Оригинальная статья
https://doi.org/10.21683/1729-2646-2021-21-3-39-46

Стегоанализ методов скрытия информации в графических контейнерах

Ярослав Л. Грачев[1*], Валентина Г. Сидоренко[1, 2]

[1]Российский университет транспорта, Москва, Российская Федерация, [2]Высшая школа экономики, Москва, Российская Федерация
[*]yaroslav446@mail.ru

Ярослав Л. Грачев

Валентина Г. Сидоренко

Резюме. Цель. *На сегодняшний день остро стоит вопрос защиты от атак на информационные системы, использующие методы стеганографии. Эти атаки опасны использованием самых обычных файлов данных – в особенности, файлов графических форматов – в качестве контейнеров для доставки в систему вредоносного кода или утечки чувствительной информации. Вопросами разработки методов обнаружения такой скрытой информации занимается специальный подраздел стеганографии – стегоанализ, а сами эти методы должны активно использоваться как компьютерными криминалистами при расследовании инцидентов безопасности, так и автоматизированными системами защиты с интегрированными модулями анализа файлов данных на предмет наличия в них вредоносной или опасной информации. Важной особенностью таких работ является необходимость исследования самых различных элементов и форм представления файлов-контейнеров. В частности, необходима проверка не только непосредственно значений цветов пикселей в изображениях, но и частотных характеристик последнего. При этом возникает ряд важных вопросов, связанных с наилучшими практиками применения алгоритмов стегоанализа и составлением корректных выводов по результатам их работы. Целью работы является краткий анализ наиболее важных и актуальных методов стегоанализа как пространственных, так и частотных, а также формирование выводов об их работе и способах анализа предоставляемых ими результатов на основе проведения тестирования программного комплекса, реализующего данные методы.* **Методы.** *Для стегоанализа скрытия, произведенного в наименее значащие биты пикселей изображения, используется метод Хи-квадрат, опирающийся на статистический способ анализа при помощи критерия Пирсона, а также метод Regular-Singular, использующий сигнатурный анализ групп пикселей и средства аналитической геометрии для формирования оценки относительного объема скрытого сообщения. Для обнаружения встраивания информации, произведенного в представлении изображения в частотной области, используется метод стегоанализа Коха-Жао, позволяющий также установить значения параметров, необходимых для извлечения скрытого сообщения.* **Результаты.** *Создан программный комплекс, объединяющий в себе программные реализации проанализированных методов. Проведен ряд тестовых испытаний данного комплекса для оценки результатов работы рассмотренных методов. Для проведения тестирования была составлена выборка изображений различных форматов, в которых различными способами произведено встраивание информации. Результаты анализа файлов выборки использованы для формирования выводов об эффективности проанализированных методов и интерпретации результатов их работы.* **Заключение.** *На основе полученных результатов тестирования сформулированы выводы о точности работы методов стегоанализа при различных объемах встроенного сообщения и методах его скрытия. Выявленные на основе анализа результатов закономерности позволили сформулировать ряд правил получения из результатов анализа выводов об установлении факта обнаружения скрытой информации и оценке ее объема.*

Ключевые слова: *стегоанализ, метод Хи-квадрат, RS-метод, метод Коха-Жао, стегоконтейнер.*

Для цитирования: *Грачев Я.Л., Сидоренко В.Г. Стегоанализ методов скрытия информации в графических контейнерах // Надежность. 2021. №3. С. 39-46. https://doi.org/10.21683/1729-2646-2021-21-3-39-46*

Поступила *12.02.2021 г.* / **После доработки** *16.07.2021 г.* / **К печати** *17.09.2021 г.*

Введение

Файлы и данные графических форматов на сегодняшний день занимают значительную долю сетевого трафика и встречаются повсеместно – не только в виде изображений, размещенных на различных сетевых ресурсах, но и в виде элементов графических интерфейсов и дизайнерских решений, составных элементов более сложных форматов данных и многих других формах.

Одновременно в последние годы показывает заметный рост доля вредоносного программного обеспечения, использующего в атаках на информационные системы методы стеганографии, то есть методы скрытия факта передачи некого тайного сообщения [1]. Как правило, такие атаки в качестве контейнеров для доставки потенциально опасных данных используют именно файлы изображений – в виду возможности скрытия в них больших объемов данных без внесения заметных человеческому глазу искажений.

Стегоанализ представляет собой один из разделов стеганографии, изучающий способы выявления тайно передаваемой информации в анализируемом информационном объекте. Под тайно передаваемой информацией обычно подразумевается информация, скрытая теми или иными стеганографическими методами. Нередко также стегоанализ изучает способы и возможности извлечения скрытой информации при ее детектировании в условиях отсутствия необходимых для этого входных данных [2].

