СИСТЕМИ ТА МЕТОДИ ОБРОБКИ ІНФОРМАЦІЇ

УДК 004.056.5

І.І. Борисенко

ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ СТЕГАНОГРАФІЧНОГО АЛГОРИТМУ НА ОСНОВІ ТЕОРІЇ ГРАФІВ

У роботі пропонується нова версія стеганографічного алгоритму, який використовує методи оптимізації для забезпечення більшої в порівнянні з базовим алгоритмом пропускної спроможності каналу прихованого зв'язку разом зі збереженням рівня збурень контейнера. Надійність сприйняття забезпечується за рахунок мінімізації впливів вбудованого повідомлення на контейнер-зображення.

Ключові слова: *стеганографічний алгоритм, повідомлення, контейнер, збурення контейнера.*

В работе предлагается новая версия стеганографического алгоритма, который использует методы оптимизации для обеспечения большей по сравнению с базовым алгоритмом пропускной способности канала скрытой связи наряду с сохранением уровня возмущения контейнера. Надежность восприятия стегоконтейнера обеспечивается за счет минимизации влияний встроенного сообщения на контейнер-изображение.

Ключевые слова: *стеганографический алгоритм, сообщение, контейнер, возмущение контейнера.*

A new version of the steganographic algorithm by using the optimization methods that provides greater bandwidth of covert communications than basic algorithm along with saving level of perturbation of the container is suggested. Reliability of the perception of stegocontainer is provided by minimizing the effects of embedded messages on the container image.

Keywords : *steganographic algorithm, message, container, embedding impact.*

Вступ. У наш час одним із ефективних засобів захисту конфіденційної інформації (далі – КІ) є цифрова стеганографія, що забезпечує приховування КІ в цифрових даних (далі – ЦД), які потім відкрито пересилаються одержувачу каналами загального користування [1; 2]. ЦД, в які вбудовується КІ за допомогою стеганографічного алгоритму, узагальнено називаються – контейнер, КІ – повідомлення, а результатом такого вбудовування є стеганоконтейнер. Будь-який стеганографічний алгоритм характеризується трьома основними властивостями: стійкістю (до стеганоаналізу або до завад), надійністю сприйняття (стеганоконтейнер візуально не повинен відрізнятися від контейнера), прихованою пропускною здатністю. Під прихованою пропускною здатністю (далі – ППЗ) розуміють максимальну кількість інформації, яка може бути вкладена в один елемент контейнера [2].

Присутність повідомлення в контейнері повинна бути таємною, тому корегування елементів контейнера повинно бути щонайменшим.

Останнім часом активно ведуться роботи по створенню стеганографічних методів та алгоритмів, розробники яких намагаються забезпечити найменш можливий вплив на контейнер як за рахунок вибору елементів контейнера для вбудовування, так і специфіки самого алгоритму [3–7].

У нашій статті "Застосування методів порівняння послідовностей в стеганографічних перетвореннях цифрових зображень" був запропонований алгоритм організації таємної передачі повідомлення, заснований на знаходженні схожих бітових послідовностей в повідомленні та контейнері, який завдяки малим збуренням контейнера дає хороші показники щодо збереження статистик контейнера після вбудовування. Але розроблений у зазначеній роботі алгоритм, який є привабливим з точки зору його стійкості до статистичного стеганоаналізу, має низьку ППЗ. Повідомленням може бути будь-яка конфіденційна інформація і, якщо це, скажімо, особисті чи медичні дані, то стеганографічний алгоритм, який вбудовує таку інформацію у контейнер, повинен забезпечувати значну ППЗ.

Багато найрізноманітніших задач формулюються в термінах теорії графів: аналіз мереж в електротехніці, в програмуванні, проектуванні електронних схем, в економіці тощо. Аналіз показав, що алгоритми, завдяки яким відшукується місце в контейнері для вбудовування повідомлення, можна оптимізувати за допомогою теорії графів.

Метою роботи є модифікація стеганографічного алгоритму [8] для підвищення його ППЗ разом із забезпеченням малих збурень контейнера за рахунок використання інструментів теорії графів.

Основна частина. В основі алгоритму, що пропонується в цій роботі, лежить побудова графової моделі взаємозв'язків між блоками повідомлення, яке вбудовується, та блоками контейнера, в які буде вбудовано повідомлення, тому наведемо декілька понять з теорії графів [9; 10].

