



УДК 004.932.2

© 2017 г. **С.В. Белим**, д-р физ.-мат. наук,

Д.Э. Вильховский

(Омский государственный университет им. Ф.М. Достоевского)

АЛГОРИТМ ВЫЯВЛЕНИЯ СТЕГАНОГРАФИЧЕСКИХ ВСТАВОК ТИПА LSB-ЗАМЕЩЕНИЯ НА ОСНОВЕ АНАЛИЗА СЛОЯ МЛАДШИХ БИТОВ

В статье предложен алгоритм выявления стеганографических вставок, реализуемых в виде подмены наименее значащих битов. Предложенный алгоритм основывается на анализе слоя младших битов. Области встраивания определяются с помощью алгоритма таксономии. Используется алгоритм предварительной обработки для повышения эффективности на участках, содержащих градиентную заливку.

Ключевые слова: поиск LSB-вставок, анализ стегоконтейнера, анализ изображений со вставками, выявление стеговставок.

DOI: 10.22250/isu.2017.54.3-11

Введение

В настоящее время наиболее известным методом встраивания стеганографических вставок является подмена наименее значащих бит (LSB-замещение) [1]. Данный метод основан на том, что замена от одного до четырех младших бит в байтах цветового представления пикселей изображения остается практически незаметной для человеческого глаза. Наибольшей возможностью для скрытия информации обладает синяя компонента, что обусловлено строением сетчатки глаза. Разработано достаточно много алгоритмов встраивания информации в изображения, аудио- и видеопотоки, однако метод LSB-замещения, являясь исторически первым, по-прежнему имеет широкое распространение. Все методы скрытия информации ориентированы на такое преобразование изображения, которое незаметно визуально. В связи с чем актуальной является задача построения алгоритмов анализа изображений на наличие стеганографических вставок.

На сегодняшний день все существующие алгоритмы ориентированы на определение самого факта присутствия или отсутствия стеганографической вставки в изображении. В работах [2, 3] приведен метод статистического анализа Хи-квадрата, основанный на предположении о случайном распределении младших битов цветных байтов изображения. Данный метод дает хорошие результаты при равномерном заполнении контейнера и слабее – при случайном выборе пикселей для замены младших битов. В статье [4] приведен метод визуального выявления наличия стеганографической вставки путем сравнения цветных срезов изображений. Показано, что данный метод дает хорошие результаты при наличии на изображении больших полей сплошной заливки. В статьях [5 – 7] стеганоанализ проводится на основе сравнения младших битов в соседних байтах с использованием формализма цепей Маркова. В статье [8] предложен метод обнаружения стеганографической вставки на основе использования искусственных нейронных сетей. Показано, что при достаточно большом объеме обучающей выборки нейронная сеть способна определять наличие вставки с ошибкой, не превышающей 15%. Все известные на сегодняшний день методы стеганоанализа метода LSB-замещения эффективны при заполнении стегоконтейнера не менее чем на 50% [9]. В статье [10] предложен метод обнаружения встроенной информации, основанный на алгоритмах сжатия информации. Основная идея метода состоит в том, что случайные данные сжимаются слабее, чем упорядоченные. Такой подход позволяет с высокой точностью определять наличие стеганографической вставки при заполнении контейнера от 40%. Данный метод получил развитие в работе [11] на основе использования предварительной обработки изображения, что позволило применять его при значительно меньших заполнениях контейнера.

Следует отметить, пока что отсутствуют алгоритмы, определяющие в байты, в которых произошла подмена младшего бита. По постановке данная задача близка к проблеме выявления пикселей, поврежденных импульсным шумом. Для импульсного шума характерно изменение цвета произвольно выбранного байта на случайно выбранную величину. Однако задача поиска встроенного сообщения является более сложной, так как величина изменения составляет всего один бит. Сегодня существует несколько методов обнаружения поврежденных пикселей. Прежде всего необходимо выделить метод SD-ROM [12], который, являясь исторически первым, послужил основой для ряда алгоритмов [13, 14]. Существуют и другие подходы, основанные на поиске ассоциативных правил [15, 16], методе кластеризации [17] и методе анализа иерархий [18].

