Міністерство освіти і науки України
Київський національний університет імені Тараса Шевченка

Факультет інформаційних технологій
Кафедра кібербезпеки та захисту інформації

ДОПУСТИТИ ДО ЗАХИСТУ:
В.о. завідувача кафедри
кібербезпеки та захисту
інформації
_____ Іван ПАРХОМЕНКО
«13» червня 2025 р.

ПОЯСНЮВАЛЬНА ЗАПИСКА

кваліфікаційної роботи

галузь знань 12 Інформаційні технології
(шифр і назва галузі знань)

спеціальність 125 Кібербезпека
(код і назва спеціальності)

освітній ступень бакалавр

освітня програма Кібербезпека
(назва освітньо-професійної програми)

на тему: «Стеганографічний метод форматування символів в документах MS Word»

Виконавець: студент IV курсу, групи КБ-42

_____ **Тимофій ХОМИЧ**
(підпис) (ім'я, прізвище)

	Підпис	**Ім'я ПРІЗВИЩЕ**
Керівник		Юрій БАБЕНКО
Нормоконтроль		Інна МИХАЛЬЧУК

Київ 2025

Міністерство освіти і науки України
Київський національний університет імені Тараса Шевченка

Факультет інформаційних технологій
Кафедра кібербезпеки та захисту інформації

ЗАТВЕРДЖЕНО:
В.о. завідувача кафедри
кібербезпеки
та захисту інформації
_____ Іван ПАРХОМЕНКО
«29» листопада 2024 р.

ЗАВДАННЯ
на виконання кваліфікаційної роботи

спеціальності 125 Кібербезпека
(код і назва спеціальності)

освітньої програми Кібербезпека
(назва освітньо-професійної програми)

Студенту **КБ-42** (група) **Хомичу Тимофію Юрійовичу** (прізвище ім'я по батькові)

Тема кваліфікаційної роботи Стеганографічний метод форматування символів в документах MS Word

1. ПІДСТАВИ ДЛЯ ПРОВЕДЕННЯ РОБОТИ

Тема кваліфікаційної роботи затверджена на засіданні кафедри кібербезпеки та захисту інформації протокол №6 від 28.11.2024 р.

2. ВИХІДНІ ДАНІ ДЛЯ ПРОВЕДЕННЯ РОБІТ

Структура форматованих документів MS Word (DOCX), принципи функціонування стеганографічних методів форматування символів

3. ЗМІСТ РОЗРАХУНКОВО-ПОЯСНЮВАЛЬНОЇ ЗАПИСКИ

Аналітичний огляд методів стеганографії, узагальнена модель стегосистеми, опис методів форматного приховування в тексті, розробка методу приховування/витягу даних, реалізація програмного рішення.

4. ВИМОГИ ДО РЕЗУЛЬТАТІВ ВИКОНАННЯ РОБОТИ

Практична цінність Створене програмне рішення для автоматизованого вбудовування та вилучення прихованої інформації з документів.

5. ДАТА ВИДАЧІ ЗАВДАННЯ

Дата видачі завдання: 29 листопада 2024 року

Завдання видав _____ (підпис) Юрій БАБЕНКО (ім'я, прізвище)

Завдання прийняв до виконання _____ (підпис) Тимофій ХОМИЧ (ім'я, прізвище)

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Найменування етапів робіт	Строки виконання робіт (початок-кінець)	Відмітка про виконання
1	Уточнення постановки задачі	29.11.2024 – 04.12.2024	*виконано*
2	Аналіз літератури	04.12.2024 – 21.12.2024	*виконано*
3	Обґрунтування вибору рішення	21.12.2024 – 16.01.2025	*виконано*
4	Аналіз застосування стеганографічних методів та їх класифікація	16.01.2025 – 01.02.2025	*виконано*
5	Дослідження текстових методів приховування інформації	03.02.2025 – 25.02.2025	*виконано*
6	Розробка методу форматування символів для приховування даних у MS Word	26.02.2025 – 25.03.2025	*виконано*
7	Реалізація програмного засобу	27.03.2025 – 26.04.2025	*виконано*
8	Оформлення пояснювальної записки	26.04.2025 – 21.05.2025	*виконано*
9	Підготовка до захисту кваліфікаційної роботи	22.05.2025 – 13.06.2025	*виконано*

Завдання видав _____ (підпис) Юрій БАБЕНКО (ім'я, прізвище)

Завдання прийняв до виконання _____ (підпис) Тимофій ХОМИЧ (ім'я, прізвище)

Термін подання кваліфікаційної роботи до ЕК 13 червня 2025 року

РЕФЕРАТ

Кваліфікаційна робота складається зі вступу, трьох розділів, загальних висновків, списку використаних джерел, додатків, має 62 сторінки основного тексту, 1 таблицю та 18 рисунків. Список використаних джерел містить 31 найменування і займає 4 сторінки.

Метою роботи є підвищення рівня прихованості передавання інформації шляхом використання стеганографічного методу у текстових документах формату MS Word.

Для досягнення зазначеної мети були поставлені наступні завдання:

1. Проаналізувати наявні стеганографічні методи захисту інформації в текстових документах та визначити їх основні переваги та недоліки.
2. Дослідити особливості форматування символів у документах MS Word та виявити параметри, які можуть бути використані для приховування інформації.
3. Розробити стеганографічний метод приховування даних у текстових документах MS Word.
4. Створити програмну реалізацію розробленого стеганографічного алгоритму.

Об'єктом дослідження є процес захисту конфіденційної інформації в електронних документах на основі стеганографічного методу.

Предметом дослідження є стеганографічний метод приховування інформації у документах MS Word.

Практична цінність виражається у створеному програмному рішенні для автоматизованого вбудовування та вилучення прихованої інформації з документів.

Ключові слова: стеганографія, текстова стеганографія, приховування інформації, захист інформації, форматування символів, Microsoft Word.

ЗМІСТ

ВСТУП

Сьогодення характеризується надзвичайно високим темпом технологічного розвитку та глобальною цифровізацією всіх сфер життєдіяльності, проблема забезпечення конфіденційності та цілісності інформації набуває першочергового значення. Цифрові документи, які містять конфіденційну інформацію, потребують надійних механізмів захисту, що здатні протистояти різноманітним кібератакам та запобігати несанкціонованому доступу до даних. В умовах, коли традиційні криптографічні методи захисту інформації стають все більш уразливими через зростання обчислювальної потужності сучасних комп'ютерних систем, виникає необхідність розробки та впровадження альтернативних підходів до вирішення питань інформаційної безпеки.

Стеганографічні методи захисту інформації, які ґрунтуються на приховуванні самого факту наявності секретного повідомлення, відкривають широкі перспективи для підвищення рівня захищеності електронних документів. Особливу увагу привертають текстові стеганографічні методи, застосування яких дозволяє здійснювати приховану передачу даних у текстових документах, зокрема, документах MS Word, шляхом маніпулювання форматуванням символів таким чином, що відповідні зміни залишаються непомітними для людського ока. Актуальність даної роботи визначається необхідністю розробки ефективних методів захисту інформації, які дозволяють приховувати конфіденційні дані в електронних документах без помітної зміни їх зовнішнього вигляду та функціональних характеристик.

На сьогоднішній день існує значна кількість досліджень, присвячених використанню стеганографічних методів для захисту інформації в мультимедійних файлах, зокрема, зображеннях, аудіо- та відеоматеріалах. Водночас, текстова стеганографія, яка використовує особливості форматування символів, наприклад, в документах MS Word, залишається недостатньо

дослідженою галуззю, що потребує подальшого розвитку та вдосконалення. Актуальність тематики даного дослідження підтверджується також зростаючою популярністю текстових процесорів, зокрема MS Word, як одного з найбільш поширених інструментів для створення та обробки електронних документів у корпоративному середовищі, освітніх установах та державних організаціях.

Метою даної роботи є підвищення рівня прихованості передавання інформації шляхом використання стеганографічного методу у текстових документах формату MS Word.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Проаналізувати стеганографічні методи захисту інформації в текстових документах та визначити їх основні переваги та недоліки.
2. Дослідити особливості форматування символів у документах MS Word та виявити параметри, які можуть бути використані для приховування інформації.
3. Розробити стеганографічний метод приховування даних у текстових документах MS Word шляхом маніпулювання параметрами форматування символів.
4. Створити програмну реалізацію розробленого стеганографічного методу.

Об'єктом дослідження є процес захисту конфіденційної інформації в електронних документах на основі стеганографічного методу.

Предметом дослідження є стеганографічний метод приховування інформації шляхом зміни параметрів форматування символів у документах MS Word.

Практичне значення одержаних результатів полягає у розробці ефективного методу та програмного рішення захисту конфіденційної інформації в електронних документах MS Word, які можуть бути впроваджені в корпоративних системах документообігу для підвищення рівня інформаційної безпеки. Розроблений стеганографічний метод дозволить створювати приховані канали передачі даних для забезпечення конфіденційності інформації в умовах,

коли використання криптографічних методів захисту є неможливим або недоцільним. Результати дослідження також можуть бути використані для підвищення ефективності системи захисту інформації в державних установах, фінансових організаціях та компаніях, які оперують конфіденційною інформацією. Запропоновані методи можуть слугувати основою для розробки нових підходів до захисту авторських прав на електронні документи та протидії несанкціонованому копіюванню текстової інформації.

Галузь застосування охоплює сферу кібербезпеки підприємств та організацій, що працюють з конфіденційною інформацією та потребують надійних механізмів прихованої передачі даних, зокрема компанії оборонного сектору, державні установи, науково-дослідні інститути, фінансові організації та підприємства мілтех-галузі, де існує необхідність захисту інтелектуальної власності, комерційних таємниць та службової документації від несанкціонованого доступу через впровадження непомітних для стороннього спостерігача стеганографічних методів у звичайні офісні документи.

Новизна полягає у розробці удосконаленого стеганографічного методу, що використовує комбінацію невидимих форматувань символів у документах MS Word для створення багаторівневого стеганографічного каналу з підвищеною стійкістю до виявлення.

Практична цінність роботи виражається у створеному програмному рішенні для автоматизованого вбудовування та вилучення прихованої інформації з документів.

Апробація роботи. Основні результати роботи доповідались на VIII Міжнародній науково-практичній конференції «Проблеми кібербезпеки інформаційно-комунікаційних систем» (PCSICS)»11 квітня 2025 року.

РОЗДІЛ 1
АНАЛІЗ ОСНОВНИХ СТЕГАНОГРАФІЧНИХ МЕТОДІВ ЗАХИСТУ ІНФОРМАЦІЇ

1.1 Концептуальні засади стеганографії як науки про приховану передачу інформації

Стеганографія, що походить від грецьких слів «steganos» (прихований) та «graphein» (писати), являє собою комплекс методів приховування конфіденційної інформації шляхом вбудовування її в нейтральні об'єкти таким чином, щоб факт наявності прихованого повідомлення залишався непомітним для сторонніх спостерігачів [1]. На відміну від криптографії, що трансформує вміст повідомлення до криптограми, з метою унеможливлення його прочитання неавторизованими суб’єктами, стеганографія зосереджує свою увагу не на змісті, а на самому факті комунікації, що зумовлює її потенціал як інструмента інформаційного маскування в умовах відкритого середовища передавання даних [2].

З теоретичного погляду, стеганографія базується на фундаментальних положеннях теорії інформації, семіотики, обчислювальної складності та системного аналізу, що дозволяє формалізувати процес прихованої передачі у вигляді математичних моделей і аналітичних структур. Основним об'єктом стеганографії виступає контейнер – носій, до якого інкапсульовано приховане повідомлення таким чином, щоб його семантична або технічна форма залишалася у межах допустимих норм і не викликала підозри в умовах типового аналізу. Як правило, для цього використовуються файли з надмірною або надлишковою структурою (текстові документи, зображення, аудіо- та відеофайли), де модифікація певних елементів не змінює сприйняття вмісту користувачем або машиною.

Концептуально стеганографія розглядається як один із напрямів інформаційного приховування, що дозволяє реалізувати так звану нульову підозру, коли навіть факт взаємодії між суб'єктами системи не фіксується або сприймається як нетривіальний. Така властивість забезпечує потенційно вищий рівень прихованості порівняно з іншими засобами інформаційного захисту, що, у свою чергу, робить стеганографію придатною для використання в умовах жорстких обмежень щодо конфіденційності або в середовищах з активним спостереженням і фільтрацією інформаційного трафіку.

Методологічно процес приховування інформації у стеганографії включає три основні компоненти: стегооб'єкт (базовий носій), секретне повідомлення (власне інформаційне навантаження) та ключ або протокол вбудовування, який визначає правила інкапсуляції й витягання даних [3]. Надійність стеганографічного методу оцінюється за кількома критеріями, серед яких: стійкість до стегоаналізу (тобто здатність методу залишатися непомітним при спробах виявлення прихованої інформації), ємність (обсяг даних, який може бути приховано), а також відсутність спотворення функціональності або змісту стегооб'єкта. В ідеальному випадку стеганографічне повідомлення не має створювати жодних відмінностей між модифікованим та оригінальним носієм у межах допустимих статистичних або перцептивних моделей.

У межах сучасної наукової парадигми стеганографія дедалі частіше розглядається не лише як технологічна дисципліна, орієнтована на засоби маскування, але і як концептуальний компонент гібридних інформаційних систем захисту, що поєднує елементи криптографії, аутентифікації, цифрового водяного маркування та обфускації. Такий підхід дає змогу створювати багаторівневі моделі захисту, в яких стеганографія виконує функцію прихованого каналу зв'язку або резервного механізму передачі критично важливої інформації в умовах порушеної комунікаційної інфраструктури.

У результаті, концептуальні засади стеганографії, з одного боку, мають глибоке теоретичне підґрунтя, що дозволяє формувати формальні моделі й алгоритми прихованої передачі інформації, а з іншого - забезпечують потужну

прикладну базу для розроблення практичних рішень, орієнтованих на конкретні інформаційно-комунікаційні середовища. Це відкриває широкі перспективи для подальших досліджень у контексті адаптації стеганографічних методів до сучасних цифрових платформ, зокрема – офісних програмних засобів, таких як MS Word, у яких особливості форматування символів можуть бути використані як стеганографічні маркери для прихованої передачі даних без зміни змісту або структури документа.

1.2 Історичний розвиток

Розвиток стеганографічних методів захисту інформації характеризується тривалою еволюцією, що бере свій початок у стародавніх цивілізаціях та продовжується до сьогодення, трансформуючись відповідно до технологічних можливостей кожної епохи. Перші згадки про використання стеганографії датуються V століттям до н.е., коли давньогрецький історик Геродот описав метод татуювання секретних повідомлень на голові раба, які ставали невидимими після відростання волосся [4]. У Стародавньому Римі широко застосовувалися невидимі чорнила на основі органічних речовин, що проявлялися лише при нагріванні або хімічній обробці.

Середньовічний період характеризувався розвитком мікрографії – техніки зменшення розміру тексту до такого ступеня, що його можна було приховати в елементах художнього оформлення рукописів або в певних літерах тексту [5]. З розвитком поліграфії в епоху Відродження з'явилися методи прихованої передачі інформації з використанням особливостей друкованого тексту, зокрема, техніка «решітки Кардано», що дозволяла читати приховане повідомлення шляхом накладання спеціального шаблону на текст.

Значний внесок у розвиток стеганографії було зроблено в період Першої та Другої світових війн, коли активно використовувалися мікрокрапки – фотографічні зменшення документів до розміру точки, які розміщували на звичайних документах. Німецькі шпигуни успішно застосовували цю

технологію для передачі секретної інформації, що спонукало розвиток відповідних методів контрстеганографії [6].

Сучасний етап розвитку стеганографії розпочався у 1980-х роках з появою цифрових технологій та персональних комп'ютерів. Цифрова стеганографія, яка використовує електронні носії інформації як стеганоконтейнери, відкрила нові можливості для прихованої передачі даних. Науковці Т. Пітерсон та Г. Боас у своїх фундаментальних працях 1990-х років сформулювали основні принципи цифрової стеганографії та запропонували перші алгоритми вбудовування інформації в цифрові зображення.

