Стойкая модификация стеганографического метода LSB matching

А.Ю.Гуфан, А.Р.Тикиджи-Хамбурьян

Задача сокрытия информации в стеганоконтейнере обычно решается с учетом специфики типа стеганоконтейнера, для которого разрабатывается метод сокрытия. Так, наиболее распространенные методы сокрытия в контейнерах, представляющих собой оцифрованные данные, исходно имеющие аналоговую природу (изображения, аудио - и видеофайлы), относятся к классу методов сокрытия в младших битах отсчетов (цифровых значений, полученных при дискретизации аналогового сигнала). Их применимость базируются на том специфическом для таких контейнеров факте, что вызываемые произвольном изменением младших бит искажения исходных объектов, как правило, не воспринимаются наблюдателем (слушателем). Однако, при статистическом анализе оказывается, что во многих случаях модификация контейнера такими методами значительно изменяет его даже самые простые статистические характеристики. Так, метод простой замены последнего бита части отсчетов на очередной бит сообщения, подлежащего сокрытию, значительно деформирует частотную гистограмму контейнера, что делает этот метод уязвимым для атак. При этом, безусловно, важны статистические особенности конкретного используемого контейнера (так, например, существенные изменения в статистических характеристиках – и соответствующие существенные изменения пригодности к использованию в качестве стегоконтейнера – вносят всевозможные виды постобработки цифровых изображений и аудиозаписей, см., например, [1]), но важен здесь принципиальный момент: деформация частотной гистограммы контейнера при использовании такого сокрытия, какова бы она ни была у незаполненного контейнера. Поэтому важной является проблема выбора альтернативных способов модификации значений отсчетов, приводящей к соответствию их младших бит очередным битам скрываемого сообщения.

Одним из популярных вариантов решения этой проблемы является метод, получивший название "LSB matching". В общем случае, модификация производится следующим образом. Значение отсчета не изменяется, если его младший бит совпадает с очередным битом скрываемого сообщения. В случае, когда младший бит отсчета и очередной бит скрываемого сообщения не совпадает, значение отсчета увеличивается или уменьшается на 1 с определенной вероятностью. Обычно вероятности событий увеличения или уменьшения отсчета фиксируются в значениях равных 0.5.

При высокой плотности сокрытия такой способ сокрытия также влияет на статистики первого порядка, что может быть обнаружено простым статистическим анализом (например, аналогичным использованному в работе [2]). Это происходит потому, что значение текущего отсчета k при изменении принимает либо значение на 1 меньше - k - 1, либо на 1 больше - k + 1 и, таким образом, частотная гистограмма "сглаживается", за счет "перехода" значений из соседних столбцов (см. рис.1). Данное "сглаживание" можно выявить статистически, например, при помощи анализа разностей между локальными экстремумами гистограммы, а также их верхней и нижней сплайновой огибающей [3].

