МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
КАФЕДРА КОМП'ЮТЕРИЗОВАНИХ СИСТЕМ ЗАХИСТУ ІНФОРМАЦІЇ

ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач кафедри
_____ С.В. Казмірчук

«_____»_____ 2020 р.

На правах рукопису
УДК 004.056.5:510.22(043.3)

КВАЛІФІКАЦІЙНА РОБОТА

ЗДОБУВАЧА ВИЩОЇ ОСВІТИ
ОСВІТНЬОГО СТУПЕНЯ «МАГІСТР»

Тема: Ентропійна стеганографічна система захисту повідомлень в комп'ютерних мережах

Автор: О.А. Микитенко

Науковий керівник: к.т.н. О.О. Високцька

Нормоконтролер: к.т.н. О.О. Високцька

Київ 2020

ВСТУП

Актуальність. В даний час істотним внеском в напрямку забезпечення достатнього рівня безпеки служить розробка і широке застосування механізмів захисту інформації, заснованих на методах криптології - науково-практичної області, що включає в себе в якості складових частин криптографію, криптоаналіз, стеганографію і стеганоаналіз. У ряді країн світу введені обмеження на використання криптографічних засобів.

Внаслідок цього посилюється затребуваність тих методів і механізмів захисту інформації, які відносяться до стеганографії. Крім того, актуальність даної роботи визначається підвищенням ступеня захисту повідомлень, що передаються в стеганографічних системах, шляхом зменшення кількості змінних елементів в стеганографічному контейнері при впровадженні в нього повідомлення, та в розробці та використанні стеганографічного алгоритму захисту інформації, при застосуванні якого кількість змінюваних елементів контейнера залежить від ймовірносно-статистичних характеристик впроваджуваного повідомлення.

Метою дипломної роботи є удосконалення ентропійної стеганографічної системи захисту повідомлень в комп’ютерних мережах. Для досягнення поставленої мети необхідно вирішити наступні завдання:

- проаналізовати сучасні стеганографічні системи та алгоритми;
- проаналізувати принципи застосування та елементи ентропійного підходу в стеганографії;
- розробити аналітичний апарат для побудови ентропійних стеганографічних алгоритмів;
- розробити оптимальний та субоптимальний ентропійні стеганографічні алгоритми для використання їх в ентропійних стеганографічних системах захисту повідомлень в комп’ютерних мережах.

Об'єкт дослідження: процес захисту інформації в інформаційних

мережах за рахунок використання стеганографічних технологій.

Предмет дослідження: алгоритми захисту конфіденціальної інформації в комп’ютерних мережах з використанням стеганографічних технологій.

Галузь застосування: алгоритми, розроблені в даній атестаційній роботі, в тому числі і методика побудови ентропійних стеганографічних систем захисту інформації, відноситься до галузі інформаційної безпеки і можуть знайти застосування при створенні стеганографічних систем для захисту повідомлень в інформаційних мережах, стати відправною точкою для подальших досліджень в цій області.

Наукова новизна:

– вперше запропоновано оптимальний ентропійний стегонографічний алгоритм захисту повідомлень, попередньо стиснений асимптотично оптимальним блоковим рівномірним кодом та субоптимальний ентропійний стеганографічний алгоритм захисту повідомлень (що не піддаються попередньому стиску).

– Вперше розроблена методика побудови ентропійних стеганографічних систем захисту повідомлень, в яких використовуються ентропійні стеганографічні алгоритми. Це дало можливість посилити захист інформації в комп'ютерних мережах за рахунок попереднього стиску повідомлень, згенерованих двійковим джерелом без пам'яті.

Практична цінність: Практична цінність роботи полягає в тому, що розроблено ентропійну стеганографічну систему захисту повідомлень в комп’ютерних мережах з використанням алгоритму та методики побудови ентропійних стеганографічних систем, що дозволяє гарантувати захист інформації при її зберіганні,обробці та передачі відкритими каналами звязку.

Апробація роботи. Основні положення роботи доповідалися та обговорювалися на наступній конференції:

- Микитенко О.А. Ентропійна стеганографічна система захисту повідомлень в комп’ютерних мережах / Микитенко О.А., Висоцька О.О. // ОБРАЗОВАНИЕТО И НАУКАТА НА XXI ВЕК - 2020, October 15-22, 2020:

Sophia. XVI МЕЖДУНАРОДНА НАУЧНА ПРАКТИЧНА КОНФЕРЕНЦИЯ – с. 59-61. URL: http://www.rusnauka.org/cgi bin/search/step7_info.cgi?id=284497&idw=jMFtM7XsRQbMw7pnFi

Розділ 1. Особливості побудови стеганографічних систем

1.1. Аналіз та визначення стеганографічної системи.

Стеганографія — це складова частина криптології , що представляє собою сукупність методів і способів, що забезпечують захист інформації від несанкціонованого доступу під час передачі, зберігання та обробки шляхом приховування самого факту її існування.

Традиційно до стеганографії відносять передачу (зберігання) інформації з використанням стеганографічного контейнера. Суть такого способу полягає в наступному. Спочатку вибирається сукупність даних, яка називається контейнером, а потім в неї певним чином закладається приховуване повідомлення (інформація), яке потім передається або зберігається в ньому.

У даній роботі під стеганографічними системами будемо мати на увазі тільки стеганографічні системи з ключами.

Розрізняють такі напрямки в стеганографії:

▫ радіоелектронна стеганографія;

▫ цифрова стеганографія;

▫ комп'ютерна стеганографія.

Під **радіоелектронною стеганографією** розуміється сукупність методів приховування інформації в аналогових потоках: широкосмугових, шумоподібних сигналах; включення сигналів в радіоповідомлення і музику; передача з використанням стрибаючих частот.

Під **цифровою стеганографією** розуміється сукупність методів приховування інформації в потоках оцифрованих (тобто, перетворених в дискретну форму) сигналів, що мають безперервну аналогову природу.

Під **комп'ютерною стеганографією** розуміється сукупність способів приховування інформації в комп'ютерних даних, що представляють собою різні файли, програми, пакети протоколів і т. п.

В даний час межа між цифровою та комп'ютерною стеганографією є

досить умовною в зв'язку з масовою комп'ютеризацією всіх областей людської діяльності, поширенням мультимедійних технологій і засобів телекомунікацій. Аналогічно тому, як в системах зв'язку аналогові сигнали (аудіо, відео) перетворюються в форму дискретних послідовностей, які діляться на пакети і передаються по мережах зв'язку, так і комп'ютерні дані, відповідні зображенням, звуковим або відеосигналам, подаються у вигляді файлів або передаються у вигляді пакетів з комп'ютерної мережі.

Повідомлення - це інформація, що підлягає захисту від несанкціонованого доступу при її передачі, зберіганні та обробці, шляхом приховування самого факту її існування.

Передбачається, що повідомлення являє собою сукупність елементів, які називаються **елементарними одиницями повідомлення** , або просто **одиницями повідомлення** (або **елементами повідомлення**).

Контейнером називається набір даних, який використовується для приховування в ньому повідомлення.

Передбачається, що контейнер являє собою сукупність елементів, які називаються **елементарними одиницями контейнера** , або просто **одиницями контейнера** (або **елементами контейнера**).

Контейнер називається заповненим контейнером, або стеганограммой (або контейнером-результатом), якщо він містить приховане повідомлення.

Контейнер, який не містить прихованого повідомлення, називається порожнім контейнером (або контейнером-оригіналом).

Існують декілька варіантів вибору контейнера:

- контейнер генерується стегосистемою;

- контейнер вибирається з деякої підмножини контейнерів за критерієм ефективності приховування даних;

- контейнер поступає ззовні;

- контейнер створюється шляхом моделювання шумових характеристик;

- контейнер вибирається із числа зарезервованих полів протоколу, які

не використовуються в процесі передачі даних.

На практиці найчастіше застосовується така процедура вибору контейнера: спочатку вибирається клас достатньо шумних контейнерів і ідентифікуються біти шуму. Потім визначається, яку частину шумових бітів контейнера можна замінити псевдовипадковими даними без значних змін в його статистичних характеристиках. Для забезпечення додаткового рівня захисту вбудовані дані зашифровують криптоалгоритмом [2].

За протяжністю розрізняють неперервні (потокові) і обмежені (фіксовані контейнери). Потоковий контейнер представляє собою неперервну послідовність біт. Повідомлення вбудовується в нього в режимі реального часу, тому невідомо заздалегідь, чи вистачить розміру контейнера для передачі всього повідомлення. Якщо розмір контейнера достатньо великий, є можливість вкладення декількох повідомлень. Інтервали між вбудовуваними бітами визначаються генератором псевдовипадкової послідовності з рівномірним розподілом інтервалів між відліками. Основна складність полягає в здійсненні синхронізації, визначенні початку і кінця повідомлення.

У фіксованого контейнера розміри і характеристики відомі наперед [2]. Це дозволяє вибрати оптимальний шлях для вкладення приховуваної інформації.

Вбудоване повідомлення, яке знаходиться в стегоконтейнері, передається від відправника до отримувача по каналу передачі, який називається стеганофонічним каналом або просто стеганоканалом.

В процесі передачі повідомлення може піддаватися різного роду трансформаціям: зменшуватися або збільшуватися, перетворюватись в інший формат і т.д. Крім того, воно може бути стиснуте, в тому числі з використанням алгоритмів стиснення з втратою даних. Саме тому стегоповідомлення повинно бути стійким до спотворень такого роду.

Задачу вбудовування і виділення повідомлення з контейнера виконує стегосистема [2]. Стегосистема складається з наступних основних

елементів (рисунок 1.2):

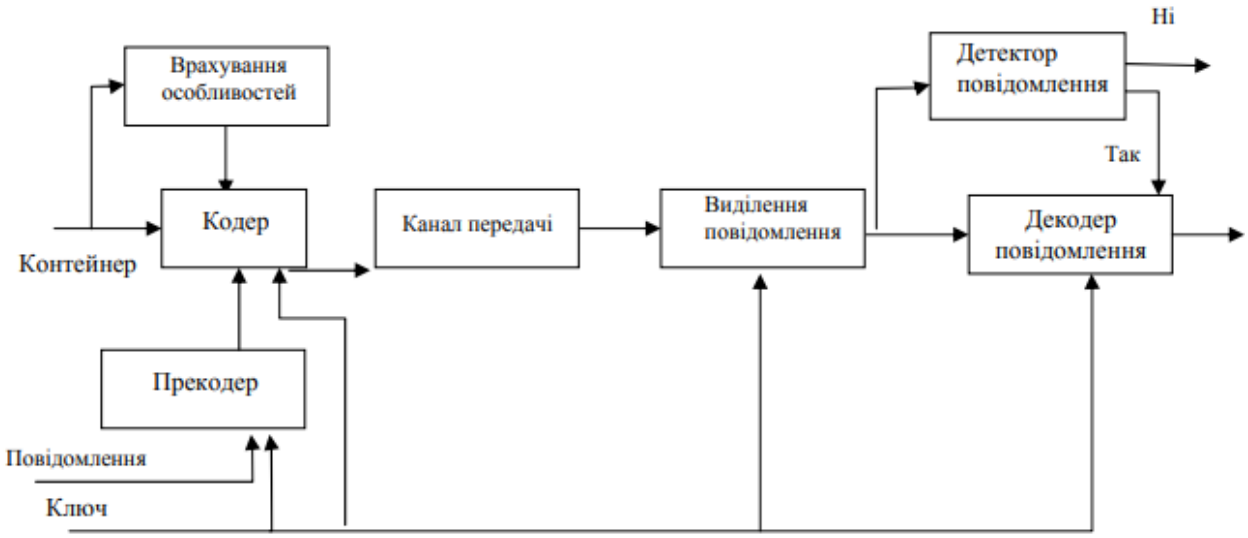

Рисунок 1.1.1 Опис стегосистеми

Вбудовування повідомлень в контейнер проходить з використанням спеціального стегоключа. Ключ – псевдовипадкова послідовність біт, яку створює генератор, що задовольняє певним вимогам (криптографічно безпечний генератор). Цей ключ визначає порядок вбудовування повідомлення в контейнер. В якості основи генератора може використовуватися, наприклад, лінійний рекурентний регістр. Тоді адресатам для забезпечення зв’язку може повідомлятися початкове значення цього регістра [3].

В залежності від кількості рівнів захисту інформації в стегосистемі може бути один або декілька стегоключів. Враховуючи всю різноманітність стеганофонічних систем, їх можна звести до 4 основних типів:

- безключеві стегосистеми;

- системи з таємним ключем;

- системи з відкритим ключем;

- змішані стегосистеми [3].

Стеганографічної системою називається п'ятірка об'єктів (M, K, C, E,

F), де

M - безліч повідомлень;

K - безліч ключів;

C - безліч контейнерів;

$E: M \times C \times K \rightarrow C$,

E - функція вбудовування (впровадження);

$F: C \times K \rightarrow M$,

F - функція витягання;

і при цьому для будь-яких $\mathbf{m} \in M$, $c \in C$, $k \in K$ виконується рівність

$F(E(\mathbf{m}, c, k), k) = \mathbf{m}$.

Так само як і в криптографії, в стеганографії передбачається виконання правила Керкгоффса, суть якого полягає в тому, що стійкість (або надійність) стеганографічної системи визначається лише секретністю ключа. Іншими словами, оцінка якості стеганографічної системи повинна проводитися за умови, що про дану стеганографічної системі відомо все, крім використаного ключа.

У роботах по стеганографії виділяють два основних типи контейнера: потоковий і фіксований.

Потоковий контейнер являє собою послідовність символів (байтів, бітів і т. п.), в яку повідомлення вбудовується в реальному масштабі часу. Тому заздалегідь невідомо, чи вистачить розмірів контейнера для передачі всього повідомлення. В один контейнер великого розміру може бути вбудовано кілька повідомлень. Інтервали між вбудованими бітами визначаються генератором псевдовипадкової послідовності. Основна проблема використання потокових контейнерів полягає в складності організації синхронізації - визначення початку і кінця послідовності. Але, з

іншого боку, з точки зору забезпечення скритності передачі, складність організації синхронізації є перевагою. Не знаючи секретного ключа, не можна не тільки дізнатися про зміст прихованої передачі, але і про сам факт її існування.

У фіксованому контейнері його розміри і характеристики заздалегідь відомі. Це дозволяє виконувати вбудовування даних, що підлягають приховуванню, оптимальним (в певному сенсі) чином.

Є й інша класифікація контейнерів. Контейнери можуть бути обраними, випадковими або нав'язаними. Обраний контейнер залежить від вбудованого повідомлення, а в граничному випадку є його функцією. Для стеганографії переважний саме такий тип контейнера. Нав'язаний контейнер з'являється, коли особа, яка надає контейнер, підозрює про можливе приховуване листування і бажає запобігти її. На практиці ж найчастіше мають справу з випадковим контейнером.

Приховування інформації, яка має певний обсяг, висуває відповідні вимоги і до розміру контейнера, який повинен для забезпечення достатнього рівня прихованості істотно перевищувати розмір вбудованих даних.

Перед вбудовуванням повідомлення в контейнер його необхідно перетворити в зручний для впровадження вид. Крім того, перед вбудовуванням в контейнер, для підвищення захищеності секретної інформації, останню можна зашифрувати чинним сертифікованим криптографічним методом.

У багатьох випадках також бажана стійкість отриманого стеганоповідомлення до спотворень (в тому числі і зловмисним).

В процесі передачі, аудіо-, відеофайли або будь-які інші цифрові сигнали, які використовуються в якості контейнера, можуть піддаватися різним трансформаціям (в тому числі з використанням алгоритмів з втратою даних), таким як: зміна обсягу, перетворення в інший формат і т. п., тому для збереження цілісності вбудованого повідомлення може знадобитися

завадостійке кодування.

Особливо відзначимо наступні два принципи синтезу і використання стеганографічних систем:

▫ визначальним в стеганографічної методі захисту інформації є спосіб вбудовування (вилучення) елементів повідомлення в контейнер;

▫ вбудовування елементів повідомлення в контейнер має здійснюватися з урахуванням тих перетворень контейнера, яким останній буде піддаватися в процесі обробки і передачі (наприклад, перетворень, що застосовуються при стисканні).

В даний час в комп'ютерній стеганографії найбільш поширеними типами контейнерів є представлені в цифровій формі зображення, звукові дані і відеопослідовності, які узагальнено називають цифровими мультимедійними сигналами. Дана обставина пояснюється тим, що мультимедійні контейнери вже за технологією свого отримання (створення) мають шумову складову, яку досить ефективно можна використовувати для маскування вбудованого повідомлення.

Зрозуміло, що найкращий контейнер - це той, наявність або передача якого в комп'ютерних мережах або мережах зв'язку більш загального профілю є типовою подією і не викликає підозр. Тому при виборі виду або типу цифрових образів мультимедійних сигналів, використовуваних як контейнери, недостатньо вимагати тільки того, щоб вони дозволяли зробити розробку стійкою до атак стеганографічної системи. Вкрай важливо ще, щоб вони були досить поширеними і широко використовувалися в практичних додатках. У зв'язку з цим доречно зауважити, що в кінці 20-го століття, з появою недорогих і потужних персональних комп'ютерів, почалася розробка і масове впровадження всіляких мультимедійних додатків і програм, в яких використовуються тексти, зображення, анімовані фрагменти і звук. Всю цю різнорідну цифрову інформацію стало можливим зберігати в комп'ютері, відображати, редагувати, програвати і передавати по каналах зв'язку. Вимоги по пам'яті для зберігання мультимедійної інформації стали критичними.

Тому в 90-х роках минулого століття проблема стиснення мультимедійної інформації стала вельми актуальною і привернула увагу багатьох дослідників і дослідницьких груп. Було розроблено велику кількість алгоритмів для стиснення (з втратами) звукових (MP3, AAC, Ogg Vorbis) [1], графічних (JPEG, JPEG 2000) [2] і відео (MPEG-2, MPEG-4, H.263) файлів, в результаті чого зберігання, обробка та передача стислих мультимедійних файлів різних форматів стали такими ж поширеними явищами всесвітньої «павутини» Інтернет, як зберігання, обробка і передача архівованих текстових файлів. Тому мультимедійні файли і їх різні стислі комп'ютерні форми подання викликають безсумнівний інтерес в якості стеганографічних контейнерів.

Поряд з факторами широкої поширеності мультимедійних файлів і відсутністю принципових складнощів при створенні їх стислих образів різних форматів, важливу роль у виборі даних об'єктів в якості контейнерів стеганографічних систем грають і такі причини, зумовлені як властивостями, властивими мультимедійним файлам і зорової та слухової систем людини, так і станом розвитку сфери інформаційних технологій:

- відносно великий обсяг цифрового уявлення мультимедійного сигналу, що дозволяє вбудовувати повідомлення значного обсягу або ж підвищувати стійкість цього вбудовування;
- можливість створення як фіксованих, так і потокових контейнерів;
- наявність в більшості реальних мультимедійних сигналів областей, що мають шумову структуру і які найкраще підходять для вбудовування інформації;
- слабка чутливість зорової та слухової систем людини до незначних змін відповідних компонент мультимедійного сигналу і невеликих їх спотворень;
- велика кількість загальнодоступних (відкритих) і стандартів і алгоритмів цифрової обробки, що набули широкого поширення

мультимедійних сигналів.

Впровадження в стеганографічний контейнер повідомлення, що підлягає приховуванню, як правило, супроводжується певними спотвореннями контейнера. Ці спотворення можуть стати демаскуючою ознакою, що призводить до непридатності відповідної стеганографічної системи для використання з метою безпечного зберігання і передачі інформації. Одним з можливих варіантів вирішення даної проблеми може служити підхід, що полягає у використанні в якості стеганографічних контейнерів стислих мультимедійних файлів, отриманих шляхом застосування до оцифрованих мультимедійних сигналів стандартних алгоритмів стиснення з втратою інформації, існуючих, як вище було зазначено, в досить великій кількості.

1.1.1. Класифікація стеганографічних систем.

За аналогією з криптографічними системами, в стеганографії розрізняють системи з секретним ключем і системи з відкритим ключем.

У стеганографічній системі з секретним ключем використовується один ключ, який повинен бути заздалегідь відомий абонентам до початку прихованого обміну секретними повідомленнями або пересланий по захищеному каналу.

У стегосистеми з відкритим ключем для вбудовування і вилучення таємного повідомлення використовуються різні ключі, причому вивести один ключ з іншого за допомогою обчислень неможливо. Один з ключів (відкритий) може передаватися вільно по незахищеному каналу зв'язку, а другий, секретний ключ, - по захищеному каналу. Дана схема добре працює при взаємній недовірі відправника і одержувача.

З огляду на все різноманіття стеганографічних систем, зведемо їх до наступних типів: безключові стегосистеми, системи з секретним ключем, системам з відкритим ключем і змішані стегосистеми.

Безключові стегосистеми. Для функціонування таких стегосістем не потрібно ніяких додаткових даних у вигляді стегоключа крім алгоритму

стеганографічного перетворення.

Безпека безключевої стегосистеми заснована на секретності використовуваних стеганографічних перетворень. Це суперечить основному принципу Керкхоффа для систем захисту інформації. Дійсно, якщо припустити, що противник знає алгоритми, які використовуються для прихованої передачі інформації, то він здатний витягти будь-яку приховану інформацію з перехоплених стеганограмм.

Найчастіше для підвищення безпеки безключової системи, перед початком процесу стеганографічного приховування попередньо виконується шифрування інформації, що приховується. Ясно, що такий підхід збільшує захищеність всього процесу зв'язку, оскільки це ускладнює виявлення прихованого повідомлення.

Стегосистеми з секретним ключем. Дотримуючись закону Керкхоффа, безпека системи повинна грунтуватися на деякій секретній інформації, без знання якої не можна витягти з контейнера секретну інформацію. У стегосистемах така інформація називається стегоключем. Відправник, вставляючи секретне повідомлення в обраний контейнер c, використовує секретний стегоключ k. Якщо використовуваний в стеганографічному перетворенні ключ k відомий одержувачу, то він зможе витягти приховане повідомлення з контейнера. Без знання такого ключа будь-який інший користувач цього зробити не зможе.

Даний тип стегосистем передбачає наявність безпечного каналу для обміну стегоключа.

Іноді стегоключ k обчислюють за допомогою секретної хеш-функції Hash, використовуючи деякі характерні особливості контейнера: k = Hash (особливості контейнера).

У деяких алгоритмах при добуванні прихованої інформації додатково потрібні відомості про вихідний контейнері або деякі інші дані, які відсутні в стеганограмі.

Стегосистеми з відкритим ключем. Ці системи не потребують

додаткового каналу ключового обміну. Для їх функціонування необхідно мати два стегоключа: один секретний, який користувач повинен зберігати в таємниці, а другий - відкритий, який зберігається в доступному для всіх місці. При цьому відкритий ключ використовується в процесі приховування інформації, а секретний - для її вилучення.

Простим засобом реалізації подібних стегосистем є використання криптосистем з відкритим ключем. Стегосистеми з відкритими ключами використовують той факт, що функція витягання прихованої інформації може бути застосовна до будь-якого контейнера незалежно від того, чи знаходиться в ньому приховане повідомлення чи ні. Якщо в контейнері відсутнє приховане повідомлення, то завжди буде відновлюватися деяка випадкова послідовність. Якщо ця послідовність статистично не відрізняється від шифртекста криптосистеми з відкритим ключем, тоді в безпечній стегосистемі можна приховувати отриманий таким чином шифртекст, а не відкритий.

Змішані стегосистеми. У більшості додатків кращими є безключові стегосистеми; хоча такі системи можуть бути відразу скомпрометовані в разі, якщо порушник дізнається застосовуване стеганографічне перетворення. У зв'язку з цим в безключовій стегосистемі часто використовують особливості криптографічних систем з відкритим і (або) секретним ключем. Розглянемо один такий приклад.

Для обміну секретними ключами стегосистеми введемо поняття протоколу, реалізованого на основі криптосистеми з відкритими ключами. Спочатку Аліса генерує випадкову пару відкритого і секретного ключа, а потім передає відкритий ключ Бобу по таємного каналу, створеного безключовою системою. Ні Боб, ні Вілі, що ведуть спостереження за каналом, не можуть визначити, яка інформація передавалася в прихованому каналі: ключ або ж випадкові біти. Однак Боб може запідозрити, що стеганограмма від Аліси може містити її відкритий ключ і постарається його виділити. Після цього він шифрує за допомогою виділеного ключа секретний стегоключ і,

проводить приховування результату шифрування в контейнер і його передачу Алісі. Вілі може спробувати витягти секретну інформацію з стеганограмми, але отримає тільки випадковий шифртекст. Аліса витягує з стеганограмми приховану криптограму і розшифровує її своїм секретним ключем. Таким чином, сторони обмінялися секретним стегоключа k для спільного використання.

