Овчинников Микола Юрійович
асистент кафедри кібербезпеки,
Національний університет «Одеська юридична академія»

ЗАХИСТ ІНФОРМАЦІЇ В МЕДІАФАЙЛАХ

DOI https://doi.org/10.36059/978-966-397-537-5-8

Актуальність теми. У сучасному цифровому середовищі обсяги створення, обміну та зберігання медіафайлів (зображень, відео, аудіо) зростають експоненційно. Мультимедійний контент відіграє ключову роль у функціонуванні соціальних мереж, електронного управління, медіаіндустрії, освіти, реклами та навіть юриспруденції. Водночас медіафайли дедалі частіше стають об'єктами кіберзлочинів, включно з підробкою, модифікацією, перехопленням та несанкціонованим поширенням конфіденційної інформації. Це створює серйозні загрози як для окремих користувачів, так і для організацій, які зберігають або обробляють велику кількість цифрового контенту.

Особливо гостро проблема захисту медіафайлів постає у світлі стрімкого розвитку технологій штучного інтелекту, зокрема генеративних моделей, здатних створювати фальшиві зображення або відео, відомі як deepfake. Такі інструменти вже активно використовуються у дезінформаційних кампаніях, кібершантажі та спробах втручання в політичні процеси. У відповідь на ці виклики постає необхідність впровадження ефективних засобів захисту, здатних забезпечити автентичність, цілісність та конфіденційність мультимедійних даних.

Традиційні підходи до захисту інформації, зокрема криптографічні методи, залишаються актуальними, однак в умовах розвитку стеганографічних атак і змішаних загроз виникає потреба у використанні комбінованих технологій. Зокрема, зростає інтерес до цифрових водяних знаків, адаптивної стеганографії, верифікації метаданих, використання блокчейн-систем для підтвердження автентичності медіа та методів штучного інтелекту для виявлення змін у контенті. Усе це формує нову галузь у сфері інформаційної безпеки – захист інформації в медіафайлах, яка потребує системного дослідження, практичного впровадження та удосконалення з урахуванням динаміки сучасних кіберзагроз.

Захист інформації в медіафайлах формується як самостійний міждисциплінарний напрям, що перебуває на стику інформаційної безпеки, цифрової криміналістики, машинного навчання та авторського права. Цей напрям охоплює не лише технічні аспекти – як-от шифрування, водяні знаки або стеганографію, – а й правові, організаційні та етичні питання, пов'язані з автентичністю, доказовістю та непорушністю цифрового контенту.

Особливість цієї галузі полягає у необхідності глибокого розуміння формату та структури медіафайлів, властивостей цифрових слідів, взаємодії між метаданими та зображенням або аудіоінформацією, а також у вмінні виявляти штучно згенеровані або змінені фрагменти. Водночас, засоби захисту повинні бути малопомітними для користувача, не впливати на якість візуального або аудіовмісту, але бути надійними з точки зору стійкості до атак і здатності підтверджувати справжність навіть у ворожому середовищі.

Інтенсивний розвиток цієї сфери стимулюється практичними потребами: від підтвердження достовірності відеозаписів у судових процесах до захисту журналістських матеріалів, отриманих у зонах конфліктів. Державні органи, ЗМІ, приватні компанії й правозахисні організації все частіше шукають надійні методи ідентифікації автентичного контенту та виявлення маніпуляцій. У зв'язку з цим міжнародні ініціативи на кшталт Content Authenticity Initiative (Adobe, BBC, New York Times) або Coalition for Content Provenance and Authenticity (C2PA) розробляють стандарти, що регулюють доказовість цифрового вмісту.

Таким чином, захист інформації в медіафайлах стає критично важливою складовою кібергігієни та національної безпеки. Формування теоретичних основ і розробка практичних інструментів у цій галузі становить актуальну дослідницьку задачу, що потребує міжгалузевого підходу, високого рівня технічної компетентності та чіткого нормативного врегулювання.

Метою даного дослідження є систематичний аналіз сучасних методів і технологій захисту інформації в медіафайлах з урахуванням динаміки розвитку кіберзагроз, – включно зі стеганографією, цифровим водяним знаком, криптографічними гібридними методами та AI-підходами – з акцентом на практичні інструменти,

здатні протистояти поточним кіберзагрозам, та на розробку рекомендацій для їхнього застосування в локальному контексті. А також обґрунтування практичної ефективності інструментів, що забезпечують автентичність, цілісність та конфіденційність мультимедійного контенту. Дослідження має на меті виявити найбільш релевантні підходи до захисту зображень, відео та аудіо у цифровому середовищі та визначити їхню придатність до використання в різних прикладних сферах – від цифрової журналістики до військової справи та судової експертизи.

У центрі уваги знаходиться як технологічна складова (алгоритми шифрування, цифрові водяні знаки, стеганографічні методи, AI/ML-підходи), так і практичний аспект – зокрема, огляд реальних інструментів, що впроваджуються у відкритих або корпоративних інформаційних системах. Також передбачено аналіз поточних тенденцій щодо інтеграції таких інструментів у процеси цифрової автентифікації, верифікації джерел, захисту авторських прав і протидії інформаційним маніпуляціям.

Окремою задачею є визначення сильних і слабких сторін існуючих рішень, ступеня їх стійкості до актуальних типів атак, а також ступеня відповідності правовим і етичним стандартам. Це дозволяє не лише підвищити рівень теоретичного осмислення проблеми, а й закласти підґрунтя для подальших розробок у цій перспективній галузі кібербезпеки.

Досягнення цілей дослідження передбачає виконання низки взаємопов'язаних завдань. Серед них, визначити основні типи кіберзагроз на медіафайли, що включають як класичні атаки на цілісність і конфіденційність, так і складніші модифікації, зокрема deepfake-маніпуляції з застосуванням штучного інтелекту. Потрібно здійснити огляд і класифікацію сучасних методів захисту: гібридних стеганографічно-шифрувальних систем, зокрема тих, що застосовують DCT/DWT-класифікацію в поєднанні з криптографією (DCT+DWT+OTP), а також інноваційних підходів coverless стеганографії, заснованих на генеративних моделях – зокрема coverless-стегіографія з DiffStega 2024 і об'єктно-орієнтовані формати на основі YOLO. Проаналізувати інструментальні реалізації, що доступні й перспективні для практичного використання, включно

з цифровими водяними знаками, стеганографічними утилітами (OpenStego, Digimarc тощо) та блокчейн-інтегрованими методами. Оцінити ефективність захисних методів за такими критеріями: стійкість до змін (JPEG, переформатування, геометричні трансформації), прихованість кодованої інформації, здатність до верифікації джерела та захисту авторського права, а також відповідність технічним можливостям українських організацій. Сформувати рекомендації щодо впровадження комплексних технологій захисту медіаінформації в умовах атак національного рівня, гібридної агресії, дезінформації та кібервійськ.

Об'єктом дослідження є процеси забезпечення інформаційної безпеки в цифрових медіафайлах.

Предметом дослідження є методи, засоби та інструменти захисту мультимедійного контенту від несанкціонованого доступу, підробки та модифікації.

Сучасний стан досліджень. За останні роки наукова спільнота активно розвиває інструменти захисту медіафайлів, зокрема шляхом поєднання стеганографії, цифрових водяних знаків та криптографічних технологій із методами глибинного навчання. Глибокі нейронні мережі, зокрема автоенкодери, дедалі активніше використовуються у стеганографії, дозволяючи поєднувати класичні методи приховування інформації з інтелектуальними алгоритмами. Це відкриває нові можливості для підвищення надійності, пропускної здатності та стійкості систем до атак завдяки адаптивному кодуванню і здатності моделі до навчання на реальних даних [1]. В іншій роботі розроблено моделі та алгоритми створення цифрових водяних знаків (ЦВЗ) для аудіофайлів на основі сучасних методів інформаційної теорії, статистичного моделювання, цифрової обробки сигналів та нейронних мереж. Показано, що застосування двошарових штучних нейронних мереж прямого поширення для стеганографічного вбудовування дозволяє зменшити спотворення сигналу та підвищити стійкість до атак. Проведено статистичний аналіз виявлення і відновлення ЦВЗ стороннім спостерігачем, що дало змогу оцінити ефективність і безпеку маркування. Результати підтверджують доцільність використання таких методів для контролю авторських прав і діагностики цифрового аудіоконтенту [2].