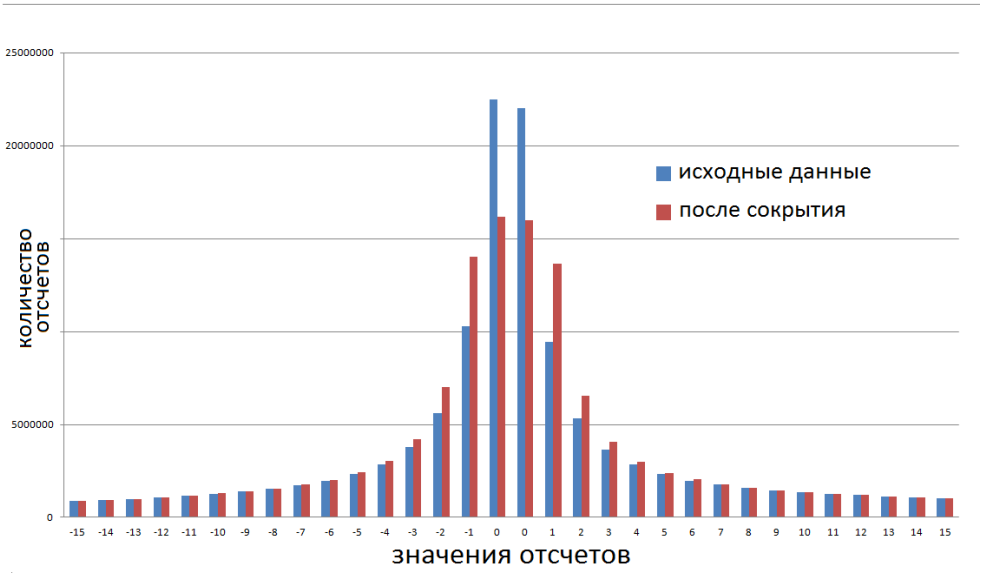

Рисунок 1 – изменение частотной гистограммы при сокрытии сообщения в 100% отсчетов аудиофайла методом LSB-matching с равными вероятностями увеличения и уменьшения значений отсчетов.

В данной работе предлагается решение описанной проблемы и некоторых других проблем метода LSB matching, родственных ей и широко описанных в литературе (см., например, [4-9]) при помощи более сложной процедуры выбора вероятностей увеличения или уменьшения каждого отсчета (вероятностей перехода отсчета со значением k в k + 1 или k – 1 при необходимости изменения его младшего бита) в зависимости от частотной гистограммы используемого стеганоконтейнера.

Обозначим $\{p_i\}_{i=1}^{N}$ - частоты встречаемости отсчетов со значениями i. Для каждого значения необходимо выбрать вероятности уменьшения и увеличения каждого отсчета: $\{q_i\}_{i=1}^{N}$ и $\{r_i\}_{i=1}^{N}$.

Ясно, что для отсчетов с максимальным для контейнера значением вероятность увеличения равна 0. Аналогично, для отсчетов с минимальным значением вероятность уменьшения также равна 0. Таким образом, вероятности на максимальном и минимальном для контейнера значениях можно записать в виде: $q_1 = 0, r_1 = 1, q_N = 1, r_N = 0$.

Чтобы искажения частотной гистограммы при трансформации значений отсчетов в соседние, были минимальны, в идеальном случае, количество прибавлений 1 к отсчету со значением k должно совпадать для любых k с количеством уменьшений на 1 отсчета со значением k + 1. Это требование формулируется в виде системы линейных уравнений

$$\begin{cases} q_1 = 0, r_1 = 1, q_N = 1, r_N = 0; \\ q_i p_i = r_{i-1} p_{i-1} \\ q_i + r_i = 1 \end{cases} \quad (1)$$

В такой системе имеется 2N-3 уравнений относительно 2N-4 неизвестных, а значит, в общем случае она не имеет решения. Разобьем данную систему на две так, чтобы каждая из них описывала задачу с учетом только одного граничного условия. Построим решение $\{q_i^1\}_{i=1}^{N-1}, \{r_i^1\}_{i=1}^{N-1}$ системы уравнений

$$\begin{cases} q_1 = 0, r_1 = 1; \\ q_i p_i = r_{i-1} p_{i-1}; \\ q_i + r_i = 1 \end{cases} \quad (2)$$

И решение $\{q_i^2\}_{i=2}^N, \{r_i^2\}_{i=2}^N$ системы уравнений

$$\begin{cases} q_N = 1, r_N = 0; \\ q_i p_i = r_{i-1} p_{i-1}; \\ q_i + r_i = 1 \end{cases} \quad (3)$$

Приближенное решение системы (1) будем искать как среднее арифметическое решений систем (2) и (3)

$$q_1 = 0, r_1 = 1, q_N = 1, r_N = 0, q_i = \frac{q_i^1 + q_i^2}{2}, r_i = \frac{r_i^1 + r_i^2}{2}$$

Такой метод выбора вероятностей, приводит к значительно более точному сохранению рельефа гистограммы и высоты пиков (см. рис.2).

