ЗАХИСТ ІНФОРМАЦІЇ

УДК 004.056.5

И.И. Бобок

ВЫЯВЛЕНИЕ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ, ПЕРЕСОХРАНЕННЫХ В ФОРМАТ БЕЗ ПОТЕРЬ ИЗ ФОРМАТА С ПОТЕРЯМИ, КАК СОСТАВНАЯ ЧАСТЬ СТЕГАНОАНАЛИЗА

В работе проведено исследование возможности решения в границах пространственной области изображения вопроса отделения цифровых изображений, сохраненных в форматах без потерь первоначально, от пересохраненных в формат без потерь после сжатия за счет анализа возмущений значений яркости пикселей в результате пересохранения. Реализация такой возможности при ее наличии может быть использована как составная часть стеганоанализа, в частности для детектирования метода модификации наименьшего значащего бита.

Ключевые слова*: стеганоанализ, цифровое изображение, формат с потерями, формат без потерь, метод модификации наименьшего значащего бита.*

У роботі проведене дослідження можливості рішення в межах просторової області зображення питання відокремлення цифрових зображень, збережених у форматах без втрат спочатку, від перезбережених у формат без втрат після стиснення за рахунок аналізу збурень значень яскравості пікселів у результаті перезбереження. Реалізація такої можливості при її наявності може бути використана як складова частина стеганоаналізу, зокрема для детектування методу модифікації найменшого значущого біта.

Ключові слова*: стеганоаналіз, цифрове зображення, формат із втратами, формат без втрат, метод модифікації найменшого значущого біта.*

The paper focuses on the solution of the problem of separation digital images, originally stored in lossless format from images which are resaved to lossless from lossy format by analyzing perturbation of the pixel brightness values as a result of re-saving. Implementation of such a possibility can be used as a part steganalysis, particularly for the detecting least significant bit method.

Keywords: *steganalysis, digital image, lossy format, lossless format, least significant bit method.*

Метод модификации наименьшего значащего бита (LSB) [1,2] был и остается на сегодняшний день одним из самых распространенных стеганографических методов, приводя к актуальности задачи стеганоанализа его многочисленных реализаций [2].

Специфика LSB такова (неустойчивость к любым возмущающим воздействиям, направленным на стеганосообщение), что после погружения дополнительной информации в контейнер, в качестве которого часто используется цифровое изображение (ЦИ), результирующее стеганосообщение (СС), сохраняется в формате без потерь (ФБП). В настоящее время хранение и передача ЦИ по каналам телекоммуникаций в связи со значительным увеличением объемов информации осуществляется в сжатом состоянии. Этот немаловажный факт не может не учитываться при разработке подхода к решению задачи стеганоанализа [3, 4], в том числе метода LSB, в силу чего в качестве контейнера имеет смысл рассматривать ЦИ в формате с потерями (ФСП), что приводит к целесообразности и желательности использования в качестве составной части стеганоанализа возможности (при наличии таковой) различать между собой ЦИ, которые сохранены в ФБП первоначально и пересохранены в ФБП после сжатия. Действительно, само по себе пересохранение из ФСП в формат без потерь никоим образом не улучшает качество ЦИ, поскольку не меняет его матрицу, однако значительно увеличивает размер хранимого ЦИ, в результате чего является косвенным показателем несанкционированных действий, проведенных над изображением (в частности, стеганопреобразования).

Анализ ЦИ в общем случае может проводиться как в пространственной, так и в области преобразования (в частности, в частотной). Однако пространственная область, не требующая дополнительных вычислительных затрат на различные преобразования для вычисления анализируемых в ходе исследования параметров, очевидно является более предпочтительной.

Целью настоящей работы является исследование возможности решения вопроса отделения ЦИ, сохраненных в ФБП первоначально, от пересохраненных в ФБП после сжатия в пространственной области ЦИ за счет анализа возмущений значений яркости пикселей в результате пересохранения ЦИ в ФСП.

При наличии такой возможности ее реализация может быть использована как составная часть стеганоанализа.

Для достижения поставленной цели в работе необходимо решить ряд *задач*.

1. Исследовать возмущения значений яркости пикселей ЦИ при сохранении в ФСП в случае, когда исходное изображение было сохранено в ФБП.

2. Исследовать возмущения значений яркости пикселей ЦИ при сохранении в ФСП в случае, когда исходное изображение было пересохранено в ФБП из ФСП.