В данной статье предложен алгоритм выявления пикселей изображения, в которых произведена подмена наименее значащего бита при стеганографическом встраивании сообщения, на основе автоматического анализа нулевого слоя.

Постановка задачи

Будем анализировать изображения, в которых может быть встроена информация в виде стеганографических вставок в младший бит синей компоненты. Анализ именно синей компоненты обусловлен тем, что встраивание в нее наименее заметно визуально, в связи с чем рекомендуется использовать ее для скрытия сообщений. Аналогичным образом могут быть проанализированы и другие компоненты без потери общности предлагаемого метода. Будем исходить из двух предположений. Во-первых, достоверно неизвестно, есть ли стеганографическая вставка или ее нет. Во-вторых, будем считать, что стеганографическая вставка заполняет некоторую прямоугольную область, размеры и положение которой неизвестны. Задача заключается не только в определении наличия стеганографической вставки, но и области, в которую произведено встраивание. Второе предположение существенно осложняет задачу, так как возможна ситуация, при которой заменены все младшие пиксели синей компоненты. Будем решать задачу, исходя из предположения о неполной подмене нулевого слоя.

Нулевой слой представляет собой матрицу из нулей и единиц. Закономерности распределения нулевых и единичных значений нулевого слоя без встраивания обусловлены структурой изображения. Встраивание сообщения вносит изменения в нулевой слой, изменяя плотность распределения единичных значений. На рис. 1 представлено изображение, где его нулевой слой без встраивания и нулевой слой с встроенным сообщением.

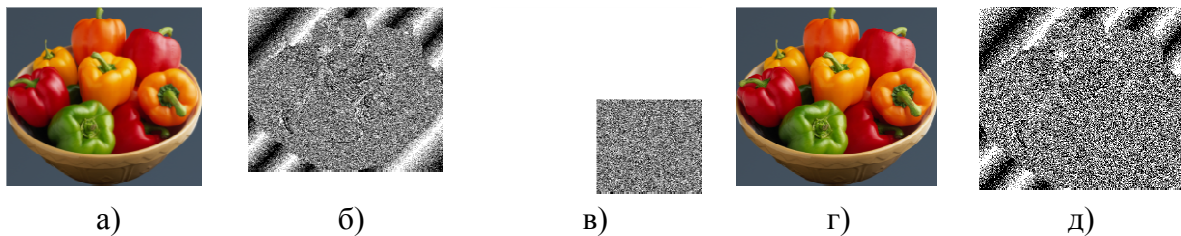


Рис. 1. Сравнение нулевых слоев: а) исходное изображение; б) нулевой слой исходного изображения; в) карта встраиваемых пикселей; г) изображение с встроенным сообщением; д) нулевой слой изображения с встроенным сообщением.

Как видно из рис. 1, наличие встроенного сообщения может быть обнаружено визуально из анализа нулевого слоя. Однако визуально можно определить только наличие встроенного сообщения, но не его расположение и размер. Поставим целью автоматически определить наличие, размеры и положение области встраивания.

Алгоритм выделения области встраивания

Для выделения области встраивания используем алгоритм, построенный на основе алгоритма таксономии FOREL [19]. В своем классическом виде FOREL

объединяет точки в таксоны, лежащие внутри окружности. В нашем случае мы будем строить таксоны прямоугольной формы.

Введем показатель плотности единичных значений p . Если в некоторой области изображения содержится N пикселей и N_1 из них имеют единичное значение, показатель плотности единичных значений $p = N_1/N$.

Будем искать прямоугольные области, имеющие плотность единичных значений заданной величины p_0 . Также в качестве входного параметра алгоритма зададим параметр R_0 , определяющий начальный размер таксона. Алгоритм состоит из следующих шагов.