Несмотря на огромное разнообразие алгоритмов скрытия факта передачи информации в графических файлах, почти все они сводятся к нескольким базовым методам стеганографии. К их числу относятся метод скрытия в наименьших значащих битах пикселей, а также метод Коха-Жао, кодирующий информацию в представлении изображения в частотной области [2, 3]. Большинство остальных стеганографических методов являются модификациями или вариациями данных двух.

Для обнаружения информации, скрытой вышеописанными способами, разработан ряд методов стегоанализа, программная реализация которых позволяет автоматизировать процесс анализа и проводить его без участия человека (в отличие от, например, разного рода визуальных атак).

1. Методы стегоанализа

1.1. Метод Хи-квадрат

Метод атаки на стегосистему с помощью анализа критерия Хи-квадрат был предложен и описан Андреасом Вестфелдом и Андреасом Пфитцманом в 1999 году. Данный способ ориентирован на обнаружение информации, скрытой методом наименьших значащих бит (НЗБ).

Первым делом вводится понятие пар значений – «PoV» (pair of values). Каждая пара значений – это пара байт, кодирующих интенсивности цветов и отличающихся лишь на один наименьший значащий бит. Метод НЗБ, фактически, производит преобразования в рамках данных пар, меняя, если это требуется, значение байта с оригинального на «смежное» в его PoV, другое значение той же пары [4].

Идея данного метода анализа основывается на предположении, что в незаполненном контейнере вероятность одновременного появления обоих значений каждой пары мала – а, следовательно, значительно разнится количество значений интенсивностей цветов таких, которые отличаются на наименьший значащий бит [4]. Иначе говоря, для незаполненного контейнера разница количества появлений двух значений одной пары значительна – и так для всех PoV. Количество появлений каждого значения также называется частотой.

В качестве теоретически ожидаемого распределения выбирается последовательность средних арифметических значений частот всех пар. Поскольку при скрытии в НЗБ происходит лишь перераспределение частот внутри пары, то сумма частот пары остается неизменной – и, следовательно, постоянным остается значение среднего арифметического частот внутри PoV.

Под наблюдаемой выборкой подразумевается последовательность, состоящая только из четных или только из нечетных значений всех PoV, что обусловлено необходимостью дальнейшего сравнения распределения этих выборок с теоретически предсказанным распределением в рамках расчетов критерия Хи-квадрат. В этой связи имеет значение исключительно разница между средним значением частот пары и любой из наблюдаемых в этой паре частотой.

Таким образом, теоретически ожидаемая последовательность значений, составленная из средних по паре, является таковой и для незаполненного, и для заполненного контейнеров. Степень сходства распределения наблюдаемой выборки с теоретически ожидаемым распределением, таким образом, становится мерой вероятности того, что произошло стеганографическое встраивание в контейнер. Если оценка по критерию Хи-квадрат позволяет сделать вывод о том, что отклонения от теоретически ожидаемого распределения незначительны, это означает, что с высокой вероятностью имело место встраивание информации.

Данный метод стегоанализа является более эффективным, если применять его алгоритм не целиком к изображению, а к его частям. Чаще всего изображение разбивается либо на блоки, составляющие примерно 1% от площади всего изображения, либо на условные строки матрицы пикселей. Последний способ позволяет увидеть приблизительные начало и конец последовательно встроенного сообщения.

Хотя наименьшие визуально заметные искажения вносят изменения пикселей синего цветового канала, использование методов скрытия в НЗБ позволяет задействовать все три канала сразу в виду низкой детектируемости человеческим глазом изменений цвета при инверсии младших бит пикселя [5]. В этой связи имеет

смысл подсчет среднего значения вероятности наличия скрытой информации по всем трем цветовым каналам одного анализируемого блока изображения. Даже если скрытие производилось лишь в одном канале, средняя вероятность будет заметно отлична от стремящейся к нулю и позволит сделать вывод о встраивании в данный блок пикселей тайной информации.

1.2. RS-метод

Метод обнаружения стеганографически скрытых сообщений Regular-Singular был предложен Андреасом Пфитцманом, Джессикой Фридрих и Мирославом Гольяном в 2001 году. Метод основывается на анализе непересекающихся групп из n смежных пикселей, n – четное [6]. После выделения групп вводится функция регулярности – функция, которая сопоставляет одной группе одно действительное число и показывает регулярность пикселей группы. Значение функции регулярности должно быть тем больше, чем более шумной является группа пикселей.

В качестве функции регулярности выбирается сумма абсолютных разностей (сумма перепадов значений) соседних пикселей группы:

$$f(G) = f(g_1, g_2, \ldots, g_n) = \sum_{i=1}^{n-1} |g_{i+1} - g_i|, \qquad (1)$$

где G – группа пикселей; g_i – i-й элемент группы пикселей G; n – количество пикселей в группе.

После подсчета значений функции регулярности для всех групп анализируемого изображения определяется группа функций переворота («функций флиппинга»). Эти функции соответствуют следующему набору свойств:

1) $\forall x \in P : F(F(x)) = x,\ P = \{0, 255\}$;
2) $F_{\text{пр}} : 0 \leftrightarrow 1, 2 \leftrightarrow 3, \ldots, 254 \leftrightarrow 255$;
3) $F_{\text{обр}} : -1 \leftrightarrow 0, 1 \leftrightarrow 2, \ldots, 255 \leftrightarrow 256$.

