УДК 004.4.277 : 004.056

ШИРОЧИН С.С.,
СУЛЕМА Є.С.

СПОСІБ СТЕГАНОГРАФІЧНОГО ЗАХИСТУ ДАНИХ В АУДІО-ФАЙЛАХ НА ОСНОВІ КОМПЛЕМЕНТАРНОГО ОБРАЗУ

Хмарні сховища є зручним засобом для збереження цифрових мультимедійних даних користувача (зображення, аудіо- та відеодані), й вони все більше набувають популярності серед широкого кола користувачів. Але при використанні хмарних сховищ, що не надають гарантії захисту даних, виникає небезпека доступу до особистих мультимедійних даних користувача сторонніх осіб. Використання існуючих способів захисту інформації для захисту мультимедійних даних користувача у хмарних сховищах не є доцільним, оскільки, по-перше, вони не враховують особливостей мультимедійних даних, а по-друге, не є орієнтованими на потреби звичайного користувача, зокрема, при розробленні засобів захисту звичайно не береться до уваги час обробки даних при їх шифруванні та дешифруванні, хоча для користувача швидкодія є важливою характеристикою програми. Тому розроблення нових способів захисту мультимедійних даних користувача є актуальною задачею.

В даній статті пропонується стеганографічний спосіб захисту даних в аудіо-файлах, який ґрунтується на збереженні комплементарного образу стегоданих замість самих стегоданих та використанні в ролі ключа модифікованого латинського квадрату. При побудові комплементарного образу використовуються особливості мультимедійних даних, а саме: багатоканальність аудіо. Уданому способі передбачається виконання паралельних обчислень, що дозволяє досягти кращих часових показників процедури кодування та декодування стегоданих порівняно з використанням алгоритму шифрування AES. Спосіб, що пропонується, може бути використаний для захисту особистих даних користувача в реальному часі при їх передаванні по мережі та зберіганні в хмарних сховищах.

Cloud storages give a convenient way for keeping user's digital multimedia data (images, audio and video data), and they are becoming more and more popular among a large number of users. However in case of using cloud storage that doesn't guarantee data protection a threat of unauthorized access to user's personal multimedia data appears. The application of existing protection methods to multimedia data in clouds isn't appropriate, because, firstly, they don't take into account the specific properties of multimedia data, and secondly, they aren't oriented to needs of a usual user. In particular, time of data processing during encryption and decryption isn't usually considered as a characteristic of data protection software, though it's important for a user. Therefore the development of specific methods of multimedia data protection is a topical task.

In this paper a new steganography method of data protection in audio-files is proposed, it's based on storing a complimentary image instead of stegodata as well as on the use of a modified Latin square as a key. Special properties of multimedia data are used while creating a complimentary image, in particular multi-channel nature of audio data is taken into account. Parallel computations are stipulated to be used in the proposed method. It lets to achieve better time characteristics of stegodata encryption and decryption in comparison with the use of AES encryption algorithm. The proposed method can be used for user's personal data protection in real time while transferring through network and keeping in cloud storages.

Вступ

Поширення використання хмарних сховищ даних призводить до того, що більшість користувачів Інтернету мають певну долю власних даних, що зберігаються в хмарних сховищах. Зокрема, користувачі зберігають в хмарних сховищах свої персональні мультимедійні дані – фотографії, аудіо- та відеозаписи. Дана інформація носить характер особистої і не має розповсюджуватися без згоди власника. Проте, зберігаючи дані поза локальним комп'ютером, користувач втрачає контроль за доступом до цих даних. Оскільки не кожен користувач є фахівцем із захисту інформації, задача розроблення простого та доступного кожному користувачу способу захисту персональних мультимедійних даних від несанкціонованого доступу в хмарних сховищах є актуальною.

Одним зі способів захисту даних є стеганографічний захист, який полягає у приховувані самого факту наявності деяких секретних даних. Позитивною якістю стеганографії для захисту мультимедійних даних є те, що стеганографічний спосіб дозволяє приховувати зображення у зображенні та аудіо-дані у аудіо-даних, що на практиці означає можливість збереження у відкритому доступі (наприклад, у хмарному

