УДК 004.056.5

І.І. Борисенко

ЗАСТОСУВАННЯ МЕТОДІВ ПОРІВНЯННЯ ПОСЛІДОВНОСТЕЙ В СТЕГАНОГРАФІЧНИХ ПЕРЕТВОРЕННЯХ ЦИФРОВИХ ЗОБРАЖЕНЬ

У статті пропонується новий стеганографічний алгоритм просторової області вбудовування в цифрове зображення. Основним принципом розробки є мінімізація впливів вбудованого повідомлення на контейнер. В основу алгоритму покладено порівняння бітових послідовностей контейнера та повідомлення, модифікація елементів контейнера виконується тільки у випадку, коли виявлено неспівпадіння відповідних бітів. Алгоритм дозволяє зменшити спотворення контейнера, зберегти статистики першого порядку та забезпечити стійкість до найбільш відомих статистичних атак.

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

В статье предлагается новый стеганографический алгоритм пространственной области встраивания в цифровое изображение. Основным принципом разработки является минимизация влияния встроенного сообщения на контейнер. В основу алгоритма положено сравнение битовых последовательностей контейнера и сообщения, модификация элементов контейнера выполняется только в случае, когда выявлено несовпадение соответствующих битов. Алгоритм позволяет уменьшить искажения контейнера, сберечь статистики первого порядка, а также обеспечить устойчивость к наиболее известным статистическим атакам.

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

The paper proposes a new steganographic algorithm for spatial-domain of digital images. The main design principle is to minimize a embedding impact by means of efficient coding algorithm. The algorithm is based on the comparing of bit sequence of the container and of the message, and change of the elements of the container are executed only if the mismatch of the appropriate bits is found. The algorithm allows to reduce the distortions of the container, to save the first order statistic and is steady against the more known statistical attacks.

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

Основною метою використання комп'ютерної стеганографії є приховування повідомлень у цифрових даних (ЦД), котрі, як правило, мають аналогову природу (мова, зображення, аудіо- або відеозапис). Це ефективний засіб захисту інформації, який стає особливо актуальним у випадку, коли застосування криптографічних методів неможливе або обмежене. Повідомленням є будь-яка конфіденційна інформація (особисті та медичні дані, банківська та комерційна інформація і т.п.), яка повинна бути вбудована таким чином, щоб навіть сам факт її присутності у контейнері був таємним. ЦД, в які вбудовується повідомлення, носять

узагальнену назву "контейнер", результатом такого вбудовування є стеганоконтейнер або стеганографічне повідомлення, яке відкрито пересилається одержувачу каналами загального користування [1; 2]. Будь-який стеганографічний алгоритм характеризується трьома основними властивостями: стійкістю, надійністю сприйняття (стеганоконтейнер візуально не повинен відрізнятися від контейнера), прихованою пропускною здатністю. Під прихованою пропускною здатністю (ППЗ) розуміють максимальну кількість інформації, яка може бути вкладена в один елемент контейнера [2].

Як правило, вбудовування повідомлення відбувається за рахунок корегування елементів контейнера, що призводить до зміни його статистичних характеристик. Ці зміни використовуються статистичними методами стеганоаналізу для розпізнавання стеганоконтейнерів [3–5]. Зрозуміло, що чим менше збурень зазнає контейнер під час вбудовування повідомлення, тим складніше стеганоаналітичним методам забезпечити низький рівень похибки при розпізнаванні. Останнім часом активно ведуться роботи зі створення 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 [6–10]. Нарівні з методами стеганографії розробляються і методи стеганоаналізу, які враховують принципи роботи таких класів алгоритмів, тому подальший розвиток стеганографії залишається актуальним.

Метою роботи є розробка стеганографічного алгоритму, стійкого до статистичних атак за рахунок забезпечення малих збурень контейнера.