Шаг 1. Выбираем начальное значение размера таксона $R = R_0$.

Шаг 2. Выбираем случайным образом точку с координатами (x_1, y_1) , которая играет роль центра таксона. Строим квадрат, левый верхний угол которого имеет координаты $(x_1 - R, y_1 - R)$, а правый нижний угол координаты $(x_1 + R, y_1 + R)$.

Шаг 3. Ищем координаты центра масс точек, лежащих внутри построенного квадрата (x_2, y_2) .

Шаг 4. Если точки (x_1, y_1) и (x_2, y_2) совпадают, то переходим к шагу 5, иначе $x_1 = x_2, y_1 = y_2$ и переходим к шагу 2.

Шаг 5. Вычисляем показатель плотности единичных значений p .

Шаг 6. Если $p > p_0$, то $R := 1.1R$ и переходим к шагу 3.

Шаг 7. Если $p < p_0$, то $R := 0.9R$ и переходим к шагу 3.

Шаг 8. Если $p = p_0$, то переходим к шагу 2.

Алгоритм выполняется до тех пор, пока все точки нулевого слоя не будут объединены в некоторые таксоны.

В качестве областей, в которые может быть встроено сообщение, выбираем таксоны, размер которых не меньше чем 10% размеров исходного изображения. Результаты работы данного алгоритма для изображения с равномерной заливкой представлены на рис. 2.

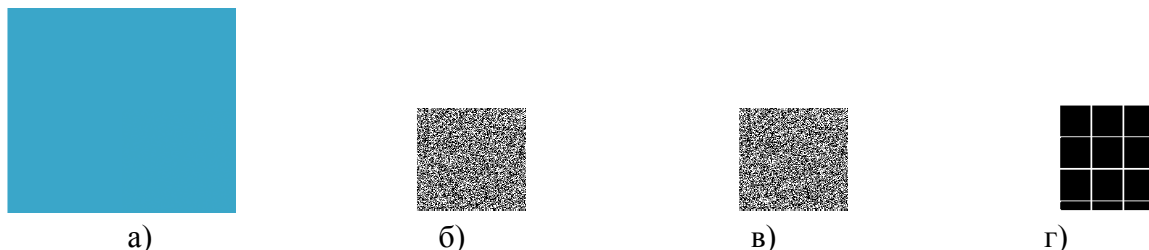


Рис. 2. Результаты работы алгоритма по автоматическому выделению области встраивания для изображения с равномерной заливкой: а) изображение со встроеной вставкой; б) карта встроенных пикселей, в) нулевой слой; г) автоматически выделенная область встраивания при $p = 0.5$.

Как видно (рис. 2), в результате работы алгоритма на искусственном изображении с равномерной заливкой область встраивания определяется достаточно

точно. Однако уже для искусственных изображений с градиентной заливкой возникают сложности, так как непрерывное изменение цвета изображения в целом проявляется как полосы одинаковых значений на нулевом слое. Появление данных полос связано с равномерным изменением цвета, которое проявляется как дискретное изменение младшего бита. Полосы могут быть устранены с помощью предварительной обработки изображения.

Алгоритм предварительной обработки изображений

Как было показано выше, нулевой слой изображения с градиентной заливкой представляет собой полосы из нулей и единиц. Используем линейное преобразование:

$$d(x, y) = ax + by - e,$$

где $a = c(x + 1, y) - c(x, y)$, $b = c(x, y + 1) - c(x, y)$, $c(x, y)$ – значения цвета пикселя, расположенного в точке с координатами (x, y) . Определим e как минимальное значение $c(x, y)$ на множестве всех пикселей изображения. В случае, когда цвет изображения представляет собой заливку с постоянным градиентом, функция $d(x, y)$ будет иметь постоянное значение ($d(x, y) = const$). Алгоритм принятия решений об изменении пикселя может быть применен к функции $d(x, y)$, но не к функции $c(x, y)$.