1.2. Аналіз стеганографічних алгоритмів

Всі алгоритми вбудовування прихованої інформації можна розділити на кілька підгруп:

• Ті, що працюють з самим цифровим сигналом. Наприклад, метод LSB.

• «вбудовування» прихованої інформації. В даному випадку відбувається накладення приховуваного зображення (звуку, іноді тексту) поверх оригіналу. Часто використовується для вбудовування цифрового водяного знаку (далі ЦВЗ).

• Використання особливостей форматів файлів. Сюди можна віднести запис інформації в метадані або в різні інші зарезервовані поля файлу.

За способом вбудовування інформації стегоалгоритму можна розділити на лінійні (адитивні), нелінійні та інші. Алгоритми адитивного впровадження інформації полягають в лінійній модифікації вихідного зображення, а її витяг в декодері проводиться кореляційними методами. При цьому ЦВЗ зазвичай складається із зображенням-контейнером, або «вбудовується» (fusion) в нього. У нелінійних методах вбудовування інформації використовується скалярне або векторне квантування. Серед інших методів певний інтерес представляють методи, які використовують ідеї фрактального кодування зображень.

- Цифровий водяний знак - це сукупність невидимих міток, які носять унікальний цифровий код. У ньому і зашифровані різні дані: авторські права, ідентифікаційний номер, керуюча інформація. Найбільш зручними

для захисту з його допомогою є нерухомі зображення, аудіо та відео файли.

Основні вимоги, що пред'являються до цифрових водяних знаків: надійність і стійкість до спотворень. Вони мають невеликий обсяг, але для виконання зазначених вище вимог, при їх встановленні використовуються більш складні методи, ніж для вбудовування звичайних заголовків або повідомлень. Такі завдання виконують спеціальні стегосистеми.

Перед переміщенням даного знака в контейнер, водяний знак потрібно перетворити до потрібного виду. Первинну обробку часто виробляють з використанням ключа - для підвищення секретності. Потім водяний знак «вкладається» в контейнер (наприклад, шляхом зміни молодших значущих бітів). Тут використовуються особливості сприйняття зображень людиною, адже відомо, що зображення мають величезну псіховізуальну надмірність. Очі людини подібні низькочастотному фільтру, який пропускає дрібні елементи зображення. Найменш помітні спотворення в високочастотній області зображень. Впровадження цифрового водяного знаку також має враховувати властивості сприйняття людини.

У багатьох стегосистемах для запису і зчитування ЦВЗ використовується ключ. Він може призначатися для обмеженого кола користувачів або ж бути секретним. Не існує таких стегосистем, в яких би при зчитуванні водяного знака була потрібна інша інформація, ніж при його запису. У стегодетекторі відбувається виявлення ЦВЗ в захищеному їм файлі, який, можливо, міг бути змінений. Ці зміни можуть бути пов'язані з впливами помилок в каналі зв'язку, або навмисними перешкодами. При цьому завдання виявлення і зчитування стегосоповідомлення вже не представляє складності, але не враховує двох факторів: невипадковість сигналу контейнера і запитів щодо збереження його якості. Облік цих параметрів дозволить будувати більш якісні стегосистеми. Для виявлення факту існування водяного знака і його зчитування використовуються спеціальні пристрої - стегодетектори. Для винесення рішення про наявність чи відсутність водяного знака використовують, наприклад, відстань по

Хеммінгу, взаімокореляцію між отриманим сигналом і його оригіналом. У разі відсутності вихідного сигналу в справу вступають більш витончені статистичні методи, які засновані на побудові моделей досліджуваного класу сигналів.

- Метод LSB (найменши значущий біт) - суть цього методу полягає в заміні останніх значущих бітів в контейнері (зображення, аудіо або відеозапису) на біти приховуваного повідомлення. Різниця між порожнім і заповненим контейнерами повинна бути не відчутна для органів сприйняття людини.

Суть методу полягає в наступному: Припустимо, є 8-бітове зображення в градаціях сірого. 00h (00000000b) позначає чорний колір, FFh (11111111b) - білий. Усього є 256 градацій (28). Також припустимо, що повідомлення складається з 1 байта - наприклад, 01101011b. При використанні 2 молодших біт в описах пікселів, нам буде потрібно 4 пікселі. Припустимо, вони чорного кольору. Тоді пікселі, що містять приховане повідомлення, будуть виглядати наступним чином: 00000001 00000010 00000010 00000011. Тоді колір пікселів зміниться: першого - на 1/255, другого і третього - на 2/255 і четвертого - на 3/255. Такі градації, мало того що непомітні для людини, можуть взагалі не відобразитися при використанні низькоякісних пристроїв виведення.

Для простоти опису можна розглянути принцип роботи цього методу на прикладі 24-бітного реєстрового RGB-зображення. Одна точка зображення в цьому форматі кодується трьома байтами, кожен з яких відповідає за інтенсивність одного з трьох складових квітів.

- Ехо-методи застосовуються в цифровій аудіостеганографії і використовують нерівномірні проміжки між ехо-сигналами для кодування послідовності значень. При накладенні ряду обмежень дотримується умова непомітності для людського сприйняття. Характеризується трьома параметрами: початкової амплітудою, ступенем загасання, затримкою. При

досягненні певного порогу між сигналом і ехом вони змішуються. У цій точці людське вухо не може вже відрізнити ці два сигнали. Наявність цієї точки складно визначити, і вона залежить від якості вихідного запису, слухача. Найчастіше використовується затримка близько 1/1000, що цілком прийнятно для більшості записів і слухачів. Для позначення логічного нуля і одиниці використовується дві різні затримки. Вони обидві мають бути менше, ніж поріг чутливості вуха слухача до одержуваного еха. Якщо з вихідного сигналу можна виділити тільки одне ехо, то може бути закодований тільки один біт секретної інформації. Отже, перед початком процесу кодування вихідний сигнал розбивається на блоки. Після виконання кодування блоки об'єднуються разом, щоб утворити остаточний вихідний сигнал.

Головні переваги даного методу - це висока швидкість передачі даних в порівнянні з іншими методами. Також ехо-методи стійкі до амплітудних і частотних атак, але нестійкі до атак за часом.

1.3. Висновки

У першому розділі розглянуто актуальні методи приховування даних в мультимедійних сигналах, які діляться на дві узагальнені групи: методи приховування в тимчасовій або просторовій області та методи приховування в частотній області. Виявлено, що методи приховування в частотній області мультимедійного сигналу є кращими з точки зору забезпечення стійкості стеганографічного алгоритму. Відмічено, що вбудовування елементів повідомлення в контейнер має здійснюватися з урахуванням тих перетворень контейнера, яким останній буде піддаватися в процесі обробки і передачі (наприклад, перетворень, що застосовуються при стисканні). Стеганографічна система повинна забезпечувати дворівневу архітектуру захисту інформації, що виконує дві основні задачі: приховування самого факту наявності інформації, що захищається в стеганоконтейнері та запобігання несанкціонованого доступу до неї.

Виявлено, що різниця значень відносини сигнал / шум квантування (метрика SQNR) мультимедійного сигналу в випадках порожнього і заповненого контейнерів є зростаючою функцією від числа змінених елементів контейнера при впровадженні в нього повідомлення, що підлягає приховуванню (рішення другої задачи даного дослідження). У зв'язку з цим актуально рішення задачі зниження числа змінених елементів контейнера при впровадженні в нього повідомлення, що підлягає приховуванню. Зазначене рішення може бути здійснено шляхом розробки і застосування відповідних стеганографічних алгоритмів, що враховують ймовірносно-статистичні характеристики приховуваного повідомлення. Це, відповідно до мети дипломної роботи, має сприяти підвищенню ефективності стеганографічної захисту інформації.

Розділ 2. Оптимальний і субоптимальний ентропійні стеганографічні алгоритми.

2.1. Первинні поняття і елементи ентропійного підходу в стеганографії.

Вбудовування (впровадження) повідомлення m (представленого у

вигляді двійкової кінцевої послідовності довжини l) в стеганографічний контейнер з n пікселів (в разі зображень або відеосигналів) або n семплів (в разі аудіосигналів) після або на етапі виконання квантування стандартним алгоритмом стиснення цифрового мультимедійного сигналу здійснюється відповідно до правила.

Додатково будемо вважати виконаним ще наступне:

- для впровадження в контейнер одного елемента $m_s \in \{0; 1\}$, (де $s \in \{1,2 \dots, l\}$) двійкового повідомлення $m = (m_1, m_2, \dots, m_l)$ використовується τ елементів контейнера (до обговорення порядку вибору значень параметра τ повернемося пізніше) і, отже, для приховування всього повідомлення m використовується $n = \tau \cdot l$ елементів контейнера;

- елементи підмножини функцій $\{H_k \mid k \in K\}$ що є структурною компонентою стеганографічного алгоритму з простором ключів K, влаштовані таким чином, що для будь-якого $k \in K$ існує набір функцій, такий що:

$$H_{kj} : \underbrace{Z \times Z \times \dots \times Z}_{\tau \text{ разів}} \to \{0; 1\}$$

Для будь-якого числа $j \in$ і справедлива рівність

$$m = H_k(g_1, g_2, \dots, g_n) =$$

$$(H_{k1}(\tilde{g}_{i_1}, \tilde{g}_{i_2}, \dots, \tilde{g}_{i_\tau}), H_{k2}(\tilde{g}_{i_{\tau+1}}, \tilde{g}_{i_{\tau+2}}, \dots, \tilde{g}_{i_{2\tau}}), \dots, H_{kl}(\tilde{g}_{i_{(l-1)*\tau+1}}, \tilde{g}_{i_{(l-1)*\tau+2}}, \dots, \tilde{g}_{i_{l*\tau}}))$$

де

$$(H_{k1}(\tilde{g}_{i_1}, \tilde{g}_{i_2}, \dots, \tilde{g}_{i_\tau}) ==$$
$$m_1, H_{k2}(\tilde{g}_{i_{\tau+1}}, \tilde{g}_{i_{\tau+2}}, \dots, \tilde{g}_{i_{2\tau}}), \dots, H_{kl}(\tilde{g}_{i_{(l-1)*\tau+1}}, \tilde{g}_{i_{(l-1)*\tau+2}}, \dots, \tilde{g}_{i_{l*\tau}}))$$

і $\tilde{g}_{i_1}, \tilde{g}_{i_2}, \dots, \tilde{g}_{i_n}$ – деяка перестановка безлічі елементів $(g_1, g_2, \dots, g_n)$; $(i_1, i_2, \dots, i_n)$ – перестановка множини чисел (1,2,…,n), яка залежить від обраного ключа k∈K;

- для будь-якого k∈K та для будь-якого j∈N існує булева функція

$f_{kj}(x_1, x_2, \dots, x_\tau)$ від τ перемінних $x_1, x_2, \dots, x_\tau$, така, що для будь-яких чисел $a_1, a_2, \dots, a_\tau$ справедлива рівність

$$H_{kj}(a_1, a_2, \dots, a_\tau) = f_{kj}(b_1, b_2, \dots, b_\tau),$$

де

$$b_r = \begin{cases} 0, \text{якщо } a_r - \text{парне число;} \\ 1 \text{ якщо } a_r - \text{непарне число;} \end{cases}$$

$r \in \{1,2, \dots, \tau\}$.

Таким чином, представлений стеганографічний алгоритм повністю визначається завданням свого набору булевих функцій f_{kj}, який назвемо набором булевих функцій впровадження(добування) інформації. Різні варіанти стеганографічного алгоритму різняться тільки наборами булевих функцій впровадження(добування) інформації. При цьому під алгоритмом з набором булевих функцій впровадження інформації $\{f_{kj} | k \in K, j \in N\}$ можна розуміти також набір алгоритмів, кожен з яких має свою одну функцію впровадження інформації f_{kj}(де $(k \in K, j \in N)$ і слугує для впровадження і вилучення повідомлення з одного біта.

Зрозуміло, що при виконанні вищевказаних додаткових умов число змін, що вносяться до контейнера в процесі впровадження в нього повідомлення m, дорівнює арифметичній сумі чисел змін, що вносяться у контейнер при впровадженні кожного біта даного повідомлення. Якщо, наприклад [29], припустити, що для будь-якого $k \in K$ та будь-якого $j \in N$

$$H_{kj}(a_1, a_2, \dots, a_\tau) = \begin{cases} 0, \text{якщо } \sum_{r=1}^{\tau} a_r - \text{парне число} \\ 1, \text{якщо } \sum_{r=1}^{\tau} a_r - \text{непарне число} \end{cases}$$

то для мулевої функції $f_{kj}(x_1, x_2, \dots, x_\tau)$ справедлива рівність

$$f_{kj}(x_1, x_2, \dots, x_\tau) = x_1 \oplus x_2 \oplus \dots \oplus x_\tau,$$

де $\oplus$ - знак операції додавання за модулем 2 [29]. В цьому випадку для вбудовування в контейнер елемента $m_s \in \{0; 1\}$ (де $s \in \{1,2, \dots, l\}$) повідомлення m за ключем генеруються відповідні номери елементів та вибираються самі елементи $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ контейнера. При цьому достатня заміна не більш ніж одного з них для отримання елементів $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$, таких, що

$$H_{ks}\left(\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}\right) = m_s.$$

Дійсно, якщо $m_s = 0$ та сума елементів $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$ дорівнює парному числу, то під час вбудовування не відбувається ніяких змін, таким чином набір $\left(\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}\right)$ збігається з набором $\left(g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}\right)$ з дотриманням порядку. Якщо ж сума елементів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ дорівнює непарному числу, то при вбудовуванні відбувається зміна одного з елементів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ шляхом додавання або віднімання одиниці в сторону, протилежну напрямку операції округлення стандартом стиску при отриманні даного елемента. В результаті сума отриманих елементів $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$ стає рівною парному числу.

Аналогічно, якщо $m_s = 1$ та сума елементів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ дорівнює непарному числу, то при вбудовуванні не відбувається ніяких змін, таким чином набір $\left(\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}\right)$ співпадає з $\left(g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}\right)$ з урахуванням порядка. Якщо сума елементів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ дорівнює парному числу, то при

вбудовуванні відбувається зміна одного з елементів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ шляхом додавання або віднімання одиниці в напрямок, протилежний напрямку операції округлення стандартом стиску при отриманні даного елемента. В результаті сума отриманих елементів $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$ стає рівною непарному числу.

Зрозуміло, що при вбудовуванні та вилученні елемента m_s можливо функцію H_{ks} замінити на булеву функцію $f_{ks} = x_1 \oplus x_2 \oplus \dots \oplus x_\tau$, використовуя замість наборів $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ та $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$ елементів контейнера, відповідно, їх набори бітів парності $b_{i_{(s-1)*\tau+1}}, b_{i_{(s-1)*\tau+2}}, \dots, b_{i_{s*\tau}}$ та $\tilde{b}_{i_{(s-1)*\tau+1}}, \tilde{b}_{i_{(s-1)*\tau+2}}, \dots, \tilde{b}_{i_{s*\tau}}$

Дана обставина дає можливість обчислити важливі з теоретичної та практичної точок зору характеристики випадкової величини $\delta(m)$, що дорівнює мінімальному числу змін, що вносяться до контейнера в результаті впровадження в нього повідомлення m.

Дійсно, припустимо, що біти парності елементів порожнього контейнера можуть бути представлені як результат роботи джерела, який генерує 0 і 1 потактно за схемою незалежних випробувань з однією і тією ж ймовірністю 0,5. А повідомлення m – результат роботи джерела, який генерує 0 і 1 потактно за схемою незалежних випробувань з вірогідністю p і q, відповідно, де $p \geq 0, q \geq 0, p + q = 1$.

Тоді для математичного очікування $E(\delta(m))$ випадкової величини $\delta(m)$ справедливий ланцюжок рівностей

$$E\big(\delta(m)\big) = \sum_{s=1}^{l} E\big(\delta(m_s)\big) = \sum_{s=1}^{l} (0 * P(\delta(m_s) = 0) + 1 * P(\delta(m_s) = 1)) =$$

$$= \sum_{s=1}^{l} 0 * P(\delta(m_s) = 0) * P(f_{ks}\left(b_{i_{(s-1)*\tau+1}}, \dots, b_{i_{s*\tau}}\right) = 0) + P(m_s = 1) * P *$$

$$* (f_{ks}\left(b_{i_{(s-1)*\tau+1}}, \dots, b_{i_{s*\tau}}\right) = 1)) + \sum_{s=1}^{l} 1 * P(\delta(m_s) = 0) * P(f_{ks} *$$

$$* \left(b_{i_{(s-1)*\tau+1}}, \dots, b_{i_{s*\tau}}\right) = 1) + P(m_s = 1) * P(f_{ks}\left(b_{i_{(s-1)*\tau+1}}, \dots, b_{i_{s*\tau}}\right) = 0)) =$$

$$= \sum_{s=1}^{l} 0 * (p * 0{,}5 + q * 0{,}5) + \sum_{s=1}^{l} 1 * (p * 0{,}5 + q * 0{,}5) = 0{,}5l,$$

де $E\big(\delta(m_s)\big)$ – математичне очікування випадкової величини $\delta(m_s)$, яка дорівнює мінімальній кількості змін, що вносяться до контейнера при вбудовуванні елемента m_s повідомлення m, . Таким чином, при вбудовуванні в контейнер одного біта повідомлення, який необхідно приховати, допускається в середньому 0,5 змін.

Так само, для дисперсії $D\big(\delta(m)\big)$ випадкової величини $\delta(m)$ є справедливим ланцюг рівностей

$$D\big(\delta(m)\big) = \sum_{s=1}^{l} D\big(\delta(m_s)\big) = \sum_{s=1}^{l} (E(\delta(m_s))^2 - (E(\delta(m_s)))^2\ ) =$$

$$= \sum_{s=1}^{l} 0^2 * (p * 0{,}5 + q * 0{,}5) + 1^2 * (p * 0{,}5 + q * 0{,}5) - (E(\delta(m_s)))^2\ ) =$$

$$= (0{,}5 - 0{,}25) * l = 0{,}25l$$

Якщо для будь-якого $k \in K$ і для будь-якого $j \in N$ булева функція f_{kj} задається через рівність (2.1.2), то значення математичного очікування $\mathrm{E}\big(\delta(m)\big)$ випадкової величини $\big(\delta(m)\big)$, яка є рівною мінімальному числу змін, внесених у контейнер (тобто $\delta(m)$ – число змінених елементів контейнера) в результаті впровадження в нього довільного двійкового повідомлення m, не залежить від значень чисел p і q, тобто не залежить від імовірнісних параметрів джерела повідомлень, що підлягають приховуванню. Однак, очевидною є

можливість зменшення значення $\mathrm{E}(\delta(m))$ шляхом обліку значень імовірнісних параметрів p і q джерела повідомлень за рахунок вибору відповідних булевих функцій $f_{kj}, k \in K,\ j \in N$. Так наприклад, якщо $\mathrm{P}(m_s = 0) = p > q = P(m_s = 1)$, то інтуїтивно можна припустити, що значення $\mathrm{E}(\delta(m))$ може бути менше, ніж значення 0,5l отримане в (2.1.3), якщо булеві функції $f_{kj}, k \in K, s \in \{1,2,\dots,l\}$ будуть мати значення 0 на більшій кількості довічних наборів, ніж значення 1. Однак це припущення вимагає відповідного дослідження для свого обґрунтування, так як при такому нерівномірному розподілі довічних наборів за значеннями 0 і 1 булевих функцій $f_{kj}, k \in K, s \in \{1,2,\dots,l\}$, може виявитися недостатньою зміною не більше ніж одного елемента для $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ отримання набору $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$.

Можуть мати місце ситуації, що вимагають зміни двох і більше елементів. У зв'язку з цим має сенс виписати загальний вираз для математичного очікування $\mathrm{E}(\delta(m))$ випадкової величини $\delta(m)$, що дорівнює мінімальному числу змін елементів контейнера при впровадженні в τ елементів контейнера повідомлення з одного біта, згенерованого джерелом повідомлень з параметрами p і q. Але перед цим звернемо увагу на те, що вище $\delta(m)$ була визначена як випадкова величина, що дорівнює мінімальному числу змін елементів контейнера при впровадженні в τ елементів контейнера повідомлення m з одного біта, згенерованого джерелом повідомлень з параметрами p і q. У цьому визначенні словосполучення «мінімальне число змін» можна замінити на словосполучення «число змін», тобто прибрати слово «мінімальне», якщо покласти, що завжди при необхідності перехід від набору $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$ до необхідного набора $\tilde{g}_{i_{(s-1)*\tau+1}}, \tilde{g}_{i_{(s-1)*\tau+2}}, \dots, \tilde{g}_{i_{s*\tau}}$ здійснюється за принципом мінімума числа змін в наборі $g_{i_{(s-1)*\tau+1}}, g_{i_{(s-1)*\tau+2}}, \dots, g_{i_{s*\tau}}$. В роботі цей принцип всюди неухильно дотримується і, відповідно $\delta(m)$ - це величина випадкова, що дорівнює числу

змінених елементів контейнера при впровадженні в контейнер повідомлення m з одного біта. У зв’язку з вищесказаним з’являється можливість порівняння між собою елементів безлічі стеганографічних алгоритмів з однією функцією впровадження інформації, призначених для впровадження одно бітових повідомлень, за значенням математичного очікування випадкової величини, яка дорівнює числу змінених елементів контейнера при впровадженні в нього повідомлення m з одного біта. Тому будемо вважати та говорити, що стеганографічний алгоритм А більш ефективний, ніж алгоритм В, якщо математичне сподівання числа змінених елементів контейнера при впровадженні одного біта повідомлення для алгоритму А менше, ніж для алгоритму В.

Зауважимо, що, повертаючись до мети магістерської роботи можна сказати, що для її досягнення необхідна розробка стеганографічних алгоритмів, більш ефективних, ніж відомі, або виявлення найбільш ефективних серед відомих шляхом їх аналізу. Для спрощення запису далі в цьому розділі і всюди в даній роботі, де мова йде про впровадження повідомлення з одного біта, індекси булевої функції f_{ks} будемо опускати і писати просто f і покладемо $\tau = n$. При фіксованій функції впровадження $f \in F_2^n$ (де F_2^n- безліч всіх двійкових функцій від n змінних) випадкову величину $\delta(m)$ будемо позначати $\delta_f(m)$, підкреслюючи тим самим її зв'язок з функцією f.

Позначимо через V_2^n безліч двійкових векторів з n координатами ($n \in N$), тобто

$V_2^n = \{(\alpha_1, \alpha_2, \dots, \alpha_n) | \alpha_i \in \{0; 1\}, i = \overline{1, n}\}$;

Припустимо

$A_f = \{\mathrm{v} \in \mathrm{V}_2^{\mathrm{n}} \mid \mathrm{f(v)} = 1\}$ і $\mathrm{B_f} = \{\mathrm{v} \in \mathrm{V}_2^{\mathrm{n}} | \mathrm{f(v)} = 0\}$.

Визначимося, що під інвертуванням координати довільного вектора з

безлічі V_2^n будемо розуміти операцію заміни нульового значення координати на 1, а одиничного значення координати на 0. З урахуванням цього можна вказати, що безлічі A_f і B_f містять відповідно системи підмножин $\{A_{1f}, \dots, A_{nf}\}$ і $\{B_{lf}, \dots, B_{nf}\}$, такі, що:

1) при $i \neq j$ справедливі рівності $A_{if} \cap A_{if} = \emptyset$ і $B_{if} \cap B_{if} = \emptyset$;

2) $A_{1f} \cup \dots \cup A_{nf} = A_f$ і $B_{1f} \cup \dots \cup B_{nf} = B_f$

3) Для будь-якого $k \in \{1,2, \dots, n\}$ безліч векторів A_{kf} таке, що якщо $A_{kf} = \emptyset$, то кожен вектор $a^{(k)} \in A_{kf}$ володіє такою властивістю, що при інвертуванні будь-яких його координат в кількості меншій ніж k і незмінності інших координат виходить вектор, що належить множині A_f; але існує набір рівно з k координат вектора $a^{(k)}$, при інвертуванні яких та незмінності інших координат виходить вектор, що належить множині \\

4) Для будь-якого $k \in \{1,2, \dots, n\}$ безліч векторів B_{kf} таке, що якщо $B_{kf} = \emptyset$, то кожен $b^{(k)} \in B_{kf}$ володіє властивістю, що при інвертуванні будь-яких його координат в кількості меншій, ніж k, і незмінності інших координат виходить вектор, що належить множині B_f; але існує набір рівно з k координат вектора $b^{(k)}$, при інвертуванні яких і незмінності інших координат виходить вектор, що належить множині A_f.