В останні десятиліття спостерігається інтенсивний розвиток методів стеганографії, орієнтованих на різні типи цифрових об'єктів: зображення, аудіо, відео, текст та мережевий трафік. Значний внесок у розвиток теоретичних засад цифрової стеганографії зробили такі дослідники, як Дж. Фрідріх, К. Каченгурас, І. Кокс, М. Міллер, які розробили математичні моделі стеганографічних систем та методи оцінки їх ефективності. Вітчизняні науковці В.О. Хорошко, О.Д. Азаров, В.Я. Задірака, А.А. Кобозєва також зробили вагомий внесок у розвиток теоретичних та практичних аспектів стеганографічного захисту інформації.

1.3 Стеганографія в сучасному інформаційному суспільстві: сфери застосування та актуальні тенденції

Інтерес до стеганогрфії в наш час обумовлюється експоненційним зростанням обсягів цифрових даних та інтенсифікацією інформаційних обмінів, стеганографічні технології набувають дедалі більшого значення як перспективний інструмент забезпечення конфіденційності та цілісності інформації. Стеганографія, яка історично розвивалася як мистецтво прихованої комунікації, трансформувалася у високотехнологічну галузь інформаційної безпеки, що інтегрується з різноманітними сферами людської діяльності та вирішує широкий спектр практичних завдань. Комплексний аналіз сучасного

стану та тенденцій розвитку стеганографії дозволяє виявити ключові напрями її застосування та окреслити перспективи подальшої еволюції цієї галузі.

Корпоративний сектор демонструє зростаючий інтерес до стеганографічних технологій як інструменту захисту комерційної таємниці та конфіденційної інформації. В умовах глобалізації бізнесу та цифровізації корпоративних процесів виникає необхідність забезпечення безпеки передачі даних між територіально розподіленими підрозділами компаній. Стеганографічні методи дозволяють створювати приховані канали комунікації, що функціонують паралельно з традиційними каналами і не привертають уваги потенційних зловмисників.

У корпоративних системах електронного документообігу стеганографія використовується для вбудовування метаданих, що містять інформацію про авторство документів, час їх створення та модифікації, а також історію доступу до них. Це забезпечує додатковий рівень контролю за рухом документів та дозволяє виявляти випадки несанкціонованого копіювання або модифікації конфіденційної інформації.

Особливу актуальність у корпоративному секторі набуває текстова стеганографія, зокрема, методи приховування інформації в документах MS Word, Excel та PDF, які є стандартними форматами для ділової комунікації. Компанії розробляють власні стеганографічні системи, інтегровані з корпоративними поштовими серверами та системами документообігу, що автоматично вбудовують у документи приховані маркери, які дозволяють ідентифікувати джерело можливого витоку інформації.

Транснаціональні корпорації, такі як IBM, Microsoft та Oracle, активно інвестують у розробку стеганографічних технологій для захисту інтелектуальної власності. Зокрема, у патентних базах даних США та Європейського Союзу спостерігається щорічне зростання кількості заявок на патентування нових стеганографічних методів, орієнтованих на корпоративне використання.

Банківська та фінансова сфера активно впроваджує стеганографічні технології для підвищення безпеки електронних транзакцій та протидії

шахрайству. Традиційні криптографічні методи, які використовуються для шифрування фінансової інформації, часто привертають увагу зловмисників саме через очевидність факту шифрування. Стеганографія дозволяє створювати додатковий рівень захисту, приховуючи сам факт наявності захищеної інформації.

У банківських системах стеганографічні методи використовуються для вбудовування прихованих водяних знаків у електронні документи, що підтверджують фінансові операції – виписки, чеки, платіжні доручення. Такі водяні знаки містять криптографічні хеш-коди, які дозволяють верифікувати цілісність документа та виявляти будь-які несанкціоновані модифікації.

Комерційні банки впроваджують стеганографічні технології для захисту персональних даних клієнтів. Наприклад, деякі європейські банки використовують стеганографію для приховування біометричних ідентифікаторів у цифрових підписах клієнтів, що дозволяє здійснювати додаткову верифікацію без явного запиту біометричних даних. Це підвищує зручність використання системи для клієнтів при одночасному підвищенні рівня безпеки.

Однією з найбільш розвинених сфер застосування стеганографії є захист авторських прав на цифровий контент. В умовах цифрової економіки, коли інтелектуальна власність стає одним із ключових активів, забезпечення захисту авторських прав на мультимедійний контент – зображення, аудіо, відео, електронні книги – набуває критичного значення. Стеганографічні технології дозволяють вбудовувати в цифрові об'єкти невидимі водяні знаки, що містять інформацію про правовласника та умови використання контенту.

Медіакомпанії та видавництва впроваджують системи цифрових водяних знаків для захисту авторських прав на фотографії, музичні композиції, фільми та електронні публікації. Провідні кіностудії, такі як Warner Bros., Disney та Universal Pictures, використовують стеганографічні технології для індивідуального маркування копій фільмів, що дозволяє ідентифікувати джерело можливого витоку контенту ще на етапі передпрокатного перегляду.

Сучасні системи цифрових водяних знаків характеризуються високою стійкістю до різноманітних трансформацій контенту, включаючи стиснення, кадрування, масштабування та фільтрацію. Це досягається шляхом вбудовування водяних знаків у частотну область зображень та аудіосигналів, що забезпечує їх збереження навіть при значних модифікаціях контейнера. У 2019 році дослідники з Каліфорнійського університету в Берклі представили систему StegaStamp – глибоконейронну стеганографічну модель, яка дозволяє вбудовувати невидимі гіперпосилання в зображення. Ці водяні знаки залишаються стійкими навіть після друку та повторного сканування, що є критично важливим для захисту авторських прав на графічні роботи [7].

Державні органи та спеціальні служби традиційно проявляють інтерес до стеганографічних технологій як інструменту забезпечення конфіденційності комунікації. У дипломатичному листуванні стеганографія використовується для створення прихованих каналів передачі особливо важливої інформації, які функціонують паралельно з офіційними, криптографічно захищеними каналами. Такий підхід дозволяє забезпечити додатковий рівень захисту в умовах, коли традиційні криптографічні методи можуть бути скомпрометовані або коли необхідно приховати сам факт передачі секретної інформації.

Військові відомства активно впроваджують стеганографічні технології для захисту тактичної інформації та забезпечення прихованої комунікації в умовах ведення бойових дій. Особливої актуальності стеганографія набуває в контексті інформаційних та кібернетичних війн, де приховування факту передачі інформації може бути критично важливим для успішного виконання операцій. Військові дослідницькі лабораторії розробляють спеціалізовані стеганографічні протоколи для прихованої передачі даних через супутникові канали зв'язку, радіочастотний спектр та мережі передачі даних.

У сфері захисту державної таємниці стеганографія використовується як додатковий рівень безпеки, що доповнює традиційні криптографічні методи. Зокрема, у системах електронного документообігу з грифом секретності впроваджуються методи стеганографічного маркування, що дозволяють

відстежувати доступ до документів та ідентифікувати потенційні джерела витоку інформації.

1.4 Види стеганографічних методів

У межах сучасної концепції захисту інформації стеганографія розглядається як дисципліна, що вивчає методи і засоби приховання факту передачі або збереження повідомлень, забезпечуючи їх недоступність для несанкціонованих суб’єктів без залучення очевидних механізмів шифрування. Її прикладне значення набуває особливої ваги в умовах, коли необхідно забезпечити латентний канал комунікації, що не виявляється стандартними засобами контролю або моніторингу. Методологічно стеганографію поділяють на дві великі групи (рис. 1.1) – технологічну та інформаційну, кожна з яких включає різнорідні способи реалізації прихованого передавання або зберігання даних, що базуються на фізичних, хімічних або структурно-семантичних властивостях носія [6].

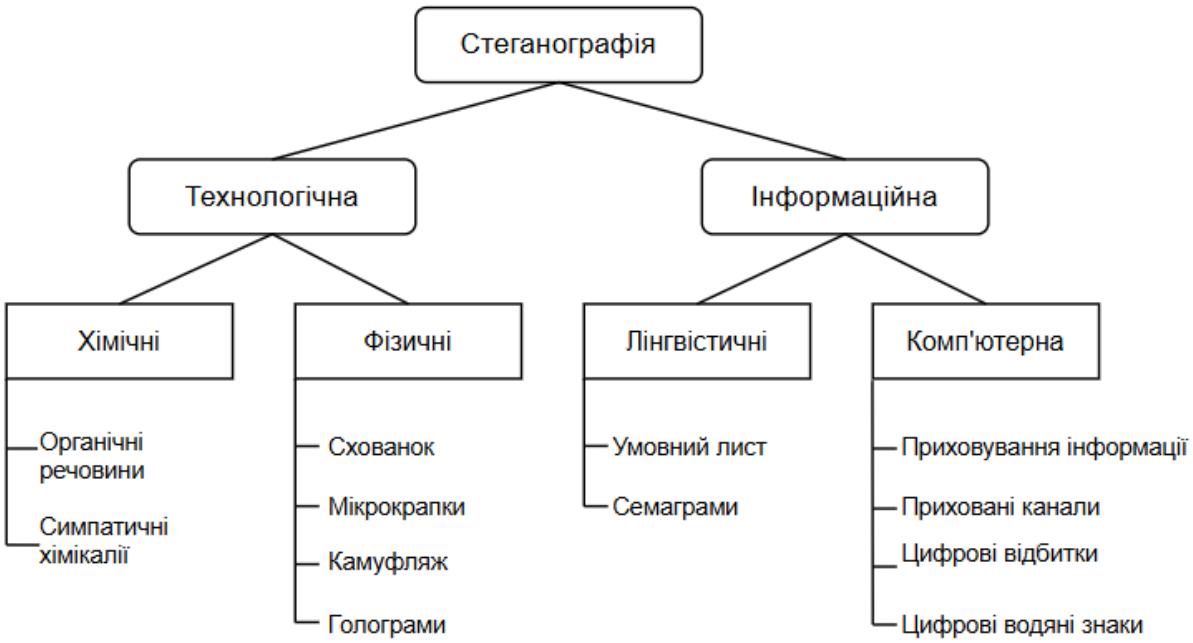


Рисунок 1.1 – Структурна схема видів стеганографічних методів

З технологічної точки зору, до стеганографічних засобів відносять як хімічні, так і фізичні методи. Перші передбачають застосування органічних речовин, здатних змінювати або відновлювати свої оптичні властивості під впливом зовнішніх чинників (наприклад, температури, світла чи вологи), що дозволяє формувати симпатичні чорнила – сполуки, які в нормальному стані є невидимими. Фізичні ж методи базуються на модифікації форми, структури або розміщення матеріального середовища: створення схованок, впровадження мікрооб'єктів (мікрокрапок), застосування камуфляжних візерунків, а також використання голографічних технологій, що дозволяють імплантувати візуальні структури у вигляді об’ємних зображень у середовищі з оптичною неоднорідністю. Зазначені підходи демонструють високу стійкість до виявлення, однак значно обмежені щодо пропускної здатності та можливості цифрової інтеграції.

Інформаційна стеганографія, своєю чергою, оперує цифровими або текстовими носіями інформації, реалізуючи приховування шляхом зміни формального представлення даних без суттєвого впливу на їх зовнішній вигляд. У межах лінгвістичних методів формуються умовні листи – тексти, що на перший погляд не викликають підозри, але за попередньо погодженими правилами містять приховані смислові конструкції або семаграми, в яких певна послідовність символів, відстаней, пунктуаційних знаків або відхилень у форматуванні виступає носієм секретної інформації. Подібні методи дозволяють використовувати звичайні документи (зокрема, текстові файли), що не виявляються при автоматизованому аналізі за критеріями криптографічної обробки.

Особливої уваги заслуговує напрям комп’ютерної стеганографії, що оперує алгоритмічними процедурами зміни структури цифрових файлів без впливу на їхню функціональність або сприйняття користувачем. Зокрема, виділяють методи приховування інформації у внутрішніх структурах файлів (наприклад, за допомогою least significant bit embedding), формування прихованих каналів у мережевих протоколах, створення цифрових відбитків, які

дозволяють ідентифікувати джерело або власника контенту, а також застосування цифрових водяних знаків для забезпечення автентичності і прав власності на мультимедійні об’єкти. На відміну від класичних підходів, ці методи мають високу щільність прихованого передавання, добру сумісність із сучасними засобами обробки інформації та потенціал для автоматизації впровадження.

Таким чином, представлена класифікаційна схема стеганографічних засобів ілюструє багатовимірність цієї галузі знань, де об’єктами впливу можуть бути як матеріальні, так і цифрові носії, а способи реалізації охоплюють як фундаментальні фізико-хімічні принципи, так і тонко організовані семантичні структури. У цьому контексті методи форматування символів у середовищі текстових процесорів – зокрема, Microsoft Word – набувають особливого значення, оскільки поєднують зручність використання, високу доступність та можливість інкапсуляції інформації в елементах форматування, які не змінюють візуального сприйняття документа. Така особливість дозволяє розглядати їх як перспективний напрям інформаційної стеганографії, що забезпечує інтеграцію захисних механізмів без зміни семантики або функціональності документа, одночасно забезпечуючи високий рівень стійкості до виявлення.

Висновки до розділу 1

Проаналізований у першому розділі теоретичний матеріал дозволяє констатувати, що стеганографія як наука про приховане передавання інформації набуває дедалі більшого значення в умовах інтенсифікації інформаційних обмінів та зростання загроз інформаційній безпеці. Ретроспективний аналіз еволюції стеганографічних методів від найдавніших часів до сучасних цифрових технологій демонструє, що кожна технологічна епоха привносила нові підходи до інкапсуляції секретної інформації з урахуванням актуальних комунікаційних засобів, технічних можливостей та соціальних потреб.

Концептуальні засади стеганографії, які базуються на фундаментальних положеннях теорії інформації, семіотики та системного аналізу, формують міцний теоретичний фундамент для реалізації прихованих каналів комунікації в умовах відкритого інформаційного простору. Високий рівень наукової розробленості проблеми стеганографічного захисту інформації досягається завдяки органічному синтезу математичних моделей з інженерними рішеннями, що дозволяє створювати ефективні механізми приховування даних у різноманітних типах носіїв без зміни їхньої функціональності або семантичного сприйняття.

Аналіз методологічних аспектів стеганографії свідчить про її структурну багатовимірність, що охоплює як технологічні підходи (хімічні та фізичні), так і інформаційні методи (лінгвістичні та комп'ютерні). Такий диверсифікований інструментарій дозволяє застосовувати стеганографічні технології в різноманітних контекстах, адаптуючи їх до специфічних потреб та обмежень конкретного середовища передавання інформації.

РОЗДІЛ 2
МЕТОДИ І ЗАСОБИ СТЕГАНОГРАФІЧНОГО ПРИХОВУВАННЯ ДАНИХ

2.1 Основна термінологія

Стеганографічна система – це сукупність методів, алгоритмів і засобів, які реалізують процес прихованого передавання інформації шляхом інкапсуляції повідомлення у нейтральний носій (контейнер) таким чином, щоб зберегти непомітність передачі для несанкціонованих спостерігачів. Типова стеганографічна система складається з підсистеми вбудовування (інкапсуляції), підсистеми витягання (декодування), стегоканалу, а також функціональних модулів керування та безпеки [8].

Контейнер у стеганографії – це нейтральний, зовні несекретний цифровий об’єкт (наприклад, текстовий документ, зображення, аудіо- або відеофайл), структура якого дозволяє внести до нього приховане повідомлення без порушення його функціональності або перцептивної цілісності.

Пустий контейнер – контейнер, який ще не містить вбудованого прихованого повідомлення; використовується як основа для інкапсуляції.