сховищі) замість мультимедійних даних користувача, які не підлягають поширенню (*стегоданих*), іншого зображення або аудіо-запису, що насправді є *контейнером*, який приховує стегодані. Таким чином, користувач може вільно користуватись відкритим сховищем, приховуючи сам факт наявності у нього деяких мультимедійних даних, які він хоче тримати у таємниці. Проте стеганографія має суттєвий недолік: якщо факт наявності стегоданих у контейнері стане відомим зловмиснику, несанкціонований доступ до них буде технічно нескладною задачею. Отже, доцільним є включення у процедуру захисту етапу попереднього шифрування стегоданих. Це шифрування може виконуватись за деяким відомим алгоритмом криптографічного захисту (наприклад, DES, AES, RSA тощо) або за спеціалізованим алгоритмом шифрування [1,2]. В даному досліджені пропонується спосіб стеганографії аудіо-даних, у якому як контейнер використовується аудіо-файл, а підвищення ступеня захисту стегоданих досягається за рахунок їх попереднього перетворення за допомогою *комплементарного перетворення*. Окремо слід підкреслити, що оскільки для користувача важливим є не лише захист даних, а й час їх шифрування та дешифрування, тому у даній статті питанню швидкодії алгоритму обробки стегоданих приділяється особлива увага.

Мета дослідження

Метою дослідження, результати якого наведено у дані статті, є створення стеганографічного способу захисту даних користувача в аудіо-файлах при їх зберіганні у хмарних сховищах зі забезпеченням підвищеної швидкодії обробки даних за рахунок виконання паралельних обчислень.

Опис способу

Спосіб, що пропонується, є модифікацією способу стеганографії зображень [2]. Однією із особливостей способу є використання природних властивостей аудіо-даних, а саме – *багатоканальності*. Однакова тривалість звукових каналів дає можливість їх комбінування для забезпечення захисту даних, що приховуються. В даному способі пропонується модифікація молодших бітів (LSB – *Least significant Bits*) [3] значень рівня звукового сигналу лише в одному з каналів, який будемо називати *контейнерним* каналом.

Іншою й основною особливістю способу є використання зашифрованого образу стегоданих, що утворюється за допомогою комплементарного перетворення, яке задає відображення множини стегоданих на множину даних одного зі звукових каналів контейнера, що не використовується в ролі контейнерного каналу. Будемо називати цей канал *комплементарним*.

Результатом цього відображення є множина зашифрованих даних, яку будемо називати *комплементарним образом (КО)*. КО вбудовується в контейнер замість стегоданих, проте тільки в один канал – контейнерний.

Комплементарне перетворення задається таблицею (*ключем*), що зберігає відповідність між байтами комплементарного каналу та байтами стегоданих. Комплементарне перетворення має задовольняти такі вимоги:

- для будь-якої пари значень даних у комплементарному каналі та стегоданих, що знаходяться в певному діапазоні значень, має існувати одне та лише одне значення ключа, що знаходиться у тому ж діапазоні;
- має існувати однозначне зворотне перетворення, що дозволяє отримати стегодані з відповідних даних КО, комплементарного каналу і ключа.

Ці вимоги задовольняє комплементарне перетворення, що ґрунтується на застосуванні латинського квадрату [4]. Створення ключа на основі латинського квадрату відбувається за алгоритмом, розробленим для захисту графічних даних [2]. Даний алгоритм передбачає, відповідно до наведених вище вимог, формування таблиці, яка заповнюється за наступними правилами:

- кожне значення може зустрічатись в рядку лише один раз;
- кожен рядок має містити всі варіанти значень у заданому діапазоні;
- послідовність значень в рядках має бути випадковою.

Таким чином, генерування ключа в способі, що пропонується, відбувається за наступним алгоритмом:

1. Створюється латинський квадрат розміром 256×256.
2. Для кожного рядка генерується псевдовипадкове число кількості перестановок у діапазоні від 128 до 255.
3. Для кожної перестановки генеруються два псевдовипадкових числа у діапазоні від 0 до 255.
4. Елементи з номерами двох отриманих чисел міняються місцями.

Отриманий ключ використовується для створення КО.

Перед застосуванням способу необхідно попередньо виконати аналіз місткості контейнера, враховуючи наступні міркування.

Для зберігання даних використовується лише один канал, при чому змінювати можна лише старший байт кожного елементу типу *double*. Через це обсяг даних, що вбудовуються у контейнер, збільшується в 4 рази, порівняно зі звичайним стеганографічним способом захисту на основі модифікації LSB. Розмір КО (тобто об'єм даних, що вбудовуються у контейнер) залежить від кількості бітів b, що модифікуються: оскільки значення b може обиратись у діапазоні від 1 до 8, це може допомогти компенсувати збільшення розміру КО.