Тоді для математичного очікування $\mathrm{E}\left(\delta_f(m)\right)$ випадкової величини $\delta_f(m)$, що дорівнює числу змінених елементів контейнера при впровадженні в n елементів контейнера повідомлення m з одного біта, згенерованого джерелом повідомлень з параметрами p і q, є вірним ланцюг рівностей

$$\mathrm{E}\left(\delta_f(m)\right) = \sum_{i=0}^{n} i \cdot P(\delta_f(m) = i) = \sum_{i=1}^{n} i \cdot (P(m = 0) \cdot$$

$$\cdot P\left(v \in A_{if}\right) + P(m = 1) \cdot P(v \in B_{if})) = p \sum_{i=1}^{n} i \cdot P\left(v \in A_{if}\right) +$$

$$+q \cdot \sum_{i=1}^{n} i \cdot P\left(v \in B_{if}\right) = \frac{p}{2^n} \cdot \sum_{i=1}^{n} i \cdot y_{if} + \frac{q}{2^n} \cdot \sum_{i=1}^{n} i \cdot z_{if},$$

де

$$y_{if} = \left|A_{if}\right|, z_{if} = \left|B_{if}\right|, i = \overline{1,n};$$

v – вектор, що складається з бітів парності елементів контейнера, в які впроваджуються повідомлення.

Стеганографічний алгоритм, в якому процедура впровадження повідомлень в контейнер здійснюється з урахуванням ймовірно статистичних характеристик джерела повідомлень, назвемо ентропійним стеганографічним алгоритмом. Слід зазначити, що тут наявна певна анологія з теорією кодування, де термін ентропія широко використовується в розділі кодування джерела повідомлень. У цій теорії за допомогою поняття ентропії джерела, що відображає його ймовірносно-статистичні характеристики, передбачається найкраще стиснення інформації, тобто, найменше в середньому числі біт, необхідне для подання кодованого повідомлення, згенерованого джерелом. Відповідні процедури кодування називають ентропійними. Прикладом такого кодування є кодування Хаффмана [21].

При фіксованому джерелі повідомлень і фіксованому числі $n \in N$, ентропійний стеганографічний алгоритм назвемо оптимальним ентропійним стеганографічним алгоритмом, якщо при цьому алгоритмі математичне очікування числа змінних елементів контейнера при впровадженні довічного повідомлення m, що складається з одного біта, в n елементів контейнера, дорівнює $\frac{min}{f \in F_2^n} E(\delta_f(m))$, де F_2^n- безліч всіх двійкових функцій від n змінних.

Таким чином, оптимальний стеганографічний алгоритм не менш

ефективний, ніж будь-який інший стеганографічний алгоритм.

Функцію g з безлічі F_2^n всіх довічних функцій від n змінних, для якої справедлива рівність

$$E\left(\delta_g(m)\right) = \frac{min}{f \in F_2^n} E(\delta_f(m))$$

назвемо nq-оптимальною функцією впровадження.

Зауважимо, що в загальному випадку при фіксованих значеннях величин n і q nq-оптимальна функція впровадження може бути визначена неоднозначно.

Відзначимо, що завдання розробки оптимального ентропійного стеганографічного алгоритму при фіксованому джерелі повідомлень є завданням комбінаторної оптимізації [51] в силу кінцівки множин F_2^n (для будь якого $n \in N$, числа елементів в множині F_2^n, тобто потужність безлічі F_2^n, дорівнює 2^{2^n}) і V_2^n (для будь-якого $n \in N$, потужність множини V_2^n дорівнює 2^n).

Оптимальний ентропійний стеганографічний алгоритм впровадження інформації в контейнер може бути отриманий шляхом перебору всіх можливих функцій $f \in F_2^n$ розбиття множини V_2^n на дві підмножини A_f і B_f з визначенням в них чисел $y_{if} = \left|A_{if}\right|, z_{if} = \left|B_{if}\right|, i = \overline{1,n}$, а по ним величини $E(\delta_f(m))$, з метою виявлення такої функції nqoptf $\in F_2^n$ на якій величина $E(\delta_f(m))$, досягає свого мінімального значення, тобто справедлива рівність

$$E\left(\delta_{nq\text{опт}f}(m)\right) = \frac{min}{f \in F_2^n} E(\delta_f(m))$$

Такий перебір функцій $f \in F_2^n$ є громіздкою процедурою, яка застосовується лише для невеликих значень параметра n. З цієї причини має сенс направити зусилля на розробку не обов'язково оптимальних, але прийнятних з практичних позицій ентропійних стеганографічних алгоритмів,

для яких значення математичного очікування числа змінених елементів контейнера при впровадженні довічного повідомлення, що складається з одного біта m, в n елементів контейнера, не перевищують деякої верхньої межі $b \in R$, де R множина дійсних чисел. Такі, в певному сенсі, «задовільні» алгоритми будемо називати субоптимальними. Різні варіанти розроблених субоптимальних алгоритмів можуть потім детально досліджуватися для визначення ступеня їх близькості до оптимального алгоритму та виявлення кращого з них по відношенню до оптимального.

2.2. Оптимальний ентропійний стеганографічний алгоритм

Даний розділ присвячений завданню побудови оптимального ентропійного стеганографічного алгоритму за умови, що впроваджені стеганографічні контейнери повідомлення попередньо кодуються (стискаються) асимптотично оптимальним блоковим рівномірним кодом. Тут вираз «асимптотично оптимальний блоковий рівномірний код» означає, що, по-перше, кодуванню піддаються повідомлення заданої кінцевої довжини (блоки однакової довжини) при її необмеженому збільшенні [63] і, по-друге, всі кодові слова в блоковому коді мають однакову довжину, і його оптимальність розуміється в сенсі мінімальності цієї довжини [33], [56], [64]. З урахуванням даних умов на впроваджувані повідомлення, сформулюємо і доведемо наступне.

Нехай повідомлення, що впроваджуються в стеганографічні контейнери попередньо кодуються (стискаються) асимптотично оптимальним блоковим рівномірним кодом. Тоді для будь-якого і для будь-якого числа q, такого, що $0 < q < 1$, ентропійний стеганографічний алгоритм з функцією впровадження-вилучення

$$g(x_1, x_2, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$$

є оптимальним ентропійним стеганографічним алгоритмом. При цьому середнє число змінюваних елементів контейнера при впровадженні

повідомлення m з одного біта дорівнює

$$E\left(\delta_g(m)\right) = \frac{min}{f \in F_2^n} E\left(\delta_f(m)\right) = 0{,}5\, H(X)$$

де X – випадкова величина приймаюча значення 0 і 1 з ймовірностями p=1-q і q, відповідно $H(X)$ - ентропія випадкової величини X (звана ентропією джерела), $H(X) = p \cdot \log_2 p + q \cdot \log_2 q$.

Для доказу цього припущення звернімося до асимптотичних методів теорії інформації [64], точніше, до деякої якісної інтерпретації теореми Шеннона про кодування при відсутності шуму [15], [21], [33], [48], [56]. Згідно з обмеженнями, прийнятими в даній роботі, займемося розглядом тільки випадку довічного джерела повідомлень, який генерує 0 і 1 потактно за схемою незалежних випробувань з вірогідністю p і q, відповідно, де

$p \geq 0$, $q \geq 0$, p + q = l .

Теорема Шеннона про кодування при відсутності шуму відповідає на наступне питання: які мінімальні ресурси необхідні для того, щоб зберігати інформацію, що отримується з джерела, так, щоб згодом можна було відновити її?

Виявляється, що для зберігання двійкового вектор-рядка довжини l потрібно (в середньому) $l \cdot H(X)$ бітів, де X – випадкова величина, що приймає значення 0 і 1 з ймовірністю p і q; $H(X)$ – ентропія випадкової величини X (ентропія джерела). $H(X) = -(p \cdot \log_2 p + q \cdot \log_2 q)$. Цей результат відомий як теорема Шеннона про кодування при відсутності шуму. У зв'язку з доказом цього корисно зрозуміти основну ідею Шеннона, на якій базується ця теорема. Вона полягає в наступному.

При досить великих значеннях натурального числа l всі повідомлення, тобто двійкові вектори-рядки довжини l, що генеруються джерелом, можна розбити на два класи: перший клас K_{1l} - це клас типових векторів-рядків, що

містять приблизно $l \cdot p$ нулів і приблизно $l \cdot q$ одиниць; другий клас K_{2l} - це клас атипових векторів-рядків, тобто тих векторів-рядків, які не потрапили в клас K_{1l}. Імовірність того, що згенероване джерелом повідомлення довжини l атипове, тобто належить класу K_{2l}, асимптотично (в межі при $l \to \infty$) мала. Таким чином, атипові вектори-рядки генеруються джерелом рідко, на відміну від типових, так як при великих значеннях l з великою ймовірністю частка символів 0 на виході джерела буде дорівнювати p, а частка символів *1* буде дорівнювати q, що узгоджується з визначенням типових векторів-рядків. Імовірність генерації джерелом будь-якого типового вектор-рядка приблизно дорівнює $p^{lp} \cdot q^{lq}$. Значення логарифма за основою 2 від цієї величини дорівнює $-l \cdot H(X)$. Отже, ймовірність генерації джерелом будь-якого одного типового вектор-рядка приблизно дорівнює $2^{-l \cdot H(X)}$. Оскільки повна вірогідність всіх типових вектор-рядків довжини l не може бути більше одиниці, то кількість типових довічних вектор-рядків не може бути більше $2^{l \cdot H(X)}$. Ідея Шеннона полягає в тому, що кодуванню з метою стиснення повинні піддаватися тільки повідомлення з класу K_{1l}. Звідси випливає досить проста схема стиснення даних на виході джерела повідомлень. Якщо джерелом згенеровано повідомлення, що відноситься до класу K_{2l}, то воно вважається помилкою і ігнорується. При великих значеннях числа l це трапляється рідко, як зазначено вище. Якщо ж згенеровано повідомлення з класу K_{1l}, то воно стискається відповідно до попередньо обраної схеми кодування шляхом його заміни на двійковий вектор-рядок з $l \cdot H(X)$ бітів. Оскільки існує не більше $2^{l \cdot H(X)}$ типових повідомлень довжини l, то для їх кодування (з можливістю подальшого однозначного декодування) довічних вектор-рядків довжини $l \cdot H(X)$ досить; більш того, встановлено, що число $l \cdot H(X)$ неможливо зменшити, тобто, всіх довічних вектор-рядків фіксованої довжини, менш, ніж $l \cdot H(X)$, не вистачає для кодування всіх типових повідомлень довжини l. При великих значеннях l дана схема стиснення працює коректно (тобто, без помилок) з ймовірністю, що наближається до одиниці

[33], [56].

Таким чином, щоб передати, по суті, всю інформацію, що переноситься вектор-рядком з l бітів, досить вибрати двійковий блоковий код В, який привласнює кодове слово довжини $l \cdot H(X)$ бітів кожного типового вектор-рядку з l бітів. Цей блоковий рівномірний код В має $2^{l \cdot H(X)}$ слів однакової довжини $l \cdot H(X)$ бітів, що з'являються з однаковою ймовірністю $2^{l \cdot H(X)}$, і називається оптимальним блоковим рівномірним кодом [33], [56], [64]. Оскільки $0 \leq H(X) \leq 1$ при $0 \leq p \leq 1$ і $H(X) = 1$ тільки при p = 0,5, то оптимальний блоковий код В стискає повідомлення при будь-якому $p \neq 0{,}5$. В силу того, що ймовірність кожного слова в коді В дорівнює $2^{-l \cdot H(X)}$, можна вважати, що кожне кодове слово генерується протягом $l \cdot H(X)$ тактів джерелом повідомлень, який генерує 0 і 1 потактно за схемою незалежних випробувань з однією і тією ж ймовірністю 0,5 [56].

Тепер розглянемо питання про обчислення величини $\frac{min}{f \in F_2^n} E\left(\delta_f(m)\right)$ при впровадженні в контейнер (з попереднім стисненням оптимальним блоковим кодом В) типового повідомлення довжини l, що представляє собою результат роботи протягом l тактів джерела, що генерує символи 0 і 1 потактно за схемою незалежних випробувань з вірогідністю p і q, відповідно, де $p \geq 0, q \geq 0$, p + q = l. В результаті кодування вихідного повідомлення з використанням оптимального коду В, отримаємо, як було зазначено вище, кодове слово довжини $l \cdot H(X)$, що представляє собою результат роботи протягом $l \cdot H(X)$ тактів джерела, що генерує символи 0 і 1 потактно за схемою незалежних випробувань з однією і тією ж ймовірністю 0,5. Впроваджуємо в контейнер отримане кодове слово. Тоді для будь-якого натурального числа n і будь-якої булевої функції f від n змінних справедливий ланцюжок співвідношень:

$$E\left(\delta_f(m)\right) = \sum_{i=0}^{n} i \cdot P\left(\delta_f(m) = i\right) = \sum_{i=0}^{n} i \cdot P(m = 0) \cdot P\left(v \in A_{if}\right) + P(m = 1) \cdot$$

$$\cdot P\left(v \in B_{if}\right) = \frac{0{,}5}{2^n} \cdot \sum_{i=1}^{n} i \cdot y_{if} + \frac{0{,}5}{2^n} \cdot \sum_{i=1}^{n} i \cdot z_{if} \geq \frac{0{,}5}{2^n} \cdot \left(\sum_{i=1}^{n} y_{if} + \sum_{i=1}^{n} z_{if}\right) = 0{,}5.$$

Отже, в цьому випадку $\frac{min}{f \in F_2^n} E\left(\delta_f(m)\right) \geq 0{,}5$ і нижня межа досягається для функції $g(x_1, x_2, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$. Дійсно, для функції $g(x_1, x_2, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$ справедливі рівності:

$$A_{2_g} = \dots = A_{ng} = B_{2g} = \dots = B_{ng} = \emptyset,$$

$$A_g = \{v \in V_2^n | g(v) = 1\} = \{v \in V_2^n | \|v\| \equiv 1 (mod 2)\} = A_{1_g},$$

$$B_g = \{v \in V_2^n | g(v) = 0\} = \{v \in V_2^n | \|v\| \equiv 0 (mod 2)\} = B_{1_g},$$

де $\|v\|$ – вага вектора $v \in V_2^n$. І тоді:

$$E\left(\delta_g(m)\right) = \sum_{i=0}^{n} i \cdot P\left(\delta_g(m) = i\right) = \sum_{i=0}^{i} i \cdot P\left(\delta_g(m) = i\right) =$$

$$= P(m = 0) \cdot P\left(v \in A_{1_g}\right) + P(m = 1) \cdot P\left(v \in B_{1_g}\right) = 0{,}5^2 + 0{,}5^2 = 0{,}5$$

Звідси випливає, що при впровадженні в контейнер всього кодового слова довжини $l \cdot H(X)$ оптимального блокового рівномірного коду В виконується рівність

$$\frac{min}{f \in F_2^n} E\left(\delta_f(m)\right) \cdot l \cdot H(X) = 0{,}5 \cdot l \cdot H(X)$$

Таким чином, при впровадженні в контейнер довічного повідомлення довжини l (з попереднім стисненням оптимальним блоковим рівномірним кодом В), спочатку представляє собою результат роботи протягом l тактів джерела, що генерує символи 0 і 1 потактно за схемою незалежних випробувань з вірогідністю p і q, відповідно ($p \geq 0, q \geq 0$, p + q = l), мінімальне значення математичного очікування числа змінених елементів контейнера одно $0{,}5 \cdot l \cdot H(X)$. Поділивши цю величину на довжину l

повідомлення, отримуємо, що мінімальне значення математичного очікування випадкової величини, яка дорівнює числу змінених елементів контейнера при впровадженні одного біта повідомлення в n елементів контейнера, так само $0{,}5H\ (X)$. Припущення 2.2.1 доведено.

Таким чином, встановлено, що ентропійний стеганографічний алгоритм з функцією впровадження-вилучення $g\ (x_1\ x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n$ є оптимальним ентропійним стеганографічним алгоритмом для захисту повідомлень, згенерованих двійковим джерелом без пам'яті і стислих асимптотично оптимальним блоковим рівномірним кодом. Тобто, для повідомлень, згенерованих двійковим джерелом без пам'яті і стислих перед впровадженням в стеганографічний контейнер з використанням асимптотично оптимального блокового рівномірного коду, оптимальним ентропійним стеганографічним алгоритмом є алгоритм, який здійснює заповнення стеганографічного контейнера побітово за ознакою парності суми обраних елементів [29].

2.3. Субоптимальний ентропійний стеганографічний алгоритм

Побудуємо пару систем підмножин $\{A_1, \ldots, A_n\}$ і $\{B_1, \ldots, B_n\}$ множини V_2^n наступним чином.

Для системи підмножин $\{A_1, \ldots, A_n\}$ покладемо: $A_1 -$ це безліч, що складається з одного довільним чином зафіксованого вектора $\{\alpha_1, \ldots, \alpha_n\} \in$

V_2^n, а множини $A_2, \dots, A_n$ - порожні безлічі, тобто:

$A_1 = \{(\alpha_1, \dots, \alpha_n)\}$, де $\alpha_i \in \{0,1\}, i = \overline{1,n}$;

$A_2 = \dots = A_n = \emptyset.$ (2.3.1)

Для системи підмножин $\{B_1, \dots, B_n\}$ покладемо: B_i - це безліч, що складається з усіх векторів, що виходять з вектора $(\alpha_1, \dots, \alpha_n)$ шляхом інвертування рівно *i* координат, де $i = \overline{1,n}$.

Побудована пара систем підмножин $(\{A_1, \dots, A_n\}, \{B_1, \dots, B_n\}) = (A_1^n, B_1^n)$ множини V_2^n є ДПСП множини V_2^n тобто

$(A_1^n, B_1^n) \in S(V_2^n)$.

Дійсно в системі підмножин $\{A_1, \dots, A_n\}$ множини V_2^n з побудови тільки підмножина A_1 є непустою, і шляхом інвертування однієї координати будь-якого вектора, що належитьA_1, отримуємо вектор, що належить множині B_1, отже, і безлічі $B = B_1 \cup \dots \cup B_n$. У свою чергу, в системі підмножин $B_1, \dots, B_n\}$ множини V_2^n всі підмножини являються не пустими по побудові. Розглянемо вектори, що відносяться до підмножини B_i, де $i = \overline{1,n}$. Будь-який з цих векторів шляхом інвертування рівно i відповідних його координат перетворюється в вектор $(\alpha_1, \dots, \alpha_n)$, що відноситься до A_1 , отже і множині $A = A_1 \cup \dots \cup B_n$.

Таким чином, побудована вище пара систем підмножин (A_1^n, B_1^n) множини V_2^n є ДПСП множини V_2^n. Позначимо її через $\sigma(\alpha_1, \dots, \alpha_n)$.

Нехай дійсне число q таке, що

0 < q < 0,5.

Поставимо питання про обчислення величини

$F^{(q)}_{\sigma(\alpha_1,\dots,\alpha_n)} = F^{(q)}(A_1^n, B_1^n)$,

де $F^{(q)}(A_1^n, B_1^n)$ задана рівністю (2.2.3) $y_i = |A_i|$, $z_i = |B_i|$, $i = \overline{1,n}$.

Має місце рівність

$$F^{(q)}_{\sigma(\alpha_1,\dots,\alpha_n)} = \frac{1-q}{2^n} + \frac{nq}{2}$$

З рівності (2.4.4) і (2.2.3) випливає, що

$$F^{(q)}_{\sigma(\alpha_1,\dots,\alpha_n)} = F^q(A_1^n, B_1^n) = \frac{1-q}{2^n} \cdot \sum_{i=1}^{n} i \cdot y_i + \frac{q}{2^n} \cdot \sum_{i=1}^{n} i \cdot z_i$$

Для потужностей $y_i = |A_i|$, $z_i = |B_i|$ множин A_i, B_i, відповідно (де $i = \overline{1,n}$), справедливі рівності:

$$y_1 = 1, y_2 = \cdots = y_n = 0; z_i = C_n^i$$

де C_n^i - число поєднань з n елементів по i елементів, $i = \overline{1,n}$

Підставивши в (2.4.6) значення потужностей множин A_i, B_i (де $i = \overline{1,n}$,) представлені равенствами (2.4.7), отримуємо:

$$F^{(q)}_{\sigma(\alpha_1,\dots,\alpha_n)} = \frac{1-q}{2^n} \cdot (1 + \sum_{i=1}^{n} i \cdot 0) + \frac{q}{2^n} \cdot \sum_{i=1}^{n} i \cdot \mathrm{C}_n^i$$

Звідси, скориставшись рівністю (див. [53])

$$\sum_{i=1}^{n} i \cdot \mathrm{C}_n^i = n \cdot 2^{n-1}$$

маємо

$$F^{(q)}_{\sigma(\alpha_1,\dots,\alpha_n)} = \frac{1-q}{2^n} + \frac{q}{2^n} \cdot n \cdot 2^{n-1} = \frac{1-q}{2^n} + \frac{nq}{2}$$

що й потрібно було довести.

Для $n \in N$ і дійсного числа q, що задовольняє подвійній нерівності (2.4.3), позначим

$$M(n,q) = \frac{1-q}{2^n} + \frac{nq}{2}$$

Нехай натуральне число t таке, що число q задовольняє подвійній нерівності

$$\frac{1}{2^t+1} < q \leq \frac{1}{2^{t-1}+1}$$

(а) Якщо число q задовольняє строгому подвійному нерівності

$$\frac{1}{2^t+1} < q \leq \frac{1}{2^{t-1}+1}$$

при деякому $t \in N$, то кінцева послідовність величин $\{M(n,q)\}_{n=1}^{t}$

є убуваючою, і при цьому максимальний член -

$$M(1,q) = 0{,}5,$$

а мінімальний член –

$$M(t,q) = \frac{1-q}{2^t} + \frac{tq}{2}$$

(б) Якщо число q задовольняє рівності

$$q = \frac{1}{2^{t-1}+1}$$

при деякому $t \in N$, то кінцева послідовність величин $\{M(n,q)\}_{n=1}^{t-1}$ є убуваючою, і при цьому максимальний член –

$$M(l,q) = 0{,}5,$$

а мінімальний член –

$$M(t-1,q) = \frac{1-q}{2^{t-1}} + \frac{t-1\cdot}{2}$$

а також справедлива рівність

$$M(t-1,q) = M(t,q)$$

(с) Послідовність величин $\{M(n,q)\}_{n=t}^{\infty}$ є зростаючою послідовністю.

Перед доведенням теореми доречно вказати на те, що в подальшому розглядаються лише ті випадки, де межі зміни параметра *n* в залежності від *t* визначені коректно.

Доведемо пункт (а). Значення $M(l,q) = 0{,}5$ отримуємо безпосередньо з рівності (2.4.8), поклавши *n*=1. Далі при $n \in \{1, \dots, t-1\}$

обчислимо різницю

$$M(n+1,q) - M(n,q) = \frac{1-q}{2^{n+1}} + \frac{(n+1)\cdot q}{q} - \frac{1-q}{2^n} - \frac{nq}{2} = \frac{q-1}{2^{n+1}} + \frac{q}{2}$$

З (2.4.14) і правої частини подвійної нерівності (2.4.11) слідує справедливість нерівності

$$M(n+1,q) - M(n,q) < \frac{2^n - 2^{t-1}}{(2^{t-1}+1)\cdot 2^{n+1}}$$

При $n \in \{1, \dots, t-2\}$ права частина нерівності (2.4.15) від'ємна, а при $n = t-1$ равна нулю. Отже, справедлива нерівність

$$M(n+1,q) - M(n,q) < 0$$

Для кожного $n \in \{1, \dots, t-1\}$, що рівносильно тому, що кінцева послідовність $\{M(n,q)\}_{n=1}^{t}$ є убуваючою, та $M(t,q) -$ її мінімальний елемент, $M(l,q) -$ максимальний елемент.