Заповний контейнер – контейнер, у структуру якого вже інтегровано приховане повідомлення за допомогою певного методу вбудовування. При цьому зовнішній вигляд або функція об’єкта залишаються незмінними або змінюються в межах статистично непомітних відхилень [9].

Вбудоване приховане повідомлення – це заздалегідь підготовлений інформаційний вміст (текстовий, числовий, бітовий тощо), що є предметом захисту, який було інтегровано до контейнера відповідно до визначеного стеганографічного методу. Повідомлення може бути попередньо зашифроване або оброблене, що забезпечує додатковий рівень безпеки в разі виявлення. Основна властивість прихованого повідомлення – його неможливість

візуального або функціонального виявлення без наявності стегоключа чи знання методу.

Стеганографічний канал – це логічна або фізична комунікаційна структура, яка забезпечує передавання контейнерів із прихованою інформацією між відправником і отримувачем із дотриманням вимог конфіденційності, автентичності та непримітності. Стегоканал має властивість бути неінтрузивним, тобто не викликати підозри в наявності секретної інформації навіть при спостереженні сторонніми особами. У загальному випадку стегоканал може функціонувати в межах стандартних інформаційних потоків, таких як електронна пошта, файловий обмін, онлайн-платформи тощо.

Стегоключ – це параметр, код або криптографічна змінна, яка регламентує процес вбудовування та/або витягання прихованого повідомлення зі стеганографічного контейнера. Стегоключ може визначати місця або правила модифікації символів, бітів, форматувальних ознак чи інших елементів контейнера. У залежності від схеми, ключ може бути симетричним (спільним для обох сторін) або асиметричним (відмінним для відправника і отримувача). Без знання або відтворення стегоключа процес коректного витягання повідомлення є або неможливим, або суттєво ускладненим.

2.2 Узагальнена модель стегосистеми

Рис. 2.1 відображає узагальнену модель стегосистеми, що демонструє повний цикл стеганографічного приховування та вилучення інформації. Модель складається з двох взаємопов’язаних частин: процесу вбудовування на стороні відправника та процесу вилучення на стороні одержувача [10].

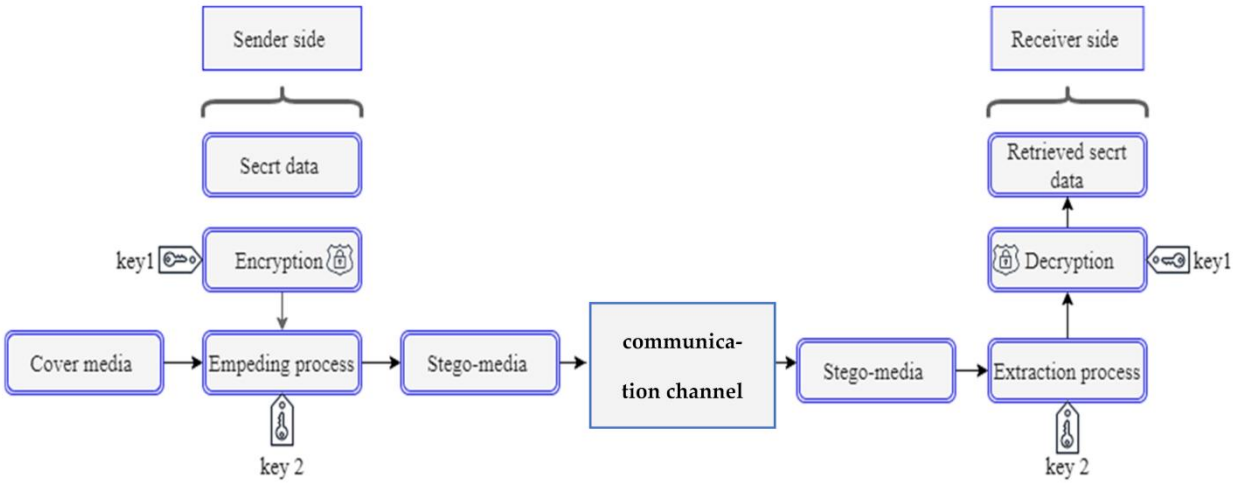


Рисунок 2.1 – Узагальнена схема для стеганосистеми

Системи, які прагнуть покращити характеристики безпеки, часто використовують ключ безпеки, структуру шифрування або обидва ці елементи під час вбудовування. Ключ може містити додаткову інформацію, таку як паролі шифрування, карти вбудовування та порогове значення, яке використовується для вибору певного коефіцієнта в процесі вбудовування. Загалом, вбудована система може бути позначена як:

$$c' = \mathrm{Em}(\mathrm{C}, \mathrm{En}(\mathrm{S}, \mathrm{k}_1), \mathrm{k}_2), \tag{2.1}$$

де c' - заповнений контейнер;

C - контейнер;

S - приховане повідомлення;

k - секретний ключ;

En - функція шифрування;

Em - функція вбудовування.

Для того, щоб функція En виконувала шифрування, необхідні два секретні ключі (k_1 та k_2), які використовуються разом із Em для виконання вбудовування. Потім c' надсилається отримувачу, і алгоритм розшифрування та вилучення повідомлення, що позначається як рівнянні 2.2:

$$s_r = D(Ex(c', k_2)k_1) \quad (2.2)$$

де s_r - це відновлене приховане повідомлення;

Ex - функція вилучення;

D - функція розшифрування;

c' - заповнений контейнер.

Ця узагальнена модель стеганосистеми фундаментально відображає симетричність процесів вбудовування та вилучення інформації, де ключ виступає критичним параметром, що забезпечує безпеку та функціональність усієї системи. Модель демонструє, що стеганографічний захист інформації ґрунтується на трансформації даних без зміни зовнішніх характеристик контейнера настільки, щоб факт наявності прихованої інформації залишався невиявленим для сторонніх спостерігачів, які не володіють відповідним ключем.

Надійність стеганографічного зв'язку полягає, головним чином, складності процедури виявлення стеганограми [6]. Якщо зловмисник має контроль над секретними даними в стеганографічній системі, то така система вважається більше не безпечною [11].

До стеганографічної системи зазвичай висувають такі основні вимоги:

1. Як стеганографічне перетворення має застосовуватися загальновідомий алгоритм і секретний стегоключ.

2. Метод приховування повинен забезпечити автентичність цілісність файлу.

3. Тільки за наявності правильного стегоключа можна виявити, витягти і довести існування прихованого повідомлення.

4. Навіть, якщо той, хто атакує, знає про факт існування прихованого повідомлення (або саме повідомлення), це не повинно дозволити йому довести даний факт третій особі і, тим паче, знайти подібні повідомлення в інших повідомленнях, поки стегоключ зберігається в таємниці.

5. Ніхто не повинен знайти який-небудь статистичний доказ існування прихованого повідомлення, його виявлення без знання ключа повинно бути обчислювальноскладною задачею [6].

2.3 Методи комп’ютерної стеганографії

Переважна кількість сучасних методів комп’ютерної стеганографії базується на застосуванні двох фундаментальних концептуальних положень. По-перше, цифрові об’єкти, які не потребують абсолютної точності відтворення (зокрема, растрові або векторні зображення, аудіосигнали, відеопотоки та інші мультимедійні формати), допускають внесення контрольованих змін в їхню структуру без істотної деградації експлуатаційних властивостей або функціональної придатності. По-друге, через обмежену чутливість сенсорних механізмів сприйняття людини до незначних відхилень в цифровому представленні таких даних, а також у зв’язку з відсутністю у більшості спостерігачів спеціалізованих засобів автоматизованого аналізу, ймовірність виявлення модифікацій зазвичай є низькою.

У контексті комп’ютерної стеганографії під середовищем передавання інформації розуміється цифровий носій, структура якого формується із мультимедійних компонентів, адаптованих до обробки, збереження та передавання засобами обчислювальної техніки та інформаційно-комунікаційних систем. Усі такі носії, включаючи, але не обмежуючись графічними файлами, аудіотреками, відеопотоками, HTML-документами тощо, піддаються формалізованому поданню у вигляді дискретного цифрового потоку, що є об’єктом стеганографічного втручання.

Структурно реалізація методів прихованого передавання даних полягає у виокремленні в межах контейнера таких інформаційних одиниць, які не мають суттєвого впливу на його семантичне навантаження або технічні характеристики, із подальшою підміною вмісту цих фрагментів на компоненти повідомлення, яке підлягає приховуванню. Конкретна процедура підміни

визначається параметрами обраного алгоритму або методології стеганографічного вбудовування.

У межах такого підходу під інформаційним кадром (структурною одиницею цифрового середовища) доцільно розуміти семантично або функціонально виділений фрагмент цифрового об'єкта, що розглядається як самостійна одиниця в структурі контейнера. Залежно від прикладного середовища, кадром може виступати окремий графічний файл, звуковий фрагмент, HTML-сторінка або інший логічно завершений елемент даних, що підлягає частковому або повному модифікуванню відповідно до стеганографічних процедур.

Для методів комп'ютерної стеганографії можна ввести певну класифікацію (рис. 2.2) [6].

Рисунок 2.2 – Класифікація методів комп'ютерної стеганографії

Спосіб вибору контейнера у методах комп’ютерної стеганографії визначає алгоритмічний підхід до визначення чи формування носія, у якому вбудовується приховане повідомлення. Він виступає ключовим чинником, що впливає на ефективність, стійкість та непомітність стеганографічного процесу, адже саме контейнер визначає можливий обсяг, якість і надійність приховування інформації.

Сурогатні методи ґрунтуються на використанні контейнерів, які створюються або підбираються таким чином, щоб імітувати природні об’єкти медіасередовища, але насправді призначені для цілеспрямованого приховування даних. Сурогатні контейнери формуються заздалегідь і містять штучно створені чи змінені елементи, що мають відповідати певним критеріям для маскування повідомлення. Важливою характеристикою є можливість підтримувати достатній рівень правдоподібності, щоб уникнути детекції з боку сторонніх.

Селективні методи передбачають, що із наявного множинного набору медіаоб’єктів (файлів, фрагментів тощо) обираються лише ті, які найбільш відповідають вимогам стеганографічної задачі з огляду на специфічні характеристики контейнера – наприклад, рівень шуму, розмір, формат або структура даних. Вибір здійснюється на основі аналізу параметрів контейнерів із метою максимального приховування, зменшення візуальних чи аудіальних спотворень, а також підвищення стійкості до атаки чи втручання.

Конструювальні методи передбачають безпосереднє створення контейнера шляхом формування чи модифікації вихідних даних з урахуванням алгоритмічних особливостей стеганографічного кодування. У цьому випадку контейнер не існує у вигляді незалежного об’єкта до початку процесу вбудовування, а конструюється таким чином, щоб забезпечити оптимальну інтеграцію прихованого повідомлення, при цьому гарантуючи його цілісність та мінімальний ризик виявлення.

Доступ до інформації характеризує спосіб, за допомогою якого прихована інформація може бути введена у контейнер або вилучена з нього. Цей параметр

суттєво визначає архітектуру стеганографічної системи, а також рівень безпеки й стійкості до несанкціонованого доступу.

Потоковий доступ передбачає безперервний або послідовний режим введення та/або вилучення прихованих даних. У цьому випадку інформація обробляється у вигляді послідовних потоків, що може бути реалізовано, наприклад, у реальному часі або в процесі передачі мультимедійних даних. Такий тип доступу є характерним для систем, де приховані повідомлення інтегруються у безперервні медіапотоки (відео, аудіо), що забезпечує відсутність явних меж у структурі вбудованих даних і підвищує непомітність стеганографічного каналу.

Доступ з довільним доступом передбачає можливість адресного або вибіркового запису та зчитування прихованої інформації у контейнері. Цей тип доступу характерний для систем, де вбудоване повідомлення розподілене по різних сегментах контейнера, до яких можна звертатися у довільному порядку, що підвищує гнучкість використання стеганографічних методів. Дозволяється як послідовне, так і вибіркове вилучення фрагментів прихованих даних, що сприяє більш ефективній обробці та зменшенню ймовірності втрати інформації.

Тип організації контейнера визначає структуру розташування та систематизації інформаційного носія, який використовується для приховування секретних даних. Даний параметр впливає на спосіб інтеграції прихованого повідомлення в контейнер, а також на ефективність і надійність стеганографічної системи.

Систематична організація контейнера характеризується структурованим, впорядкованим розміщенням інформації, що базується на заздалегідь визначеному правилі або схемі. Такий підхід передбачає чітке визначення позицій, в яких вбудовується приховане повідомлення, з урахуванням структурних особливостей контейнера. Використання систематичної організації забезпечує передбачуваність та однозначність доступу до вбудованої інформації, що спрощує процес кодування і декодування, а також дозволяє оптимізувати

алгоритми захисту і підвищує стійкість стегоканалу до випадкових спотворень чи цілеспрямованих атак.

Несистематична організація контейнера передбачає нерегулярне, випадкове або псевдовипадкове розташування прихованої інформації всередині контейнера без суворої структури. Такий підхід ускладнює виявлення та вилучення прихованих даних сторонніми особами, підвищуючи рівень секретності. Водночас це може призводити до збільшення складності алгоритмів кодування та декодування, а також знижувати ефективність використання ємності контейнера, оскільки відсутність систематичності у розміщенні інформації може спричиняти втрати корисного об’єму для вбудованих даних.

Тип інформаційного середовища визначає категорію цифрових даних або медіа, що виступають у ролі контейнера для вбудованого прихованого повідомлення. Вибір конкретного типу інформаційного середовища є ключовим фактором, який визначає особливості реалізації стеганографічних алгоритмів, їхню ефективність, а також рівень захищеності прихованої інформації.

Текстове середовище – передбачає використання цифрових документів, що містять символічну інформацію у вигляді тексту, наприклад, файли форматів TXT, DOC, DOCX, PDF тощо. Особливістю текстових середовищ є високий рівень структурованості і лінгвістичних закономірностей, що накладає певні обмеження на вбудовування прихованої інформації без помітних змін. При цьому методи стеганографії в текстових документах часто базуються на зміні форматування символів, пропусків, інтервалів, а також на використанні кодових замін.

Звукове середовище – включає цифрові аудіофайли, що містять звукову інформацію, наприклад, WAV, MP3, AAC. Особливості звукового середовища пов’язані з високою чутливістю людського слуху до спотворень, що накладає обмеження на глибину модифікації. Методи стеганографії тут ґрунтуються на використанні психологічних особливостей сприйняття звуку, наприклад, шляхом вбудовування даних у менш помітні частоти або у варіанти амплітудних коливань.

Стоп-кадр/відео середовище – включає цифрові графічні зображення та відеозаписи у форматах JPEG, PNG, BMP, AVI, MP4 тощо. Такі середовища характеризуються значною обсягом інформації та надмірністю даних, що дозволяє розміщувати приховані повідомлення у малозначущих частинах медіафайлу (наприклад, у найменш значущих бітах пікселів). Відео середовище також має особливості часової послідовності кадрів, що відкриває додаткові можливості для приховування даних.

Використовуваний принцип у методах комп'ютерної стеганографії визначає базову концепцію або підхід, на основі якого здійснюється вбудовування прихованої інформації у контейнер та організація стеганографічного каналу. Цей параметр класифікації відображає методологію, що зумовлює вибір конкретних алгоритмів і технік приховування даних, а також впливає на стійкість, ємність і непомітність стеганографічної системи.

Принцип на основі надлишковості середовища – базується на використанні тих елементів контейнера, які містять надлишкову, або несуще важливу інформацію, зайву для відтворення чи сприйняття медіа без помітних спотворень. Ці надлишкові елементи слугують «носіями» для вбудовування прихованого повідомлення без суттєвого впливу на якість або функціональність контейнера. Наприклад, у зображеннях такими надлишковими елементами можуть бути найменш значущі біти пікселів, у аудіофайлах – несуттєві гармоніки або частотні компоненти.

