microwave Week. 2020. P. 105 – 109.

55. Powercast Corporation. TX91501b Powercaster®Transmitter. [Электронний ресурс] // Режим доступу: https://www.powercastco.com/prod ucts/powercaster-transmitter (4 січня 2021).

56. Powercast Corporation. Powercaster® Powerspot. [Электронний ресурс] // Режим доступу: https://www.powercastco.com/products/powersp (4 січня 2021).

57. Ossia. Ossia's Cota: Real Wireless Power. [Электронний ресурс] // Режим доступу: https://www.ossia.com/cota (4 січня 2021).

58. Ren J., Hu J., Zhang D., Guo H., Zhang Y., Shen X. RF energy harvesting and transfer in cognitive radio sensor networks: Opportunities and challenges // IEEE Commun. Mag. 2018. Vol. 56. P. 104 – 110.

59. Stoopman M., Keyrouz S., Visser H., Philips K., Serdijn W. A self-calibrating RF energy harvester generating 1V at 26.3 dBm // Proceedings of the 2013 Symposium on VLSI Circuits, Kyoto, Japan. 2013. P. 226 – 227.

60. Stoopman M., Keyrouz S., Visser H.J., Philips K., Serdijn W.A. Co-design of a CMOS rectifier and small loop antenna for highly sensitive RF energy harvesters // IEEE J. Solid-State Circuits. 2014. Vol. 49. P. 622 – 634.

61. Sample A.P., Parks A.N., Southwood S., Smith J.R. Wireless ambient radio power. In Wirelessly Powered Sensor Networks and Computational RFID // Springer: Berlin/Heidelberg, Germany, 2013. P. 223 – 234.

62. Papotto G., Carrara F., Finocchiaro A., Palmisano G. A 90-nm CMOS 5-Mbps crystal-less RF-powered transceiver for wireless sensor network nodes // IEEE J. Solid-State Circuits. 2013. Vol. 49. P. 335 – 346.

63. Gubbi J., Buyya R., Marusic S., Palaniswami M. Internet of Things (IoT): A vision, architectural elements, and future directions // Future Gener. Comput. Syst. 2013. Vol. 29. P. 1645 – 1660.

64. Asghari P., Rahmani A.M., Javadi H.H.S. Internet of Things applications: A systematic review // Comput. Netw. 2019. Vol. 148. P. 241 – 261.

65. Choudhary P., Bhargava L., Sing, V., Choudhary M., Kumar Suhag A. A survey–Energy harvesting sources and techniques for internet of things devices // Mater. Today Proc. 2020. Vol. 30. P. 52 – 56.

66. Alsharif M.H., Kim S., Kuruo˘ glu, N. Energy Harvesting Techniques for Wireless Sensor Networks/Radio-Frequency Identification: A Review // Symmetry. 2019. Vol. 11. P. 865.

67. Adila A.S., Husam A., Husi G. Towards the self-powered Internet of Things (IoT) by energy harvesting: Trends and technologies for green IoT // Proceedings of the 2018 2nd International Symposium on Small-Scale Intelligent Manufacturing Systems (SIMS), Cavan, Ireland, 2018. P. 1 – 5.

68. Krupitzer C., Müller S., Lesch V., Züfle M., at al. A Survey on Human Machine Interaction in Industry 4.0. arXiv 2020, arXiv:2002.01025.

69. Sherazi H.H.R., Grieco L.A., Imran M.A., Boggia G. Energy-efficient LoRaWAN for Industry 4.0 Applications // IEEE Trans. Ind. Inform. 2020. Vol. 17. P. 891 – 902.

70. Tahir M.A., Ferrer B.R., Luis J., Lastra M. An Approach for Managing Manufacturing Assets through Radio Frequency Energy Harvesting // Sensors. 2019. Vol. 19. p. 1 – 21.

71. Tang X., Wang X., Cattley R., Gu F., Ball A.D. Energy harvesting technologies for achieving self-powered wireless sensor networks in machine condition monitoring: A review // Sensors. 2018. Vol. 18. P. 1 – 39.

72. Zungeru A.M., Ang L.M., Prabaharan S., Seng K.P. Radio frequency energy harvesting and management for wireless sensor networks. In Green Mobile Devices and Networks: Energy Optimization and Scavenging Techniques; Number 13 in 0 // CRC Press: New York, NY, USA. 2012. P. 341 – 368.

73. Visser H.J., Vullers R.J.M. RF Energy Harvesting and Transport for Wireless Sensor Network Applications: Principles and Requirements // Proc. IEEE. 2013. Vol. 101. P. 1410 – 1423.

74. Boisseau S., Despesse G. Energy harvesting, wireless sensor networks & opportunities for industrial applications // EE Times. 2012.

75. Iyengar A., Kundu A., Pallis G. Healthcare Informatics and Privacy // IEEE Internet Comput. 2018. Vol. 22. P. 29 – 31.

76. Yang L., Zhou Y.J., Zhang C., Yang X.M., Yang X.X., Tan C. Compact multiband wireless energy harvesting based battery-free body area networks sensor for mobile healthcare // IEEE J. Electromagn. Microwaves Med. Biol. 2018. Vol. 2. P. 109 – 115.

77. Anwar M., Abdullah A.H., Qureshi K.N., Majid A.H. Wireless body area networks for healthcare applications: An overview // Telkomnika. 2017. Vol. 15. P. 1088 – 1095.

78. Luo Y., Pu L., Zhao Y. RF Energy Harvesting Sensor Networks for Healthcare of Animals: Opportunities and Challenges // arXiv 2018, arXiv:1803.00106.

79. Saraereh O.A., Alsaraira A., Khan I., Choi B.J. A hybrid energy harvesting design for on-body Internet-of-Things (IoT) networks // Sensors. 2020. Vol. 20. p. 1 – 17.

80. Haghi M., Thurow K., Stoll R. Wearable devices in medical internet of things: Scientific research and commercially available devices // Healthc. Inform. Res. 2017. Vol. 23. P. 4 – 15.

81. Borges L.M., Chávez-Santiago R., Barroca N., Velez F.J., Balasingham I. Radio-frequency energy harvesting for wearable sensors // Healthc. Technol. Lett. 2015. Vol. 2. P. 22 – 27.

82. Lin C., Chiu C., Gong J. A Wearable Rectenna to Harvest Low-Power RF Energy for Wireless Healthcare Applications // In Proceedings of the 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics(CISP-BMEI), Beijing, China. 2018. P. 1 – 5.

83. Hande A., Bridgelall R., Bhatia D. Energy harvesting for active RF sensors and ID tags. In Energy Harvesting Technologies // Springer: Berlin/Heidelberg, Germany. 2009. P. 459 – 492.

84. Aparicio M.P., Bakkali A., Pelegri-Sebastia J., Sogorb, T., Bou V. Radio frequency energy harvesting-sources and techniques. In Renewable Energy: Utilisation and System Integration // Intechopen: London, UK. 2016.

85. Cui L., Zhang Z., Gao N., Meng Z., Li Z. Radio frequency identification and sensing techniques and their applications: A review of the state-of-the-art // Sensors. 2019. Vol. 19. p. 1 – 23.

86. Mhatre P., Duche R., Nawale S., Patil P. RF power harvesting system for RFID applications in multiband systems // Proceedings of the 2015 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Denton, TX, USA. 2015. P. 1 – 5.

87. Olgun U., Chen C., Volakis J.L. Wireless power harvesting with planar rectennas for 2.45 GHz RFIDs // Proceedings of the 2010 URSI International Symposium on Electromagnetic Theory, Berlin, Germany. 2010. P. 329–331.

88. Pellerano S., Alvarado J., Palaskas Y. A mm-Wave Power-Harvesting RFID Tag in 90 nm CMOS // IEEE J. Solid-State Circuits. 2010. Vol. 45. P. 1627 – 1637.

89. Bakhtiar A.S., Jalali M.S., Mirabbasi S. An RF power harvesting system with input-tuning for long-range RFID tags // Proceedings of 2010 IEEE International Symposium on Circuits and Systems, Paris, France. 2010. P. 4085 – 4088.

90. Slesinski R.J. Power Harvesting for Actively Powered RFID Tags and Other Electronic Sensors // U.S. Patent App. 12/039,691, 3 September 2009.

91. Sony S., Laventure S., Sadhu A. A literature review of next-generation smart sensing technology in structural health monitoring // Struct. Control. Health Monit. 2019. Vol. 26. P. 1 – 22.

92. Srinivasan R., Ali U.H.H. Energy harvesting wireless sensor for achieving self-powered structural health monitoring system // Circuit World. 2020. Vol. 46. P. 307 – 315.

93. Loubet G., Takacs A., Gardner E., De Luca A., Udrea F., Dragomirescu D. LoRaWAN Battery-Free Wireless Sensors Network Designed for Structural Health Monitoring in the Construction Domain // Sensors. 2019. Vol. 19. P. 1 – 26.

94. Loubet G., Takacs A., Dragomirescu D. Implementation of a battery-free wireless sensor for cyber-physical systems dedicated to structural health monitoring applications // IEEE Access. 2019. Vol. 7. P. 24679 – 24690.

95. Sidibe A., Takacs A., Okba A., Aubert H. Design and Characterization of a Compact Rectenna for Structural Health Monitoring Applications // Proceedings of the 2019 IEEE International Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting, Atlanta, GA, USA. 2019. P. 1803 – 1804.

96. Cao S., Li J. A survey on ambient energy sources and harvesting methods for structural health monitoring applications // Adv. Mech. Eng. 2017. Vol. 9. P. 1 – 14.

97. Шифрин Я.С., Лучанинов А.И., Шокало В.М. Приемно-выпрямительные элементы ректенных систем. Харьков : Харьк. ин-т радиоэлектроники: Деп. в УкрНИИНТИ. 31.03.89. № 941–Ук89, 1988. 182 с.

98. Шокало В.М., Лучанинов А.И., Рыбалко А.М., Грецких Д.В. Крупноапертурные антенны-выпрямители систем беспроводной передачи энергии микроволновым лучом. Харьков : Коллегиум, 2006. 308 с.

99. Шокало В.М., Правда В.І., Усін В.А., Вунтесмері В.С., Грецьких Д.В. Електродинаміка та поширення радіохвиль. Ч.2. Випромінювання та поширення електромагнітних хвиль. Харків : Колегіум, 2010. 435 с.

100. Gretskih D.V. Electrodynamic Model of a Wireless Power Transmission System / D.V. Gretskih, A.I. Luchaninov, J.V. Vishniakova, V.A. Katrich, M.V. Nesterenko // XXIII International Seminar. Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory. 2018. P. 80 – 85.

101. Luchaninov A.I. Electrodynamic Approach to Designing WPT Systems with Accounting for Non-system Interactions / A.I. Luchaninov, D.V. Gretskih, A.V. Gomozov, V.A. Katrich, M.V. Nesterenko // IEEE 2nd Ukraine Conference on Electrical and Computer Engineering. 2019. P. 80 – 85.

102. Gretskih D. Electrodynamic Approach to Designing Wireless Power Transfer Systems (Internal System Processes) / D. Gretskih, A. Luchaninov, V. Katrich, M. Nesterenko // IV International Conference on Information and Telecommunication Technologies and Radio Electronics. 2019. P. 1 – 6.

103. Gretskih D. External Parameters of Wireless Power Transmission Systems / D. Gretskih, A. Luchaninov, A. Gomozov, V. Katrich, M. Nesterenko // XXIV International Seminar. Workshop on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory. 2019. P. 117 – 121.

104. Грецких Д.В., Лихограй В.Г., Щербина А.А., Гомозов А.В. Внешние параметры систем беспроводной передачи энергии // Радиотехника. 2019. №199. С. 59 – 66.

105. Gretskih D. Nonlinear integral equations for multi-input radiating structures / D. Gretskih, A. Luchaninov, V. Alieksieiev, V. Katrich, M. Nesterenko // Proceedings of the XXV International Seminar on Direct and Inverse Problems of Electromagnetic and Acoustic Wave Theory. 2020. P. 97 – 102.

106. Luchaninov A. Two-level Iterative Algorithm for Solving State Equations of the WPT System / A. Luchaninov, D. Gretskih, V. Alieksieiev at all // 16th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering. 2022. P. 352 – 357.

107. Hoang H., Bien F. Maximizing Efficiency of Electromagnetic Resonance Wireless Power Transmission Systems with Adaptive Circuits // Chapters 11 in: Wireless Power Transfer – Principles and Engineering Explorations Ed. by K.Y. Kim, InTech. 2012. P. 207 – 226.

108. Hirayama H. Equivalent Circuit and Calculation of Its Parameters of Magnetic-Coupled-Resonant Wireless Power Transfer // Chapters 6 in: Wireless Power Transfer – Principles and Engineering Explorations Ed. by K.Y. Kim, InTech. 2012. P. 117 – 132.

109. Dionigi M., Mongiardo M. Magnetically Coupled Resonant Wireless Power Transmission Systems with Relay Elements // IEEE MTT-S International Microwave Workshop Series on Innovative Wireless Power Transmission: Technologies, Systems, and Applications. 2012. P. 223 – 226.

Надійшла до редколегії 04.10.2022

Відомості про авторів:

Алєксєєв Василь Олександрович – Харківський національний університет радіоелектроніки, аспірант кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; email: vasyl.alieksieiev@nure.ua; ORCID: https://orcid.org/0000-0002-3282-5985

Грецьких Дмитро Вячеславович – д-р техн. наук, доцент, Харківський національний університет радіоелектроніки, професор кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; email: dmytro.gretskih@nure.ua; ORCID: https://orcid.org/0000-0002-2645-7872

Гавва Дмитро Сергійович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; email: dmytro.gavva@nure.ua; ORCID: https://orcid.org/0000-0002-4033-7746

Лихограй Василь Григорович – канд. техн. наук, доцент, Харківський національний університет радіоелектроніки, доцент кафедри комп'ютерної радіоінженерії та систем технічного захисту інформації, Україна; email: vasyl.lykhograi@nure.ua; ORCID: https://orcid.org/0000-0002-9226-1309

УДК 531.21 **DOI:10.30837/rt.2022.4.211.10**

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

СТРУКТУРНЕ МОДЕЛЮВАННЯ І РОЗРАХУНОК ТЕПЛОПРОВІДНОСТІ ПОЛІІМІДНИХ КОМПОЗИТНИХ МАТЕРІАЛІВ

Вступ

Прогнозування ефективної теплопровідності наповнених полімерних систем вимагає знання не тільки теплових властивостей складових компонентів, але й ряд інших факторів. Першочергове значення має при цьому вивчення структурних змін, які відбуваються у композиті. Дуже важливим з цієї точки зору є розробка моделі наповненого полімеру, а також моделювання граничних шарів, які дуже впливають на властивості наповненої полімерної композитної системи. Необхідно враховувати, що механізми впливу матриці та наповнювача на теплопровідність різні за різних концентрацій. Управління властивостями полімерних композитних матеріалів має забезпечуватися не тільки збільшенням концентрації, а також збільшенням модифікуючої дії наповнювача (за рахунок дисперсності та інших факторів).

Для розробки нових теплопровідних поліімідних (ПІ) композитних плівок слід вибирати такі сполучні та наповнювачі, які забезпечують оптимальні значення теплофізичних, фізико-механічних та діелектричних характеристик поліімідних матеріалів [1].

При використанні поліімідних композитних матеріалів у системах електричної ізоляції та комутуючих елементів електронних вузлів актуальним є завдання покращення теплопередачі з одночасним збереженням високих діелектричних характеристик теплопровідних поліімідних шарів.

Полімерна матриця (у тому числі поліімідна), як правило, має знижені теплопровідні та високі діелектричні властивості. Щоб покращити теплопередачу, можна вводити в поліімідну матрицю діелектричні наповнювачі мікронних і менших розмірів, у тому числі і нанорозмірів, що мають високі теплофізичні характеристики. При такому підході не можна виключити зниження електроізоляційних властивостей композитних поліімідних плівок, характер зміни яких визначається електрофізичними властивостями частинок та їх концентрацій у композиті. Можливості підвищення ефективності теплопередачі у такому матеріалі будуть визначатися правильно підібраними властивостями, розмірами та концентрацією частинок наповнювача в поліімідній матриці.

Оскільки поліімідний композит є двокомпонентною системою, для розрахунку теплопровідності доцільно використовувати аналітичні моделі для вирішення теплофізичних задач. Це дозволяє для наближених розрахунків ефективної теплопровідності двокомпонентних сумішей для широкого діапазону вмісту порошкових наповнювачів у композитах використовувати прості моделі, наприклад формулу Бургера [2, 3]:

$$\lambda_{еф} = \frac{V_{св}\lambda_{св} + CV_{н}\lambda_{н}}{V_{св} + CV_{н}},$$

де $V_н$ – об'ємна частка наповнювача; $V_{св}$ – об'ємна частка сполучного; $\lambda_н$, $\lambda_{св}$ і $\lambda_{еф}$ – теплопровідності наповнювача, сполучного та ефективна теплопровідність суміші відповідно; C – розрахунковий коефіцієнт. Під визначенням «сполучна» розуміється безперервна фаза, а під «наповнювачем» – уривчаста.

Безрозмірний коефіцієнт C залежить, по-перше, від характеристик розподілу фаз у суміші: $C = 1$ – гомогенна суміш, $C << 1$ – ізольована фаза наповнювача, $C > 1$ – безперервна фаза наповнювача. По-друге, від відношення $\lambda_н / \lambda_{св}$ (чим більше це відношення, тим менший коефіцієнт C).

Наявність коефіцієнта C у формулі Бургера робить її зручною для опису теплопровідності двокомпонентних композитних матеріалів. Величина цього коефіцієнта враховує такі характеристики матеріалу, як безперервність/уривчастість фаз кожного з компонентів, форму і розмір частинок наповнювача і відношення їх коефіцієнтів теплопровідності. Саме в цьому полягає його фізичне значення. Недоліком цієї формули є необхідність підбору значення коефіцієнта C для кожного конкретного типу композитного матеріалу, що знижує її універсальність і впливає на точність розрахунків.

У роботах [2, 3] зроблено висновок про суттєвий вплив твердої фази на теплопровідність композитів та підтверджується факт узгодження експериментальних даних з розрахунками за формулою Бургера при значенні коефіцієнта $C > 1$. Оскільки коефіцієнт C розташований при доданку, який враховує теплопровідність твердих частинок, то великі значення коефіцієнту C відповідають збільшенню вкладу твердої фази в теплопровідність суміші, а менші – його зниженню.

Розрахунки, виконані за формулою Бургера для різних наповнювачів і сполучного, відповідають передбачуваній залежності коефіцієнта C від відношення $\lambda_н / \lambda_{св}$ в широкому діапазоні значень ефективної теплопровідності композитів $\lambda_{еф}$ і об'ємного вмісту наповнювача $V_н$.

Однак на практиці теплопровідність композиту при одній і тій же концентрації наповнювача може змінюватись у великих межах. Це пов'язано з особливостями розподілу наповнювача у матриці. Особливо ця залежність проявляється при великій різниці коефіцієнтів теплопровідності матриці та наповнювача. Тому для підтвердження застосування розроблених структурних моделей та програм розрахунку теплопровідності за формулою Бургера необхідно виготовити та дослідити ефективну теплопровідність невиправдано великої кількості експериментальних зразків різних типів нових високонаповнених поліімідних композитних плівок. Визначити величину відхилень експериментально одержаних значень коефіцієнта теплопровідності композитного матеріалу від теоретично розрахованих за моделлю. Здійснити розрахунок та провести корекцію значень коефіцієнтів C за експериментальними даними. Для проведення таких досліджень потрібні значні кошти, матеріальні та трудові витрати, оскільки серед полімерних композитних матеріалів поліімідні матеріали відрізняються відносно високою ціною. При цьому ціна високодисперсних порошків наповнювачів поліімідних композитних плівок може бути вищою, ніж у крупнодисперсних у кілька разів, а ціна нанопорошків вище ціни крупнодисперсних (десятки мкм) порошків майже на два порядки. Таким чином, якщо кількість наповнювача в композиті складатиме понад 50 мас. %, то вартість нового композитного матеріалу, що розробляється, значною мірою визначатиметься вартістю наповнювачів. Загальна ж вартість матеріалу при цьому може різко зрости, тим самим зводячи нанівець одну з головних можливих переваг нових композитних матеріалів, що розробляються, перед їх аналогами – нижчу вартість. Крім того, в тих випадках, коли потрібно виконання розрахунку з можливістю зміни концентрацій частинок наповнювача в широкому діапазоні або варіювання форми та розміри частинок (одиниці мікрон, ультрамікронні або наночастинки), аналітичні моделі не можуть забезпечити достатню достовірність результатів.

