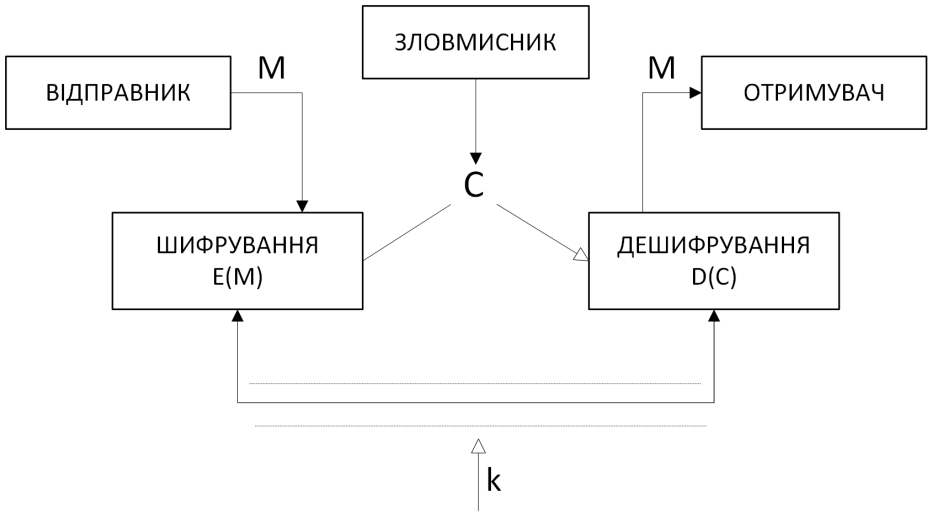
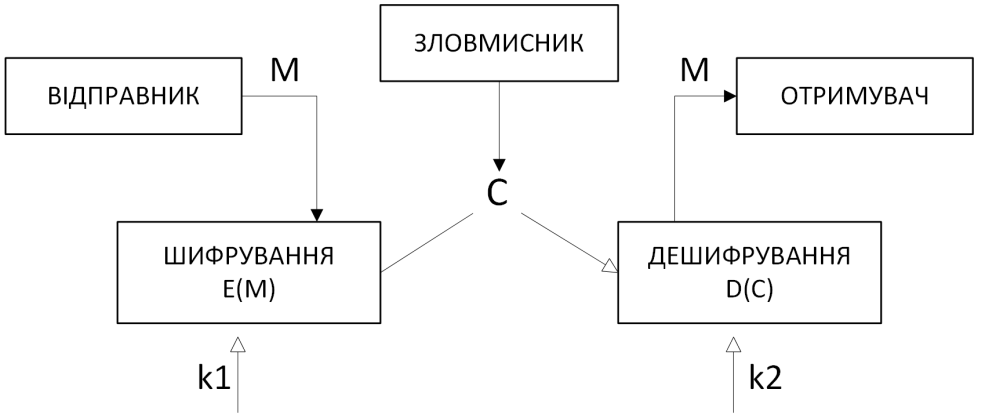
Львівський національний університет імені Івана Франка

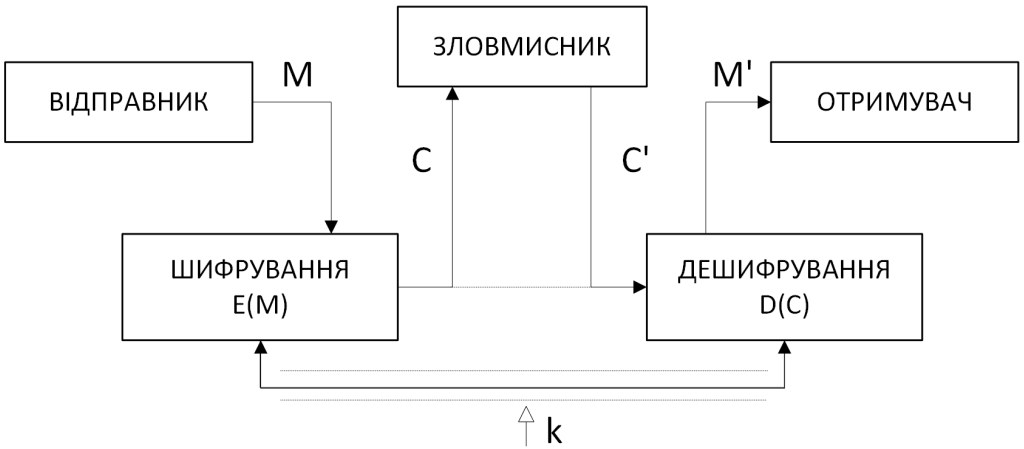
Порівняння ефективності класів стеганографічних алгоритмів