Граф G становить структуру (V,E), де V – множина його вершин, а $E \subseteq V \times V$ – множина ребер. Неорієнтований граф – це граф, усі ребра якого не мають орієнтації, тобто $(x,y)\in E$ і $(y,x)\in E$ – це одне і те саме ребро. Дві вершини суміжні, якщо вони з'єднані ребром, тобто, якщо $(y,x)\in E$ – це ребро, то вершини y та x суміжні. Два ребра називаються суміжними, якщо вони інциндентні одній і тій самій вершині, тобто такі ребра мають спільну вершину. Будь-який граф однозначно визначається матрицею суміжності. Матриця суміжності $A[a_{ij}]$ з елементами a_{ij} – це матриця виду: $a_{ij} = \begin{cases} 1, \textit{ якщо існує ребро } (x_i, x_j) \\ 0, \textit{ якщо вершини } x_i, x_j \textit{ не суміжні} \end{cases}$. Якщо граф зважений, тобто вага ребра більше одиниці, то будується матриця ваги $W[w_{ij}]$ графа, яка відрізняється від $A[a_{ij}]$ лише тим, що якщо існує ребро (x_i, x_j), то w_{ij} – це його вага.

Граф H називається дводольним, якщо множину його вершин V можна розбити на дві підмножини V_1 і V_2 таким чином, що існують тільки ребра, які з'єднують вершини з різних підмножин, і не існує жодного ребра, яке з'єднувало б вершини однієї й тієї самої підмножини V_i. Граф є повним дводольним, якщо будь-яка вершина $v_i \in V_1$ суміжна з усіма вершинами V_2.

Паросполученням P (або незалежною множиною ребер) в дводольному графі H називають множину ребер, у якій жодні два ребра не суміжні.

Як контейнер будемо використовувати цифрове зображення, яке представляється бінарною матрицею M, біти якої послідовно групуються в підмножини M_i довжиною m кожна. Біти повідомлення, яке позначимо літерою N, послідовно групуються в підмножини N_i довжиною n, де $m>n$. Вбудовування повідомлення N у контейнер M у базовому алгоритмі Seek-Place [8] проводилося послідовно, тобто біти поточного N_i порівнювалися з бітами відповідного блоку M_i і таким чином відшукувалося входження N_i в M_i. Входжень N_i в M_i з точним збігом (тобто з якоїсь позиції l усі біти N_i збігаються з бітами M_i) дуже мало, тому, як правило, більшість блоків M_i корегувались. Між ППЗ та рівнем збурень, яких зазнає контейнер після вбудовування повідомлення, завжди повинен існувати розумний компроміс, інакше повідомлення буде легко виявлено стеганоаналітичними методами, тому було введено параметр d, значення якого відповідало кількості пікселів, які корегувалися, тобто це кількість неспівпадінь N_i з бітами M_i, які ми дозволяли допустити. Якщо ж співпадіння з точністю до d не було виявлено при всіх можливих положеннях N_i відносно M_i, то такий M_i не використовувався і N_i вбудовувався в наступний блок M_{i+1}, якщо він відповідав умові d. Якщо в сукупності виявлялося, що невикористаних блоків контейнера достатньо багато, то корегувалися параметри m та n (збільшувалося m або (і) зменшувалося n), що зменшувало об'єм інформації, яка передавалася.

У новій версії алгоритму, назвемо його Seek-Place-optim, пропонується шукати входження поточного N_i у всіх блоках контейнера. У результаті створюється зважений повний дводольний граф H, в якому вершинами V_1 є блоки N_i, а вершинами V_2 – блоки M_i. Вага будь-якого ребра (N_i, M_j) – це максимальна кількість k бітів, що співпали при входженні N_i з деякої позиції l в M_j. Далі на одержаному графі ставиться оптимізаційна задача знаходження паросполучення максимальної ваги, тобто задача оптимального розподілу блоків повідомлення N_i по блокам контейнера M_j.

Для вирішення поставленої задачі будується матриця ваги $W[w_{ij}]$ графа H, і використовується алгоритм пошуку оптимальної перестановки π [11] на $W[w_{ij}]$. Як приклад, наведемо фрагмент $W[w_{ij}]$, який має вигляд, зображений на рис. 1.

N_i	M_i					
	1	2	3	4	5	6
1	**2**	**6**	**4**	**7**	**1**	**2**
2	**8**	**7**	**2**	**4**	**3**	**5**
3	**1**	**5**	**3**	**4**	**3**	**3**
4	**8**	**7**	**1**	**2**	**0**	**1**
5	**9**	**11**	**1**	**3**	**3**	**6**
6	**12**	**9**	**9**	**4**	**5**	**7**

Рис. 1. Фрагмент матриці ваги графа

Використовуючи алгоритм пошуку оптимальної перестановки, знаходимо, що для наведеної матриці існує три оптимальні перестановки: $\pi_1 = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 4\,6\,5\,2\,1\,3 \end{pmatrix}$, $\pi_2 = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 4\,2\,5\,1\,6\,3 \end{pmatrix}$, $\pi_3 = \begin{pmatrix} 1\,2\,3\,4\,5\,6 \\ 4\,1\,5\,2\,6\,3 \end{pmatrix}$. Таким чином, відповідно маємо три оптимальні