У цій ситуації найбільш доцільним і надійним методом (крім експериментального) визначення ефективної теплопровідності нових поліімідних матеріалів є метод прямого моделювання, тобто чисельний метод розрахунку ефективного коефіцієнта теплопровідності поліімідного композитного матеріалу з урахуванням граничних і початкових умов [4]. Суть методу полягає у безпосередньому вирішенні рівняння теплопровідності з урахуванням граничних та початкових умов кінцево-різницевими методами. Використання чисельних методів для знаходження ефективної теплопровідності композиційної системи дозволяє проводити більш точні розрахунки для різних розподілів наповнювача в матриці, що дозволяє враховувати вплив структури на теплові властивості композиту, а також наявність граничного шару.

Метою роботи було дослідження надійного, відносного простого та автоматизованого теоретичного методу визначення ефективної теплопровідності нових поліімідних теплопро-

відних композитних плівок за допомогою програмного комплексу COMSOL MULTIPHYSICS®.

У зв'язку з цим було проведено чисельне моделювання теплопровідності композитних матеріалів при введенні в поліімідну матрицю високотеплопровідних частинок різних порошкових наповнювачів з урахуванням граничних та початкових умов. При цьому варіювалася об'ємна концентрація сумішей наповнювачів у них. Було проаналізовано вплив теплопровідності частинок наповнювача та їх розмірів на ефективну теплопровідність композитних поліімідних матеріалів.

Запропоновано конкретні рекомендації підвищення ефективної теплопровідності від 0,12 до 1 – 4 Вт/(м•К) поліімідних композитних плівок шляхом зміни розмірів, концентрації, теплопровідності частинок наповнювача, типів матеріалів порошків наповнювача з урахуванням особливостей розподілу сумішей наповнювачів у поліімідній матриці.

Численний метод розрахунку ефективного коефіцієнта теплопровідності поліімідного композитного матеріалу з урахуванням граничних і початкових умов

Застосування методу прямого моделювання теплопровідності середовищ зі складною структурою дозволяє проводити розрахунки ефективної теплопровідності для різних розподілів частинок порошків наповнювачів у поліімідних сполучних. При цьому є можливість зміни концентрації частинок наповнювачів у матриці у широкому діапазоні або варіювання розміру частинок, що дозволяє врахувати вплив структури та граничних шарів на теплові властивості композитів.

У цьому випадку композитна система (поліімідний композитний матеріал), теплопровідність якої необхідно розрахувати, моделюється у вигляді куба, розбитого на комірки [4 – 7].

На рис. 1 зображено модельне уявлення композитної системи у вигляді куба, розбитого на елементарні комірки.

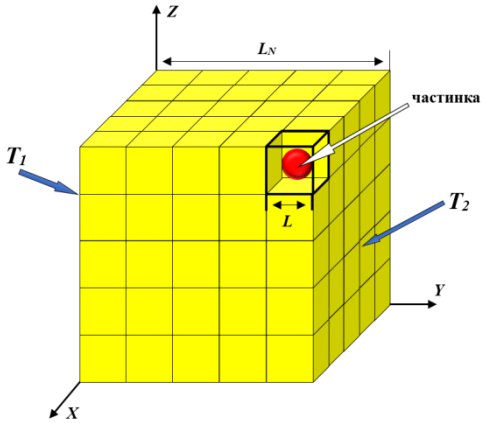
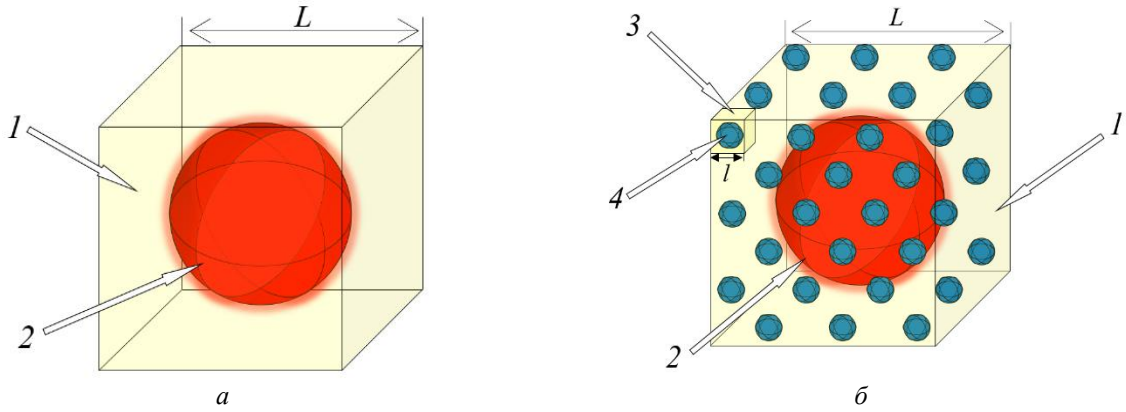


Рис. 1. Модель композитної системи у вигляді куба, розбитого на елементарні комірки

На протилежних бокових гранях куба композитної системи з довжиною ребра L_N задаються різні температури T_1 та T_2. В результаті вздовж одного із напрямків створюється температурний градієнт. Тепловий потік через інші грані куба відсутній.

При моделюванні розглядався рівномірний розподіл частинок. Об’єм матриці композиту можна уявити у вигляді декількох елементарних комірок, у кожній з яких поміщена одна або кілька частинок наповнювача необхідної форми (рис. 2). На рис. 2, *а* зображено елементарну кубічну комірку с довжиною ребра L, яка заповнена однією сферичною частинкою мікронного розміру наповнювача з діаметром $Ø = L$. При цьому максимальна об'ємна концентрація матеріалу сферичної частинки наповнювача в такій елементарній кубічній комірці композитної системи становить не більше 52,4 об. %.

На рис. 2, *б* зображена елементарна кубічна комірка з довжиною ребра *L*, заповнена однією сферичною частинкою з діаметром *Ø = L* і додатковими кількома елементарними кубічними комірками з довжиною ребра *l*, які заповнені однією сферичною часткою субмікронного розміру з діаметром *Ø = l*.



а
Комірка, яка заповнена однією сферичною частинкою мікронного розміру

б
Комірка, яка заповнена однією сферичною частинкою мікронного розміру та додатковими кількома частинками субмікронного розміру

Рис. 2. Модельне уявлення рівномірного розподілу частинок наповнювачів:
1,3 – елементарні комірки матриці композитної поліімідної плівки; *2* – наповнювач у формі сферичної частинки мікронного розміру; *4* – наповнювач у формі сферичної частинки субмікронної розміру

Для спрощення розрахункової моделі було прийнято ряд основних допущень:

- задачі розглядаються як стаціонарні, оскільки вирівнювання температури в теплопровідному об’ємі відбувається значно швидше, ніж зміна зовнішніх умов;
- матеріали, складові композита, вважаються ізотропними. При цьому їх теплопровідність не залежить від температури;
- розміри частинок кожного типу наповнювача вважаються однаковими. Поодинокі елементарні кубічні комірки з довжиною ребра *L* і *l* заповнені однією частинкою наповнювача сферичної форми мікронного або субмікронного розміру відповідно. Елементарні кубічні комірки рівномірно розподілені у всьому об’ємі композитної системи. Фракційний склад частинок не враховується;
- тепловий контакт між частинками та полімером приймається ідеальним (дане припущення тягне за собою зростання теплопровідності композиту внаслідок нульового теплового опору на межі середовищ полімер – частинка);
- наявність мікродефектів не враховується.

Параметризація чисельної моделі дозволяє варіювати властивості композитної системи; в даному випадку це розміри частинок, властивості їх матеріалів та властивості поліімідної матриці. Відповідно до моделі зв'язок між довжиною *L* ребра куба елементарної комірки (вона визначає розмір розрахункової області та розмір частинки) та об’єм частинки V_f задаються формулою

$$L = \left(\frac{V_f}{N_v} \right)^{\frac{1}{3}}, \tag{1}$$

де N_v – об'ємна концентрація частинок.

Ефективна теплопровідність поліімідного композитного матеріалу

$$\lambda_{eff} = q_m \frac{L}{|T_1 - T_2|}, \tag{2}$$

де λ_{eff} – ефективна теплопровідність композитного матеріалу; q_m – середній тепловий потік; L – довжина ребра куба; $T_1 - T_2$ – різниця температур протилежних граней.

З метою визначення величини q_m виконується чисельне рішення тривимірної стаціонарної задачі теплопровідності:

$$\cdot \frac{\partial}{\partial x}\left(\lambda \frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(\lambda \frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(\lambda \frac{\partial T}{\partial z}\right)=0 \quad (3)$$

Коефіцієнт теплопровідності у цьому рівнянні залежить від координат: $\lambda = \lambda\ (x, y, z)$.

На двох поверхнях куба задаються граничні умови першого роду, на інших чотирьох накладаються умови теплоізоляції [7]. Сучасні обчислювальні засоби дозволяють проводити пряме моделювання теплопровідності середовищ зі складною структурою. При цьому автори використали для автоматизації розрахунку середнього теплового потоку q_m програмний комплекс COMSOL MULTIPHYSICS.

Результати розробки теоретичних моделей поліімідних композитних плівок на основі поліімідних лаків і теплопровідних порошків наповнювачів

Відповідно до класифікації дисперсні наповнювачі за розміром частинок d поділяються на крупнодисперсні (d більше 40 мкм), середньодисперсні (d від 10 до 40 мкм), високодисперсні (d від 1 до 10 мкм) та ультрадисперсні (d менше 1 мкм). До цієї класифікації належить група нанодисперсних порошків (d менше 0,1 мкм).

У розрахунках використовувалися наповнювачі, що належать до групи високодисперсних та ультрадисперсних порошків. Вибір таких розмірів частинок наповнювачів обумовлений прагненням отримати на практиці необхідні значення електрофізичних і механічних характеристик поліімідних композитів при мінімальній вартості матеріалів, оскільки вартість компонентів може істотно впливати на підсумкову вартість композитів.

При аналізі характеру теплоперенесення у дисперсно-наповнених композитах слід враховувати, що теплоперенесення через великі частинки мікронного розміру відбувається легше, ніж через дрібні частинки. Це пов'язано з меншою протяжністю меж між великими частинками.

З іншого боку, дрібні частинки при ідентичних ступенях наповнення розташовані ближче один до одного в композиті, що означає формування більш тонкого міжфазного полімерного шару. Це дозволяє зменшити граничний тепловий опір наповнювач-матриця і, як результат, теплопередача через сітку частинок наповнювача стає більш ефективною.

При високій концентрації теплопровідних частинок механічні властивості композиту різко погіршуються, і матеріал стає жорстким і крихким. Тому створення поліімідного композиту, що поєднує в собі хороші механічні та теплопровідні властивості, є досить непростою задачею. Основна ідея ефективного управління теплофізичними характеристиками композитного матеріалу при високих ступенях наповнення полягає у максимізації теплопровідних шляхів поряд з мінімізацією граничного теплового опору наповнювач-наповнювач і наповнювач-матриця.

Таким чином, для розробки структурних моделей варіантів високонаповнених теплопровідних електроізоляційних композитних поліімідних матеріалів були обрані порошки наповнювачів SiO_2, SiC, Al_2O_3 і AlN з дрібнодисперсними і ультрадисперсними розмірами частинок. В якості поліімідного сполучного був обраний поліімідний термореактивний лак типу Pyre ML RC 5069, який є базовим матеріалом для виготовлення поліімідних плівок типу DuPont™ Kapton® HN, що мають високу термічну стабільність, але низьку теплопровідність (0,12 Вт/(м·К).

Такий вибір обумовлений необхідністю підтримки допустимих електрофізичних і механічних властивостей при розробці нових теплопровідних поліімідних композитних матеріалів.

В даний час теплопровідні електроізолюючі поліімідні плівки, що промислово випускаються, мають типову теплопровідність від 0,36 до 0,8 Вт/(м·К) при товщинах від 25 до 75 мкм (1 – 3 mil, 1 mil ~ 25 мкм) і при цьому вони зберігають високу напругу електричного пробою, механічну стійкість та гнучкість [8, 9].

Тому для побудови структурних моделей та оцінки теплопровідності нових типів поліімідних композитних матеріалів з очікуваною теплопровідністю до 1,0 Вт/(м·К) і більше товщина поліімідної плівки вибиралася не менше 25 мкм. Частинки порошків теплопровідних наповнювачів, що мають форму кулі, мали діаметри, кратні товщині композиту та становлять для дрібнодисперсних порошків 8 мкм, а для ультрадисперсних порошків – 0,8 і 0,4 мкм.

Такий підхід дозволяє уявити модель ділянки композитної системи у вигляді куба з довжиною ребра L_N = 40 мкм, що має структуру кубічних ґрат 5 х 5 х 5 з елементарними кубічними комірками з довжиною ребра L = 8 мкм, які рівномірно заповнені частинками. При цьому загальна кількість елементарних комірок у кубічних ґратах такої ділянки композиту становить 125 комірок (рис. 3).

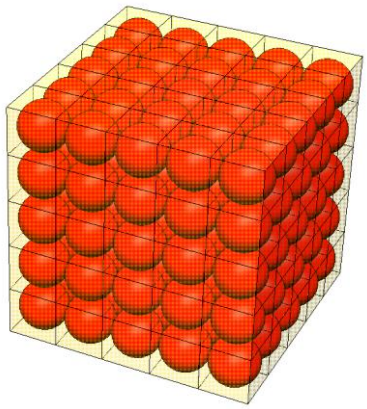


Рис. 3. Модель кубічних грат ділянки композитної системи

У табл. 1 представлено вихідні дані для розгляду різних варіантів структурних моделей поліімідних теплопровідних композитних плівок.

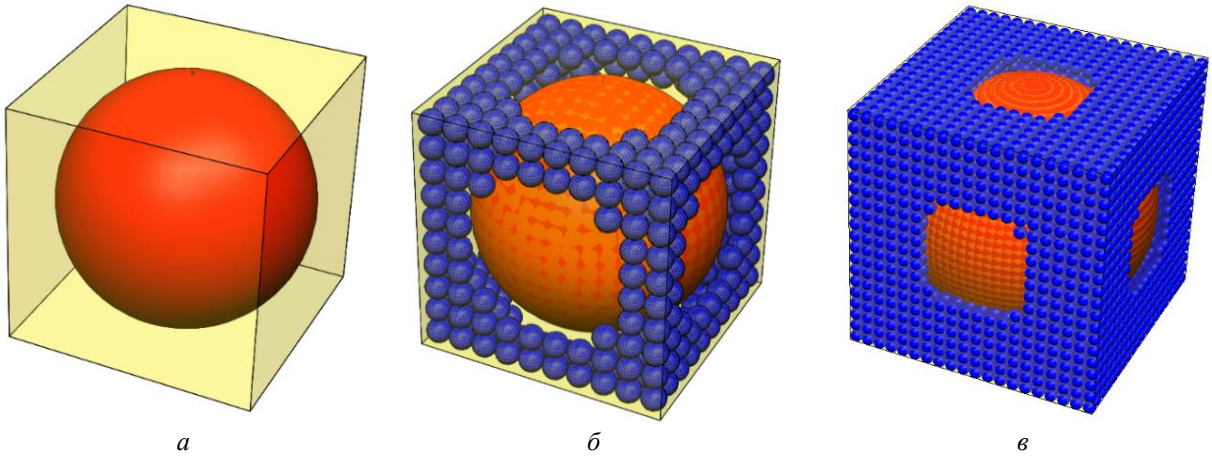
Температурний градієнт ($T_1 – T_2$) протилежних граней кубічних ґрат композитів задавався при припущенні, що температура T_2 відповідає температурі навколишнього середовища T_a = 25 °С. Температура T_1 відповідає рекомендованій максимальній робочій температурі області *p-n*-переходу напівпровідникових джерел тепла і становить T_j = 80 °С при стабільному стані теплового розподілу на кожній досліджуваній ділянці композитної системи.

Таблиця 1

Структурні моделі поліімідних теплопровідних композитних плівок
з різною об'ємною концентрацією наповнювачів у композитній системі

Типи порошків наповнювачів	$\lambda_н$, Вт/(м•К)	$\lambda_{св}$, Вт/(м•К)	T_1, °С	T_2, °С	Об'ємна концентрація частинок N_v,об. %		
					Варіант 1	Варіант 2	Варіант 3
					8 мкм	8 мкм + 0.8 мкм	8 мкм + 0.4 мкм
SiO_2	7	0,12	80	25	0 – 52,4	0 – 69,5	0 – 73,5
Al_2O_3	11	0,12					
SiC	25	0,12					
AlN	55	0,12					

На рис. 4 представлено варіанти досліджуваної структурної 3D-моделі елементарних комірок поліімідних композитних плівок.



а *б* *в*

Рис. 4. Досліджувана структурна 3D модель комірок поліімідних композитних плівок:
а – варіант 1; *б* – варіант 2; *в* – варіант 3

Для варіанта 1 одна елементарна комірка кубічної ґратки з довжиною ребра 8 мкм має об'єм рівний 512 мкм^3, а одна сферична частка діаметром 8 мкм в комірці має об'єм рівний 268 мкм^3 (рис. 4, *а*). При цьому максимальна об'ємна концентрація матеріалу частинки наповнювача у такій елементарній комірці композитної системи складе не більше 52,4 об. %.

Для досягнення об'ємної концентрації частинок наповнювача більш ніж 52,4 об. % у структурну 3D-модель елементарної комірки були додані додаткові ультрадисперсні частинки. Таким чином, для варіанта 2 одна елементарна комірка кубічної ґратки з довжиною ребра 8 мкм має об'єм рівний 512 мкм^3, одна сферична частка діаметром 8 мкм в комірці має об'єм 268 мкм^3. Додані сферичні частинки діаметром 0,8 мкм (328 шт.) у комірці мають об'єм рівний 88 мкм^3 (рис. 4, *б*). При цьому максимальна об'ємна концентрація матеріалу частинок наповнювача в такому комбінованому варіанті елементарної комірки композитної системи становитиме 69,5 об. %.

Для варіанта 3 одна елементарна комірка кубічної ґратки з довжиною ребра 8 мкм має об'єм 512 мкм^3, одна сферична частка діаметром 8 мкм в комірці має об'єм 268 мкм^3. Додані сферичні частинки діаметром 0,4 мкм (3224 шт.) у комірці мають об'єм 108 мкм^3 (рис. 4,*в*). При цьому максимальна об'ємна концентрація матеріалу частинок наповнювача в такому комбінованому варіанті елементарної комірки композитної системи становитиме 73,5 об. %.

За допомогою програмного комплексу COMSOL MULTIPHYSICS® та формули (2) були виконані 3D-структурне моделювання комірок поліімідних композитних плівок і розрахунки середніх теплових потоків та ефективних коефіцієнтів теплопровідності поліімідних композитних плівок для варіантів 1, 2, 3.

На рис. 5 показано результати розрахунку ефективної теплопровідності поліімідних композитних плівок для різних наповнювачів та об'ємної концентрації частинок наповнювачів.

