МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Г. Ф. Конахович, Д. О. Прогонов, О. Ю. Пузиренко

# КОМП'ЮТЕРНА СТЕГАНОГРАФІЧНА ОБРОБКА Й АНАЛІЗ МУЛЬТИМЕДІЙНИХ ДАНИХ

ПІДРУЧНИК

Київ
«Центр учбової літератури»
2018

УДК 004.[056.5+624+93]
К 338

*Рекомендовано до друку*
*Вченою радою Навчально-наукового інституту аеронавігації*
*Національного авіаційного університету*
*(протокол № 4 від 18.12.2017 р.)*

*та Вченою радою Національного авіаційного університету*
*(протокол № 5 від 24.01.2018 р.)*

**Рецензенти:**
***Мачуський Є. А.*** — д-р техн. наук, професор (Національний технічний університет "Київський політехнічний інститут" ім. І. Сікорського);
***Юдін О. К.*** — д-р техн. наук, професор (Національний авіаційний університет);
***Толюпа С. В.*** — д. т. н., професор (Київський національний університет імені Тараса Шевченка).

**К 338** **Комп'ютерна стеганографічна обробка й аналіз мультимедійних даних**: підручник. / Г. Ф. Конахович, Д. О. Прогонов, О. Ю. Пузиренко. – Київ: «Центр учбової літератури», 2018. – 558 с.

**ISBN 978-617-673-741-4**

Розглянуто теоретичні й практичні основи обробки мультимедійних даних з використанням комп'ютерної стеганографії, а також методи стеганографічного аналізу графічного контенту інфокомунікаційних систем. Наочно показано особливості використання універсальної математичної системи Mathcad у цілях стеганографічної обробки даних.

Представлено приклади практичної реалізації приховання даних у статичних зображеннях, аудіосигналах і тексті.

Системно викладені проблеми надійності та стійкості довільної стеганографічної системи по відношенню до різноманітних типів атак, а також оцінки пропускної здатності каналу прихованого обміну даними.

Представлені результати інформаційно-практичного дослідження проблеми стеганографічного аналізу цифрових зображень.

ISBN 978-617-673-741-4

# ЗМІСТ

# ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

| | |
|---|---|
| АК | ансамбль класифікаторів |
| АКФ | автокореляційна функція |
| АС | автоматизована система |
| АСА | активний стеганографічний аналіз |
| БК | базовий класифікатор |
| БЧХ | код Боуза-Чоудгурі-Хоквінгема |
| ВА | варіограмний аналіз |
| ВКФ | взаємокореляційна функція |
| ВЧ | високочастотний (сигнал) |
| ГПВП | генератор псевдовипадкової перестановки |
| ГПВФ | генератор псевдовипадкової функції |
| ДВП | дискретне вейвлет перетворення |
| ДКП | дискретне косинусне перетворення |
| ДМО | демаскуюча ознака |
| ДПФ | дискретне перетворення Фур'є |
| ЗСД | загальний стеганографічний детектор |
| ЗСЛ | зорова система людини |
| ІКМ | імпульсно-кодова модуляція |
| КВЗ | канал відкритого зв'язку |
| КЛР | класифікатор на основі лінійної регресії |
| КПЗ | канал прихованого зв'язку |
| КППД | канал передавання прихованих даних |
| КС | комп'ютерна стеганографія |
| ЛДФ | лінійний дискримінант Фішера |
| ЛРЗЗЗ | лінійний регістр зсуву зі зворотним зв'язком |
| МФА | мультифрактальний аналіз |
| МФС | мультифрактальний спектр |
| МФФА | мультифрактальний флуктуаційний аналіз |
| МЯ | метрика якості |
| НЗБ | найменший значущий біт |
| НЧ | низькочастотний (сигнал) |
| ОПЦЗК | область перетворення цифрового зображення-контейнера |
| ПА | перетворення Арнольда |