В дослідженні Бондарь І. В. розроблено моделі та алгоритми створення цифрових водяних знаків (ЦВЗ) для аудіофайлів на основі сучасних методів інформаційної теорії, статистичного моделювання, цифрової обробки сигналів та нейронних мереж. Показано, що застосування двошарових штучних нейронних мереж прямого поширення для стеганографічного вбудовування дозволяє зменшити спотворення сигналу та підвищити стійкість до атак. Проведено статистичний аналіз виявлення і відновлення ЦВЗ стороннім спостерігачем, що дало змогу оцінити ефективність і безпеку маркування. Результати підтверджують доцільність використання таких методів для контролю авторських прав і діагностики цифрового аудіоконтенту [3]. Пропонується і комплексний огляд застосувань блокчейн-технологій для захисту мультимедійного контенту, розробляючи системну таксономію, що охоплює технічні аспекти блокчейну, методи захисту – такі як шифрування, цифрове управління правами, водяні знаки та відстеження транзакцій – а також критерії ефективності, й окреслюють ключові виклики та перспективи подальших досліджень у цій галузі [4]. У роботах різних авторів зустрічається систематизація підходів до прихованого вбудовування даних у зображення, аудіо та відео за допомогою нейромереж, зазначаючи вагомі переваги таких моделей у забезпеченні непомітності та надійності захисту. Подальший розвиток цих напрямків продовжився в гібридних методах, де використовуються перетворення DCT або DWT у поєднанні з GAN, що дозволяє вбудовувати дані в частотну область з підвищеною стійкістю до атак і мінімальним погіршенням візуальної якості. Алгоритми, що комбінують DCT, DWT, а також криптографічні підходи (наприклад, OTP чи Playfair), довели свою ефективність у задачах захисту авторських прав та забезпечення цілісності через вимірювання PSNR, SNR та MSE. Hiroki Tomosada та співавтори у своїй роботі пропонують високоякісний метод усунення розмиття зображень, який ґрунтується на використанні дискретного косинусного перетворення (DCT) і відзначається зниженою обчислювальною складністю. Модель була навчена на спеціально підготовленому наборі даних, що містить зображення з інтенсивними розмиттями, зумовленими рухом. У той час як традиційні підходи до відновлення зображень, засновані на згорткових нейронних мережах

(CNN) і генеративних змагальних мережах (GAN), часто використовують багатомасштабну або багатофрагментну архітектуру для зменшення артефактів і покращення чіткості, запропонований метод *DeblurDCTGAN* досягає подібних або кращих результатів без цих ускладнень. Ключовою інновацією є використання функції втрат на основі DCT, яка дозволяє порівнювати відновлене зображення з еталонним у частотній області. Завдяки цьому підходу вдається ефективно зменшити блокові шуми та кільцеві артефакти, зберігаючи при цьому якість розфокусованих областей. Результати експериментів демонструють, що *DeblurDCTGAN* перевершує традиційні методи за показниками PSNR, SSIM і часом обробки. Крім того, при тестуванні на реальних наборах даних (GoPro, DVD, NFS, HIDE) метод показав підвищену ефективність завдяки кастомізованому набору тренувальних зображень [5]. На Рисунку 1 представлено використання семантичної карти обличчя як апріорної для нейронної мережі для усунення розмиття [6]. Цей метод спочатку отримує семантичні мітки розмитих зображень обличчя за допомогою мережі розбору облич. Потім семантична карта об'єднується з розмитим зображенням у вимірі каналу як вхідними даними мережі для усунення розмиття. Відтворено з дозволу посилання © IEEE 2018.

Окрім того, у 2024 р. з'явилися приклади застосування гібридних підходів до відео – де поєднуються стеганографія та криптографія для вбудовування секретних даних у відеофрейми, що забезпечує подвійний рівень захисту: спершу шифрування, потім – приховане вбудовування у послідовність кадрів. У сфері стеганалітики демонструють ефективні результати невикористовувані CNN-моделі, що здатні виявляти навіть технологічно вдосконалені приховані зв'язки в медіафайлах. У дослідженні C. Ravichandran та співавторів розглянуто сучасні підходи до відеостеганографії як методу прихованого передавання інформації в умовах загроз з боку третіх осіб. Основна увага приділяється методу приховування даних на основі найменш значущого біта (LSB), модифікованого із застосуванням хеш-функцій для підвищення стійкості та безпеки. Автори провели ґрунтовний аналіз існуючих методів стеганографії, спрямованих на збереження якості відео при одночасному підвищенні захищеності прихованих даних, а також вказали на актуальність теми в умовах зростання загроз цифровій безпеці.

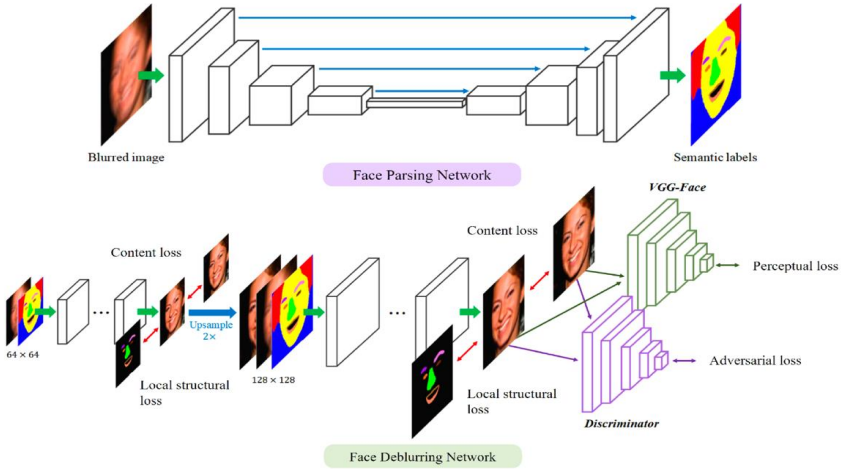

Рисунок 1 – Використання семантичної карти обличчя як апріорної для нейронної мережі для усунення розмиття [6]

Результати моделювання, проведеного у середовищі MATLAB, засвідчили перевагу запропонованого алгоритму над іншими сучасними рішеннями. Запропонований підхід забезпечує надійне приховування даних із мінімальними спотвореннями, що зберігає високу якість відеопотоку та ускладнює виявлення вбудованої інформації сторонніми спостерігачами [7].

Одним із ключових трендів стає розвиток так званої “coverless” стеганографії, при якій дані не вставляються безпосередньо до медіафайла, а реконструюються на основі моделі, що забезпечує високу непомітність для систем детекції. У роботі, присвяченій огляду алгоритмів безвбудовувальної (coverless) стеганографії, автори аналізують понад 90 наукових праць, що охоплюють основні етапи розвитку цієї технології у сфері зображень та відео. На відміну від традиційних методів стеганографії, які вносять незначні зміни в носій інформації (що може бути виявлено за допомогою алгоритмів стегоаналізу), безвбудовувальна стеганографія не модифікує сам носій, а встановлює відповідність між переданим повідомленням і характеристиками обраного об’єкта. Завдяки відсутності змін у структурі носія, така технологія значно ускладнює або навіть унеможливлює

виявлення прихованої інформації сторонніми засобами. У статті узагальнено внески існуючих методів, а також детально розглянуто актуальні проблеми обмеженої місткості, стійкості до атак та забезпечення конфіденційності. Особливу увагу приділено безпеці безвбудовувальної стеганографії – вперше виконано її аналіз як з теоретичної, так і з прикладної точки зору [8]. Таким чином, сучасні дослідження демонструють активний рух у напрямі мультидисциплінарних гібридних рішень, що поєднують крипто-, стего- та машинно-навчальні технології, створюючи новий рівень захисту інтелектуальної та конфіденційної інформації в медіафайлах. Стрімкий розвиток цифрових технологій та широке поширення мультимедійного контенту зумовлює необхідність постійного вдосконалення засобів захисту інформації в медіафайлах. На тлі зростання кількості кібератак, витоків персональних даних, атак типу "Man-in-the-Middle", а також масового використання deepfake-технологій, особливої актуальності набуває проблема забезпечення автентичності та цілісності аудіовізуального контенту. У статті, присвяченої атакам типу «людина посередині» (Man-In-The-Middle, MITM), які становлять одну з найсерйозніших загроз для комп'ютерної безпеки, основна увага приділяється аналізу способів реалізації MITM-атак у контексті моделі взаємодії відкритих систем (OSI), а також у популярних мережевих технологіях GSM та UMTS. Атаки класифікуються за такими параметрами, як розташування зловмисника в мережі, тип комунікаційного каналу та методи імітації (імперсонації). Для кожного з класів MITM-атак наводяться типові етапи їх виконання. Окремо розглянуто існуючі захисні засоби, проведено їх порівняльний аналіз, а також запропоновано систематизацію механізмів запобігання MITM-атакам. На завершення автори окреслюють перспективні напрямки подальших досліджень у цій сфері [9].