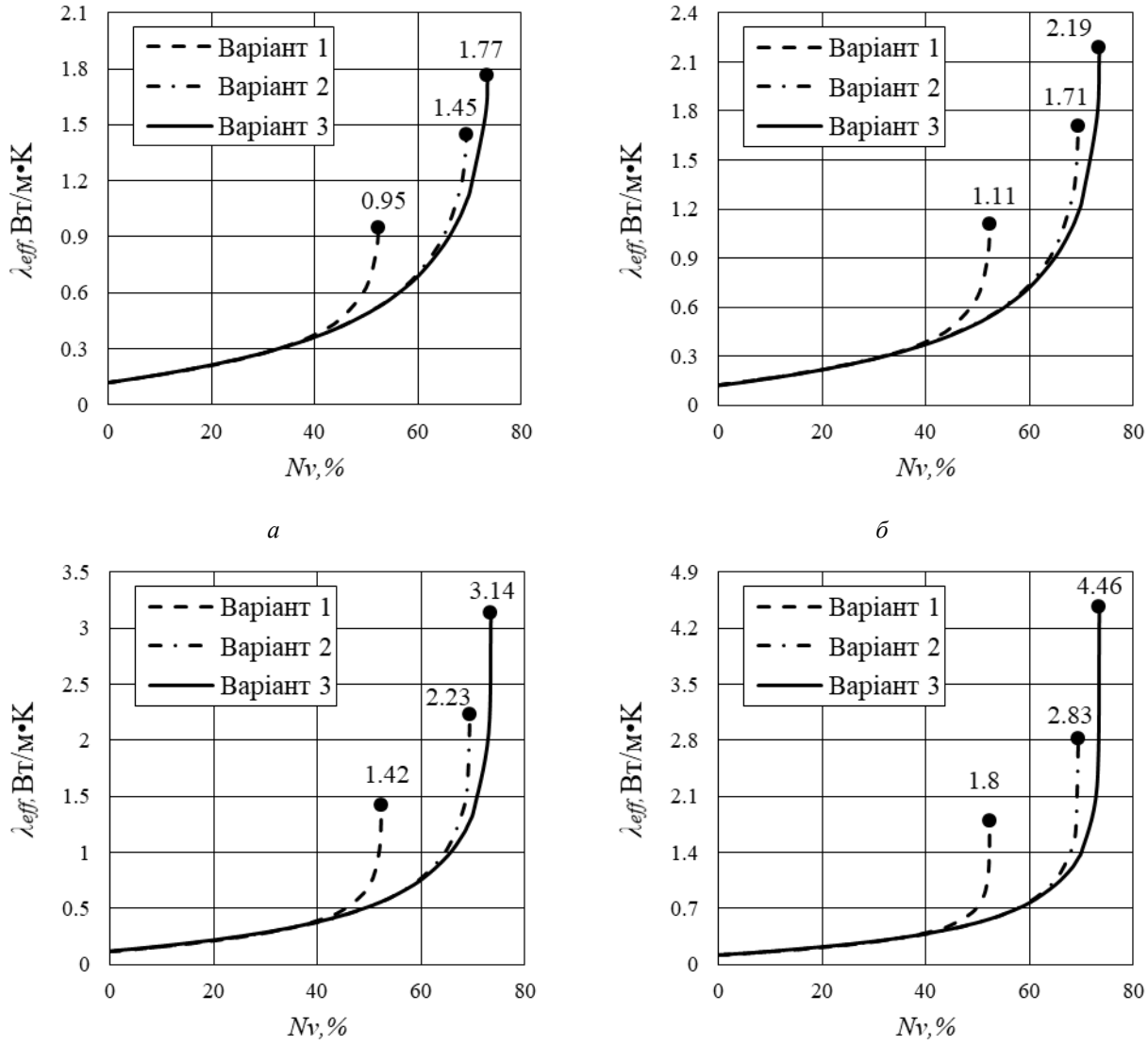


Рис. 5. Розрахункові значення ефективної теплопровідності поліімідних композитних систем для різних порошкових наповнювачів залежно від об'ємної концентрації частинок наповнювачів: *а – SiO_2; б – Al_2O_3; в – SiC; г – AlN*

Варіювання кількістю наповнювача у поліімідній матриці дозволило встановити, що більш високі теплопровідні властивості досягаються при вмісті частинок наповнювача в поліімідному композиті не менше 52,4 об. % та зростають при подальшому збільшенні вмісту наповнювача.

Максимальна ефективна теплопровідність спостерігалася для варіантів композитних матеріалів, в яких у якості наповнювача використовується суміш мікронних і ультрамікронних частинок порошків наповнювачів з переважним вмістом мікрочастинок.

При загальному об'ємі концентрації наповнювачів від 52,4 до 69,5 об. % у композитах при складі сумішей порошків близько 75 % із середніми розмірами частинок до 8 мкм та 25 % із середніми розмірами частинок до 0,8 мкм підтверджена чисельними розрахунками очікувана максимальна ефективна теплопровідність поліімідних плівок в діапазоні значень від 0,95 до 2,83 Вт/(м•К) для досліджених порошків наповнювачів із SiO_2, SiC, Al_2O_3 та AlN.

При загальному об'ємі концентрації наповнювачів від 52,4 до 73,5 об. % у композитах при складі сумішей порошків близько 71 % із середніми розмірами частинок до 8 мкм та 29 % із середніми розмірами частинок до 0,4 мкм підтверджена чисельними розрахунками

очікувана максимальна ефективна теплопровідність поліімідних плівок в діапазоні значень від 0,95 до 4,46 Вт/(м•К) для досліджених порошків наповнювачів із SiO_2, SiC, Al_2O_3 та AlN.

При цьому комбінована поліімідна композитна система із застосуванням суміші високотеплопровідних високодисперсних (8 мкм) та ультрадисперсних (0,4 мкм) порошкових наповнювачів з AlN показала кращі оціночні результати максимальної ефективної теплопровідності (λ_{eff} = 4,46 Вт/(м•К), варіант 3).

Висновки

Авторами розроблено структурні моделі та проведено чисельне моделювання теплопровідності композитних матеріалів при введенні в поліімідну матрицю високотеплопровідних частинок порошкових наповнювачів із SiO_2, SiC, Al_2O_3, AlN, з урахуванням граничних та початкових умов за допомогою програмного комплексу COMSOL MULTIPHYSICS.

В результаті чисельного моделювання було вирішено такі задачі:

- вивчено вплив теплопровідності частинок SiO_2, SiC, Al_2O_3, AlN, які включені у матрицю поліміду, на ефективну теплопровідність композитів;

- вивчено вплив об'ємної концентрації сферичних частинок з мікронними розмірами наповнювачів SiO_2, SiC, Al_2O_3, AlN на ефективну теплопровідність поліімідних композитних матеріалів;

- встановлено суттєвий вплив об'ємної концентрації сумішей сферичних частинок з мікронними та ультрамікронними розмірами наповнювачів SiO_2, SiC, Al_2O_3, AlN, на ефективну теплопровідність поліімідних композитних матеріалів.

Аналіз отриманих результатів дозволив зробити висновки, що:

– ефективна теплопровідність поліімідних композитів визначається величинами теплопровідності як поліімідної матриці, так і наповнювача;

– максимальна ефективна теплопровідність спостерігалася для варіантів моделей композитних матеріалів, в яких в якості наповнювача використовуються суміші мікронних і ультрамікронних частинок порошків наповнювачів з переважним вмістом мікронних частинок;

– при загальній об'ємній концентрації наповнювачів від 52,4 до 69,5 об. % у композитах при складі сумішей порошків близько 75 % із середніми розмірами частинок до 8 мкм та 25 % із середніми розмірами частинок до 0,8 мкм розрахунками підтверджена очікувана максимальна ефективна теплопровідність композитних поліімідних плівок у діапазоні значень від 0,95 до 2,83 Вт/(м•К) для досліджених порошків наповнювачів із SiO_2, SiC, Al_2O_3 та AlN.

– при загальній об'ємній концентрації наповнювачів від 52,4 до 73,5 об. % у композитах при складі сумішей порошків близько 71 % із середніми розмірами частинок до 8 мкм та 29 % із середніми розмірами частинок до 0,4 мкм розрахунками підтверджена очікувана максимальна ефективна теплопровідність композитних поліімідних плівок у діапазоні значень від 0,95 до 4,46 Вт/(м•К) для досліджених порошків наповнювачів із SiO_2, SiC, Al_2O_3 та AlN.

Таким чином, за результатами роботи запропоновано конкретні рекомендації щодо проведення прямого моделювання теплопровідності полімерних середовищ зі складною структурою та чисельних розрахунків з достатньою достовірністю ефективної теплопровідності поліімідних композитних плівок, що розробляються, з метою збільшення їх теплопровідності від 0,12 до 1 – 4 Вт/(м•К) шляхом варіювання концентрації та теплопровідності сумішей частинок порошкових наповнювачів мікронних та ультрамікронних розмірів. Комбінована поліімідна композитна система із застосуванням суміші високотеплопровідних високодисперсних (8 мкм) та ультрадисперсних (0,4 мкм) порошкових наповнювачів з AlN показала кращі оціночні результати максимальної ефективної теплопровідності – λ_{eff} = 4,46 Вт/(м•К).

Список літератури:

1. New approaches to creating promising heat-conductive electrical insulating polyimide nanocomposite materials / Borshchov V.M., Listratenko O.M., Protsenko M.A., Tymchuk I.T., Kravchenko O.V., Syddia O.V., Slipchenko M.I., Chichkov B.M. // Functional Materials. 2022. Vol.29, No.1.P. 20 – 29

2. Міхєєв В.А. Забезпечення якості нових функціональних матеріалів для теплопровідних покриттів на стадії розробки та виробництва : дис. … канд. техн. наук. 2018. 173 с. (рос. мовою).

3. Сулаберідзе В.Ш., Скорнякова Є.А. Оцінка параметрів розрахункових моделей теплопровідності композиційних матеріалів з полімерним сполучним за експериментальними даними // Вісник Магнітогор. держ. техн. ун-ту ім. Г.І. Носова. 2020. Т.18. №4. С. 57 – 64. https://doi.org/10.18503/1995-2732-2020-18-4-57-64. (Рос. мовою).

4. Нікітін А. В. Чисельний метод розрахунку коефіцієнта теплопровідності наповнених полімерів / А. В. Нікітін, А. Ю. Бачуріна // Вісник Гродн. держ. ун-ту імені Янки Купали. Сер. 2. Математика. фізика. інформатика, обчислювальна техніка та управління. 2011. С. 106 – 111. (Рос. мовою).

5. Нікітін Д.А. Моделювання структури композиційних систем та розрахунок їх коефіцієнта теплопровідності // Матеріали. Технології. Інструменти. 2004. Т. 9, № 2. С. 11 – 15. (Рос. мовою).

6. Бачуріна А.Ю., Нікітін А.В. Чисельний метод розрахунку коефіцієнта теплопровідності композиційної системи // Вісник Гродн. держ. ун-ту. Сер. 2. Математика. фізика. Інформатика, обчислювальна техніка та управління. Біологія 2010. №2. С.93 – 99. (Рос. мовою).

7. Степанов В.В. Петреня Ю.К., Андрєєв А.М., Костєльов А. М., Маннанов Е.Р., Талалов В.А. Вплив властивостей компонентів на ефективну теплопровідність полімерних композитних матеріалів // Наук.-техн. відомості СПбГПУ. Фіз.-мат. науки. 2018. Т. 11. № 4. С. 85 – 94. DOI: 10.18721/JPM.11408. (Рос. мовою).

8. Теплопровідна поліімідна плівка DuPont™ Kapton® MT, https://www.dupont.com/products/kapton-mt.html // офіційний сайт (дата звернення 03.10.2022).

9. Теплопровідна поліімідна плівка DuPont™ Kapton® MT +, https://www.dupont.com/products/kapton-mt-plus.html. // офіційний сайт (дата звернення 03.10.2022).

Надійшла до редколегії 07.10.2022

Відомості про авторів:

Борщов Вячеслав Миколайович – д-р техн. наук, професор, ТОВ «Науково-виробниче підприємство «ЛТУ», перший заступник директора – головний конструктор; Україна; e-mail: viatcheslav.borshchov@cern.ch; ORCID: https://orcid.org/0000-0002-5579-8932

Лістратенко Олександр Михайлович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», провідний науковий співробітник; Україна; e-mail: sasha.listratenko.12@gmail.com; ORCID: https://orcid.org/0000-0001-7643-5295

Проценко Максим Анатолійович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», начальник відділення – заступник головного конструктора; Україна; e-mail: max.protsenko.1978@gmail.com; ORCID: https://orcid.org/0000-0001-9313-1701

Тимчук Ігор Трохимович – канд. техн. наук, ТОВ «Науково-виробниче підприємство «ЛТУ», головний технолог; Україна; e-mail: ihortymchuk78@gmail.com; ORCID: https://orcid.org/0000-0002-6436-7253

Кравченко Олександр Вікторович – ТОВ «Науково-виробниче підприємство «ЛТУ», заступник начальника відділу; Україна; e-mail: kravcenkoaleksandr671@gmail.com; ORCID: https://orcid.org/0000-0002-7145-4304

Суддя Олександр Валерійович – ТОВ «Науково-виробниче підприємство «ЛТУ», науковий співробітник; Україна; e-mail: 4el1195@gmail.com; ORCID: https://orcid.org/0000-0002-2403-979X

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 623.373.072.9 **DOI:10.30837/rt.2022.4.211.11**

V.V. RAPIN, Doctor of Science

THEORETICAL INVESTIGATION OF INJECTION-LOCKED DIFFERENTIAL OSCILLATOR

Introduction

Application of ingection-locked differential LC oscillator began at the beginning of this century, simultaneously with the advent of an autonomous differential LC oscillator. Currently, synchronized differential LC oscillators are an essential part of not only communication systems, but also many automation and measuring devices. The synchronization mode allows significantly expanding the functional posibility of oscillators. Practical application was found not only for separate fundamentally synchronized oscillators but also for the frequency division and multiplication modes, and more complex devices including several coupled oscillators [1 – 12].

The study of the synchronized differential oscillators is a difficult problem, and a satisfactory method for performing this kind of work has not yet been developed. The lack of an adequate mathematical model, permitting the application of rigorous mathematical methods of the nonlinear theory of electrical oscillations, was the reason for using a simplified approach. Usually it was reduced to a transition to a simpler single- circuit LC oscillator, being equivalent to a differential one. However, no justification for such a transition was provided.

The circuit diagram of a single-circuit LC-oscillator was shown as a parallel LC circuit connected to a bipolar active element, being a current generator. At first it was considered that the current had the form of square wave, according to the operating mode of the differential oscillator transistors, which are represented as ideal switches, distributing the direct current I_0 (the bias current, which is considered known) between differential oscillator branches. The current of this source varied from $+I_0$ to $-I_0$, that was an idealization, but made it possible to determine the amplitude of the first harmonic and, as a consequence, the amplitude of the oscillator signal. Such a range of current variation provided an approximate equality of the signals amplitudes of the differential oscillator and its equivalent.

Many authors proposed a mathematical model of the equivalent oscillator in the form of nonlinear differential equations, representing the nonlinear characteristic of the active element by a piecewise constant function I_0 sgn(), where the argument was the oscillator signal. Instead of such an expression for the nonlinear characteristic, the first harmonic current component of the active element could also be used directly, which was easily determined. Such representations were used in many papers, for example [1, 2, 5].

Improving the differential oscillator study results is also associated with the equivalent oscillator, as in the previous cases, but with the representation of the nonlinear characteristic of the active element by a polynomial of the third degree [3 – 9, 12]. In the works, it was indicated without details, that this nonlinear characteristic was determined by modeling the differential oscillator in the Spice simulator and was quite accurately approximated by the above mentioned polynomial. However, the method of experimental determination was not presented, despite the fact that this nonlinear characteristic had a very specific form. Besides, a relation of the the equivalent oscillator active element nonlinear characteristic and the nonlinear characteristics of the two amplifying elements of the differential oscillator has not been established.

The above problems, solved for an autonomous differential oscillator [13], are partially used in this paper.

The analysis of publications devoted to synchronized differential oscillators leads to a conclusion that the problem of a rigorous theoretical study of such devices has not been satisfactorily resolved.

Thus, the purpose of the paper is to study the fundamentally injected differential LC -oscillator by rigorous mathematical methods of the nonlinear theory of electrical oscillations.

Synchronized oscillator equation

Let us consider the circuit diagram of the oscillator shown in Fig. 1, which consists of two connected identical single-circuit LC- oscillators. The equations describing the operation of the differential oscillator are derived under the following commonly used assumptions: u_0=const, I_0=const, the transistors are identical and are inertialess amplifying elements, the influence of their input and output resistances can be neglected. The nonlinear characteristics of transistors are approximated by a polynomial of the third degree $i = a_0 + a_1u_y + a_2u_y^2 + a_3u_y^3$, where u_y is the control voltage. For transistor M_1 $u_y = u_{g1} = V_{DD} - u_0 + e$ and for transistor M_2 $u_y = u_{g2} = V_{DD} - u_o + u$, where u_{g1} is the gate voltage of transistor M_1 and u_{g2} is the voltage at the transistor gate M_2, u *and* e are variable voltage components.

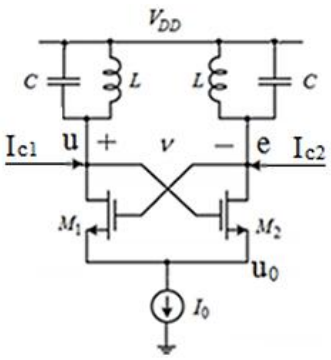


Fig. 1 Circuit diagram of the oscillator

Using Kirchhoff's laws, the system of nonlinear differential equations of the differential oscillator can be represented as:

$$\begin{aligned} &\frac{d^2u}{dt^2} + \frac{1}{C}\frac{d}{dt}\left(\frac{u}{R} - i_1\right) + \omega_0^2 u = \frac{1}{C}\frac{d}{dt} i_{c1}, \\ &\frac{d^2e}{dt^2} + \frac{1}{C}\frac{d}{dt}\left(\frac{e}{R} - i_2\right) + \omega_0^2 e = \frac{1}{C}\frac{d}{dt} i_{c2}, \end{aligned} \tag{1}$$

where i_1 and i_2 are the currents flowing through the nonlinear elements M_1 and M_2, R, ω_0 are the resonant resistance and frequency of the resonant circuits, $i_{c1} = I_{c1}cos*(\omega_c t + \varphi_1)$, $i_{c2} = I_{c2}cos*(\omega_c t + \varphi_2)$, $I_{c1} = I_{c2} = const$, $\omega_c \approx \omega_0$.

In accordance with the algorithm of the oscillator operation, the variable components of the voltages on the transistors drains are antiphase, i.e. $e = -u$, and the information parameter is their difference $v = u - e$. Subtracting the second equation from the first equation of the system (1), we obtain the differential equation for the transition to an equivalent oscillator

$$\frac{d^2v}{dt^2} + \frac{1}{C}\frac{d}{dt}\left(\frac{v}{R} - \left(i_{1(u_{g1})} - i_{2(u_{g2})}\right)\right) + \omega_0^2 v = \frac{1}{C}\frac{d}{dt}\left(i_{c1} - i_{c2}\right) \tag{2}$$

where $i_{1(u_{g1})} - i_{2(u_{g2})} = i$ is the current of the equivalent oscillator active element. The problem is to find a function $i_{(v)}$, so that $i_{(v)}$=$i_{1(u_{g1})}$−$i_{2(u_{g2})}$.

After solving this problem, the equation (2) will describe the operation of the oscillator which is an equivalent to the differential one, but having one resonant circuit and one amplifying element. Moreover, this circuit is identical to the resonant circuits of the differential oscillator, and the nonlinear characteristic of the active element is different from the nonlinear characteristics of the differential oscillator amplifying elements and depends on them. Thus, equation (2) turns into the Van Der Pol equation and can be studied by rigorous mathematical methods of the theory of nonlinear electrical oscillations.

