УДК 004.056 (045)

ВИЯВЛЕННЯ ПРИХОВАНИХ КАНАЛІВ ПЕРЕДАЧІ ІНФОРМАЦІЇ НА БАЗІ МЕТОДІВ СТЕГАНОАНАЛІЗУ

О. К. Юдін, д-р техн. наук, проф., ***Я. А. Симониченко***

Національний авіаційний університет

e-mail: yaroslavsim@ukr.net

У статті запропоновано та описано метод виявлення наявності прихованих каналів передачі інформації, утворених з використанням стеганографічної системи, на базі методів стеганографічного аналізу. Даний метод базується на виявленні наявності прихованого текстового повідомлення, вбудованого в цифрове зображення методом модифікації молодшого біту колірної компоненти моделі RGB, шляхом порівняння розподілу кількості 1-х бітів в умовних блоках бітових площин колірних компонентів цифрового зображення та кодів символів текстового повідомлення при їх двійковому представленні. Наведено результати дослідження описаного методу виявлення та перевірки наявності прихованого текстового повідомлення в цифровому зображенні з використанням утворених матриць умовних блоків колірних компонентів, а також, можливого апріорного визначення мови, що використовується в прихованому повідомленні.

Ключові слова: стеганографічна система, стеганоконтейнер, стеганографічний аналіз.

The article proposes and describes a method of identifying secure communication channels which have been formed with the use of a steganographic system on the basis of methods of the steganographic analysis. This method is based on identification hidden message built in the digital image using modifications of the least significant bit of the color component of the RGB model by comparing the distribution of the number of 1 bits in the conditional blocks of bit planes of the color components of the digital image and the character codes of text message when they submitted in binary representation. Listed the results of the research described method to detect and validate the presence of a hidden text message in digital image using the formed matrixes of conditional blocks of the color components, and also a priori language identification used in the hidden text message.

Keywords: steganographic system, steganographic container, stegoanalysis.

Вступ

Широке використання інформаційних технологій (використання автоматизованих систем, систем електронного документообігу, електронних платежів та ін.) в діяльності держави та житті сучасної людини, призвело до необхідності забезпечення інформаційної безпеки та захисту інформації в інформаційному просторі [1]. Однією з сучасних технологій, що використовується для вирішення даного питання є стеганографія. Велика кількість стеганографічних засобів, які дозволяють приховувати інформацію та факт її подальшого передавання каналами зв'язку, зробили зазначену вище технологію поширеним методом захисту інформації.

Як наслідок — реалізація зазначених вище методів приховування інформації може виконуватися з метою організації захисту інформації або з метою організації прихованого витоку цінної інформації, з використання діючих інформаційних систем, як на рівні державних установ та підприємств, так і громадянина. На сьогодні, розвиток методів стеганографічного аналізу є актуальним завданням. Реалізація стеганографічного аналізу дає можливість здійснювати дослідження інформаційного об'єкта з метою виявлення та встановлення факту наявності прихованої інформації.

Постановка завдання

Реалізація методів стеганографічного захисту призводить до створення спеціальних стеганографічних систем. Під стеганографічною системою слід розуміти об'єднання методів і засобів, які використовуються для створення прихованого каналу передачі інформації [2]. Стеганографічна система виконує вбудовування контейнера із повідомленням в інформаційний об'єкт, його передавання стеганографічним каналом та декодування приховуваного повідомлення. Найчастіше, з огляду на функціонал сучасного стегано-

графіного програмного забезпечення, прихованою інформацією є текстове повідомлення. Найчастіше як контейнер використовують цифрове зображення, у якому виконується вбудовування прихованого текстового повідомлення одним із стеганографічних методів.

Передавання цифрового зображення із прихованою інформацією може відбуватися, як мережею інформаційно-телекомунікаційної системи, так і людиною на носії інформації (карти пам'яті, флеш-пам'ять та ін.) у вигляді графічного файлу.

Для виявлення зазначеного вище каналу передачі інформації та факту наявності прихованого текстового повідомлення можуть використовуватися методи стеганографічного аналізу. Також, даний аналіз дає змогу виявити джерела повідомлення для подальшого його контролю в разі його протиправних дій або витоку інформації, що може завдати збитків. Наприклад, джерелом може бути адреса електронної пошти, з якої виконується передача зображення, або обліковий запис соціальної мережі, від імені якого були викладені цифрові зображення із вбудованою прихованою інформацією.