Для применения данного линейного преобразования к фотографическим изображениям необходимо определить области градиентной заливки. Будем вычислять вторые производные от функции $c(x, y)$ и выявлять области, в которых они имеют нулевое значение. В связи с тем, что области идеальной градиентной заливки на фотографических изображениях встречаются крайне редко, будем требовать выполнения трех более «мягких» условий:

$$\left| \frac{\partial^2 c(x, y)}{\partial x^2} \right| \leq 2, \quad \left| \frac{\partial^2 c(x, y)}{\partial y^2} \right| \leq 2, \quad \left| \frac{\partial^2 c(x, y)}{\partial x \partial y} \right| \leq 2.$$

Нестрогое неравенство вместо равенства нулю введено для того, чтобы учесть небольшие отклонения от градиентной заливки и не потерять встроенные биты.

После выявления связных областей, удовлетворяющих условиям на вторые производные, необходимо определить коэффициенты функции $d(x, y)$. Для их нахождения был использован метод наименьших квадратов, после чего вычислялось значение функции $d(x, y)$ для каждой из выделенных областей. После этого производилось построение карты для поиска области встраивания $map(x, y)$.

1. Если точка (x, y) находится в области градиентной заливки, то $map(x, y) = d(x, y)$.

2. Для всех остальных точек $map(x, y) = 0$.

К карте $map(x, y)$ применялся алгоритм поиска областей встраивания сообщений. Присваивание нулевых значений $map(x, y)$ в точках, не относящихся к областям градиентной заливки, позволяет снизить влияние границ областей изображения на результаты работы алгоритма.

Компьютерный эксперимент и результаты

Компьютерный эксперимент проводился для различных цветных изображений – как искусственных, так и фотографических. Встраиваемое сообщение представляло собой текстовую строку, представленную в виде последовательности битов. Встраивание производилось в синюю компоненту как наименее различимую человеческим глазом. Встраивание осуществлялось в прямоугольную область. В качестве задачи стегаанализа ставилось определение области, в которой расположены подмененные биты.

Первоначально алгоритм тестировался на прямоугольном искусственном изображении с градиентной заливкой. Результаты представлены на рис. 3.

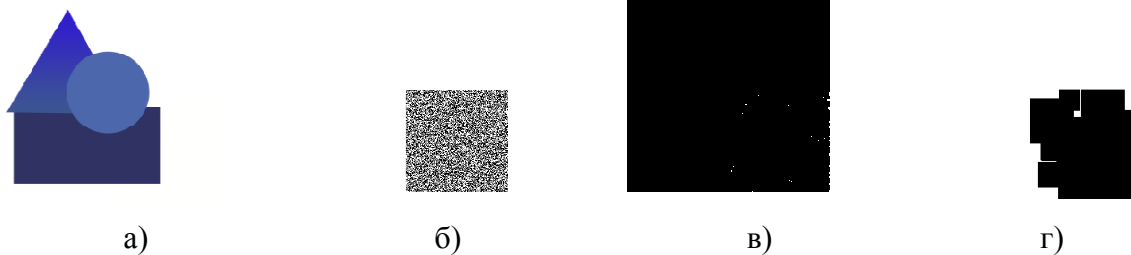


Рис. 3. Результаты работы алгоритма по автоматическому выделению области встраивания для изображения с градиентной заливкой: а) изображение со встроенной вставкой; б) карта встроенных пикселей; в) нулевой слой после предварительной обработки изображения; г) автоматически выделенная область встраивания при $p = 0.5$.

Как видно (рис. 3), алгоритм предварительной обработки изображений позволяет обрабатывать изображения с градиентной заливкой так же эффективно, как и одноцветные изображения. На рис. 4 представлены аналогичные результаты для фотографического изображения.