Структурний принцип – передбачає зміну або модифікацію структурних компонентів контейнера, які мають семантичне або функціональне значення, але можуть бути варіативними або допускають альтернативні представлення. До структурних елементів можуть належати, зокрема, форматування тексту, розташування абзаців, порядок кадрів, тимчасові інтервали між елементами, структура метаданих. Вбудовування інформації відбувається за рахунок маніпуляцій цими структурними характеристиками, що може бути складнішим для виявлення, проте вимагає більш складних алгоритмів кодування і декодування.

Призначення у контексті стеганографічних систем визначає функціональне спрямування та роль, яку відіграє метод або технологія прихованої передачі інформації в інформаційному просторі, зокрема в системах захисту даних. Це поняття відображає мету застосування стеганографічних методів, яка полягає у забезпеченні конфіденційності, цілісності та невиявності інформації в процесі її передачі або зберігання.

Основними аспектами призначення є:

- забезпечення прихованості комунікації, що полягає у маскуванні факту існування переданої інформації, що істотно підвищує рівень безпеки в порівнянні з традиційними криптографічними методами, які лише шифрують вміст повідомлення.
- підтримка цілісності та автентичності інформації за рахунок використання додаткових механізмів виявлення змін або несанкціонованого втручання у приховане повідомлення.
- інтеграція з наявними цифровими середовищами, що дозволяє безпечно вбудовувати секретні дані в мультимедійні об’єкти без порушення їх функціональних характеристик і якості.
- реалізація додаткових рівнів інформаційного захисту у комплексі заходів кібербезпеки, спрямованих на захист інформаційних систем від несанкціонованого доступу, аналізу та перехоплення.

2.4 Текстова стеганографія

У сучасних умовах розвитку інформаційних технологій переважна більшість стеганографічних методів функціонує в інформаційних середовищах, що характеризуються високим рівнем надлишковості, яка може бути використана як ресурс для прихованого впровадження додаткових даних без істотного порушення функціональної цілісності основного повідомлення [12]. Однак, на відміну від таких медіаоб'єктів, як аудіосигнали чи зображення, які за своєю природою містять значну кількість шумових компонентів та структурної

надлишковості, письмовий текст демонструє суттєво нижчий рівень таких характеристик, що істотно ускладнює процес вбудовування прихованої інформації без помітного впливу на його лінгвістичну або візуальну структуру [13].

Водночас, методи лінгвістичної стеганографії, що передбачають інтеграцію прихованих повідомлень у текстові структури, мають давню історію та відомі ще з епохи середньовіччя. Їх принципова основа полягає у використанні або природної надлишковості мови, що проявляється на морфологічному, синтаксичному чи семантичному рівнях, або у використанні візуально-структурних особливостей текстових повідомлень, таких як форматування, розміщення символів, міжсимвольні інтервали тощо [14].

З урахуванням інтенсивного розвитку засобів комп'ютерної обробки текстової інформації, традиційні підходи до лінгвістичної стеганографії зазнали трансформації, що дозволило реалізувати сучасні методи приховування даних на якісно новому технічному рівні. Застосування таких методів уможливлює формування комунікаційних каналів із високим рівнем латентності, які здатні залишатися непоміченими як для автоматизованих систем контент-аналізу, що функціонують у телекомунікаційних мережах, так і для людини, що здійснює ручну перевірку.

Сучасні реалізації лінгвістичних стеганографічних алгоритмів базуються на низці ключових підходів (табл. 2.1).

Таблиця 2.1

Класифікація методів текстової стеганографії

Метод	**Короткий опис**
Модифікації форматної структури текстових документів	Передбачає використання різного роду візуально непомітних змін, пов'язаних з маніпуляціями над шрифтами, інтервалами, регістром символів та іншими параметрами оформлення тексту

Продовження табл. 2.1

Синтаксичні методи	Реалізуються шляхом переформулювання речень або зміни їх граматичної структури з метою вбудовування прихованої інформації без порушення граматичних норм
Семантичні методи	Засновані на варіативності смислових конструкцій природної мови, що дозволяє приховувати повідомлення через синонімічну заміну, перебудову висловлювань або використання багатозначних мовних одиниць
Генеративного типу	Стеганограма формується на основі вхідного секретного повідомлення шляхом створення тексту, який одночасно несе задану приховану інформацію та відповідає нормам природної мови

Методи модифікації форматної структури текстових документів передбачають використання прихованих повідомлень через візуально непомітні зміни, які не впливають на змістовне сприйняття тексту з боку користувача. До таких змін належать маніпуляції зі шрифтовими параметрами (наприклад, незначне збільшення або зменшення розміру символів, вибір схожих гарнітур шрифтів), варіації міжрядкових та міжсимвольних інтервалів, регістрів окремих літер або вставка спеціальних символів (таких як символи нульової ширини - zero-width space) [15]. Перевагою цього підходу є його високий ступінь непомітності для людського зору, однак чутливість до форматного перетворення (наприклад, при копіюванні тексту між різними редакторами або збереженні у форматах з різною підтримкою стилізації) може негативно впливати на цілісність прихованого повідомлення.

Синтаксичні методи лінгвістичної стеганографії базуються на трансформаціях граматичної структури речень із метою кодування бітів інформації у виборі синтаксичних конструкцій. Наприклад, те саме

повідомлення може бути сформульоване за допомогою активного або пасивного стану, прямого чи інверсованого порядку слів, без порушення граматичних норм мови [16]. Стійкість таких методів значною мірою визначається природністю побудованих речень і відсутністю лінгвістичних аномалій, що можуть бути виявлені під час автоматичного аналізу. Водночас ефективність цього підходу залежить від глибини морфосинтаксичного аналізу та ступеня варіативності мовного матеріалу.

Семантичні методи використовують смислову надлишковість природної мови для приховування інформації. Найчастіше застосовується синонімічна заміна ключових слів і словосполучень, перестановка фраз із збереженням загального змісту висловлювання або використання багатозначних лексем [17]. Наприклад, вибір між синонімами «помилка» і «похибка» може сигналізувати про логічну «одиницю» або «нуль» у прихованому повідомленні. Зазначені методи потребують наявності розгалужених лексичних баз, які дозволяють гарантувати семантичну еквівалентність варіантів і водночас варіативність для кодування інформації. При цьому зберігається загальна стилістична й лексична природність тексту, що ускладнює виявлення прихованого впливу при поверхневому читанні або статистичному аналізі.

Окрему групу становлять методи генеративного типу, які передбачають синтез нових текстових об’єктів на основі заздалегідь заданого секретного повідомлення. У цьому випадку стеганограма створюється не шляхом модифікації наявного тексту, а як результат генерації, де процес вибору наступного слова або фрази обумовлюється не лише статистичними мовними моделями (наприклад, n-грамами або трансформерними архітектурами), а й значеннями бітів, що кодують приховану інформацію. Такі алгоритми дозволяють досягати високої природності синтезованого тексту, водночас залишаючи мінімальні сліди впливу зовнішньої інформації на його структуру. Метод потребує значних обчислювальних ресурсів, натомість забезпечує високу стійкість до виявлення, особливо у випадку використання адаптивних моделей мовного моделювання.

Таким чином, кожен із розглянутих підходів до лінгвістичного стеганографічного приховування має як переваги, так і обмеження, що зумовлені як характеристиками природної мови, так і властивостями стеганографічної системи в цілому. Вибір конкретного методу здійснюється з урахуванням необхідного балансу між обсягом переданої інформації, стійкістю до модифікацій та виявлення, а також збереженням легітимного вигляду текстового контейнера.

2.5 Недоліки та виклики текстової стеганографії

Попри концептуальну привабливість текстової стеганографії як методу прихованої передачі інформації в умовах контролю цифрового простору, її практична реалізація супроводжується низкою суттєвих обмежень, недоліків і технічних ускладнень, що знижують загальну ефективність застосування в реальних інформаційних системах. Перш за все, варто акцентувати увагу на тому, що текст, на відміну від мультимедійних даних (зображень, аудіо- та відеопотоків), не характеризується значним рівнем надлишковості, що обмежує доступний діапазон змін без істотного впливу на смислову або структурну цілісність повідомлення. У результаті, потенціал вбудовування прихованої інформації у текстові об’єкти є відносно низьким, що безпосередньо впливає на корисну ємність стеганоканалу [18].

Одним із фундаментальних недоліків текстової стеганографії є її обмежена стійкість до редакторських модифікацій. Більшість методів, заснованих на візуально-структурних параметрах, таких як форматування, регістрові відмінності, інтервали або використання нульової ширини символів, виявляються надзвичайно вразливими до автоматичного переформатування тексту, копіювання між різними програмними середовищами, перекодування в інші формати, а також до дій користувача, пов’язаних із редагуванням. У таких випадках структура прихованого повідомлення може бути безповоротно

пошкоджена або повністю знищена, що унеможливлює його коректне вилучення на приймальній стороні.

Крім того, значною проблемою є висока чутливість до стилістичних аномалій. Синтаксичні та семантичні методи, що базуються на трансформації речень або заміні слів синонімами, часто призводять до появи мовних конструкцій, які не характерні для природного мовлення. Навіть незначне зниження стилістичної природності тексту може бути розпізнане як маркер наявності зовнішнього втручання, особливо в умовах застосування сучасних систем обробки природної мови (Natural Language Processing, NLP), здатних виявляти відхилення від типової лінгвістичної моделі [19]. Така уразливість особливо критична у випадках, коли стеганотекст має проходити через автоматизовані системи перевірки в умовах цензури або інформаційного моніторингу.

Наступним обмежувальним фактором є низька пропускна здатність текстових контейнерів. Навіть у випадку граничного використання допустимих методів модифікації без порушення природності тексту, обсяг даних, що можуть бути приховані в одиниці інформаційного об’єкта, залишається істотно меншим порівняно з графічними чи аудіо файлами. Зокрема, типові реалізації дозволяють приховати лише декілька бітів на кожне речення, що обмежує загальну продуктивність каналу прихованої передачі і робить такі методи непридатними для передавання значних обсягів даних без істотного збільшення розміру контейнера.

Окрему категорію проблем складає лінгвістична і культурна залежність алгоритмів, що реалізують текстову стеганографію. Багато методів вимагають глибокого розуміння граматичних, синтаксичних та семантичних особливостей мови, для якої вони були розроблені. Це означає, що адаптація таких методів до іншомовних середовищ потребує окремої розробки лінгвістичних моделей і словників, що значно ускладнює масштабування або міжнародне застосування відповідних систем. Зокрема, методи, ефективні для англійської мови, часто

виявляються непридатними для мов із вільним порядком слів (наприклад, української), багатою флексією або складною морфологією [20].

Ще одним істотним обмеженням є уразливість до лінгвістично-статистичного аналізу, особливо у випадках, коли стеганограма містить великі обсяги вбудованої інформації. Оскільки природні тексти підкоряються певним статистичним закономірностям (наприклад, закону Ципфа щодо розподілу частот слів), порушення таких закономірностей може бути виявлене за допомогою алгоритмів машинного навчання або евристичних методів. Зокрема, порівняння частотного профілю підозрілого тексту із профілем референтної вибірки дозволяє виявити аномалії, які можуть слугувати індикаторами прихованого вмісту [21].

У випадку застосування методів генеративного типу, що передбачають створення тексту на основі секретного повідомлення, постає проблема забезпечення достатнього рівня природності синтезованого тексту. Хоча сучасні мовні моделі (наприклад, GPT-подібні архітектури) демонструють високий ступінь семантичної та синтаксичної природності, контрольоване вбудовування інформації в процес генерації тексту призводить до компромісу між якістю тексту та обсягом закодованих даних. Крім того, потреба в потужних обчислювальних ресурсах і складність забезпечення детермінованої декодування вимагають спеціалізованого програмного забезпечення як на боці відправника, так і на боці отримувача.

Зазначені недоліки актуалізують потребу в розробці нових методів підвищення стійкості текстових стеганографічних систем до виявлення та модифікацій, з урахуванням динамічного розвитку технологій обробки текстової інформації. Перспективними напрямами у цьому контексті є використання адаптивних лінгвістичних моделей, глибоке контекстне кодування, залучення елементів криптографії для ускладнення виявлення сигналу, а також інтеграція текстової стеганографії з іншими формами інформаційного маскування, зокрема, у мультимодальних середовищах.

2.6 Метод найменш значущих бітів

Метод найменш значущих бітів (LSB) є фундаментальною технікою стеганографічного приховання інформації, яка ґрунтується на принципі мінімальної модифікації носія даних [9]. Суть методу полягає в заміні найменш значущих бітів контейнера (зображення, аудіофайлу, текстового документа) бітами секретного повідомлення без помітних змін у структурі вихідного файлу.

Теоретичні основи методу базуються на властивості найменш значущих бітів мати мінімальний вплив на загальне сприйняття інформації. У цифрових системах кожен елемент даних представлений бінарною послідовністю, де старші біти визначають основні характеристики, а молодші біти мають мінімальний вплив на інформативність.

Метод LSB описується наступним алгоритмом:

1. Представлення вихідного контейнера у бінарному форматі.
2. Розкладання секретного повідомлення на бітову послідовність.
3. Послідовна заміна найменш значущих бітів контейнера бітами повідомлення.

Для прикладу, використання методу в колірній моделі RGB зображення полягає в розподіленні бітів вкладення між трьома колірними каналами пікселя зображення, що забезпечує високу інформаційну місткість, мінімальні спотворення вихідного зображення, складність несанкціонованого виявлення прихованої інформації.

Головна перевага метода LSB у стеганографії з використанням колірної моделі RGB (рис. 2.3) – здатність приховувати значні обсяги інформації практично без помітних змін у структурі контейнера, що робить метод надзвичайно ефективним для конфіденційних комунікацій [22].

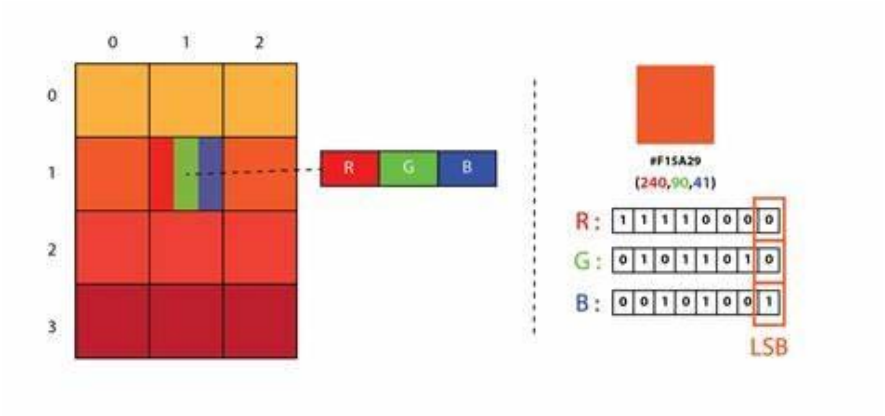


Рисунок 2.3 – Використання методу LSB у колірній моделі RGB

2.7 Документ MS Word як стегаконтейнер

Документ Microsoft Word (DOCX) являє собою комплексну структуру, котра надає значний потенціал для імплементації стеганографічних методів захисту інформації. Файловий формат DOCX, впроваджений корпорацією Microsoft з версії Word 2007, фактично представляє собою архів, скомпресований за технологією ZIP, що містить низку XML-файлів та директорій, що забезпечують структуроване зберігання текстової інформації, форматування, метаданих та додаткових мультимедійних компонентів (рис. 2.4) [23, 24]. Дана архітектура документа створює обширний простір для вбудовування конфіденційної інформації в різноманітні структурні елементи, що забезпечує високий ступінь прихованості повідомлень від несанкціонованого доступу.