Грициндишина Віталія Любомировича

24 травня 2022 р.





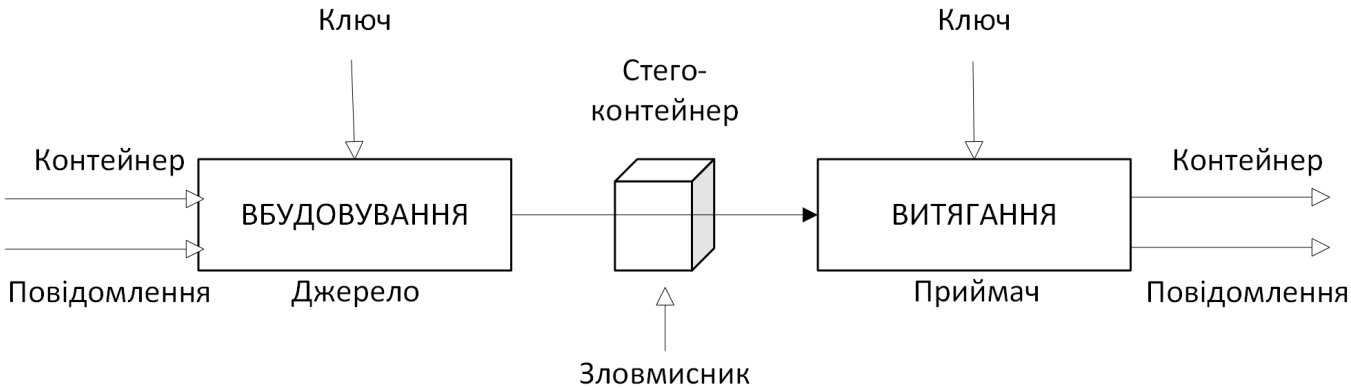


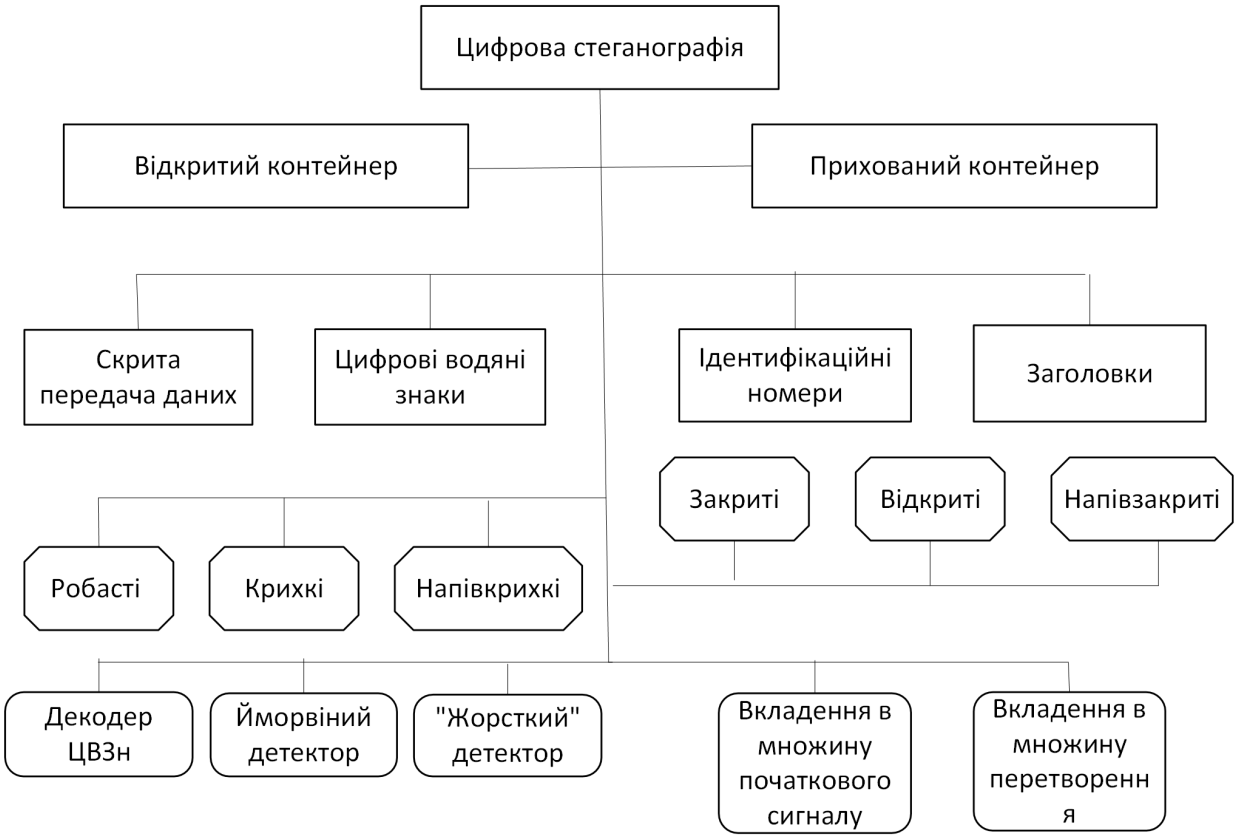
$$C = E_k(M)$$

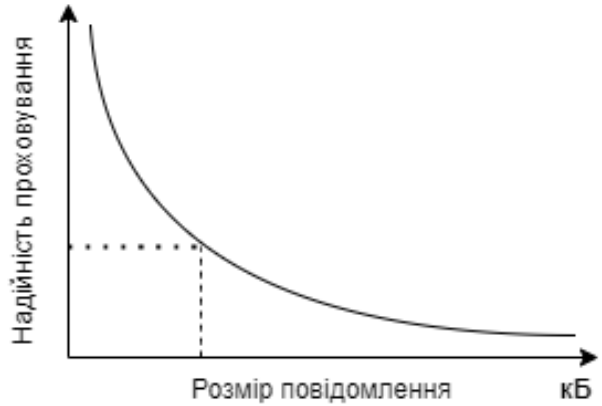
$$D = E_k^{-1}$$

$$D_k(C) = E_k^{-1}(E_k(M)) = M$$

$$E_k : \overline{M} \to \overline{C}$$


Ключ
Стего-
контейнер
Ключ
Контейнер
ВБУДОВУВАННЯ
ВИТЯГАННЯ
Контейнер
Повідомлення
Джерело
Приймач
Повідомлення
Зловмисник


Цифрова стеганографія
Відкритий контейнер
Прихований контейнер
Скрита передача даних
Цифрові водяні знаки
Ідентифікаційні номери
Заголовки
Закриті
Відкриті
Напівзакриті
Робасті
Крихкі
Напівкрихкі
Декодер ЦВЗн
Йморвіний детектор
"Жорсткий" детектор
Вкладення в множину початкового сигналу
Вкладення в множину перетворення