паросполучення графа H, а саме: P_1= {(N_1, M_4), (N_2, M_6), (N_3, M_5), (N_4, M_2), (N_5, M_1), (N_6, M_3)}, P_2={(N_1, M_4), (N_2, M_2), (N_3, M_5), (N_4, M_1), (N_5, M_6)}, (N_6, M_3), P_3={(N_1, M_4), (N_2, M_1), (N_3, M_5), (N_4, M_2), (N_5, M_6), (N_6, M_3)}. Цінність призначення будь-якого з P_i становить: $w_{14}+w_{26}+w_{35}+w_{42}+w_{51}+w_{63}=40$. Отже, якщо вважати, що n=12, тобто довжина повідомлення становить 72, то з 72 бітів контейнера потрібно корегувати лише 32. Якщо ж блоки повідомлення вбудовувати за алгоритмом Seek-Place, тобто послідовно, то цінність призначення становила б $w_{11}+w_{22}+w_{33}+w_{44}+w_{55}+w_{66}=24$ і в цьому випадку корегувалося б уже 48 бітів.

Для вбудовування може бути використано будь-яке з одержаних призначень P_i.

Алгоритм створення стеганографічного контейнера

Вхід: матриця цифрового зображення – контейнер (M) та повідомлення (N) у бінарному представленні.

Вихід: Стеганографічний контейнер

1. Розбити контейнер на підмножини M_i довжини m. Розбити повідомлення на підмножини N_i довжини n.

2. Для кожного N_i та M_j виконувати підпрограму Seek [8], для знаходження максимальної кількості k бітів, що співпали при входженні N_i з деякої позиції l в M_j, одночасно створюючи матрицю ваги $W[w_{ij}]$, де $w_{ij}=k$ і матрицю $Pos[p_{ij}]$, де p_{ij}=l.

3. Використовуючи алгоритм пошуку оптимальної перестановки, знайти паросполучення {(N_i, M_j)}.

4. Відповідно до знайденого паросполучення на кроці 3 вбудувати N_i в M_j, одночасно створюючи елемент ключа K_i, який містить номер блоку контейнера j та номер позиції l в блоці, з якої починається вбудована інформація.

Алгоритм декодування повідомлення

Вхід: Матриця стеганографічного контейнера в бінарному представленні (S).

Ключ (K).

Вихід: Повідомлення.

1. Розбити S на підмножини S_i довжини m.

2. Для кожного K_i звернутися до блоку S_j, починаючи з позиції l виписати послідовність його елементів довжини n.

Оцінка властивостей побудованого алгоритму

Для порівняння ППЗ двох алгоритмів базового та його нової версії використовувалося середнє значення довжини повідомлення, яке одержувалося таким чином. У контейнер-зображення вбудовувалися різні бінарні послідовності. У кожному конкретному випадку визначалася довжина повідомлення. Одержані результати усереднювалися для всіх контейнерів, які тестувалися. Так, наприклад, при m=48 і n=8 середня довжина повідомлення для базового алгоритму становила 24 800 бітів, а для його нової версії 25 600.

Для оцінки збурень контейнера, які були викликані вбудованим повідомленням, та якісного порівняння алгоритмів Seek-Place та Seek-Place-optim у роботі було використано метод дослідження збурень, яких зазнали сингулярні