Nonlinear characteristic of the equivalent oscillator amplifying element

Let us start with the nonlinear characteristics of the differential oscillator amplifying elements being identical and forming nonlinear characteristic of the equivalent oscillator hypothetical amplifying element. Due to this identity, as well as for simplicity and clarity, we will proceed from the nonlinear characteristic of one amplifying element, shown in Fig. 2. However, we take into account that the parameters u and e are antiphase. First of all, according to [13], it is necessary to determine positions of the differential oscillator operating points.

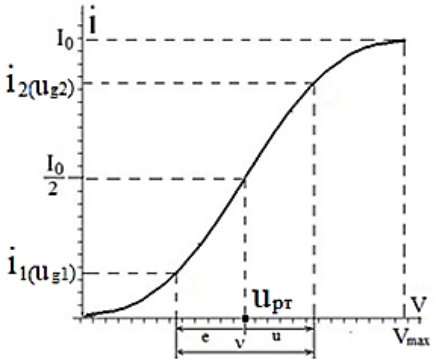


Fig. 2 Nonlinear characteristic of the amplifying element

Obviously, $\nu = 0$, only if $u = 0$ and $e = 0$. In this case, the voltages at the gates of the transistors are equal to the voltages on their drains and are described by the expression $u_{g10} = u_{g20} = V_{DD} - u_0$. Then, the transistor currents are also the same and equal to $I_0/2$. Consequently, the operating points of the amplifying elements of the differential oscillator are described by the parameters $u_{\text{pr}} = u_{g10} = u_{g20} = V_{DD} - u_0$ and $I_{\text{pr}} = I_0/2$. In Figure 2, the operating point is indicated by the symbol u_{pr}.

Next, in Fig. 2 we mark points $u_{g1} = V_{DD} - u_0 - e$ and $u_{g2} = V_{DD} - u_o + u$, where $u_{g1} < u_{g2}$, and e and u are absolute values, then we find the corresponding currents $i_{1(u_{g1})}$ and $i_{2(u_{g2})}$ and determine their difference $i_{1(u_{g1})} - i_{2(u_{g2})} < 0$. This difference corresponds to a certain parameter $\nu = u_{\text{M1}} - u_{\text{M2}} = u_{g2} - u_{g1} = u + e > 0$, which allows writing $i_{1(\text{ug}_1)} - i_{2(\text{ug}_2)} = i_{(\nu)}$. These values permit to find the position of the point corresponding to the nonlinear characteristic of the equivalent oscillator amplifying element. Thus, for $u_{g2} - u_{g1} > 0$, $i_{(\nu)} = i_{1(\text{ug}_1)} - i_{2(\text{ug}_2)} < 0$. This case is shown in Fig. 2. So, the negative values of the difference $i_{1(u_{g1})} - i_{2(u_{g2})}$ correspond to the positive value of the parameter ν.

It is easy to see that a similar dependence can be obtained for the case $u_{g2} - u_{g1} < 0$ i.e. $\nu < 0$, where $i_{(\nu)} = i_{1(u_{g1})} - i_{2(u_{g2})} > 0$. This case will be shown in Fig. 2, if we swap the symbols e and u, as well as the symbols $i_{1(u_{g1})}$ and $i_{2(u_{g2})}$.

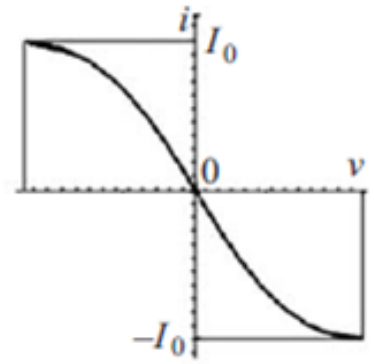


Fig. 3 Nonlinear characteristic of the equivalent oscillator amplifying element

Acting in accordance with this algorithm one sets different values of the parameters u_{g1} and u_{g2}, determines current values $i_{1(u_{g1})}$ and $i_{2(u_{g2})}$, as well as the difference $i_{1(\text{ug}_1)} - i_{2(\text{ug}_2)} = i_{(\nu)}$ and corresponding values of the parameter ν. As a result the function $i = f(\nu)$ is obtained, which describes the nonlinear characteristic of the active element. It is presented graphically in Fig. 3.

Obviously, this nonlinear characteristics belongs to an electronic device with negative differential resistance, and the equivalent oscillator is an oscillator with internal positive feedback. This circumstance will be taken into account in the differential equation (2), if the sign "minus", in front of the term containing the currents, is changed into the sign "plus". Then, the differential equation of the equivalent oscillator takes the form:

$$\frac{\text{d}^2\nu}{\text{dt}^2} + \frac{1}{\text{C}}\frac{\text{d}}{\text{dt}}\left(\frac{\nu}{\text{R}} + i_{(\nu)}\right) + \omega_0^2\nu = \frac{1}{\text{C}}\frac{\text{d}}{\text{dt}}i_{\text{c}}, \qquad (3)$$

where $i_{\text{c}} = i_{\text{c1}} - i_{\text{c2}} = 2I_{c1}\sin\left(\frac{\varphi_1-\varphi_2}{2}\right)\sin\left(\omega_c t + \frac{\varphi_1+\varphi_2}{2}\right)$.

It is easy to see that the values of the synchronizing signals phases have a very strong effect on the amplitude of the equivalent oscillator synchronizing signal. Let us consider the case when the phase difference is equal 180^0. Then

$$i_{\text{c}} = i_{\text{c1}} - i_{\text{c2}} = 2I_{c1}\cos(\omega_c t + \varphi_1) = I_c\cos(\omega_c t + \varphi_1),$$

where $I_c = 2I_{c1}$.

The obtained nonlinear characteristic, shown in Fig. 3, is symmetric about the origin, and it is an odd function. This also means that the bias is zero, which simplifies the investigation. For re-

search, this nonlinear characteristic is approximated by a polynomial of the third degree $i = -a_1 v + a_3 v^3$.

The equivalent oscillator mathematical model

Having an analytical expression describing the nonlinear characteristic of the amplifying element, the study is simplified because the equation is the Van der Pol equation with a positive feedback coefficient equal to one. The mathematical methods used in this case are rigorous and the methodology has been worked out and well tested. Then, the equation (3) can be represented in the form

$$\frac{d^2v}{dt^2} - \varepsilon\omega_0 \frac{dv}{dt}(v + \gamma v^3) + \omega_0^2 v = R\delta\omega_0 \frac{\mathrm{d}}{\mathrm{dt}} i_c, \tag{4}$$

where $\varepsilon = \delta\alpha$ is a small parameter, $\alpha = Ra_1 - 1$ is the regeneration coefficient, $\alpha_0' = -a_1 + 1/R$, $\gamma = a_3/\alpha_0'$, $\delta = 1/Q$, ω_0, R, Q are the resonant frequency of the oscillator circuit, its resistance and quality factor.

This differential equation in the most general form describes the processes in the oscillator and, to simplify the study, we present it in a dimensionless form. In the first approximation, we can consider that the oscillations are harmonic $v = A_0 \cos(\tau + \varphi_0)$ and find the oscillation amplitude of the autonomous oscillator in the steady state. Substituting this expression into the original equation (4) and taking into account only the components of the fundamental frequency, after simple transformations, we obtain:

$$A_0 = \sqrt{-4/(3\gamma)}$$

When choosing a method for solving equation (4), it is necessary to estimate the values of its terms. To do this we introduce dimensionless variables $\tau = \omega_c t$, $v_n = v/A_0 \leq 1$, $I_{cn} = I_c/I_0 \ll 1$, where $I_0 = A_0/R$, $i_c = I_{cn}\, cos(\omega_c t + \varphi_1)$, and take into account that $\omega_c \approx \omega_0$. Then, it is possible to write:

$$\frac{d^2v_n}{d\tau^2} - \varepsilon\left[\frac{d}{d\tau}\left(v_n - \frac{4}{3}v_n^3\right)\right] + \frac{\omega_0^2}{\omega_c^2} v_n = \delta \frac{\mathrm{d}}{\mathrm{dt}} i_c, \tag{5}$$

Now, when the values of the variable v_n are known, it is easy to see that, for small values of the small parameter and the small synchronization signal, this equation describes the behavior of a weakly nonlinear system and its solution, as is known, can be represented in the form $v_n = A_{(\tau)} \cos(\tau + \varphi_{(\tau)})$. A rigorous, well-established methods, such as the slowly varying amplitude method or the averaging method, can be used to find the amplitude and phase of oscillations. However, the above mentioned methods lead to systems of nonlinear shortened differential equations, presenting significant difficulties, when solving. An approximate analytical method, satisfying the needs of practice in most cases, developed recently, is given in [14].

Thus, the study of the synchronized differential oscillator is reduced to the study of the synchronized Van der Pol oscillator.

Conclusion

The paper presents the study of the fundamentally injected differential LC- oscillator by rigorous methods of the nonlinear theory of electrical oscillations in the case of small values of the small parameter. A research methodology has been proposed, an adequate mathematical model has been obtained. It makes it possible to study the differential oscillator as easily as the Van Der Pol oscillator. This model allows studying small but important effects, such as fluctuations of the amplitude and phase of oscillations and other significant parameters. It can be useful when developing devices using differential oscillators.

References:

1. Ahmad Mirzaei, Mohammad E. Heidari, Rahim Bagheri. Saeed Chehrazi, Asad A. Abidi. The Quadrature LC Oscillator: A Complete Portrait Based on Injection Locking // IEEE Journal of Solid-State Circuits, vol. 42, no. 9, pp. 1916 – 1932, 2007, doi: 10.1109/JSSC.2007.903047.
2. Antonio Buonomo, Michael Peter Kennedy, Alessandro Lo Schiavo. On the Synchronization Condition for Superharmonic Coupled QVCO // IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 58, no. 7, pp. 1637 – 1646, 2011, doi: 10.1109/TCSI.2011.2123370.
3. Antonio Buonomo, Alessandro Lo Schiavo. Modeling, Analysis, and Experimental Validation of Frequency Dividers with Direct Injection // Journal of Electrical and Computer Engineering, article ID 365692, 7 p., 2013, doi:org/10.1155/2013/365692.
4. A. Buonomo, A. Lo Schiavo. Divide-by-Three Injection-Locked Frequency Dividers with Direct Forcing Signal // Hindawi Publishing Corporation Journal of Electrical and Computer Engineering, article ID 145314, 9 p., 2013, doi: org/10.1155/2013/145314.
5. Antonio Buonomo, Alessandro Lo Schiav. Analytical Approach to the Study of Injection-Locked Frequency Dividers // IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 60, no. 1, pp. 51 – 62, 2013, doi:10.1109/TCSI.2012.2215716.
6. Mihaela Izabela Ionita, David Cordeau, Jean Marie Paillot, Mihai Iordache LAII. Analysis and Design of an Array of Two Differential Oscillators Coupled Through a Resistive Network // 20th European Conference on Circuit Theory and Design, Linkцping : Sweden. 2012. doi: 10.1109/ECCTD.2011.6043612.
7. M. Ionita, D. Cordeau, J. Paillot, S. Bachir, M. Iordache. A CAD tool for an array of differential oscillators coupled through a broadband network // Physics International Journal of RF and Microwave Computer Aided Engineering, vol. 23, no. 2, pp. 178 – 187, 2012, doi:10.1002/mmce.20663.
8. Mihaela-Izabela Ionita, Mihai Iordache, Dumitriu Lucia, David Cordeau, Jean-Marie Paillot. Generation of the coupling circuit parameters for the coupled oscillators used in antenna arrays // Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD). 2012 International Conference on, Sep.. S_eville. Spain. pp. 237 – 240, 2012, doi: 10.1109/SMACD.2012.6339383.
9. David Cordeau, Mihaela-Izabela Ionita, Jean-Marie Paillot, Mihai Iordache. New Formulation of the Equations Describing the Locked States of Two Van der Pol Oscillators Coupled through a Broadband Network – Application to the Design of Two Differential Coupled VCOs // Frequenz, vol. 67, no. 7-8, pp. 237-247, 2013,doi:10.1515/freq-2012-0089.
10. Sanmitra Bharat Naik, R. K. Siddharth, Anirban Chatterjee, Y. B. Nithin Kumar, M. H. Vasantha, Ramnath Kini. A 1.8 V Quadrature Phase LC Oscillator for 5G Applications // 2020 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS), pp. 345 – 349, 2020, doi: 10.1109/iSES50453.2020.00068.
11. Hao-En Liu, Wei-Cheng Chen, Hong-Yeh Chang. A W-band Quadrature Voltage-Controlled Oscillator with an Injection-Locked Frequency Divider in 40-nm CMOS Process // 2021 IEEE International Symposium on Radio-Frequency Integration Technology (RFIT). Hualien, Taiwan, pp. 1 – 3, 2021, doi: 10.1109/RFIT52905.2021.9565268.
12. Kaouthar Djemel, Rahma Aloulou, David Cordeau, Hassene Mnif, Jean-Marie Paillot, Dorra Mellouli, Mourad Loulou. An original determination of the maximum phase shift range obtained for an array of N coupled oscillators // Analog Integrated Circuits and Signal Processing, vol. 106, no. 3, pp. 683 – 696, 2021, doi: org/10.1007/s10470-020-01791-x.
13. Рапин В. Теоретичне дослідження дифференційного автогенератора // Известия высших учебных заведений. Сер. Радиоэлектроника, т. 65. no. 5. с. 309 – 319, 2022, doi: org/10.20535/S0021347022050041
14. Рапин В. Решение укороченных уравнений синхронизированного автогенератора // Известия высших учебных заведений. Сер. Радиоэлектроника, т. 62, no. 6, с. 335 – 347, 2019, doi: 10.20535/S0021347019060037.

Received 11.10.2022

About the author:

Rapin Vladimir Vasilyevich – Doctor of Science, associate professor, Kharkiv National University of Radioelectronics, professor of the Information and Network Engineering Department; Ukraine; e-mail volodymyr.rapin@nure.ua; ORCID: https://orcid.org/0000-0002-9773-7695

УДК 004.056.53 **DOI:10.30837/rt.2022.4.211.12**

О.Й. ДОВНАР, канд. техн. наук, М.Ф. БАБАКОВ, канд. техн. наук, В.І. ЧЕРКІС

ВИКОРИСТАННЯ СКАНЕРУ ВІДБИТКІВ ПАЛЬЦІВ ДЛЯ ЗАХИСТУ ДАНИХ У МЕДИЧНИХ ІНФОРМАЦІЙНИХ СИСТЕМАХ

Вступ

В умовах інформаційних війн та постійних DoS нападів бази та сервери підпадають під постійну загрозу. Більшість баз захищені лише стандартними системами авторизації, що може стати фатальним, особливо для закладів охорони здоров'я. Так, на офіційному інформаційному ресурсі Державної служби спеціального зв'язку та захисту інформації України [1] зазначено, що у ІІІ кварталі 2022 р. кількість кібератак на критичну інформаційну інфраструктуру України значно зросла і з 15-го лютого Україна зазнала понад 3000 DDoS-атак.

Таким чином, актуальність обмеження доступу до коитичної інформації наразі має найвищу степінь, а розробка систем або проистроїв може повністю змінити погляд на інформаційне устаткування та його захист. Задля убезпечення захисту інформації наразі застосовують DNS маски та блокування надмірних запитів. Це певною мірою покращує ситуацію або дозволяє швидше переналаштувати сервер, що знаходиться під атакою. Альтернативне рішення цієї проблеми (наприклад, для закладів охорони здоров'я) – це просте устаткування для авторизації сесії, а у всіх інших випадках – ігнорування або блокування доступу.

Рішення, яке може спростити та покращити заходи безпеки, – заміна технології, що використовує звичайну пару логін/пароль, зчитування відбитків пальців.

Ідентифікація за допомогою відбитків пальців – найбільш поширений біометричний спосіб визначення особистості. У сучасному світі відбитки пальців застосовується все більше, а саме: в криміналістиці, в медичній сфері, науці, в сучасній оборонній промисловості.

Метою розробки є забезпечення надійного захисту для серверів та баз даних шляхом ідентифікації користувачів до створення сесії.

Об'єктом для убезпечення доступу до аккаунту стане пристрій з можливостями сканування відбитків.

Вибір прототипу

Відбиток пальця – це відбиття, що залишає папілярні лінії (або візерунок) людського пальця. Відбиток пальця дуже легко залишається на відповідних поверхнях (наприклад, скло, метал або шліфований камінь) через фізіологічну властивість виділенню поту з екзокринних залоз, які присутні на епідермальній нерівності. Інколи існує можливість отримати неякісне зображення відбитка. Причинами цього є: жирна або ж суха шкіра рук, випадкові рухи під час зняття відбитка та великі пори. Використання неякісних зображень відбитків знижує прохідну здатність біометричних систем. Для уникнення таких проблем існують алгоритми визначення та поліпшення якості зображення відбитків пальців [2 – 7].

Біометричні прилади для доступу до баз даних зазвичай використовують FTIR сенсори або NFC датчики при авторизації за телефоном. Найчастіше використовують такі методи контроля доступу [6 – 8]:

- застосування ключ-карти;
- ідентифікація за райдужкою ока;
- ідентифікація за обличчям;
- застосування NFC датчика телефону.

У всіх цих методів є свої недоліки [9]. Так, вадою ключ-карти є простота копіювання, ідентифікація за райдужкою ока – складний та дорогий метод, ідентифікація скануванням обличчя не є достатньо достовірним методом. Окрім того, ці методи потребують драйверів для своєї працездатності. Таким чином, можна стверджувати, що використання сканеру відбитка пальця є найбільш придатним для рішення поставленої задачі ще й тому, що більшість питань стосовно його роботи та безпеки вже розглянуті.

Основним недоліком системи з авторизацією за відбитком є проблеми зі зчитуванням (у випадку відсутності відбитків людина не зможе пройти авторизацію), а також муляжі відбитків (залежить від схеми та типу сканеру).

Подібні схеми вже наявні на мобільних пристроях з скануванням відбитків. Їх основна мета – спростити та забезпечити безпеку, коли створюється нова сесія. При завантаженні додатка на мобільний пристрій записується відбиток пальця, а при наступній спробі відбувається ідентифікація – зчитування відбитку та порівняння його з записаним еталоном і у випадку успішної авторизації – створення сесії.

Що стосується комп’ютерів, то для подібної авторизації буде необхідний додатковий портативний пристрій – біометричний сканер відбитку та спеціальна програма для роботи з цим обладнанням, що ускладнює наш критерій простоти для роботи медичного персоналу. А влаштовані подібні рішення є лише у системі Windows 10, 11 у ideapad та thinkpad.

Таким чином, за технічними характеристиками необхідно розробити подібний до вбудованого пристрою сканер відбитків. Розглянемо основні типи сканерів, що застосовуються на даний час [10 – 15].

Оптичні сканери. Такі сканери працюють на основі використання оптичних методів отримання зображення. Існує кілька основних способів реалізації оптичного методу.

1. Оптичний метод відображення. У цьому методі використовується ефект повного внутрішнього відображення (Frusted Total Internal Reflection). Провідними виробниками таких сканерів є BioLink, Digital Persona, Identix.

2. Оптичний метод на просвіт. Сканери даного типу є оптоволоконною матрицею, в якій всі хвилеводи на виході з'єднані з фотодатчиками. Чутливість кожного датчика дозволяє фіксувати залишкове світло, що проходить через палець, у точці зіткнення пальця з поверхнею матриці. Цей метод характеризується високою надійністю зчитування та стійкістю до використання муляжів. Даний тип сканерів випускається американською компанією Security First Corp.