Надійність проховування
Розмір повідомлення
кБ

$$E : C \times M \to S; \qquad (1)$$

$$D : S \to M, \qquad (2)$$

$$E(c_a, m_a) \cap E(c_b, m_b) = \emptyset \qquad sim[c, E(c, m)] = 1, \forall m \in M, c \in C \text{ де}$$

$$(c_a, m_a), (c_b, m_b) \in S \qquad c = max(sim[x, E(x, m)]) \qquad (3)$$

$\dot{S} = \{(c_1, m_1), (c_2, m_2), ..., (c_q, m_q), \} = \{s_1, s_2, ..., s_q\}$ - множина стеганограм (заповнених контейнерів), $m_a, m_b \in M$,
$(c_a, m_a), (c_b, m_b) \in S$
$E(c, m) \approx E(c + \delta, m)$ або $D[E(c, m)] \approx D[E(c + \delta, m)]m$

Означення

Нехай C - непорожня множина, тоді функція $sim(C) \to (-\infty, 1]$ є функцією подібності на множині C, якщо $\forall x, y \in C$ справедливо, що $sim(x, y) = 1$ у випадку $x = y$ і $sim(x, y) < 1$ при $x \neq y$.

$E(c, m) \approx E(c + \delta, m)$ або $D[E(c, m)] \approx D[E(c + \delta, m)]m$

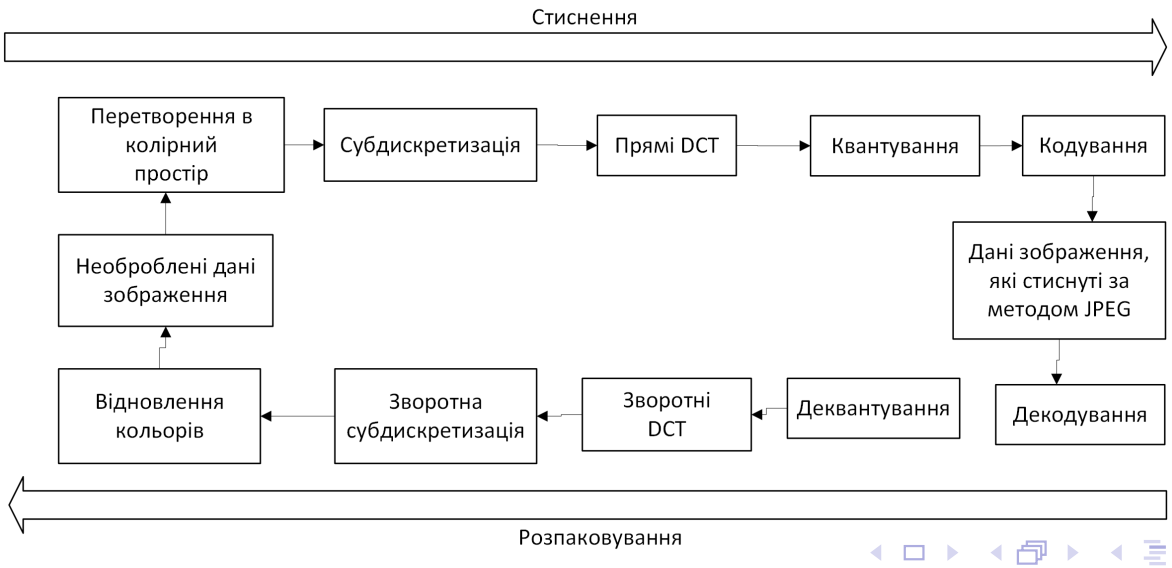
GIF формат:

- Ідентифікатор GIF;
- дескриптор екрана;
- дескриптор зображення;
- локальна таблиця кольорів;
- растрові дані;
- термінатор GIF;
- розширений блок GIF.

BMP формат:

- заголовок;
- інформаційний заголовок ;
- палітра;
- дані зображення.