Пункт (а) доведений

Доведемо пункт (б). Так само, як і в пункті (а), доводиться справедливість $M(l,q) = 0{,}5$. Далі при $n \in \{1, \dots, t-2\}$ отримаємо

$$M(n+1,q) - M(n,q) = \frac{q-1}{2^{n+1}} + \frac{q}{2}$$

Звідси і з рівності (2.4.12) випливає справедливість рівності

$$M(n+1,q) - M(n,q) = \frac{2^n}{(2^{t-1}+1)\cdot 2^{n+1}}$$

При $n \in \{1, \dots, t-2\}$ права частина рівності (2.4.16) від'ємна, а при $n = t-1$ дорівнює нулю. Отже, справедлива нерівність

$$M(n+1,q) - M(n,q) < 0$$

Для кожного $n \in \{1, \dots, t-2\}$, а при $n = t-1$ вірна рівність

$$M(n+1,q) - M(n,q) = 0$$

що рівнозначно тому, що кінцева послідовність $\{M(n,q)\}_{n=1}^{t-1}$ є убиваючою, $M(t-l,q) -$ її мінімальний елемент, $M(l,q)$ – максимальний елемент і, крім того

$$M(t,q) = M(t-l,q).$$

Пункт (б) доведений

Доведемо пункт (в). При $n \geq t$ аналогічно (2.4.14) отримуємо

$$M(n+1,q) - M(n,q) = \frac{q-1}{2^{n+1}} + \frac{q}{2}$$

Звідси і з лівої частини подвійної нерівності (2.4.9) слідує справедливість нерівності

$$M(n+1,q) - M(n,q) > \frac{2^n - 2^t}{(2^{t-1}+1)\cdot 2^{n+1}}$$

При $n \geq t$ права частина рівності (2.4.17) позитивна і, отже, справедлива нерівність

$$M(n+1,q) - M(n,q) > 0$$

для кожного $n \geq t$, що рівнозначно тому, що послідовність $\{M(n,q)\}_{n=t}^{\infty}$ є зростаючою послідовністю.

Пункт (в) доведений.

Для наочності приклада послідовності $\{M(n,q)\}_{n=1}^{\infty}$ для різних значень q представлені графічно на рис. 2.4.18 і 2.4.20; необхідні обчислення наведені в таблицях 2.3.1 і 2.3.3, відповідно.

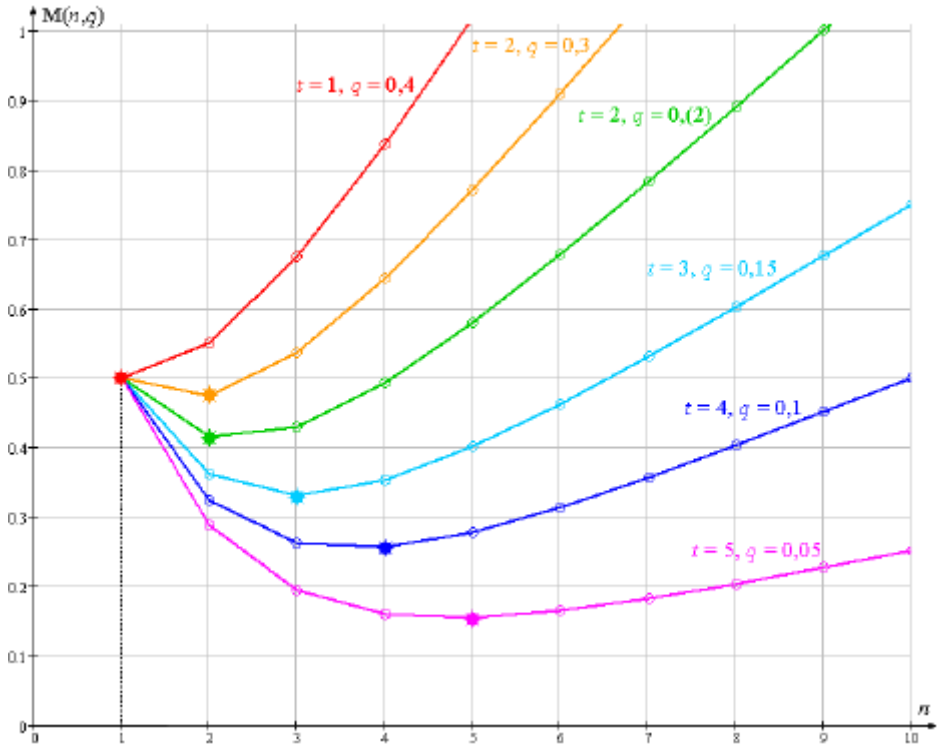

Рисунок 2.3.1. Графічні зображення початкових елементів послідовності $\{M(n,q)\}_{n=1}^{\infty}$ при $\frac{1}{2^t+1} < q < \frac{1}{2^{t-1}+1}$ для різних значень t.

Таблиця 2.3.1

Результати обчислення значень величини $M(n,q)$ при $\frac{1}{2^t+1} < q < \frac{1}{2^{t-1}+1}$

$M(n,q)=\frac{1-q}{2^n}+\frac{nq}{2}$						
n \ q	0.4	0.3	0.(2)	0.15	0.1	0.05
1	0.5	0.5	0.5	0.5	0.5	0.5
2	0.55	0.475	0.417	0.363	0.325	0.288
3	0.675	0.538	0.431	0.331	0.263	0.194
4	0.838	0.644	0.493	0.353	0.256	0.159
5	1.019	0.772	0.580	0.402	0.278	0.155
6	1.209	0.911	0.679	0.463	0.314	0.165
7	1.405	1.055	0.784	0.532	0.357	0.182
8	1.602	1.203	0.892	0.603	0.404	0.204

Доречно звернути увагу на те, що при $\frac{1}{2^t+1} < q < \frac{1}{2^{t-1}+1}$ є один мінімальний елемент послідовності $\{M(n,q)\}_{n=1}^{\infty}$ (на рисунку і в таблиці точка мінімуму зафарбована).

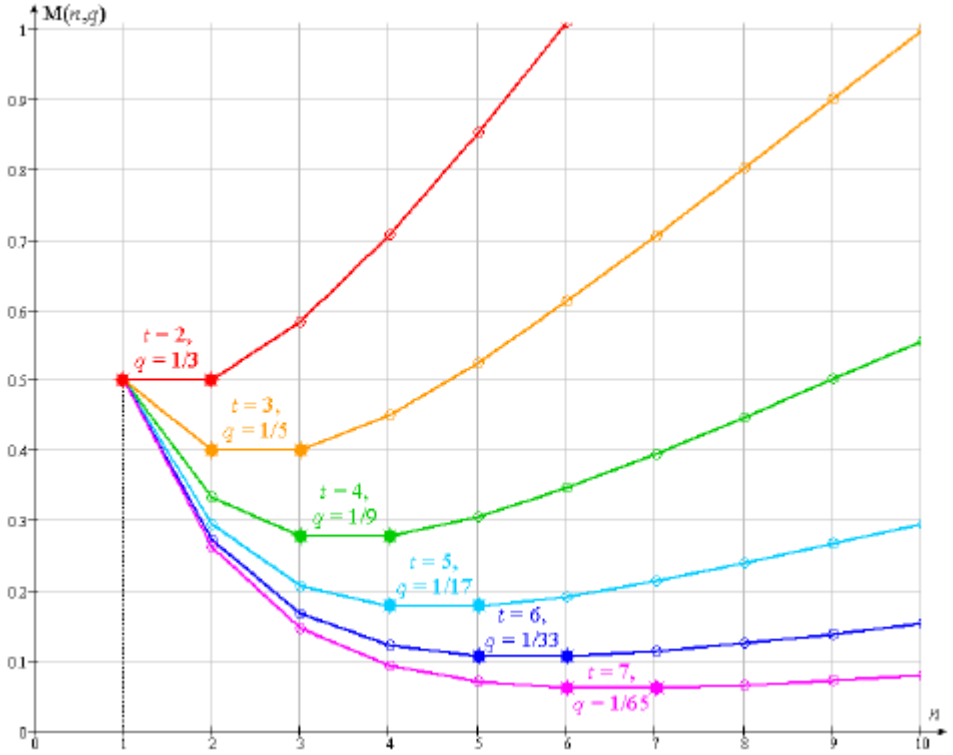

Рисунок 2.3.2. Графічні зображення початкових елементів послідовності $\{M(n,q)\}_{n=1}^{\infty}$ при $q = \frac{1}{2^{t-1}+1}$ для різних значень t

Таблиця 2.3.2

Результати обчислення значень величини $M(n,q)$ при $q = \frac{1}{2^{t-1}+1}$

$$M(n,q)=\frac{1-q}{2^n}+\frac{nq}{2}$$

n \ q	1/3	1/5	1/9	1/17	1/33	1/65
1	0.5	0.5	0.5	0.5	0.5	0.5
2	0.5	0.4	0.333	0.294	0.273	0.262
3	0.583	0.4	0.278	0.206	0.167	0.146
4	0.708	0.45	0.278	0.176	0.121	0.092
5	0.854	0.525	0.306	0.176	0.106	0.069
6	1.010	0.613	0.347	0.191	0.106	0.062
7	1.172	0.706	0.396	0.213	0.114	0.062
8	1.336	0.803	0.448	0.239	0.125	0.065

Звернемо увагу на те, що при

$$q=\frac{1}{2^{t-1}+1}$$

є два мінімальних елемента послідовності $\{M(n,q)\}_{n=1}^{\infty}$ (На рисунку і в таблиці точки мінімуму зафарбовані).

Приступимо тепер до вирішення наступних завдань: побудова ентропійного стеганографічного алгоритму впровадження в контейнер повідомлень, що генеруються двійковим джерелом без пам'яті, на виході якого з'являються символи 0 і 1 з ймовірностями p і q, відповідно, де $p \geq 0, q \geq 0$, p + q = l, $\frac{1}{2^t+1} < q < \frac{1}{2^{t-1}+1}$, $t \in N$ і доказу його субоптимальності. При цьому побудову самого стеганографічного алгоритму в цьому розділі здійснимо в тій мірі повноти, яка достатня для з'ясування питання його субоптимальності. Загальний опис стеганографічного алгоритму приведено в розділі 2.1. З цього опису для побудови ентропійного стеганографічного алгоритму необхідно, зокрема, визначити відповідну булеву функцію *f*. Задамо цю булеву функцію $f(x_1,\ldots,x_t)$, поклавши

$$f(x_1,\ldots,x_t)={x_1}^{\alpha}\cdot{x_2}^{\alpha}\cdot\ldots\cdot{x_t}^{\alpha},$$

де

$$x_i^{\alpha_i} = \begin{cases} 1, \text{якщо } x_i = \alpha_i \\ 0 \text{ в протилежному випадку'} \end{cases}$$

$i \in (1, \dots, t), (\alpha_i, \dots, \alpha_t) -$ деякий довільним чином зафіксований вектор з безлічі V_2^t. Тут число змінних булевої функції, позначене в розділі 2.1 через *n*, відразу взято рівним *t* з причини, яка буде пояснена нижче. Тоді як зазначено (2.1.5) - (2.1.8), введеними в розділі 2.1, маємо

$$A_{1f} = \{(\alpha_1, \dots, \alpha_t)\}, A_{2f} = \dots = A_{tf} = \emptyset$$

$$A_f = \{v \in V_2^t | f(v) = 1\} = A_{1f} \cup \dots \cup A_{tf} = \{(\alpha_1, \dots, \alpha_t)\}$$

Для системи підмножин $\{B_{1f}, \dots, B_{tf}\}$ покладемо: B_{if} - це безліч, що складається з усіх векторів, що виходять з вектора $(\alpha_1, \dots, \alpha_t)$ шляхом інвертування рівно *i* координат, де $i = \overline{1, t}$.

Для впровадження в контейнер одного елемента $m \in \{0; 1\}$ повідомлення, згенерованого джерелом повідомлень, по ключу генеруються відповідні номери елементів і вибираються самі елементи $g_{i_1}, g_{i_2}, \dots, g_{i_t}$ контейнера. В процесі впровадження m в контейнер вибрані елементи контейнера змінюються, і виходять елементи $\bar{g}_{i_1}, \bar{g}_{i_2}, \dots, \bar{g}_{i_t}$, таки що

$$f(\tilde{b}_{i_1}, \tilde{b}_{i_2}, \dots, \tilde{b}_{i_t}) = m$$

де

$$\tilde{b}_{i_r} = \begin{cases} 0, \text{якщо } \tilde{g}_{i_r} - \text{парне число;} \\ 1, \text{ якщо } \tilde{g}_{i_r} - \text{непарне число} \end{cases}$$

$r \in \{1,2, \dots, t\}$. Тоді з рівності для математичного очікування випадкової величини, яка дорівнює числу змінених елементів контейнера при впровадженні в *t* елементів контейнера повідомлення *m* з одного біта, згенерованого джерелом повідомлень з параметрами *p* і *q*, вірна наступна рівність

$$\mathrm{E}\left(\delta_f(m)\right) = \frac{1-q}{2^t} + \frac{tq}{2}$$

Звідси і отримуємо

$$\mathrm{E}\left(\delta_f(m)\right) = M(t, q) = \min_{n \in N} M(n, q)$$

де мінімум береться по параметру *n*, що приймає значення з безлічі натуральних чисел N.

Рівності (2.4.24) в сукупності з тією обставиною, що в даній роботі процес оптимізації стеганографічних алгоритмів здійснюється з позиції зменшення величини $\mathrm{E}\left(\delta_f(m)\right)$, і прояснюють причину того, чому вище число змінних булевої функції *f* належить рівним *t*. Дійсно, з (2.1.9) випливає, що при числі змінних булевої функції, що дорівнює *t*, отримуємо мінімальне значення величини $\mathrm{E}\left(\delta_f(m)\right)$, для способу побудови стеганографічного алгоритму, обраного в цьому розділі.

Покажемо, що обговорюваний в даному розділі алгоритм є субоптимальних алгоритмом. Для цього, як випливає з 2.1, досить довести, що величина $\mathrm{E}\left(\delta_f(m)\right)$ не перевищує числа 0,5, тобто необхідно довести справедливість нерівності

$$\frac{1-q}{2^t}+\frac{tq}{2}\leq\frac{1}{2}$$

при умові що

$$\frac{1}{2^t+1}<q<\frac{1}{2^{t-1}+1},\ t\in N$$

Однак має сенс привести і незалежний доказ нерівності.

Спершу окремо розглянемо випадки *t* = l і *t* = 2.

Нехай *t* = l. тоді:

$$\frac{1-q}{2^t}+\frac{tq}{2}=\frac{1-q}{2}+\frac{q}{2}=\frac{1}{2}$$

і, отже, нерівність (2.4.25) при *t* = l виконується.

Нехай *t* = 2. Тоді з правої частини подвійного нерівності (2.4.26) слід:

$$\frac{1-q}{2^t}+\frac{tq}{2}=\frac{1-q}{2^2}+\frac{2q}{2}=\frac{1}{4}+\frac{3q}{4}\leq\frac{1}{4}+\frac{3}{4}\cdot\frac{1}{3}=\frac{1}{2}$$

і, отже, нерівність (2.4.25) при *t* = 2 також виконується.

Тепер розглянемо випадок $t\geq 3$. З подвійної нерівності (2.4.26) випливає, що

$$\frac{1-q}{2^t}+\frac{tq}{2}<\frac{1-\left(\frac{1}{2^t+1}\right)}{2^t}+\frac{t\cdot\left(\frac{1}{2^{t-1}+1}\right)}{2}=\frac{2^t}{2^t\cdot(2^t+1)}+\frac{t}{2^t+2}=$$

$$=\frac{1}{2^t+1}+\frac{t}{2^t+2}<\frac{1}{2^t}+\frac{t}{2^t}=\frac{t+1}{2^t}$$

тобто, справедлива нерівність

$\frac{1-q}{2^t}+\frac{tq}{2}<\frac{t+1}{2^t}$

Розглянемо функцію

$y=\frac{t+1}{2^t}$

визначену на множині всіх дійсних чисел, отже, і при $t \geq 3$. Похідна

$$y'=\frac{(t+1)'\cdot 2^t-(t+1)\cdot(2^t)}{(2^t)^2}=\frac{2^t-(t+1)\cdot 2^t\cdot\ln 2}{(2^t)^2}=\frac{1-(t+1)\cdot\ln 2}{2^t}$$

цієї функції існує всюди на R і негативна при $t>\log_2 e-1$. Отже (див. [45]), функція (2.4.28) є спадною функцією при $t>\log_2 e-1$. Так як $3>\log_2 e-1$, то функція (2.4.28) є убуваючою і при $t \geq 3$. Значення функції (2.4.28) при *t*=3 дорівнює 0,5. З урахуванням цього, з рівності (2.4.27) отримаємо

$\frac{1-q}{2^t}+\frac{tq}{2}<0{,}5$

при $t \geq 3$.

Отже, доведена справедливість нерівності (2.4.25) для всіх $t \geq 1$, тобто вірна нерівність $\mathrm{E}\left(\delta_f(m)\right)\leq 0{,}5$, що рівносильно субоптимальності побудованій вище в цьому розділі ентропійного стеганографічного алгоритму.

Представляє теоретичний і практичний інтерес порівняння даного алгоритму з алгоритмом впровадження повідомлення з попереднім стисненням останнього асимптотично оптимальним блоковим рівномірним кодом. Певне уявлення про переваги побудованого алгоритму можна отримати із таблиці 2.4.29 і графіків залежності від $t \in N$ при $q=\frac{1}{2^t}$ математичного очікування числа змінених елементів контейнера при впровадженні одного біта повідомлення для трьох стеганографічних алгоритмів: алгоритму впровадження по парності суми елементів контейнера,

субоптимального ентропійного стеганографічного алгоритму і оптимального ентропійного стеганографічного алгоритму (рис. 2.4.30). При цьому вимога $t \in N$ виправдано і тим, що для q справедливо нерівність (2.4.3).

Таблиця 2.3.3.

Чисельні результати для порівняння стеганографічних алгоритмів

t	1	2	3	4	5	6	7	8	9	10
$q = 1/2^t$	0,5	0,25	0,125	0,063	0,031	0,016	0,008	0,004	0,002	0,001
$\mathbb{E}_f$	0,5	0,438	0,297	0,184	0,108	0,062	0,035	0,02	0,011	0,006
H/2	0,5	0,406	0,272	0,169	0,1	0,058	0,033	0,019	0,01	0,006
Δ_H	0	0,032	0,025	0,015	0,008	0,004	0,002	0,001	0,001	$<10^{-3}$
∂_H, %	0	7,9	9,2	8,9	8,1	7,2	6,5	5,9	5,3	4,9
Δ_2	0	0,062	0,203	0,316	0,392	0,438	0,465	0,48	0,489	0,494
∂_2, %	0	12,5	40,6	63,3	78,3	87,6	93	96,1	97,8	98,8

Тут:

$E_f = \frac{1-q}{2^t} + \frac{tq}{2}$ – математичне очікування числа змін для субоптимального алгоритму;

$H = -\big((1-q)\cdot \log_2(1-q) + q \cdot \log_2 q\big)$ – ентропія двійкового джерела без пам’яті;

Δ_H - абсолютна різниця між математичним очікуванням числа змін для субоптимального алгоритму і оптимального алгоритму (для захисту повідомлень з попередніми стисненням асимптотично оптимальним блоковим рівномірним кодом);

∂_H – відносна різниця, що показує, на скільки відсотків ефективність субоптимального алгоритму нижче ефективності оптимального алгоритму: $\partial_H = \frac{\Delta_H}{H/2} \cdot 100\%$;

Δ_2 – абсолютна різниця між математичним очікуванням числа змін для субоптимального алгоритму і алгоритму впровадження за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення;

$\partial_2 -$ відносна різниця, що показує, на скільки відсотків ефективність субоптимального алгоритму вище ефективності алгоритму впровадження за ознакою парності: $\partial_2 = \frac{\Delta_2}{0{,}5} \cdot 100\%$.

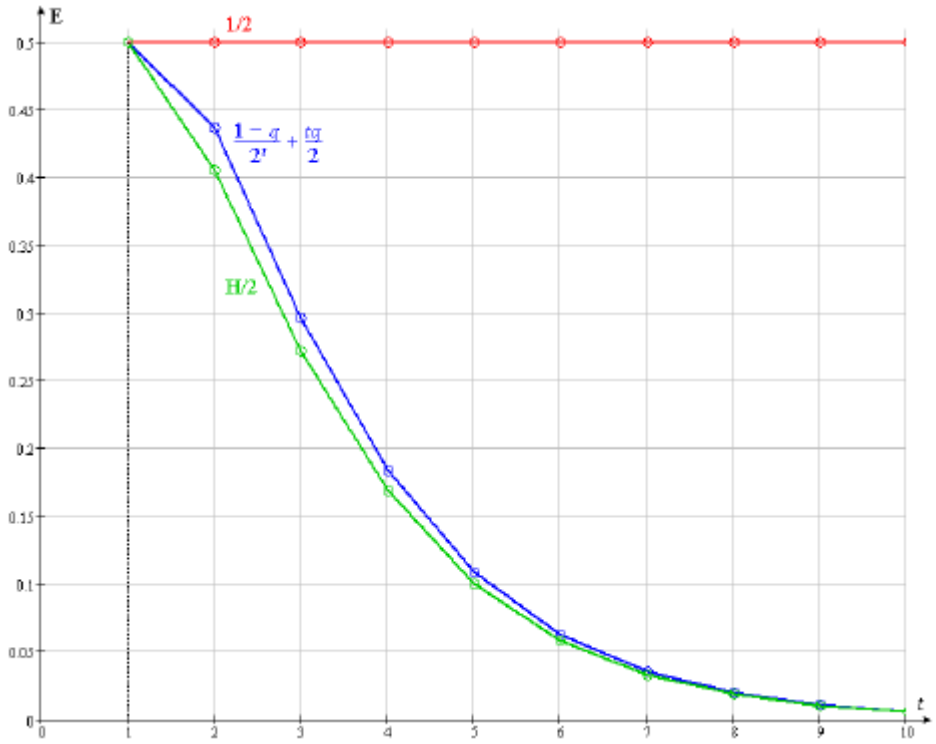

Рисунок 2.3.3 Графічне зображення залежності від $t \in N$ при $q = 1/2^t$ математичного очікування числа змінених елементів контейнера при впровадженні одного біта повідомлення для трьох стеганографічних алгоритмів.

З таблиці 2.4.29 і рисунка 2.4.3 випливає, що викладений в даному розділі субоптимальний ентропійний стеганографічний алгоритм впровадження в контейнер повідомлення без попереднього стиснення досить близький за своїми характеристиками до оптимального ентропійного стеганографічного алгоритму впровадження в контейнер повідомлення з попереднім стисненням асимптотично оптимальним блоковим рівномірним кодом . Ця близькість посилюється в міру зменшення параметра q двійкового джерела повідомлень.

2.4 Висновки

Ймовірносно-статистичні характеристики джерела повідомлень можуть служити основою оптимізації процедури впровадження повідомлень в стеганографічний контейнер з позиції зменшення середнього числа змінних елементів контейнера на один біт впроваджуваного повідомлення. Оптимізовані зазначеним підходом стеганографічні алгоритми називаються оптимальними стеганографічними алгоритмами.

При певних обмеженнях на стеганографічний алгоритм завдання його оптимізації може бути зведена до задачі перебору (з обчисленням певного параметра) в множині всіх булевих функцій з заданим числом аргументів *n*, що є завданням комбінаторної оптимізації, ефективно розв'язуваної тільки для невеликих значень n. У зв'язку з цим з'являється необхідність визначення і побудови субоптимальних алгоритмів, що не припускають при своїй побудові громіздких обчислювальних процедур.

Для субоптимальних алгоритмів верхня межа для середнього числа змінних елементів контейнера на один біт впроваджуваного повідомлення дорівнює 0,5. Це значення є досяжним, і воно - найменше значення, при якому безліч субоптимальних ентропійних стеганографічних алгоритмів не є порожнім для будь-якого фіксованого двійкового джерела без пам'яті.

Виявлено набір структурних ознак області визначення булевої функції впровадження-вилучення оптимального ентропійного стеганографічного алгоритму, який можна ефективно використовувати при побудові оптимальних і субоптимальних ентропійних стеганографічних алгоритмів.

Отримано асимптотичний результат, що полягає в тому, що середнє число змінюваних елементів контейнера на один біт вихідного двійкового повідомлення при його впровадженні з попередніми стисненням асимптотично оптимальним блоковим рівномірним кодом дорівнює половині ентропії довічного джерела повідомлень без пам'яті.