В наш час є велика кількість алгоритмів для порівняння стрічок символів з метою визначення їх схожості: пошук за допомогою графів, застосування динамічного програмування, алгоритми точного співпадіння Бойєра-Мура, Кнутта-Моріса-Пратта й їх модифікації та інші. Таке розмаїття методів пояснюється тим, що завдання, які виникають в реальному житті, наприклад порівняння біологічних послідовностей [11], є складними і не завжди можна відразу відшукати їх точне вирішення. Окремим завданням виступає побудова алфавіту, якому будуть належати елементи послідовностей (залежно від прикладної задачі).

Якщо ж маємо дві бінарні послідовності, то завдання спрощується і задачу побітового пошуку входження послідовності N в послідовність M можна звести до виконання арифметичних операцій, використовуючи формули

$f(N)=\sum_{k=1}^{n}2^{n-k}N(k)$ та $f(M_j)=\sum_{k=1}^{n}2^{n-k}M(j+k-1)$ для переходу від двійкової до десяткової системи числення, де n – довжина послідовності N, j – позиція послідовності M, з якої починається порівняння N та M. Наприклад, визначимо входження послідовності N=0110 в послідовність M=10001101. Один раз обчислюємо $f(N)$, це значення дорівнює 6, а потім для кожної позиції j послідовності M обчислюємо $f(M_j)$. Зрозуміло, якщо є входження N в M з позиції j, то $f(N)=f(M)$, оскільки будь-яке ціле число однозначно представлене у вигляді суми додатних ступенів двійки. Така позиція для цього прикладу є, це – j=3.

Як контейнер будемо використовувати цифрове зображення, яке представлене бінарною матрицею M, біти якої послідовно групуються в підмножини M_i довжиною m кожна. Біти повідомлення N послідовно групуються у підмножини N_i довжиною n, де $m > n$. Вбудовування N_i в M_i, зважаючи на

алгоритм, наведений вище, можливе лише у випадку, коли об'єм повідомлення незначний і можна дозволити будувати M_i такі, що $m >> n$, щоб забезпечити повну відповідність бітів послідовності N_i деякій підмножині бітів підпослідовності M_i. Наприклад, як показали дослідження, якщо ППЗ становить 0.16 біт/піксель, то точна відповідність становить приблизно 16 % з усіх M_i, при зменшенні ППЗ до 0.125 біт/піксель маємо вже 20 %.

Зрозуміло, що чим більшого об'єму повідомлення треба вбудувати в контейнер, тим меншого значення повинен набувати параметр m, а отже, стеганографічний алгоритм повинен забезпечити можливість корегування яскравості пікселів, двійкове значення яких не збіглося з бітами повідомлення. Між ППЗ та рівнем збурень, яких зазнає контейнер після вбудовування повідомлення, завжди повинен існувати розумний компроміс, інакше повідомлення буде легко виявлено стеганоаналітичними методами. Введемо параметр *d*, значення якого буде відповідати кількості пікселів, які будуть корегуватися, тобто це кількість неспівпадінь N_i з бітами M_i, які ми дозволяємо допустити. Надалі алгоритм, який запропонований в роботі, будемо називати Seek-Place.

Ядром алгоритму Seek-Place є підпрограма Seek, основне призначення якої – пошук послідовності N_i в послідовності M_i з позиції *j* з точністю до *d*. На кожній ітерації роботи Seek обчислюється кількість бітів mm послідовності N_i і послідовності M_i, які збіглися, починаючи з позицій *r* для N_i та *j_rab* для M_i до першої розбіжності.

Підпрограма Seek *пошуку послідовності* N_i *в послідовності* M_i *з позиції j*

1. Покласти *r*=1, *j_rab=j*, *number = 0*.

2. Обчислити кількість *num* бітів, що збіглися, починаючи з позицій *r* для N_i і *j_rab* для M_i за правилом $num = \sum_k \overline{N_i(k) \oplus M_i(k)}$, де $\oplus$ – операція додавання за модулем 2, $\overline{\bullet}$ – операція логічного заперечення.

3. Якщо $r+num=n+1$, то знайдено збіг з точністю до *d* послідовностей N_i та M_i, яке починається в *j*; вихід з підпрограми.