JPEG формат:



Деякі методи приховування даних:

- Метод заміни найменш значущого біта;

$$\begin{array}{lll} (00100111 & 11101001 & 11001000) \\ (00100111 & 11001000 & 11101001) \\ (11001000 & 00100111 & 11101001) \end{array}$$

Коли вставляється символ A, двійкове значення якого дорівнює 10000001, виходить така сітка:

$$\begin{array}{lll} (00100111 & 11101000 & 11001000) \\ (00100110 & 11001000 & 11101000) \\ (11001000 & 00100111 & 11101001) \end{array}$$

- Метод маскування та фільтрації;
  Методи обмежують 24 бітами. Їх порівнюють до водяних знаків, створюючи позначки, трохи змінюючи яскравість на окремих частинах зображення.

- Метод перетворення.;

$$F(u,v) = \frac{1}{4}C(u)C(v)\left[\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)\cdot \cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16}\right],$$

де $C(x) = 1/\sqrt{2}$, коли $x = 0$ та $C(x) = 1$ в інших випадках. Після обчислення коефіцієнтів виконують операцію квантування:

$$F^{Q}(u,v) = \left\lfloor \frac{F(u,v)}{Q(u,v)} \right\rfloor,$$

де $Q(u,v)$ - 64 елементна таблиця квантування.

До методів приховування інформації у текстових файлах належать:

- синтаксичні методи;
- семантичні методи, які обумовлюють два синоніми, що відповідають значенням прихованих бітів;
- методи довільного інтервалу.

Метод	**Знаків стега**	**Щільність,%**
Чергування маркерів кінця	267	0.21
Вирівнювання пробілами	411	0.32
Двійкові нулі	740	0.58
Хвостові пропуски	1071	0.85
Знаки одного накреслення	4065	3.21

Дослідження ефективності цих методів на щільність було проведене на одному з томів книги «Мертва вода» розміром понад 126729 тисяч байтів та налічує 2143 рядки, що вирівняли на 65-символьну, по 4-символьному абзацному відступі.

Вимоги до стегоситеми:

- стійкість даних до фільтрації, різних шумів, втрат зі стистом, аналогових та цифрових модифікацій;
- у випадку ЦВДЗ бажання позбутись даних може пошкодити контейнер;
- дані не мають докладати в сигнал спотворення помітні людині;
- дані не мають значно впливати на сатистику контейнера.

Класи аудіосигналів:

- розмова телефонної якості (300 – 3400 Гц);
- широкосмугова мова (50 – 7000 Гц);
- широкосмугові аудіосигнали (20 – 20000 Гц).

Методів приховування інформації в аудіосигналах:

- заміни найменш значущого біта;
- фазового кодування;
- розширення спектру;
- приховування інформації за допомогою вставки тонів;
- приховування інформації використовуючи ехо-сигнал.

Стегосистему вважають **зламаною**, якщо зловмисник зміг довести, що в перехопленому контейнері наявна прихована інформація. Якщо йому це не вдалось, то кажуть, що система **стійка**. Види атак на стегосистему:

- атака, що грунтується на знанні заповненого контейнера;
- атака, що грунтується на відомому повідомленні ;
- атака, що грунтується на прихованих даних чи заповнених контейнерів, що обирались зловмисником;
- атака, що грунтується на відомому пустому контейнері ;
- атака, що грунтується на виборі порожнього контейнера зловмисником;
- атака, що грунтується на знанні математичної моделі контейнера.

Етапи зламування стегосистеми:

- розкриття факту існування прихованої інформації;
- добування прихованих даних;
- перетворення секретних даних;
- заборона поширення скритого повідомлення.