3. Исследовать зависимость/независимость результатов решения задач 1,2 от значений элементов матрицы квантования, используемой при сохранении ЦИ в ФСП.

Независимо от конкретики непосредственной реализации сжатия в силу специфики человеческого зрения сжатие происходит таким образом, что его результат приводит к исключению из сигнала его высокочастотных (а возможно, и среднечастотных) составляющих за счет обнуления соответствующих коэффициентов [5]. В силу этого матрицы ЦИ в ФБП (далее обозначаемые F_T) и в ФСП (далее обозначаемые F_J) различны. Кроме того, матрица изображения, сохраненного в ФБП первоначально (F_T) и пересохраненного в ФБП после сжатия (обозначим последнюю $F_{J\to T}$, она не отличается от F_J) качественно отличаются друг от друга по своим характеристикам [6]. Действительно, F_T отвечает представлению сигнала, у которого все частотные коэффициенты в невозмущенном (неквантованном) виде, в то время как $F_{J\to T}$ соответствует представлению сигнала, у которого уже практически “отсутствуют” высокочастотные (возможно, среднечастотные) составляющие – коэффициенты при них если и ненулевые, то малые. Это приведет к тому, что изображения с матрицами F_T и F_J по-разному отреагируют на пересохранение ЦИ в ФСП. Существование возможности выявления отличий в возмущениях F_T и F_J при пересохранении в ФСП даст возможность для отделения ЦИ, сохраненного первоначально в ФБП, от пересохраненного в ФБП из ФСП.

Пускай сжатие происходит с достаточно высоким качеством (здесь и ниже в качестве ФСП, не ограничивая общности рассуждений, для определенности рассматривается самый распространенный в настоящий момент формат – JPEG, основанный на дискретном косинусном преобразовании (ДКП)). Это требование обусловлено требованием сохранения надежности восприятия СС. Сжатие для F_T – первое (матрицу результата обозначим $F_{J\to T}$), свойства $F_{J\to T}$ очевидно аналогичны F_J. При высоком качестве сжатия ожидаемым результатом здесь является незначительные возмущения значений яркости большинства пикселей изображения (за счет обнуления при квантовании высокочастотных коэффициентов ДКП). Для $F_{J\to T}$ (F_J) очередное сжатие является вторым (матрицу результата обозначим $F_{J\to T\to J}$ (или, что по сути то же самое, $F_{J\to J}$)), причем его характеристики могут как совпадать, так и не совпадать с характеристиками первого сжатия, в силу чего результат возмущений значений яркости пикселей $F_{J\to T}$ при переходе к $F_{J\to T\to J}$ в общем случае принципиально предсказать невозможно.

Действительно, пусть u – некоторый произвольный коэффициент ДКП ЦИ в ФБП с матрицей F_T, $\bar{u}$ – соответствующий ему коэффициент ДКП того же ЦИ в ФСП (с матрицей $F_{T\to J}$ (или, по сути, то же самое F_J, или $F_{J\to T}$)), полученный после квантования u с коэффициентом квантования [5] q_1:

$$\left[\frac{u}{q_1}\right]q_1 = \bar{u},$$

а $\bar{\bar{u}}$ – значение u после повторного пересохранения ЦИ в ФБП в ФСП с коэффициентом квантования q_2 (с матрицей $F_{J\to T\to J}$):

$$\bar{\bar{u}} = \left[\frac{\bar{u}}{q_2}\right]q_2 = \left[\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}\right]q_2,$$

где [] — операция округления до ближайшего целого. Предположим здесь, что $\left[\frac{u}{q_1}\right] \neq 0$.

При округлении до целого значения [7] имеет место соотношение:

$$\left|\left[\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}\right] - \left[\frac{u}{q_1}\right]\frac{q_1}{q_2}\right| \leq \frac{1}{2},$$

тогда

$$\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}-\frac{1}{2}\le\left[\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}\right]\le\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}+\frac{1}{2},$$

а

$$\left[\frac{u}{q_1}\right]q_1-\frac{q_2}{2}\le\left[\left[\frac{u}{q_1}\right]\frac{q_1}{q_2}\right]q_2=\overset{=}{u}\le\left[\frac{u}{q_1}\right]q_1+\frac{q_2}{2}. \tag{1}$$