Группа функций переворота F состоит из прямой $F_{\text{пр}}$, обратной $F_{\text{обр}}$ и нулевой F_0.

Применение функций переворота эмулирует добавление обратимого шума, усиливая всплески значений в группе и уменьшая ее регулярность [2].

Для применения данных функций (также называемого «флиппингом») к значениям пикселей группы вводится маска, описывающая группу функций флиппинга, применяемых к группе пикселей. Маска – это группа из n значений, каждое из которых выбирается из числа трех: –1, 0 или 1. Каждое из них кодирует одну из трех функций переворота: значение «–1» соответствует $F_{\text{обр}}$, «0» – F_0, «1» – $F_{\text{пр}}$. К пикселю группы при проведении флиппинга, таким образом, применяется соответствующая ему закодированная в маске функция переворота.

После применения функций переворота к группе производится сравнение значений функции регулярности со значениями до флиппинга. На основе данного сравнения группа относится к одному из классов: обычные (*regular*), необычные (*singular*), непригодные (*unusable*):

если $f(F(G)) > f(G)$, то $f \in R$ (группа обычная);
если $f(F(G)) < f(G)$, то $f \in S$ (группа необычная);
если $f(F(G)) = f(G)$, то $f \in U$ (группа непригодная).

Для каждой группы флиппинг производится два раза: с прямой и с инвертированной маской. После проведения операций классификации для всех групп выполняется подсчет ряда количественных характеристик:

– количество обычных групп для маски M: R_M;
– количество необычных групп для маски M: S_M;
– количество обычных групп для инверсной маски $-M$: R_{-M};
– количество необычных групп для инверсной маски $-M$: S_{-M}.

Все описанные характеристики задаются как относительные величины, то есть в процентах от общего числа групп k. Таким образом, $R_M + S_M \le 1$ и $R_{-M} + S_{-M} \le 1$. Основной гипотезой данного метода является предположение о том, что в незаполненном контейнере количества групп одного класса для обычной и инверсной маски равны: $R_M \cong R_{-M}$ и $S_M \cong S_{-M}$.

На рис. 1 представлен типичный вид так называемой RS-диаграммы – графиков значений R_M, S_M, R_{-M} и S_{-M} в зависимости от количества пикселей с инвертированными НЗБ в изображении [6]. Под p на рисунке 1 и далее подразумевается процент заполнения стегоконтейнера скрытым сообщением (относительная длина сообщения). На оси абсцисс отмеряется процент пикселей с инвертированными НЗБ, на оси ординат – процент групп обычных и необычных классов прямой и инвертированной масок (от общего числа групп).

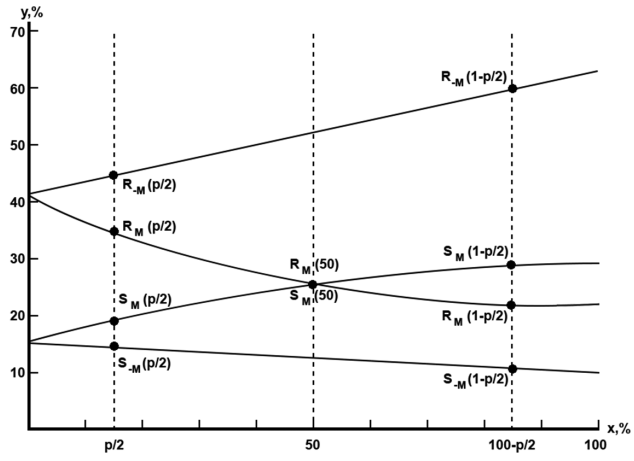


Рис 1. RS-диаграмма

На основе информации о данном типичном поведении графиков количественных характеристик групп производятся дальнейшие вычисления, позволяющие оценить относительную длину скрытого сообщения.

Если относительная длина скрытого сообщения равна p, то, так как p является случайным битовым потоком, в среднем в изображении инвертируется $p/2$ НЗБ. В таком случае, 100%-я заполненность стегоконтейнера ведет к тому, что $p/2$=50%. Инверсия половины НЗБ означает,

что разница между количеством обычных и необычных групп сведется к нулю. Тогда $R_M \cong S_M$, что отражено RS-диаграмме.

Измерения численных характеристик групп соответствуют точкам RS-диаграммы $R_M(p/2)$, $S_M(p/2)$, $R_{-M}(p/2)$ и $S_{-M}(p/2)$. Вычисленные численные характеристики групп для того же изображения после инвертирования всех его НЗБ будут соответствовать точкам $R_M(1–p/2)$, $S_M(1–p/2)$, $R_{-M}(1–p/2)$ и $S_{-M}(1–p/2)$.

