8 МЕТОДЫ И СИСТЕМЫ ЗАЩИТЫ ИНФОРМАЦИИ

УДК 003.26.09

РАЗРАБОТКА СТЕГАНОАЛГОРИТМА НА БАЗЕ ФОРМАТНЫХ И ПРОСТРАНСТВЕННЫХ ПРИНЦИПОВ СОКРЫТИЯ ДАННЫХ

А.Г. Коробейников, С.С. Кувшинов, С.Ю. Блинов, А.В. Лейман, С.И. Нестеров

Рассмотрена задача разработки стеганоалгоритма на базе форматных и пространственных методов сокрытия данных. Разработаны алгоритмы встраивания и извлечения информации из стегоконтейнера. Предложена общая схема работы системы скрытой передачи.

Ключевые слова: стеганография, стегоконтейнер, дискретное косинусное преобразование, вейвлет-преобразование.

Введение

В настоящее время в связи с бурным развитием информационных технологий возникают задачи как по криптографической защите данных, так и по сокрытию факта передачи информации [1, 2]. Задача сокрытия данных во внешне безобидных контейнерах с целью скрытой передачи встает, например, при необходимости защиты переписки в сети, которая заставляет абонентов сети иногда использовать методы компьютерной стеганографии (КС) для сокрытия факта переписки. Использование цифровой графики в качестве стегоконтейнеров обусловлено следующими причинами:

- высокая степень распространения цифровой графики;
- популярность и простота процессов обмена цифровыми фотографиями и опубликования их в сети Интернет;
- удобный объем контейнера с точки зрения операций работы с файлами (аудиофайлы и видеофайлы, как правило, в среднем имеют больший объем, чем цифровые изображения);
- особенности системы человеческого зрения, не позволяющие визуально определить наличие незначительных изменений контейнера.

В настоящей работе предложены методы для встраивания сообщений в изображения формата JPEG. Рассмотрим основные методы КС, работающие с данным форматом.

Классификация методов стеганографии

Форматные методы. Форматные методы, по мнению специалистов, не относятся к КС в чистом виде, поскольку не связаны с цифровой обработкой сигналов. Они основаны на избыточности форматов компьютерных данных, например, структуре файлов, IP-пакетов. Цифровые изображения также являются сигналами, имеющими, однако, «застывший» характер. С этой точки зрения при работе с алгоритмом, дописывающим в конец файла JPEG-байты файла RAR, нельзя строго говорить о КС изображений, поскольку это не что иное, как форматный метод в КС.

Предлагается следующее определение: форматный метод в КС изображений – это метод, осуществляющий такое преобразование изображения, при котором вносимые изменения не инициируют визуализацию артефактов встраивания данных и в то же время учитываются и используются в процессе декодирования файла изображения программой просмотра изображений в соответствии со спецификацией формата JPEG.

Данное определение не делает классификацию метода зависимой от конкретной программы просмотра при условии соблюдения программами просмотра спецификации формата JPEG.

Стеганографический алгоритм, представляемый в работе, использует именно такой форматный подход к сокрытию информации в графическом изображении. Отметим, что сокрытие в формате осуществляется не в потоке собственно сообщения, а в сигнальной и служебной информации, не несущей смысловой нагрузки. Это позволяет говорить о том, что в представляемом комбинированном стегоалгоритме форматная составляющая и исходное сообщение развязаны.

Стеганоалгоритмы пространственной области. Алгоритмы данного типа внедряют информацию в области самого изображения. Их преимуществом является то, что для внедрения нет необходимости выполнять вычислительно громоздкие линейные преобразования изображений. Данные внедряются за счет манипуляций цветовыми составляющими или яркостью.

Стеганоалгоритмы области преобразования. Наиболее популярны в стеганографии следующие преобразования:

1. дискретное косинусное преобразование (ДКП);
2. вейвлет-преобразование (ВП).

ДКП используется в алгоритме сжатия JPEG, что является большим стимулом использования ДКП в стеганографии JPEG. ВП, в свою очередь, – основа сжатия в алгоритме JPEG 2000.