Отже, розміщення стегоданих в контейнері є можливим за виконання наступної умови:

$$\frac{C \cdot b}{8} \geq 4 \cdot I,$$

що спрощується до умови:

$$C \geq \frac{32 \cdot I}{b},$$

де b – кількість біт, що модифікуються;

I – розмір стегоданих;

C – розмір контейнера;

Процес створення КО полягає у заміні пари значень $< \mathrm{S_i}, \mathrm{C_j} >$, де S_i – i-й байт стегоданих, $\mathrm{C_j}$ – j-й байт контейнера, значенням X_i, яке являє собою i-й байт КО, відповідно до ключа T. При цьому S_i та $\mathrm{C_j}$ використовуються в ролі координат комірки зі значенням X_i у ключі (таблиці) T.

Отриманий блок даних (тобто КО) вбудовується в контейнерний канал шляхом модифікації LSB. Зміна даних у комплементарному каналі не відбувається. Оскільки для збереження аудіо-даних в форматі *wave*, до якого в даному способі перетворюються всі вхідні аудіо-формати, використовується тип даних *double*, що займає 2 байти, то модифікація LSB відбувається, як було зазначено вище, лише у старшому байті кожного елементу типу *double*: байти в *double* розташовані в зворотному порядку і модифікація старшого байту менше впливає на значення *double*. В більшості випадків старший байт містить нулі в молодших бітах, тому для модифікації може бути використано більше, ніж 1 біт старшого байта. При реалізації даного способу програмним шляхом була забезпечена можливість модифікації від 1 до 8 біт старшого байта.

Після вбудовування даних у контейнер, аудіо-дані ущільнюються без втрат. Для збереження контейнеру можуть бути використані будь-які формати з ущільненням без втрат, наприклад, FLAC.

Ключ передається окремо від основних даних згідно зі загальноприйнятими правилами передачі секретного ключа.

Схема захисту даних в аудіо-файлі представлена на рис. 1.

Декодування стегоданих відбувається аналогічно декодуванню у способі стеганографії зображень на основі комплементарного образу [2] і складається з двох етапів. Спочатку відбувається зчитування комплементарного образу, що для біту n байту k комплементарного образу визначається наступним чином:

$$X_k = X_k | (C_i \& (255 \gg (8 - b))) \ll (8 - n - b),$$

де X_k – k-й байт комплементарного образу;

C_i – i-й байт контейнера;

b – кількість стегобіт в байті контейнера;

n – порядковий номер стегобіта в послідовності стегобіт, $n \leq b$.

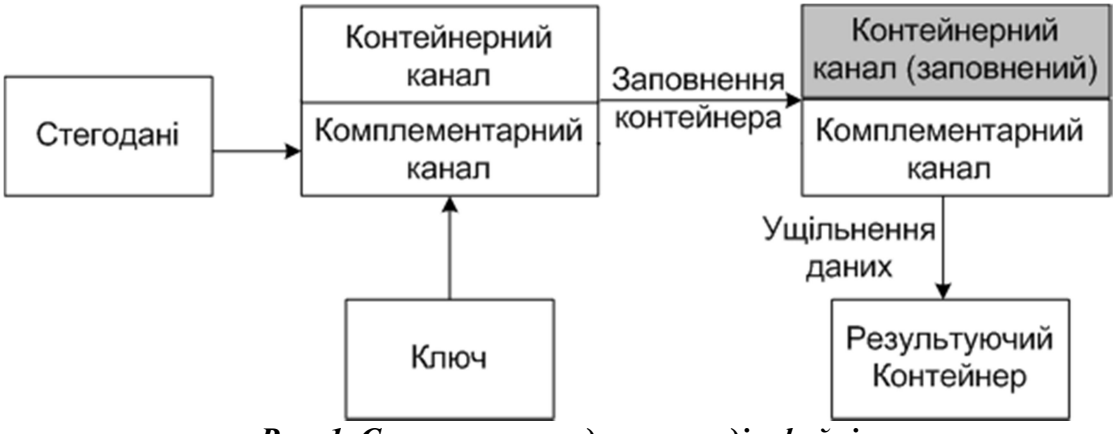

Рис. 1. Схема захисту даних в аудіо-файлі

Далі з КО отримується приховане зображення. З пари значень $< C_i, X_i >$, де C_i – i-й байт контейнерного каналу, X_i – i-й байт комплементарного каналу, за допомогою ключа T отримується значення S_i, що являє собою i-й байт прихованого зображення.