Медіафайли дедалі частіше використовуються як доказова база у судовій практиці, у сферах журналістики, правозастосування, розслідувань, військової розвідки, що вимагає надійних механізмів підтвердження їхнього джерела, захисту від маніпуляцій та несанкціонованого копіювання. На відміну від традиційного текстового контенту, медіа мають складну внутрішню структуру, що створює додаткові виклики в контексті прихованого вбудовування або

захисту від змін. Такі характеристики як колірні простори, частотні області, рівні стиснення та втрата якості під час передачі можуть істотно впливати на ефективність методів захисту. У зв'язку з цим все більшої уваги заслуговують алгоритми, які поєднують кілька методик: шифрування даних, цифрові водяні знаки, динамічну стеганографію, детекцію маніпуляцій із застосуванням машинного навчання. Методи захисту, засновані на мультимедійному шифруванні, стають дедалі актуальнішими, оскільки дозволяють безпечно обмінюватися контентом на цифрових платформах. У роботі, присвяченій шифруванню різних типів мультимедійних даних, розглянуто стан безпечних та конфіденційних схем шифрування, що застосовуються до цифрового мультимедіа, такого як цифрові зображення, цифрове відео та цифрове аудіо. Результати огляду сприятимуть кращому розумінню ефективності та надійності наявних рішень, а також стануть підґрунтям для розробки більш досконалих та безпечних схем мультимедійного шифрування в майбутньому [10].

Крім того, особливе значення ця тема має в умовах війни, коли медіаінформація використовується не лише для інформування, а й як інструмент психологічного впливу та дезінформації. Наявність технологій верифікації цифрових зображень та відео, захисту метаданих, а також можливості відстеження джерела походження медіа є критично важливими для забезпечення інформаційної безпеки держави, військових структур і цивільного суспільства.

Таким чином, вибір теми дослідження обумовлений не лише її високою науковою значущістю, але й чіткою практичною необхідністю розробки інструментів, здатних забезпечити цілісний захист інформації в медіафайлах в умовах сучасних гібридних і кіберзагроз.

У межах дослідження визначено проблему підвищення ефективності захисту інформації в медіафайлах, що включає необхідність обрання адаптивних та стійких до сучасних кіберзагроз методів. Актуальними є комбінації криптографічних протоколів, цифрових водяних знаків, стеганографії та машинного навчання з метою забезпечення цілісності, прихованості, незмінності джерела й атрибутації медіаконтенту. Головна задача – сформувати комплекс практичних рекомендацій щодо впровадження захисних технологій у системах, що обробляють цифрові зображення, аудіо- й відеофайли,

враховуючи наукову перспективу та реальні обмеження впровадження в Україні.

Захист інформації в медіафайлах реалізується через чотири основні напрями: криптографічні методи, цифрові водяні знаки, стеганографія та гібридні підходи, часто із застосуванням технологій штучного інтелекту.

Криптографічні технології зберігають конфіденційність даних, однак сам факт шифрування медіафайлу може привертати увагу потенційного зловмисника. Для подолання цих обмежень застосовують цифровий водяний знак – незначні модифікації медіа, що приховані від ока, але дозволяють перевірити походження й цілісність. Наприклад, M. Hamidi та співавтори описали комбіновану техніку на основі DWT, DCT та SIFT, яка забезпечує стійкість до геометричних трансформацій та компресії. У даній роботі запропоновано надійний гібридний метод водяного знакування зображень, що базується на дискретному вейвлет-перетворенні (DWT), дискретному косинусному перетворенні (DCT) та масштабно-інваріантному перетворенні ознак (SIFT). Основна мета дослідження – розробка стійкої до атак схеми маркування, здатної витримувати як обробку зображень, так і геометричні спотворення при збереженні високої непомітності. Для цього водяний знак вбудовується в середній діапазон коефіцієнтів DCT, отриманих з піддіапазону HL1 DWT, що забезпечує захист від стандартної обробки зображень. Алгоритм SIFT застосовується для виявлення ключових точок, які дозволяють компенсувати геометричні викривлення під час вилучення водяного знака. Результати численних експериментів підтверджують високу стійкість методу до атак та спотворень при збереженні високої якості зображення, а також демонструють його переваги порівняно з альтернативними підходами [11]. При цьому стеганографія – це приховане вбитовування повідомлення в сам файл. Класичні методи використовують LSB або частотні перетворення, а сучасні – нейронні мережі і “coverless” підходи, що не змінюють сам медіафайл. Серед них особливо перспективною є DiffStega, яка на основі дифузійної моделі створює stego-зображення без потреби оригінала, з високою стійкістю до атак та сильною прихованістю завдяки Noise-Flip і пароль-залежним запитом. DiffStega – це метод стеганографії зображень без покриття, який використовує

моделі дифузії для вбудовування секретних зображень. Стеганографія без покриття підвищує непомітність, усуваючи необхідність використання зображення-покриття. Замість того, щоб приховувати дані в існуючому зображенні, DiffStega генерує нове зображення, яке містить секретну інформацію. Цей підхід потенційно може запропонувати покращену безпеку та ємність порівняно з традиційними методами стеганографії [12; 13]. Інший напрям – DDIM-моделі для coverless стеганографії з реальним ключем, що дозволяє оновлювати stego-зображення без повторної передачі промтів. DDIM (Denoising Diffusion Implicit Models) – це варіант дифузійних моделей генерації зображень, який використовується для швидшого та керованого створення високоякісного контенту. У контексті coverless стеганографії – це сучасний підхід, коли секретна інформація не вбудовується у файл-носій, а відображається через генерацію або вибір певного медіаоб'єкта, що пов'язаний з цією інформацією. У дослідженні, проведеному авторами статті, присвяченій судово-медичному аналізу дипфейків, запропоновано нову форензичну методику ідентифікації навчального датасету (наприклад, *CelebA* або *FFHQ*), використаного для генерації зображень за допомогою Generative Adversarial Networks (GANs). Запропонована система ґрунтується на інтерпретованому аналізі ознак, який включає спектральні перетворення (Fourier/DCT), метрики колірного розподілу та дескриптори локальних ознак (SIFT). Застосування класифікаторів на основі машинного навчання (Random Forest, SVM, XGBoost) дало змогу досягти 98–99 % точності як у бінарній класифікації (реальні/синтетичні зображення), так і у багатокласовому визначенні джерела датасету. Дослідження підкреслює важливість частотних ознак у виявленні артефактів, характерних для конкретних датасетів, таких як патерни апсемплінгу та спектральні спотворення, тоді як колірні гістограми виявляють непрямі ознаки регуляризації, властиві тренуванню GAN. Окремо розглянуто правові та етичні аспекти, пов'язані з авторським правом, захистом персональних даних та дотриманням нормативних актів (наприклад, GDPR і каліфорнійського закону AB 602). Запропонований підхід сприяє підвищенню відповідальності та прозорості у сфері генеративного ШІ й може бути застосований у цифровій криміналістиці, модерації контенту та вирішенні спорів з інтелектуальної власності [14].

Інші приклади – системи, що командуються через YOLO і хвильові хеші. YOLO (You Only Look Once – Ви тільки дивитесь один раз) – це система виявлення об'єктів, а вейвлет-хеші пов'язані зі стеганографією зображень, яка може використовувати глибокі згорткові нейронні мережі. Виявлення об'єктів, для якого розроблена YOLO, включає ідентифікацію та визначення місцезнаходження об'єктів на зображенні або відео. Стеганографія, з іншого боку, зосереджена на приховуванні інформації в носії, такому як зображення, щоб запобігти виявленню. В одному із досліджень розглянуто сучасний підхід до стеганографії зображень на основі глибокого навчання. Зокрема, робота базується на розвитку методу **DeepWaveletFusion** шляхом інтеграції згорткових нейронних мереж (CNN) з дискретним вейвлет-перетворенням (DWT) для підвищення якості вбудовування та витягання прихованих даних. Запропонований метод, під назвою **DeepWaveletFusionToo**, вирізняється легкістю архітектури та ефективністю, оскільки використовує спеціалізований датасет зображень, трансформованих за допомогою DWT, і навчається з використанням функції втрат середньоквадратичної похибки (MSE), що суттєво знижує складність моделі та обчислювальні витрати. Результати експериментів підтверджують, що **DeepWaveletFusionToo** забезпечує вищу непомітність приховання інформації порівняно з попередником та демонструє конкурентну точність відновлення, що підкреслює простоту та ефективність підходу [15]. У другому дослідженні запропоновано новий підхід до відновлення стереозображень з використанням дифузійних моделей (DM), які раніше майже не застосовувалися в цій галузі через ряд обмежень. Зокрема, проблема подвійного відновлення зображень призводить до зростання обчислювальної складності, а більшість латентних дифузійних моделей орієнтовані на семантичний зміст і при цьому відкидають високочастотні компоненти, які є критично важливими для задач реставрації. Щоб подолати ці обмеження, було запропоновано **DiffStereo** – першу високочастотну обізнану дифузійну модель, спеціально розроблену для стереозображень. Метод передбачає навчання латентного представлення високочастотних компонентів (LHFR) високоякісних зображень. Дифузійна модель використовується для оцінки LHFR у парі стереозображень, після