ДКП может применяться как ко всему изображению в целом, так и к отдельным блокам пикселей изображения. Обычно же контейнер разбивается на блоки размером 8×8 пикселей. ДКП применяется к каждому блоку, в результате чего получаются матрицы коэффициентов ДКП, также размером 8×8 [2]. Коэффициенты будем обозначать через $c_b(j, k)$, где b – номер блока, (j, k) – позиция коэффициента внутри блока. Если блок сканируется в зигзагообразном порядке (как это имеет место в JPEG), то коэффициенты будем обозначать через $c_{b,j}$. Коэффициент в левом верхнем углу, $c_b(0,0)$, обычно называется DC-коэффициентом. Он содержит информацию о яркости всего блока. Остальные коэффициенты называются AC-коэффициентами. Иногда выполняется ДКП всего изображения, а не отдельных блоков [3].

Рассмотрим процесс внедрения/извлечения информации в области ДКП на примере алгоритма Koch [4]. В данном алгоритме в блок размером 8×8 осуществляется встраивание 1 бита цифрового водяного знака (ЦВЗ). Известны две реализации алгоритма: псевдослучайно могут выбираться два или три коэффициента ДКП. Предлагается модификация алгоритма с двумя выбираемыми коэффициентами (s_i).

Встраивание информации осуществляется следующим образом: для передачи бита 0 добиваются того, чтобы разность абсолютных значений коэффициентов была бы больше некоторой положительной величины ε, а для передачи бита 1 эта разность делается меньше некоторой отрицательной величины –ε :

$$\left|c_b(j_{i,j}, k_{i,1})\right| - \left|c_b(j_{i,2}, k_{i,2})\right| > \varepsilon, \quad \text{если } s_i = 0,$$
$$\left|c_b(j_{i,j}, k_{i,1})\right| - \left|c_b(j_{i,2}, k_{i,2})\right| < -\varepsilon, \quad \text{если } s_i = 1.$$

Таким образом, исходное изображение искажается за счет внесения изменений в коэффициенты ДКП. Для чтения ЦВЗ в декодере выполняется та же процедура выбора коэффициентов, и решение о переданном бите принимается согласно следующему правилу:

$$s_i = 0, \text{ если } \left|c_b(j_{i,j}, k_{i,1})\right| > \left|c_b(j_{i,2}, k_{i,2})\right|,$$
$$s_i = 1, \text{ если } \left|c_b(j_{i,j}, k_{i,1})\right| < \left|c_b(j_{i,2}, k_{i,2})\right|.$$

Разработка стеганоалгоритма

Для успешной работы с JPEG-файлом необходимо знать его структуру, а также алгоритм сжатия с потерями [5]. Поведение программ просмотра JPEG напрямую зависит от структуры файла, и внесение в него изменений стеганографическим алгоритмом не должно провоцировать визуализацию факта изменений. Исходя из структуры формата файлов JPEG, можно утверждать, что следующие маркеры определяют сегменты, не участвующие в JPEG-преобразовании и не влияющие на визуализацию изображения, а потому игнорируемые программами просмотра:

- COM;
- APP15;
- DAC;
- DNL;
- SOF2 – SOF10;
- неспецифицированные сегменты.

Форматная составляющая разработанного метода предполагает работу как раз с такими сегментами. В качестве подготовительной работы были разработаны и реализованы в виде программ-парсеров алгоритмы разбора (парсинга) файлов JPEG и BMP. Возможности данных модулей позволяют получить структурированное представление блоков, для дальнейшего анализа потенциальных мест для внедрения битов сообщения.

Данная система обеспечивает скрытую передачу сообщения с использованием в качестве контейнера передачи неподвижное изображение формата JPEG. Сообщение встраивается в графический файл-контейнер, далее файл доставляется адресату, адресат извлекает из полученного изображения текст сообщения (рис. 1).