Далее RS-метод предлагает аппроксимацию кривых, проходящих через точки $R_{-M}(p/2)$, $R_M(1–p/2)$ и $S_{-M}(p/2)$, $S_{-M}(1–p/2)$ соответственно, прямыми линиями. Кривые, проходящие через точки $S_M(p/2)$, $S_M(1–p/2)$ и $R_M(p/2)$, $R_M(1–p/2)$, аппроксимируются квадратичными параболами с учетом наличия точек пересечения прямых и парабол: соответствующие одной маске парабола и прямая имеют точку пересечения на оси ординат *RS*-диаграммы, а параболы, как сказано выше, пересекаются при $p/2$=50%.

Для оценки относительной длины сообщения p надо решить систему 11 уравнений с 11 неизвестными: по два коэффициента для каждой из двух прямых, по три коэффициента для каждой из двух парабол и значение p. Система включает в себя 8 уравнений прямых или парабол для 8 найденных ранее точек, 2 уравнения точек пересечений парабол с соответствующими прямыми и уравнение для точки пересечения парабол. Решение системы позволяет найти значение p [6].

Ключевая особенность RS-метода состоит в том, что он анализирует количественные характеристики небольших групп пикселей. В связи с чем он, хотя и не способен детектировать область потенциального встраивания, может обнаружить скрытие, произведенное в случайные биты, а не последовательно.

1.3. Стегоанализ метода Коха-Жао

Данный вид анализа предназначен для детектирования встраивания в изображение-контейнер сообщения по методу Коха-Жао. Этот метод ищет информацию, закодированную в частотном представлении изображения.

Представление изображения в частотной области формируется путем вычисления коэффициентов дискретного косинусного преобразования (ДКП), для чего изображение разбивается на блоки размером 8×8 пикселей, после чего над каждым блоком производится двумерное ДКП, в результате чего формируется матрица из 64-х коэффициентов [7].

В полученной матрице коэффициент в левом верхнем углу, соответствующий нулевой частоте (элемент матрицы с индексами (0; 0)), называется *DC*-коэффициентом. Он определяет основной цветовой оттенок (среднюю интенсивность цвета) всего блока. Все остальные полученные коэффициенты называются *AC*-коэффициентами и выражают частоту изменения интенсивности цвета по разным направлениям выбранного блока (по горизонтали и вертикали) [8].

Таким образом, каждая матрица коэффициентов ДКП делится на три подмножества: низкочастотные, среднечастотные и высокочастотные (от левого верхнего к правому нижнему углу матрицы).

Низкочастотные коэффициенты имеют большее влияние на интенсивность цвета пикселей. В связи с этим любые изменения и преобразования над коэффициентами ДКП производятся в средне- или высокочастотных областях.

Одной из важнейших задач, требующих решения при попытке детектировать встраивание методом Коха-Жао, является получение верного вывода о том, за счет каких коэффициентов ДКП происходило встраивание. Поскольку применение метода Коха-Жао предполагает скрытие информации в одном наборов среднечастотных компонент, основные операции анализа выполняются для каждого из таких наборов отдельно [9].

Так как кодирование бит скрываемого сообщения осуществляется за счет разницы абсолютных значений выбранных коэффициентов, сперва следует вычислить абсолютные значения данных разниц для всех блоков изображения:

$$C_i = \left| |D_i(k1,k2)| - |D_i(k2,k1)| \right|, \qquad (2)$$

где $D_i(x,y)$ – значение коэффициента ДКП с индексами (x, y) в i-м блоке.