4. Якщо $number \le d$, то виконати: $number = number + 1$, $r = r + num + 1$, $j\_rab = j\_rab + num + 1$; перейти до кроку 2.

Якщо $number = d + 1$, то збіг з точністю до *d* в позиції *j* послідовностей N_i та M_i немає; вихід з підпрограми.

Алгоритм створення стеганографічного контейнера

Вхід: матриця цифрового зображення – контейнер (*M*) та повідомлення (*N*) в 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 M_i та N_i виконувати підпрограму Seek, починаючи з *j*=1 та збільшуючи *j* на одиницю при кожному новому її викликові доти, поки в позиції

j знайдеться збіг з точністю до d, і тоді N_i вважається вбудованою з корегуванням відповідних d пікселів, а j є елементом ключа K_i, який передається разом зі стеганоконтейнером, або ж збіг із точністю до d не виявлено при всіх можливих положеннях N_i відносно M_i і тоді такий M_i пропускається, елемент ключа K_i для такого M_i дорівнює нулю, а N_i вбудовується в M_{i+1}.

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 S, починаючи з позиції K_i, якщо $K_i \neq 0$, виписати послідовність його елементів довжини n.

Оцінка властивостей побудованого алгоритму

Одним із основних завдань стеганоаналізу є оцінка спотврень зображень-контейнерів, які вносяться при вбудовуванні повідомлення. Таку оцінку можна дати за допомогою різницевих та кореляційних показників [12]. Для порівняльної оцінки ефективності алгоритму Seek-Place використовувався алгоритм [10], який був реалізований в стеганосистемі Steghide і є стійким до статистичної атаки x^2 та її модифікацій за рахунок того, що при вбудовуванні повідомлення елементи контейнера більшою мірою обмінюються місцями ніж модифікуються. Показники відображені в таблиці 1 і свідчать на користь Seek-Place.

Таблиця 1

Показники візуального спотворення

Назва показника	Оригінал	Steghide	Seek-Place
Середня абсолютна різниця	0	0,0823	0,0228
Нормована середня абсолютна різниця	0	3,7577e-04	1,8760e-04
Середньоквадратична помилка	0	0,0823	0,0228
Відношення “сигнал/шум” (SNR)	∞	6,4724e+05	9,3503e+05
Пікове відношення “сигнал/шум” (PSNR)	∞	7,1958e+05	2,6918+06

Для оцінки збурень контейнера, які викликані вбудовуванням повідомлення, та якісного порівняння довільних стеганографічних алгоритмів у роботі також було використано новий підхід, заснований на теорії збурень та матричному аналізі [13], в основі якого лежить аналіз наборів параметрів матриці контейнера, котрі однозначно його визначають, в ролі яких виступають сингулярні числа (СНЧ) та ортонормовані сингулярні вектори (СНВ), а також власні числа (ВЧ) та ортонормовані власні вектори (ВВ). Для визначеності будемо використовувати перший набір параметрів, а саме СНЧ та СНВ. Метод порівняльної оцінки властивостей стеганоалгоритмів [13] полягає в дослідженні збурень, яких зазнали

СНВ матриці контейнера і (чи) її ортонормовані СНВ в результаті вбудовування повідомлення. Застосуємо зазначений метод для дослідження Steghide та Seek-Place. Позначимо матрицю контейнера літерою *M*, а стеганоконтейнери, одержані Steghide та Seek-Place, – *S_Steg* та *S_Seek* відповідно. Основні кроки дослідження:

1) для матриць *M*, *S_Steg*, *S_Seek* побудувати нормальний сингулярний розклад [14]: $M=USV^{\mathrm{T}}$, $S\_Steg=U_{\mathrm{Steg}}S_{\mathrm{Steg}}V^{\mathrm{T}}{}_{\mathrm{Steg}}$, $S\_Seek=U_{\mathrm{Seek}}S_{\mathrm{Seek}}V^{\mathrm{T}}{}_{\mathrm{Seek}}$;