3. Оптичні безконтактні сканери. В оптичних безконтактних сканерах (touchless scanners) не потрібно безпосередньо контакту пальця з поверхнею скануючого пристрою. Провідний виробник сканерів цього типу Touchless Sensor Technology. Перевагами оптичних сканерів є відносно низька ціна та компактність. Недоліки: чутливість до забруднення, стану шкіри та слабка захищеність від муляжів та інших способів обману [8, 18 – 20]

Ємнісні сканери є найбільш поширеними напівпровідниковими пристроями для отримання зображення відбитка пальця. Їхня робота заснована на ефекті зміни ємності *p*-*n*-переходу напівпровідника при дотику гребеня папілярного візерунка з елементом напівпровідникової матриці. Перевагами таких сканерів є низька собівартість та висока надійність, а недоліками – слабка захищеність від муляжів. Провідними виробниками сканерів цього типу є Veridicom та STMicroelectronics

Радіочастотні сканери використовують матрицю елементів, що працюють як мініантени. Оскільки аналізуються фізіологічні властивості шкіри, ймовірність обману даного сканера прагне до нуля, але при поганому контакті з пальцем робота такого сканера може бути нестійкою. Відомим виробником радіочастотних сканерів є компанія Authentec.

Сканери, які використовують метод тиску у своїй конструкції, мають матрицю п'єзоелектричних елементів, чутливих до натискання. Чутливі до тиску сканери випускає компанія BMF. Недоліками таких сканерів є низька чутливість, неефективний захист від муляжів та схильність до пошкоджень при надмірно докладних зусиллях.

Термосканери використовують датчики, що складаються з піроелектричних елементів, та дозволяють фіксувати різницю температури і перетворювати її на напругу. Такий метод має безліч переваг: висока стійкість до електростатичного розряду, стійка робота в широкому температурному діапазоні, ефективний захист від муляжів. До недоліків цього методу можна віднести те, що зображення швидко зникає через те, що палець і датчик приходять до температурної рівноваги.

Ультразвукові сканери сканують поверхню пальця ультразвуковими хвилями. Перевагами таких сканерів є підвищена якість зображення та повний захист від муляжів. Недоліком є висока собівартість.

Таким чином, у якості прототипу було використано оптичний сканер завдяки його розповсюдженості та відносно невисокої вартості, а саме FPM10A, зображений на рис. 1.

Рис. 1. Сенсор сканування відбитків FPM10A

Технічні рішення

Оскільки доступ до серверу відбувається через браузер, то використовувати COM порт проблематично. Окрім того, є труднощі із драйверами. Тобто, потрібно мати пристрій, сумісний з стандартним набором драйверів. Саме таким апаратним рішенням є Arduino Pro Micro рис. 2, котра може імітувати мишку, клавіатуру, або джойстик.

Рис. 2. Arduino Pro Micro

Для збільшення об’єму пам’яті тут можлива комбінація з Micro sd, а для передачі даних через мережу – ESP8266. Схема підключення елементів показана на рис. 3.

Для взаємодії зі сканером задіяна стандартна бібліотека Arduino Fingerprint. Записати до пристрою можливо 255 відбитків, а перевірку правильності буде виконувати саме Arduino Pro Micro.

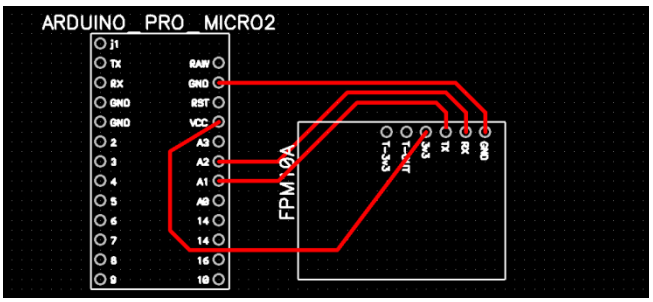
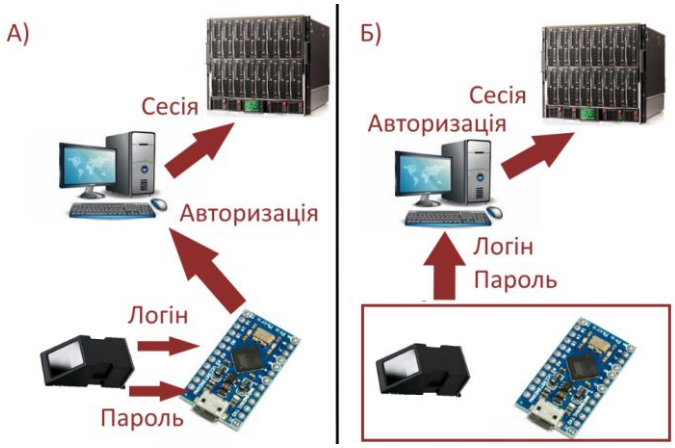


Рис. 3. Принципіальна електрична схема підключення

Можливі два варіанти взаємодії (рис. 4). Перший має на увазі перевірку та авторизацію прямо на Arduino Pro Micro, при цьому використовуються одразу два відбитки: один відповідає за логін, а інший – за пароль. У другому випадку авторизацію виконує комп'ютер за запитом програмного пристрою (рис. 4, *а*), або усі запити пристрою перевіряє комп'ютер (рис. 4, *б*). В обох варіантах потрібна спеціальна програма, котра перевірить особу у внутрішній базі та дозволить отримати доступ до веб-ресурсу. Оскільки локальні бази наявні майже у всіх закладах охорони здоров'я, то перехід до такого устаткування не викликатиме суттєвих змін, а лише дозволить забезпечити високий рівень захисту серверу, а також ідентифікувати користувачів.



а *б*

Рис. 4. Спосіб авторизації: *а* – сканер – пристрій; *б* – пристрій – комп'ютер

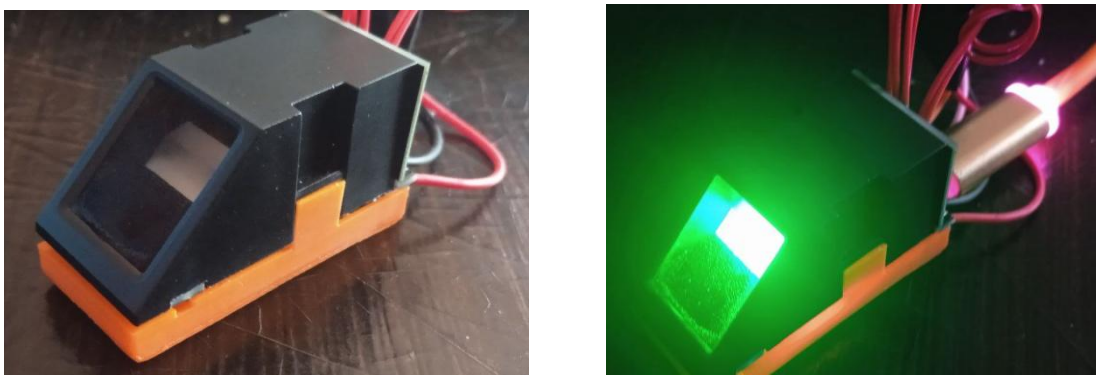
Більш надійним та зручним вважається другий спосіб, оскільки отримати доступ до серверу буде складніше за умови імітації одразу трьох складових частин.

Що стосується способів підробки та муляжів подібних пристроїв – захисний механізм використання передбачає наявність запрограмованої дати компіляції і весь програмний код на Arduino Pro Micro підпорядкований цьому формату дати та часу, що унеможливлює спосіб використання муляжу, а дата при повторній компіляції буде переналаштована, що зробить пристрій нефункціонуючим. Таким чином, пристрій важко підробити та легко налаштувати під конкретний заклад.

Перевірка технічних рішень

Для простоти збірки була розроблена та надрукована 3D модель, що покращує властивості передачі та збільшує строк роботи пристрою. Результат зображено на рис. 5. Загальний розмір прототипу склав 46x25x23 мм, що робить його аналогом кишенькового носія пам'яті, а також економним та зручним у застосуванні. Пристрій не потребує зовнішнього джерела електроенергії та підключається звичайним USB кабелем зарядки до користувацького

комп’ютеру. Для застосування потрібно лише перейти до програми та прикласти спочатку палець-логін, а після підтвердження – палець-пароль, що забезпечує максимальну надійність.



а *б*

Рис. 5. Прототип пристрою сканування: *а* – загальний вигляд пристрою; *б* – пристрій у робочому стані

Тестування прототипу проводилось на серверній програмі VinGo v2.3 (рис. 6) та VsLabLite v4.2. Авторизація пройшла успішно у 28 випадках із 30, а у категорії з помірними знаннями та навичками у комп’ютерній сфері успішних було 26 спроб з 30. При цьому не було жодної помилкової авторизації.

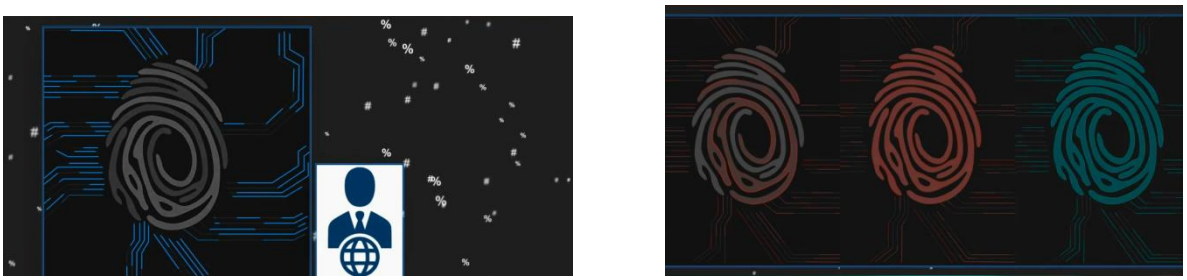


Рис. 6. Вікна тестування прототипу

Розробка функціонує строго за заданим алгоритмом та заблокує пристрій у разі декількох невдалих спроб авторизації, а спосіб авторизації дозволить отримати доступ до акаунту та сесії лише у випадку успішного зчитування обох пальців, що майже унеможливлює напад на сервери та бази даних.

Висновки

Розроблений простий та ефективний пристрій для попередньої авторизації до цільового об’єкту за відбитком. В основу входить плата, що надає стандартний набір драйверів та зумовлює простоту використання на будь-якій платформі (Windows, Linux, Mac). Розробка не потребує батареї та постійного джерела живлення, що робить її економною. Виріб є компактним та зручним у застосуванні, що для медичного персоналу є суттєвим. Пристрій має високий рівень захисту від помилкової авторизації завдяки двохетапній автентифікації, а також надійний захист від муляжів та спроб злому, що дозволяє застосовувати його у медичних та військових сферах.

Список літератури:

1. Державна служба спеціального зв’язку та захисту інформації України [Електронний ресурс] Режим доступу: https://cip.gov.ua/ua/news/kilkist-kiberatak-na-ukrayinu-prodovzhuye-zrostati

2. Л. Монастирський, В. Лозинський, Я. Бойко, Б. Соколовський. Розпізнавання відбитків пальців у недорогій біометричній системі // Електроніка та інформаційні технології. 2018. Випуск 9. С. 120 – 124.

3. Шаньгин В. Ф. Защита компьютерной информации. Эффективные методы и средства. Москва : ДМК Пресс, 2012. 544 С.
4. Патент 67772 України МПК 7G06K9/20, A61B5 / 117. Спосіб та пристрій для ідентифікації особи шляхом безконтактного розпізнавання ліній руки і пальців / Хауке Рудольф, DE, Айнігхаммер Хайнс Й., DE, Айнігхаммер Йєнс, DE., заявник і патентовласник – ТСТ-ТАЧЛЕСС СЕНСОР ТЕКНОЛОДЖИ СЕЙЛЗ ЕНД МАРКЕТІНГ АГ, СН
5. Унгул В., Проценко М. Метод поліпшення якості зображення відбитків пальців за допомогою фільтра Габора // Междунар. научн. журнал // 2016. № 6, т. 2. С. 15 – 18.
6. Fingerprint Sensor FPC1011F [Електронний ресурс] // Fingerprint. Режим доступу: http://www.fingerprint.se/en/Products/All%20products%20overview.aspx?sc_lang=en.
7. Product Specifications TCS5 TouchStrip® Fingerprint Sensor (TCEEA4 (TCS4C+TCD50A)) [Електронний ресурс] // Upek. Режим доступу: http://www.upek.com/solutions/productfinder/
8. AuthenTec Fingerprint Sensors AES2660 [Електронний ресурс] // Authentec. Режим доступу: http://www.authentec.com/products-pcsandperipherals.cfm.
9. Ворона В. А., Тихонов В. А. Системы контроля и управления доступом. Москва : Горячая линия Телеком, 2010. 272 с.
10. NI Vision Assistant Tutorial [Електронний ресурс]. Режим доступу: http://www.ni.com/pdf/manuals/372228a.pdf
11. International biometrics and idenity association [Електронний ресурс]. Режим доступу: www.ibia.org.
12. Biolink біометричні системи [Електронний ресурс].
13. Biometric terminals add security to a variety of processes [Електронний ресурс]. Режим доступу: www.bioscrypt.com
14. From identity and secure access to biometric identity [Електронний ресурс]. Режим доступу: www.crossmatch.com
15. Everywhere Identity Matters [Електронний ресурс]. Режим доступу: www.identix.com

Надійшла до редколегії 20.10.2022

Відомості про авторів:

Довнар Олександр Йосипович – канд. техн. наук, доцент, Національний аерокосмічний університет ім. М.Є Жуковського «Харківський авіаційний інститут», доцент кафедри Радіоелектронних та медичних комп'ютеризованих засобів та технологій, Україна; email: a.dovnar@khai.edu; ORCID: https://orcid.org/0000-0001-7171-0024

Бабаков Михайло Федорович – канд. техн. наук, доцент, Національний аерокосмічний університет ім. М.Є Жуковського «Харківський авіаційний інститут», професор кафедри Радіоелектронних та медичних комп'ютеризованих засобів та технологій, Україна; email: m.babakov@khai.edu; ORCID: https://orcid.org/0000-0001-8642-3693

Черкіс Владислав Ігорович – студент, Національний аерокосмічний університет ім. М.Є Жуковського «Харківський авіаційний інститут», Україна, email: v.i.cherkis@student.khai.edu

УДК 621.396.967.2 **DOI:10.30837/rt.2022.4.211.13**

І.В. СВИД, канд. техн. наук, А.О. СЄРІКОВ, І.І. ОБОД, д-р техн. наук

ЗАСТОСУВАННЯ MATLAB ДЛЯ МОДЕЛЮВАННЯ РАДІОЛОКАЦІЙНИХ СИСТЕМ

Вступ

Радіолокаційні системи (РЛС) займають важливе місце у багатьох сферах людської діяльності. РЛС застосовуються для вирішення складних та важливих задач суспільства: організація та забезпечення зв'язку, контроль та управляння повітряним рухом, геодезія, картографія тощо. Зокрема, моделювання: роботи системи у цілому [1 – 5], ймовірнісних характеристик систем пізнавання державної приналежності [6 – 8], ситуацій у каналах приймання та передавання польотної інформації [9 – 11], алгоритмів обробки радіолокаційної інформації [12 – 14], алгоритмів щодо оптимізації параметрів у радіолокаційних системах [15, 16], тощо.

Наявні РЛС постійно удосконалюються та модернізуються, також сучасна індустрія спрямована на проєктування і розробку новітніх РЛС. Для якісного проєктування та розробки РЛС вкрай важливі програмні засоби для моделювання та дослідження проєктованих систем.

MATLAB це пакет прикладних програм (ППП) для числового аналізу, що також включає мову програмування. MATLAB, створено компанією The MathWorks та є зручним засобом для роботи з математичними матрицями, малювання функцій, роботи з алгоритмами, створення робочих оболонок (user interfaces) з програмами іншими мовами програмування. Серед багатьох програмних засобів для числового аналізу MATLAB є найбільш комерційно успішним в світі [17 – 21].

MATLAB має доволі широкий спектр реалізованих прикладних застосувань для різних галузей: автоматизовані системи водіння; обчислювальна біологія; системи управління; Data Science; глибоке навчання; електрифікація; вбудовані системи; підприємство та IT-системи; розробка FPGA, ASIC і SoC; обробка зображень і комп'ютерний зір; інтернет речі; машинне навчання; мехатроніка; системи змішаних сигналів; прогнозне технічне обслуговування; робототехніка; обробка сигналів; випробування та вимірювання; бездротовий зв'язок [21 – 25].

Основні можливості MATLAB: аналіз даних; графіка; програмування; створення додатків; зовнішні мовні інтерфейси (Python, C/C++, Fortran, Java тощо); спряження з обладнанням; паралельні обчислення; робота у веб-версії продукту; версія продукту, що інсталюється; робота в хмарних середовищах від MathWorks Cloud до публічних хмар, включаючи AWS і Azure.

Можна вважати, що на сьогодні MATLAB є потужним програмним засобом для проєктування та дослідження радіолокаційних систем, що може забезпечити широкий спектр можливостей для моделювання і дослідження.

Метою запропонованої роботи є аналіз особливостей застосування MATLAB для моделювання радіолокаційних систем.

Приклади моделювання РЛС у MATLAB

Для проведення якісного моделювання та проєктування РЛС у MATLAB необхідно: мати базові та поглиблені знання, у відповідності до наявних задач щодо побудови і роботи радарів; розбиратися в особливостях параметрів радарів, які залежать від технічного завдання та вимог до поточного проєкту; вміти аналізувати запропоновані рішення та пропонувати варіанти щодо їх оптимізації.

За допомогою функцій прикладного пакету програм MATLAB можна змоделювати для радіолокаційних систем, наприклад: стиснення/розширення імпульсу, обробка імпульсів, узгоджений фільтр, обчислення ймовірності виявлення за всіма моделями Сверлінга, високу роздільну здатність системи, ступінчастий аналіз частотної форми сигналу, фільтр відстеження, фільтр Калмана, фазовані антенні решітки, обчислення завад, функції радіолокаційної

неоднозначності, навмисні/ненавмисні завади, корельовані/некорельовані завади та багато іншого.

Далі розглянемо деякі приклади, що дозволять продемонструвати наведені прикладні задачі для моделювання радіолокаційних систем.

Для прикладу розглянемо рівняння радіолокаційних втрат (або співвідношення сигнал/завада на виході приймача):

$$\left(SNR\right)_O = \frac{P_t G^2 \lambda^2 \sigma}{\left(4\pi\right)^3 kT_e BFLR^4}, \tag{1}$$

де P_t – пікова потужність, що передається; G – підсилення антени; λ – довжина хвилі; σ – радіолокаційний поперечний переріз; $k = 1{,}38 \cdot 10^{-23}$ джоуль / градус Кельвіна – постійна Больцмана; T_e – ефективна шумова температура в градусах Кельвіна; B – робоча смуга радіолокатора; F – коефіцієнт шуму; L – радіолокаційні втрати; R – відстань від радіолокатора.

Наведемо та розглянемо один з варіантів врахування в MATLAB параметрів рівняння (1) для моделювання:

$$\text{snr = radar (pt, freq, g, sigma, te, b, nf, loss, range)}, \tag{2}$$

де pt – пікова потужність, Вт; freq – центральна частота радара, Гц; g – коефіцієнт підсилення антени, дБ; sigma – радіолокаційний поперечний переріз, дБм^2; te – ефективна шумова температура, К; b – робоча смуга пропускання, Гц; nf – коефіцієнт шуму, дБ; loss – радіолокаційні втрати, дБ; range – відстань від радіолокатора (може бути або одним значенням, або вектором), м; snr – SNR (одне значення або вектор, залежно від вхідного діапазону), дБ.

Формула (2) може приймати одне значення для вхідного діапазону або вектор, що містить багато значень діапазону.

З використанням форми запису (2) отримано графіки, що представлено на рис. 1. Особливістю моделювання при використанні запису (2) – є дублювання запису із внесенням змін у параметри.