При этом $\left[\frac{u}{q_1}\right]q_1=\bar{u}$ – восстановленный коэффициент ДКП u после первого квантования, удовлетворяющий аналогичному (1) соотношению:

$$u-\frac{q_1}{2}\le\left[\frac{u}{q_1}\right]q_1=\bar{u}\le u+\frac{q_1}{2}. \tag{2}$$

Из (1) и (2) получаем:

$$\left|\bar{u}-u\right|\le\frac{q_1}{2} \tag{3}$$

$$\left|\overset{=}{u}-\bar{u}\right|\le\frac{q_2}{2} \tag{4}$$

Из (3) и (4) вытекает, что соотношение возмущений, которые претерпевают коэффициенты ДКП (а вследствие этого и матрицы ЦИ в целом) при пересохранении в ФСП изображения, хранимого первоначально в ФБП, и изображения, хранимого первоначально в ФСП, будет вытекать из соотношения коэффициентов квантования q_1, q_2. В общем случае вид матрицы квантования, значения коэффициентов квантования q_1, q_2 зависит от технических характеристик фотокамер, при помощи которых делается ЦИ. Необходимо отметить, что оценки (3) и (4) − это оценки сверху для возмущений соответствующих коэффициентов ДКП. В реалиях эти возмущения в обоих случаях могут и не достигать верхних границ, а в случае $q_1=q_2$ даже максимальные абсолютные погрешности не смогут отделить ЦИ, первоначально находящиеся в ФСП от ФБП. Таким образом, в общем случае при помощи проделанного анализа принципиально невозможно сделать вывод о том, когда матрица ЦИ претерпевает большее возмущение: при переходе из F_T в $F_{T\to J}$ (или, что то же самое, в F_J) или при переходе F_J ($F_{J\to T}$) в $F_{J\to J}$ $\left(F_{J\to T\to J}\right)$.

Из (1) и (2) также вытекает, что

$$u-\frac{q_1}{2}-\frac{q_2}{2}\le\left[\frac{u}{q_1}\right]q_1-\frac{q_2}{2}\le\overset{=}{u}\le\left[\frac{u}{q_1}\right]q_1+\frac{q_2}{2}\le u+\frac{q_1}{2}+\frac{q_2}{2},$$

откуда видно, что если

$$u << \frac{q_1}{2} + \frac{q_2}{2},$$

то принципиально возможна ситуация, когда повторное квантование приведет к обнулению (сравнимости с нулем) $\overline{\overline{u}}$, даже в случае, когда $\left[\frac{u}{q_1}\right] \neq 0$. Такая ситуация может возникнуть для среднечастотных коэффициентов ДКП. Коэффициенты квантования для среднечастотных коэффициентов ДКП значительны [5]. Это значит, что квантование среднечастотных коэффициентов (имеющих немалое по модулю значение), осуществляемое дважды, может привести к их значительному возмущению и обнулению, т.е. в ЦИ с матрицей $F_{J \to T}$ (F_J) эти среднечастотные составляющие еще присутствуют, а в $F_{J \to T \to J}$ их уже нет (заметим, что повторное квантование на высокочастотных коэффициентах ДКП, которые после первого уже имели нулевые (малые по модулю за счет округлений) значения, снова обнулит их). Среднечастотные коэффициенты ДКП имеют по модулю, в большинстве своем, большие значения, чем высокочастотные, поэтому их возможное обнуление при повторном квантовании может привести к возникновению более сильных возмущений при переходе от $F_{J \to T}$ к $F_{J \to T \to J}$, чем при переходе от F_T к $F_{T \to J}$.

В ходе проведения исследований тестировались два множества I_1, I_2, каждое из которых состояло из 200 ЦИ размером 1024×1024 пикселей, сгенерированных различными фотоаппаратами (использующими разные матрицы квантования). Результаты тестирования следующие.

Обозначим:

$$R = abs\left(F_T - F_{T \to J}\right), \tag{5}$$

$$\overline{R} = abs\left(F_{J \to T} - F_{J \to T \to J}\right) \tag{6}$$

как матрицы абсолютных значений разностей соответствующих элементов F_T и $F_{T \to J}$, $F_{J \to T}$ и $F_{J \to T \to J}$ соответственно.

Пусть $M(A)$− значение, которое встречается среди элементов произвольной матрицы A с максимальной частотой, $\max(A)$ − максимальное значение среди элементов матрицы A. Рассматривается первая группа ЦИ − множество I_1.