чого ці представлення інтегруються у трансформерну мережу для відновлення зображень, забезпечуючи збагачення високочастотною інформацією. Важливо, що роздільна здатність LHFR зберігається на рівні з вхідними зображеннями, що дозволяє зберегти текстури, а зменшення кількості каналів знижує обчислювальні витрати. Також реалізовано спеціальний механізм позиційного кодування для забезпечення контекстної глибини під час відновлення. Результати експериментів демонструють, що поєднання дифузійної генеративної моделі з трансформерною архітектурою дозволяє **DiffStereo** перевершити сучасні підходи за точністю реконструкції та сприйняттям якості при вирішенні задач суперроздільності, деблурінгу та покращення зображень у слабкому освітленні [16] Zhiwei Kang описал метод, який за допомогою YOLO вибирає об'єкти для формування послідовності і ключів, що мають високий ступінь стійкості навіть при геометричних атаках. З огляду на переваги розпізнавання орієнтації у просторі скінченного ріджлет-перетворення було запропоновано метод вбудовування водяного знаку в зображення, стійкий до геометричних атак. На першому етапі оригінальне зображення піддається одномірному ріджлет-перетворенню. Далі виконується тривимірне вейвлет-декодування, після чого водяний знак вбудовується у коефіцієнти ріджлет-перетворення. Після цього проводиться зворотне відновлення зображення. Експериментальні результати демонструють високу стійкість запропонованого методу до геометричних атак, таких як обертання, масштабування, зсув і зсув по зрезу (shearing), а також до стандартних атак, включаючи JPEG-компресію, медіанну фільтрацію, додавання гаусового шуму та імпульсного шуму (сіль і перець) [17]. Ranjan та Kumar запропонували метод, що використовує DWT-хешування та Huffman-кодування для формування сховища, яке стійке до масштабування, шуму, JPEG-упаковки та інших атак. Стиснення даних, зокрема за допомогою дискретного вейвлет-перетворення (DWT) та кодування Хаффмана, має вирішальне значення для ефективного зберігання та передачі даних. Метою дослідження була розробка методу для покращення якості зображення та стиснення. Завдяки поєднанню PCA, DWT та канонічного кодування Хаффмана, запропонований метод перевершив існуючі методи з точки зору бітової швидкості, значень PSNR,

CR та SSIM. Об'єктивні та суб'єктивні тести доводять, що новий підхід є ефективнішим для стиснення зображень, що має потенціал для покращення якості зберігання та передачі даних зображень через цифрові мережі [18].

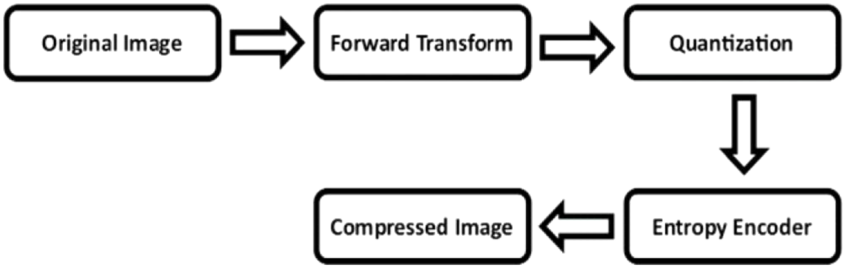

Рисунок 2 – Блок-схема стиснення зображень з втратами в загальному вигляді [18]

Обробка сигналів у зашифрованому середовищі – це перспективна технологія, що має на меті забезпечення захисту цінної інформації під час обробки у потенційно незахищених середовищах. У цьому дослідженні запропоновано метод реалізації дискретного вейвлет-перетворення (DWT) та багатороздільного аналізу (MRA) у гомоморфно зашифрованому середовищі. В одній з робіт спочатку представлено концептуальну модель для виконання DWT та оберненого DWT (IDWT) у зашифрованому вигляді, після чого проведено аналіз впливу розширення даних та квантувальних похибок у межах цієї моделі. З метою зменшення розширення даних – одного з критично важливих факторів для практичного застосування – автори запропонували спеціальний підхід до оптимізації виконання DWT та IDWT. Запропонований метод дозволяє здійснювати багаторівневе DWT/IDWT із мінімальним збільшенням об'єму даних у гомоморфно зашифрованому середовищі. Крім того, розроблено нову процедуру обробки сигналів, у якій на завершальному етапі застосовується метод оберненого множення для обмеження розширення даних. На прикладі 2D- вейвлет-перетворення Хаара було проведено низку експериментів, що продемонстрували ефективність запропонованого підходу для захищеної обробки зображень.

Також виконано аналіз обчислювальної складності та порівняльне дослідження. Наскільки відомо авторам, це перше повідомлення про реалізацію DWT та MRA у зашифрованому середовищі [19]. Значну наукову увагу також привертають гібридні стего-шифрувальні методи, де застосовують DCT/DWT + GAN. Наприклад, поєднання DCT із GAN для прихованого передавання даних у медіа з підвищеною стійкістю до атак та покращення співвідношення швидкість/якість.

Таким чином, сучасні дослідження демонструють поступ до мультидисциплінарних рішень, які комбінують криптографію, стеганографію, цифрові водяні знаки, метадані та машинне навчання. Ці гібриди дозволяють забезпечити ключові характеристики захисту: прихованість, стійкість до атак, перевірку походження контенту та збереження високої якості медіа. Для практики українських дослідників це означає не лише освоєння теоретичних алгоритмів, а й їхню адаптацію до локальних умов: обмеження по обчислювальних ресурсах, законодавчі рамки, інтеграцію в наявні IT-системи.

Класифікація медіафайлів та типів загроз.

Цифрові медіафайли умовно поділяють на зображення, аудіо та відео – кожен із них має унікальні характеристики, що впливають на уразливості. Зображення зазвичай містять дані в пікселях або частотних компонентах (DCT, DWT), аудіо – у тимчасових та спектральних доменах, відео – це послідовність кадрів із додатковими метаданими та руховими векторами. Загрози include стеганографії, де вміст файлу змінюється для прихованого передавання інформації, та stegomalware – випадки, коли зловмисний код вбудовується в мультимедіа й запускається після завантаження. Крім того, є атаки на цілісність: зміни кадрів, компресія, геометричні модифікації, deepfake-маніпуляції, що викликають серйозну загрозу автентичності.

У контексті інформаційної безпеки медіафайли – це цифрові об’єкти, які зберігають візуальну або аудіовізуальну інформацію. Їхня специфіка полягає в складній структурі (пікселі, кодеки, частотні компоненти, метадані), що обумовлює велику кількість векторів атак. Для ефективного захисту перш за все необхідно класифікувати як типи самих файлів, так і загрози, що їм притаманні.

Класифікація медіафайлів:

– Цифрові зображення (JPEG, PNG, BMP, TIFF та ін.). Мають структуру, засновану на матриці пікселів, із можливим використанням частотних перетворень (наприклад, JPEG-компресія використовує DCT). Сприйнятливі до змін у просторовому та частотному доменах.

– Аудіофайли (MP3, WAV, FLAC). Містять звукову інформацію, представлену у часовій або частотній формі. Часто піддаються втратній компресії, що ускладнює вбудовування прихованої інформації без спотворень.

– Відеофайли (MP4, AVI, MKV). Є поєднанням зображень (кадрів) та аудіо, іноді – субтитрів. Складаються з великої кількості компонентів (ключові кадри, рухові вектори, метадані), що відкриває численні вектори для атак і прихованого передавання даних.

– Гібридні медіа (інтерактивні, 3D-файли, AR/VR-контент). Новий тип медіа, що поєднує зображення, відео, аудіо, навігаційні дані, сенсорні метадані. Зокрема, файли формату USDZ або glTF мають складну структуру й передбачають нові ризики – як на рівні контенту, так і метаданих.

Тепер перейдемо до більш детального розгляду кожного типу медіафайлів. Систематизуємо ґрунтовний опис основних форматів мультимедійних даних, з урахуванням їхніх структурних особливостей, способів кодування інформації та специфіки використання в контексті інформаційної безпеки. Це дозволить краще зрозуміти, які характеристики визначають вразливість кожного формату до можливих загроз або маніпуляцій.

Цифрові зображення – це мультимедійні об’єкти, які зберігають інформацію у вигляді двовимірних масивів пікселів. Кожен піксель має колірні або яскравісні значення, що формують візуальне представлення зображення. Існує кілька поширених форматів цифрових зображень, кожен із яких має власні особливості структури та способи кодування даних:

BMP (Bitmap Image File) – формат із простим представленням зображення у вигляді незжатого масиву пікселів, що зберігає інформацію у растровому форматі. Підходить для високої якості, але займає багато пам’яті.

JPEG (Joint Photographic Experts Group) – широко використовуваний формат із стисненням із втратами, що застосовує дискретне косинусне перетворення (DCT) для ефективного стиснення. Забезпечує баланс між якістю та розміром файлу.

PNG (Portable Network Graphics) – формат із безвтратним стисненням, підтримує прозорість та зберігає якість зображення без втрат, використовуючи методи стиснення на основі алгоритмів DEFLATE.

TIFF (Tagged Image File Format) – гнучкий формат, який підтримує як стиснення без втрат, так і з втратами, широко застосовується у професійній графіці та медичних зображеннях.