Таким чином, **метою даної статті** є реалізація запропонованого методу виявлення та перевірки наявності прихованих каналів передачі інформації з використанням цифрового зображення шляхом порівняння розподілу кількості 1-х бітів в умовних блоках бітових площин колірних компонентів зображення та кодів символів текстового повідомлення при їх двійковому представленні. На основі проведених досліджень буде визначено метод виявлення прихованого текстового повідомлення в цифровому зображенні з використанням утворених матриць умовних блоків колірних компонентів, а також, можливого апріорного визначення мови, що використовується в прихованому повідомленні.

Розв'язання проблеми

З огляду на те, що як приховану інформацію використовують текстове повідомлення, було проведено аналіз розподілу кількості 1-х бітів (далі — РКБ) для кодів символів тексту англійської та російської мов.

Для представлення та зберігання інформації в пам'яті комп'ютера використовується двійковий спосіб кодування, що передбачає використання лише двох можливих значення бітів — «0» або «1».

Людина розрізняє символи за їх виглядом, а комп'ютер — за їх кодом. Кожен символ представляється 8-розрядним двійковим кодом (8 бітів). Кодування полягає в тому, що кожному символу відповідає унікальний двійковий код від «0000 0000» до «1111 1111» або відповідний унікальний десятковий код від «0» до «255».

Міжнародним стандартом для кодування текстових символів та інформаційного обміну між користувачами комп'ютерів є таблиця ASCII.

Розглянемо використання таблиці ASCII на прикладі кодування текстового повідомлення англійською (латинськими символами таблиці) та російською мовами (символами російського алфавіту).

Як текст російською мовою оберемо повідомлення – «Национальный авиационный университет» та «National Aviation University» — для повідомлення англійською мовою.

Виконаємо відображення кодування символів перших слів обох повідомлень (табл. 1–2).

Таким чином, було отримано представлення частини повідомлення з використанням таблиці ASCII та їх значень у двійковому вигляді.

Виконаємо дослідження РКБ для кодів символів зазначених текстових повідомлень при їх двійковому представленні (табл. 3).

Таблиця 1

Кодування символів першого слова повідомлення російською мовою

Символ	ASCII код	Двійкове представлення ASCII коду
Н	205	1100 1101
а	224	1110 0000
ц	246	1111 0110
и	232	1110 1000
о	238	1110 1110
н	237	1110 1101
а	224	1110 0000
л	235	1110 1011
ь	252	1111 1100
н	237	1110 1101
ы	251	1111 1011
й	233	1110 1001

Таблиця 2

Кодування символів першого слова повідомлення англійською мовою

Символ	ASCII код	Двійкове представлення ASCII коду
N	78	0100 1110
a	97	0110 0001
t	116	0111 0100
i	105	0110 1001
o	111	0110 1111
n	110	0110 1110
a	97	0110 0001
l	108	0110 1100

Таблиця 3

Підрахунок бітів із значенням «1»

Символ	ASCII код	Двійкове представлення ASCII коду							
Н	205	1	1	0	0	1	1	0	1
а	224	1	1	1	0	0	0	0	0
ц	246	1	1	1	1	0	1	1	0
и	232	1	1	1	0	1	0	0	0
о	238	1	1	1	0	1	1	1	0
н	237	1	1	1	0	1	1	0	1
а	224	1	1	1	0	0	0	0	0
л	235	1	1	1	0	1	0	1	1
ь	252	1	1	1	1	1	1	0	0
н	237	1	1	1	0	1	1	0	1
ы	251	1	1	1	1	1	0	1	1
й	233	1	1	1	0	1	0	0	1
Номер бітового блоку		**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
Кількість 1-х бітів в блоці		12	12	11	3	9	6	4	6

Для реалізації даного дослідження виконаємо такі дії:

– виконаємо представлення кожного символу частини повідомлення російської мови «Национальный» у відповідний двійковий код (по 8 бітів);