С учетом поставленных задач в качестве предмета исследования рассматривались гистограммы значений матриц R, $\overline{R}$.

При проведении вычислительных экспериментов для определенности на этой стадии исследований в качестве ФБП использовался TIF, а как ФСП − JPEG, основанный на ДКП.

На первом этапе эксперимента устанавливались характерные особенности возмущений элементов матрицы изображения в пространственной области при переходе от F_T к $F_{T \to J}$. В результате вычислительного эксперимента было получено, что для различных ЦИ

$$\max(R) \in \{7,8,9,\ldots,32\}$$

причем для подавляющего большинства изображений $\max(R) \in \{18,19,\ldots,24\}$.

Для изображений из I_1:

$$M(R) \le 1.$$

Большинство пикселей ЦИ при переходе от F_T к $F_{T \to J}$, как и предполагалось, претерпевают незначительные возмущения или не возмущаются вовсе.

Во второй части вычислительного эксперимента устанавливались характерные особенности возмущений элементов матрицы изображения при переходе от $F_{J \to T}$ к $F_{J \to T \to J}$. В результате вычислительного эксперимента было установлено, что

$$\max(\overline{R}) \in \{36,37,\ldots,47\}$$

для разных ЦИ, причем для подавляющего большинства изображений $\max(\overline{R}) \in \{40,\ldots,44\}$. Для всех протестированных ЦИ $M(\overline{R}) > 1$, причем для абсолютного большинства изображений

$$M(\overline{R}) >> 1.$$

Таким образом, при тестировании ЦИ из I_1 для матриц (5), (6) были установлены следующие соотношения:

$$\max(R) \le \max(\overline{R}), \tag{7}$$

$$M(R) \le M(\overline{R}). \tag{8}$$

Тестирование ЦИ из группы I_2, полученной другим техническим средством, с отличной от первой матрицей квантования, дали противоположную качественную картину:

$$\max(R) \ge \max(\overline{R}), \tag{9}$$

$$M(R) \ge M(\overline{R}). \tag{10}$$

Однако, учитывая существование конечного множества матриц квантования и базы таких матриц, существует принципиальная возможность построения метода отделения ЦИ, сохраненных в ФБП первоначально, от тех, которые пересохранены в ФБП из ФСП, с учетом условий (7)–(8) и (9)–(10), работающего в пространственной области ЦИ.

В работе на основании проведенного исследования доказана возможность решения в пространственной области изображения задачи отделения ЦИ, сохраненных в формате без потерь первоначально, от ЦИ, пересохраненных в формат без потерь после сжатия с потерями. Алгоритм, реализующий данную возможность, может быть использован в качестве составной части любого стеганоаналитического метода, используемого для детектирования наличия конфиденциальной информации в СС, сохраненном в ФБП, внедренной в ЦИ-контейнер, хранимый в ФСП, в частности, для детектирования реализаций LSB-метода.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

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

2. Стеганография, цифровые водяные знаки и стеганоанализ : монография / А.В. Аграновский, А.В. Балакин, В.Г. Грибунин, С.А. Сапожников. – М. : Вузовская книга, 2009. – 220 с.

3. *Бобок И.И.* Стеганоаналитический метод для цифрового сигнала-контейнера, хранящегося в формате с потерями / И. И. Бобок // Сучасний захист інформації. – 2011. – № 2. – С. 50–60.

4. *Бобок И.И.* Стеганоаналитический алгоритм для основного сообщения, хранимого в форматах с потерями / И.И. Бобок // Вісник Національного технічного університету "ХПІ". – 2012. – № 29. – С. 41–49.

5. *Гонсалес Р.* Цифровая обработка изображений / Р. Гонсалес, Р. Вудс ; пер. с англ. П.А. Чочиа. — М. : Техносфера, 2006. — 1070 с.

6. *Бобок И.И.* Детектирование наличия возмущений матрицы цифрового изображения как составная часть стеганоанализа / И.И. Бобок // Вісник Східноукр. нац. ун-ту ім. В. Даля. – 2011. – № 7(161). – С. 32–41.

7. *Бахвалов Н.С.* Численные методы : учебное пособие для студ. физико-математических спец. вузов ; реком. МО РФ / Н.С. Бахвалов, Н.П. Жидков, Г.М. Кобельков. – 6-е изд. – М. : БИНОМ. Лаборатория знаний, 2008. – 636 с.

Отримано 17.09.2013