GIF (Graphics Interchange Format) – формат з обмеженою палітрою кольорів (до 256), що підтримує анімацію, але використовується переважно для простих графічних елементів.

RAW – формат, що містить необроблені дані, отримані безпосередньо з сенсора камери, забезпечує максимальну якість і гнучкість для подальшої обробки.

HEIF/HEIC (High Efficiency Image Format) – сучасний формат із високоефективним стисненням, що дозволяє зберігати зображення з кращою якістю при меншому розмірі файлу, часто використовується на мобільних пристроях.

Ці формати можуть базуватись на різних типах кодування, зокрема:

Просторове кодування – безпосереднє збереження пікселів у матриці (BMP, RAW).

Частотне кодування – застосування перетворень, таких як дискретне косинусне перетворення (DCT, JPEG) або дискретне вейвлет-перетворення (DWT, TIFF), що дозволяє більш ефективно стиснути дані.

Структура і формат впливають на стійкість зображень до змін і на вибір методів захисту, таких як цифрове водякування, стеганографія, шифрування та інші. З позиції авторського права, вбудовані у медіафайли дані слугують надійним засобом запобігання претензіям третіх осіб на право власності. Водночас цифровий водяний знак має залишатися конфіденційним, доступним лише автору, та зберігати свою стійкість до різноманітних видів атак і спроб видалення [20]. Крім того, характеристики форматів визначають рівень

втрат при стисненні, можливість збереження метаданих (EXIF, XMP), підтримку прозорості, анімації або багатошаровості. Розуміння цих особливостей є критично важливим для розробки ефективних систем захисту інформації, оскільки різні формати зображень по-різному реагують на атаки і втручання у просторовому або частотному домені. Серед різноманітних підходів до захисту зображень за допомогою цифрових водяних знаків, особливу увагу приділяють методам, що базуються на частотних перетвореннях, таких як дискретне косинусне перетворення. В одній з робіт розглянуто підхід, у якому з метою забезпечення стійкості цифрових водяних знаків до різноманітних спотворень та стиснення зображень, при збереженні прийнятної візуальної якості, було запропоновано нові критерії вибору низькочастотних коефіцієнтів DCT, придатних для модифікації, а також умови, за якими здійснюється вбудовування цифрових водяних знаків [21].

Аудіофайли є важливим типом медіа, що містить звукову інформацію у цифровій формі, представлено у часовому або частотному домені. У контексті інформаційної безпеки аудіоформати відіграють особливу роль у задачах стеганографії, цифрового водякування, автентифікації й виявлення підробок. Залежно від формату, аудіофайли відрізняються типом стиснення, якістю відтворення, структурою та можливостями прихованого вбудовування даних [22].

WAV (Waveform Audio File Format) – це один із найпростіших та найбільш придатних для аналізу аудіоформатів, що зберігає сигнал у незжатому вигляді. Завдяки відсутності втрат під час збереження, WAV забезпечує повне збереження амплітудно-часової структури сигналу, що робить його ідеальним кандидатом для застосування цифрових підписів або вбудовування водяних знаків без спотворень.

MP3 (MPEG-1 Audio Layer III) є найпоширенішим форматом аудіо з втратами. Він застосовує складні методи перетворення, зокрема модифіковане дискретне косинусне перетворення (MDCT), щоб зменшити розмір файлу, зберігаючи прийнятну якість звуку. Проте під час такого стиснення втрачається частина звукової інформації, що ускладнює приховане вбудовування даних і робить такі файли менш надійними для задач захисту інформації.

FLAC (Free Lossless Audio Codec) – це формат зі стисненням без втрат, який зберігає високу якість сигналу при значно меншому обсязі, ніж WAV. FLAC підтримує додаткові метадані й є придатним для безпечного зберігання і обробки, що робить його ефективним для вбудовування захищеної інформації або використання в системах контролю автентичності.

AAC (Advanced Audio Coding) є сучасним форматом із втратами, який забезпечує кращу якість звуку порівняно з MP3 при однаковому бітрейті. Він широко застосовується у потоковому аудіо та відео (наприклад, YouTube, iTunes). Через стиснення з втратами цей формат також обмежено придатний до надійної стеганографії, однак сучасні методи дозволяють частково компенсувати ці обмеження за допомогою адаптивних алгоритмів.

OGG (Ogg Vorbis) – відкритий формат стиснення з втратами, відомий своєю ефективністю та високою якістю при низьких бітрейтах. Він є альтернативою MP3 та AAC і використовується в ігрових движках, стримінгових платформах, вільному програмному забезпеченні. Його гнучка структура теоретично дозволяє вбудовування прихованих даних, однак, як і інші втратні формати, він потребує стійких методів до перетворень і рекомпресії [23].

OPUS – це формат нового покоління, оптимізований для голосового та музичного сигналу з низькою затримкою, що використовується у VoIP, відеозв'язку, месенджерах (наприклад, Discord, Zoom). Він автоматично адаптується до типу звуку та каналу. У сфері інформаційної безпеки OPUS розглядається як потенційна платформа для захищеної комунікації, однак використання в стеганографії обмежене через динамічну структуру та активну компресію.

AIFF (Audio Interchange File Format) – це формат без втрат, подібний до WAV, розроблений компанією Apple. Він застосовується переважно в професійних аудіостудіях та продуктах Apple. Завдяки збереженню повної звукової інформації він придатний для вбудовування цифрових водяних знаків, автентифікації та інших задач ІБ, де критичною є точність і достовірність сигналу.

Ці характеристики мають вирішальне значення при виборі методу захисту звукової інформації, зокрема у задачах аудіостеганографії, цифрового підпису та детектування змін. Формати зі стисненням без

втрат (WAV, FLAC, AIFF) забезпечують кращу сумісність із надійними ІБ-методами, тоді як втратні формати потребують спеціалізованих рішень, стійких до спотворень [24].

Відеофайли представляють собою складний медіаформат, який об'єднує послідовність зображень (кадрів), аудіодоріжки та іноді субтитри, створюючи інтегрований мультимедійний контент. Найпоширенішими форматами є MP4, AVI, MKV, кожен із яких має власну структуру і механізми стиснення, що впливають на якість та обсяг файлу. Відео містить велику кількість складових елементів, серед яких ключові кадри, рухові вектори, звукові доріжки, інформація про кадрування і метадані, що описують технічні параметри та контекст відтворення. Така багатошарова структура відкриває численні можливості для різноманітних кібератак, зокрема підміни, модифікації або прихованого передавання інформації. У контексті інформаційної безпеки це створює як виклики, так і потенціал для застосування методів захисту – цифрових водяних знаків, стеганографії, криптографічних засобів, які можуть бути вбудовані не лише у візуальну або аудіодоріжку, а й у метадані чи параметри стиснення. Водночас складність формату відео та велика об'ємність даних вимагають розробки спеціалізованих алгоритмів, які забезпечують ефективний захист без втрати якості та сумісності з пристроями відтворення. Та через високу ємність та складну структуру відео, їм часто надається перевага над іншими носіями, такими як зображення, текст або аудіо, для стеганографії. Існує безліч методів приховування секретної інформації у відео, кожен з яких має свої якості та недоліки. Однак у літературі бракує достатньої кількості оглядових статей, які б розглядали всі методи. На основі методу вбудовування методи відеостеганографії класифікуються на дві категорії, а саме: методи просторової області та методи частотної області [25].

MP4 (MPEG-4 Part 14) є одним із найпоширеніших форматів для зберігання відео, який підтримує широкий спектр кодеків, включаючи H.264 і H.265. Цей формат відомий своєю здатністю ефективно компресувати відеодані з мінімальними втратами якості, що робить його популярним для потокового передавання, мобільних пристроїв і вебплатформ. MP4 підтримує інтеграцію аудіодоріжок, субтитрів та метаданих, що дозволяє створювати багатофункціональний

мультимедійний файл. З огляду на складність структури, захист інформації в MP4 вимагає врахування можливих точок впливу – від окремих кадрів і аудіо до метаданих, а також механізмів стиснення і шифрування.

AVI (Audio Video Interleave) – це один із перших форматів відеофайлів, розроблений компанією Microsoft, який зберігає аудіо та відео в одному контейнері без сильного стиснення. Через відсутність ефективних методів стиснення, файли AVI зазвичай мають великий розмір, що обмежує їх застосування у потоковому відео. Проте цей формат часто використовується в архівуванні та обробці відео завдяки своїй простоті. Через відкриту структуру AVI може бути вразливим до модифікацій метаданих і заміни відеопотоку, що ставить завдання розробки надійних засобів цифрового підпису та водяних знаків для забезпечення цілісності та автентичності.