Оскільки канали розміщуються в аудіо-файлі паралельно, зі зсувом на 2 байти, адреси КО можна виразити через адреси контейнера, тоді значення S_i буде отримуватись з пари $< C_i, C_{i+2} >$, якщо правий канал є комплементарним, а лівий – контейнерним і з пари $< C_{i+2}, C_i >$ в протилежному випадку.

Розпаралелювання

Для покращення часових характеристик алгоритму реалізації запропонованого способу було застосоване розпаралелювання між ядрами багатоядерного процесора. Частина алгоритму, що має бути розпаралелена, обрана аналогічним чином до способу стеганографії зображень на основі комплементарного образу [2]. Розпаралелювання було застосовано до запису даних в контейнер та до зчитування їх з контейнеру.

Програмно алгоритм запису даних в контейнер (рис. 2) реалізується у вигляді двох вкладених циклів, де зовнішній цикл відповідає за побайтову обробку стегоданих, а внутрішній – за вбудовування кожного окремого біту. Кількість ітерацій зовнішнього циклу залежить від об'єму стегоданих, кількість ітерацій внутрішнього – від кількості стегобіт. Для обрання оптимального способу розпаралелювання було розглянуто декілька варіантів паралельних реалізації алгоритму:

1. Розпаралелювання лише внутрішнього циклу.
2. Розпаралелювання як зовнішнього, так і внутрішнього циклів.
3. Розпаралелювання лише зовнішнього циклу.

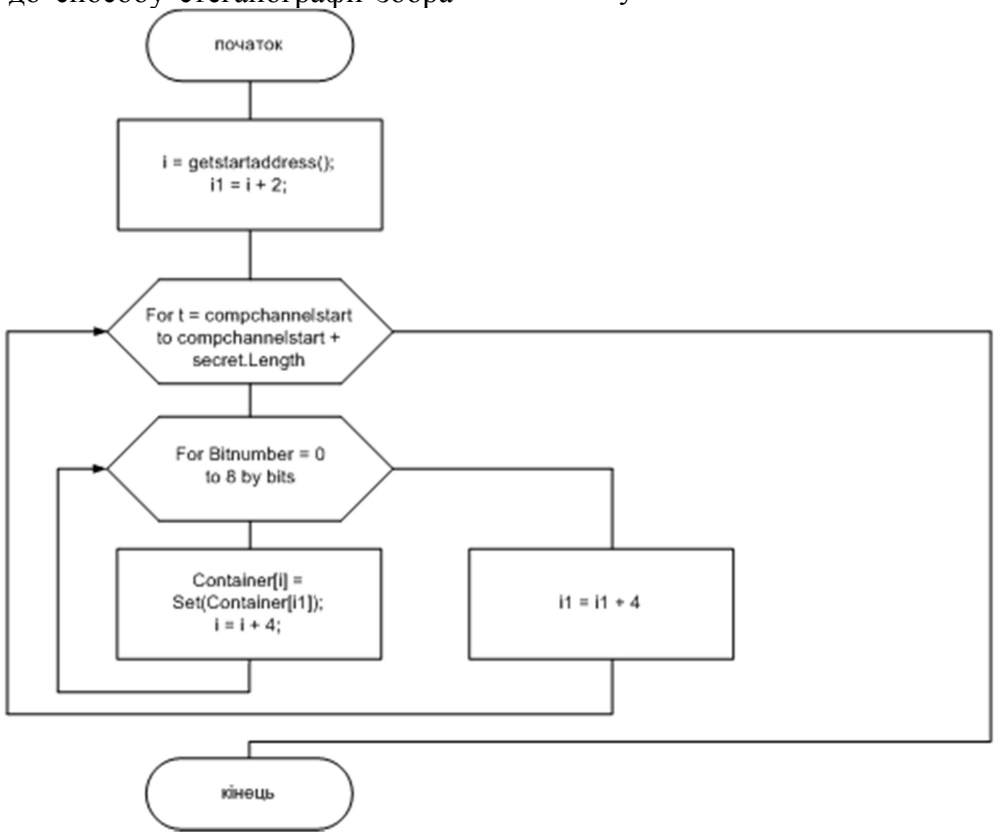

Рис. 2. Блок-схема алгоритму запису даних в контейнер без використання розпаралелювання

Оскільки внутрішній цикл має обмежену кількість ітерацій (від 1 до 8), то у першому та другому варіантах розпаралелювання є неефективним через те, що навіть у випадку 2 ядер та 8 стегобіт кожен паралельний потік буде виконаний лише 4 рази. В більшості випадків використовується мінімальна кількість стегобіт, тому розпаралелювання буде лише збільшувати обчислювальну складність через необхідність організації паралельних потоків, у яких фактично буде виконуватись лише 1 ітерація.