2) знайти збурення матриць СНЧ $\Delta S_{Steg}=S-S_{Steg}$, $\Delta S_{Seek}=S-S_{Seek}$ та СНВ $\Delta U_{Steg}=U-U_{Steg}$, $\Delta U_{Seek}=U-U_{Seek}$;

3) оцінити значення: $\delta_{Steg}=\max_i|(\Delta S_{Steg})_{ii}|$ та $\delta_{Seek}=\max_i|(\Delta S_{Seek})_{ii}|$, де $(\Delta S_{[\bullet]})_{ii}$ – діагональні елементи матриць $\Delta S_{[\bullet]}$;

4) оцінити збурення СНВ матриць U_{Steg} та U_{Seek}, використовуючи яку-небудь векторну норму.

Обчислювальний експеримент проводився в середовищі MatLab. У контейнерах виділялися блоки розміром $n\times n$, в які вбудовувалась однакова кількість бітів повідомлення алгоритмами Steghide та Seek-Place, потім досліджувалися параметри одержаних стеганоконтейнерів. Наведемо результати одного з дослідів для n-48 та пропускною здатністю блока 0,17 біт/піксель: $\delta_{Steg}=0{,}9853$, $\delta_{Seek}=0{,}4213$; норми перших чотирьох сингулярних векторів U_{Steg} – 2,0361e-04 0,0279 0,0410 0,1050, для U_{Seek} – 8.0490e-05 0,0082 0,0236 0,0592 для інших СНВ, які відповідають СНЧ, починаючи з п'ятого для матриць U_{Steg} та U_{Seek}, значення норм є порівнюваними і майже однаковими. Таким чином, порівнюючи збурення матриць СНЧ та СНВ, внесені алгоритмами Steghide та Seek-Place, можна зробити висновок, що збурення контейнера – зображення менші при використанні Seek-Place.

У роботі побудовано стеганографічний алгоритм Seek-Place, заснований на знаходженні схожих бітових послідовностей в повідомленні та контейнері. Порівняно з Steghide представлений алгоритм забезпечує менші візуальні спотворення контейнера за всіма різницевими показниками за рахунок менших збурень параметрів його матриці. Пояснити одержані результати можна, напевно, тим, що моделлю контейнера в Steghide є граф, а ребра між вузлами є тільки тоді, коли хоча б один піксель одного вузла можна обміняти хоча б на один піксель іншого. Отже, наявні ізольовані вузли, пікселі яких доводиться модифікувати. Два вузла можуть з'єднуватися більш ніж одним ребром, тому відшукується максимальне паросполучення, відповідно до якого пікселі одного вузла обмінюється на пікселі іншого. Не для кожного вузла знаходиться пара, тому це ще один випадок, коли модифікуються пікселі. Загальна кількість модифікованих пікселів алгоритмом Steghide виявилася більшою ніж Seek-Place. Завдяки можливості настроювання параметрів алгоритму Seek-Place, таких як n, m і d зберігаються статистики першого порядку контейнера, алгоритм стійкий до статистичних атак на основі методу оцінки числа переходів значень НЗБ в сусідніх елементах контейнера, x^2 та інших.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. *Хорошко В.О.* 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. *Барсуков В.С.* Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации / В.С. Барсуков, А.П. Романцов // Специальная Техника. – 2000. – № 1.

4. *Дрюченко М.А.* Алгоритмы выявления стеганографического скрытия информации в jpeg-файлах / М.А. Дрюченко // Вест. Воронеж. гос. ун. Системный анализ и информационные технологии. – 2007. – № 1. – С. 21–30.

5. *Иванов М.А.* Теория, применение и оценка качества генераторов псевдослучайных последовательностей / М.А. Иванов, И.В. Чугунков. – М. : КУДИЦ-ОБРАЗ, 2003.

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

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

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

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

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

11. *Д. Гасфилд.* Строки, деревья и последовательности в алгоритмах / Д. Гасфилд. – Санкт-Петербург : Невский диалект, БХВ-Петербург, 2003.

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

13. *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.

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

Отримано 29.07.2014