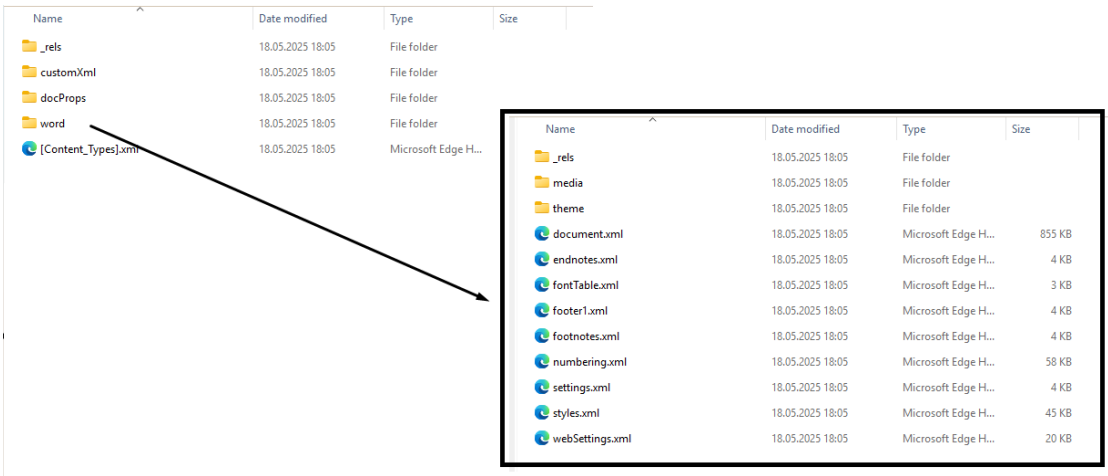


Рисунок 2.4 – Структура Word документа як zip-архіва

При декомпресії документа DOCX виявляється розгалужена ієрархія файлів, основними компонентами якої є: директорія «_rels», що містить дані про взаємозв'язки між частинами документа, каталог «docProps», який зберігає метадані та налаштування документа, директорія «word», де розміщується основна інформація про вміст документа. Зокрема, у директорії «word» містяться такі ключові файли, як «document.xml», що зберігає основний текстовий вміст, «styles.xml», який відповідає за визначення стилів форматування, «settings.xml», що фіксує користувацькі параметри, «numbering.xml», котрий визначає нумерацію елементів, «theme1.xml», який встановлює тематичні кольори та шрифти, директорія «media», де зберігаються вбудовані зображення, аудіо та відео матеріали. Кожен з означених компонентів становить потенційний стеганографічний контейнер, здатний приховувати інформацію при незначній модифікації його атрибутів та параметрів.

Стеганографічні методи, орієнтовані на використання документів MS Word у якості контейнерів, характеризуються значною ефективністю внаслідок надзвичайно великої кількості параметрів форматування, що можуть бути застосовані до текстових фрагментів. Ці параметри дозволяють імплементувати різноманітні підходи до приховування даних, зокрема: мікромодифікації міжсимвольних інтервалів (трекінг); варіювання міжрядкових інтервалів (інтерліньяж); маніпулювання параметрами шрифтів (незначні зміни розміру, насиченості, нахилу); впровадження невидимих розривів рядків та додаткових пробільних символів; приховування даних у метаінформації документа; алгоритмічне модифікування параметрів XML-тегів у структурі файлів контейнера [25]. Означені методи надають можливість впроваджувати значні об'єми стеганографічної інформації без візуальної детекції модифікації параметрів документа автентичним користувачем.

Фундаментальною особливістю використання документів MS Word як стегаконтейнерів є можливість безпосереднього втручання в XML-структуру файлів, що становлять документ. Кожен XML-файл представляє собою комплекс ієрархічно впорядкованих елементів, які визначають специфічні характеристики

документа. XML-теги мають численні атрибути, котрі визначають параметри форматування та різноманітні властивості елементів. Зміна значень певних несуттєвих атрибутів дозволяє впроваджувати приховані повідомлення за допомогою різноманітних алгоритмів стеганографічного кодування інформації. При цьому, імплементація стеганографічних методів на рівні модифікації XML-структури характеризується високою стійкістю до стеганоаналізу, оскільки передбачає втручання в низькорівневі параметри форматування документа, котрі зазвичай не відображаються у візуальному представленні та не підлягають типовим процедурам редагування.

Висновки до розділу 2

У другому розділі здійснено комплексний аналіз теоретичних та прикладних аспектів стеганографічного приховування інформації, що дозволило сформувати цілісне уявлення про сучасний стан, методологічні засади й перспективи розвитку відповідного напрямку інформаційної безпеки. Розглянута термінологічна база забезпечила формування уніфікованого поняттєво-категоріального апарату, необхідного для подальшого структурованого аналізу.

На основі узагальненої моделі стегосистеми встановлено, що стеганографічна система являє собою багатокомпонентну функціональну структуру, в межах якої здійснюється процес імплантації прихованого повідомлення у вибраний контейнер за допомогою алгоритмічно визначених процедур, що базуються на використанні стегоключів, інформаційних перетворень та форматно-залежних особливостей носія.

Проведена класифікація стеганографічних методів дозволила виокремити основні напрямки впровадження прихованих даних, які відрізняються не лише за типом контейнера (текстовий, графічний, мультимедійний), а й за використовуваними математичними та алгоритмічними принципами. Зокрема,

текстова стеганографія реалізується через модифікацію структурно-семантичних властивостей документа.

Особливу увагу приділено методу найменш значущих бітів (LSB), який належить до базових і водночас найпоширеніших підходів приховання інформації у візуальних контейнерах, та дозволяє імплантувати дані зі збереженням візуальної непомітності. Крім того, розглянуто структуру та функціональні властивості файлів формату Microsoft Word, які завдяки архітектурі, заснованій на XML-представленні, можуть бути ефективно використані як носії прихованої інформації.

Сформоване у цьому розділі теоретичне підґрунтя, систематизація методів та ідентифікація їхньої прикладної значущості є основою для подальшого розроблення та практичного застосування стеганографічної технології, орієнтованої на захист інформації шляхом приховування в текстових документах MS Word.

РОЗДІЛ 3
РОЗРОБКА ПРОГРАМНОГО РІШЕННЯ ДЛЯ РЕАЛІЗАЦІЇ ЗАПРОПОНОВАНОГО МЕТОДУ

Під час виконання практичної частини кваліфікаційної роботи було розроблено програмне забезпечення для прихованого впровадження інформації у документи формату MS Word шляхом стеганографічної модифікації форматування символів. Реалізацію алгоритмів кодування та декодування прихованих повідомлень здійснено мовою програмування C# з інтеграцією бібліотек для обробки документів Open XML.

3.1 Запропонований метод

Метод стеганографічного перетворення базується на модифікації колірних характеристик текстових елементів у документах з використанням тривимірного RGB-простору як носія прихованої інформації. У MS Word кожен символ може мати колір, визначений трьома компонентами: R (Red), G (Green), B (Blue). Кожна з цих складових представлена 8-бітним значенням (від 0 до 255) [22].

Особливість запропонованого підходу полягає в використанні найменш значущих бітів колірних каналів символів для приховування секретних даних безпосередньо в структуру текстового документа. Механізм ґрунтується на властивості людського зору сприймати мікроскопічні колірні варіації як практично ідентичні, що забезпечує високий рівень конспіративності при передачі інформації [26].

Технічна реалізація методу передбачає трансформацію секретного повідомлення в послідовність бітів, які розподіляються між трьома колірними компонентами: червоний (R), зелений (G) та синій (B). Математична модель кодування визначається системою нелінійних перетворень, де вхідний байт повідомлення піддається складним бітовим операціям маскування та зсуву.

Загальна схема приховування повідомлень у RGB каналі символу документа представлена на рис. 3.1.

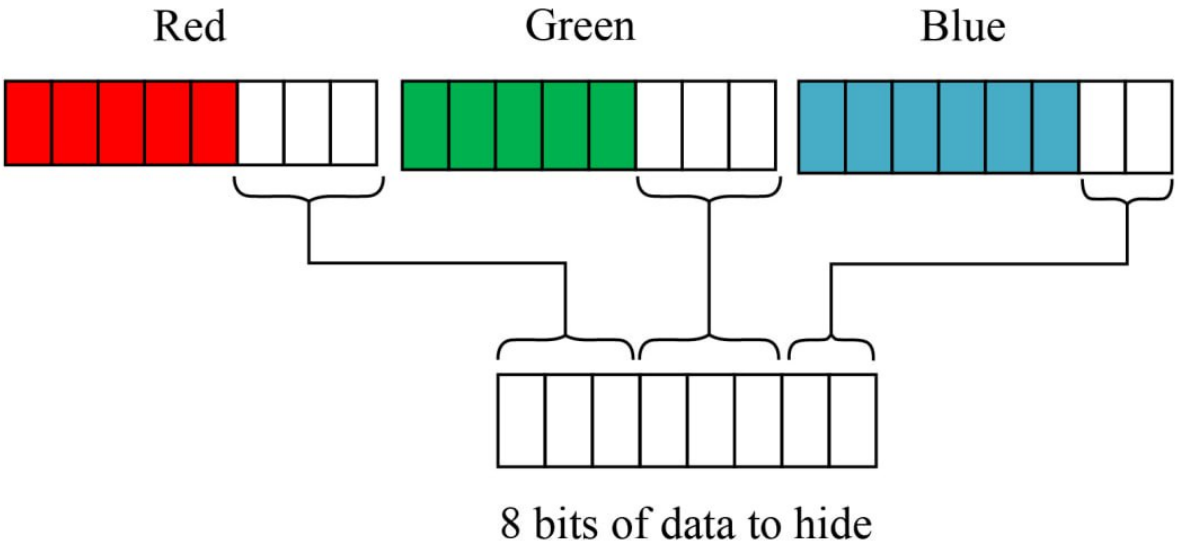


Рисунок 3.1 – Загальна схема приховування повідомлень у RGB каналі символу

Приховання інформації шляхом модифікації кольорів символів з використанням метода LSB має низку суттєвих обмежень та потенційних вразливостей, які можуть знизити ефективність стеганографічного захисту в середовищі текстових документів MS Word. Це:

1. Висока чутливість до перетворень документа. Будь-які зміни в структурі або форматуванні документа можуть повністю зруйнувати приховану інформацію. Операції редагування тексту, зміна шрифтів, конвертація між форматами або навіть автоматичне коректування можуть призвести до повної втрати вбудованих даних.

2. Простота виявлення спеціалізованими засобами. Сучасні стеганографічні аналізатори можуть легко детектувати незначні зміни в колірних каналах символів. Використання складних алгоритмів комп'ютерного зору та статистичного аналізу дозволяє ідентифікувати приховану інформацію з досить високою ймовірністю [27].

3.2 Середовища розробки Microsoft Visual Studio

Visual Studio – це потужний інструмент розробника, який можна використовувати для завершення всього циклу розробки в одному місці. Це комплексне інтегроване середовище розробки (IDE), яке можна використовувати для написання, редагування, налагодження та створення коду [28].

Microsoft Visual Studio є провідним інтегрованим середовищем розробки, що забезпечує комплексне рішення для створення програмних додатків різного рівня складності та призначення. Платформа розроблена корпорацією Microsoft і представляє собою інструментарій для професійної розробки програмного забезпечення, що підтримує широкий спектр мов програмування, технологій та платформ розгортання.

Середовище розробки підтримує мультиплатформний підхід, дозволяючи створювати додатки для операційних систем Windows, macOS, Linux, а також мобільних платформ iOS та Android. Інтеграція з хмарними сервісами Microsoft Azure забезпечує можливості для розробки та розгортання масштабованих веб-додатків, мікросервісів та корпоративних рішень. Visual Studio включає розвинені засоби для роботи з базами даних, системами контролю версій, інструментами безперервної інтеграції та автоматизованого тестування.

Архітектура Visual Studio базується на модульній структурі з розширюваною системою компонентів, що дозволяє адаптувати середовище під специфічні потреби проектів та команд розробників. Інтелектуальна система IntelliSense забезпечує розширені можливості автодоповнення коду, аналізу синтаксису в реальному часі та контекстуальної допомоги при написанні програм. Вбудований компілятор Roslyn для мов C# та Visual Basic.NET надає передові можливості статичного аналізу коду та оптимізації продуктивності.

Visual Studio забезпечує повноцінну підтримку екосистеми Microsoft .NET . Середовище включає шаблони проектів для різних типів додатків, включаючи консольні програми, Windows Forms додатки, WPF застосунки, веб-додатки ASP.NET, веб-API сервіси та бібліотеки класів. Інтеграція з менеджером пакетів

NuGet спрощує управління залежностями та дозволяє легко інтегрувати сторонні бібліотеки в проекти.

3.3 Мова програмування C# для розробки додатку

Для реалізації стеганографічного методу обрано мову програмування C#.

Мова C# – це багатопарадигмова об'єктноорієнтована та компонентно-орієнтована мова програмування зі строгою типизацією, розроблена для платформи .NET Framework [29]. Використання мови визначене стандартами ECMA-334 [30] та ISO/IEC 23270:2006 [31]. Розроблена командою Microsoft Research під керівництвом Андерса Гейлсберга.

Дана мова програмування є об'єктно-орієнтованою мовою високого рівня, розробленою компанією Microsoft у рамках ініціативи .NET Framework. C# поєднує потужність та гнучкість мов програмування C++ з простотою використання мови Visual Basic, забезпечуючи оптимальний баланс між продуктивністю розробки та ефективністю виконання програм. Мова характеризується строгою типізацією, автоматичним управлінням пам'яттю через збирач сміття та розширеними можливостями об'єктно-орієнтованого програмування.

C# надає природну інтеграцію з Microsoft Office Object Model та OpenXML SDK, що критично важливо для роботи з документами Word. Мова включає розвинені засоби для роботи з файловими системами, обробки винятків та багатопоточного програмування. Система типів C# підтримує узагальнення (generics), делегати, події та LINQ-запити, що дозволяє створювати елегантний та підтримуваний код для складних алгоритмів стеганографічної обробки. Компіляція в проміжний код MSIL забезпечує високу продуктивність виконання та можливість оптимізації під час виконання.

3.4 Windows Forms як платформа користувацького інтерфейсу

Для створення графічного інтерфейсу стеганографічного додатку використовується технологія Windows Forms, яка представляє класичний підхід до розробки настільних додатків для операційної системи Windows. Windows Forms забезпечує багатий набір візуальних компонентів та засобів для створення інтуїтивно зрозумілого інтерфейсу користувача з використанням традиційної подійно-орієнтованої моделі програмування. Технологія базується на обгортці навколо Win32 API, що забезпечує нативну продуктивність та повну інтеграцію з операційною системою Windows.

Архітектура Windows Forms включає систему форм як контейнерів для елементів управління, розвинену ієрархію класів для різних типів компонентів інтерфейсу та гнучку систему обробки подій. Технологія підтримує візуальне проектування інтерфейсу через дизайнер Visual Studio, що значно спрощує процес створення складних діалогових вікон та основних форм програми. Система прив'язки даних дозволяє автоматично синхронізувати елементи інтерфейсу з об'єктами предметної області, що особливо корисно для відображення параметрів стеганографічного кодування.

Windows Forms надає широкий спектр стандартних елементів управління, включаючи текстові поля, кнопки, списки, меню, панелі інструментів та діалогові вікна для роботи з файлами.

3.5 Програмна реалізація запропонованого методу приховування даних у документах MS Word

Розроблене програмне рішення призначено для практичної імплементації стеганографічного методу, що базується на модифікації колірних характеристик символів у документах формату Microsoft Word (.docx). Архітектура програми реалізує комплексний підхід до приховування цифрової інформації шляхом

застосування методу LSB (Least Significant Bit) стеганографії з використанням RGB-колірної моделі як носія стеганограми.

Програмна реалізація структурована у вигляді об'єктно-орієнтованої системи, центральним компонентом якої виступає клас Steganography, що інкапсулює функціональність кодування та декодування стеганографічної інформації (Додаток Б). Система використовує бібліотеку DocumentFormat.OpenXml для маніпулювання структурними елементами документів Office Open XML формату, забезпечуючи програмний доступ до XML-метаданих документа та безпосередню модифікацію атрибутів форматування тексту.