Розроблено (див. розділ 2.3) субоптимальний ентропійний алгоритм. Для цього алгоритму встановлено оптимальне значення числа елементів контейнера, задіяних для впровадження одного біта повідомлення.

Шляхом порівняльного аналізу виявлено переваги розробленого в рохділі 2.3 субоптимального ентропійного стеганографічного алгоритму по відношенню до стеганографічного алгоритму з лінійною булевою функцією впровадження-добування інформації.

Розділ 3. РОЗРОБКА ЗАХИЩЕНОЇ СТЕГАНОГРАФІЧНОЇ СИСТЕМИ

3.1. Методика побудови ентропійних стеганографічних систем з контейнерами, що представляють собою цифрові мультимедійні файли

Як випливає з назви даного параграфа, тут буде викладена методика побудови ентропійних стеганографічних систем з контейнерами, що представляють собою отримані за допомогою стандартних алгоритмів стиснення образи мультимедійних файлів. При цьому під методикою ми розуміємо сукупність методів і прийомів доцільного проведення будь-якої роботи [49].

До основних завдань стеганографії належать розробка нових і вдосконалення наявних стеганографічних методів і способів захисту інформації та створення на їх основі високоефективних стеганографічних систем для безпечного зберігання та передачі інформації, в тому числі і конфіденційної з гарантованими значеннями за параметрами стійкості, а також допускають надійні програмні і апаратні реалізації.

Для успішного вирішення цих завдань особливе значення має застосування сучасних наукових методів, які дозволяють глибоко проаналізувати останні зміни в стеганографічних методах захисту інформації та розкрити закономірності розвитку цих методів. Це дає можливість на якісно новому рівні підійти до вирішення завдань забезпечення безпечного зберігання та передачі конфіденційної інформації на основі створення та використання сімейства стеганографічних систем з контейнерами, що представляють собою підмножини квантових коефіцієнтів частотної області мультимедійних сигналів. Дані квантовані коефіцієнти є результатом застосування до мультимедійних файлів стандартних алгоритмів стиснення. До них відносяться алгоритми MP3, AAC, Ogg Vorbis, AC3 (для стиснення звукових файлів), JPEG, JPEG 2000 (для стиснення графічних файлів), MPEG-2, MPEG-4, H.263, H.264 (для стиснення відеофайлів) та ін. Коло стандартних алгоритмів з кожним роком все більше розширюється, поліпшуючись якісно і кількісно в плані характеристик стисненого уявлення мультимедійної інформації. Однак постійним в них, серед іншого, залишається (незалежно від

того, що стискається: звукові, графічні або відеофайли) використання процедур переходу в частотну область і подальшого квантування. Дана обставина є основою для досліджень в напрямку створення в певному сенсі «всеїдних» універсальних стеганографічних алгоритмів, що допускають контейнери, побудовані на основі будь-яких типів мультимедійних файлів.

Методику будемо представляти у вигляді послідовно виконуваних етапів.

Етап 1. Визначення архітектури стеганографічної системи.

Сформована до теперішнього часу архітектура стеганографічних систем, заснованих на вищезазначених універсальних алгоритмах, передбачає (в режимі впровадження в контейнер повідомлення, що підлягає стеганографічного захисту) наступну двохблокову структуру (рис. 3.1.1):

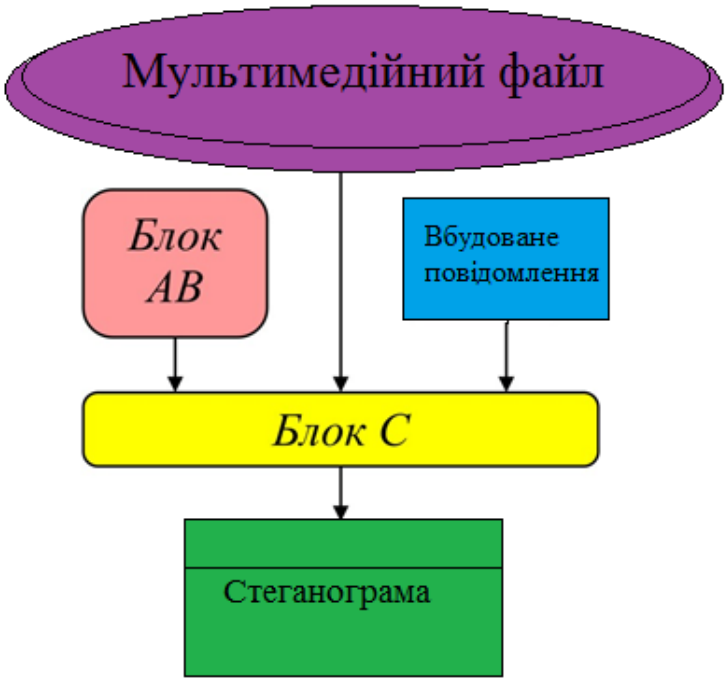

Рисунок 3.1.1. Структурна схема стеганографічної системи в режимі вбудовування інформації

Схему стеганографічної системи (при функціонуванні в режимі вилучення з стеганограмми прихованого в ній повідомлення), побудованої на основі вищезазначених універсальних алгоритмів, можна представити в наступному блоковому вигляді (рис. 3.1.2):

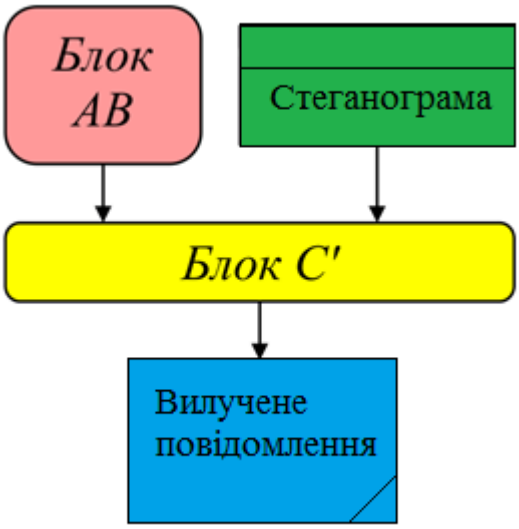

Рисунок 3.1.2. Структурна схема стеганографічної системи в режимі вилучення інформації

Етап 2. Побудова блоку AB, який виробляє послідовність номерів елементів контейнера по Стеганографічному ключу.

Блок AB (на рисунках 3.1.1 і 3.1.2) складається з двох подблоків (рис. 3.1.3): A і B.

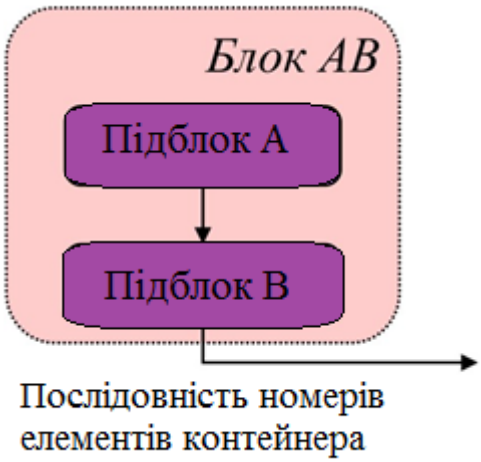

Рисунок 3.1.3. Генератор номерів використовуваних елементів контейнера для впровадження повідомлення, що підлягає стеганографічної захисту

Етап 2.1. Побудова подблока A - генератора псевдовипадкових чисел, початковий стан якого є ключовим елементом стеганографічної системи.

Підблок A є програмної або апаратної реалізацією генератора псевдовипадкових чисел. Як нього може бути обраний будь-який генератор з

безлічі криптографічних генераторів [14], [15], [42], [61], [65]. Найбільш широко відомі і досліджені представники безлічі генераторів псевдовипадкових чисел - це лінійні регістри зсуву і їх ускладнення, такі, як фільтруючі генератори і комбінують генератори. Ключові елементи подблока А складають ключове простір стеганографічної системи. В якості ключових елементів подблока А в залежності від конкретної реалізації можуть бути обрані початкові стану і многочлени зворотного зв'язку лінійних регістрів зсуву, функції ускладнення фільтруючих і комбінує генераторів. Одне з основних вимог до ключового простору стеганографічної системи - це його досить велика потужність, яка повинна забезпечити, зокрема, стійкість стеганографічної системи проти такої «силовий атаки», як виявлення в стеганограмме прихованого повідомлення шляхом «тотального» перебору всіх ключів стеганографічної системи. Після початкової установки ключів подблок А виробляє вихідну послідовність, яка надходить на вхід подблока В, генеруючого:

- в разі, представленому на рис. 3.1.1, - послідовність номерів елементів стеганографічного контейнера, в які будуть вбудовані (впроваджені) елементи повідомлення, що підлягає приховану;

- в разі, представленому на рис. 3.1.2, - послідовність номерів елементів заповненого стеганографічного контейнера, з яких будуть вилучені елементи прихованого в стеганограмме повідомлення.

При цьому вихідна послідовність подблока А для кожної установки ключів повинна задовольняти вимогам криптологічну характеру, таким, як, наприклад, великий період, великий лінійний розмах і відповідні якості ймовірносно-статистичного плану [61], [65]. Практичні завдання програмної і апаратної реалізації криптографічних генераторів отримали своє рішення у багатьох роботах, присвячених розробці програмно апаратних засобів криптографічного захисту інформації. Досягнення відкритих досліджень доступні, наприклад, в монографії Б. Шнайера [65].

Етап 2.2. Побудова подблока В, який виробляє послідовність номерів

елементів контейнера на основі вихідної послідовності подблока А.

Підблок В являє собою програмну або апаратну компоненту, що виробляє послідовність номерів елементів контейнера, які використовуються для впровадження повідомлення, що підлягає стеганографічної захисту. У роботах, присвячених розробці програмно-апаратних засобів стеганографічної захисту інформації, пропонуються різні варіанти побудови подблока В. Багато з них наводяться в відомих монографіях [29], [31], [71], [72] та ін. Однак автору даної роботи представляється найбільш вдалою реалізація подблока В, представлена в роботах [3], [4], [6], [39], [40], [41], основним складовим елементом якої є автомат Мілі [18] спеціального типу. Його детальний опис наведено в роботі [6]. Вихідні послідовності цього автомата визначають перетворення ковзної перестановки, характерною рисою яких є їх придатність до символьних послідовностей потокового виду незалежно від характеру символів і довжини послідовностей (теоретично допускаються навіть нескінченні послідовності). На вхід подблока В надходить вихідна послідовність подблока А. Вихідна послідовність подблока В - це послідовність номерів елементів стеганографічного контейнера, в які будуть вбудовані (впроваджені) елементи повідомлення, що підлягає приховану (рис. 3.1.1); або послідовність номерів елементів заповненого стеганографічного контейнера, з яких будуть вилучені елементи прихованого в стеганограмме повідомлення (рис. 3.1.2). При цьому послідовність номерів така, що реалізується довільне поєднання випадкового вибору і перестановки [43], [64]. Як результуючої вихідний послідовності блоку АВ виступає вихідна послідовність подблока В.

Етап 3. Побудова блоків впровадження повідомлення в контейнер і вилучення повідомлення з контейнера. Вихідна послідовність блоку АВ надходить на вхід блоку С в режимі вбудовування інформації в стеганографічний контейнер (рис. 3.1.1) або на вхід блоку С 'в режимі вилучення з стеганограмми прихованої в ній інформації (рис. 3.1.2).

Етап 3.1. Реалізація блоку С - спеціальної модифікованої версії кодера

стандарту стиснення мультимедійних файлів.

На вхід блоку С (рис. 3.1.1) надходять три послідовності:

- мультимедійний файл;

- повідомлення, яке підлягає стеганографічної захисту (впроваджувана інформація);

- вихідна послідовність блоку АВ.

Результатом роботи блоку С, тобто виходом блоку С, є стеганограмма.

Блок С – це програмна або апаратна реалізація спеціальної модифікованої версії кодера стандарту стиснення мультимедійних файлів. Узагальнене опис модифікованої версії кодера стандарту стиснення мультимедійних файлів може бути представлене в такий спосіб (рис. 3.1.4):

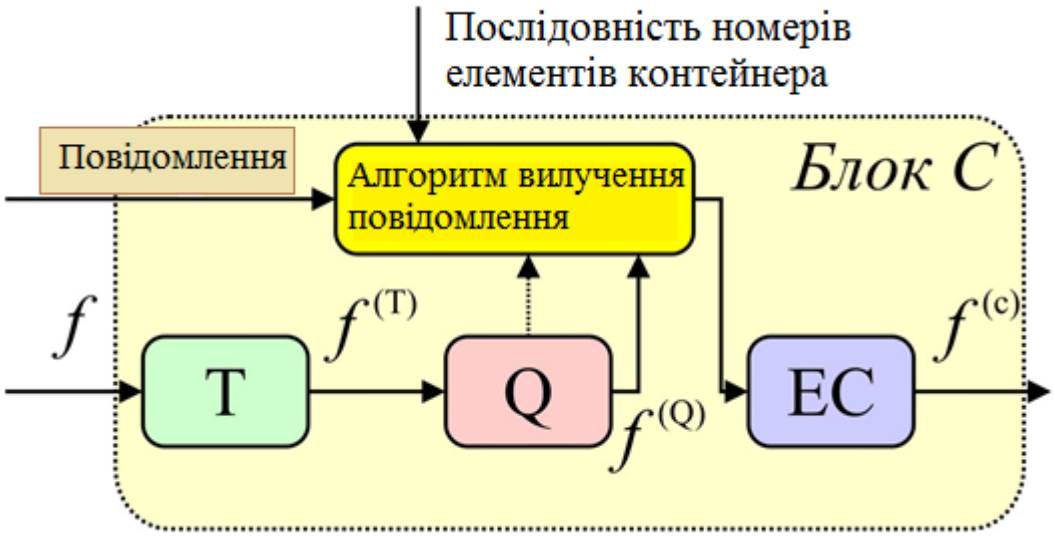

Рисунок 3.1.4. Блок С - модифікована версія кодера стандарту стиснення мультимедійних файлів

Тут же відзначимо наступне.

Аналіз конвеєра операцій, використовуваних в стандартах стиснення мультимедійних файлів, показує, що з точки зору стеганографічних додатків представляють особливий інтерес квантовані коефіцієнти частотної області мультимедійних файлів, тобто результати, отримані після етапу квантування алгоритму стиснення мультимедійних файлів.

Конвеєр операцій модифікованої версії кодера стандартів стиснення

мультимедійних файлів відрізняється від конвеєра операцій кодера стандартів стиснення мультимедійних файлів наявністю ще одного додаткового етапу, що виконується після закінчення етапу квантування, перед початком етапу ентропійного кодування. Призначення зазначеного додаткового етапу полягає в реалізації впровадження елементів повідомлення, що підлягає приховану, в квантовані коефіцієнти частотної області мультимедійного сигналу, обрані відповідно до номерами, виробленими блоком АВ. Впровадження полягає в попередньо певному зміні вибраних квантованих сигналів в залежності від значення впроваджуваного елемента повідомлення. При цьому за значеннями змінених квантованих сигналів можна однозначно відновити значення впровадженого елемента повідомлення. Цей момент дуже важливий для забезпечення процедури вилучення повідомлення законним одержувачем, при наявності у останнього відповідного ключа для генерації номерів елементів стеганограмми, в які вбудовані елементи приховуваного повідомлення.

Відзначимо, що для впровадження одного біта повідомлення використовуються n елементів контейнера, де число n вибирається таким чином, що виконується умова Pε, n = α, де Pε, n = 1 - (1 - 2ε) n, а числа α і ε встановлюються виходячи з практичних міркувань. Така вимога до числа n пов'язано з тим, що при впровадженні біта повідомлення в обрані елементи контейнера (досить часто це здійснюється по парності суми обраних n елементів контейнера [29]) один з обраних квантованих коефіцієнтів може бути змінений шляхом додавання або віднімання 1 в сторону, протилежну тому, що було зроблено стандартом стиснення мультимедійного сигналу (± 1 embedding [72]). Практичні вимоги щодо зменшення демаскирующих ознак можуть бути такими, що для зміни може знадобитися квантований коефіцієнт з частотної області мультимедійного сигналу, дрібна частина якого належить проміжку від 0,5−ε до 0,5+ε при попередньо заданому значенні параметра ε. Вищевказане вимога до числа *n* забезпечує наявність такого квантованного коефіцієнта серед *n* випадковим чином обраних з ймовірністю $P_{\varepsilon,n} = \alpha$. Тут

передбачається, що дробові частини квантованих коефіцієнтів до округлення рівномірно розподілені на полуінтервале [0; 1). Таким чином, в термінології даної роботи, в стеганографічної системі використовується булева функція впровадження-вилучення $g(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n$. При цьому змінюваних елементів контейнера при впровадженні одного біта m повідомлення дорівнює:

$$\mathrm{E}\left(\delta_g(m)\right) = 0{,}5, \qquad (3.1)$$

що відповідно вказує на те, що при впровадженні не враховуються ймовірносно-статистичні характеристики впроваджуваного повідомлення (якщо воно не піддавалося попередньому стиску). Відзначимо, що саме в цій частині можуть бути застосовані результати цієї роботи для поліпшення характеристик стеганографічної системи. До цього питання повернемося в даному параграфі нижче. Тут же зазначимо ще, що квантовані сигнали мультимедійного файлу, не використані для впровадження повідомлення, проходять етап впровадження повідомлення без змін («транзитом»). Результати виконання етапу впровадження повідомлення далі обробляються відповідно до стандартного етапом алгоритму ентропійним кодуванням (наприклад, в JPEG і MP3 кодуються по Хаффману).

Етап 3.2. Реалізація блоку С '- спеціальної модифікованої версії декодера стандарту стиснення мультимедійних сигналів.

На вхід блоку С '(рис. 3.1.2) надходять дві послідовності:

- стеганограмма;
- вихідна послідовність блоку АВ.

Результатом роботи блоку С ', тобто виходом блоку С', є повідомлення (інформація), витягнуте з стеганограмми.

Блок С '- це програмна або апаратна реалізація спеціальної модифікованої версії декодера стандарту стиснення мультимедійних сигналів. Узагальнений опис модифікованої версії декодера стандарту стиснення мультимедійних сигналів можна представити таким чином (рис. 3.1.5):

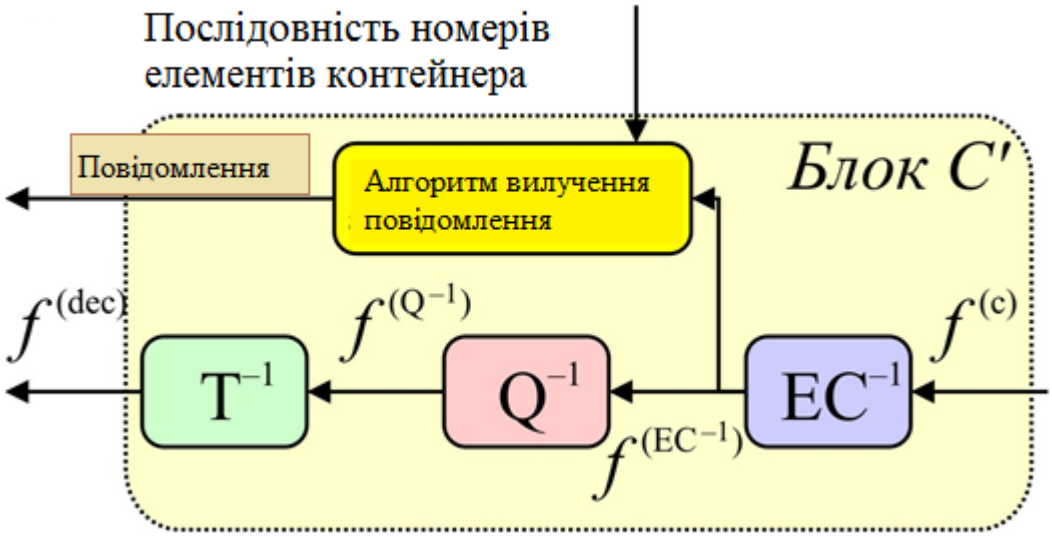

Рисунок 3.1.5. Блок C '- модифікована версія декодера стандарту стиснення мультимедійних сигналів

Конвеєр операцій модифікованої версії декодера стандарту стиснення мультимедійних сигналів відрізняється від конвеєра операцій декодера стандарту стиснення мультимедійних сигналів наявністю ще одного додаткового етапу, що виконується після закінчення етапу ентропійного декодування, перед початком етапу деквантування. Призначення зазначеного додаткового етапу полягає в реалізації процедури вилучення елементів повідомлення, прихованого в стеганограмме, з елементів заповненого контейнера, обраних відповідно до номерами, виробленими блоком АВ. Тобто кожен біт витягується повідомлення обчислюється як значення булевої функції внедренія- вилучення $g(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \ldots \oplus x_n$ при підстановці замість змінних x_1, x_2, …, x_n бітів парності квантованих коефіцієнтів в частотної області мультимедійного сигналу, обраних відповідно до номерами, виробленими блоком АВ. Однак при застосуванні досягнень даної роботи функція впровадження-вилучення може бути змінена, про що буде сказано нижче.

Етап 4. Реалізація ентропійного алгоритму впровадження повідомлення в контейнер і вилучення повідомлення з контейнера.

Тепер зупинимося на можливості поліпшення в бік зменшення середньої кількості змінюваних елементів контейнера на один біт впроваджуваного

повідомлення розглянутих вище стеганографічних алгоритмів і відповідних стеганографічних систем шляхом використання результатів даної роботи в разі, коли впроваджується повідомлення довжини l бітів є реалізацією послідовності довжини l незалежних випадкових величин:

$$\eta_1, \eta_2, \dots, \eta_l, \tag{3.2}$$

де $l \in N, P(\eta_i = 0) = p_i, P(\eta_i = 1) = q_1, p_i + q_i = 1, i = \overline{1,l}$, тобто кожен і-й біт повідомлення є результатом роботи двійкового джерела без пам'яті, який генерує 0 і 1 потактно за схемою незалежних випробувань з вірогідністю p_i и q_i, відповідно.

Основою цього поліпшення може служити заміна булевої функції впровадження-вилучення $g(x_1, x_2, \dots, x_n) = x_1 \oplus x_2 \oplus \dots \oplus x_n$ на набір булевих функцій:

$$\{f_i(x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}, \dots, x_{(t(i)-1)n+1}, \dots, x_{t(i)n} | i = \overline{1,l}\}, \tag{3.3}$$

Де функція

$$f_i(x_1, \dots, x_n, x_{n+1}, \dots x_{2n}, \dots, x_{(t(i)-1)n+1}, \dots, x_{t(i)n}) = x_1 \oplus \dots \oplus x_n)^{a_{li}}(x_{n+1} \oplus \dots \oplus x_{2n})^{a_{2i}} \dots (x_{(t(i)-1)n+1} \oplus \dots \oplus x_{t(i)n})^{a_{t(i)i}} \tag{3.4}$$

використовується як булева функція впровадження-вилучення в стеганографічний контейнер і-го біта повідомлення, і при цьому, якщо покласти

$$Y_1 = x_1 \oplus \dots \oplus x_n,$$

$$Y_2 = x_{n+1} \oplus \dots \oplus x_{2n},$$

...

$$Y_{t(i)} = x_{(t(i)-1)n+1} \oplus \dots \oplus x_{t(i)n}$$ То

$$f_i(x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}, \dots, x_{(t(i)-1)n+1}, \dots, x_{t(i)n}) = F_i(Y_1, Y_2, \dots, Y_{t(i)}), \tag{3.6}$$

І бульова функція

$$F_i(Y_1, Y_2, \dots, Y_{t(i)}) = Y_1^{a_{1i}} Y_2^{a_{2i}} \dots Y_{t(i)}^{a_{t(i)i}} \tag{3.7}$$

є в разі $0 < q_i \leq 0{,}5$ булевої функцією впровадження-вилучення субоптимального ентропійного стеганографічного алгоритму впровадження

в контейнер повідомлень, що генеруються двійковим джерелом без пам'яті, на виході якого з'являються символи 0 і 1 з ймовірностямии p_i и q_i, відповідно; (α_{1i}, …, $\alpha_{t(i)i}$) — деякий довільним чином зафіксований вектор з безлічі $V_2^{t(i)}$ всіх довічних векторів довжини t(i), де t(i) визначається з подвійної нерівності

$$\frac{1}{2^{t(i)}+1} < q_i \leq \frac{1}{2^{t(i)-1}+1} \tag{3.8}$$

У випадку $0<p_i\leq 0{,}5$

$$F_i\left(Y_1, Y_2, \dots, Y_{t(i)}\right) = Y_1^{a_{1i}} Y_2^{a_{2i}} \dots Y_{t(i)}^{a_{t(i)i}} \oplus 1 \tag{3.9}$$

І число t (i) визначається з подвійного нерівності

$$\frac{1}{2^{t(i)}+1} < p_i \leq \frac{1}{2^{t(i)-1}+1} \tag{3.10}$$

Звернемо увагу на те, що булеві функції впровадження-вилучення та відповідні булеві функції впровадження-вилучення субоптимального ентропійного стеганографічного алгоритму, певні в параграфі 2.4, мають деякі відмінності. У розділі 2.4 змінні функцій були незалежними, а кожна змінна представляє собою двійкову суму *n* незалежних змінних. Це пов'язано з тим, що в субоптимальних ентропійних стеганографічних алгоритмах в параграфі 2.4 допускається можливість зміни всіх елементів контейнера, обраних для впровадження біта повідомлення, що підлягає стеганографічної захисту. У той же самий час, як було сказано вище, дрібна частина змінюваного елемента повинна належати проміжку від 0,5−ε до 0,5+ε при попередньо заданому значенні параметра ε з ймовірністю $P_{\varepsilon,n} = \alpha$. Вимога одночасного виконання цих умов тягне за собою необхідність заміни незалежних змінних на виконавчі суми *n* незалежних змінних булевої функції впровадження-вилучення. Саме таким шляхом вдається при впровадженні повідомлення в контейнер одночасно і контролювати величину спотворення кожного змінюваного елемента контейнера, і зменшувати загальну кількість змінюваних елементів контейнера. Однак при цьому математичне сподівання кількості змінених елементів контейнера на

один біт вбудованого в контейнер повідомлення не змінюється, так як воно не залежить від значення натурального числа n.