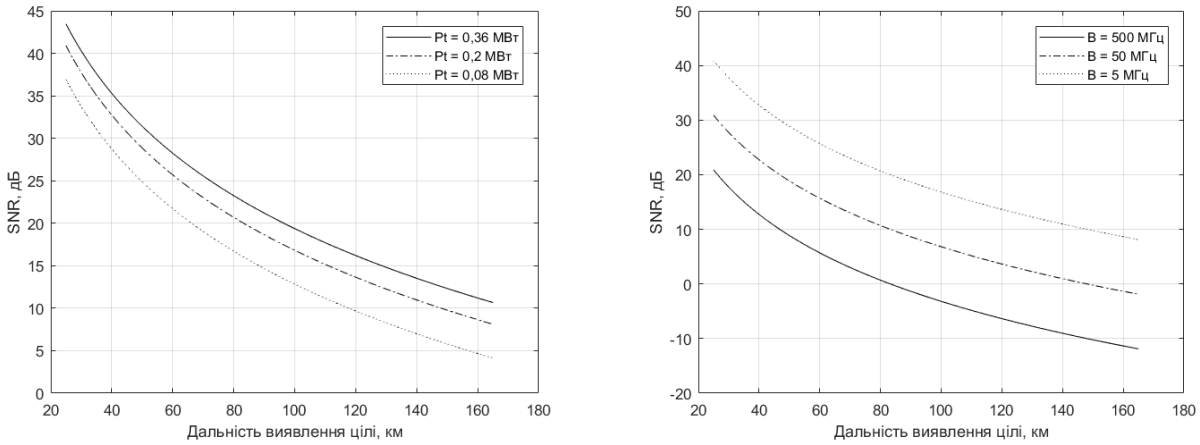


Рис. 1. Приклад побудови графіків у MATLAB за виразом (2)

У наступному прикладі розглянемо один із варіантів моделювання ймовірності виявлення в порівнянні з одним імпульсом SNR для кількох значень ймовірності помилкової тривоги. Моделювання проведемо у припущенні, що: радіолокаційний сигнал є синусоїдальним сигналом; імовірність помилкової тривоги є малою відносно імовірності виявлення. Тоді можна скористатися наближеним виразом [26]:

$$P_D \approx 0,5 \times erfc\left(\sqrt{-\ln P_{fa}} - \sqrt{SNR + 0,5}\right), \quad (3)$$

де P_{fa} – ймовірність помилкової тривоги; SNR – одно імпульсне співвідношення сигнал/завада; $erfc(q)$ – додаткова функція помилки, яка визначається як

$$erfc(q) = 1 - \frac{2}{\sqrt{\pi}} \int_0^q e^{-v^2} dv. \quad (4)$$

На рис. 2 показано графіки ймовірності виявлення від одиничного імпульсу SNR для кількох значень ймовірності помилкової тривоги, при використанні для моделювання виразів (3) і (4).

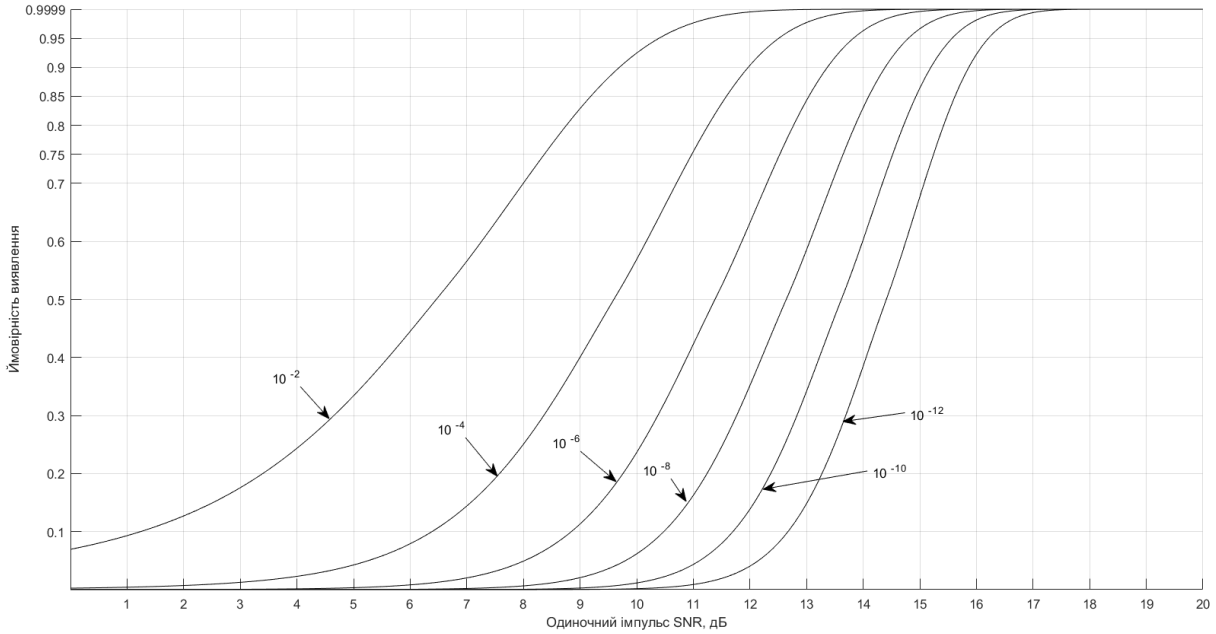


Рис. 2. Приклад побудови графіків у MATLAB за виразами (3) і (4)

До особливостей моделювання можна віднести використання спрощеного інженерного представлення ймовірності виявлення, яке реалізується шляхом використання наступних записів у MATLAB:

Лістінг

```
b = sqrt(-2.0 * log(10^(-nfa)));
a = sqrt(2.0 * 10^(.1*snr));
pro(index) = marcumsq(a,b);
```

Функція MATLAB marcumsq – це узагальнена Q-функція Маркума. Q = marcumq(a,b) обчислює Q-функцію Маркума першого порядку для параметра нецентральності *a* та аргументу *b*. Цей синтаксис можна використовувати для обчислення особливого випадку узагальненої Q-функції Маркума порядку m з m = 1.

Вище наведена невелика кількість варіантів можливого використання пакета прикладних програм MATLAB для дослідження та моделювання радіолокаційних систем.

Radar Toolbox – засіб моделювання і дослідження РЛС

Radar Toolbox містить алгоритми та інструменти для проєктування, моделювання, аналізу та тестування багатофункціональних радіолокаційних систем. Довідкові приклади є відп-

равною точкою для впровадження бортових, наземних, корабельних і автомобільних радіолокаційних систем. Radar Toolbox підтримує кілька робочих процесів, включаючи аналіз вимог, проєктування, розгортання та аналіз польових даних. Radar Toolbox включає: радіолокаційні додатки; інженерію радіолокаційних систем; синтез радіолокаційних даних; радіолокаційні сигнали і обробку даних [27].

За допомогою програми Radar Designer можна в інтерактивному режимі виконувати аналіз бюджету зв'язку та оцінювати компроміси проєктування на рівні радара. Набір інструментів містить моделі для передавачів, приймачів, каналів розповсюдження, цілей, засобів завад і самих завад. Можна моделювати радари на різних рівнях абстракції, використовуючи імовірнісні моделі та моделі рівня сигналу. Можна обробляти виявлення, створені на основі цих моделей або даних, зібраних із радіолокаційних систем, використовуючи алгоритми обробки сигналів і даних, надані в наборі інструментів. Можна розробити когнітивні радари, які працюють у переповнених середовищах із спільним радіочастотним спектром.

Висновки

Проведений аналіз показує, що пакет прикладних програм MATLAB є потужним засобом для моделювання, дослідження та проєктування радіолокаційних систем різного призначення. Пакет Radar Toolbox є готовим до застосування віконним механізмом, який забезпечує проєктування, моделювання та тестування багатофункціональних радіолокаційних систем. Radar Toolbox забезпечує швидке моделювання, модернізацію та прототипування стандартних та модернізованих елементів радіолокаційних систем. Застосування MATLAB для моделювання радіолокаційних систем неможливе без чіткого розуміння фундаментальних основ побудови та функціонування радіолокаційних систем. Також необхідно знати та вміти застосовувати функції MATLAB для опису, представлення та моделювання структурних елементів і процесів у радіолокаційних системах.

Наведено основні функціональні можливості та можливі варіанти застосування моделей у MATLAB для моделювання та дослідження радіолокаційних систем. Наведений перелік варіантів моделей не є вичерпним і остаточним. Цей напрямок має багато можливостей для досліджень і моделювання.

Список літератури:

1. Свид І., Обод І. Завадостійкість радіолокаційних систем ідентифікації за ознакою «свій-чужий». Харків : Друкарня Мадрид, 2021. 253 с.
2. Обод І., Свид І., Мальцев О. Обробка даних радіолокаційних систем спостереження повітряного простору : навч. посібник. Харків : Друкарня Мадрид, 2021. 255 с.
3. Li J., Stoica P.. 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. Обод І.І., Свид І.В., Штих І.А. Завадозахищеність запитальних систем спостереження повітряного простору : монографія ; за заг. ред. І.І. Обода. Харків : ХНУРЕ, 2014. 312 с.
6. Svyd I., Obod I. and Maltsev O. Interference Immunity Assessment Identification Friend or Foe Systems // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 287 – 306, 2021. doi: 10.1007/978-3-030-71892-3_12.
7. Semenets V., Svyd I., I Obod I., Maltsev O. and Tkach M. Quality Assessment of Measuring the Coordinates of Airborne Objects with a Secondary Surveillance Radar // Ageyev D., Radivilova T., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 69. Springer, Cham, pp. 105 – 125, 2021. doi: 10.1007/978-3-030-71892-3_5.
8. 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.
9. Obod I., Svyd I., Maltsev O., Vorgul O., Maistrenko G. and Zavolodko G. Optimization of the Quality of Information Support for Consumers of Cooperative Surveillance Systems // Radivilova T., Ageyev D., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 48. Springer, Cham, pp. 133 – 155, 2021. doi: 10.1007/978-3-030-43070-2_8.

10. Obod I., Svyd I., Maltsev O., Zavolodko G., Pavlova D. and Maistrenko G. Fusion the Coordinate Data of Airborne Objects in the Networks of Surveillance Radar Observation Systems // Radivilova T., Ageyev D., Kryvinska N. (eds) Data-Centric Business and Applications. Lecture Notes on Data Engineering and Communications Technologies, vol 48. Springer, Cham, pp. 731 – 746, 2021. doi: 10.1007/978-3-030-43070-2_31.

11. 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.

12. You H., Jianjuan X., Xin G. Radar Data Processing with Applications. Publishing House of Electronics Industry, 2016. doi: 10.1002/9781118956878.

13. Neindre F. L., Ferre G., Dallet D., Letellier F. and Pitois K. A Successive Interference Cancellation-based Receiver for Secondary Surveillance Radar // IEEE Transactions on Aerospace and Electronic Systems, 2022, doi: 10.1109/TAES.2022.3193649.

14. Leonardi M. and Fausto D. D. Secondary Surveillance Radar Transponders classification by RF fingerprinting // 2018 19th International Radar Symposium (IRS), 2018, pp. 1 – 10, doi: 10.23919/IRS.2018.8448244.

15. Xu J., Dai X.-Z., Xia X.-G., Wang L.-B., Yu J. and Peng Y.-N. 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.

16. Barbary M., Hafez A. S. and Crew T. An Industrial Design and Implementation Approach of Secondary Surveillance Radar System // International Telecommunications Conference (ITC-Egypt), 2021, pp. 1 – 9, doi: 10.1109/ITC-Egypt52936.2021.9513961.

17. Забара С., Гагарін А., Кузьменко І., Щербашин Ю. Моделювання систем у середовищі MATLAB. Київ : Вид-во “Університету “Україна”, 2011. 136 с.

18. Гоблик Н.М., Гоблик В.В. MATLAB в інженерних розрахунках. Комп'ютерний практикум : навч. посібник. Львів : Нац. ун-т "Львів. політехніка", 2020. 190 с.

19. Гаєв Є.О., Нестеренко Б.М. Універсальний математичний пакет MATLAB і типові задачі обчислювальної математики : навч. посібник. Київ : НАУ, 2004. 176 с.

20. MATLAB Book 2017a. Programming Fundamentals © Copyright 1984-2017 by The MathWorks, Inc.

21. Trauth M. H. and Sillmann E. Collecting, processing and presenting Geoscientific Information: MATLAB and design recipes for earth sciences. Berlin : Springer, 2018.

22. Gonzalez R. C., Woods R. E., Eddins S.L. Digital Image Processing Using MATLAB. 4th ed. Gatesmark Publishing. 2020. ISBN 9780982085417.

23. Mahafza, Bassem R. MATLAB simulations for radar systems design / Bassem R. Mahafza, Atef Z. Elsherbeni. Chapman & Hall/CRC CRC Press LLC. Boca Raton London New York Washington, D.C. 706 p.

24. Paluszek M., Thomas S. J. and Ham E. Practical matlab deep learning: A projects-based approach. New York : Apress, 2022. doi: https://doi.org/10.1007/978-1-4842-7912-0.

25. Рибальченко М.О., Єгоров О.П., Зворикін В.Б. Цифрова обробка сигналів : навч. посібник. Дніпро : НМетАУ, 2018. 79 с.

26. North, D. O. An Analysis of the Factors which Determine Signal/Noise Discrimination in Pulsed Carrier Systems // Proc. IEEE 51, No. 7, July 1963, pp. 1015 – 1027.

27. “Radar Toolbox,” MathWorks. [Online]. Available: https://www.mathworks.com/products/radar.html. [Accessed: 03-Dec-2022].

Надійшла до редколегії 13.11.2022

Відомості про авторів:

Свид Ірина Вікторівна – канд. техн. 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

Сєріков Антон Олександрович – аспірант кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: anton.sierikov1@nure.ua; ORCID: https://orcid.org/0000-0002-3917-2008

Обод Іван Іванович – д-р техн. наук, професор, професор кафедри мікропроцесорних технологій і систем, Харківський національний університет радіоелектроніки, Україна; email: ivan.obod@nure.ua; ORCID: https://orcid.org/0000-0002-9898-0937.

ABSTRACTS РЕФЕРАТИ

SYSTEMS AND METHODS OF INFORMATION PROTECTION
СИСТЕМИ І МЕТОДИ ЗАХИСТУ ІНФОРМАЦІЇ

UDC 004.056.5

Classification and analysis of vulnerabilities of modern information systems from classical and quantum attacks / *Ye.V. Ostrianska, S.O. Kandiy, I.D. Gorbenko, M.V. Yesina* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 7 – 21.

Recent advances in quantum technology and the potential that practical quantum computers may become a reality in the future have led to renewed interest in developing cryptographic technologies that are secure against conventional and quantum attacks. Currently, virtually all asymmetric cryptographic schemes in use are threatened by the potential development of powerful quantum computers. Post-quantum cryptography is one of main the ways to combat this threat. Its security is based on the complexity of mathematical problems that are currently considered unsolvable efficiently, even with the help of quantum computers. The security of information systems is ensured through protection against various threats that use system vulnerabilities. Security protocols are the building blocks of secure communication. They implement security mechanisms to provide security services. Security protocols are considered abstract when analyzed, but may have additional vulnerabilities in implementation. This work contains a holistic study of security protocols. Basics of security protocols, taxonomy of attacks on security protocols and their implementation are considered, as well as various methods and models of protocol security analysis. In particular, the differences between information-theoretic and computational security, computational and symbolic models are specified. In addition, an overview of the computational security models for Authenticated Key Exchange (AKE) and Password Authentication Key Exchange (PAKE) protocols is provided. The most important security models for the AKE and PAKE protocols were also described. With the emergence of new technologies that may have different security requirements, as well as with increased opportunities for competition, there is always a need to develop new protocols. Thus, the purpose of this article is to review, classify, analyze, and research the vulnerabilities of information systems from classical, quantum, and special attacks, performed taking into account the forecast regarding the possibilities of attacks on post-quantum cryptographic transformations; studying security assessment models for existing cryptographic protocols, as well as reviewing and benchmarking security models and providing suggestions for protection against existing potential attacks.

Key words: post-quantum cryptography; encryption scheme; security protocol; security model; AKE protocol; PAKE protocol.

1 fig. Ref: 42 items.

УДК 004.056.5

Класифікація та аналіз вразливостей сучасних інформаційних систем від класичних та квантових атак / *Є.В. Острянська, С.О. Кандій, І.Д. Горбенко, М.В. Єсіна* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 7 – 21.

Завдяки останнім досягненням у квантових технологіях та потенціалу того, що практичні квантові комп'ютери можуть стати реальністю в майбутньому, відновився інтерес до розробки криптографічних технологій, захищених від звичайних та квантових атак. Наразі практично всім асиметричним криптографічним схемам, які зараз використовуються, загрожує потенційна розробка потужних квантових комп'ютерів. Постквантова криптографія є одним із способів боротьби з цією загрозою. Її безпека базується на складності математичних проблем, які вважаються нерозв'язними ефективно – навіть за допомогою квантових комп'ютерів. Безпека інформаційних систем досягається через захист від різноманітних загроз, що використовують вразливості системи. Протоколи безпеки є будівельними блоками безпечного зв'язку. Вони реалізують механізми безпеки для надання послуг безпеки. Протоколи безпеки вважаються абстрактними під час аналізу, але вони можуть мати додаткові вразливості у реалізації. Ця стаття містить цілісне дослідження протоколів безпеки. Розглядаються основи протоколів безпеки, таксономія атак на протоколи безпеки та їх впровадження, а також різні методи та моделі аналізу безпеки протоколів. Зокрема уточнюються відмінності між інформаційно-теоретичною та обчислювальною безпекою, обчислювальними та символьними моделями. Крім того, надано огляд моделей обчислювальної безпеки для автентифікованого обміну ключами (AKE) і протоколів обміну ключами з автентифікацією пароля (PAKE). Також було описано найважливіші моделі безпеки для протоколів AKE і PAKE. З появою нових технологій, які можуть мати інші вимоги до безпеки, а також завдяки збільшеним можливостям змагальності, завжди виникає потреба в розробці нових протоколів. Таким чином, метою статті є огляд, класифікація, аналіз та дослідження вразливостей інформаційних систем від класичних, квантових та спеціальних атак, виконані з урахуванням прогнозу щодо можливостей здійснення атак на постквантові криптографічні перетворення; вивчення моделей для оцінки безпеки для існуючих криптографічних протоколів, а також огляд та порівняльний аналіз моделей безпеки та надання пропозицій щодо захисту від існуючих потенційних атак.

Ключові слова: постквантова криптографія; схема шифрування; протокол безпеки; модель безпеки; AKE протокол; PAKE протокол.

Іл. 1. Бібліогр.: 42 назв.

UDC 004.056.5

Analysis of DSTU 8961:2019 in random oracle model / *S.O. Kandiy* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 22 – 36.

The paper provides a proof in the IND-CCA2 random oracle model of the security of the asymmetric encryption scheme described in the DSTU 8961:2019 standard, and the IND-CCA2 security of the corresponding key encapsulation mechanism. Since the standard contains only a technical description of transformations, a formalized mathematical model was introduced in Chapter 4 without unnecessary technical details that do not affect safety assessments. Since the system-wide parameters in the standard were chosen in such a way that the scheme did not contain decryption errors, it was possible to simplify significantly the proof. Section 5 provides a schematic overview of possible attack vectors on the DSTU 8961:2019, but a detailed analysis is the subject of further research. In addition to safety, the analysis also showed that the DSTU 8961:2019 has a certain disadvantage in terms of safety. The design can be significantly simplified and accelerated without loss of safety. Security, on the contrary, can be significantly increased.

Key words: post-quantum cryptography; algebraic lattices; DSTU 8961:2019 "Skelya"; Random Oracle Model.

3 tab. 7 fig. Ref: 23 items.

УДК 004.056.5

Аналіз безпеки ДСТУ 8961:2019 у моделі випадкового оракула / *С.О. Кандій* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 22 – 36.