MKV (Matroska Video) – це сучасний відкритий формат-контейнер, який здатен зберігати необмежену кількість аудіо, відео та субтитрових потоків у одному файлі. MKV підтримує високоякісне відео з різними кодеками та дозволяє додавати багатошарові метадані, що робить його популярним у середовищах, де потрібна гнучкість і висока функціональність. Його багатокомпонентна структура створює широкі можливості для вбудовування захисних механізмів – від цифрових водяних знаків у відеопотоках до складних стеганографічних рішень у метаданих. Водночас складність і варіативність MKV вимагає ретельної розробки засобів контролю доступу і цілісності для забезпечення безпеки в мультимедійному середовищі.

Інші менш поширені формати, такі як MOV (формат, розроблений Apple), WMV (Windows Media Video), а також WEBM, часто застосовуються у специфічних сферах, наприклад, у професійному відеомонтажі або вебвідео. Вони мають власні особливості структури, компресії та підтримки метаданих, що також визначає особливості їх захисту від несанкціонованого доступу і модифікацій. Мета полягає в тому, щоб приховати дані таким чином, щоб їх присутність була невиявною. Це досягається шляхом зміни характеристик відео, таких як значення пікселів або бітові потоки, без суттєвого впливу на його візуальну якість [26].

Усі ці формати відрізняються за структурою, методами компресії, підтримкою метаданих і складністю, що безпосередньо впливає на вибір та ефективність методів інформаційного захисту. Забезпечення цілісності, автентичності і конфіденційності відеоконтенту вимагає комплексного підходу, що охоплює як криптографічні механізми, так і специфічні мультимедійні технології, адаптовані до особливостей кожного формату.

Гібридні медіафайли представляють собою сучасний та складний тип мультимедійного контенту, який поєднує у собі різні види даних – зображення, відео, аудіо, а також додаткові навігаційні та сенсорні метадані, що забезпечують інтерактивність і глибоку взаємодію користувача з контентом. Такі файли використовуються в технологіях доповненої (AR) і віртуальної реальності (VR), тривимірній графіці та інтерактивних презентаціях. Формати, зокрема USDZ і glTF, розроблені для забезпечення високої сумісності та зручності візуалізації складних 3D-моделей разом із пов'язаними елементами, такими як текстури, анімації і фізичні властивості. Через багатокомпонентність і складну ієрархію структури, гібридні медіа несуть нові виклики для інформаційної безпеки. Загрози можуть проявлятися як на рівні безпосереднього вмісту – наприклад, уразливості в 3D-моделях, можливість впровадження шкідливого коду чи маніпуляцій із сенсорними даними, так і на рівні метаданих, що часто містять інформацію про місцезнаходження, користувача або сценарії використання. Забезпечення захисту таких медіа вимагає розробки комплексних методів, які поєднують криптографічні засоби, контроль цілісності, автентифікацію елементів контенту та механізми безпечного обміну даними, адаптовані до специфіки інтерактивних і тривимірних форматів.

Класифікація загроз.

Усі загрози можна умовно поділити на п'ять основних категорій, кожна з яких має підвиди, специфічні для типу медіафайлу:

Загрози конфіденційності. Незаконне копіювання або передача медіафайлів (наприклад, перехоплення під час передачі через нешифрований канал). Несанкціонований доступ до локального або хмарного сховища.

– Загрози цілісності. Маніпуляції зі вмістом (зміна пікселів, редагування аудіодоріжки, вставка або видалення кадрів). Атаки типу

“content tampering” (підробка відео- або фотофактів). Вплив через стиснення, обрізання, перекодування.

– Загрози автентичності. Видавання підробленого медіафайлу за оригінальний. Відсутність або підміна цифрового підпису. Фальсифікація метаданих (дати створення, GPS-координати, пристрій зйомки).

– Загрози прихованого передавання інформації (стеганозагрози). Вбудовування прихованих повідомлень (текстів, команд, шкідливого ПЗ) у пікселі, спектральні компоненти, звук або рухи відео.

Використання форматів із надлишковими метаданими або каналами передачі (наприклад, невидимі субтитри чи просторове шумове кодування).

– Шкідливі програми в медіафайлах. Stegomalware: використання медіафайлів як носіїв шкідливого коду, що активується під час відкриття. Polyglot-файли: комбіновані формати (наприклад, PNG + JS), які інтерпретуються по-різному в різних середовищах.

Таблиця 1 – Вектори атак за типами медіафайлів

Тип медіафайла	**Основні загрози**	**Приклади атак**
Зображення	Стеганографія, зміна метаданих, deepfake	LSB-атаки, деструктивне редагування
Аудіо	Вбудовування даних, прихована команда	Audio-steganography, reverse speech
Відео	Модифікація ключових кадрів, вставка фальшивок	Deepfake-відео, insertion attacks
Гібридні медіа	Зміна логіки сцени, викривлення AR/VR-ефекту	GPS-підміна, атаки на сценографію

Джерело: авторська розробка

Методи захисту інформації в медіафайлах

Захист інформації у медіафайлах базується на принципах конфіденційності, цілісності, автентичності та стійкості до несанкціонованого доступу або маніпуляцій. Залежно від типу загроз, застосовують різні технічні підходи – від шифрування до штучного інтелекту [27].

1. Шифрування.

Шифрування – найпоширеніший метод забезпечення конфіденційності. До медіафайлів застосовують як симетричні (AES, Twofish),

так і асиметричні (RSA, ECC) алгоритми. Особливою є проблема обробки великих обсягів даних, адже відео або зображення можуть мати значний розмір, що впливає на продуктивність.

Наприклад, AES-256 широко застосовується в інструментах типу VeraCrypt [28], де можливо створити захищені контейнери або шифрувати цілі логічні диски. Хоча такий метод добре захищає від витоку, він не забезпечує прихованості чи підтвердження.

2. Цифрові підписи.

Цифрові підписи використовуються для перевірки автентичності джерела та цілісності файлу. Вони ґрунтуються на асиметричній криптографії: хеш медіафайлу підписується закритим ключем автора і може бути перевірений будь-ким із публічним ключем.

Стандарти, як-от X.509, або формати електронного підпису для зображень/відео (наприклад, CMS/PKCS#7, Cryptographic Message Syntax) дозволяють інтегрувати підписи у робочі процеси цифрової журналістики, правових доказів та електронного документообігу.

3. Цифрові водяні знаки.

Це метод вбудовування унікального ідентифікатора (водяного знака) в медіафайл, часто – приховано. Залежно від області вставлення, розрізняють:

– просторові методи – зміни пікселів або амплітуд;
– частотні методи – вбудовування у коефіцієнти DCT, DWT, SVD.

Водяні знаки бувають видимими (наприклад, логотип телеканалу) або невидимими (що не впливають на візуальне/звукове сприйняття). Один із популярних інструментів – Digimarc, що реалізує частотне кодування для стійкості до атаки JPEG, масштабування тощо [29].

Сучасні системи використовують і адаптивні методи, де водяний знак підлаштовується під особливості контенту для підвищення непомітності.

4. Стеганографія.

Цей метод дозволяє приховано передавати інформацію в межах звичайного медіафайлу. На відміну від шифрування, стеганографія приховує сам факт існування повідомлення.

Основні методи стеганографії включають різноманітні підходи, що забезпечують приховане вбудовування інформації у цифрові медіафайли, з урахуванням їх структури та особливостей. Один із

найпоширеніших методів – LSB (Least Significant Bit), який полягає у зміні найменш значущих бітів пікселів у зображеннях або аудіосемплах. Цей спосіб відзначається простотою реалізації та високою швидкістю, але може бути вразливим до деяких видів атак і обробки файлів.

Інший потужний напрям – стеганографія у домені перетворень (transform domain). Тут прихована інформація вбудовується не у безпосередньо пікселі чи аудіодані, а у коефіцієнти перетворень, таких як дискретне косинусне перетворення (DCT), дискретне вейвлет-перетворення (DWT) або інші частотні аналоги. Такий підхід забезпечує кращу стійкість до стиснення, фільтрації та різних маніпуляцій із медіа, адже модифікації розподілені у більш стійкому до змін частотному просторі.

Сучасним інноваційним напрямком є coverless стеганографія – метод, що принципово відрізняється від традиційних. У цій технології секретне повідомлення не вставляється безпосередньо у носій, а відтворюється за допомогою моделей штучного інтелекту, таких як генеративні моделі або дифузійні моделі (наприклад, DiffStega). Такий підхід усуває традиційні сліди втручання в носій і значно підвищує захищеність інформації від виявлення.

Крім того, існують і інші важливі методи, серед яких частотне модуляційне кодування, фазове кодування, методи на основі хаотичних систем, а також стеганографія, заснована на змінах текстурних або статистичних властивостей медіаоб'єктів. Всі ці методи мають свої переваги і недоліки, а їх вибір залежить від конкретних завдань захисту, типу носія і вимог до стійкості, ємності та непомітності вбудованої інформації.

Такі інструменти, як OpenStego (LSB, прості водяні знаки) або нейронні системи на базі GAN (наприклад, SteganoGAN, DiffStega) використовуються як для безпечної передачі, так і для обхідних цілей зловмисниками.