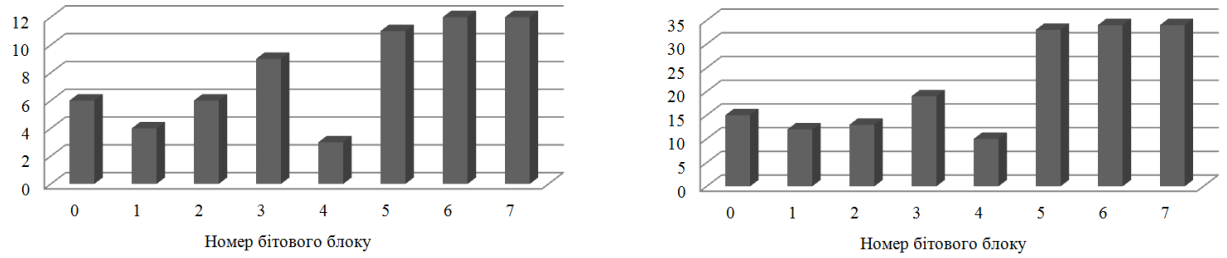
– виконаємо розбиття утвореного бітового представлення на блоки, що відповідають бітовим розрядом кожного двійкового представлення (від 0 до 7);

– виконаємо підрахунок кількості бітів із значенням «1» у кожному бітовому блоці.

Виконаємо графічне представлення РКБ для кодів символів зазначеної першої частини текстового повідомлення на російській мові, при її двійковому представленні (рис. 1, *а*), та повного текстового повідомлення (рис. 1, *б*).

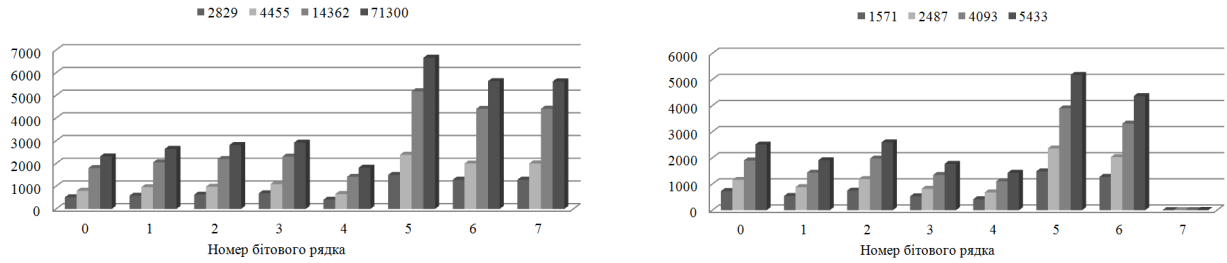
Для більш детальнішого дослідження було обрано чотири повідомлення російською мовою довжиною 2829, 4455, 14362 та 71300 символів та виконано представлення їх РКБ (рис. 2, *а*).

Також, було обрано чотири повідомлення англійською мовою довжиною 1571, 2487, 4093 та 5433 символів та виконано дослідження вищезазначеного розподілу (рис. 2, *б*).



а *б*

Рис. 1. Представлення РКБ першої частини повідомлення (*а*) та повного (*б*) повідомлення, де 0...7 — номер бітового блоку



а *б*

Рис. 2. Представлення РКБ повідомлень російською (*а*) та англійською (*б*) мовою, де 0...7 — номер бітового блоку

Після аналізу результатів досліджень РКБ для кодів символів обраних текстових повідомлень при їх двійковому представленні можна зробити висновок про наявність збереження певного співвідношення кількості 1-х бітів у бітових блоках кожного розряду. Зокрема наявність характерних максимальних та мінімальних кількостей 1-х бітів у бітових блоках, що відповідають розрядам двійкових кодів відповідних символів. При аналізі рис. 5, *а* можна зробити висновок, що характерною особливістю для повідомлення російського тексту є зменшення кількості 1-х бітів у такій послідовності стовпців: 5, 6, 7, 3, 2, 1, 0 та 4. При аналізі рис. 5, *б* можна побачити, що характерною особливістю для повідомлення англійського тексту є зменшення кількості 1-х бітів у такій послідовності: 5, 6, 2, 0, 1, 3, 4 та 7.

Таким чином, при аналізі кожних 4-х повідомлень відповідних текстів дана властивість зберігається.

Для дослідження методу виявлення наявності прихованого текстового повідомлення було використано 24-х бітове растрове зображення. Збереження зображення відбувалося у BMP-форматі, оскільки він є оптимальнішим форматом при виконанні стеганоперетворення [3]. Вбудовування в зображення виконувалось методом модифікації молодшого біту в компоненті синього кольору при використанні колірної моделі RGB. Для кодування градацій кольору кожної компоненти моделі RGB використовується 8 бітів (загалом 24 біти для кодування 3-х кольорів компонент).

Для дослідження було обрано зображення розміром — 568×500 пікселів (рис. 3, *а*). Ступінь модифікації контейнера при вбудовуванні текстового повідомлення англійської мови до компоненти синього кольору становив 8, 16, 24 та 32 %.