Наведено доказ в моделі випадкового оракула IND-CCA2 безпеки схеми асиметричного шифрування, що описана в стандарті ДСТУ 8961:2019, та IND-CCA2 безпека відповідного механізму інкапсуляції ключів. Оскільки стандарт містить тільки технічний опис перетворень, у розд. 4 введена формалізована математична модель без зайвих технічних деталей, що не впливають на оцінки безпеки. Оскільки загальносистемні параметри в стандарті обрані так, щоб схема не містила помилок дешифрування, вдалося значно спростити доказ. У розд. 5 наведено схематичний огляд можливих векторів атак на ДСТУ 8961:2019, проте детальний аналіз є предметом подальших досліджень. Показано, що ДСТУ 8961:2019 має певну збитковість у сенсі безпеки. Конструкція може бути значно спрощена та пришвидшена без втрати безпеки. Безпеку, навпаки, можливо значно підвищити.

Ключові слова: постквантова криптографія; алгебраїчні решітки; ДСТУ 8961:2019 “Скеля”; модель випадкового оракула.

Табл. 3. Іл. 7. Бібліогр.: 23 назв.

UDC 004.065

The main categories of NewSQL databases and their features / *V.I. Yesin, V.V. Vilihura* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 37 – 66.

In the modern world, the problem of working with big data and workloads is becoming more and more acute. For more than forty years, relational databases have been the main leading systems for storing, searching and managing data. However, despite their great popularity, application experience and universality, traditional relational DBMS, due to the growing needs for scalability and performance, often cannot meet modern requirements. This has led to the emergence of new alternative data management systems, including NewSQL systems. NewSQL is a class of modern relational database management systems that provide performance comparable to NoSQL systems while maintaining the data consistency guarantees inherent in traditional database systems. The growing interest in NewSQL technology in recent times has led to an increase number of evaluations and comparisons among competing NewSQL technologies. However, today there is still a certain lack of work devoted to the study of the features of NewSQL solutions and their capabilities in comparison with other technologies. This paper discusses the main features of the most famous NewSQL products of different categories and the identified problems associated with them. To overcome the certain ambiguity in the names and translations of some terms related to the subject area under consideration, which takes place in numerous relevant sources, additional explanations are given. For comparison, the paper presents the values of important characteristics inherent in NewSQL, traditional relational and NoSQL database systems. This paper can help researchers and people from the industry choose the best storage solutions for their needs.

Key words: database; relational database; database management system; NoSQL; NewSQL.

2 tab. 4 fig. Ref: 87 items.

УДК 004.065

Основні категорії NewSQL баз даних та їх особливості / *В.І. Єсін, В.В. Вілігура* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 37 – 66.

У сучасному світі все гостріше постає проблема роботи з великими обсягами даних та навантаженнями. Понад сорок років основними провідними системами, що забезпечують зберігання, пошук та управління даними, були реляційні бази даних. Однак, незважаючи на велику популярність, досвід застосування та універсальність, традиційні реляційні СКБД через зростаючі потреби в масштабованості та продуктивності часто не можуть задовольнити сучасним вимогам. Це призвело до появи нових альтернативних систем керування даними, зокрема NewSQL систем. NewSQL – це клас сучасних систем керування реляційними базами даних, які забезпечують продуктивність, яку можна порівняти з NoSQL системами, не послаблюючи у своїй гарантій узгодженості даних, властивих традиційним системам баз даних. Зростаючий інтерес до технології NewSQL призвів до збільшення кількості оцінок та порівнянь між конкуруючими технологіями NewSQL. Однак сьогодні все ще

існує певний дефіцит робіт, присвячених вивченню особливостей NewSQL, рішень та їх можливості щодо інших технологій. У роботі розглядаються основні особливості найбільш відомих продуктів NewSQL різних категорій та виявлені проблеми, пов'язані з ними. Для подолання наявної в численних релевантних джерелах певної неоднозначності в назвах і перекладах деяких термінів, пов'язаних з тематикою, що розглядається, надаються додаткові пояснення. Для порівняння наводяться значення важливих характеристик, властивих NewSQL, традиційним реляційним і NoSQL системам баз даних. Робота може допомогти дослідникам та фахівцям-практикам у виборі кращих рішень для зберігання даних відповідно до їхніх потреб.

Ключові слова: база даних; реляційна база даних; система керування базами даних; NoSQL; NewSQL.

Табл. 2. Іл. 4. Бібліогр.: 87 назв.

UDC 004.056.55

Analysis of the Falcon signature compared to other signatures. GPV and Rabin frameworks / *D.V. Harmash* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 67 – 71.

The article discusses the analysis of the essence and protection possibilities of the Falcon post-quantum signature. The main properties of the Falcon signature are considered. An estimate of what resources and computing power is required to use successfully the Falcon signature. A structural analysis of the Falcon signature is performed. The GPV and Rabina frameworks are analyzed. Detailed conclusions are made regarding the conducted analyses. The stability and complexity of the GPV and Rabin frameworks are evaluated, the main structures and protocols of these frameworks are considered. A detailed analysis of the main properties of NTRU lattices is carried out, the main rules of factorization of the GPV and Rabin frameworks are considered. Fast Fourier sampling is investigated. Conclusions are made regarding each conducted study.

Key words: Falcon; cryptanalysis; vulnerability; scheme; algorithm.

Ref: 8 items.

УДК 004.056.55

Аналіз підпису FALCON в порівнянні з іншими підписами. Фреймворки GPV та Рабіна / *Д.В. Гармаш* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 67 – 71.

Розглядається аналіз сутності та можливостей захисту постквантової сигнатури Falcon. Розглянуто основні властивості сигнатури Falcon. Оцінка того, які ресурси та обчислювальна енергія потрібні для успішного використання підпису Falcon. Виконано структурний аналіз сигнатури Falcon. Аналізуються фреймворки GPV і Рабіна. Робляться детальні висновки щодо проведених аналізів. Дано оцінку стійкості та складності фреймоврркам GPV та Рабіна, розглянуто основні структури та протоколи цих фреймворків. Проведено детальний аналіз основних властивостей NTRU решіток, розглянуто основні правила факторизації фреймворків GPV та Рабіна. Досліджується швидка вибірка Фур'є. Робляться висновки стосовно кожного проведеного дослідження.

Ключові слова: Falcon; криптоаналіз; вразливість; схема; алгоритм.

Бібліогр.: 8 назв.

RADIO PHYSICS
РАДІОФІЗИКА

UDC 621.372(075.8)

Multifractal analysis of model fractal and multifractal signals / *O.V. Lazorenko, A.A. Onishchenko, L.F. Chernogor* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 72 – 83.

One of the topical directions of modern fractal radio physics is the multifractal analysis of signals and processes of various origins. A set of deterministic and stochastic models of monofractal and multifractal signals and processes in the time domain is proposed. New multifractal signal characteristics, namely, the coefficient of asymmetry of the multifractal spectrum function, the relative width of the multifractal spectrum and the dimension of the multifractal support, are introduced, the necessity of their use is demonstrated on examples. Using Wavelet Transform Modulus Maxima Method and Multi-Fractal Detrended Fluctuation Analysis Method, a detailed multifractal analysis of model signals is performed. The features of multifractal analysis of monofractal, multifractal and non-fractal signals are established, the appropriate recommendations for practitioners are formulated. Convenient formats for presenting analysis results have been developed. It was found that during the transition of the multifractal signal to the monofractal regime, the function of the multifractal spectrum of the physical fractal does not collapse into a point, as it should happen in theory for a mathematical fractal. Threshold values of multifractal characteristics, which are indicators of the appearance of the monofractal, are proposed. It has been shown that multifractal analysis of non-fractal signals leads to the appearance of multifractal spectra with anomalous values of multifractal characteristics. The correction function method is modified for the methods of multifractal analysis of signals and processes. It is proved that its usage makes it possible to reduce the deviation of the obtained estimate of the generalized Hurst exponent from the true known value of the Hölder exponent of the analyzed signal from 5 – 90% to 3 – 8%.

Key words: fractal; multifractal; signal; process; analysis; method; dimension; estimation; accuracy; correction.

3 fig. Ref: 19 items.

УДК 621.372(075.8)

Мультифрактальний аналіз модельних фрактальних і мультифрактальних сигналів / *О.В. Лазоренко, А.А. Онищенко, Л.Ф. Чорногор* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 72 – 83.

Одним із актуальних напрямків сучасної фрактальної радіофізики є мультифрактальний аналіз сигналів і процесів різного походження. Запропоновано набір детермінованих і стохастичних моделей монофрактальних і мультифрактальних сигналів і процесів у часовій області. Введено нові мультифрактальні характеристики сигналів, а саме – коефіцієнт асиметрії функції мультифрактального спектру, показник відносної ширини мультифрактального спектру та розмірність носія мультифракталу, на прикладах продемонстровано необхідність їх використання. Із використанням методів Wavelet Transform Modulus Maxima та Multi-Fractal Detrended Fluctuation Analysis проведено докладний мультифрактальний аналіз модельних сигналів. Встановлено особливості мультифрактального аналізу монофрактальних, мультифрактальних і нефрактальних сигналів, сформульовано відповідні рекомендації для практиків. Розроблено зручні формати представлення результатів аналізу. Встановлено, що під час переходу мультифрактального сигналу до монофрактального режиму функція мультифрактального спектру фізичного фрактала не колапсує у точку, як це має відбуватися у теорії для математичного фрактала. Запропоновано порогові значення мультифрактальних характеристик, що є індикаторами появи монофрактального режиму. Продемонстровано, що мультифрактальний аналіз нефрактальних сигналів призводить до появи мультифрактальних спектрів із аномальними значеннями мультифрактальних характеристик. Метод коригуючої функції модифіковано для методів мультифрактального аналізу сигналів і процесів. Доведено, що його застосування дозволило знизити відхилення отримуваної оцінки узагальненого показника Херста від істинної відомої величини показника Гьольдера аналізованого сигналу з 5 – 90 % до 3 – 8 %.

Ключові слова: фрактал; мультифрактал; сигнал; процес; аналіз; метод; розмірність; оцінювання; точність; коригування.

Іл. 3. Бібліогр.: 19 назв.

RADIO LOCATION AND RADIO NAVIGATION
РАДІОЛОКАЦІЯ І РАДІОНАВІГАЦІЯ

UDC 621.396.96, 621.397.48:004.932.2

Features of the tasks of identifying and observing groups of unmanned letter vehicles / *V.M. Kartashov, V.A. Pososhenko, A.I. Kapusta, M.V. Rybnykov, E.V. Pershin* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 84 – 92.

The current trend towards increasing the efficiency of UAV use is the transition from their single use to group use. In accordance with this, when building a complex integrated system for detecting and monitoring UAVs, including various information channels, it is advisable to take into account the features of the system functioning associated with the detection of groups of UAVs.

The article discusses the information, energy and search capabilities of individual detection tools being a part of the integrated UAV surveillance system in order to build an effective algorithm for joint processing of incoming input signals, taking into account various capabilities of individual channels (in terms of range, recognition, etc.).

An optimal algorithm for detecting groups of UAVs in a complex integrated system combining detection decisions made in private channels is synthesized. According to the synthesized algorithm, complex processing consists in summing up the solutions of individual detectors with some weights determined by the quality of the decisions made in the channels. The quality of solutions, in turn, depends on the technical means used in the channels and the conditions of observation.

A sequence for solving a set of interrelated tasks in a complex integrated UAV surveillance system as the group target approaches the protected object is proposed. The sequence includes the following operations: identification of a group target (energy detection); estimation of coordinates of a group of objects; spatial resolution and determination of the number of vehicles in a group; recognition (type determination) of each individual device; assessment of the coordinates of each aircraft separately; determination of the composition of the group (homogeneous, heterogeneous); determination of the specialization of the group and revealing the nature of its tasks.

Key words: unmanned aerial vehicle; detection; observation; resolution; coordinate estimation; recognition; algorithm; aggregation.

1 tabl. 2 fig. Ref: 45 items.

УДК 621.396.96, 621.397.48:004.932.2

Особливості задач виявлення і спостереження груп безпілотних літальних апаратів / *В.М. Карташов, В.О. Посошенко, А.І. Капуста, М.В. Рибников, Є.В. Першин* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 84 – 92.

Сучасна тенденція підвищення ефективності використання БПЛА полягає у переході від їхнього одиночного застосування до групового. Відповідно до цього при побудові комплексної інтегрованої системи виявлення та спостереження за БПЛА, що включає різні інформаційні канали, доцільно враховувати особливості функціонування системи, пов'язані з виявленням груп БПЛА.

Розглянуто інформаційні, енергетичні та пошукові можливості окремих засобів виявлення, що входять до складу інтегрованої системи спостереження БПЛА, з метою побудови ефективного алгоритму спільної обро-

бки вхідних сигналів, що надходять, з урахуванням різних можливостей окремих каналів (за дальністю, розпізнаванням тощо).

Синтезовано оптимальний алгоритм виявлення груп БПЛА у комплексній інтегрованій системі, що поєднує рішення про виявлення, винесені у приватних каналах. Відповідно до синтезованого алгоритму комплексна обробка полягає у підсумовуванні рішень окремих виявлювачів з деякими вагами, що визначаються якістю рішень, прийнятих у каналах. Якість рішень залежить від технічних засобів, що використовуються у каналах, і умов спостереження.

Запропоновано послідовність вирішення сукупності взаємопов'язаних завдань у комплексній інтегрованій системі спостереження БПЛА у міру наближення групової цілі до об'єкта, що охороняється. Послідовність включає наступні операції: виявлення групової цілі (енергетичне виявлення); оцінка координат групи об'єктів; просторове розрізнення та визначення кількості апаратів у групі; розпізнавання (визначення типу) кожного окремого апарату; оцінка координат кожного літального апарату окремо; визначення складу групи (однорідна, неоднорідна); визначення спеціалізації групи та розтин характеру її завдань.

Ключові слова: безпілотний літальний апарат; виявлення; спостереження; роздільна здатність; оцінка координат; розпізнавання; алгоритм; комплексування.

Табл. 1. Іл. 2. Бібліогр.: 45 назв.

UDC 621.396.96, 621.397.48

Algorithm for estimating the energy distribution of radar signals scattering on acoustic disturbances created by UAVs / *V.M. Kartashov, V.A. Pososhenko, K.V. Kolesnik, V.I. Kolesnik, R.I. Bobnev, A.I. Kapusta* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 93 – 100.

The task of estimating the energy distribution over the observation interval of radar signals scattered on atmospheric inhomogeneities, arising as a result of UAV operation, is considered. The solution to this problem is necessary to improve detection algorithms, to classify the detected UAVs according to additional informational features, to improve the resolution when detecting several devices located at the same range during the group application of UAVs, to clarify the time parameters of the evolution of the movement of UAVs in time and space. A similar problem arises due to the need to process useful radar signals with a low signal-to-noise ratio in order to achieve the maximum possible range of reliable UAV detection. Because of this, it becomes impossible to estimate directly the energy of useful signals by the method of comparison with reference physical quantities due to a large measurement error. Therefore, an evaluation algorithm is proposed, based on the methods of the theory of ordinal statistics, which provide, instead of comparing numerical realizations with a certain standard, to form a variational series from them under the condition of a priori knowledge of the distribution function of these realizations. At the same time, the fact is used that for certain distributions of a random variable, among which there are normal and all limited ones, the variance of the estimate in the form of a mathematical expectation of certain ordinal statistics is significantly less than the variance of a direct measurement at a low signal-to-noise ratio. In order to save time and computing resources during real-time processing of received signals, it is proposed to use pre-calculated arrays of numerical values of mathematical expectation and dispersion of ordinal statistics for various parameters of the density distribution of a random variable.

Key words: UAV detection; radar signals; assessment of energy distribution; observation interval; order statistics; central chi-square distribution; non-central chi-square distribution; variation series; evaluation algorithm; acoustic disturbances.

5 tabl. Ref: 16 items.

УДК 621.396.96, 621.397.48

Алгоритм оцінювання розподілу енергії радіолокаційних сигналів, які розсіюються на акустичних збуреннях, створених БПЛА / *В.М. Карташов, В.О. Посошенко, К.В. Колісник, В.І. Колісник, Р.І. Бобнєв, А.І. Капуста* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 93 – 100.

Розглянуто задачу оцінювання розподілу на інтервалі спостереження енергії радіолокаційних сигналів, розсіяних на атмосферних неоднорідностях, які виникають внаслідок функціонування БПЛА. Рішення цієї задачі необхідно для удосконалення алгоритмів виявлення, класифікації виявлених БПЛА за додатковими інформаційними ознаками, підвищення роздільної здатності при виявленні декількох апаратів, розташованих на одній дальності при груповому застосуванні БПЛА, з'ясування часових параметрів еволюції руху БПЛА у часі та у просторі. Подібна задача виникає через необхідність обробки корисних радіолокаційних сигналів при малому співвідношенні сигнал/шум для досягнення максимально можливої дальності впевненого виявлення БПЛА. Через це стає неможливим пряме оцінювання енергії корисних сигналів методом порівняння з еталонними фізичними величинами через велику похибку вимірювання. Тому запропоновано алгоритм оцінювання, заснований на методах теорії порядкових статистик, які передбачають замість порівняння чисельних реалізацій з певним еталоном формування з них варіаційного ряду за умови апріорного знання функції розподілу цих реалізацій. При цьому використано той факт, що для певних розподілів випадкової величини, серед яких є нормальний та всі обмежені, дисперсія оцінки у вигляді математичного очікування певної порядкової статистики суттєво менше дисперсії прямого вимірювання при малому співвідношенні сигнал/шум. Для заощадження часу та обчислювального ресурсу при обробці у реальному масштабі часу сигналів, що приймаються, запропоновано використовувати попередньо розраховані масиви чисельних значень математичного очікування та дисперсії порядкових статистик для різних параметрів щільності розподілу випадкової величини.

Ключові слова: виявлення БПЛА; радіолокаційні сигнали; оцінювання розподілу енергії; інтервал спостереження; порядкові статистики; центральний розподіл «*xi*-квадрат»; нецентральний розподіл «*xi*-квадрат»; варіаційний ряд; алгоритм оцінювання; акустичні збурення.

Табл. 5. Бібліогр.: 16 назв.

UDC 621.396.967.2

Comparative analysis of interference protection of "Friend-Foe" radar identification systems / *I.V. Svyd, M.G. Tkach, I.I. Obod* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. № 211. P. 101 – 113.

The paper examines the existing systems of radar identification based on the "friend-foe" feature from the point of view of immunity evaluation. Currently, there are two systems of the radar identification based on the characteristic of "friend-foe", namely, the "Password" and MkXIIA. They are the radar identification systems based on the "home-foreign" feature. The former of the specified systems operates in a frequency range that differs from the frequency range of secondary radar systems, while the second one operates in the frequency range of secondary radar. These systems are ones of the main information resources of the airspace control system and are built on the principles of a one-channel or two-channel information transmission system. It allows the interested party, both an unauthorized use of this information resource for remote determination of the coordinates of aerial objects, on the one hand, and twisting the information of these information resources, on the other hand, which leads to unpredictable results. The purpose of the work is to assess the immunity of existing radar identification systems based on the "friend-foe" feature. The analysis of the interference protection of existing systems for the object radar identification based on the "friend-foe" feature, built on the principles of interrogative and non-interrogative information systems presented in this work, showed that the use of rectangular radio signals with time-pulse modulation as request and response signals emitting by air objects, has low immunity and excludes the energy stealthiest of the respondents of aerial objects. And, as a result, it allows for unauthorized calculation of the coordinates of air objects by the interested party based on the emitted identification signals on the basis of "friend-foe".

Key words: system; airspace; radar identification system; surveillance system; MkXII; "Parol"; "friend-foe"; immunity; obstacle; energy stealthness.

3 fig. Ref: 40 items.

УДК 621.396.967.2

Порівняльний аналіз завадозахищеності радіолокаційних систем ідентифікації «свій-чужий» / *І.В. Свид, М.Г. Ткач, І.І. Обод* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 101 – 113.