| | |
|---|---|
| ПВП | псевдовипадкова послідовність |
| ПВЧ | псевдовипадкове число |
| ПЗ | пропускна здатність |
| ПКЛ | перетворення Карунена-Лоева |
| ППЗ | прихована пропускна здатність |
| ПСА | пасивний стеганографічний аналіз |
| РГБ | розмірність Гаусдорфа-Безіковича |
| РС | розширення спектра сигналу |
| РСПП | розширення спектра сигналу прямою послідовністю |
| СА | стеганографічний алгоритм |
| СД | стеганографічні дані |
| СМ | стеганографічний метод |
| СМК | статистична модель контейнера |
| ССЛ | слухова система людини |
| СУЕГ | спектр узагальнених експонент Герста |
| СУФР | спектр узагальнених фрактальних розмірностей |
| СЧ | середньочастотний (сигнал) |
| УЕГ | узагальнені експоненти Герста |
| УСД | універсальний стеганографічний детектор |
| ФҐ | фільтр Ґауса |
| ЦВЗ | цифровий водяний знак |
| ЦЗ | цифрове зображення |
| ЦЗК | цифрове зображення-контейнер |
| ЦОС | цифрова обробка сигналів |
| ЦС | цифрова стеганографія |
| ШПФ | швидке перетворення Фур’є |

| | |
|---|---|
| *ASCİİ* | американський стандартний код для обміну інформацією (*American Standard Code for İnformation İnterchange*) |
| *BD* | формат оптичних дисків, що використовуються для зберігання відео високої чіткості і даних (*Blu-ray Disc*) |
| *BMP* | формат бітового відображення графічного об'єкта (*BitMaP*), в якому растрове зображення зберігається у вигляді двовимірного масиву пікселів |
| *CR* | службовий *ASCİİ*-код, що позначає операцію повернення курсору (каретки) — переведення його до лівого краю аркуша при виведенні тексту на символьний пристрій (*Carriage Return*) |
| *CRC* | контроль циклічним надмірним кодом (*Cyclic Redundancy Check*) |
| *DCT* | дискретне косинусне перетворення (*Discrete Cosine Transform*) — математичне перетворення, використовуване в алгоритмах компресії зображень (наприклад, у *JPEG*) |
| *FDCT* | пряме дискретне косинусне перетворення (*Forward Discrete Cosine Transform*) |
| *GİF* | формат обміну графічними даними (*Graphics İnterchange Format*), широко використовуваний для зберігання простих растрових зображень, що містять великі поля одного кольору |
| *İDCT* | зворотне дискретне косинусне перетворення (*İnverse Discrete Cosine Transform*) |
| *İTU* | Міжнародний союз електрозв'язку (*İnternational Telecommunication Union*) |
| *JPEG* | розроблений групою експертів з машинної обробки фотографічних зображень (*Joint Photographic Experts Group*) стандарт стиснення зі втратами повноколірних нерухомих зображень на основі алгоритму дискретного косинусного перетворення |
| *LF* | службовий *ASCİİ*-код, який викликає переведення курсору на екрані до тієї ж самої колонки на один рядок нижче (*Line Feed*) |
| *LFSR* | лінійний регістр зсуву зі зворотним зв'язком (*Linear Feedback Shift Register*) |
| *LSB* | молодший значущий біт (розряд) двійкового числа (*Least Significant Bit*) |

| | |
|---|---|
| *MPEG* | група стандартів на стиснення рухомих зображень і звуку (*Motion Picture Experts Group*) |
| *MSB* | старший значущий біт (розряд) двійкового числа (*Most Significant Bit*) |
| *PCM* | імпульсно-кодова модуляція (*Pulse Code Modulation*) |
| *RGB* | основна палітра: «червоний, зелений, синій» (*Red-Green-Blue*), що використовується у комп'ютерній графіці та програмуванні |
| *SS* | розширення спектра сигналів (*Spread Spectrum*) |
| VPN | віртуальна приватна мережа (*Virtual Private Network*), підмережа корпоративної мережі, яка забезпечує безпечне входження в неї віддалених користувачів |
| *XOR* | виключна диз'юнкція (нееквівалентність, додавання за модулем 2) (*eXclusive OR*) — бінарна логічна операція, результат якої є істинним лише тоді, коли значення операндів не збігаються |

# ВСТУП

Інформація є однією з найцінніших речей у сучасному житті. З появою глобальних комп'ютерних мереж отримання доступу до неї надзвичайно спростилося. Утім, за відсутності заходів щодо захисту даних, легкість і швидкість такого доступу значно підвищили й рівні таких загроз, як неавторизований доступ, фальсифікація, «піратство» тощо.

Задачі надійного захисту авторських прав, прав інтелектуальної власності або конфіденційних даних (які здебільшого мають цифровий формат) є одними з найстаріших і повністю не вирішених на сьогодні. У зв'язку з інтенсивним розвитком і поширенням технологій, які дозволяють за допомогою персонального комп'ютера інтегрувати, обробляти і синхронно відтворювати різноманітні типи сигналів (так звані *мультимедійні технології),* питання захисту інформації, представленої у цифровому форматі, є надзвичайно актуальними. Переваги представлення і передавання саме цифрових даних (простота відновлення, висока потенційна завадостійкість, перспективи використання універсальних апаратних і програмних рішень) можуть бути перекреслені з тією ж легкістю, з якою можливі їх викрадення і модифікація. Тому в усьому світі давно назрілим є питання розробки методів і заходів захисту інформації організаційного, методологічного і технічного характеру, основними серед яких є методи криптографії і стеганографії.