Найбільш раціональним є варіант розпаралелювання лише зовнішнього циклу, оскільки в такому випадку у кожному потоці буде виконуватись достатньо велика кількість ітерацій.

Схема роботи розпаралеленого алгоритму запису даних в контейнер зображена на рис. 3.

Розпаралелювання алгоритму зчитування даних з контейнеру було виконане аналогічним чином

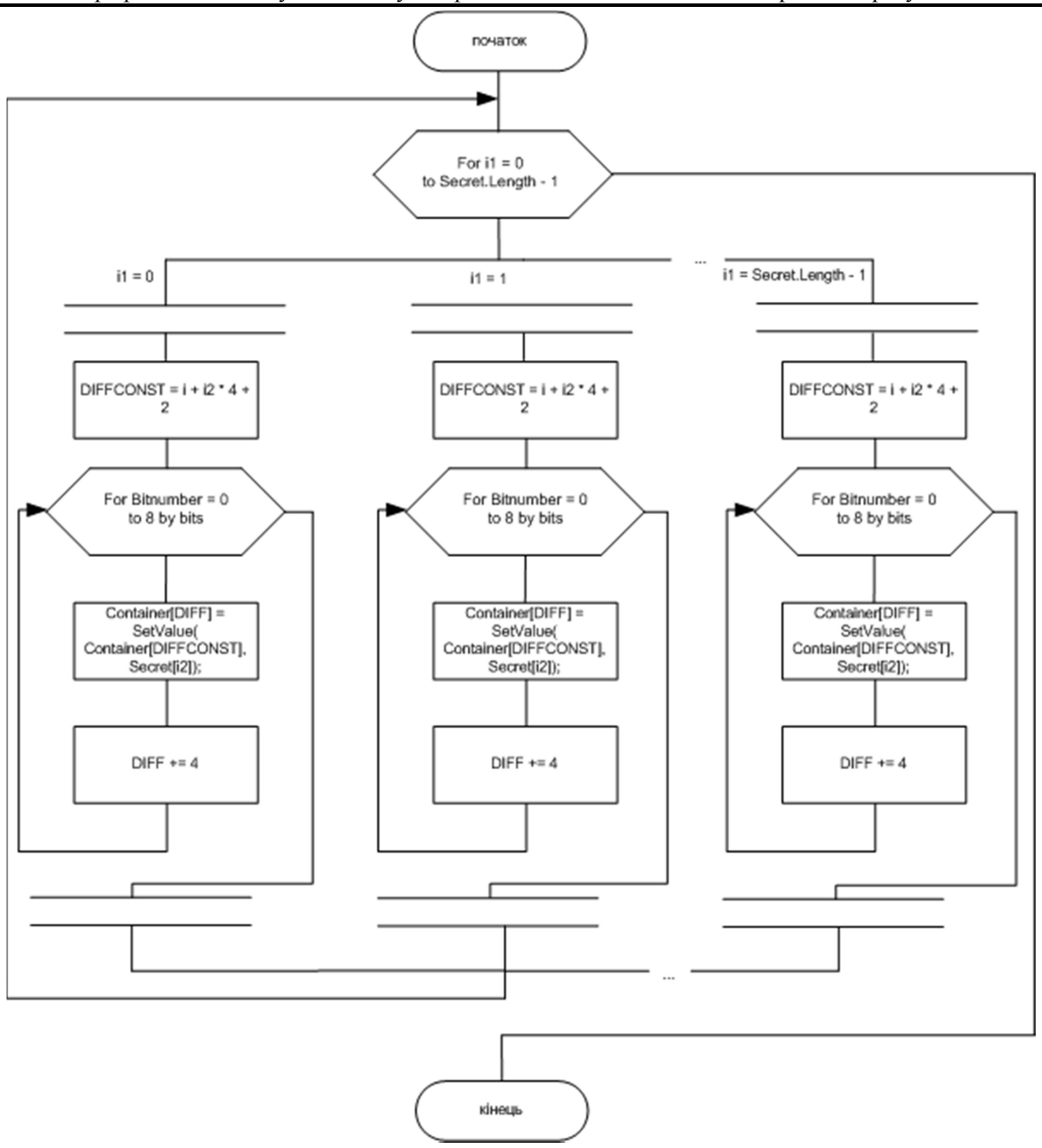

Рис. 3. Блок-схема алгоритму запису даних в контейнер з використанням розпаралелювання

Результати

Запропонований спосіб оцінювався за двома критеріями: ступінь захисту стегоданих та швидкодія алгоритму.