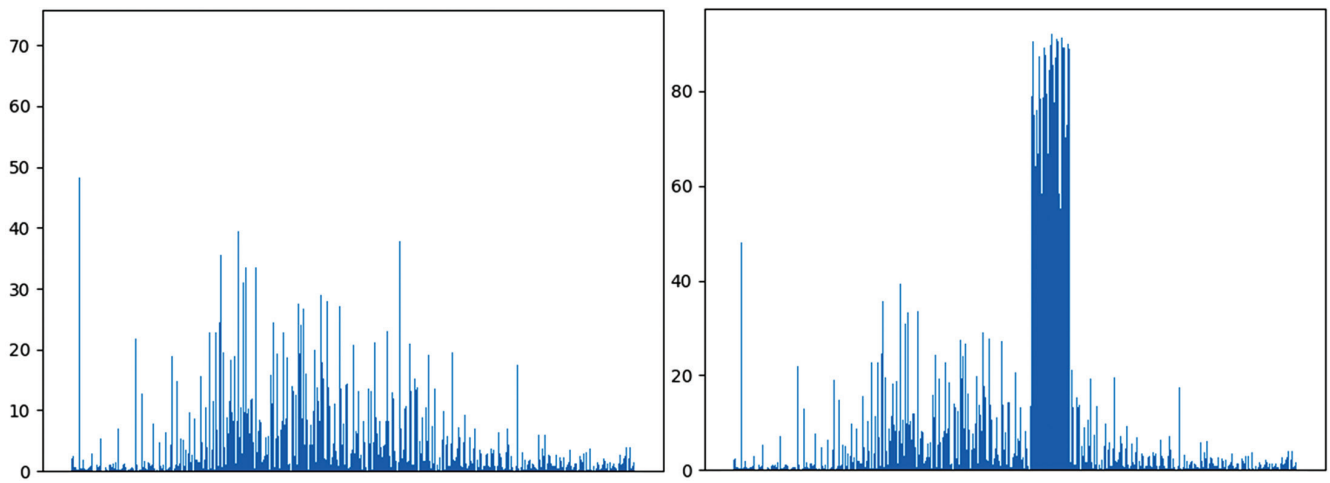


Рис. 2 – Гистограммы значений C_i пустого и заполненного контейнера

На данном этапе вычисляется не просто разность абсолютных значений коэффициентов, а их абсолютные значения – это связано с тем, что кодирование бит определяется преодолением порогового значения P для нуля и $–P$ – для единицы [9].

Несмотря на возможность появления флуктуаций в виде различной величины пиковых значений C_i в блоках, которые не использовались для кодирования бит скрытого сообщения, реально использованные для встраивания блоки отличает сравнительно длинный непрерывный последовательный участок пиковых значений.

На рис. 2 представлены гистограммы значений C_i для одного и того же пустого и заполненного контейнера соответственно. По оси абсцисс расположены индексы блоков, по оси ординат – значения C_i соответствующих блоков.

Как видно, гистограммы зависимости значения C_i от номера блока для пустого и заполненного контейнеров будут отличаться наличием на последнем ступенчатого участка.

Для получения границ области встраивания составляется последовательность модулей разниц значений гистограммы, где два наибольших значения должны соответствовать границам области встраивания.

Однако при больших размерах изображения детектирование скрытого сообщения малого размера (а также, в особенности, закодированное с малым значением порога) может быть значительно затруднено за счет «случайных» пиковых значений последовательности C_i. Это может происходить за счет использования графических форматов, допускающих сжатие, за счет различий в точности вычисления ДКП на кодирующем и анализирующем устройствах, а также за счет возможных искажений и потерь при передаче изображений.

Для того, чтобы уменьшить вероятность сбоев работы алгоритма определения области встраивания из-за подобного «шума», следует произвести предварительный анализ гистограммы с учетом того факта, что встроенное сообщение соответствует непрерывному участку пиковых значений – то есть найти наиболее длинный такой участок.

Таким образом, вывод о наличии встроенного сообщения может быть сделан, а также определены границы скрытого сообщения в значениях индексов блоков. Для вычисления порогового значения кодирования следует найти минимальное из значений C_i в детектированной области встраивания. По данному значению можно произвести извлечение информации.

Описанные процедуры анализа должны быть проведены для каждой из предполагаемых пар коэффициентов ДКП. Как правило, это следующие наборы пар коэффициентов: (3;4) и (4;3), (3;5) и (5;3), а также (4;5) и (5;4) [8]. Из полученных результатов следует выбирать тот, что соответствует наибольшему из обнаруженных значений порога кодирования P.

2. Результаты тестирования эффективности методов

Тестирование вышеописанных методов стегоанализа производилось при помощи разработанного авторами программного комплекса, в рамках которого были реализованы все три алгоритма обнаружения.

В рамках тестирования был проведен анализ выборки из 1600 файлов изображений, в которых одним из трех способов (последовательная вариация метода НЗБ, псевдослучайная вариация метода НЗБ, метод Коха-Жао) был скрыт один из девяти подготовленных текстовых файлов либо информация не скрывалась в принципе.

Для каждого файла выборки анализ производился при помощи всех трех реализованных методов: методом Хи-квадрат, RS и методом стегоанализа скрытия, произведенного по Коха-Жао. Результаты работы методов, представляющие собой относительную (для первых двух) или абсолютную (для последнего) предполагаемую длину скрытого в изображении сообщения, были сравнены с реальными объемами скрытого в каждом контейнере сообщения.

Под средним отклонением подразумевается среднее арифметическое значение модулей разниц между детектированным относительным объемом скрытого изображения и реальным относительным объемом. Средние отклонения методов Хи-квадрат и RS представлены в табл. 1.

Табл. 1. Средние отклонения методов Хи-квадрат и RS

Название метода	Для пустых контейнеров, %	Последовательный НЗБ, %	Псевдослучайный НЗБ, %
Хи-квадрат	0,105	1,411	22,225
RS	4,055	22,996	4,403

Как видно, метод Хи-квадрат практически не имеет ложных срабатываний на пустых контейнерах, а RS-метод показывает для них в среднем крайне низкий уровень заполненности – около 4%.