Дійсно, в разі $0 < q_i \leq 0{,}5$ маємо:

$$A_{lf} = \left\{\left(x_1, \dots, x_n, x_{n+1}, \dots, x_{2n}, \dots, x_{(t(i)-1)n+1}, \dots, x_{t(i)n}\right) \in \mathbb{V}_2^{t(i)n} | x_{(k-1)n+1} \oplus \dots \oplus x_{kn} = a_k, k \in \{1,2, \dots, t(i)\}\right\},$$

(3.11)

$$A_{2f} = \dots = A_{t(i)f} = \varnothing$$

(3.12)

$$A_f = \left\{v \in \mathbb{V}_2^{t(i)n} | f(v) = 1\right\} = A_{1f} \cup \dots \cup A_{t(i)f} = A_{1f};$$

(3.13)

система підмножин $\{B_{1f}, \dots, B_{t(i)f}\}$ Влаштована таким чином, що: B_{kf} — це безліч, що складається з усіх векторів, що виходять з векторів безлічі A_{1f} шляхом інвертування в їх параметрах $(\alpha_1, \dots, \alpha_{t(i)})$ рівно k координат, де $k \in \{1, 2, \dots, t(i)\}$. Звідси слідує що

$$y_{lf} = |A_{lf}| = 2^{(n-1)t(i)}, y_{2f} = |A_{2f}| = 0, \dots, y_{t(i)f} = |A_{t(i)f}| = 0,$$

(3.14)

$$z_{kf} = |B_{kf}| = C_{t(i)}^k \cdot 2^{(n-1)t(i)}, k \in \{1,2, \dots, t(i)\}.$$

(3.15)

Тоді отримаємо

$$\mathbb{E}\left(\xi_f(m)\right) = \sum_{k=0}^{t(i)} k \cdot P\left(\xi_f(m) = k\right)$$
$$= \sum_{k=1}^{t(i)} k \cdot \left(P(m=0) \cdot P\left(v \in A_{kf}\right) + P(m=1) \cdot P\left(v \in B_{kf}\right)\right)$$
$$= p_i \cdot \sum_{k=1}^{t(i)} k \cdot P\left(v \in A_{kf}\right) + q_i \cdot \sum_{k=1}^{t(i)} k \cdot P\left(v \in B_{kf}\right)$$
$$= \frac{p_i}{2^{nt(i)}} \cdot \sum_{k=1}^{t(i)} k \cdot y_{kf} + \frac{q_i}{2^{nt(i)}} \cdot \sum_{k=1}^{t(i)} k \cdot z_{kf} = \frac{p_i}{2^{nt(i)}} \cdot 2^{(n-1)t(i)} + q_i / 2^{nt(i)} \cdot \sum_{k=1}^{t(i)} k \cdot C_{t(i)}^{k} \cdot 2^{(n-1)t(i)}$$
$$= \frac{p_i}{2^{t(i)}} + \frac{q_i}{2^{t(i)}} \cdot \sum_{k=1}^{t(i)} k \cdot C_{t(i)}^{k} = \frac{p_i}{2^{t(i)}} + \frac{q_i}{2^{t(i)}} \cdot t(i) \cdot 2^{t(i)-1}$$
$$= \frac{p_i}{2^{t(i)}} + t(i) \cdot \frac{q_i}{2} = \frac{1 - q_i}{2^{t(i)}} + t(i) \cdot \frac{q_i}{2}, \tag{3.16}$$

що відповідає рівності для функції fi (звернемо увагу на те, що щоб уникнути появи багатоповерхових індексів в останніх викладках в нижніх індексах формул замість fi писалося просто f).

Для випадку $0 < p_i \leq 0{,}5$ відповідні викладки проводяться аналогічно. Зауважимо також, що набір векторів

$$\left\{(a_{1i}, \dots, a_{t(i)i} | i = \overline{1, l}\right\}$$

може бути включений в безліч ключових елементів стеганографічної системи і тим самим служити ще одним потенційним напрямком розширення її ключового простору і, як наслідок, підвищення її стійкості. Далі в параграфах 3.2 і 3.3 запропонований в даному параграфі підхід для поліпшення характеристик універсальних стеганографічних систем шляхом використання

ентропійних стеганографічних алгоритмів застосований відповідно до випадків, коли повідомлення, що підлягає стеганографічної захисту, є текстом російською мовою, представленим в коді Windows-1251 (або CP1251), або англійською мовою, представленим в коді UTF-8.

Покроковий опис алгоритму впровадження повідомлення в контейнер (реалізованого жовтим блоком «Алгоритм впровадження повідомлення» на рис. 3.1.4) і алгоритму вилучення повідомлення з стеганограмми (реалізованого жовтим блоком «Алгоритм вилучення повідомлення» на рис. 3.1. 5). Вважаємо, що $0 < qi \leq 0{,}5$ За умов, прийнятих в даній роботі щодо процесу впровадження повідомлення ($m1$, $m2$, …, ml) із l бітів в контейнер, біти повідомлення вбудовуються послідовно і незалежно один від одного. Впровадження і-го біта mi (i=1, l) повідомлення полягає у виконанні процедури, що складається з наступних кроків.

Крок 1. По згенеровуваним блоком АВ номерами $j_1^{(i)}, j_2^{(i)}, \dots, j_{t(i)\cdot n}^{(i)}$ елементів контейнера, які використовуються для впровадження біта mi, визначаються самі елементи $g_1^{(i)}, g_2^{(i)}, \dots., g_{t(i)\cdot n}^{(i)}$ і відповідні їм дійсні числа$F_1^{(i)}, F_2^{(i)}, \dots, F_{t(i)\cdot n}^{(i)}$, де $g_k^{(i)}$ — элемент з номером $j_k^{(i)}$, a $F_k^{(i)}$ — дійсне число, з якого шляхом округлення згідно стандартним алгоритмом стиснення отриманий елемент$g_k^{(i)}$. Позначимо молодші біти $g_k^{(i)}$ через $b_k^{(i)}$, $k = 1, t(i) \cdot n$.

Крок 2. Обчислюється значення функції впровадження-вилучення $f_1(x_1, x_2, \dots, x_{t(i)\cdot n}$ при $x_1 = \left|b_1^{(i)}, x_2 = b_2^{(i)}, \dots, x_{t(i)\cdot n} = b_{t(i)\cdot n}^{(i)}\right.$.

Якщо $f_1(b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}) = m_i$, то перехід до кроку 7.

Якщо $f_i\left(b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}\right) = 1$ и $m_i = 0$, то перехід до кроку 3.

Якщо $f_i\left(b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}\right) = 0$ и $m_i = 1$, то перехід до кроку 5.

Крок 3. Серед елементов $g_1^{(i)}, g_2^{(i)}, \dots., g_{t(i)\cdot n}^{(i)}$ находим такий елемент $g_k^{(i)}$,

що дробова частина числа $F_k^{(i)}$ найбільш близька до 0,5.

Крок 4. Отриманий на кроці 3 елемент шляхом додавання або віднімання 1 (± 1 embedding [72]) змінюємо в бік, протилежний тому, що було зроблено застосовуваним стандартним алгоритмом стиснення мультимедійного сигналу при округленні в процесі виконання процедури квантування. Перехід до кроку 7.

Крок 5. Для вектора $(b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})})$ обчислюємо блокові (по n елементів у блоці) суми по модулю 2 і отримуємо вектор $B_{1i}, \dots., B_{t(i)i}$. Отриманий вектор порівнюємо покоординатно з вектором $a_{1i}, \dots., a_{t(i)i}$. Кількість відрізняються координат позначимо через $k(i)$, а їх номера — як $s_1^{(\mathrm{i})}, s_2^{(\mathrm{i})}, \dots, s_{k(\mathrm{i})}^{(\mathrm{i})})$. Далі в цих номерах верхній індекс і для стислості будемо опускати і писати $s_1, \dots, s_{k(i)}$.

Крок 6. Для кожного $\mu \in \{1, \dots, k(i)\}$ серед елементів $g^{(i)}_{(g_\mu - 1)\cdot n + z_\mu}$, що дрібна частина відповідного числа $F^{(i)}_{(g_\mu - 1)\cdot n + z_\mu}$ найбільш близька до 0,5. Цей елемент шляхом додавання або віднімання 1 (± 1 embedding [72]) змінюємо в бік, протилежний тому, що було зроблено застосовуваним стандартним алгоритмом стиснення мультимедійного сигналу при округленні в процесі виконання процедури квантування. Крок 7. Процедура впровадження біта m_i завершена.

Молодші біти елементів контейнера, задіяних для вбудовування біта m_i, на кроці 7 позначимо$\widetilde{b_1^{(i)}}, \widetilde{b_2^{(i)}}, \dots, \widetilde{b_{t(i)\cdot n}^{(i)}}$. Тоді очевидно, що справедливо рівність $f_i\left(\widetilde{b_1^{(i)}}, \widetilde{b_2^{(i)}}, \dots, \widetilde{b_{t(i)\cdot n}^{(i)}}\right) = m_i$.

Нарешті, викладемо покроково алгоритм вилучення повідомлення з стеганограмми.

За умов, прийнятих в даній роботі щодо процесу вилучення

повідомлення $(m_1, m_2, \dots, m_l)$ із l бітів з стеганограмми, біти повідомлення витягуються послідовно і незалежно один від одного. Витяг і-го біта $m_i (i = \overline{1, l})$ повідомлення полягає у виконанні процедури, що складається з наступних кроків.

Крок 1. За згенерував блоком АВ номерами$f_1^{(\mathrm{i})}, f_2^{(\mathrm{i})}, \dots, f_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}$ елементів контейнера, які використовуються для впровадження біта *mi*, визначаються самі елементи$g_1^{(\mathrm{i})}, g_2^{(\mathrm{i})}, \dots, g_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}$, де $g_k^{(i)}$ — елемент з номером $j_k^{(i)}$. Позначимо їх молодші біти через$b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})}$.

Крок 2. Біт повідомлення m_i знаходимо шляхом обчислення значення функції впровадження-вилучення $f_i(x_1, x_2, \dots, x_{t(i)\cdot n})$ при $x_1 = b_1^{(i)}, x_2 = b_2^{(i)}, \dots, x_{t(i)\cdot n} = b_{t(i)\cdot n}^{(i)}$, т. е. $m_i = f_i(b_1^{(\mathrm{i})}, b_2^{(\mathrm{i})}, \dots, b_{t(\mathrm{i})\cdot n}^{(\mathrm{i})})$.

Крок 3. Процедура вилучення біта m_i завершена.

Блок-схеми описаних процедур впровадження та вилучення представлені на рис. 3.1.13 і 3.1.14, відповідно.

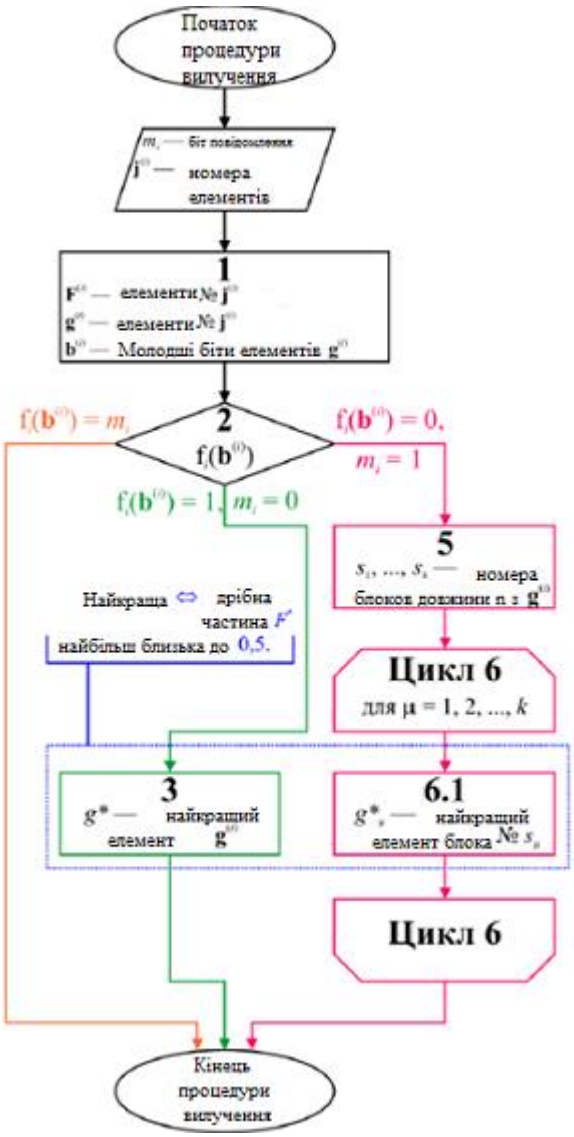

Рисунок 3.1.6. Процедура впровадження і-го біта mi

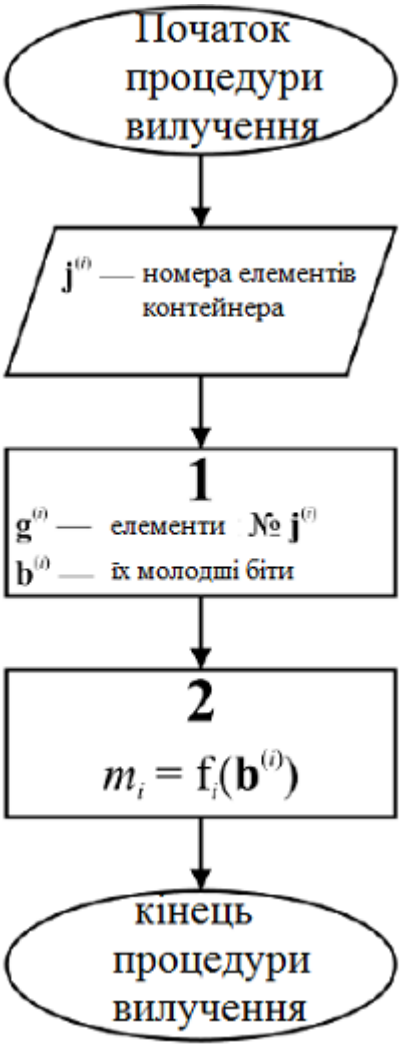

Рисунок 3.1.7. Щоб вийняти і-го біта mi

Підсумовуючи даний параграф, можна констатувати, що отримана сукупність методів і прийомів побудови ентропійних стеганографічних систем з контейнерами, що представляють собою цифрові мультимедійні файли. Це і становить зміст методики, заявленої в заголовку параграфа.

3.2. Значення параметрів і чисельні характеристики субоптимального стеганографічного алгоритму для захисту повідомлень

Windows-1251, або CP1251, - стандартна кодова сторінка для всіх російських версій Microsoft Windows. Перша («нижня») половина таблиці збігається з ASCII, російський алфавіт має коди C0-DF («А» - «Я»), E0-FF («а» - «я»), а також A8 (буква «Е») і B8 (буква «е»). Windows-1251 вигідно відрізняється від інших 8-бітних кириличних кодувань наявністю практично

всіх символів, що використовуються в російській типографике. За даними статистики http://gs.statcounter.com за 2012-2013 роки, майже 90% російських інтернет-користувачів використовують одну з версій операційної системи Windows. При поданні повідомлень російською мовою з урахуванням розділових знаків в коді Windows-1251 значущими є всі вісім розрядів кодових слів. Виходячи з цього, при вирішенні завдання стеганографічної захисту повідомлення російською мовою в коді Windows-1251 може бути представлено у вигляді двійковій послідовності:

$$m_{71}m_{61}m_{51}m_{41}m_{31}m_{21}m_{11}m_{01}m_{72}m_{62}m_{52}m_{42}m_{32}m_{22}m_{12}m_{02}\ldots$$

$$\ldots m_{7i}m_{6i}m_{5i}m_{4i}m_{3i}m_{2i}m_{1i}m_{0i}\ldots m_{7k}m_{6k}m_{5k}m_{4k}m_{3k}m_{2k}m_{1k}m_{0k}$$

довжина $8k$, де k — довжина повідомлення в символах, m_{ri} - двоичное значення r-го розряду кодового слова, відповідного і-му символу повідомлення,
$r \in \{0,1,\ldots,7\}, i \in \{1,2,\ldots,k\}$. Будемо вважати, що такі повідомлення задовольняють таким умовам:
- розряди кодових слів незалежні в сукупності, тобто можливі ймовірносно-статистичні залежності між розрядами не враховуються;
- двоичная послідовність$\{m_{ri}m_{r2}\ldots m_{rk}\}$

довжини k, складена з значень r-го розряду кодового слова кожного символу повідомлення, породжена двійковим джерелом без пам'яті, генеруючим 0 з ймовірністю p_r, а 1 – із ймовірністю $q_r, p_r + q_r = 1$, де $r \in \{0,1,\ldots.,7\}$.

Зроблені припущення спрощують мовну модель до того ступеня, що в рамках цієї моделі можна ефективно застосувати результати, отримані в даній роботі.

Параметри p_r і q_r(где $r \in \{0,1,\ldots,7\}$) довічного джерела без пам'яті визначені шляхом обробки текстів повідомлень російською мовою (тексти взяті з літературних творів: Лев Толстой «Анна Кареніна», Сергій Лук'яненко

«Нічний дозор»). Відповідні результати представлені у вигляді такої таблиці (табл. 3.2.1).

Таблиця 3.2.1.

Параметри джерел повідомлень російською мовою

r	p_r	q_r
0	0,64	0,36
1	0,61	0,39
2	0,60	0,40
3	0,58	0,42
4	0,77	0,23
5	0,03	0,97
6	0,23	0,77
7	0,23	0,77

Маючи ці вихідні дані, поставимо завдання обчислення наступних величин щодо повідомлень російською мовою:

- математичне очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального алгоритму без попереднього стиснення повідомлення і відповідний варіант двійковій функції впровадження-вилучення;
- математичне очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при його попередньому стисненні за допомогою асимптотично оптимального рівномірного блокового коду і відповідний варіант двійковій функції впровадження-вилучення. При цьому завдання контролю величини спотворення елемента контейнера при його зміні не розглядається.

Розглянемо випадок r = 0.

з рівності$q_0 = 0{,}36$ і подвійної нерівності маємо

$$\frac{1}{2^{t_0}+1} < 0{,}36 \leq \frac{1}{2^{t_0-1}+1}.$$

що тягне справедливість рівності t0 = 1. Тоді з рівності (2.4.23) для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму маємо:

$$E\left(\delta_f(m_{0i})\right) = \frac{1-q_0}{2^{t_0}} + \frac{t_0 q_0}{2} = \frac{1-q_0}{2} + \frac{q_0}{2} = 0{,}5 \quad \text{где } i \in \{1,2,\dots,k\}.$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E(\delta_f(m_{0i}))}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%.$$

Двійкова функція впровадження-вилучення має наступний вигляд:

$$f_{0i}(x_1) = x_1^{a_1}, \text{ где } a_1 \in \{0,1\}, i \in \{1,2,\dots,k\}.$$

При впровадженні в контейнер повідомлення з попередніми стисненням за допомогою асимптотично оптимального рівномірного блокового коду для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення маємо:

$$E\left(\delta_f(m_{0i})\right) = 0{,}5H = -0{,}5 \cdot (p_0 \cdot \log_2 p_0 + q_0 \cdot \log_2 q_0) = 0{,}5 \cdot (0{,}64 \cdot \log_2 0{,}64 + 0{,}36 \cdot \log_2 0{,}36) = 0{,}5 \cdot 0{,}943 = 0{,}4715.$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_f(m_{0i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}943}{0{,}5} \cdot 100\% = 5{,}7\%$$

Двійкова функція впровадження-вилучення для кожного біта стисненого повідомлення має наступний варіант загального вигляду:

$$g(x_1) = x_1^{a_1},$$

де $a_1 \epsilon \{0,1\}$.

Розглянемо випадок $r = 1$.

З рівності q1 = 0,39 і подвійного нерівності маємо

$$\frac{1}{2^{t_1}+1} < 0{,}39 \leq \frac{1}{2^{t_1-1}+1}$$

що тягне справедливість рівності t1 = 1. Тоді з рівності для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму маємо:

$$E\left(\delta_f(m_{0i})\right) = \frac{1-q_1}{2} + \frac{q_1}{2} = 0{,}5, \text{ где } i \epsilon \{1,2,\dots,k\}.$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_f(m_{0i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%$$

Двійкова функція впровадження-вилучення має наступний вигляд:

$$f_{1i}(x_1) = x_1^{a_1}, \text{ де } a_1 \epsilon \{0,1\}, i \epsilon \{1,2,\dots,k\}.$$

При впровадженні в контейнер повідомлення з попередніми стисненням

за допомогою асимптотично оптимального рівномірного блокового коду для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення маємо:

$$E\left(\delta_f(m_{0i})\right) = 0{,}5H = -0{,}5 \cdot (p_1 \cdot \log_2 p_1 + q_1 \cdot \log_2 q_1)$$
$$= -0{,}5 \cdot (0{,}61 \cdot \log_2 0{,}61 + 0{,}39 \cdot \log_2 0{,}39) = 0{,}5 \cdot 0{,}965$$
$$= 0{,}4825$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_f(m_{1i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}965}{0{,}5} \cdot 100\% = 3{,}7\%$$

Двійкова функція впровадження-вилучення для кожного біта стисненого повідомлення має наступний варіант загального вигляду:

$$g(x_1) = x_1^{a_1},\ \textit{де}\ a_1 \in \{0,1\}.$$

Аналогічні обчислення можна провести і для інших розрядів кодового слова. Наведемо відповідні результати в короткій формі.