Ступінь захисту може бути оцінений наступним чином. Ключ має фіксований розмір 64 Кб, тобто 524288 біт, отже складність підбору ключа буде 2^{524288}. Для порівняння – максимальний розмір ключа AES дорівнює 256 бітам, отже складність його підбору складає 2^{256}. Оскільки у запропонованому способі передбачається використання ключа лише один раз, це гарантує захист від фальсифікації ключа.

При оцінці швидкодії програмної реалізації запропонованого способу часові характеристики алгоритму порівнювались з часовими характеристиками таких алгоритмів:

- алгоритм стеганографічного захисту на основі модифікації LSB (базовий спосіб захисту),
- алгоритм стеганографічного захисту на основі модифікації LSB з використанням шифрування AES (виконується попереднє шифрування стегоданих за алгоритмом AES перед виконанням модифікації LSB),
- алгоритм стеганографічного захисту з використанням фрагментації [5],
- алгоритм на основі КО в двох варіантах: звичайному та розпаралеленому.

Всі заміри проводились на ПК з двоядерним процесором Intel Centrino, під операційною сис-

темою Windows 7. В експериментах використовувався однаковий набір файлів. В ролі контейнера використовувався звуковий файл розміром 16 Мб, а в ролі стегоданих – файл розміром 350 Кб.

На рис. 4 наведені результати замірів часу запису стегоданих в контейнер. Аналіз отриманих експериментальних даних дозволяє зробити наступні висновки:

1. Включення у процедуру стеганографічного захисту будь-якого шифрування уповільнює час роботи алгоритму.
2. Збільшення кількості стегобіт прискорює запис при всіх способах шифрування.
3. Швидкодія алгоритму захисту на основі комплементарного образу залежить від кількості стегобіт: при 1 та 2 стегобітах він дає гірші результати, ніж інші способи; при 4 і 8 стегобітах він є швидшим, ніж алгоритм з використанням шифрування AES.
4. Розпаралелювання алгоритму захисту на основі комплементарного образу дозволяє досягнути збільшення швидкодії. Чим менша кількість стегобіт, тим більший ефект від розпаралелювання.

Також було заміряно час зчитування стегоданих (рис. 5). Аналіз отриманих даних дозволяє зробити такі висновки:

1. Алгоритм захисту на основі комплементарного образу показує кращі часові результати, ніж алгоритм з використанням шифрування AES.
2. Розпаралелювання алгоритму захисту на основі комплементарного образу при зчитуванні дає прискорення лише при 1 та 2 стегобітах. В інших випадках значення приблизно однакові (4 стегобіти) або паралельна реалізація вимагає більше часу (8 стегобіт).
3. Збільшення кількості стегобіт прискорює всі алгоритми, оскільки модифікується менша кількість байт контейнера.

Також було проаналізовано залежність швидкодії алгоритму від кількості стегобіт. Було виявлено, що при збільшенні кількості стегобіт в кожному байті контейнера швидкодія алгоритму зростає, оскільки для вбудовування тієї ж самої кількості стегоданих потрібно модифікувати меншу кількість байт контейнера, отже обробка даних виконується за меншу кількість ітерацій. Значення в табл. 1 показують, наскільки швидше алгоритм виконується у випадку 2, 4 та 8 стегобіт, ніж при 1 стегобіті за однакових інших умов.

Отримані результати дозволяють зробити висновок, що збільшення швидкодії при записі спостерігається для всіх способів, особливо для способу на основі комплементарного образу. При цьому паралельна реалізація має менші відносні показники приросту швидкодії, ніж послідовна. При збільшенні кількості стегобіт ефективність розпаралелювання зменшується, проте воно все ще є доцільним.