При этом видно, что метод Хи-квадрат показывает неплохую точность при обнаружении последовательного встраивания в наименее значащие биты (ошибка оценки размера сообщения в среднем составляет всего лишь около 1,4%), а при псевдослучайном встраивании ошибается на объем, составляющий чуть меньше четверти размера контейнера – то есть, в лучшем случае, позволяет лишь обнаружить или заподозрить факт встраивания.

Метод Regular-Singular показывает противоположные результаты, с неплохой точностью определяя псевдослучайное скрытие и крайне плохо оценивая размер последовательно скрытых сообщений.

Для уточнения данных об эффективности оценки размера скрытого сообщения методами Хи-квадрат и

RS были подсчитаны значения средних отклонений для различных степеней относительной заполненности контейнера. Эти значения приведены в табл. 2 и табл. 3.

Табл. 2. Средние отклонения при различных степенях заполненности для метода Хи-квадрат

Степень заполненности	Среднее отклонение при последовательном НЗБ, %	Среднее отклонение при псевдослучайном НЗБ, %
100%	0,609	0,000
90% – 100%	2,107	26,665
80% – 90%	3,678	61,536
70% – 80%	2,851	69,184
60% – 70%	5,013	54,700
50% – 60%	3,248	52,638
40% – 50%	1,783	44,272
30% – 40%	1,761	35,669
20% – 30%	1,635	24,164
10% – 20%	1,019	13,543
0% – 10%	0,366	5,384

Табл. 3. Средние отклонения при различных степенях заполненности для метода RS

Степень заполненности	Среднее отклонение при последовательном НЗБ, %	Среднее отклонение при псевдослучайном НЗБ, %
100%	70,501	11,593
90% – 100%	67,184	8,060
80% – 90%	57,385	4,112
70% – 80%	50,704	1,806
60% – 70%	40,795	1,638
50% – 60%	29,291	1,264
40% – 50%	21,286	2,977
30% – 40%	11,476	2,985
20% – 30%	6,277	2,166
10% – 20%	5,254	3,832
0% – 10%	3,894	3,722

Исходя из представленных результатов, можно выделить несколько закономерностей. Точность определения размера сообщения, скрытого последовательным способом, довольно высока для метода Хи-квадрат при любых объемах скрытого сообщения. Наибольшее отклонение – около 5% – метод показывает на контейнерах, заполненных на 60-70%.

В то же время метод Хи-квадрат способен с крайне высокой точностью определять размер скрытого псевдослучайно сообщения, если он составляет 100% объема контейнера. При заполненности менее 100%, но более 90% данный метод позволяет однозначно идентифицировать встраивание, но ошибка в оценке размера сообщения значительна. При степени заполненности менее 90% метод Хи-квадрат не всегда способен в принципе детектировать скрытие – его среднее отклонение в некоторых случаях примерно эквивалентно относительной длине сообщения. При заполненности ниже 60% метод Хи-квадрат показывает наихудшие результаты при псевдослучайном скрытии, а при заполненности от 60% до 90% способен определить встраивание, но с огромными ошибками в оценке размера сообщения.

Метод RS показывает неплохую точность при обнаружении псевдослучайного встраивания. Однако если степень заполненности менее 10%, он не всегда способен обнаруживать скрытие (величина среднего отклонения около 3,7%, а отклонение для пустых контейнеров составляет около 4,4%). При заполненности от 20% до 80% метод показывает хорошую точность (менее 3% среднего отклонения). Если размер псевдослучайно скрытого сообщения более 80%, RS-метод ошибается уже заметно сильнее – тем больше, чем больше объем встроенной информации. Более высокую величину ошибки метод показывает при заполненности в 10-20%.

При оценке размера последовательно скрытого сообщения метод RS показывает куда худшие результаты и имеет среднее отклонение тем выше, чем выше степень заполненности. При этом отклонение при объеме сообщения менее 10% не позволяет в большинстве случаев однозначно идентифицировать факт скрытия (аналогично псевдослучайно скрытой информации и с учетом среднего отклонения для пустых контейнеров).

При степени заполненности более 10% метод RS будет оценивать размер скрытого сообщения в среднем примерно в 15-30% объема контейнера, позволяя сделать однозначный вывод о факте встраивания, но крайне неточно определяя его объем.

Для оценки результатов работы алгоритма стегоанализа скрытия, произведенного по методу Коха-Жао, было определено количество корректно распознанных изображений, для которых алгоритм правильно оценил размер скрытого сообщения, некорректно распознанных, для которых алгоритм вернул неправильное, но отличное от нуля значение размера сообщения, и нераспознанных, для которых алгоритм вернул 0 в то время, как в изображении действительно скрыта информация.

Вышеописанные характеристики представлены в табл. 4.

Табл. 4. Характеристики различных видов результата анализа скрытия по Коха-Жао

Вид результата	Количество	Относительное количество
Корректно распознаны	277	92,3%
Некорректно распознаны	19	6,3%
Не распознаны	4	1,3%

Все пустые контейнеры были корректно идентифицированы методом Коха-Жао как изображения, не содержащие скрытой информации.