числа (СНЧ) та сингулярні вектори (СНЧ) матриці контейнера [12]. Як і раніше матриця контейнера позначається літерою M, а стеганоконтейнери, одержані алгоритмами Seek-Place та Seek-Place-optim – *S_Seek та S_Seek_opt* відповідно. Основні кроки дослідження полягали в тому, що для матриць M, *S_Seek та S_Seek_opt*, було побудовано нормальний сингулярний розклад [13]: $M=USV^T$, $S\_Seek = U_{Seek} S_{Seek} V_{Seek}^T$, $S\_Seek\_opt = U_{Seek\_opt} S_{Seek\_opt} V_{Seek\_opt}^T$; знайдено збурення матриць СНЧ $\Delta S_{Seek} = S - S_{Seek}$, $\Delta S_{Seek\_opt} = S - S_{Seek\_opt}$ та СНВ $\Delta U_{Seek} = U - U_{Seek}$, $\Delta U_{Seek\_opt} = U - U_{Seek\_opt}$; оцінені значення $\delta_{Seek} = \max_i \left|(\Delta S_{Seek})_{ii}\right|$ та $\delta_{Seek\_opt} = \max_i \left|(\Delta S_{Seek\_opt})_{ii}\right|$, де $(\Delta S_{[\bullet]})_{ii}$ – діагональні елементи матриць $\Delta S_{[\bullet]}$; Обчислювальний експеримент проводився в середовищі MatLab. У контейнерах виділялися блоки розміром m×m, в які вбудовувалась однакова кількість бітів повідомлення алгоритмами Seek-Place та Seek-Place-optim, потім досліджувалися параметри одержаних стеганоконтейнерів. Наведемо результати одного з дослідів для m=48 та пропускною здатністю блока 0,17 біт/піксель: $\delta_{Seek} = 0{,}4213$, $\delta_{Seek\_opt} = 0{,}2129$; норми перших чотирьох сингулярних векторів U_{Seek} – 8.0490e-05 0,0082 0,0236 0,0592 для U_{Seek_opt} – 6.0491e-06 0,0002 0,0036 0,0292 для інших СНВ, які відповідають СНЧ починаючи з п'ятого для матриць U_{Seek} та U_{Seek_opt} значення норм зрівнянні і майже однакові. Таким чином, порівнюючи збурення матриць СНЧ та СНВ, внесені алгоритмами Seek-Place та Seek-Place-optim, можна зробити висновок, що збурення контейнера – зображення менші при використанні Seek-Place-optim, але для повідомлення з однаковою кількістю бітів. Завдяки вимогам параметра d, призначення якого було згадано вище, не всі блоки контейнера можуть використовуватися для вбудовування алгоритмом Seek-Place, тому об'єм вбудованого повідомлення може бути меншим ніж об'єм повідомлення вбудований Seek-Place-optim. Тобто у цьому разі рівень збурень, привнесений у контейнер алгоритмом Seek-Place-optim, повинен бути вищим, але, як показали дослідження, за рахунок оптимального розподілу блоків повідомлення по блокам контейнера збурення СНЧ та СНВ зрівнянні.

Висновки. У роботі розглянуто алгоритм Seek-Place-optim, який є покращеною версією алгоритму Seek-Place, представленого в роботі [8]. У порівнянні з Seek-Place алгоритм Seek-Place-optim має більшу ППЗ і при цьому зберігається майже однаковий рівень збурень за рахунок оптимізації процесу вбудовування повідомлення. Оптимізація була досягнута завдяки знаходженню оптимального паросполучення блок повідомлення → блок контейнера на дводольному графі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Основи комп'ютерної стеганографії : навчальний посібник для студентів і аспірантів / В.О. Хорошко, О.Д. Азаров, М.Є. Шелест, Ю.Є. Яремчук. – Вінниця : ВДТУ, 2003. – 143 с.

2. *Грибунин В.Г.* Цифровая стеганография / В.Г. Грибунин, И.Н. Оков, И.В. Туринцев. – М. : Солон-Пресс, 2002. – 272 с.

3. *T. Filler, J. Judas, J. Fridrich* Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes// Forensics and Security, vol. 6(1), pp. 920–935, 2011.

4. *J. Kodovsky, J. Fridrich, V. Holub* On Dangers of Overtraining Steganography to an Incomplete Cover Model// Proc. ACM Multimedia & Security Workshop, Niagara Falls, New York, September 29–30, pp. 69–76, 2011.

5. *T. Filler, J. Fridri*ch Gibbs construction in Steganography //*Forensics and Security*, 5(4), pp. 705–720, 2010.

6. *Fridrich J., Filler T.* Practical methods for minimizing embedding impact in steganography // Proceedings SPIE, Electronic Imaging, Steganography, and Watermarking of Multimedia Contents IX.– 2007.– 6505.– pp. 2–3.

7. *Hetzl S., Mutzel P.* A graph-theoretic approach to steganography// Proc. Communication and Multimedia security. –2005. pp.119-128.

8. *Борисенко І.І.* Застосування методів порівняння послідовностей в стеганографічних перетвореннях цифрових зображень / І.І. Борисенко // Сучасна спеціальна техніка. – 2014. – № 2. – С. 110–115.

9. *Харари Ф.* Теория графов / Ф. Харари. – М. : "Мир", 1993. – С. 203.

10. *Нікольський Ю.В.* Дискретна математика / Ю.В. Нікольський, В.В. Пасічник, Ю.М. Щербина. – К. : Видавнича група BHV, 2007. – С. 354.

11. *Иванов Б.Н.* Дискретная математика. Алгоритмы и программы : учеб. пособие / Б.Н. Иванов. – М. : Лаборатория Базовых Знаний, 2001. – 288 с.

12. *Кобозєва А.А.* Загальний підхід до оцінки властивостей стеганографічного алгоритму, заснованого на теорії збурень / А.А. Кобозєва // Информационные технологии и компьютерная инженерия. – 2008. – № 1. – С. 164–171.

13. *C. Bergman, J. Davidson* Unitary embedding for data hiding with the SVD // Security, steganography, and watermarking of multimedia contents VII, SPIE Vol.5681, 2005.

Отримано 27.04.2016
Рецензент Рибальський О.В., д.т.н.