Виконаємо видобування кожної колірної компоненти заповненого зображення колірної моделі RGB (рис. 3, *б*, 3, *в* та 3, *г*) та виконаємо дослідження бітових площин компоненти синього кольору зображення із 8 % заповненням на наявність прихованого повідомлення. Для цього виконаємо наступні дії (рис. 4):

– виконаємо перетворення матриці компоненти синього кольору зображення в вектор-стовпець, утвореного із значень градації кольору синьої компоненти, та переведемо значення вектора у двійковий код, за аналогією методу табл. 3 (отримаємо матрицю двійкового представлення, кожний стовпець якої відповідає бітовій площині);

– виконаємо розбиття кожного стовпчика бітової площини на вісім умовних блоків, що утворені шляхом підрахунку бітів із значенням «1» наступним чином: кожний умовний блок (0…7), розміром 1×8, буде містити значення кількості 1-х бітів відповідного стовпчика матриці двійкового представлення (0…7) із кроком зсуву «+8»: для 0-го блоку 0-ї бітової площини — елементи з індексами рядків 0, 8, 16, 24…; для 1-го блоку 0-ї бітової площини — елементи з індексами рядків 1, 9, 17, 25…; для 2-го блоку 0-ї бітової площини — елементи з індексами рядків 2, 10, 18, 26… і т. д.;

– отримаємо матрицю умовних блоків (далі — МУБ), розміром 8×8, стовпці якої будуть відповідати номеру бітової площини компоненти зображення, а рядки — номеру умовного блоку із підрахунком розподілу кількості 1-х бітів у кожному блоці.

а *б* *в* *г*

Рис. 3. Зображення для приховування (*а*) повідомлення та його колірні компоненти (*б* — *RED*; *в* — *GREEN*; *г* — *BLUE*)

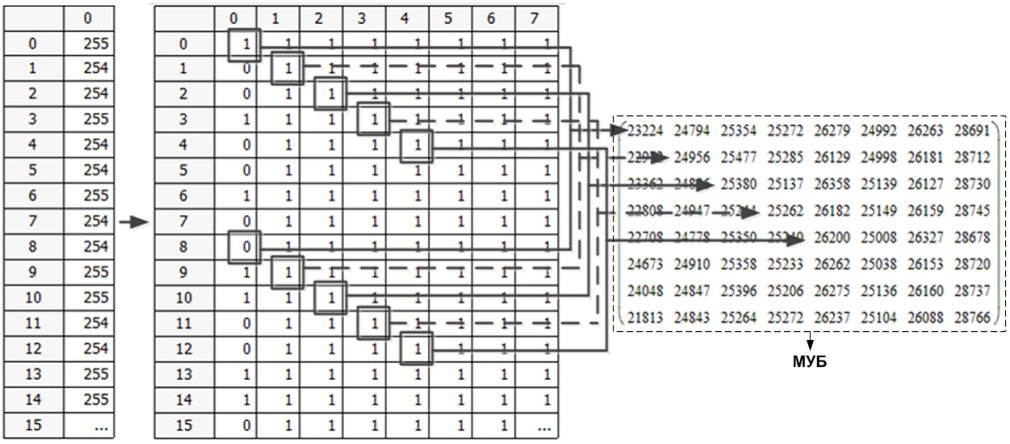


Рис. 4. Утворення матриці умовних блоків РКБ в бітових площинах синьої компоненти

Виконаємо графічне представлення значень утвореної матриці умовних блоків (рис. 5).

Як можна побачити із графічного представлення утворених матриць, наявні зміни у бітовій площині молодшого біту зображення із 8 % заповненням, оскільки значення РКБ у даних умовних блоках змінюється, приблизно, в межах від 21000 до 24000. Значення інших умовних блоків бітових площин з індексами від 1 до 7 відносно рівномірне: для 1-го умовного блоку в

межах 24000; для 2-го блоку — 25000; для 3-го блоку — 25000 і т.д. Розмах варіації, що визначається за формулою: $R = X_{max} - X_{min}$, де X_{min} та X_{max} — відповідні мінімальне та максимальне значення РКБ в умовних блоках бітової площини молодшого біту зображення становить 2860. Для 1-го стовпця МУБ — 178, для 2-го стовпця МУБ — 233, для 3-го стовпця МУБ — 148 і т. д. Можна зробити висновок, що підвищене значення розмаху варіації для значень РБК в умовних блоках бітової площини молодшого біту синьої компоненти зображення свідчить про те, що дана бітова площина може містити приховане повідомлення. Для оцінювання та порівняння відповідності РКБ в бітових площинах зображення та кодів символів текстового повідомлення будемо виконувати їх порівняння на основі показника коефіцієнта кореляції Пірсона, що визначається за такою формулою:

$$r_{xy} = \frac{\frac{1}{n} \times \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{S_x^2} \times \sqrt{S_y^2}},$$

де $\bar{x}, \bar{y}$ — середні значення вибірки x та y; S — середньоквадратичне відхилення.