Как видно из таблицы, более, чем в 90% случаев алгоритм смог корректно определить размер встроенного сообщения, а значит, с большой долей вероятности, его можно извлечь. В наибольшем числе остальных случаев размер оценен неправильно, что не позволяет извлечь скрытые данные (по крайней мере, без ручного вмешательства), однако позволяет сделать однозначный вывод о наличии скрытия в частотной области. Не распознанными вовсе остались 1,3% изображений, что может говорить о низком выбранном пороге встраивания в них информации или их высокой зашумленности, вызывающей всплески значений разниц коэффициентов матрицы ДКП, осложняющие идентификацию области встраивания.

Анализ результатов тестирования также показал, что ни в одном из случаев стегоанализ Коха-Жао не выявил скрытие в пространственной области изображения (в наименьших значащих битах), а методы Хи-квадрат и RS не могли показать результат, позволяющий установить факт встраивания в представление изображения в частотной области.

3. Выводы по использованию методов для стегоанализа

Исходя из результатов, представленных выше, для получения однозначного вывода о наличии встраивания информации в младшие биты пикселей требуется проводить анализ и методом Хи-квадрат, и методом RS. Сравнение результатов работы обоих методов позволяет, чаще всего, установить вид встраивания – последовательное или псевдослучайное, а также получить более-менее точную оценку размера сообщения.

Если оценка по Хи-квадрат меньше 0,1%, а по RS – меньше 4%, то встраивание отсутствует и контейнер пуст.

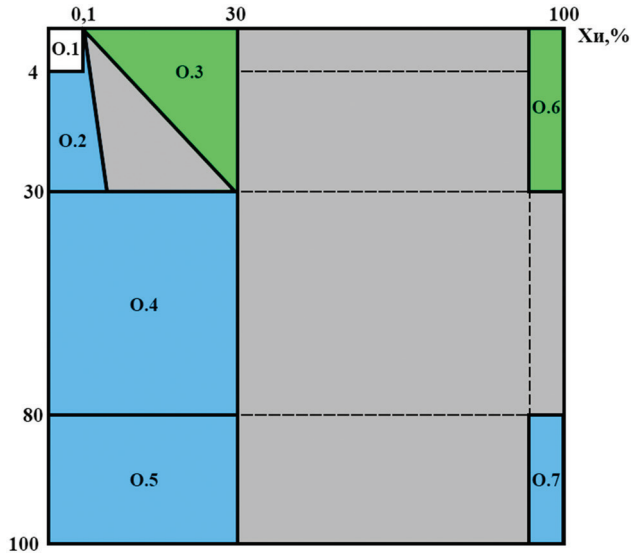
Если оценка по Хи-квадрат близка к 100% – контейнер действительно полон или почти полон. Если оценка по RS при этом составляет около 30% или меньше – информация встроена последовательно, если около 80% и более – псевдослучайно.

Если оценка по RS составляет более 30%, а оценка по Хи-квадрат – не более 30%, то обнаружено псевдослучайное встраивание. Если результат по RS при этом составляет менее 80%, то его можно считать почти точной оценкой размера скрытого сообщения. В противном случае достоверно можно считать, что объем скрытой информации превышает 80%.

Если же оценка RS-методом составляет менее 30%, то ее следует сравнить с оценкой метода Хи-квадрат. Если последняя в разы меньше или близка к нулю, то обнаружено псевдослучайное встраивание с размером сообщения, соответствующим оценке RS-метода. Однако если результат анализа по Хи-квадрат примерно равен оценке RS-метода или превышает ее, то обнаружено последовательное скрытие информации – и, соответственно, размером скрытых данных следует считать оценку размера, предоставленную методом Хи-квадрат.

Использование вышеописанных правил сравнения и анализа результатов в приведенной последовательности позволяет покрыть подавляющее большинство сочетаний результатов методов Хи-квадрат и RS. Получение результатов, не описанных данными правилами, говорит о возникновении флуктуаций, которые могут являться последствиями самых различных факторов (высокая зашумленность изображения, крайне малые его размеры или другие) и требует ручного вмешательства специалиста для формирования окончательного вывода.

Графическое представление полученных выводов представлено на рис. 3. Под *p* подразумевается финальная оценка размера скрытого сообщения. Синим цветом обозначены области, соответствующие псевдослучайному встраиванию (О.2, О.4, О.5, О.7). Зеленым цветом обозначены области, соответствующие последовательному встраиванию (О.3, О.6). Серым цветом обозначены области значений критериев, требующие дополнительного исследования. Область белого цвета (О.1) соответствует пустому контейнеру.