*Криптографічний* (з грецької *χρυπτός* — «таємний», *γράφω* — «пишу») захист інформації (система зміни останньої з метою зробити її незрозумілою для непосвячених, приховання змісту повідомлень за рахунок їх шифрування) не знімає зазначену вище проблему повністю, оскільки наявність шифрованого повідомлення сама по собі привертає увагу і зловмисник, заволодівши, наприклад, захищеним криптографічно файлом, маючи підозри про розміщення в ньому певної секретної інформації і за наявності належної зацікавленості цілком здатен скористатися наявним у його розпорядженні обчислювальним ресурсом для дешифрування даних.

Приховання ж самого факту існування секретних даних при їх передаванні, зберіганні чи обробці є задачею *стеганографії* (з грецької *στεγανός* — «прихований») — науки, що вивчає способи і методи приховання конфіденційних відомостей. Задача видобування інформації при цьому відступає на другий план і в більшості випадків розв'язується стандартними криптографічними методами. Інакше кажучи, під приховання існування розуміється не лише унеможлив-

лення виявлення в перехопленому повідомленні наявності іншого (прихованого) повідомлення, але й взагалі зробити неможливим викликання підозр стосовно цього, оскільки в останньому випадку проблема інформаційної безпеки повертається до стійкості криптографічного коду. Таким чином, займаючи свою нішу в забезпеченні інформаційної безпеки, стеганографія не замінює, а, скоріше, доповнює криптографію [1]. Процедура стеганографічного захисту може здійснюватися найрізноманітнішими способами, загальною рисою яких є те, що приховане повідомлення вбудовується в деякий не привабліюючий увагу об'єкт (контейнер), який згодом відкрито транспортується (надсилається) адресатові.

Історично напрямок стеганографічного приховання інформації був першим [2], але згодом з багатьох причин він був витіснений криптографією. Інтерес до стеганографії відродився наприкінці XX ст. і був пов'язаний з масовим розповсюдженням технологій мультимедіа (що є цілком закономірним, з огляду на зазначені вище проблеми, пов'язані з захистом інформації). Не менш важливою стала і поява нових способів здійснення інформаційного обміну, що в сукупності з першим фактором надало нового імпульсу розвиткові та удосконаленню стеганографії, сприяло виникненню нових стеганографічних методів, в основу яких було закладено особливості представлення інформації в комп'ютерних файлах, обчислювальних мережах і т. п. Це, у свою чергу, дає можливість казати про існування окремого напрямку у сфері захисту інформації — *комп'ютерної стеганографії* (КС) [3–5, 19].

З 1996 р. проводяться міжнародні симпозіуми, присвячені проблемам приховання даних (*İnformation Workshop on İnformation Hiding*). Перша конференція, присвячена стеганографії, відбулася у липні 2002 р. На сьогодні стеганографія вже є наукою, яка продовжує швидко і динамічно розвиватися, використовуючи при цьому методи і досягнення криптографії, цифрової обробки сигналів, теорії зв'язку та інформації.

Приховане передавання даних (так звана *класична стеганографія*) є не єдиною сферою застосування КС. Методи стеганографії дозволяють успішно вирішувати задачі завадостійкої автентифікації, захисту інформації від несанкціонованого копіювання, відстеження шляхів поширення інформації мережами зв'язку, пошуку інформації в мультимедійних базах даних тощо. Усі ці обставини дозволяють в межах традиційно існуючих інформаційних потоків чи інформаційного середовища вирішувати і певні важливі питання захисту інформації низки прикладних галузей.

Існують два ключових напрямки використання КС: пов'язаний з цифровою обробкою сигналів (ЦОС) і не пов'язаний. У першому випадку секретні повідомлення вбудовуються у цифрові дані, які, як правило, мають аналогову природу свого походження (мова, зображення, аудіо- і відеозаписи) [1, 3–5]. У другому — конфіденційна інформація розміщується в заголовках файлів чи пакетів даних. Але цей напрямок не знайшов широкого застосування через відносну легкість розкриття і/або знищення прихованої інформації. Переважна більшість поточних досліджень в галузі стеганографії так або інакше пов'язана са́ме з ЦОС. Що, у свою чергу, дозволяє вже говорити про *цифрову стеганографію* (ЦС) [5, 19].