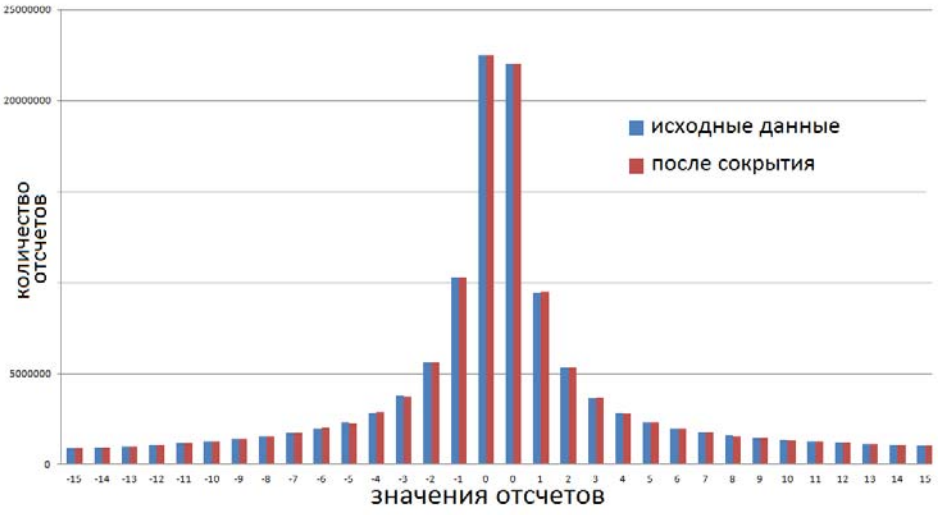

Рисунок 2 – изменение частотной гистограммы при сокрытии сообщения в 100% отсчетов аудиофайла методом LSB-matching с использованием предлагаемого способа выбора вероятностей изменений отсчетов.

Важной особенностью такого подхода к выбору изменений, которым подвергается стегоконтейнер, является его обобщаемость на случай противодействия более сложным методам анализа – таким, как, например, анализ пар значений [10].

Список литературы:

1. Д.А.Хрящев Д.А. Повышение качества изображений, полученных в условиях недостаточной освещенности [Электронный ресурс] // «Инженерный вестник Дона», 2013, №3 – Режим доступа: http://ivdon.ru/magazine/archive/n3y2013/1796 (доступ свободный) – Загл. с экрана. – Яз. рус.

2. И.М. Ажмухамедов, А.Н. Марьенков Поиск и оценка аномалий сетевого трафика на основе циклического анализа [Электронный ресурс] // «Инженерный вестник Дона», 2012, №2 – Режим доступа: http://ivdon.ru/magazine/archive/n2y2012/742 (доступ свободный) – Загл. с экрана. – Яз. рус.

3. Zhang J., Hu Y., Yuan Z. Detection of LSB Matching steganography using the Envelope of Histogram // Journal of Computers, vol. 4, No. 7, July 2009.

4. Andrew D. Ker Steganalysis of LSB Matching in Grayscale Images // IEEE Signal processing letters, vol. 12, No. 6, June 2005, p. 441 – 444.

5. Fangjun Huang, Sun Yat-Sen, Bin Li, Jiwu Huang, Attack LSB Matching Steganography by Counting Alteration Rate of the Number of Neighbourhood Gray Levels // Image Processing, 2007. ICIP , vol. 1, p. 401 – 404.

6. Jun Zhang, Dan Zhang Detection of LSB Matching Steganography in Decompressed Images // Signal Processing Letters, Vol. 17, Iss. 2, p. 141 – 144.

7. Jiaohua Qin, Xuyu Xiang, Meng Xian Wang A Review on Detection of LSB Matching Steganography // Information Technology Journal 2010, vol. 9, Iss. 8, p. 1725 - 1738.

8. Q. Liu, A. Sung, J. Xu, B. Ribeiro, Image complexity and feature extraction for steganalysis of LSB," in ICPR06, pp. II:267–270, 2006

9. Jun Zhang, I. J. Cox, G. Doerr, "Steganalysis for LSB matching in images with high-frequency noise," Proc. IEEE Workshop on Multimedia Signal Processing, pp. 385-388, 2007

10. S. Dumitrescu, X. Wu, Z. Wang, "Detection of LSB steganography via sample pair analysis," IEEE Trans. Signal Process, 51 (7), pp.1995–2007, 2003