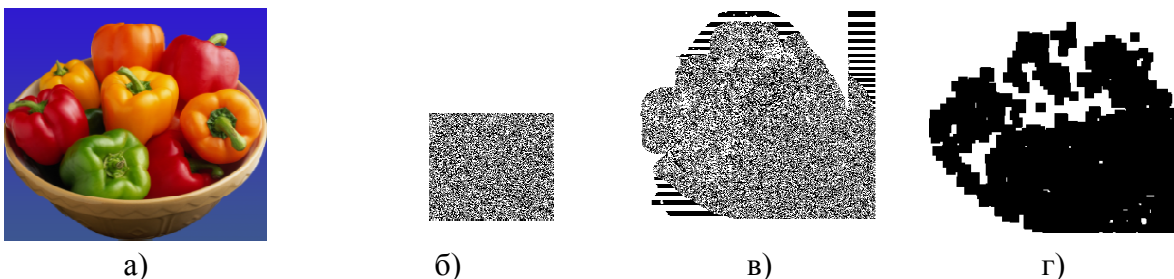


Рис. 4. Результаты работы алгоритма по автоматическому выделению области встраивания для фотографического изображения «Перцы»: а) изображение со встроенной вставкой; б) карта встроенных пикселей; в) нулевой слой после предварительной обработки изображения; г) автоматически выделенная область встраивания при $p = 0.5$.

Как видно (рис. 4), значительное влияние на результаты оказывают области небольшого размера с резким изменением цвета, обусловленным наличием четких границ. Эксперимент на большом количестве изображений показал, что в среднем правильно идентифицируется 88% пикселей, в которые производилось встраивание. Ложные срабатывания составляют в среднем 34% от выделенных пикселей.

Значительно лучших результатов можно добиться, если есть априорная информация о геометрической форме области, в которую производится встраивание. Например, если известно, что область прямоугольная, то необходимо построить прямоугольники, встроенные в закрашенную область (рис. 4г). Прямоугольники наибольших размеров представлены на рис. 5.



Рис. 5. Результаты выделения вписанных прямоугольных областей в контур, полученный на рис. 4г.

Сравнение изображений на рис. 5 показывает, что прямоугольник на рис. 5а практически точно совпадает с областью встраивания на рис. 4б.

Обсуждение результатов и выводы

Таким образом, предложенный в данной статье алгоритм позволяет определять не только наличие стеганографической вставки, но также ее положение и размер. Причем алгоритм основан не на оценке статистических характеристик изображения, а на детальном рассмотрении каждого пикселя, что расширяет область его применимости. Статистические методы оценки дают достаточно надежный результат для изображений, обладающих «средними» характеристиками. Если же структура изображения имеет ряд особенностей, то статистические методы неприменимы. Предложенный подход, основанный на анализе отдельных пикселей, дает хорошие результаты не на всех изображениях. Необходимым условием является наличие пересечения области встраивания с достаточно большой областью градиентной заливки на исходном изображении. Следует отметить, что другие методы определения наличия стеганографических вставок предъявляют, по сути, те же требования. Если на изображении нет областей равномерной заливки, то встраивание сообщения без корреляций между битами не влияет на статистические характеристики нулевого слоя.

Преимуществом предложенного алгоритма является возможность определения размера и положения стеганографической вставки. Компьютерный эксперимент показал, что предложенный алгоритм позволяет верно определить положение в среднем 88% встроенных пикселей, при уровне ложных срабатываний 34%.

Отметим, что эффективность работы алгоритма существенно зависит от количества мелких деталей на изображении. В связи с тем, что нам неизвестны работы, решавшие задачу поиска пикселей с подмененным младшим битом, сравним результаты с близкой, по сути, задачей поиска пикселей, поврежденных импульсным шумом. Эта задача значительно проще, так как изменения, вносимые в изображение, более значительны. Например, алгоритм SD-ROM [12] позволяет определять полностью поврежденные пиксели с эффективностью до 94%, при ложных срабатываниях до 36%.