Даний коефіцієнт кореляції вимірюється в межах від –1,00 до +1,00. Значення коефіцієнта кореляції «–1,00» буде стверджувати про відсутність кореляції між величинами, «0» — про нульову кореляцію, а «+1,00» — про повну кореляцію величин. Тобто, чим ближче значення коефіцієнта кореляції до «+1,00», тим сильніший зв'язок між двома досліджуваними величинами.

Виконаємо графічне відображення значень аналогічно утворених МУБ описаним вище методом для зображень із 16, 24 та 32 % заповненням (рис. 6–8). Графічне представлення утворених МУБ бітових площин синіх компонент зображень із заповненням 16, 24 та 32 %, також дає можливість виявлення наявності змін у бітових площинах молодших бітів обраних колірних компонентів зображень. Значення РКБ в умовних блоках бітової площини молодшого біту синьої компоненти зображення із 16 % заповненням змінюється, приблизно, в межах від 19000 до 24000.

У зображенні із 24 % заповненням — від 17000 до 25000. У зображенні із 32 % заповненням — від 15000 до 26000. Отже, у разі підвищення заповнення синьої компоненти зображення, виконується підвищення показника розмаху варіації для значень РБК у МУБ бітових площин зображення. Це свідчить про можливий факт використання стеганографічного перетворення в даних колірних компонентах цифрового зображення.

Попередньо, можна зробити висновок про наявність у синій компоненті даного зображення прихованого текстового повідомлення із символами англійської мови. Виконаємо порівняння РКБ у бітових площинах синьої компоненти зображення із 8, 16, 24 та 32 % заповненням та кодів символів англійського текстового повідомлення при їх двійковому представленні на основі показника коефіцієнта кореляції Пірсона (табл. 4).

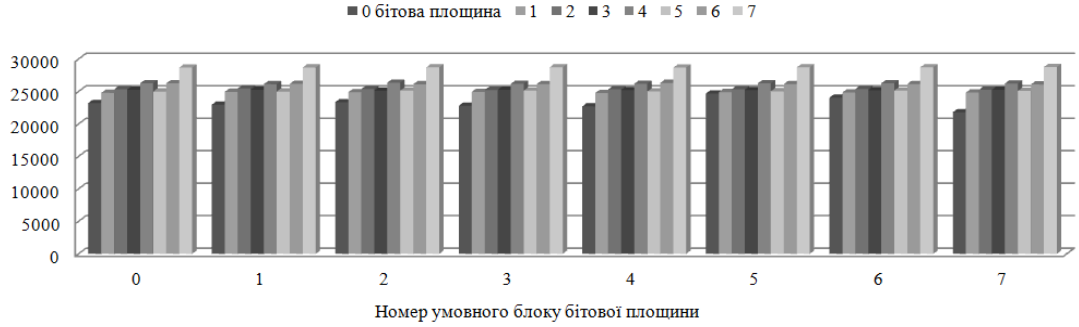


Рис. 5. Відображення значень утвореної МУБ для зображення із 8 % заповненням, де 0...7 — номер умовного блоку відповідної бітової площини