Алгоритм приховування інформації реалізовано через метод Hide(), який здійснює покрокове перетворення вхідного файлу-контейнера. Процес характеризується наступною послідовністю операцій: ініціалізація генератора байтової послідовності через функцію GetContent(), що формує структуровану послідовність даних, включаючи магічне число ідентифікації, прапорці розмітки структури (formatFlag, sizeFlag, contentFlag), метадані файлу та безпосередньо корисне навантаження.

Кодування здійснюється шляхом бітового розподілу кожного байта інформації відповідно до заданого пароля-ключа [3, 3, 2], що визначає розподіл бітів між компонентами RGB-моделі (рис. 3.2). Біти розподіляються у такий спосіб: старші 3 біти кодуються у червоній компоненті, наступні 3 біти - у зеленій, молодші 2 біти - у синій компоненті кольору. Математичне представлення процесу екстракції компонент виражається через бітові операції зсуву та маскування.

```
var colorRed = (b >> (password[1] + password[2])) & ((int)Math.Pow(2, password[0]) - 1);
var colorGreen = (b >> password[2]) & ((int)Math.Pow(2, password[1]) - 1);
var colorBlue = b & ((int)Math.Pow(2, password[2]) - 1);
```

Рисунок 3.2 – Розподіл бітів у програмному застосунку

Декодування реалізовано через метод Decode(), що виконує зворотну трансформацію колірних атрибутів у вихідну байтову послідовність. Процес передбачає парсинг усіх елементів типу Run документа, екстракцію значень кольорів та реконструкцію початкових байтів через бітові операції об'єднання, що зображено на рис. 3.3.

```
var color = colorElement.Val!.Value!;
var colorRed = int.Parse(color.Substring(0, 2), NumberStyles.HexNumber);
var colorGreen = int.Parse(color.Substring(2, 2), NumberStyles.HexNumber);
var colorBlue = int.Parse(color.Substring(4, 2), NumberStyles.HexNumber);

var b = (colorRed << (password[1] + password[2])) | (colorGreen << password[2]) | colorBlue;
```

Рисунок 3.3 – Відтворення початкових байтів для декодування

Графічний інтерфейс користувача, реалізований у класі Form1, забезпечує інтуїтивну взаємодію з функціональністю системи через Windows Forms API. Інтерфейс структуровано навколо чотирьох основних операцій: вибір файлу-контейнера, визначення цільового файлу результату, селекція контенту для приховування та ініціація процесів кодування/декодування.

Система забезпечує валідацію цілісності даних через верифікацію магічного числа при декодуванні, що гарантує коректність процесу екстракції та виявлення потенційних пошкоджень стеганограми. Обробка винятків реалізована для запобігання критичним збоям системи при роботі з некоректними вхідними даними або недостатнім обсягом носія для приховування заданого контенту.

Алгоритм роботи програми можна описати таким чином (Додаток А):

- Інтерактивна селекція файлових ресурсів: користувацька взаємодія з системою передбачає поетапний вибір необхідних файлових компонентів через стандартизовані діалогові вікна Windows Forms, що зображено на рис. 3.4.

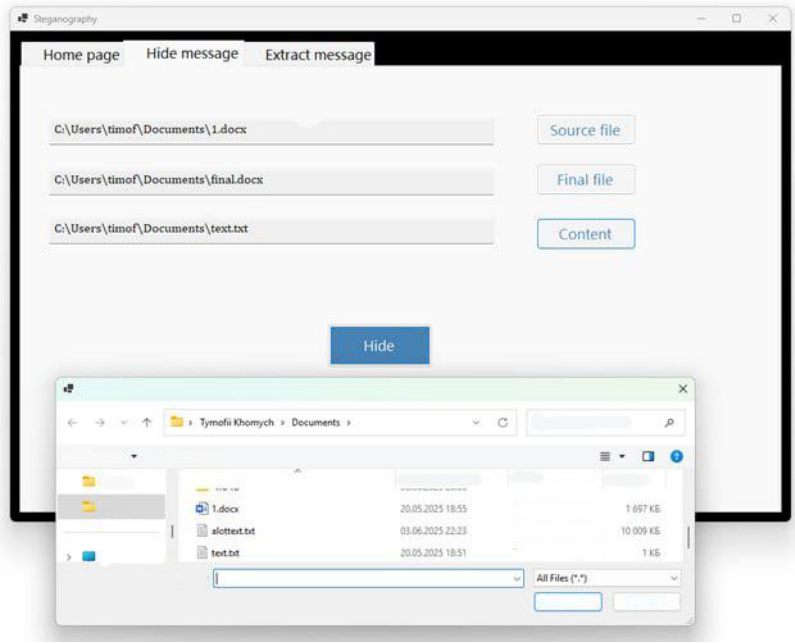


Рисунок 3.4 – Діалогове вікно з вибором файлових ресурсів

Спершу пропонується вибір документа-контейнера. Натискання на відовідну кнопку ініціює створення екземпляра OpenFileDialog з фільтром «Word Documents (.docx)|.docx», що обмежує селекцію виключно документами Microsoft Word формату Office Open XML. Початкову сторінку файла-контейнера, який є звичайним документом MS Word, зображено на рис. 3.5.

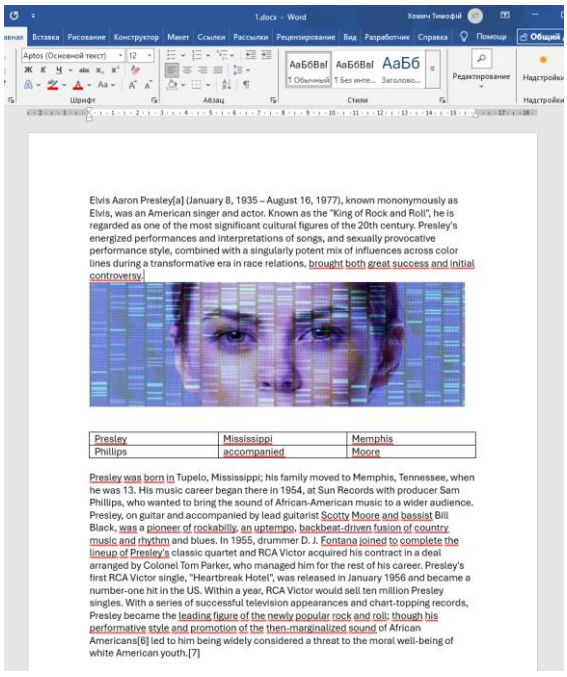


Рисунок. 3.5 – Документ Microsoft Word як стегаконтейнер

Далі визначається цільовий файл результату. Обробник подіє активує діалог зі збереження файлу з ідентичним фільтром, дозволяючи користувачу специфікувати локацію та ім'я результуючого документа, що міститиме приховану інформацію. Останнім полем є селекція контенту для стеганографічного приховування. Натискання на кнопку ініціює діалог OpenFileDialog з універсальним фільтром «All Files (.)|.», забезпечуючи можливість вибору файлів довільного формату та розширення. Файл з повідомленням зображено на рис. 3.6.

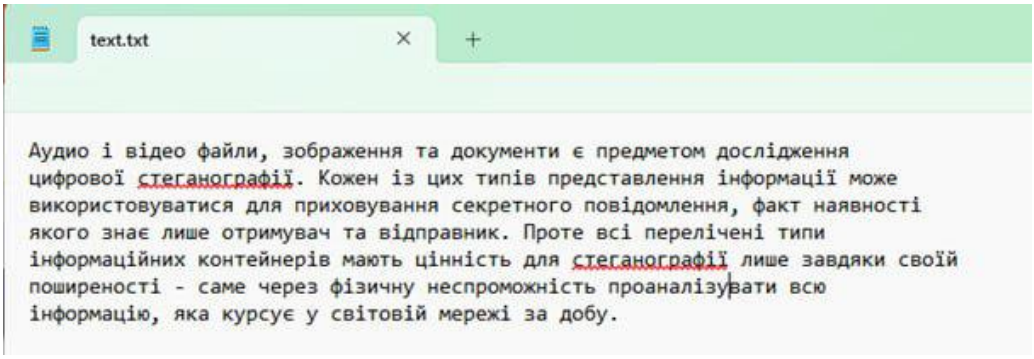


Рисунок 3.6 – Повідомлення, яке буде приховано

- Алгоритм приховування інформації: процедура кодування активується через обробник (рис. 3.7) при натискання кнопки Hide, що виконує первинну валідацію заповненості всіх текстових полей з подальшим викликом методу steganography.Hide().

```
private void HideButton_Click(object sender, EventArgs e)
{
    var sourceFilePath = SourceFileTextBox.Text;
    var finalFilePath = FinalFileTextBox.Text;
    var contentFilePath = ContentFileTextBox.Text;

    if (string.IsNullOrWhiteSpace(sourceFilePath) ||
        string.IsNullOrWhiteSpace(finalFilePath) ||
        string.IsNullOrWhiteSpace(contentFilePath))
    {
        MessageBox.Show("Please provide all required file paths.", "Warning", MessageBoxButtons.OK, MessageBoxIcon.Warning);
        return;
    }

    try
    {
        steganography.Hide(sourceFilePath, finalFilePath, contentFilePath);
        MessageBox.Show("Content successfully hidden in the file.", "Success", MessageBoxButtons.OK, MessageBoxIcon.Information);
    }
    catch (Exception ex)
    {
        MessageBox.Show($"Error hiding content: {ex.Message}", "Error", MessageBoxButtons.OK, MessageBoxIcon.Error);
    }
}
```

Рисунок 3.7 – Обробник HideButton_Click

Далі виконуються підготовчі операції та валідація параметрів. Система здійснює перевірку існування цільового файлу з його подальшим видаленням, після чого створює точну копію вихідного документа-контейнера. Наступним кроком є ітеративна обробка структурних елементів документа. Система послідовно обходить усі елементи тіла документа. Основним етапом є посимвольне кодування на рівні текстових абзаців - для кожного символу в межах Run елементів виконується бітовий розподіл байта відповідно до алгоритму зображеного на рисунку 3.2. Фінальним кроком даного етапу є генерація модифікованих елементів форматування, де кожен символ перетворюється на індивідуальний Run з унікальними колірними атрибутами у шістнадцятковому RGB-представленні. Це продемонстровано на рис. 3.8.

```
var newRun = new Run(text);
newRun.RunProperties =
    new RunProperties(new Color { Val = $"{colorRed:X2}{colorGreen:X2}{colorBlue:X2}" });
newParagraph.AppendChild(newRun);
```

Рисунок 3.8 – Генерація модифікованих елементів форматування

- Процедура збереження та фіналізації. Після завершення процесу система очищує вихідну структуру документа та замінює її модифікованими елементами, зберігаючи результат через wordDoc.MainDocumentPart.Document.Save(). У разі неможливості розміщення всього контенту генерується виняток переповнення, а при успішному завершенні користувач отримує повідомлення, як на рис. 3.9, про успішне приховування інформації.

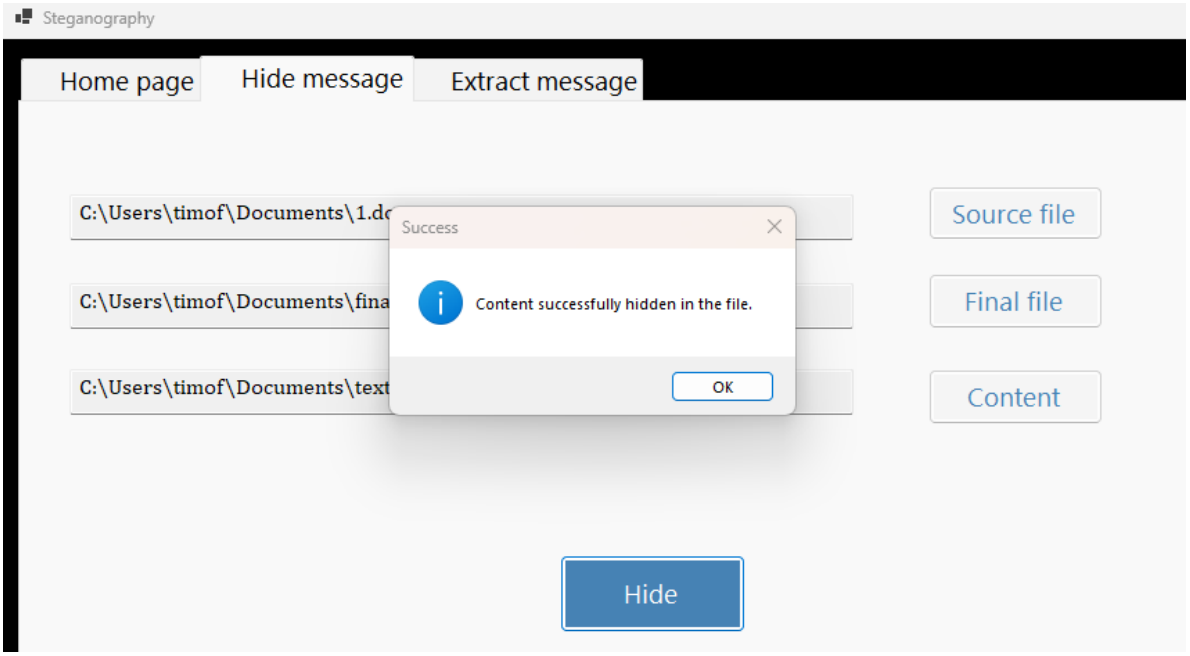


Рисунок 3.9 – Успішне приховування повідомлення

Якщо повідомлення завелике і його неможливо приховати в файл-контейнер, на екрані відобразиться повідомлення, як на рис. 3.10.

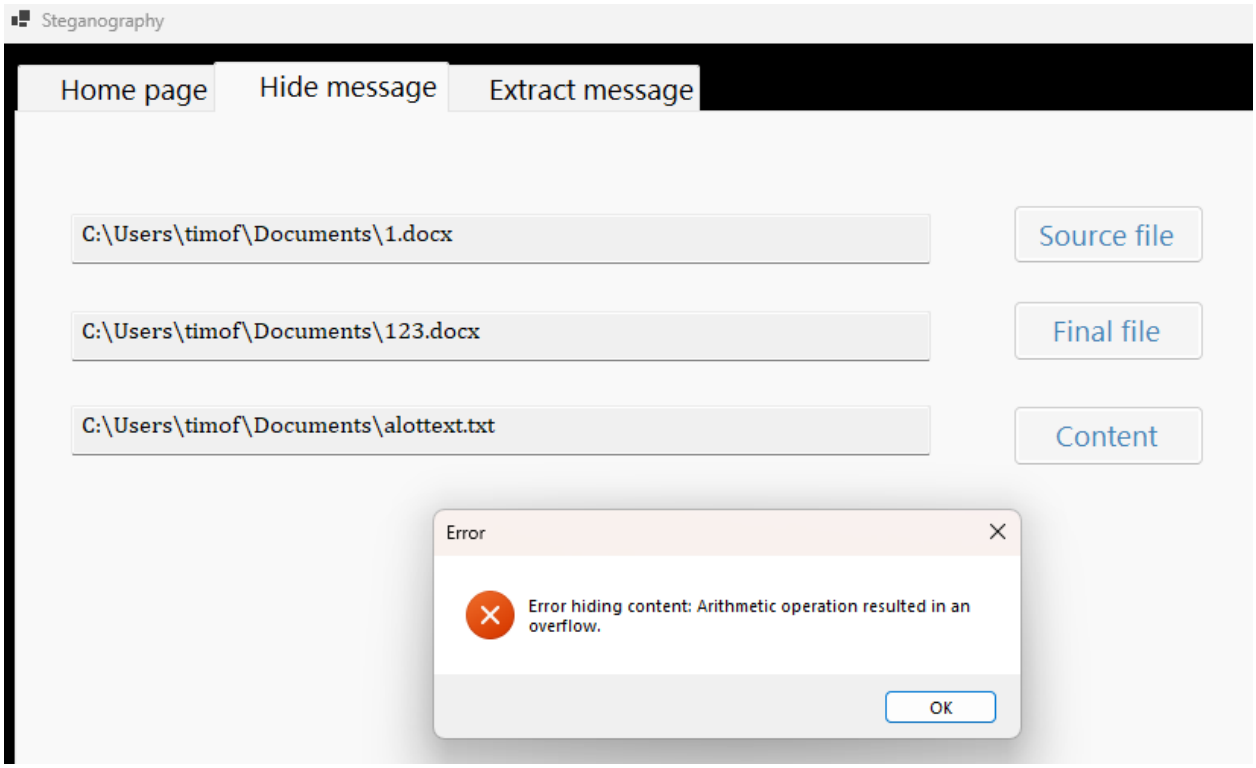


Рисунок 3.10 – Повідомлення про помилку

- Алгоритм екстракції прихованої інформації. Процедура декодування ініціюється натискання на кнопку Extract (рис. 3.11) з подвійною діалоговою взаємодією для вибору джерельного файлу та визначення локації збереження.