Можна виділити щонайменше дві причини популярності в наш час досліджень в галузі стеганографії: обмеження на використання криптографічних засобів в низці країн світу і поява проблеми захисту прав власності на інформацію, представлену у цифровому вигляді.

Перша причина спричинила велику кількість досліджень у дусі класичної стеганографії (тобто приховання власне факту здійснення передавання), друга — не менш чисельні роботи в галузі так званих цифрових водяних знаків (ЦВЗ) — спеціальних міток, приховано вбудовуваних до різноманітного мультимедійного контенту з метою можливості контролювання його використання.

Приховання інформації лише завдяки факту невідомості зловмиснику методу або методів, закладених до основи приховання, на сьогоднішній день є малоефективним. Ще у 1883 р. фламандський криптограф Оґюст Керкгоффс *(A. Kerckhoffs)* писав про те, що система захисту інформації повинна виконувати покладені на неї функції навіть за повної поінформованості супротивника про її структуру і алгоритми функціонування [6]. Уся секретність системи захисту передаваних повідомлень має міститися лише в ключі — фрагменті інформації, що є попередньо (як правило) розділеним між адресатами. Незважаючи на те, що цей принцип відомий вже більше 100 років, і досі зустрічаються розробки, що ними зневажають. Очевидно, що вони можуть застосовуватися лише з навчальною метою.

В основі багатьох підходів до вирішення задач стеганографії лежить загальна з криптографією методична база, яку заклав ще всередині минулого століття Клод Шеннон *(C. E. Shannon)* [45, 60]. Але й дотепер теоретичні основи стеганографії є недостатньо висвітленими у сучасній літературі.

Беручі до уваги вищесказане, можна зробити висновок, що на сьогодні існує актуальна науково-технічна проблема удосконалення алгоритмів і методів проведення стеганографічного приховання конфіденційних даних або захисту авторських прав на певну

інформацію мультимедійного типу. На сьогодні абсолютно не бракує стеганографічних програм як початкового, так і професійного рівня (*S-Tools*, *Steganos Security Suite*, *bmpPacker* та ін.). Але захищеність їх коду (особливо це стосується програм професійного рівня) не дозволяє простежити методи, закладені в основу алгоритмів їх дії. Викладені ж на *İnternet*-ресурсах численні тексти програм, через свою низьку інформативність для непрограмістів, мало чим зараджують, оскільки їх компіляція має своїм результатом вже виконувану програму з вкрай важко простежуваним алгоритмом дії, оскільки видає вже готовий результат — заповнений стеганоконтейнер, — заздалегідь встановити достатність рівня прихованості конфіденційної інформації в якому практично не можливо. Отже, є цілком очевидною нестача са́ме програм навчального (початкового) рівня, які б крок за кроком наочно демонстрували весь процес стеганографічного перетворення, що можна було б використати в учбовому процесі при підготовці фахівців у сфері захисту мультимедійного контенту в сучасних телекомунікаційних системах і мережах.

Стан порушеного питання у сфері стеганографії характеризується наступними основними досягненнями. Питання стеганографічного приховання секретних повідомлень, включно з побудовою ефективних алгоритмів приховання, свого часу розглядали у своїх роботах Ґуставус Сіммонс *(G. J. Simmons)*, Джесіка Фрідріх *(J. Fridrich)*, Рос Андерсон *(R. J. Anderson)*, Даніель Ґрул *(D. Gruhl)*, Норішіґе Морімото *(N. Morimoto)*, Крістіан Кешін *(C. Cachin)*, Іоаніс Пітас *(İ. Pitas)* та ін. [7–9, 13–16]. Результати досліджень стеганографічних алгоритмів на стійкість приводять у своїх працях *J. Fridrich*, Річард Попа *(R. Popa)*, Ніл Джонсон *(N. F. Johnson)*, Сушіль Джаджодіа *(S. Jajodia)*, Святослав Волошиновський *(S. Voloshynovskiy)* та ін. [9, 17, 18, 20, 40, 41]. Також необхідно відзначити праці авторів Біргіт Фіцманн *(B. Pfitzmann)*, Брюс Шнаєр *(B. Schneier)* і Скотт Крейвер *(S. Craver)* з питань узгодження термінології та формування основних стеганографічних протоколів [10–12].