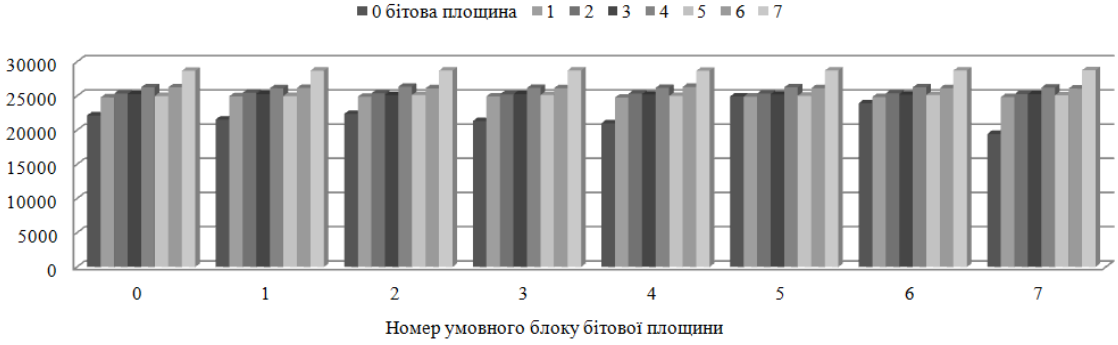


Рис. 6. Відображення значень утвореної МУБ для зображення із 16 % заповненням

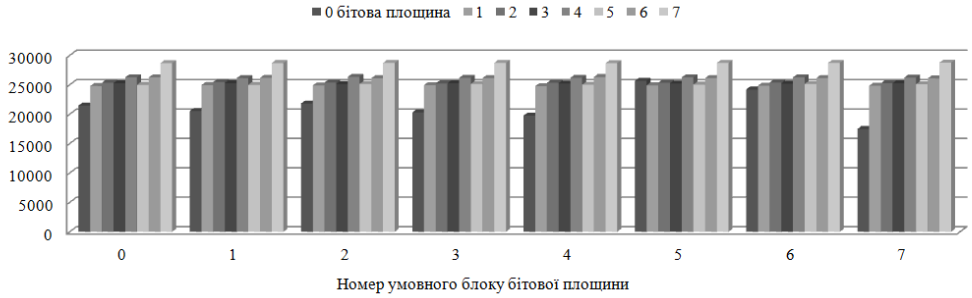


Рис. 7. Відображення значень утвореної МУБ для зображення із 24 % заповненням

Рис. 8. Відображення значень утвореної МУБ для зображення із 32% заповненням

Таблиця 4

Значення коефіцієнта кореляції РКБ зображення та текстового повідомлення

Ступінь заповнення компоненти, %	Номер бітової площини синьої компоненти зображення							
	0	1	2	3	4	5	6	7
8	0,995235	0,1625518	0,3881639	–0,3856865	0,3643116	0,0122231	–0,029609	–0,1530243
16	0,9986903	0,1625518	0,3881639	–0,3856865	0,3643116	0,0122231	–0,029609	–0,1530243
24	0,9991113	0,1625518	0,3881639	–0,3856865	0,3643116	0,0122231	–0,029609	–0,1530243
32	0,9993043	0,1625518	0,3881639	–0,3856865	0,3643116	0,0122231	–0,029609	–0,1530243

Після визначення коефіцієнта кореляції РКБ МУБ для кожної бітової площини синьої компоненти зображення можна побачити, що нульова бітова площина має дуже високу кореляцію при її порівнянні з РКБ кодами символів англійського текстового повідомлення. При підвищенні ступеню заповнення стеганоконтейнера (8–32 %) виконується підвищення показника коефіцієнта кореляції (0,995235–0,999304), що дає змогу впевнитися в тому, що в бітовій площині молодшого біта синьої компоненти досліджуваних зображень наявне приховане текстове повідомлення із символами англійської мови. Показники коефіцієнтів кореляції інших РКБ бітових площин синьої компоненти мають дуже слабку кореляцію, що свідчить про відсутність прихованих повідомлень у даних бітових площинах.

Таким чином, була виявлена наявність прихованого повідомлення в бітових площинах молодших бітів синіх компонент наявних зображень із 8, 16, 24 та 32% заповненням, що свідчить про можливу наявність прихованого каналу передачі інформації.

Висновок

Запропонований метод виявлення наявності прихованих каналів передачі інформації, утворених з використанням стеганографічної системи, дає можливість дослідження цифрового зображення на наявність прихованого текстового повідомлення, вбудованого методом модифікації молодшого біту колірної компоненти моделі RGB, та визначення мови прихованого повідомлення. Виявлення наявності прихованої інформації в зображенні виконується шляхом порівняння РКБ МУБ бітових площин колірних компонентів зображення та кодів символів повідомлення при їх двійковому представленні.

ЛІТЕРАТУРА

1. *Юдін О. К.* Інформаційна безпека держави / О. К. Юдін, В. М. Богуш. — К. : Консум, 2005. — 576 с.

2. *Юдін О. К.* Удосконалення стеганографічних методів на базі аналізу колірних моделей зображення / О. К. Юдін, Я. А. Симониченко // Наукоємні технології. — 2012. — №1 (13). — С. 70-75.

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

Стаття надійшла до редакції 07.11.2016