ЛИТЕРАТУРА

1. Пат. 4939515 США. Digital Signal Encoding and Decoding Apparatus / Е.Н. Adelson. Оpubл. 03.07. 90.
2. *Provos N., Honeyman P.* Detecting steganographic content on the internet // Technical Report CITI 01-1a, University of Michigan, 2001.
3. *Westfeld A., Pfitzmann A.* Attacks on Steganographic Systems: Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and STools and Some Lessons Learned // 3rd International Workshop on Information Hiding, 2000.
4. *Алиев А.Т.* О применении стеганографического метода LSB к графическим файлам с большими областями монотонной заливки // Вестник ДГТУ. – Ростов н/Д. – 2004. – Т. 4, №4 (22). – С.454-460.
5. *Барсуков В.С., Романцов А.П.* Оценка уровня скрытности мультимедийных стеганографических каналов хранения и передачи информации // Специальная техника.– 2000. – № 1.
6. *Кустов В.Н., Параскевопуло А.Ю.* Простые тайны стегоанализа // Защита информации, INSIDE. – 2005. – № 4. – С.72-78.
7. *Голуб В.А., Дрюченко М.А.* Комплексный подход для выявления стеганографического скрытия в JPEG-файлах // Инфокоммуникационные технологии. – 2009. – Т. 7, № 1. – С.44-50.
8. *Абденов А.Ж., Леонов Л.С.* Использование нейронных сетей в слепых методах обнаружения встроенной стеганографической информации в цифровых изображениях // Ползуновский вестник. – 2010. – № 2. – С.221-225.
9. *Westfeld A., Pfitzmann A.* Attacks on Steganographic Systems. Breaking the Steganographic Utilities EzStego, Jsteg, Steganos and S-Tools – and Some Lessons Learned // Lecture Notes in Computer Science. – 2000. – № 1768. – P.61-75.
10. *Жилкин М.Ю.* Стегоанализ графических данных в различных форматах // Доклады ТУСУРа – 2008. – № 2. – С. 63-64.
11. *Монарев В.А.* Сдвиговой метод обнаружения скрытой информации // Вестник СибГУТИ. – 2012. – № 4. – С.62-68.

12. *Abreu E., Lightstone M., Mitra S.K., Arakawa S.K.* A new efficient approach for the removal of impulse noise from highly corrupted images // IEEE Transactions on Image Processing. – 1996. – Vol. C-5, № 6. – P.1012-1025.
13. *Garnett R., Huegerich T., Chui C., He W.* A Universal Noise Removal Algorithm with an Impulse Detector // IEEE Trans Image Process. – 2005. – Vol. C-14, № 11. – P.1747-54.
14. *Сорокин С.В., Щербачев М.А.* Реализация SD-ROM фильтра на основе концепции нечеткой логики // Известия высших учебных заведений. Поволжский регион. – 2007. – №3. – С.56-65.
15. *Белим С.В., Майоров-Зильбернагель А.О.* Алгоритм поиска поврежденных пикселей и удаления импульсного шума на изображениях с использованием метода ассоциативных правил // Наука и образование: электронное научно-техническое издание 12.12.2014. URL: <http://technomag.bmstu.ru/doc/744983.html> (дата обращения: 15.06.17).
16. *Белим С.В., Майоров-Зильбернагель А.О.* Восстановление изображений со статическими пропусками на основе метода ассоциативных правил // Вестник компьютерных и информационных технологий. – 2014. – №12. – С.18-23.
17. *Белим С.В., Кутлуни П.Е.* Выявление поврежденных пикселей на изображении с помощью алгоритма кластеризации // Вестник компьютерных и информационных технологий. – 2016. – №3. – С.3-10.
18. *Белим С.В., Селиверстов С.А.* Использование метода анализа иерархий для выявления импульсного шума в графических объектах // Информационные технологии. – 2015. – №4. – С.251-258.
19. *Загоруйко Н.Г.* Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 1999.

Статья представлена к публикации членом редколлегии А.А. Шелупановым.

E-mail:

Белим Сергей Викторович – sbelim@mail.ru;

Вильховский Данил Эдуардович – vilkhovskiy@gmail.com.