5. Штучний інтелект та машинне навчання (AI/ML)

AI-системи застосовуються як для захисту, так і для атак. Для захисту використовують глибокі згорткові мережі (CNN), autoencoder-архітектури, GAN для вбудовування інформації або розпізнавання фальсифікацій.

Diffusion-based methods дозволяють створювати високо приховані й природні образи з інтегрованими повідомленнями.

AI-антифейкові системи використовуються для детекції змінених зображень/відео, фальсифікацій deepfake.

На практиці це дозволяє захистити файли не лише від модифікацій, а й забезпечити автоматичне виявлення підробок у цифрових потоках.

6. Блокчейн

Блокчейн дозволяє забезпечити незмінність метаданих про медіафайл. Наприклад, при завантаженні фото у блокчейн зберігається:

– його хеш;
– ідентифікатор автора;
– часовий штамп;
– геолокація;
– цифровий підпис.

Такі рішення застосовуються в цифровій журналістиці (проєкти типу Content Authenticity Initiative або CAI – це міжгалузева спільнота, заснована компанією Adobe у 2019 році, метою якої є боротьба з дезінформацією в інтернеті шляхом додавання до цифрового контенту інформації про походження та історію змін.) та судовій експертизі. Наприклад, система TruePic документує створення фото/відео верифікованими способами із прив'язкою до блокчейну. TruePic Vision: Цифрова платформа для інспекції, що забезпечує простий та ефективний спосіб отримання перевірених, достовірних зображень та відео безпосередньо від клієнтів. Вона використовується в процесах розгляду бізнес-претензій та страхових вимог. Вона замінює дорогі та трудомісткі візити на місце, пропонуючи рішення для віртуальної інспекції [29].

У сучасній практиці захисту медіаконтенту активно використовуються спеціалізовані інструменти, що реалізують різні підходи – від криптографії до водяного маркування та стеганографії. Розглянемо найбільш відомі з них з огляду на функціональність, ефективність та сферу застосування.

VeraCrypt – це програмне забезпечення з відкритим кодом для шифрування файлів, томів та цілих дисків. Його перевагою є використання стійких алгоритмів (AES, Serpent, Twofish) і підтримка прихованих томів, що додає додатковий рівень обфускації. VeraCrypt дозволяє

ефективно захищати великі обсяги медіафайлів від несанкціонованого доступу, особливо у випадках передачі через зовнішні носії або хмарні сервіси. Завдяки підтримці Windows, macOS і Linux, VeraCrypt є кросплатформним інструментом широкого застосування [28].

Таблиця 2 – Методи захисту інформації у медіафайлах

Метод	**Ціль**	**Інструмент / Приклад**	**Переваги**	**Обмеження**
Шифрування	Конфіденційність та цілісність даних	VeraCrypt, AES-256	Сильний захист, широко підтримується	Високі вимоги до обчислювальних ресурсів
Цифровий підпис	Перевірка автентичності та цілісності	PKCS#7, X.509	Юридична значущість, інтеграція в документообіг	Не приховує дані, потребує інфраструктури ключів
Цифровий водяний знак	Маркування авторства або джерела	Digimarc, DWT/DCT	Стійкість до перетворень, непомітність	Може бути видалений або змінений
Стеганографія	Прихована передача інформації	OpenStego, SteganoGAN	Невидимість передавання, можна комбінувати із шифруванням	Вразлива до стегоаналізу, обмежена місткість
AI/ML	Автоматична детекція/ вбудовування даних	DiffStega, CNN, GAN	Гнучкість, масштабованість, антифейкові системи	Складність реалізації, потребує навчання моделей
Блокчейн	Фіксація походження та незмінності	TruePic, Content Authenticity Initiative	Незмінність записів, прозорість	Висока складність, низька швидкодія

Джерело: авторська розробка

Digimarc – комерційна платформа цифрового водяного маркування, яка пропонує стійке та малопомітне вбудовування маркерів у зображення, відео та аудіо. Ці маркери можуть містити метадані, авторські права або дані автентифікації. Перевага Digimarc – високий рівень захисту від видалення чи фільтрації, що робить її ефективною для захисту цифрових медіа у видавництвах, фотоагентствах і художніх проєктах. Платформа інтегрується з Adobe Photoshop, що спрощує робочі процеси дизайнерів [30].

OpenStego – безкоштовний інструмент для стеганографії, що дозволяє приховано вбудовувати текстові повідомлення у зображення. OpenStego підтримує як класичну стеганографію, так і вбудовування водяних знаків. Це робить його корисним для базової дослідницької та освітньої діяльності, однак алгоритм може бути вразливим до стегоаналізу. OpenStego поширюється з відкритим кодом і може адаптуватися під різні потреби [31].

TruePic – сучасне комерційне рішення, що базується на фіксації походження зображень і відео через хешування та збереження у блокчейні. Система дозволяє автоматично перевіряти, чи не змінювався файл з моменту створення. Це критично важливо для цифрової журналістики та судової експертизи, де доказове значення медіа залежить від їхньої цілісності. Компанія TruePic є частиною ініціативи *Content Authenticity Initiative*, яка підтримується Adobe, Microsoft та іншими [29].

SteganoGAN – один із найсучасніших підходів до глибокої стеганографії на основі генеративних нейронних мереж (GAN). Завдяки машинному навчанню SteganoGAN забезпечує високу якість зображення при збереженні прихованого повідомлення. Це рішення актуальне для захищеного обміну інформацією в умовах високої загрози перехоплення. Використання глибокого навчання дозволяє системі адаптуватися до виявлення нових стегоаналізаторів.

Таким чином, кожен інструмент має свою нішу застосування. Реальна ефективність кожного інструмента залежить не лише від технічних характеристик, а й від контексту використання, підготовки користувача та середовища розгортання.

З огляду на стрімке зростання ролі цифрових медіа в критично важливих сферах, таких як журналістика, оборона та судочинство, особливу увагу слід приділити практичному застосуванню методів захисту інформації в медіафайлах.

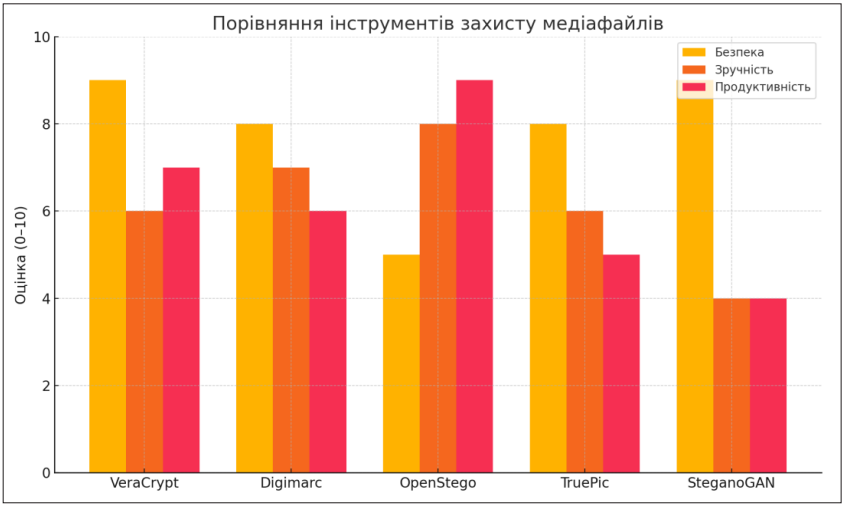

Рисунок 3 – Порівняння інструментів захисту медіафайлів

Джерело: авторська розробка

У цьому розділі розглянуто приклади використання сучасних інструментів і технологій, що забезпечують автентичність, цілісність і конфіденційність цифрового контенту в реальних умовах. У сфері сучасної журналістики захист медіаконтенту є критично важливим не лише для збереження авторських прав, а й для гарантування достовірності джерела в умовах дезінформації. Використання цифрових водяних знаків (Digimarc, Adobe CAI) забезпечує маркування зображень на рівні метаданих, що дозволяє верифікувати їхнє походження у процесі поширення. Системи перевірки достовірності фото- та відеоматеріалів, як-от Project Origin або Content Authenticity Initiative, стали відповіддю на поширення фейкових новин та deepfake-контенту. Журналісти-розслідувачі використовують стеганографію для прихованої передачі чутливої інформації в репресивних країнах, де пряме шифрування може викликати підозру.

У контексті гібридної війни та інформаційних операцій медіа є не лише способом передачі новин, а й об'єктом технічної атаки. Шифрування файлів (VeraCrypt, LUKS, BitLocker) використовується для захисту зображень і відео з дронів, які передаються через польові

станції. AI/ML-інструменти для виявлення маніпуляцій зображеннями використовуються у військовій розвідці для верифікації матеріалів з відкритих джерел (OSINT). Блокчейн для логування метаданих (наприклад, через OpenTimestamps) забезпечує незмінність походження відео, що фіксують військові злочини.