$$IF = 1 - \frac{\sum_{x=1}^{n}\sum_{y=1}^{m}(C_{x,y} - S_{x,y})^2}{\sum_{x=1}^{n}\sum_{y=1}^{m}(C_{x,y})^2};$$

$$NAD = \frac{\sum_{x=1}^{n}\sum_{y=1}^{m}|C_{x,y} - S_{x,y}|}{\sum_{x=1}^{n}\sum_{y=1}^{m}|C_{x,y}|};$$

$$SNR = \frac{\sum_{x=1}^{n}\sum_{y=1}^{m}(C_{x,y})^2}{\sum_{x=1}^{n}\sum_{y=1}^{m}(C_{x,y} - S_{x,y})^2};$$

$$AD = \frac{1}{X \cdot Y}\sum_{x=1}^{n}\sum_{y=1}^{m}|C_{x,y} - S_{x,y}|$$

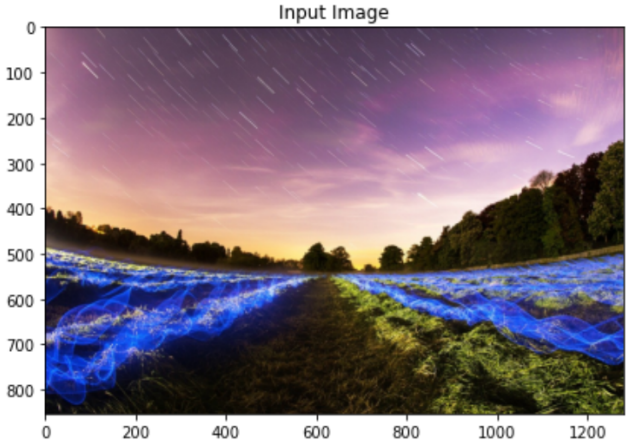
$$MSE = \frac{1}{X \cdot Y}\sum_{x=1}^{n}\sum_{y=1}^{m}(C_{x,y} - S_{x,y})^2$$

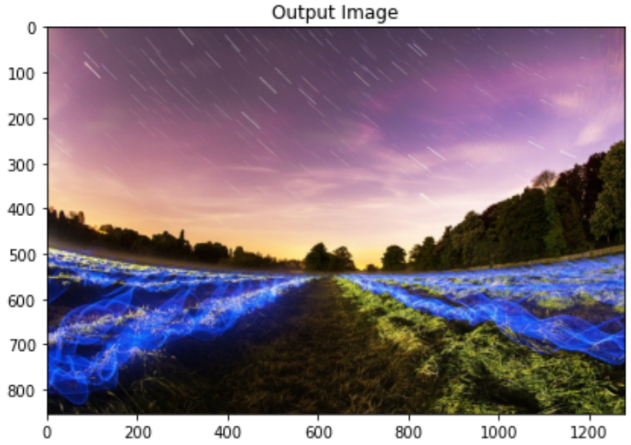
Загальний алгоритм приховування даних в BMP форматі зображення:

- вікриття файлу контейнера для прихованих даних;
- вибір кольору палітри контейнера (червона, синя, зелена);
- відкриття файлу з даними, які треба приховати;
- конвертування даних в двійковий формат для засекречення в масиви по 8 біт (всі характеристики зображення та інтенсивність кольорів пікселів);
- заміна найменш значущого біта пікселя контейнера бітом прихованих даних відповідно до величини ключа;
- зберігання заповненого контейнера, тобто зміненого файлу, у форматі BMP відповідно до правил зберігання.

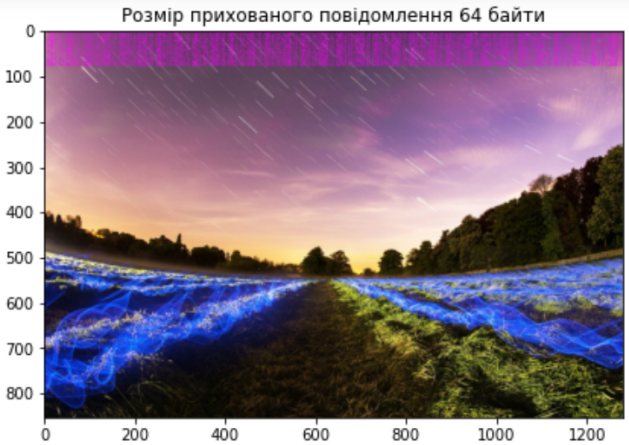
Для видобування даних потрібно:

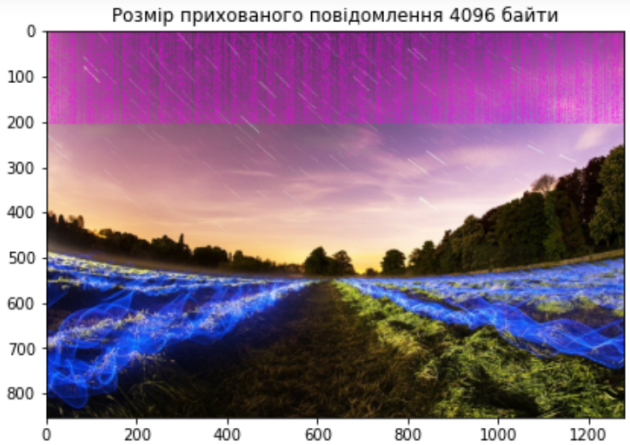
- знати ключ, з яким в контейнері розміщувались дані;
- відкрити файл з секретною інформацію, щоб отримати значення кольорів пікселів повного контейнера;
- використовуючи ключ, вибрати найменш значущі біти кольорів пікселів та згенерувати з них секретні дані.





Розмір повідомлення, байт	MSE	PSNR
64	8.273495506057053e-05	88.95391325528558
128	0.00016699638530676044	85.90373290069535
256	0.00031964390386869873	83.08413934724122
512	0.0006307395466979289	80.13230299219362
1024	0.0012688061742868308	77.09685077517824
4096	0.004987299726455647	71.15214891660129





Метод Коха-Жао можна описати наступним алгоритмом:

Крок 1: Спочатку початкове зображення розділюють на блоки розміром 8×8 пікселів, до яких застосовують дискретно косинусні перетворення, причому в одному блоці приховують один біт даних.

Крок 2: Застосувати дискретно-косинусне перетворення до кожного блоку. В результаті отримаємо набір матриць коефіцієнтів D_i розмірності 8.

Крок 3: Обираємо блоки для вбудовування та записуємо в кожен блок по одному біту даних.

Крок 4: У всіх блоках потрібно вибрати два коефіцієнти перетворення, які симетричні головній діагоналі.

Крок 5: Щоб передати біт «0», необхідно, щоб різниця абсолютних значень коефіцієнтів Д-К перетворень була більша ніж якась порогова величина. Щоб передати біт «1», необхідно, щоб різниця була меншою ніж якась порогова величинаю. Тому, наприклад, для вбудовування нульового біту збільшуть перший, але зменшують другий коефіцієнт на однакову величину та навпаки.
Крок 6: Четвертий та 5 пункт потрібно виконати для всіх блоків
Крок 7: Виконати обернене дискретно косинусне перетворння для всіх блоків.
Для розшифровування інформації:
Крок 4: Потрібно виділити матриці колірних компонент зображення
Крок 5: Виконуємо розбиття матриці на блоки Ic
Крок 6: До кожного блоку застосовуємо дискретно-косинусне перетворення.

- Стеганографія в поєднанні з криптографією, є потужним захистом інформації.
- Метою стеганографічних алгоритмів є забезпечити приховування факту наявності даних, що потребує захисту.
- Надійним алгоритмом захисту даних в зображеннях є алгорит маскування, що використовують видимі аспекти зображення, ніж, скажімо, модифікація модетів заміни найменш значущого біта щодо різних видів обробки зображень.
- Найбільш надійним для захисту даних в текстових форматах є метод знаків одного накреслення.
- Для аудіо-форматів найефективнішим вважають метод заміни найменш значущого біта.
- Метод ЗНЗБ для приховування даних в BMP форматі є простим та може приховувати великий об'єм даних.
- Збільшуючи розмір вбудованих даних зменшується якість зображення.
- Метод Коха-Жао є стійким до більшості відомих стегоатак, але, має низьку пропускну здатність

Дякую за увагу!

https://github.com/vitalikkk19/Coursework