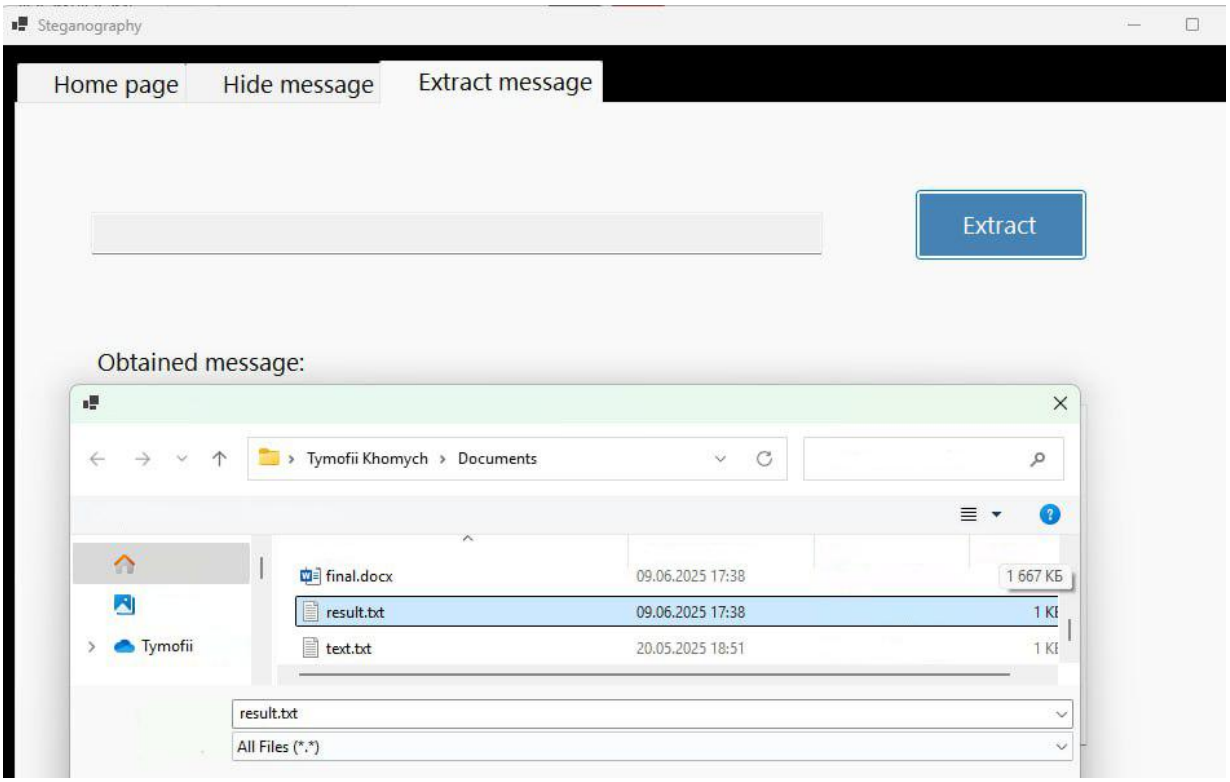


Рисунок 3.11 – Діалогове вікно для вилучення прихованого повідомлення

Система отримує доступ до структури документа у режимі читання, здійснюючи послідовну ітерацію по всіх параграфах та текстових прогонах з колірними атрибутами. Для кожного Run з визначеними колірними властивостями виконується декомпозиція RGB-значення та зворотна бітова трансформація для відновлення початкового байта. Екстрагований контент записується у вказаний користувачем файл через File.WriteAllBytes(), після чого система надає підтвердження успішного завершення операції (рис. 3.12).

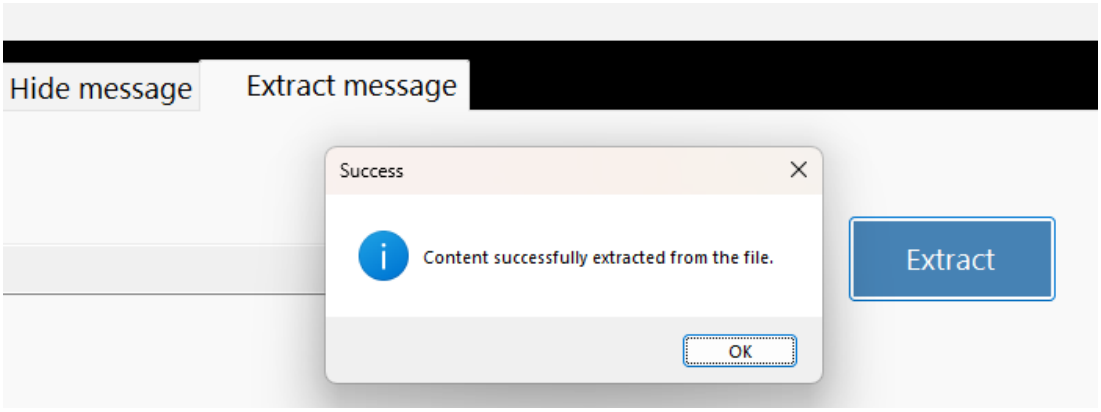


Рисунок. 3.12 – Повідомлення про успішне виконання процесу вилучення

Вміст відтвореного повідомлення можна побачити у файлі, який буде знаходитись за обраним шляхом. Результат виконання, який повністю ідентичний повідомленню з рис. 3.6, зображено на рис. 3.13.

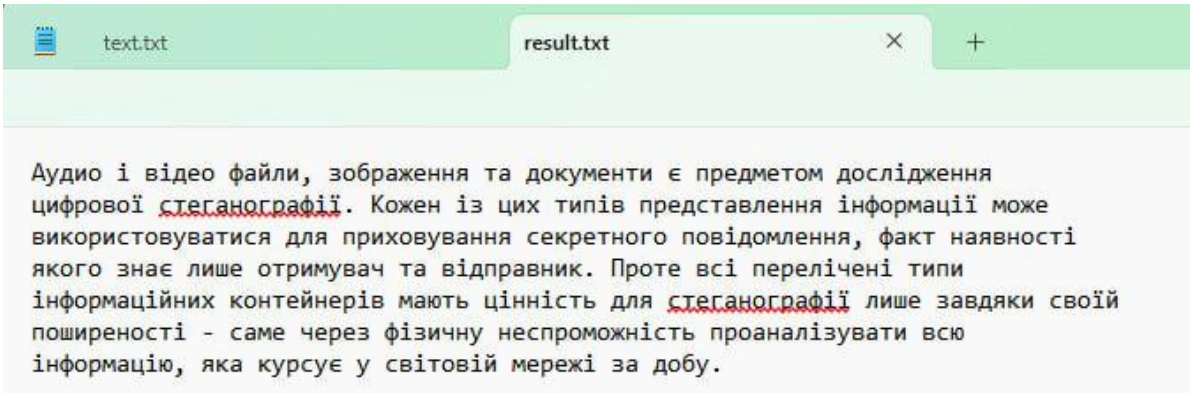


Рисунок 3.13 – Вміст вилученого файлу

Висновок до розділу 3

Розроблена програмна реалізація стеганографічного методу приховування інформації у документах Microsoft Word демонструє ефективне поєднання теоретичних принципів цифрової стеганографії з практичними аспектами програмної інженерії. Архітектура системи, побудована на основі об'єктно-орієнтованої парадигми з використанням бібліотеки DocumentFormat.OpenXml,

забезпечує надійний програмний доступ до внутрішньої структури документів формату Office Open XML.

Імплементований алгоритм кодування, що базується на модифікованому методі найменш значущих бітів у колірному просторі RGB, характеризується оптимальним співвідношенням між ємністю стеганографічного каналу та візуальною непомітністю модифікацій. Використання асиметричного розподілу бітів за схемою [3:3:2] для RGB-компонент дозволяє досягти максимальної ефективності кодування при збереженні прийнятного рівня колірної дисторсії.

Графічний інтерфейс користувача, реалізований засобами Windows Forms API, забезпечує інтуїтивну взаємодію з функціональністю системи через стандартизовані діалогові компоненти операційної системи. Комплексна система обробки винятків гарантує стабільність функціонування програмного продукту при роботі з некоректними вхідними даними або в умовах обмежених системних ресурсів.

ВИСНОВКИ

У ході виконання кваліфікаційної роботи було здійснено комплексне дослідження концептуальних, теоретичних і практичних аспектів стеганографічного приховування інформації в контексті використання текстових документів як стегаконтейнерів, зокрема зосереджено увагу на експлуатації структурно-функціональних характеристик форматованих документів MS Word.

У першому розділі було викладено концептуальні засади стеганографії як галузі знань, що досліджує приховані методи передавання даних у середовищах з відкритим доступом до інформаційних каналів. Проаналізовано історичну еволюцію основних підходів до приховування повідомлень. Розглянуто сучасні області застосування стеганографії. Окреслено основні методи стеганографії, класифіковані за типом носія, із виокремленням особливостей їх реалізації в умовах цифрового середовища. Зазначена структура першого розділу створює теоретичне підґрунтя для подальшого формалізованого опису методів і засобів приховування даних, що стали предметом дослідження в наступних розділах.

У другому розділі було систематизовано основну термінологію та викладено загальну модель функціонування стегосистем, яка охоплює етапи вбудовування, передачі та вилучення прихованого повідомлення. Наведено класифікацію сучасних стеганографічних методів із виділенням окремих підходів, серед яких особливу увагу зосереджено на текстовій стеганографії. Проведено порівняльний аналіз методів, заснованих на модифікації лінгвістичної структури, синтаксису, семантики та форматних характеристик, а також розглянуто ключові проблеми та обмеження, притаманні цій групі методів. Окремий акцент зроблено на документах MS Word, внутрішня структура яких у вигляді XML-архіву забезпечує технологічну придатність до використання як повноцінних стегаконтейнерів.

У третьому розділі було розроблено прикладне програмне рішення, реалізоване засобами мови C# у середовищі Microsoft Visual Studio з

використанням архітектури Windows Forms. Створено функціональний прототип, здатний виконувати приховування та вилучення інформації за рахунок модифікації параметрів форматування кольору окремих символів у тексті. Завдяки використанню OpenXML SDK забезпечено доступ до структурних елементів документа без необхідності його відкриття у середовищі текстового редактора. Запропонований підхід продемонстрував стабільну працездатність, збереження зовнішньої цілісності документів та здатність до кодування повідомлень із прийнятною інформаційною ємністю, що засвідчує його ефективність та практичну доцільність.

Виходячи із поставленої мети кваліфікаційної роботи були виконані наступні завдання:

- Проаналізувано наявні стеганографічні методи захисту інформації в текстових документах та визначено їх основні переваги та недоліки.
- Досліджено особливості форматування символів у документах MS Word та виявити параметри, які можуть бути використані для приховування інформації.
- Розроблено стеганографічний алгоритм приховування даних у текстових документах MS Word шляхом маніпулювання параметрами форматування символів.
- Створено програмну реалізацію розробленого стеганографічного алгоритму.

Апробація роботи. Основні результати роботи доповідались на VIII Міжнародній науково-практичній конференції «Проблеми кібербезпеки інформаційно-телекомунікаційних систем» (PCSICS)» 11 квітня 2025 року.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Арделян І. С. Використання стеганографії в сучасних кібератаках // Актуальні питання забезпечення кібербезпеки та захисту інформації : матеріали IX Міжнар. наук.-практ. конф., м. Київ, 30 берез. 2023 р.

2. Majeed M. A., Sulaiman R. An improved LSB image steganography technique using BIT-inverse in 24 BIT colour image // Journal of Theoretical and Applied Information Technology. – 2015. – Vol. 80, No. 2.

3. Elbeji R. Friendly Introduction To Steganography [Електронний ресурс]. – Режим доступу: https://medium.com/@rabi3elbeji/friendly-introduction-to-steganography-4cf032240904 (дата звернення: 13.01.2025).

4. Guinness World Records. First use of steganography [Електронний ресурс]. – Режим доступу: https://www.guinnessworldrecords.com/world-records/first-use-of-steganography (дата звернення: 16.01.2025).

5. Digital micrography / R. Maharik, M. Ben-Ezra, Y. Wexler, M. Werman // ACM Transactions on Graphics. – 2011. – Vol. 30, No. 4. – P. 1–12.

6. Хорошко В. О., Яремчук Ю. Є., Карпінець В. В. Комп’ютерна стеганографія. – Вінниця : ВНТУ, 2017. – 149 с.

7. Tancik M., Mildenhall B. StegaStamp: Invisible Hyperlinks in Physical Photographs // arXiv. – 2020.

8. Intellect icu. Steganography, classification of types of steganography [Електронний ресурс]. – Режим доступу: https://intellect.bond/steganography-classification-of-types-of-steganography-5824 (дата звернення: 24.01.2025).

9. Денисюк В. О. Стеганографічний алгоритм захисту даних з використанням файлів зображень // Ефективна економіка. – 2017.

10. Cheddad A., Condell J., Curran K., Mc Kevitt P. Digital image steganography: Survey and analysis of current methods // Signal Processing. – 2010. – Vol. 90. – P. 727–752.

11. Johnson N. F., Jajodia S. Exploring steganography: Seeing the unseen // Computer. – 1998. – Vol. 31. – P. 26–34.

12. Digital Watermarking and Steganography / J. Bloom, I. Cox, M. Miller, J. Fridrich. – Elsevier Science & Technology Books, 2007.

13. Provos N., Honeyman P. Hide and seek: an introduction to steganography // IEEE Security & Privacy. – 2003. – Vol. 1, No. 3. – P. 32–44. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1109/msecp.2003.1203220 (дата звернення: 02.02.2025).

14. A Review on Text Steganography Techniques / M. A. Majeed et al. // Mathematics. – 2021. – Vol. 9, No. 21. – P. 2829. [Електронний ресурс]. – Режим доступу: https://doi.org/10.3390/math9212829 (дата звернення: 05.02.2025).

15. Information Hiding Techniques for Steganography and Digital Watermarking / ред. S. K. Katsikas, F. A. Petitcolas. – Artech House Publishers, 2000. – 220 p.

16. Petitcolas F. A. P., Anderson R. J., Kuhn M. G. Information hiding – a survey // Proceedings of the IEEE. – 1999. – Vol. 87, No. 7. – P. 1062–1078. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1109/5.771065 (дата звернення: 13.02.2025).

17. Khan M., Shahab A., Asghar Z. Introduction to Linguistic Steganography // Nonlinear Engineering. – 2015. – Vol. 4, No. 3. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1515/nleng-2015-0013 (дата звернення: 16.02.2025).

18. Hamzah A. A., Bayomi H. Text Steganography with High Embedding Capacity Using Arabic Calligraphy // Advances in Intelligent Systems and Computing. – Cham, 2019. – P. 127–138. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1007/978-3-030-33582-3_13 (дата звернення: 25.02.2025).