Випадок $r = 2$.

$$q_2 = 0{,}40, \frac{1}{2^{t_2}+1} < 0{,}40 \le \frac{1}{2^{t_2-1}+1} \Longrightarrow t_2 = 1 \Longrightarrow E\left(\delta_f(m_{2i})\right)$$
$$= \frac{1-q_2}{2} + \frac{q_2}{2} = 0{,}5, i \in \{1,2,\dots,k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{2i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%.$$

$$f_{2i}(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}, i \in \{1,2,\dots,k\}.$$

$$E\left(\delta_f(m_{2i})\right) = 0{,}5H = -0{,}5 \cdot (p_1 \cdot \log_2 p_2 + q_2 \cdot \log_2 q_2) = -0{,}5 \cdot (0{,}60 \cdot \log_2 0{,}60 + 0{,}40 \cdot \log_2 0{,}40) = 0{,}5 \cdot 0{,}971 = 0{,}4855.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{2i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}971}{0{,}5} \cdot 100\% = 2{,}9\%.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Випадок r = 3.

$$q_3 = 0{,}42, \frac{1}{2^{t_3}+1} < 0{,}42 \le \frac{1}{2^{t_3-1}+1} \Longrightarrow t_3 = 1 \Longrightarrow E\left(\delta_f(m_{3i})\right) = \frac{1-q_3}{2} + \frac{q_3}{2} = 0{,}5, i \in \{1,2,\dots,k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{3i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%.$$

$$f_{3i}(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}, i \in \{1,2,\dots,k\}.$$

$$E\left(\delta_f(m_{3i})\right) = 0{,}5H = -0{,}5 \cdot (p_3 \cdot \log_2 p_3 + q_3 \cdot \log_2 q_3) = -0{,}5 \cdot (0{,}58 \cdot \log_2 0{,}58 + 0{,}42 \cdot \log_2 0{,}42) = 0{,}5 \cdot 0{,}981 = 0{,}4905.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{3i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}981}{0{,}5} \cdot 100\% = 1{,}9\%.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Випадок *r* = 4.

$$q_4 = 0{,}23, \frac{1}{2^{t_4}+1} < 0{,}23 \le \frac{1}{2^{t4-1}+1} \Longrightarrow t_4 = 2 \Longrightarrow$$

$$E\left(\delta_f(m_{4i})\right) = \frac{1-q_4}{2^{t_4}} + \frac{t_4 q_4}{2} = \frac{1-q_4}{2^2} + \frac{2q_4}{2} = 0{,}4225, i \in \{1,2,\dots,k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{4i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}4225}{0{,}5} \cdot 100\% = 15{,}5\%.$$

$$f_{4i}(x_1, x_2) = x_1^{a_1} x_2^{a_2}, a_1, a_2 \in \{0,1\}, i \in \{1,2,\dots,k\}.$$

$$E\left(\delta_f(m_{4i})\right) = 0{,}5H = -0{,}5 \cdot (p_4 \cdot \log_2 p_4 + q_4 \cdot \log_2 q_4) = -0{,}5 \cdot$$
$$(0{,}77 \cdot \log_2 0{,}77 + 0{,}23 \cdot \log_2 0{,}23) = 0{,}5 \cdot 0{,}778 = 0{,}3890.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{4i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}778}{0{,}5} \cdot 100\% = 22{,}2\%.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Випадок $r = 5$.

Тут ймовірність одиниці $q_5 = 0{,}97$ більше, ніж ймовірність нуля $p_5 = 0{,}03$. Тому для обчислень даного розряду замість ймовірності одиниці підставляємо ймовірність нуля. З урахуванням цього маємо:

$$\frac{1}{2^{t_5}+1} < 0{,}03 \le \frac{1}{2^{t_5-1}+1} \Longrightarrow t_5 = 6 \Longrightarrow$$

$$E\left(\delta_f(m_{5i})\right) = \frac{1-q_5}{2^{t_5}} + \frac{t_5 q_5}{2} \approx 0{,}1052, i \in \{1,2,\dots,k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{5i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}1052}{0{,}5} \cdot 100\% \approx 79\%.$$

Двійкова функція впровадження-вилучення з урахуванням того, що в цьому розряді ймовірність нуля менше ймовірності одиниці) для і-го біта повідомлення має наступний вигляд:

$$f_{5i}(x_1, x_2, x_3, x_4, x_5, x_6) = x_1^{a_1} x_2^{a_2} x_3^{a_3} x_4^{a_4} x_5^{a_5} x_6^{a_6} \oplus 1,$$

Де $a_1, a_2, \dots, a_6 \in \{0,1\}\ i \in \{1,2, \dots, k\}$.

$$E\left(\delta_f(m_{5i})\right) = 0{,}5H = -0{,}5 \cdot (p_5 \cdot \log_2 p_5 + q_5 \cdot \log_2 q_5) = -0{,}5 \cdot (0{,}03 \cdot \log_2 0{,}03 + 0{,}97 \cdot \log_2 0{,}97) = 0{,}5 \cdot 0{,}194 = 0{,}0970.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{5i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}194}{0{,}5} \cdot 100\% = 80{,}6\%.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Випадок r = 6.

Тут ймовірність одиниці q_6 = 0,77 більше, ніж ймовірність нуля p_6 = 0,23. Тому для обчислень даного розряду замість ймовірності одиниці підставляємо ймовірність нуля. З урахуванням цього маємо:

$$\frac{1}{2^{t_6}+1} < 0{,}23 \leq \frac{1}{2^{t_6-1}+1} \Longrightarrow t_6 = 2 \Longrightarrow$$

$$E\left(\delta_f(m_{6i})\right) = \frac{1-q_6}{2^{t_6}} + \frac{t_6 q_6}{2} = 0{,}4225, i \in \{1,2, \dots, k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{6i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}4225}{0{,}5} \cdot 100\% = 15{,}5\%.$$

$$f_{6i}(x_1, x_2) = x_1^{a_1} x_2^{a_2} \oplus 1, a_1, a_2 \in \{0,1\}\ i \in \{1,2, \dots, k\}.$$

$$E\left(\delta_f(m_{6i})\right) = 0{,}5H = -0{,}5 \cdot (p_6 \cdot \log_2 p_6 + q_6 \cdot \log_2 q_6) = -0{,}5 \cdot (0{,}23 \cdot \log_2 0{,}23 + 0{,}77 \cdot \log_2 0{,}77) = 0{,}5 \cdot 0{,}778 = 0{,}3890.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Випадок r = 7.

Тут ймовірність одиниці q_6 = 0,77 більше, ніж ймовірність нуля p_6 = 0,23. Тому для обчислень даного розряду замість ймовірності одиниці підставляємо

ймовірність нуля. З урахуванням цього маємо:

$$\frac{1}{2^{t_7}+1} < 0{,}23 \le \frac{1}{2^{t_7-1}+1} \Longrightarrow t_7 = 2 \Longrightarrow$$

$$E\left(\delta_f(m_{7i})\right) = \frac{1-q_7}{2^{t_7}} + \frac{t_7 q_7}{2} = 0{,}4225, i \in \{1,2,\dots,k\}.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{7i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}4225}{0{,}5} \cdot 100\% = 15{,}5\%.$$

$$f_{7i}(x_1, x_2) = x_1^{a_1} x_2^{a_2} \oplus 1, a_1, a_2 \in \{0,1\}\, i \in \{1,2,\dots,k\}.$$

$$E\left(\delta_f(m_{7i})\right) = 0{,}5H = -0{,}5 \cdot (p_7 \cdot \log_2 p_7 + q_7 \cdot \log_2 q_7) = -0{,}5 \cdot$$
$$(0{,}23 \cdot \log_2 0{,}23 + 0{,}77 \cdot \log_2 0{,}77) = 0{,}5 \cdot 0{,}778 = 0{,}3890.$$

$$\frac{0{,}5 - E\left(\delta_f(m_{7i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}778}{0{,}5} \cdot 100\% = 22{,}2\%.$$

$$g(x_1) = x_1^{a_1},\ a_1 \in \{0,1\}.$$

Позначимо через Vr відміну в процентах в сторону зменшення математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта r -го розряду повідомлення. через W_rпозначимо відміну в процентах в сторону зменшення (т. е. виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта r-го розряду повідомлення, де $r \in \{0,1,\dots,7\}$. Для порівняння двох способів стеганографічного захисту повідомлення (способу

захисту без стиснення повідомлення із застосуванням субоптимального стеганографічного алгоритму, представленого в параграфі 2.4, і способу захисту з попередніми стисненням асимптотично оптимальним рівномірним блоковим кодом) по математичному очікуванню числа змінних елементів контейнера на один біт повідомлення у зручний спосіб деякі результати вищенаведених обчислень представити у вигляді такої таблиці (3.2.2):

Таблиця 3.2.2.

Кількісні характеристики порівнюваних алгоритмів

r	$\mathbb{E}(\xi_{\mathrm{c}}(m_{ri}))$	$\mathbb{E}(\xi_{\mathrm{f}}(m_{ri}))$	$\mathbb{E}(\xi_{\mathrm{g}}(m_{ri}))$	V_r	W_r	$W_r - V_r$
0	0,5	0,5	0,4715	0%	5,7%	5,7%
1	0,5	0,5	0,4825	0%	3,5%	3,5%
2	0,5	0,5	0,4855	0%	2,9%	2,9%
3	0,5	0,5	0,4905	0%	1,9%	1,9%
4	0,5	0,4225	0,3890	15,5%	22,2%	6,7%
5	0,5	0,1052	0,0970	79,0%	80,6%	1,6%
6	0,5	0,4225	0,3890	15,5%	22,2%	6,7%
7	0,5	0,4225	0,3890	15,5%	22,2%	6,7%

У таблиці 3.2.2 через $E(\delta_c(m_{ri})), E\left(\delta_f(m_{ri})\right)$ і $E(\delta_g(m_{ri}))$ позначені математичні очікування числа змінних елементів контейнера на один біт повідомлення відповідно при його впровадженні без стиснення за ознакою парності суми елементів контейнера, за допомогою субоптимального ентропійного стеганографічного алгоритму (див. параграф 2.4) і з попередніми стисненням за допомогою асимптотично оптимального рівномірного блокового коду.

Наочно порівняльні характеристики розглянутих алгоритмів представлені у вигляді такої діаграми (рис. 3.2.1).

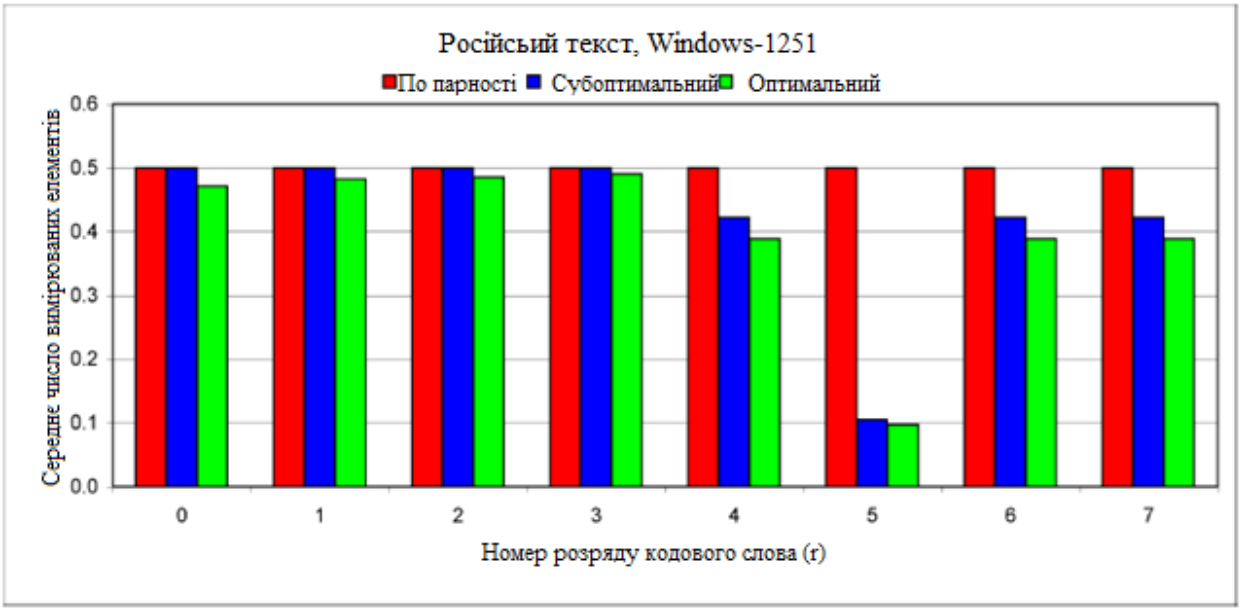

Рисунок 3.2.1. Діаграма значень середнього числа змінних елементів контейнера на один біт російськомовного повідомлення для трьох стеганографічних алгоритмів

Тут червоні стовпчики відповідають впровадженню без стиснення за ознакою парності суми елементів контейнера, сині - за допомогою субоптимального ентропійного стеганографічного алгоритму, зелені - з попередніми стисненням за допомогою асимптотично оптимального рівномірного блокового коду.

3.3. Значення параметрів і чисельні характеристики субоптимального стеганографічного алгоритму для захисту повідомлень англійською мовою, представлених в двійковому коді UTF-8

При поданні повідомлень англійською мовою з урахуванням розділових знаків в коді UTF-8 значущими є тільки перші сім розрядів кодових слів. Виходячи з цього при вирішенні завдання стеганографічної захисту таких повідомлень має сенс обмежитися розглядом для кожного символу тексту (літера або розділовий знак) тільки перших семи розрядів відповідного кодового слова. Отже, при такому обмеженні повідомлення англійською мовою в коді UTF-8 може бути представлено у вигляді двійковій послідовності

$$m_{61}m_{51}m_{41}m_{31}m_{21}m_{11}m_{01}m_{62}m_{52}m_{42}m_{32}m_{22}m_{12}m_{02}\ \ldots$$

$$\dots m_{6i}m_{5i}m_{4i}m_{3i}m_{2i}m_{1i}m_{0i} \dots m_{6k}m_{5k}m_{4k}m_{3k}m_{2k}m_{1k}m_{0k}$$

довжини 7k, де k - довжина повідомлення в символах, m_{ri} - бінарне значення r-го розряду кодового слова, відповідного і-му символу повідомлення, $r \in \{0,1, \dots ,6\}, i \in \{1,2, \dots , k\}$. Будемо вважати, що такі повідомлення задовольняють таким умовам:

- розряди кодових слів незалежні в сукупності, тобто можливі ймовірносно-статистичні залежності між розрядами не враховуються;
- двоичная послідовність

$$\{m_{ri}m_{r2} \dots m_{rk}\}$$

довжини k, складена з значень r-го розряду кодового слова кожного символу повідомлення, породжена двійковим джерелом без пам'яті, генеруючим 0 з ймовірністю p_r, а 1 – з ймовірністю $q_r, p_r + q_r = 1,$ де $r \in \{0,1, \dots ,6\}$.

Зроблені припущення спрощують мовну модель до того ступеня, що в рамках цієї моделі можна ефективно застосувати результати, отримані в даній роботі.

Параметри p_r и q_r(где $r \in \{0,1, \dots ,6\}$) довічного джерела без пам'яті визначені шляхом обробки текстів повідомлень англійською мовою (тексти взяті з літературних творів: Mark Twain «Tom Sawyer», Herbert Wells «The Time Machine»). Відповідні результати представлені у вигляді такої таблиці (табл. 3.3.1).

Таблиця 3.3.1.

Параметри джерел повідомлень англійською мовою

r	p_r	q_r
0	0,56	0,44
1	0,68	0,32
2	0,52	0,48
3	0,68	0,32
4	0,75	0,25
5	0,03	0,97
6	0,23	0,77

Маючи ці вихідні дані, поставимо завдання обчислення наступних величин щодо повідомлень англійською мовою:

- математичне очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального алгоритму без попереднього стиснення повідомлення і відповідний варіант двійковій функції впровадження-вилучення;
- - математичне очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при його попередньому стисненні за допомогою асимптотично оптимального рівномірного блокового коду і відповідний варіант двійковій функції впровадження-вилучення. При цьому завдання контролю величини спотворення елемента контейнера при його зміні не розглядається.

Розглянемо випадок r = 0.

з рівності q_0 = 0,44 і подвійної нерівності маємо

$$\frac{1}{2^{t_0}+1} < 0{,}44 \leq \frac{1}{2^{t_0-1}+1},$$

що тягне справедливість рівності $t_0 = 1$. Тоді для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму маємо:

$$E\left(\delta_f(m_{0i})\right) = \frac{1-q_0}{2^{t_0}} + \frac{t_0 q_0}{2} = \frac{1-q_0}{2} + \frac{q_0}{2} = 0{,}5, i \in \{1,2,\dots,k\}.$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E(\delta_f(m_{0i}))}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%.$$

Двійкова функція впровадження-вилучення має наступний вигляд:

$$f_{0i}(x_1) = x_1^{a_1}, \text{ де } a_1 \in \{0,1\}, i \in \{1,2,\dots,k\}.$$

При впровадженні в контейнер повідомлення з попередніми стисненням за допомогою асимптотично оптимального рівномірного блокового коду для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення маємо:

$$E\left(\delta_f(m_{0i})\right) = 0{,}5H = -0{,}5 \cdot (p_0 \cdot \log_2 p_0 + q_0 \cdot \log_2 q_0) = -0{,}5 \cdot$$

$$(0{,}56 \cdot \log_2 0{,}56 + 0{,}44 \cdot \log_2 0{,}44) = 0{,}5 \cdot 0{,}998 = 0{,}4950.$$

Відмінність у відсотках в сторону зменшення (т. Е. Виграш) математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_g(m_{0i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}990}{0{,}5} \cdot 100\% = 1\%$$

Двійкова функція вбудовування – вилучення для кожного біта стисненого повідомлення має наступний варіант класичного зразка:

$$g(x_1) = x_1^{\alpha_1}$$

де $\alpha_1 \in \{0,1\}$

Розглянемо випадок r=1.

З рівності $q_1 = 0{,}32$ і подвійної нерівності маємо

$$\frac{1}{2^{t_1}+1} < 0{,}32 \leq \frac{1}{2^{t_1-1}+1}$$

що тягне справедливість рівності $t_1 = 2$. Тоді з рівності для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму маємо:

$$E\left(\delta_f(m_{1i})\right) = \frac{1-q_1}{2^{t_1}} + \frac{t_1 q_1}{2} = 0{,}49$$

де $i \in \{1,2,\dots,k\}$

Відмінність у відсотках в сторону зменшення математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_f(m_{1i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}49}{0{,}5} \cdot 100\% = 2\%$$

Двійкова функція вбудовування-вилучення має наступний вигляд:

$$f_{1i}(x_1, x_2) = x_1^{\alpha_1} x_2^{\alpha_2}$$

Де $\alpha_1, \alpha_2 \in \{0,1\}$, $i \in \{1,2,\dots,k\}$

При впровадженні в контейнер повідомлення з попереднім стисненням за допомогою асимптотично оптимального рівномірного блокового коду для математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення маємо:

$$E\left(\delta_g(m_{1i})\right) = 0{,}5H = -0{,}5 \cdot (p_1 \cdot \log_2 p_1 + q_1 \cdot \log_2 q_1) =$$

$$= -0{,}5 \cdot (0{,}68 \cdot \log_2 0{,}68 + 0{,}32 \cdot \log_2 0{,}32) = 0{,}5 \cdot 0{,}904 = 0{,}4520$$

Відмінність у відсотках в сторону зменшення математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта повідомлення, так само

$$\frac{0{,}5 - E\left(\delta_g(m_{1i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}904}{0{,}5} \cdot 100\% = 9{,}6\%$$

Двійкова функція впровадження-вилучення для кожного біта стисненого повідомлення має наступний варіант загального вигляду:

$$g(x_1) = {x_1}^{\alpha_1}$$

Аналогічні обчислення можна провести і для інших розрядів кодового слова. Наведемо відповідні результати в короткій формі.

Випадок r=2

$$q_2 = 0{,}48, \frac{1}{2^t + 1} < 0{,}48 \leq \frac{1}{2^{t-1} + 1} \rightarrow t_2 \rightarrow 1 \rightarrow 0$$

$$E\left(\delta_f(m_{2i})\right) = \frac{1 - q_2}{2} + \frac{q_2}{2} = 0{,}5, i \in \{1{,}2, \ldots, k\}$$

$$\frac{0{,}5 - E\left(\delta_g(m_{1i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5}{0{,}5} \cdot 100\% = 0\%$$

$$E\left(\delta_f(m_{2i})\right) = 0{,}5H = -0{,}5 \cdot (p_2 \cdot \log_2 p_2 + q_2 \cdot \log_2 q_2) =$$

$$= -0{,}5 \cdot (0{,}52 \cdot \log_2 0{,}52 + 0{,}48 \cdot \log_2 0{,}48) = 0{,}5 \cdot 0{,}999 = 0{,}4995$$

$$\frac{0{,}5 - E\left(\delta_g(m_{1i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}999}{0{,}5} \cdot 100\% = 0{,}1\%$$

Двійкова функція вбудовування і з урахуванням того, що в цьому розряді вірогідність 0 менше вірогідності 1 для *i*-го біта повідомлення має наступний вигляд:

$$f_{5i}(x_1, x_2, x_3, x_4, x_5, x_6) = {x_1}^{a_1}{x_2}^{a_2}{x_3}^{a_3}{x_4}^{a_4}{x_5}^{a_5}{x_6}^{a_6}$$

Випадок r=3

$$q_3 = 0{,}48, \frac{1}{2^t + 1} < 0{,}48 \leq \frac{1}{2^{t-1} + 1} \rightarrow t_2 \rightarrow 2$$

$$E\left(\delta_f(m_{3i})\right) = \frac{1 - q_3}{2} + \frac{q_3}{2} = 0{,}5, i \in \{1,2, \dots, k\}$$

$$\frac{0{,}5 - E\left(\delta_g(m_{3i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}49}{0{,}5} \cdot 100\% = 2\%$$

$$E\left(\delta_f(m_{3i})\right) = 0{,}5H = -0{,}5 \cdot (p_3 \cdot \log_2 p_3 + q_3 \cdot \log_2 q_3) =$$

$$= -0{,}5 \cdot (0{,}68 \cdot \log_2 0{,}68 + 0{,}32 \cdot \log_2 0{,}32) = 0{,}5 \cdot 0{,}904 = 0{,}4520$$

$$\frac{0{,}5 - E\left(\delta_g(m_{3i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}904}{0{,}5} \cdot 100\% = 9{,}6\%$$

Випадок r=4

$$q_4 = 0{,}25, \frac{1}{2^t + 1} < 0{,}25 \leq \frac{1}{2^{t-1} + 1} \rightarrow t_4 \rightarrow 2$$

$$E\left(\delta_f(m_{4i})\right) = \frac{1 - q_4}{2} + \frac{q_4}{2} = 0{,}5, i \in \{1,2, \dots, k\}$$

$$\frac{0{,}5 - E\left(\delta_g(m_{4i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}4375}{0{,}5} \cdot 100\% = 12{,}5\%$$

$$E\left(\delta_f(m_{4i})\right) = 0{,}5H = -0{,}5 \cdot (p_4 \cdot \log_2 p_4 + q_4 \cdot \log_2 q_4) =$$

$$= -0{,}5 \cdot (0{,}75 \cdot \log_2 0{,}75 + 0{,}25 \cdot \log_2 0{,}25) = 0{,}5 \cdot 0{,}811 = 0{,}4055$$

$$\frac{0{,}5 - E\left(\delta_g(m_{4i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}811}{0{,}5} \cdot 100\% = 18{,}9\%$$

Випадок r=5

$$q_5 = 0{,}03, \frac{1}{2^t + 1} < 0{,}03 \leq \frac{1}{2^{t-1} + 1} \rightarrow t_4 \rightarrow 2$$

$$E\left(\delta_f(m_{5i})\right) = \frac{1 - q_5}{2} + \frac{q_5}{2} = 0{,}1052, i \in \{1,2, \dots, k\}$$

$$\frac{0{,}5 - E\left(\delta_g(m_{5i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}1052}{0{,}5} \cdot 100\% = 79\%$$

$$E\left(\delta_f(m_{5i})\right) = 0{,}5H = -0{,}5 \cdot (p_5 \cdot \log_2 p_5 + q_5 \cdot \log_2 q_5) =$$

$$= -0{,}5 \cdot (0{,}03 \cdot \log_2 0{,}03 + 0{,}97 \cdot \log_2 0{,}97) = 0{,}5 \cdot 0{,}194 = 0{,}0970$$

$$\frac{0{,}5 - E\left(\delta_g(m_{5i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}194}{0{,}5} \cdot 100\% = 80{,}6\%$$

Випадок r=6

$$q_6 = 0{,}23, \frac{1}{2^t + 1} < 0{,}23 \leq \frac{1}{2^{t-1} + 1} \rightarrow t_6 \rightarrow 2$$

$$E\left(\delta_f(m_{6i})\right) = \frac{1 - q_6}{2} + \frac{q_6}{2} = 0{,}4225, i \in \{1{,}2, \dots, k\}$$

$$\frac{0{,}5 - E\left(\delta_f(m_{6i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}4225}{0{,}5} \cdot 100\% = 15{,}5\%$$

$$E\left(\delta_f(m_{6i})\right) = 0{,}5H = -0{,}5 \cdot (p_6 \cdot \log_2 p_6 + q_6 \cdot \log_2 q_6) =$$

$$= -0{,}5 \cdot (0{,}23 \cdot \log_2 0{,}23 + 0{,}77 \cdot \log_2 0{,}77) = 0{,}5 \cdot 0{,}788 = 0{,}3890$$

$$\frac{0{,}5 - E\left(\delta_g(m_{4i})\right)}{0{,}5} \cdot 100\% = \frac{0{,}5 - 0{,}5 \cdot 0{,}811}{0{,}5} \cdot 100\% = 22{,}2\%$$

Позначимо через V_r відміну в процентах в сторону зменшення математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення при використанні субоптимального ентропійного стеганографічного алгоритму в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта r -го розряду повідомлення. Через W_r позначимо відміну в процентах в сторону зменшення математичного очікування числа змінних елементів контейнера на один біт впроваджуваного повідомлення з попередніми стисненням повідомлення за допомогою асимптотично оптимального рівномірного блокового коду в порівнянні з впровадженням за ознакою парності суми елементів контейнера, обраних для впровадження одного біта r-го розряду повідомлення, де $r \in \{0{,}1, \dots, 6\}$. Для порівняння двох способів стеганографічного захисту повідомлення (способу захисту без стиснення повідомлення із застосуванням субоптимального

стеганографічного алгоритму, представленого в розділі 2.3, і способу захисту з попереднім стисненням асимптотично оптимальним рівномірним блоковим кодом) по математичному очікуванню числа змінних елементів контейнера на один біт повідомлення у зручний спосіб деякі результати вищенаведених обчислень представити у вигляді такої таблиці (3.3.2):

Таблиця 3.3.2

Кількісні характеристики порівнюваних алгоритмів

r	$\mathbb{E}(\xi_c(m_{ri}))$	$\mathbb{E}(\xi_f(m_{ri}))$	$\mathbb{E}(\xi_g(m_{ri}))$	V_r	W_r	$W_r - V_r$
0	0,5	0,5	0,4950	0%	1,0%	1,0%
1	0,5	0,49	0,4520	2,0%	9,6%	7,6%
2	0,5	0,5	0,4995	0%	0,1%	0,1%
3	0,5	0,49	0,4520	2,0%	9,6%	7,6%
4	0,5	0,4375	0,4055	12,5%	18,9%	6,4%
5	0,5	0,1052	0,0970	79,0%	80,6%	1,6%
6	0,5	0,4225	0,3890	15,5%	22,2%	6,7%

В таблиці 3.3.2 через $E(\delta_c(m_{ri}))$, $E\left(\delta_f(m_{ri})\right)$ і $E\left(\delta_g(m_{ri})\right)$ позначені математичні очікування числа змінних елементів контейнера на один біт повідомлення відповідно при його впровадженні без стиснення за ознакою парності суми елементів контейнера, за допомогою субоптимального ентропійного стеганографічного алгоритму і з попереднім стисненням за допомогою асимптотично оптимального рівномірного блокового коду.