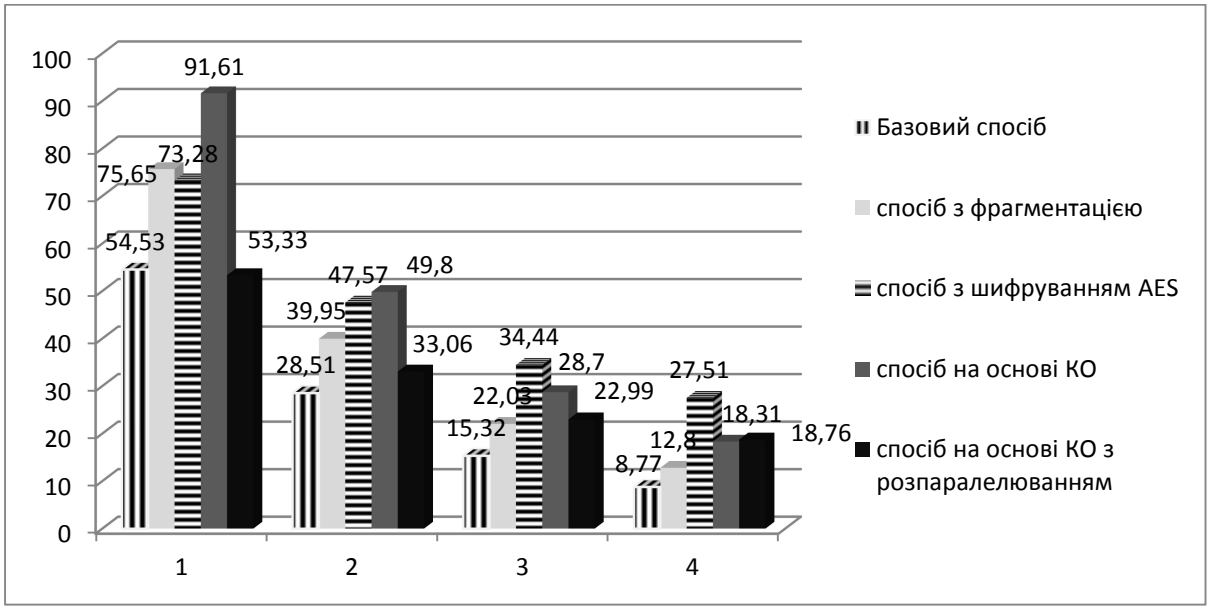

Рис. 4. Час запису стегоданих (мс): 1 – 1 біт, 2 – 2 біта, 3 – 4 біта, 4 – 8 біт

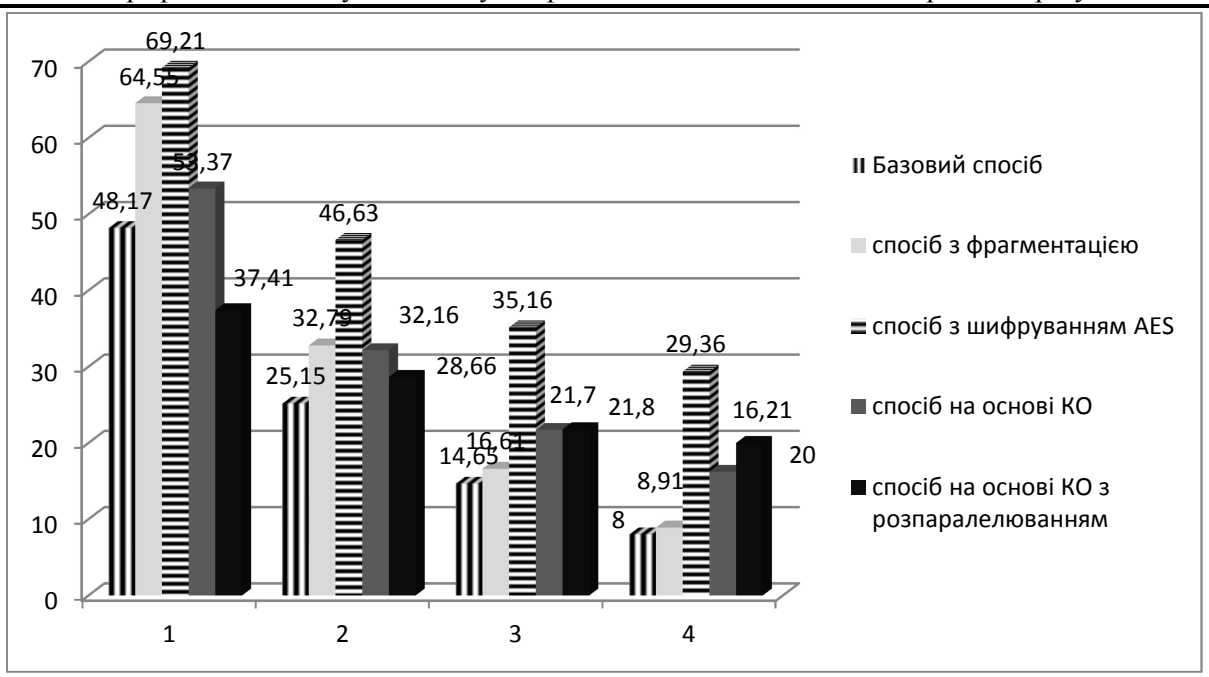

Рис. 5. Час зчитування стегоданих (мс): 1 – 1 біт, 2 – 2 біта, 3 – 4 біта, 4 – 8 біт.