19. Linguistic Steganography Detection Using Statistical Characteristics of Correlations between Words / Z. Chen et al. // Information Hiding. – Berlin, Heidelberg, 2008. – P. 224–235. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1007/978-3-540-88961-8_16 (дата звернення: 01.03.2025).

20. Maji G., Mandal S. A forward email based high capacity text steganography technique using a randomized and indexed word dictionary // Multimedia Tools and Applications. – 2020. – Vol. 79, No. 35–36. – P. 26549–26569. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1007/s11042-020-09329-z (дата звернення: 07.03.2025).

21. Dulera S., Jinwala D., Dasgupta A. Experimenting with the Novel Approaches in Text Steganography // International Journal of Network Security & Its Applications. – 2011. – Vol. 3, No. 6. – P. 213–225. [Електронний ресурс]. – Режим доступу: https://doi.org/10.5121/ijnsa.2011.3616 (дата звернення: 13.03.2025).

22. Хомич Т., Бабенко Ю. Захист інформації з використанням стеганографічних методів при форматуванні символів в документах MS Word // Проблеми кібербезпеки інформаційно-комунікаційних систем : матеріали VIII Міжнар. наук.-практ. конф., м. Київ, 2025. – С. 133–134.

23. Introducing the Office (2007) Open XML File Formats // Microsoft Learn [Електронний ресурс]. – Режим доступу: https://learn.microsoft.com/en-us/previous-versions/office/developer/office-2007/aa338205(v=office.12) (дата звернення: 18.03.2025).

24. ECMA-376 – Ecma International [Електронний ресурс]. – Режим доступу: https://ecma-international.org/publications-and-standards/standards/ecma-376/ (дата звернення: 20.03.2025).

25. Morkel T., Eloff J. H. P., Olivier M. S. An overview of image steganography // Proceedings of the ISSA 2005 Conference. – Sandton, South Africa, 29 June – 1 July 2005.

26. Prasad S., Pal A. K. An RGB colour image steganography scheme using overlapping block-based pixel-value differencing // Royal Society Open Science. – 2017. – Vol. 4, No. 4. – P. 161066. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1098/rsos.161066 (дата звернення: 28.03.2025).

27. Avcibas I., Memon N., Sankur B. Steganalysis using image quality metrics // IEEE Transactions on Image Processing. – 2003. – Vol. 12, No. 2. –

P. 221–229. [Електронний ресурс]. – Режим доступу: https://doi.org/10.1109/tip.2002.807363 (дата звернення: 01.04.2025).

28. What is the Visual Studio IDE? // Microsoft Learn [Електронний ресурс]. – Режим доступу: https://learn.microsoft.com/uk-ua/visualstudio/get-started/visual-studio-ide?view=vs-2022 (дата звернення: 03.04.2025).

29. Коноваленко І. В., Марущак П. О. Платформа .NET та мова програмування C# 8.0. – Тернопіль, 2020. – 322 с.

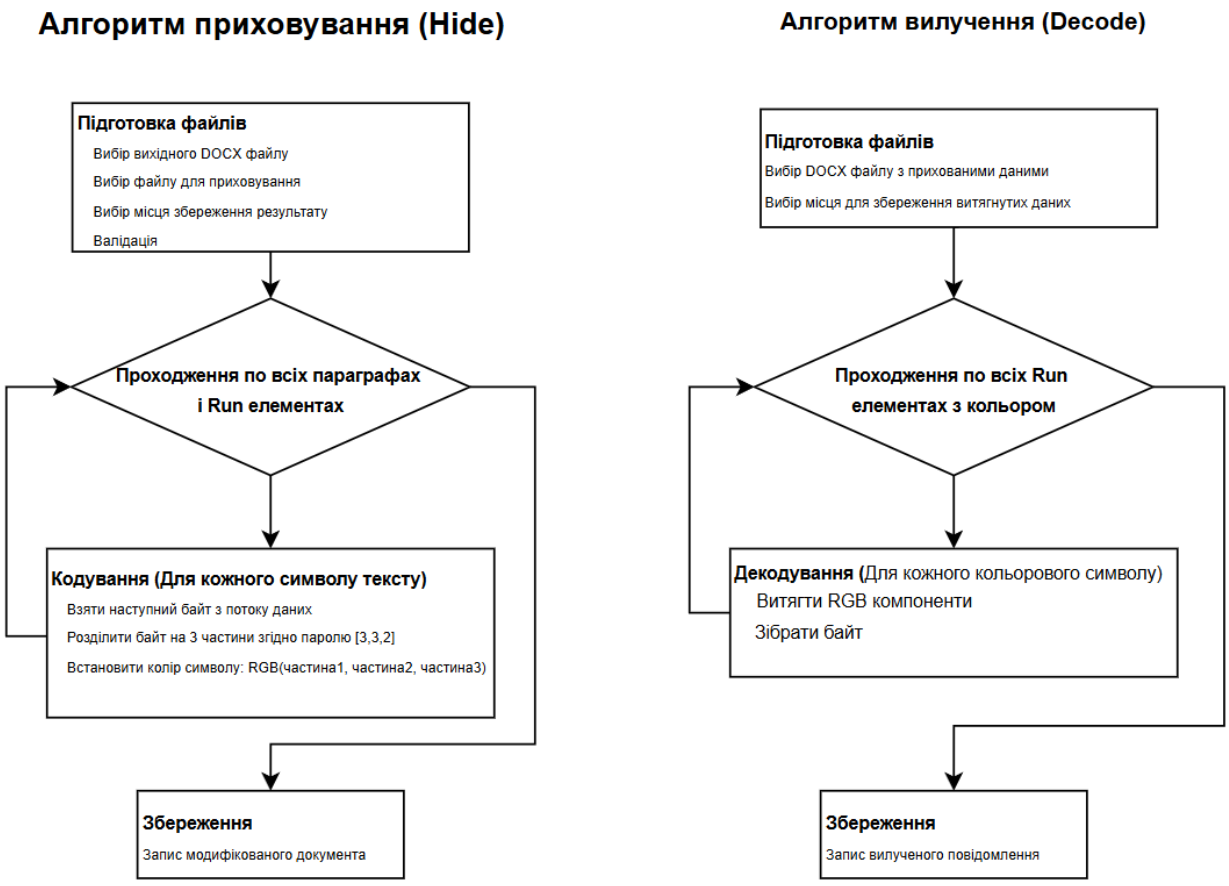
30. Standard ECMA-334. C# Language Specification. – 4th ed. – Geneva : Ecma International, 2006.

31. International standard ISO/IEC 23270:2006. Information technology – Programming languages – C#. – 2nd ed. – Geneva : ISO/IEC, 2006.

ДОДАТКИ

Додаток А

Формальне відображення роботи програмного рішення



Додаток Б

Лістинг коду програмного рішення

```
using System.Text;
using System.Windows.Forms;
using WinFormsApp1.Interfaces;

namespace WinFormsApp1
{
   public partial class Form1 : Form
   {
      private readonly Steganography steganography = new Steganography();
      public Form1()
      {
         InitializeComponent();
      }

      private void SourceFileButton_Click(object sender, EventArgs e)
      {
         using var openFileDialog = new OpenFileDialog { Filter = "Word Documents (*.docx)|*.docx" };
         if (openFileDialog.ShowDialog() == DialogResult.OK)
         {
            SourceFileTextBox.Text = openFileDialog.FileName;
         }
      }

      private void FinalFileButton_Click(object sender, EventArgs e)
      {
         using var saveFileDialog = new SaveFileDialog { Filter = "Word Documents (*.docx)|*.docx" };
         if (saveFileDialog.ShowDialog() == DialogResult.OK)
         {
            FinalFileTextBox.Text = saveFileDialog.FileName;
         }
      }

      private void ContentButton_Click(object sender, EventArgs e)
      {
         using var openFileDialog = new OpenFileDialog() { Filter = "All Files (*.*)|*.*", };
         if (openFileDialog.ShowDialog() == DialogResult.OK){
            ContentFileTextBox.Text = openFileDialog.FileName;
         }
      }

      private void HideButton_Click(object sender, EventArgs e)
      {
         var sourceFilePath = SourceFileTextBox.Text;
         var finalFilePath = FinalFileTextBox.Text;
         var contentFilePath = ContentFileTextBox.Text;

         if (string.IsNullOrWhiteSpace(sourceFilePath) ||
            string.IsNullOrWhiteSpace(finalFilePath) ||
            string.IsNullOrWhiteSpace(contentFilePath))
         {
            MessageBox.Show("Please provide all required file paths.", "Warning", MessageBoxButtons.OK,
MessageBoxIcon.Warning);
            return;
         }

         try
         {
```

Продовження додатку Б

```
            steganography.Hide(sourceFilePath, finalFilePath, contentFilePath);
            MessageBox.Show("Content successfully hidden in the file.", "Success", MessageBoxButtons.OK,
MessageBoxIcon.Information);
        }
        catch (Exception ex)
        {
            MessageBox.Show($"Error hiding content: {ex.Message}", "Error", MessageBoxButtons.OK,
MessageBoxIcon.Error);
        }
    }

    private void ExtractButton_Click(object sender, EventArgs e)
    {
        using var openFileDialog = new OpenFileDialog { Filter = "Word Documents (*.docx)|*.docx" };
        if (openFileDialog.ShowDialog() != DialogResult.OK)
        {
            return;
        }

        var sourceFilePath = openFileDialog.FileName;
        using var saveFileDialog = new SaveFileDialog { Filter = "All Files (*.*)|*.*" };
        if (saveFileDialog.ShowDialog() != DialogResult.OK)
        {
            return;
        }

        var outputFilePath = saveFileDialog.FileName;

        try
        {
            var extractedData = steganography.Decode(sourceFilePath);
            File.WriteAllBytes(outputFilePath, extractedData.ToArray());
            MessageBox.Show("Content successfully extracted from the file.", "Success", MessageBoxButtons.OK,
MessageBoxIcon.Information);
        }
        catch (Exception ex)
        {
            MessageBox.Show($"Error extracting content: {ex.Message}", "Error", MessageBoxButtons.OK,
MessageBoxIcon.Error);
        }
    }
  }
}


using DocumentFormat.OpenXml;
using DocumentFormat.OpenXml.Drawing.Charts;
using DocumentFormat.OpenXml.Packaging;
using DocumentFormat.OpenXml.Wordprocessing;
using System.Globalization;
using System.Text;
using WinFormsApp1.Interfaces;
using Color = DocumentFormat.OpenXml.Wordprocessing.Color;

public class Steganography
{
  int[] password = [3, 3, 2];
  const byte formatFlag = 0xde;
  const byte sizeFlag = 0xed;
  const byte contentFlag = 0xbe;
```

Продовження додатку Б

```
public void Hide(string srcFile, string destFile, string contentPath) {

    if (File.Exists(destFile)) { File.Delete(destFile); }
    File.Copy(srcFile, destFile);
    using var wordDoc = WordprocessingDocument.Open(destFile, true);

    var contentLength = new FileInfo(contentPath).Length;
    if (contentLength > 32 * 1024 * 1024) // 32 MB
    {
        throw new OverflowException();
    }

    var fileFinished = false;
    var enumerable = GetContent();
    using var enumerator = enumerable.GetEnumerator();

    var body = wordDoc.MainDocumentPart!.Document.Body!;
    var newElements = new List<OpenXmlElement>();

    foreach (var element in body.Elements())
    {
        if (fileFinished)
        {
            newElements.Add(element.CloneNode(true));
            continue;
        }

        if (element is not Paragraph paragraph)
        {
            newElements.Add(element.CloneNode(true));
            continue;
        }

        if (paragraph.Descendants<Drawing>().Any())
        {
            newElements.Add(paragraph.CloneNode(true));
            continue;
        }

        var newParagraph = new Paragraph();
        foreach (var pEl in paragraph.Elements())
        {
            if (fileFinished)
            {
                newParagraph.AppendChild(pEl.CloneNode(true));
                continue;
            }

            if (pEl.Descendants<Drawing>().Any())
            {
                newParagraph.AppendChild(pEl.CloneNode(true));
                continue;
            }

            if (pEl is not Run run)
            {
                newParagraph.AppendChild(pEl.CloneNode(true));
                continue;
            }
```

Продовження додатку Б

```
        for (var index = 0; index < run.InnerText.Length; index++)
         {
            if (!enumerator.MoveNext())

            {
               fileFinished = true;
               newParagraph.AppendChild(new Run(new Text(run.InnerText[index..])));
               break;
            }

            var b = enumerator.Current;

            var colorRed = (b >> (password[1] + password[2])) & ((int)Math.Pow(2, password[0]) - 1);
            var colorGreen = (b >> password[2]) & ((int)Math.Pow(2, password[1]) - 1);
            var colorBlue = b & ((int)Math.Pow(2, password[2]) - 1);

            var c = run.InnerText[index];
            var text = c == ' '
               ? new Text(" ")
               {
                  Space = new EnumValue<SpaceProcessingModeValues>(SpaceProcessingModeValues.Preserve)
               }
               : new Text(c.ToString());
            var newRun = new Run(text);
            newRun.RunProperties =
               new RunProperties(new Color { Val = $"{colorRed:X2}{colorGreen:X2}{colorBlue:X2}" });
            newParagraph.AppendChild(newRun);
         }
      }

      newElements.Add(newParagraph);
   }

   // Clear the body and add the new paragraphs
   body.RemoveAllChildren();
   foreach (var element in newElements)
   {
      body.AppendChild(element);
   }

   wordDoc.MainDocumentPart.Document.Save();

   if (!fileFinished)
   {
      throw new OverflowException();
   }

   IEnumerable<byte> GetContent()
   {
      foreach (var b in magicNumber)
      {
         yield return b;
      }

      yield return formatFlag;

      foreach (var b in Path.GetExtension(contentPath).Skip(1).Take(3))
      {
         yield return (byte)b;
      }
```

Продовження додатку Б

```
            yield return sizeFlag;

            yield return (byte)contentLength;
            yield return (byte)(contentLength >> 8);

            yield return (byte)(contentLength >> 16);
            yield return (byte)(contentLength >> 24);
            yield return contentFlag;
            using var fs = new FileStream(contentPath, FileMode.Open);

            var result = 0;
            while ((result = fs.ReadByte()) != -1) yield return (byte)result;
        }
    }

    public List<byte> Decode(string srcFile)
    {
        using var wordDoc = WordprocessingDocument.Open(srcFile, false);
        var body = wordDoc.MainDocumentPart!.Document.Body!;
        var symbols = new List<byte>();

        foreach (var paragraph in body.Elements<Paragraph>())
        {
            foreach (var run in paragraph.Elements<Run>())
            {
                if (run.RunProperties is null)
                {
                    continue;
                }

                var colorElement = run.RunProperties.GetFirstChild<DocumentFormat.OpenXml.Wordprocessing.Color>();
                if (colorElement is null)
                {
                    continue;
                }

                var color = colorElement.Val!.Value!;
                var colorRed = int.Parse(color.Substring(0, 2), NumberStyles.HexNumber);
                var colorGreen = int.Parse(color.Substring(2, 2), NumberStyles.HexNumber);
                var colorBlue = int.Parse(color.Substring(4, 2), NumberStyles.HexNumber);

                var b = (colorRed << (password[1] + password[2])) | (colorGreen << password[2]) | colorBlue;
                symbols.Add((byte)b);
            }
        }
        if (!symbols.Take(magicNumber.Length).SequenceEqual(magicNumber))
        {
            throw new Exception("Magic number mismatch.");
        }
        var contentStartIndex = symbols.IndexOf(contentFlag) + 1;
        return symbols.Skip(contentStartIndex).ToList();
    }
}
```

Додаток В

Апробація результатів дослідження

Хомич Т., Бабенко Ю. Захист інформації з використанням стеганографічних методів при форматуванні символів в документах MS Word // Проблеми кібербезпеки інформаційно-комунікаційних систем : матеріали VIII Міжнар. наук.-практ. конф., м. Київ, 2025. – С. 133–134.