Застосування цифрових медіа у криміналістиці потребує не лише правильного збирання, а й доказового ланцюга автентичності. Цифрові підписи й гешування медіафайлів (SHA-256, PGP) дозволяють зафіксувати момент створення й гарантувати, що файл не змінювався. Інструменти як TruePic або Amber Authenticate дозволяють здійснити верифікацію фото/відео ще на етапі зйомки. Судові експерти застосовують стегоаналіз і цифрову атрибуцію, щоб визначити джерело підроблених фото або факт фальсифікації відео.

Таблиця 3 – Приклади інструментів захисту медіафайлів за сферами застосування

Сфера застосування	Інструмент / Технологія	Призначення
Журналістика	Digimarc, CAI (Adobe)	Верифікація автентичності зображень
Журналістика	SteganoGAN	Приховане передавання інформації
Військова справа	VeraCrypt, BitLocker	Захист відео з дронів
Військова справа	OpenTimestamps	Незмінність походження відео
Судова експертиза	TruePic, Amber Authenticate	Криптографічна фіксація моменту зйомки
Судова експертиза	PGP, SHA-256	Доказова автентифікація медіаданих

Джерело: авторська розробка

Висновки. У межах проведеного дослідження було здійснено системний аналіз сучасних методів захисту інформації в медіафайлах, зокрема шифрування, цифрових водяних знаків, стеганографії, блокчейн-технологій та засобів штучного інтелекту. Встановлено, що ефективний захист мультимедійного контенту в умовах зростаючих кіберзагроз вимагає використання комбінованих підходів, здатних

забезпечити автентичність, цілісність та конфіденційність цифрових об'єктів.

Наукову новизну роботи становить узагальнення та структуризація існуючих методів захисту медіафайлів, запропонована класифікація загроз, а також аналіз можливостей інтеграції технологій глибокого навчання з традиційними інструментами безпеки. Окрему увагу приділено використанню цифрових водяних знаків для аудіофайлів, автоенкодерів для стеганографії та інтелектуальних систем детектування змін у зображеннях.

Практична цінність дослідження полягає у формулюванні рекомендацій щодо захисту цифрових архівів, медіаресурсів у журналістиці, правозастосовній діяльності, кіберрозвідці, військовій справі та цифровій експертизі. Запропоновані методи та інструменти можуть бути використані як для підвищення інформаційної безпеки організацій, так і для особистого захисту цифрового контенту.

Подальші дослідження доцільно зосередити на питаннях глибшої інтеграції штучного інтелекту з криптографічними і стеганографічними підходами, розробці адаптивних методів захисту для нових форматів медіа (наприклад, AR/VR-контенту), а також на просуванні принципів цифрової кібергігієни серед користувачів як ключового чинника протидії дезінформації та технічному втручанню.

Список використаних джерел

1. Прокопович-Ткаченко Д. Глибокі автоенкодери для приховування інформації: сучасні підходи та перспективи розвитку. *Кібербезпека: освіта, наука, техніка* : електронне фахове наукове видання. 2025. Т. 4. №. 28. С. 150–161.

2. Світловський Є. В. Моделі і алгоритми створення цифрових водяних знаків для аудіо-файлів. *Перспективні технології та прилади.* 2024. Т. 1. №. 24. С. 99–106.

3. Бондарь І. В. Метод цифрової стеганографії мультимедіа-об'єктів : дис., Тернопіль : ЗУНУ, 2023.

4. Qureshi A., Megías Jiménez D. (). Blockchain-based multimedia content protection: Review and open challenges. *Applied Sciences.* 2020. № 11 (1), 1. DOI: 10.3390/app11010001 [in English].

5. Tomosada H., Kudo T., Fujisawa T., Ikehara M. GAN-Based Image Deblurring Using DCT Loss With Customized Datasets. *IEEE Access.* 2021. Vol. 9. P. 135224–135233. DOI: 10.1109/access.2021.3116194

6. Shen Z. et al. Deep semantic face deblurring. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018. P. 8260–8269.

7. Ravichandran C., Vajravelu A., Panda S., Degadwala S. D. Data hiding using video steganography. *International Journal of Electronic Security and Digital Forensics.* 2024. Vol. 16 (1). P. 112–123. DOI: 10.1504/ijesdf.2024.136018

8. Meng L., Jiang X., Sun T. A review of coverless steganography. *Neurocomputing*. 2024. Vol. 566. P. 126945.

9. Conti M., Dragoni N., Lesyk V. A survey of man in the middle attacks. *IEEE communications surveys & tutorials*. 2016. Vol. 18 (3). P. 2027–2051.

10. Hosny K. M., Zaki M. A., Lashin N. A., Fouda M. M., Hamza H. M. Multimedia Security Using Encryption: A Survey. *IEEE Access.* 2023. Issue 11. P. 63027–63056. DOI: 10.1109/access.2023.3287858

11. Hamidi M. et al. A hybrid robust image watermarking method based on DWT-DCT and SIFT for copyright protection. *Journal of Imaging.* 2021. Vol. 7 (10). P. 218.

12. Yang Y., Liu Z., Jia J., Gao Z., Li Y., Sun W., Liu X., Zhai G. DiffStega: Towards Universal Training-Free Coverless Image Steganography with Diffusion Models (Version 1). arXiv. 2024. DOI: 10.48550/ARXIV.2407.10459

13. Kheddar H., Hemis M., Himeur Y., Megías D., Amira A. Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. *Neurocomputing.* 2024. Issue 581. P. 127528. DOI: 10.1016/j.neucom.2024.127528

14. Cassia M., Guarnera L., Casu M., Zangara I., Battiato S. Deepfake Forensic Analysis: Source Dataset Attribution and Legal Implications of Synthetic Media Manipulation. In Computer Vision and Pattern Recognition. 2025.

15. Khalifa A., Yadav Y. Wavelet-Based Fusion for Image Steganography Using Deep Convolutional Neural Networks.

Electronics. 2025. Vol. 14, No. 14. P. 2758. DOI: https://doi.org/10.3390/electronics14142758.

16. Cao H., Shi Y., Xia B., Jin X., Yang W. DiffStereo: High-Frequency Aware Diffusion Model for Stereo Image Restoration. 2025.

17. Zhi-wei K. Image Watermarking Against Geometric Attack Based on Finite Ridgelet Transform. In Microcomputer Information. 2012.

18. Ranjan R., Kumar P. An Improved Image Compression Algorithm Using 2D DWT and PCA with Canonical Huffman Encoding. *Entropy*. 2023. Vol. 25 (10). P. 1382. DOI: 10.3390/e25101382

19. Chen D., Wan S., Xiang J., Bao F. S. (2017). A high-performance seizure detection algorithm based on Discrete Wavelet Transform (DWT) and EEG. *PLOS ONE*. 2017. Vol. 12 (3). P. e0173138. DOI: 10.1371/journal.pone.0173138

20. Зубко І. С. и др. Огляд методів нанесення цифрових водяних знаків для захисту зображень / Національний університет «Полтавська політехніка імені Юрія Кондратюка», 2024. С. 121.

21. Яремчук Ю. Є., Карпінець В. В. Аналіз стійкості стеганографічного перетворення до вбудовування цифрових водяних знаків у зображення. *Інформаційні технології та компютерна інженерія.* 2007. № 1. С. 212–217.

22. Овчинников М. Ю., Слатвінська В. М. Технічний та нормативно-правовий аспекти захисту об'єктів інтелектуальної власності на ринку аудіо-візуальної продукції. *Вісник Херсонського національного технічного університету.* 2024. №. 3 (90). С. 259–27.

23. Vorbis audio compression. URL: https://xiph.org/vorbis/

24. Овчинников М. Ю. Додавання та обробка звуку в метаданих файлів за допомогою PHP. *Наука і техніка сьогодні.* 2025. №. 4 (45). С. 1332–1344. DOI: 10.52058/2786-6025-2025-4(45)-1332-1344

25. Disha, Saini, K. (2017). A review on video steganography techniques in spatial domain. 2017 Recent Developments in Control, Automation & Power Engineering (RDCAPE). 2017. P. 366–371. DOI: 10.1109/rdcape.2017.8358298

26. Eltahir M. E., Kiah L. M., Zaidan B. B., Zaidan A. A. High Rate Video Streaming Steganography. *International Conference on Information Management and Engineering.* 2009. P. 550–553. DOI: 10.1109/icime.2009.13

27. Kadian P., Arora S. M., Arora N. Robust Digital Watermarking Techniques for Copyright Protection of Digital Data: A Survey. *Wireless Personal Communications.* 2021. Vol. 118 (4). P. 3225–3249. DOI: 10.1007/s11277-021-08177-w

28. VeraCrypt. Безкоштовне програмне забезпечення з відкритим вихідним кодом для шифрування дисків під Windows, Mac OSX та Linux. URL: https://veracrypt.io/en/Home.html

29. Truepic: The Trusted Virtual Inspection. URL: https://www.truepic.com

30. Product Digitization for Authentication & Consumer Engagement. URL: https://www.digimarc.com

31. OpenStego, the free steganography solution. URL: https://www.openstego.com