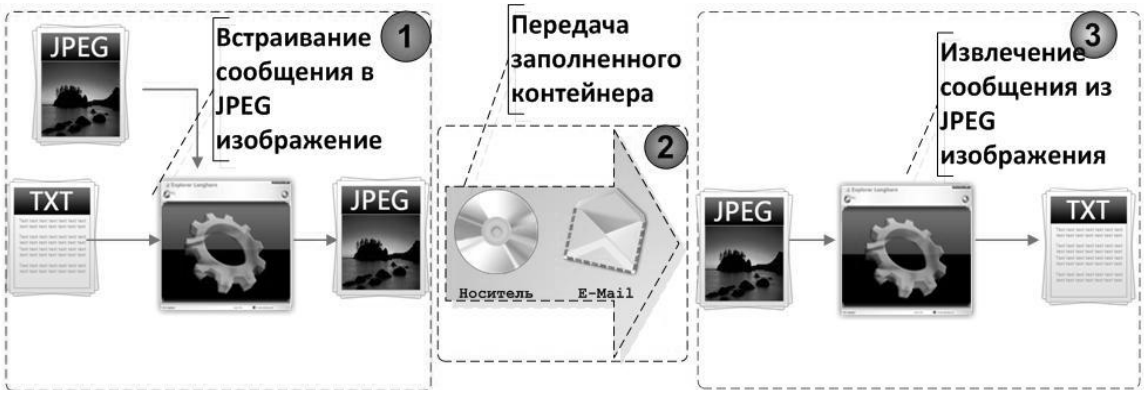

Рис. 1. Общая схема работы системы скрытой передачи

Процессы встраивания и извлечения автоматизированы и выполняются с помощью разработанного приложения. Этап 2 является неконтролируемым, и методы извлечения и встраивания не зависят от канала передачи. В результате получается извлеченное сообщение, идентичное передаваемому, при условии, что передаваемое изображение с внедренным сообщением (стегопосылка) не менялось с момента завершения процесса встраивания данных отправителем до момента начала их извлечения получателем. При несоблюдении указанной целостности программа не гарантирует тот факт, что получатель сможет извлечь полезную для него информацию (сообщение). Схема алгоритма внедрения представлена на рис. 2.

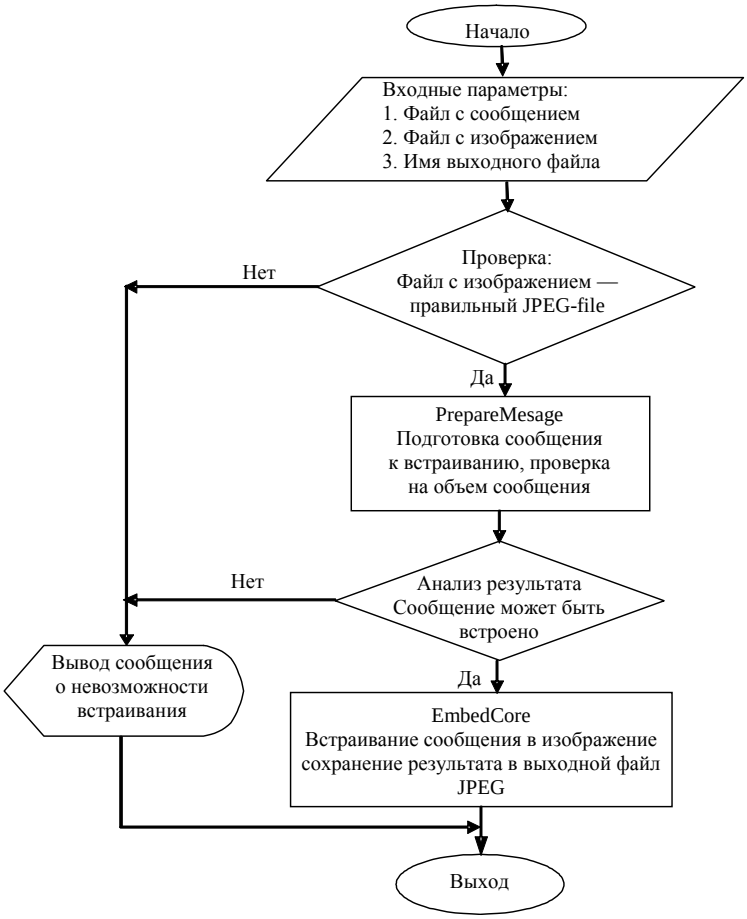

Рис. 2. Общая схема алгоритма внедрения

Система принимает входные данные, проверяет их корректность:

- существование входных файлов на носителе;
- соотношение размеров входных файлов;
- соответствие файла-контейнера формату JPEG.

Алгоритм внедрения включает следующие основные шаги.

На первом этапе производится преобразование потока данных JPEG в поток данных BMP. При этом увеличивается размер потока за счет изменения принципа кодирования информации о цветовых свойствах участков изображения. За счет того, что в формате BMP каждая точка изображения кодируется тремя байтами, отвечающими за вклад основных цветов (R – красного, G – зеленого и B – синего) в целевой цвет точки, изменение размера потока в большую сторону значительно и позволяет встроить необходимый объем информации. Известно, что человеческий глаз наименее чувствителен к изменениям в оттенках синего цвета, поэтому для встраивания используются B-составляющие RGB-структур. Для минимизации объема изменения в алгоритме по умолчанию используется только младший бит такого байта, что до минимума снижает вероятность обнаружения изменения даже на изображениях с большой площадью заливки синего цвета. Далее реализуются режимы, позволяющие использовать 2, 3 и 4 LSB. Простейший способ замены битов – последовательная замена в каждом *b*-байте.