Табл. 1. Ступінь прискорення при збільшенні кількості стегобіт

Алгоритм	Запис			Зчитування		
	2 біти	4 біти	8 біт	2 біти	4 біти	8 біт
Базовий спосіб	1,91	3,56	6,22	1,92	3,29	6,02
Спосіб з фрагментацією	1,89	3,43	5,91	1,97	3,89	7,24
Спосіб з шифруванням AES	1,54	2,13	2,66	1,48	1,97	2,36
Спосіб на основі КО	1,84	3,19	5,00	1,66	2,46	3,29
Спосіб на основі КО з розпаралелюванням	1,61	2,32	2,84	1,31	1,72	1,87

Аналіз приросту швидкодії при зчитуванні даних дає інші результати. Найбільший приріст швидкодії спостерігається для базового способу та для способу з фрагментацією, найменший – для способу на основі комплементарного образу з розпаралелюванням. При 1 та 2 стегобітах розпаралелювання є доцільним, а при 4 та 8 ефект втрачається і стає зворотним. Проте у всіх випадках абсолютні часові показники способу на основі комплементарного образу є кращими, ніж при використанні способу з шифруванням AES, особливо при паралельній реалізації алгоритму.

Висновки

Порівняння запропонованого способу зі способом стеганографічного захисту з використанням шифрування AES дозволяє зробити висновок про те, що оскільки розмір ключа на основі латинського квадрату є більшим, ніж розмір ключа при шифруванні AES, то підбір ключа ускладнюється, що підвищує стійкість способу, але водночас великий розмір ключа на основі латинського квадрату спричинює збільшення часу, що витрачається на його генерування.

Для покращення часових характеристик використовується розпаралелювання алгоритму для виконання на багатоядерних процесорах. Аналіз швидкодії алгоритму реалізації запропонованого способу показав його перевагу порівняно з алгоритмами реалізації інших способів, зокрема з алгоритмом реалізації способу з використанням шифрування AES. Час запису стегоданих у всіх випадках є меншим, ніж при шифруванні, незважаючи на те, що частина часу витрачається на генерування ключа. При зчитуванні стегоданих даний алгоритм є також більш швидким, ніж при застосуванні шифрування, проте його паралельна реалізація є доцільною лише при 1 та 2 стегобітах. При дослідженні виконувалось розпаралелювання на 2 ядра, отже при використанні даного способу на

процесорах з більшою кількістю ядер ефективність розпаралелювання буде більшою. Розпаралелювання обчислень на 2 ядра не дає двократного приросту швидкодії, тому що частина ресурсів витрачається на організацію взаємодії роботи паралельних потоків.

Отже, даний спосіб забезпечує достатній ступінь захисту даних та високі часові показники обробки даних і може бути застосований для захисту користувацьких даних в реальному часі. Зокрема запропонований спосіб може бути використаний для захисту персональних мультимедійних даних користувача при їх зберіганні в хмарних сховищах.

Список посилань

1. Сулема Є.С., Широчин С.С. Метод захисту зображень на основі шифрування палітри // Науковий журнал «Вісник Хмельницького національного університету». – Хмельницький : ХНУ. – 2014. – №3. – С. 114-119.
2. Сулема Є.С., Широчин С.С. Спосіб стеганографії зображень на основі комплементарного образу // Журнал «Захист інформації». – Київ : НТУУ «КПІ». – 2013. – Випуск 4. – С. 345-353.
3. Owens, M. A discussion of covert channels and steganography // SANS Institute, 2002.
4. Латинские квадраты [Електронний ресурс] // Режим доступу: http://www.e-olimp.com/en/problems/2105 – [02.03.2012].
5. Сулема Є.С., Широчин С.С. Спосіб стеганографії зображень з фрагментацією стегоданих та розділенням закритого ключа. // Науково-технічний збірник «Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні». – Київ : НТУУ «КПІ». – 2012. – Випуск 1 (22). – С. 64-68.