О.1 - Пустой контейнер
О.2 - p ≈ оценка по RS, псевдослучайное встраивание
О.3 - p ≈ оценка по Хи, последовательное встраивание
О.4 - p ≈ оценка по RS, псевдослучайное встраивание
О.5 - p > 80%, псевдослучайное встраивание
О.6 - p ≈ 100%, последовательное встраивание
О.7 - p ≈ 100%, псевдослучайное встраивание

Рис. 3. Графическое представление выводов

Реализованный метод стегоанализа в частотной области позволяет обнаруживать скрытие, произведенное по методу Коха-Жао, в подавляющем большинстве случаев, однако размер скрытого сообщения может быть в ряде ситуаций определен неверно. Следовательно, в некоторых случаях для извлечения скрытой в частотной области информации может потребоваться ручное вме-

шательство в виде анализа специалистом гистограмм разниц коэффициентов ДКП.

Таким образом, данные правила и выводы могут быть использованы в комплексных системах защиты или отдельных модулях анализа изображений, в автоматическом режиме обнаруживая потенциально опасные графические файлы, которые могут нести в себе встроенный вредоносный код или чувствительные данные.

Заключение

Проведенный анализ методов стегоанализа позволил выполнить их программную реализацию в рамках единого комплекса. Тестирование полученного комплекса на выборке из 1600 файлов изображений позволило оценить общую эффективность методов, их средние ошибки в оценке длины встроенных сообщений. Тестирование позволило выявить ряд закономерностей, зависимость результатов работы методов от объемов скрытых сообщений. Выявленные закономерности позволили сформировать по работе рассмотренных методов стегоанализа ряд выводов, позволяющих корректно оценивать получаемые оценки объемов скрытой информации и на их основе делать выводы.

Эти закономерности и выводы позволяют упростить работу компьютерных криминалистов, использующих данные методы стегоанализа при анализе изображений, а также могут служить критериями при исследовании графических файлов на наличие скрытой информации в автоматизированных системах защиты информации или их стегоаналитических модулях, направленных на предотвращение утечек информации по стеганографическим каналам или предупреждение атак, использующих методы стеганографии.

В рамках разработки программного комплекса была произведена адаптация проанализированных методов стегоанализа к современным графическим форматам файлов, использующих схему хранения цветов *TrueColor*. Помимо обнаружения скрытия в НЗБ, оценки области и примерного порога встраивания информации в представление изображения в частотной области, комплекс позволяет также в автоматическом режиме производить попытку извлечения информации даже в случаях с заметно зашумленными изображениями. Выводы о специфике работы рассмотренных методов стегоанализа, сформированные в виде правил сравнения результатов проведения совместного анализа, позволяют тонко настраивать критерии систем защиты или анализаторов трафика для предотвращения инцидентов информационной безопасности.

Библиографический список

1. Варновский Н.П., Голубев Е.А., Логачев О.А. Современные направления стеганографии // Материалы конференции «Математика и безопасность информационных технологий» в МГУ 28-29 октября 2004 г. М.: МЦНМО, 2005. С. 32-64.

2. Аграновский А.В., Балакин А.В., Грибунин В.Г и др. Стеганография, цифровые водяные знаки и стеганоанализ: Монография. М.: Вузовская книга, 2009. 217 с.

3. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. Киев: «МК-Пресс», 2006. 286 с.

4. Westfeld A., Pfitzmann A. Attacks on Steganographic Systems. Dresden University of Technology, Department of Computer Science. Dresden, Germany, 1999. DOI: 10.1007/10719724_5

5. Гонсалес Р.С., Вудс Р.Е. Цифровая обработка изображений. М.: Техносфера, 2005. 1105 с.

6. Fridrich J., Goljan M., Du R. Reliable Detection of LSB Steganography in Color and Grayscale Images. Binghamton University. New York, USA, 2001. DOI: 10.1145/1232454.1232466

7. Syed Ali Khayam. The Discrete Cosine Transform (DCT): Theory and Application. Department of Electrical & Computer Engineering, Michigan State University. Michigan, USA, 2003.

8. Farid H. Digital Image Forensics. Scientific American, 2008. 199 с.

9. Белим С.В., Вильховский Д.Э. Стегоанализ алгоритма Коха-Жао // Математические структуры и моделирование. 2018. № 4(48). С. 113–119. DOI: 10.25513/2222-8772.2018.4.113-119

Сведения об авторах

Ярослав Леонидович Грачев – обучающийся, Российский университет транспорта, Москва, Российская Федерация, e-mail: yaroslav446@mail.ru

Валентина Геннадьевна Сидоренко – доктор технических наук, профессор, профессор кафедры «Управление и защита информации», Российский университет транспорта, профессор Департамента бизнес-информатики, Высшая школа бизнеса, Высшая школа экономики, Москва, Российская Федерация, e-mail: valenfalk@mail.ru

Вклад авторов в статью

Грачев Я.Л. Разработка программного комплекса, реализующего проанализированные методы стегоанализа, составление тестовой выборки, тестирование и анализ результатов, формирование выводов о работе протестированных методов.

Сидоренко В.Г. Анализ методов и принципов стеганографии, обзор методов стегоанализа.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.