Розглянуто існуючі системи радіолокаційної ідентифікації за ознакою «свій-чужий» з точки зору оцінки завадозахищеності. В теперішній час існують дві системи радіолокаційної ідентифікації за ознакою «свій-чужий» «Пароль» та MkXIIA. Перша з вказаних систем радіолокаційної ідентифікації за ознакою «свій-чужий» працює в частотному діапазоні, який відрізняється від частотного діапазону роботи систем вторинної радіолокації, а друга працює в частотному діапазоні вторинної радіолокації. Вказані системи є одним з головних інформаційних ресурсів системи контролю повітряного простору та побудовані на принципах одноканальної чи двоканальної системи передачі інформації. Це дозволяє зацікавленій стороні як несанкціоноване використання цього інформаційного ресурсу для дальнього визначення координат повітряних об’єктів, з одного боку, та перекручування інформації цих інформаційних ресурсів, з другого боку, що призводить до непередбачуваних результатів. Метою роботи є оцінка завадозахищеності існуючих систем радіолокаційної ідентифікації за ознакою «свій-чужий». Наведений аналіз завадозахищеності існуючих систем радіолокаційної ідентифікації об’єктів за ознакою «свій-чужий», побудованих на принципах запитальних та беззапитальних інформаційних систем, показав, що використання прямокутних радіосигналів з часово-імпульсною модуляцією у якості сигналів запиту та відповіді, які випромінюють повітряні об’єкти, має низьку завадозахищеність та виключає енергетичну прихованість відповідачів повітряних об’єктів. І, як наслідок, надає можливість здійснювати несанкціоноване обчислення координат повітряних об’єктів зацікавленою стороною на основі випромінених сигналів ідентифікації за ознакою «свій-чужий».

Ключові слова: радіолокаційна система; повітряний простір; система радіолокаційної ідентифікації; система спостереження; MkXII; «Пароль»; «свій-чужий»; завадозахищеність; завада; енергетична прихованість.

Іл. 3. Бібліогр.: 40 назв.

PHYSICS OF DEVICES, ELEMENTS AND SYSTEMS
ФІЗИКА ПРИЛАДІВ, ЕЛЕМЕНТІВ І СИСТЕМ

UDC 662.396.67: 621.314.6

Wireless power transmission technologies / *V.O Alieksieiev. D.V. Gretskih, D.S. Gavva, V.G. Lykhograi* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 114 – 132.

The article consists of three parts. The analysis of existing technologies of wireless power transfer (WPT) is carried out in the first part. It is noted that one of the factors that determines the choice of one or another WPT technology is the distance over which the power is transmitted and the type of electromagnetic (EM) energy used. The essence of WPT technologies in the near zone, Fresnel zone and Fraunhofer zone is explained. A generalized block diagram of the WPT system is presented. Areas of application and trends in the further development of the WPT technologies over short distances using induction and resonance methods, the WPT technologies over long distances, the technology of

EM energy harvesting from the surrounding space and its conversion into direct current for powering low-power devices are considered.

The achievements of the team of the antenna laboratory of the Kharkiv National University of Radio Electronics (KhNURE) in the area of WPT are presented in the second part of the article. Namely, the electrodynamics' approach is considered which is based on a single idea about the functioning of WPT systems and which include antennas and their circuits and ways of excitation with nonlinear elements. The stages of building a nonlinear mathematical model (MM) of the electrodynamics' level of the WPT system are presented, according to which the entire WPT system, which generally includes the transmitting subsystem and the receiving subsystem, is considered as a single multi-input antenna system with nonlinear characteristics. The proposed MM provides a complete representation of the WPT systems operation of a wide class and purpose, in which fundamentally different WPT technologies are used.

The third part of the article presents new results related to continued research. The analysis of the adequacy of the developed MM of WPT system is carried out. The results of simulation of WPT systems with the induction method of energy transfer (near zone) and their comparison with theoretical and experimental data of other authors showed the reliability and universality of the proposed approach and the MM of WPT system developed on its basis.

Key words: wireless power transfer; energy transmission technology; wireless energy transmission system; energy harvesting from the surrounding space; rectenna; scattering matrix; non-system interactions; internal system processes.

1 tabl. 19 fig. Ref.: 109 items.

УДК 662.396.67: 621.314.6

Технології безпровідної передачі енергії / *В.О. Алєксєєв, Д.В. Грецьких, Д.С. Гавва, В.Г. Лихограй* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 114 – 132.

Стаття складається з трьох частин. У першій частині проведено аналіз існуючих технологій безпровідної передачі енергії (БПЕ). Зазначено, що одним з чинників, що визначають вибір тієї або іншої технологій БПЕ, є відстань, на яку передається енергія, та вид використовуваної електромагнітної (ЕМ). Пояснюється суть технологій БПЕ в ближній зоні, зоні Френеля та зоні Фраунгофера. Наведено узагальнену структурну схема системи БПЕ. Розглянуто галузі застосування та тенденції подальшого розвитку технології передачі енергії на малі відстані за допомогою індукційного та резонансного методів, технології передачі енергії на великі відстані, технології збору ЕМ енергії з навколишнього простору та її перетворення у постійний струм для живлення малопотужних пристроїв.

У другій частині роботи наведено здобутки колективу лабораторії антен ХНУРЕ в галузі БПЕ. А саме, розглянуто електродинамічний підхід, в основі якого лежить єдине уявлення про функціонування систем БПЕ, до складу яких входять антени і тракти їх збудження, що містять нелінійні елементи. Наведено етапи побудови нелінійної математичної моделі електродинамічного рівня системи БПЕ, згідно з якою вся система БПЕ, що включає в загальному випадку передавальну підсистему та приймальну підсистему, розглядається як єдина багатовходова антенна система з нелінійними характеристиками. Запропонована модель дає повне уявлення про функціонування систем БПЕ широкого класу та призначення, в яких використовуються принципово різні технології БПЕ.

У третій частині наведено нові результати, які пов'язані з продовженням досліджень колективу ХНУРЕ. Проведено аналіз адекватності розробленої математичної моделі системи БПЕ. Результати моделювання систем БПЕ з індукційним способом передачі енергії (ближня зона) та порівняння їх з теоретичними та експериментальними даними інших авторів показали достовірність та універсальність запропонованого в ХНУРЕ підходу та розробленої на його основі математичної моделі системи БПЕ.

Ключові слова: безпровідна передача енергії; технологія передачі енергії; система безпровідної передачі енергії; збір енергії з навколишнього простору; ректена; матриця розсіяння; позасистемна взаємодія; внутрішньосистемні процеси.

Табл. 1. Іл. 19. Бібліогр.: 109 назв.

UDC 536.21

Structural modeling and calculation of thermal conductivity of polyimide composite materials / *V.M. Borshchov, O.M. Listratenko, M.A. Protsenko, I.T. Tymchuk, O.V. Kravchenko, O.V. Suddia, I.V. Borshchov, M.I. Slipchenko* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 133 – 142.

Issues of direct modeling effective thermal conductivity of two-component thermally conductive polyimide composite films based on polyimide thermosetting varnishes and thermally conductive powder fillers are considered.

3D-structural modeling of elementary cubic cells of polyimide composites has been performed.

Calculations of average heat fluxes and effective thermal conductivity of variants of polyimide composite films with the introduction of highly thermally conductive highly dispersed and ultradispersed powder fillers into the polyimide matrix were carried out, including those from SiO2, SiC, Al2O3, AlN, taking into account boundary and initial conditions using COMSOL MULTIPHYSICS software.

Specific recommendations are proposed for direct modeling of the thermal conductivity of environments with a complex structure and for carrying out with sufficient reliability numerical calculations of the effective thermal conductivity of polyimide composite films in order to increase their thermal conductivity from 0,12 W/(m•K) up to 1-4 W/(m•K) by changing concentration and thermal conductivity of mixtures of filler particles of micron and ultramicron sizes.

Key words: thermally conductive polyimide composite; effective thermal conductivity; structural modeling; numerical calculations.

1 tab. 5 fig. Ref: 9 items.

УДК 536.21

Структурне моделювання і розрахунок теплопровідності поліімідних композитних матеріалів / *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. 2022. Вип. 211. С. 133 – 142.

Розглянуто питання прямого моделювання ефективної теплопровідності двокомпонентних теплопровідних поліімідних композитних плівок на основі поліімідних лаків та теплопровідних порошкових діелектричних наповнювачів.

Виконано 3D-структурне моделювання елементарних кубічних комірок поліімідних композитів. Проведено розрахунки середніх теплових потоків та ефективної теплопровідності варіантів поліімідних композитних плівок при введенні в поліімідну матрицю високотеплопровідних високодисперсних та ультрадисперсних порошкових наповнювачів, у тому числі із SiO2, SiC, Al2O3, AlN з урахуванням граничних та початкових умов за допомогою програмного комплексу COMSOL MULTIPHYSICS.

Запропоновано конкретні рекомендації щодо прямого моделювання теплопровідності середовищ зі складною структурою та проведення з достатньою достовірністю чисельних розрахунків ефективної теплопровідності поліімідних композитних плівок з метою збільшення їх теплопровідності від 0,12 до 1–4 Вт/(м•К) шляхом зміни концентрації та теплопровідності сумішей частинок наповнювачів мікронних та ультрамікронних розмірів.

Ключові слова: теплопровідні поліімідні композити; ефективна теплопровідність; структурне моделювання; чисельні розрахунки.

Табл. 1. Іл. 5. Бібліогр.: 9 назв.

UDC 623.373.072.9

Theoretical investigation of injection-locked differential oscillator / *V.V. Rapin* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 143 – 147.

A preliminary analysis of published works on this topic showed that at present there is no sufficiently substantiated theory of such devices, and the approximate approaches used are rough and do not always meet the requirements of practice. The proposed transition from a differential self-oscillator to an equivalent single-circuit oscillator has not received a convincing justification.

This article presents a methodology for studying a synchronized differential oscillator using rigorous methods. A mathematical model of such oscillator is presented in the form of nonlinear differential equations obtained on the basis of Kirchhoff's laws. Their analysis made it possible to substantiate the transition to the model of a single circuit LC oscillator, equivalent to a differential one. A technique for such a transition is proposed, including the procedure for determining the nonlinear characteristics of the amplifying element of this self-oscillator, based on the nonlinear characteristics of two amplifying elements of the differential oscillator.

The mathematical model of an equivalent oscillator is represented by a non-linear differential Van der Pol equation in a dimensionless form, it is simple and accurate. This form of representation made it possible to single out a small parameter and estimate its value. In the case of small values of the small parameter, as is known, traditional methods can be used for its analysis. Thus, the task of studying the synchronization process of a differential oscillator is reduced to the study of the synchronization process of a Van der Pol oscillator. The presented results can be useful in the development of various devices based on synchronized differential oscillators.

Key words: differential oscillator; synchronization; main tone; mathematical model.

3 fig. Ref.: 14 items.

УДК 623.373.072.9

Теоретичне дослідження синхронізованого диференціального автогенератора / *В.В. Рапін* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 143 – 147.

Аналіз опублікованих робіт з даної теми показав, що в даний час відсутня досить обґрунтована теорія таких пристроїв, а наближені підходи, що використовуються, грубі і не завжди задовольняють вимогам практики. Запропонований перехід від диференціального автогенератора до еквівалентного одноконтурного автогенератора не отримав переконливого обґрунтування.

Наведено методику дослідження синхронізованого диференціального автогенератора суворими методами нелінійної теорії електричних коливань. Подано математичну модель такого автогенератора у вигляді нелінійних диференціальних рівнянь, отриманих на основі законів Кірхгофа. Їхній аналіз дозволив обґрунтувати перехід до моделі одноконтурного LC автогенератора, еквівалентного диференціальному. Запропоновано методику такого переходу, що включає процедуру визначення нелінійної характеристики підсилювального елемента цього автогенератора, виходячи з нелінійних характеристик двох підсилювальних елементів диференціального автогенератора.

Математична модель еквівалентного автогенератора представлена нелінійним диференціальним рівнянням Ван дер Поля у безрозмірній формі, вона проста та точна. Така форма уявлення дозволила виділити малий параметр та оцінити його величину. Що стосується малих величин малого параметра, як відомо, для її аналізу

можуть бути використані традиційні методи. Таким чином, завдання дослідження процесу синхронізації диференціального автогенератора зведено до дослідження процесу синхронізації автогенератора Ван дер Поля. Наведені результати можуть бути корисними при розробці різних пристроїв на базі синхронізованих диференціальних автогенераторів.

Ключові слова: диференціальний автогенератор; синхронізація; основний тон; математична модель.

Іл. 3. Бібліогр.: 14 назв.

BIOMEDICAL RADIO ELECTRONICS
БІОМЕДИЧНА РАДІОЕЛЕКТРОНІКА

UDC 004.056.53

Using a fingerprint scanner to protect data in medical information systems / *O.I. Dovnar, M.F. Babakov, V.I. Cherkis* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. №211. P. 148 – 153.

Since the beginning of hostilities, medical institutions have suffered from cyber attacks, the number of which is not decreasing. The load on the servers even before the start of the war was already quite significant. Since Ukraine suffered severe restrictions due to Covid-19, in part, most institutions were transferred to global medical databases, but with a serious breakthrough, digital capacities did not have time to grow. Even small cyber attacks in the medical field can be critical, so it is necessary to identify accurately medical personnel for access and implement a special application, which will not allow other people to get it, this makes it possible to block immediately all other requests without processing, and this is one of the ways to optimize and increasing security.

The task can be solved by standard means of authorization, but they are quite easy to bypass, ordinary software can also be reprogrammed, identification by phone number can be too expensive, so a combined method is proposed: special equipment for reading fingerprints and the corresponding software application. This method is sufficiently reliable and simple to implement (medical personnel will only need to put their finger to the sensor), as well as economical and not demanding.

The developed prototype for authorization on the Arduino hardware platform provides the necessary functionality and meets the task. The sensor recognizes more than 100 people, and the convenient mini USB allows you to connect conveniently the device to any computer port. A special algorithm will block the device in case of intervention by a person who does not have access to the corresponding resource. Can be easily reprogrammed and configured to the required equipment, and the case can be printed according to the necessary requirements.

Key words: DoS attacks; information protection; sensor; fingerprint scanner; 3D printing.

6 fig. Ref.: 15 items.

УДК 004.056.53

Використання сканеру відбитків пальців для захисту даних у медичних інформаційних системах / *О.Й. Довнар, М.Ф. Бабаков, В.І. Черкіс* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 148 – 153.

Від початку воєнних дій медичні установи потерпають від кібератак, число яких не зменшується. Навантаження на сервери ще до початку війни було вже досить значним. Оскільки Україна зазнала серйозних обмежень через Covid-19, частково більшість установ були переведені на глобальні медичні бази даних, але попри серйозний прорив цифрові потужності не встигали зростати. Навіть невеликі кібератаки у медичній сфері можуть бути критичними, тому є необхідність точно ідентифікувати медичний персонал для доступу та впровадити спеціальний застосунок який не дозволить отримати його іншим особам, що дозволяє одразу блокувати всі інші запити без обробки, а це один з шляхів оптимізації та нарощування безпеки.

Поставлена задача може бути вирішена стандартними засобами авторизації, але їх досить легко обійти, звичайні програмні засоби також можуть бути перепрограмовані, ідентифікація за номером телефону може бути занадто дорогою, тому пропонується комбінований спосіб: спеціальне устаткування для зчитування відбитків і відповідний програмний застосунок. Такий метод є достатньо надійним і простим у реалізації (медичному персоналу достатньо буде прикласти палець до сенсору), а також економічним та невибагливим.

Розроблений прототип для авторизації на апаратній платформі Arduino забезпечує необхідний функціонал та відповідає поставленому завданню. Сенсор розпізнає понад 100 осіб, а зручний mini USB дозволяє зручно підключити пристрій до будь-якого порту комп'ютера. Спеціальний алгоритм заблокує пристрій у разі втручання особи, що не має доступу до відповідного ресурсу. Може бути легко перепрограмований та налаштований до необхідного устаткування, а корпус може бути надрукований під необхідні вимоги.

Ключові слова: DoS напади; захист інформації; сенсор; сканер відбитків пальців; 3D-друк.

Іл. 6. Бібліогр.: 15 назв.

INFORMATION METHODS OF RADIO ENGINEERING, SIGNAL PROCESSING
ІНФОРМАЦІЙНІ МЕТОДИ РАДІОТЕХНІКИ, ОБРОБКА СИГНАЛІВ

UDC 621.396.967.2

Applying MATLAB to Radar Systems Modeling / *I.V. Svyd, A.O. Sierikov, I.I. Obod* // Radiotekhnika : All-Ukr. Sci. Interdep. Mag. 2022. № 211. P. 154 – 158.

The paper demonstrates the important role of using radar systems in many spheres of human activity to solve complex and important problems of society. The paper examines and analyzes the capabilities of the MATLAB

program package applying to the design, modernization and research of radar systems. The analysis performed shows that the MATLAB software package is a powerful tool for modeling, researching and designing radar systems for various purposes. Also, attention is paid to the MATLAB Radar Toolbox, which includes a wide range of available models that can be modified. The Radar Toolbox provides rapid modeling, upgrading and prototyping of standard and upgraded radar systems. Using the MATLAB to research radar systems is impossible without understanding the principles of construction and operation of radar systems. It is also necessary to know the features of using the MATLAB function to describe, present and model structural elements and processes in radar systems. The paper presents the main functionality and possible options for implementing models in the MATLAB for modeling and studying radar systems. The given list of model options is not exhaustive and final and can be expanded and supplemented depending on specific tasks and requirements for implementation.

Key words: radar system; modeling; model; signal; interference; MATLAB; Radar Toolbox; Radar Designer.

2 fig. Ref: 27 items.

УДК 621.396.967.2

Застосування MATLAB для моделювання радіолокаційних систем / *І.В. Свид, А.О. Сєріков, І.І. Обод* // Радіотехніка : Всеукр. міжвід. наук.-техн. зб. 2022. Вип. 211. С. 154 – 158.

Продемонстровано важливу роль застосування радіолокаційних систем у багатьох сферах людської діяльності для вирішення складних та важливих задач суспільства. Розглянуто та проаналізовано можливості пакету прикладних програм MATLAB для проєктування, модернізації й дослідження радіолокаційних систем. Аналіз показує, що пакет прикладних програм MATLAB є потужним засобом для моделювання, дослідження та проєктування радіолокаційних систем різного призначення. Приділено увагу інструментарію MATLAB Radar Toolbox, який включає широкий спектр наявних моделей, що мають можливість до внесення змін. Radar Toolbox забезпечує швидке моделювання, модернізацію та прототипування стандартних та модернізованих елементів радіолокаційних систем. Використання MATLAB для дослідження радіолокаційних систем неможливе без розуміння принципів побудови та функціонування радіолокаційних систем. Також необхідно знати особливості застосовування функції MATLAB для опису, представлення та моделювання структурних елементів і процесів у радіолокаційних системах. Наведено основні функціональні можливості та можливі варіанти реалізації моделей у MATLAB для моделювання та дослідження радіолокаційних систем. Наведений перелік варіантів моделей не є вичерпним і остаточним та може бути розширений і доповнений в залежності від конкретних завдань і вимог до реалізації.

Ключові слова: радіолокаційна система; моделювання; модель; сигнал; завада; MATLAB; Radar Toolbox; Radar Designer.

Іл. 2. Бібліогр.: 27 назв.

COLLECTION OF SCIENTIFIC PAPERS
RADIOTEKHNIKA
Issue 211
In English and Ukrainian Russian

ЗБІРНИК НАУКОВИХ ПРАЦЬ
РАДІОТЕХНІКА
Випуск 211
Англійською та українською мовами

Коректор Л.І. Сащенко

Підп. до друку 30.12.2022. Формат 60x90/8. Папір офсет. Гарнітура Таймс. Друк. ризограф.
Ум. друк. арк**.** 11,9. Обл.-вид. арк. 10,3. Тираж 300 прим. Зам. № 501. 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.