Разработанный механизм компенсирования потерь при межформатных преобразованиях предназначен для успешного извлечения данных на принимающей стороне. В рамках этого механизма после проведения встраивания выполняется следующая последовательность действий. Измененные *b*-байты выделяются из потока и копируются во временный буфер – буфер 1. В другой временный буфер (буфер 2) сохраняется информация о том, как встраивались биты сообщения – в каждый байт подряд или через

один, два, т.д. Эта процедура может не выполняться, если порядок встраивания – последовательная замена бита в каждом *b*-байте. Далее модифицированный поток байтов BMP подвергается JPEG-сжатию с наивысшим коэффициентом качества. Полученный JPEG-поток на этой стадии еще не сохраняется в выходной файл. Для обеспечения возможности последующего извлечения производятся дополнительные действия. Производится попытка извлечения сообщения, которая заключает в себе:

- клонирование потока JPEG во временный поток;
- декодирование временного потока JPEG в поток байтов BMP;
- извлечение модифицированных *b*-байтов, в которые были встроены биты сообщения с учетом информации из буфера 2.

Далее следует анализ извлеченного сообщения, т.е. сравнение извлеченных *b*-байтов с *b*-байтами, сохраненными в буфере 1, и сохранение разностей в новый буфер – буфер 3. Эти разности являются неизбежным результатом потерь при межформатных преобразованиях JPEG – RGB, BMP – JPEG.

При реализации представленного стеганоалгоритма в виде программного обеспечения необходимо учитывать следующие особенности:

- сохранение сообщения в зоне LSB-файла BMP позволяет скрыть большое количество байтов сообщения, однако факт такого сокрытия легко обнаружить;
- сохранение сообщения в игнорируемых сегментах файла JPEG позволяет скрыть большое количество байтов сообщения, однако факт такого сокрытия легко обнаружить;
- количество битов, пригодных для встраивания, зависит от исходного JPEG-файла (характер зависимости определяется JPEG-преобразованием) и линейно зависит от числа LSB, используемых для встраивания.

Заключение

В работе представлен стеганоалгоритм внедрения сообщения с использованием в качестве контейнера передачи неподвижное изображение формата JPEG. Кроме того, разработан механизм компенсирования потерь при межформатных преобразованиях, предназначенный для извлечения данных на принимающей стороне.

Дальнейшее развитие данного направления сопряжено с разработкой систем принятия решений, необходимых для программной реализации стеганографических дополнений к web-браузерам и почтовым клиентам.

Литература

1. Коробейников А.Г., Воробьев А.О., Сидоркина И.Г., Пылин В.В. Анализ криптографической стойкости алгоритмов асимметричного шифрования информации // Изв. вузов. Приборостроение. – 2007. – Т. 50. – № 8. – С. 28–32.
2. Коробейников А.Г., Прохожев Н.Н., Михайличенко О.В., Хоанг З. Выбор коэффициентов матрицы дискретно-косинусного преобразования при построении стеганографических систем // Вестник компьютерных и информационных технологий. – 2008. – № 11. – С. 12–17.
3. Коробейников А.Г., Михайличенко О.В., Прохожев Н.Н. Оценка устойчивости ЦВЗ к внешним воздействиям, внедренных с помощью алгоритмов пространственной области встраивания // Научно-технический вестник СПбГУ ИТМО. – 2008. – Вып. 51. – С. 168–172.
4. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. – Киев: МК-Пресс, 2006. – 288 с.
5. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. – М.: ДИАЛОГ-МИФИ, 2003. – 384 с.

Коробейников Анатолий Григорьевич – Институт земного магнетизма, ионосферы и распространения радиоволн им. Н.В.Пушкова РАН, Санкт-Петербургский филиал, доктор технических наук, профессор, зам. директора, Korobeynikov_A_G@mail.ru

Кувшинов Станислав Сергеевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, кандидат технических наук, доцент, ss.kuvshinov@gmail.com

Блинов Станислав Юрьевич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, stasblino@yandex.ru

Лейман Альберт Владимирович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, аспирант, allxxl@yandex.ru

Нестеров Сергей Игоревич – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, nesterov.serge@gmail.com