Тривалий час у вітчизняній літературі і літературі країн СНД стеганографії було присвячено лише декілька оглядових журнальних статей [1, 4, 22–24, 48]. Крім того, заслуговують на увагу роботи [3] під авторством Володимира Хорошка, Олексія Азарова, Михайла Шелеста і Юрія Яремчука, а також [5] під авторством Вадима Грибуніна, Ігоря Окова та Ігоря Турінцева, заслугою яких є чи не перша спроба системного викладення стеганографічних методів, узагальнення найостанніших результатів досліджень у сфері комп'ютерної стеганографії. Вже згодом з'явилися більш вузькоспеціалізовані видання на зразок [109–111,123].

У 2006 р. вийшла друком і монографія [112] Георгія Конаховича і Олександра Пузиренка, метою якої було викладення теоретичних і, що не менш важливо в освітніх цілях, практичних основ комп'ютерної стеганографії, для чого були розглянуті особливості і перспективи використання в цілях стеганографічного захисту інформації вельми наочної і доступної для розуміння пересічними користувачами системи символьної математики *Mathcad*. Книга, що її ви наразі тримаєте у своїх руках, є другим, дещо переробленим і більш розгорнутим у плані стеганографічного аналізу виданням.

Проведено аналіз спеціалізованих літературних джерел та ресурсів мережі *Internet* щодо перспективних напрямків, за якими можливе використання стеганографії як інструменту захисту інформації в автоматизованих системах обробки даних. Шляхом дослідження відомих публікацій вітчизняних і закордонних авторів здійснено системне викладення проблем надійності і стійкості довільної стеганографічної системи по відношенню до видів здійснюваних на неї атак, а також оцінки пропускної здатності каналу прихованого обміну даними, яким, по суті, і є стеганографічна система. Наведено результати існуючих інформаційно-теоретичних досліджень проблеми інформаційного приховання у випадку активної протидії порушника.

Здійснено системне викладення відомих стеганографічних методів, спрямованих на приховання конфіденційних даних у комп'ютерних файлах графічного, звукового і текстового форматів.

Наведено приклади програмних комплексів для демонстрації принципів, закладених в основу методів стеганографічного приховання інформації у просторовій (часовій) або частотній областях використовуваного мультимедійного контейнера.

Використання під час комп'ютерного моделювання універсальної математичної системи *Mathcad* дозволяє використовувати потужні засоби реалізації чисельних методів розрахунку і математичного моделювання у поєднанні з можливістю виконання операцій символьної математики [25, 26]. Сторони, що здійснюють прихований обмін даними, практично позбавляються необхідності у програмуванні власне розв'язку задач, на них лише покладається коректний опис алгоритму розв'язку на вхідній мові *Mathcad*, що є мовою дуже високого рівня. Зазначене є суттєвою перевагою у порівнянні з існуючими на сьогодні додатками, написаними за допомогою вузькоспеціалізованих мов програмування на зразок *C/C++*, *Java* тощо. Останні, хоча і відрізняються більш високим рівнем гнучкості з точки зору можливостей реалізації тих або інших методів стеганографії, проте характеризуються незрівнянно тривалішим внесенням змін до вже написаної програми і наступної її компіляції. Час, затрачуваний на

внесення модифікацій, стає особливо важливим у випадку багатоетапних досліджень, що мають місце при використанні програми в навчальному процесі.

Завдяки своїй наочності та можливості швидкого проведення модифікацій програмних модулів, розроблені комплекси відповідають вимогам, що ставляться до програм, використовуваним у навчальних цілях. Підхід поєднання теоретичного викладення матеріалу з демонстрацією його практичного використання дозволяє позбавитися абстрактності формулювань, прийнятої у спеціалізованій і довідковій літературі з інформаційної безпеки, і сприяє розвитку у студентів здорового інтересу до практичних аспектів вирішення науково-технічних задач із захисту інформації. Книга може використовуватися в якості довідникового посібника з питань комп'ютерної стеганографії при використанні сучасних комп'ютерно-математичних систем.

Також одним з основних завдань даної книги є демонстрація ключових принципів, закладених в основу поширених на сьогодні методів стеганографічної обробки мультимедійних даних з можливістю проведення стеганоаналізу.

Книгу призначено для фахівців, які працюють у сфері захисту інформації і зацікавлені в ефективному використанні можливостей сучасних обчислювальних систем, а також для студентів і викладачів ВНЗ, які навчаються чи спеціалізуються у сфері інформаційної безпеки в телекомунікаційних системах і мережах.