Наочно порівняльні характеристики розглянутих алгоритмів представлені у вигляді такої діаграми (рис. 3.3.1).

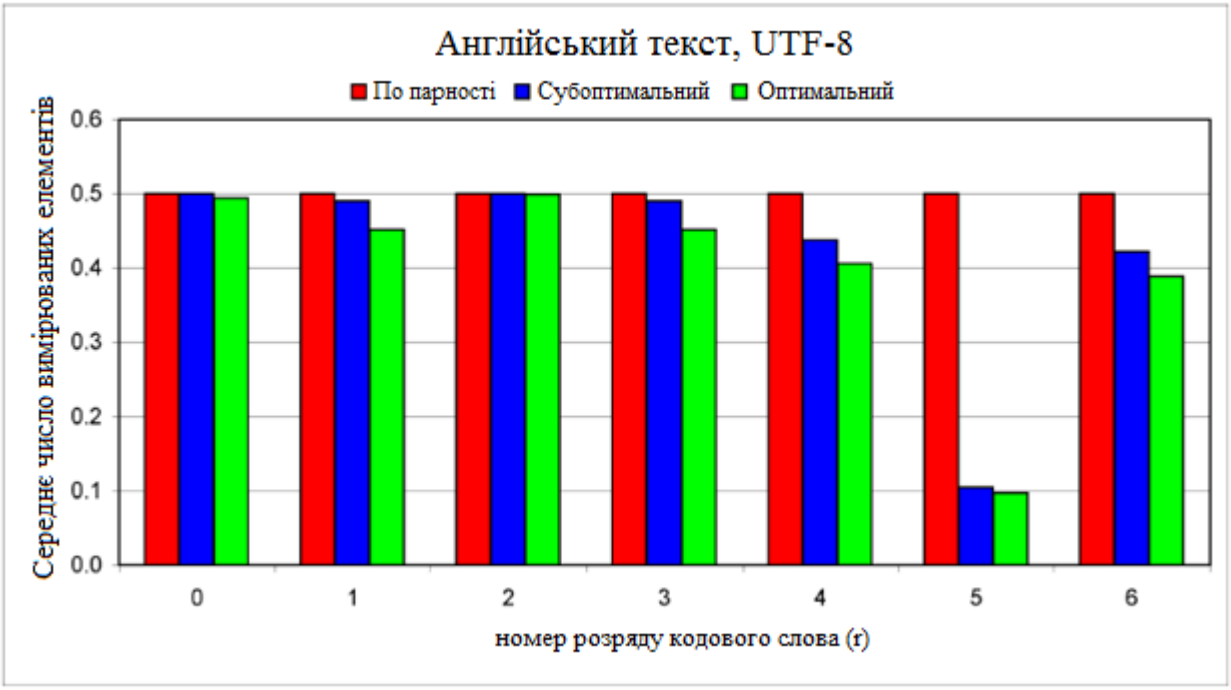

Рисунок 3.3.1. Діаграма значень середнього числа змінних елементів контейнера на один біт англомовного повідомлення для трьох стеганографічних алгоритмів.

Тут червоні стовпчики відповідають впровадженню без стиснення за ознакою парності суми елементів контейнера, сині - за допомогою субоптимального ентропійного стеганографічного алгоритму, зелені - з попередніми стисненням за допомогою асимптотично оптимального рівномірного блокового коду.

3.4 Висновки

Облік ймовірносно-статистичних характеристик впроваджуваного повідомлення, що підлягає стеганографічному захисту, може служити основою для розробки субоптимальних стеганографічних алгоритмів, що забезпечують зменшення, в порівнянні з відомими стеганографічними алгоритмами, середнього числа змінних елементів контейнера на один біт впроваджуваного повідомлення. Даний підхід може бути ефективним в тих додатках, в яких не допускається або небажано стиснення повідомлення перед його впровадженням в стеганографічний контейнер.

Розроблений і представлений в цьому дослідженні субоптимальний стеганографічний алгоритм можна модифікувати для спільного вирішення відразу двох завдань: завдання контрольованого обмеження величини спотворення елементів контейнера при їх зміні і завдання зменшення середнього числа змінних елементів на один біт впроваджуваного повідомлення. Відповідний модифікований стеганографічний алгоритм зберігає характеристики вихідного субоптимального стеганографічного алгоритму. При цьому можливе збільшення числа елементів контейнера для впровадження одного біта повідомлення і зміна двійковій функції впровадження-вилучення.

ВИСНОВКИ

Мета роботи, яка полягає в підвищенні ефективності стеганографічної захисту інформації шляхом модифікації існуючих ентропійних стеганографінчих алгоритмів, досягнута. У процесі виконання роботи були отримані наступні результати:

– Описано особливості побудови стеганографічних систем, а саме проведений аналіз та визначене поняття стеганографічної системи, проведена класифікація стеганографічних систем та методика їх побудови, також проведений порівняльний аналіз існуючих стеганографічних алгоритмів.

– Розглянуто первинні поняття і елементи ентропійного підходу в стеганографії, виявлено ентропійний механізм, в певному математичному сенсі загальний для всіх стандартних алгоритмів компресії цифрових мультимедійних файлів. Він запропонований в якості основи розробки і побудови алгоритмів впровадження повідомлень, що підлягають приховану, в стеганографічний контейнер, який представляє собою безліч квантованих коефіцієнтів частотної області мультимедійного сигналу. Розглянутий підхід не залежить від природи сигналу (аудіо-, відеосигнал або нерухоме зображення).

– Розроблено та запропоновано аналітичний апарат для побудови ентропійних стеганографічних алгоритмів. Введено і математично обгрунтовані поняття оптимального ентропійного стеганографічного алгоритму і субоптимального ентропійного стеганографічного алгоритму. Виявлено і досліджено властивості області визначення булевої функції впровадження-вилучення ентропійного стеганографічного алгоритму, які можуть бути використані при побудові оптимальних і субоптимальних ентропійних стеганографічних алгоритмів.

– Введено і математично обгрунтовані поняття оптимального ентропійного стеганографічного алгоритму і субоптимального ентропійного стеганографічного алгоритму. Виявлено і досліджено властивості області

визначення булевої функції впровадження-вилучення ентропійного стеганографічного алгоритму, які можуть бути використані при побудові оптимальних і субоптимальних ентропійних стеганографічних алгоритмів. За рахунок цього запропоновано оптимальний ентропійний стеганографічний, при якому заповнення стеганографічного контейнера здійснюється побітово за ознакою парності суми обраних елементів для повідомлень, згенерованих двійковим джерелом без пам'яті і перед впровадженням в стеганографічний контейнер стислих з використанням асимптотично оптимального блокового рівномірного коду. Та субоптимальний ентропійний стеганографічний алгоритм захисту повідомлень (що не піддаються попередньою стиску), згенерованих двійковим джерелом без пам'яті.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Аграновский, А. В. Основы компьютерной стеганографии / А. В. Аграновский, П. Н. Девянин, Р. А. Хади, А. В. Черемушкин. — М.: Радио и связь, 2003. — 151 с.

2. Аграновский, А. В. Стеганография, цифровые водяные знаки и стеганоанализ / А. В. Аграновский, А. В. Балакин, В. Г. Грибунин, С. А. Сапожников. — М.: Вузовская книга, 2009. — 220 с.

3. Алиев, Ф. К. О сходимости матриц переходных вероятностей автоматов, определяющих преобразования скользящей перестановки / Ф. К. Алиев // Третья Всероссийская школа-коллоквиум по стохастическим методам. — М.: ТВП, 1996. — С. 13–14.

4. Алиев, Ф. К. О реализуемости вероятностных автоматов, определяющих преобразования скользящей перестановки, автоматами Мили со случайным входом / Ф. К. Алиев // Третья Всероссийская школа-коллоквиум по стохастическим методам. — М.: ТВП, 1996. — С. 14–15.

5. Алиев, Ф. К. Курс лекций по математической логике и теории алгоритмов / Ф. К. Алиев, И. А. Юров. — М.: МИФИ, 2003. — 198 с.

6. Алиев, Ф. К. О возможности генерации подстановок с использованием генераторов скользящих перестановок / Ф. К. Алиев, А. В. Зайцева, В. А. Киселенко, А. Г. Сенцов // Обозрение прикладной и промышленной математики. — М.: ТВП, 2012. — Т. 19, в. 3. — С. 421.

7. Алиев, Ф. К. О возможности изменения плотности распределения вероятностей амплитуды шума квантования цифрового мультимедийного сигнала при использовании его сжатых образов в стеганографических приложениях / Ф. К. Алиев, А. В. Зайцева, В. А. Киселенко // Обозрение прикладной и промышленной математики. — М.: ТВП, 2012. — Т. 19, в. 5. — С. 654–655.

8. Алиев, Ф. К. Элементы энтропийного подхода в стеганографии / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк // Обозрение прикладной и промышленной математики. — М.: ТВП, 2012. — Т. 19, в. 5. — С. 655–659.

9. Алиев, Ф. К. Допустимые пары систем подмножеств множества двоичных векторов / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк, А. Г. Сенцов // Обозрение прикладной и промышленной математики. — М.: ТВП, 2012. — Т. 19, в. 5. — С. 659–660.
10. Алиев, Ф. К. Об оптимальных энтропийных стеганографических алгоритмах / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк, А. Г. Сенцов, И. А. Шеремет // Сборник научных трудов по итогам работы 6–7 Всероссийских научно-технических конференций школы-семинара «Информационная безопасность — актуальная проблема современности». — Краснодар: ФВАС (г. Краснодар), 2013. — Т. 1. — С. 56–60.
11. Алиев, Ф. К. О субоптимальном энтропийном стеганографическом алгоритме защиты сообщений, сгенерированных двоичным источником без памяти / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк, А. Г. Сенцов, И. А. Шеремет // Сборник научных трудов по итогам работы 6–7 Всероссийских научно-технических конференций школы-семинара «Информационная безопасность — актуальная проблема современности». — Краснодар: ФВАС (г. Краснодар), 2013. — Т. 1. — С. 50–52.
12. Алиев, Ф. К. Оптимальный и субоптимальный энтропийные стеганографические алгоритмы защиты сообщений, сгенерированных двоичным источником без памяти / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк, А. Г. Сенцов, И. А. Шеремет // Известия Института инженерной физики, 2013. — №4 (30). — С. 2–9.
13. Алиев, Ф. К. Энтропийные стеганографические системы с контейнерами, представляющими собой цифровые мультимедийные файлы / Ф. К. Алиев, А. В. Зайцева, И. В. Костенюк, А. Г. Сенцов, И. А. Шеремет // Известия Института инженерной физики, 2014. — №1 (31). — С. 2–10.
14. Алферов, А. П. Основы криптографии: Учебное пособие / А. П. Алферов, А. Ю. Зубов, А. С. Кузьмин, А. В. Черемушкин. — М.: Гелиос АРВ, 2005. — 480 с.
15. Бабаш, А. В. Криптография / А. В. Бабаш, Г. П. Шанкин. — М.: СОЛОН-Р,

2002. — 512 с.

16. Балакин, А. В. Современная стеганография: модели и методы преобразования информации / А. В. Балакин, С. А. Репалов, Г. Н. Шагов. — Ростов-на-Дону: СКНЦ ВШ, 2004. — 238 с.

17. Боровков, А. А. Теория вероятностей / А. А. Боровков. — М.: Наука, 1976. — 352 с.

18. Брауэр, В. Введение в теорию конечных автоматов / Вильфрид Брауэр. — М.: Радио и связь, 1987. — 392 с.

19. Быков, Р. Е. Цифровое преобразование изображений / Р. Е. Быков, Р. Фрайер, К. В. Иванов, А. А. Манцветов. — М.: Горячая линия – Телеком, 2012. — 228 с.

20. Ватолин, Д. Методы сжатия данных / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. — М.: ДИАЛОГ-МИФИ, 2003. — 384 с.

21. Вернер, М. Основы кодирования / Мартин Вернер. — М.: Техносфера, 2006. — 288 с.

22. Гантмахер, Ф. Р. Теория матриц / Ф. Р. Гантмахер. — М.: Физматлит, 2004. — 560 с.

23. Генне, О. В. Основные положения стеганографии / О. В. Генне // «Защита информации. Конфидент», 2000. — № 3. — С. 20–25.

24. Гмурман, В. Е. Руководство к решению задач по теории вероятностей и математической статистике / В. Е. Гмурман. — М.: Высшая школа, 2000. — 400 с.

25. Голубев, Е. А. Стелсографические угрозы / Е. А. Голубев // Материалы 2-ой межведомственной конференции «Научно-техническое и информационное обеспечение деятельности спецслужб». — Москва, 1998. — С. 58–60.

26. Голубев, Е. А. Стеганография как одно из направлений обеспечения информационной безопасности / Е. А. Голубев, Г. В. Емельянов // T-Comm. Телекоммуникации и транспорт. — Москва, 2009. — Спецвыпуск, август, ч. 3. — С. 185–186.

27. Гонсалес, Р. Цифровая обработка изображений / Рафаэль Гонсалес, Ричард Вудс. — М.: Техносфера, 2012. — 1104 с.

28. Грибунин, В. Г. Стеганографическая защита речевых сигналов в каналах открытой телефонной связи / В. Г. Грибунин, И. Н. Оков, И. В. Туринцев // Сборник тезисов Российской НТК «Методы и технические средства обеспечения безопасности информации». — СПб.: ГТУ, 2001. — С. 83–84.

29. Грибунин, В. Г. Цифровая стеганография / В. Г. Грибунин, И. Н. Оков, И. В. Туринцев. — М.: «Солон-Пресс», 2002. — 272 с.

30. Грибунин, В. Г. JPEG 2000: десять лет спустя / В. Г. Грибунин // Технологии защиты, 2007. — С. 50–52.

31. Грибунин, В. Г. Основы стеганографии. Учебно-методическое пособие / В. Г. Грибунин, О. А. Жердин, А. П. Мартынов, Д. Б. Николаев, А. Г. Силаев, В. Н. Фомченко. — Трехгорный: изд-во СарФТИ, 2012 г. — 130 с.

32. Грибунин, В. Г. Стеганографический анализ изображений, представленных в оптимальном базисе вейвлет-пакетов / В. Г. Грибунин, Д. А. Токарев // Известия Института инженерной физики, 2014. — №1 (31). — С. 6–11.

33. Духин, А. А. Теория информации / А. А. Духин. — М.: Гелиос АРВ, 2007. — 248 с.

34. Зайцева, А. В. Современное состояние стеганографии / А. В. Зайцева // Интеллектуальные системы: Труды Девятого Международного симпозиума / под ред. К.А. Пупкова. — М.: РУСАКИ, 2010. — С. 568–570.

35. Зайцева, А. В. О подходах к построению энтропийных стеганографических систем защиты сообщений в информационных сетях / А. В. Зайцева // Вопросы защиты информации, 2014. — №2. — С. 57–64.

36. Карякин, В. Л. Цифровое телевидение / В. Л. Карякин. — М.: СОЛОН-ПРЕСС, 2008. — 272 с.

37. Кирко, И. Н. Вопросы защиты авторских прав электронных документов / И. Н. Кирко, В. В. Бедусенко // Инновационные недра Кузбасса. IT-технологии: сборник научных трудов. — Кемерово: ИНТ, 2007. — 420 с.

38. Киселенко, В. А. Об одном способе встраивания информации в

стегосистему / В. А. Киселенко // Обозрение прикладной и промышленной математики, 2005. — Т. 12, в. 4. — С. 984–985.

39. Киселенко, В. А. О периодических свойствах преобразований скользящей перестановки в связи со стеганографическими приложениями / В. А. Киселенко // Обозрение прикладной и промышленной математики, 2005. — Т. 12, в. 4. — С. 985–986.

40. Киселенко, В. А. О неопределенности ключа в стегосистемах со схемой встраивания информации на основе скользящей перестановки / В. А. Киселенко, И. А. Рождественский // Обозрение прикладной и промышленной математики, 2005. — Т. 12, в. 4. — С. 986–987.

41. Киселенко, В. А. Об удельной энтропии выходных последовательностей автомата, определяющего преобразования скользящей перестановки, в связи со стеганографическими приложениями / В. А. Киселенко // Обозрение прикладной и промышленной математики, 2006. — Т. 13, в. 5. — С. 865–866.

42. Киселенко, В. А. Об использовании фильтрующих генераторов для гарантирования периодов последовательностей встраивания информации в стеганографические контейнеры / В. А. Киселенко // Обозрение прикладной и промышленной математики, 2006. — Т. 13, в. 5. — С. 863–864.

43. Конахович, Г. Ф. Компьютерная стеганография. Теория и практика. / Г. Ф. Конахович, А. Ю. Пузыренко. — К.: «МК-Пресс», 2006. — 288 с.

44. Корн, Г. Справочник по математике для научных работников и инженеров / Гранино Корн, Тереза Корн. — М.: Наука, 1984. — 833 с.

45. Кудрявцев, Л. Д. Курс математического анализа (в двух томах): Учебник для студентов университетов и втузов / Л. Д. Кудрявцев. — М.: Высш. школа, 1981. — Т. 1. — 687 с.

46. Кустов, В. Н. Методы встраивания скрытых сообщений / В. Н. Кустов, А. А. Федчук // Конфидент, 2000. — № 3. — С. 68–72.

47. Леонтьев, В. П. Цифровое фото, музыка и звук. / В. П. Леонтьев, И. В. Прокошев. — М.: ОЛМА-ПРЕСС, 2005. — 384 с.

48. Нильсен, М. Квантовые вычисления и квантовая информация / Майкл

Нильсен, Исаак Чанг. — М.: Мир, 2006. — 824 с.

49. Новиков, А. М. Методология: Словарь системы основных понятий / А. М. Новиков, Д. А. Новиков. — М.: ЛИБРОКОМ, 2013. — 208 с.

50. Оппенгейм, А. Цифровая обработка сигналов / Алан Оппенгейм, Рональд Шафер — М.: Техносфера, 2006. — 856 с.

51. Пападимитриу, Х. Комбинаторная оптимизация. Алгоритмы и сложность / Христос Пападимитриу, Кеннет Стайглиц. — М.: Мир, 1985. — 512 с.

52. Приоритетные проблемы научных исследований в области обеспечения информационной безопасности Российской Федерации (утверждена Исполняющим обязанности Секретаря Совета Безопасности Российской Федерации, председателя научного совета при Совете Безопасности Российской Федерации 7 марта 2008 г.). [Электронный ресурс] Режим доступа: http://www.scrf.gov.ru/documents/93.html.

53. Риордан, Дж. Комбинаторные тождества / Джон Риордан. — М.: Наука, 1982. — 256 с.

54. Ричардсон, Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения / Ян Ричардсон. — М.: Техносфера, 2005. — 368 с.

55. Смирнов, А. В. Цифровое телевидение / А. В. Смирнов, А. Е. Пескин. — М.: Горячая линия – Телеком, 2005. — 352 с.

56. Стратонович, Р. Л. Теория информации / Р. Л. Стратонович. — М.: «Сов. радио», 1975. — 424 с.

57. Стрельцов, А. А. Обеспечение информационной безопасности России / А. А. Стрельцов. — М.: МЦНМО, 2002. — 296 с.

58. Сэломон, Д. Сжатие данных, изображений и звука / Дэвид Сэломон. — М.: Техносфера, 2004. — 368 с.

59. Тюхтин, М. Ф. Системы Интернет-телевидения / М. Ф. Тюхтин. — М: Горячая линия – Телеком, 2008. — 236 с.

60. Феллер, В. Введение в теорию вероятностей и её приложения / Вильям Феллер. — М.: Мир, 1967. — Т. 2. — 751 с.

61. Фомичев, В. М. Дискретная математика и криптология / В. М. Фомичев. — М.: ДИАЛОГ-МИФИ, 2003. — 400 с.

62. Челноков, А. С. О чем не рассказал Сноуден / А. С. Челноков. — М.: Яуза-пресс, 2013. — 256 с.

63. Чечета, С. И. Введение в дискретную теорию информации и кодирования / С. И. Чечета. — М.: МЦНМО, 2011. — 224 с.

64. Шеннон, К. Работы по теории информации и кибернетике / Клод Шеннон. — М.: ИЛ, 1963. — 869 с.

65. Шнайер, Б. Прикладная криптография / Брюс Шнайер. — М.: ТРИУМФ, 2003 г. — 816 с.

66. Штарк, Г.-Г. Применение вейвлетов для ЦОС / Ганс-Георг Штарк. — М.: Техносфера, 2007. — 192 с.

67. Эндрюс, Дж. Теория разбиений / Джордж Эндрюс. — М.: Наука, 1982. — 256 с.

68. Ярославский, Л. П. Введение в цифровую обработку изображений / Л. П. Ярославский. — М.: Сов. радио, 1979. — 312 с.

69. ATSC A/52:2010. Digital Audio Compression Standard (AC-3, E-AC-3). 22 November 2010.

70. Bцhme, R. Advanced Statistical Steganalysis / Rainer Bцhme — Springer-Verlag Berlin Heidelberg, 2010. — 285 p.

71. Cox, I. Digital Watermarking and Steganography / Ingemar Cox, Matthew Miller, Jeffrey Bloom, Jessica Fridrich, Ton Kalker. — San Francisco: Morgan Kaufmann Publishers, 2008. — 594 p.

72. Fridrich, J. Steganography in Digital Media / Jessica Fridrich. — Cambridge University Press, 2010. — 437 p.

73. ISO/IEC (1993) International Standard IS 11172-3 «Information Technology – Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbits/s – Part 3: Audio».

74. ISO/IEC (1994) International Standard IS 13818-1 «Information Technology – Coding of Moving Pictures and Associated Audio Information. Part 1:

Systems./Ed/1, JTS 1/SC 29, 1994.