УДК 003.26.09

ЦИФРОВЫЕ ВОДЯНЫЕ ЗНАКИ В ГРАФИЧЕСКИХ ФАЙЛАХ

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

Рассмотрена задача генерации цифровых водяных знаков для графических файлов. Представлена математическая модель генерации цифровых водяных знаков. Проанализирован алгоритм внедрения сообщений. Предложено применение разработанного стеганоалгоритма для решения задачи проверки авторского права на конкретный файл мультимедиа.

Ключевые слова: ЦВЗ, форматные методы, стеганоалгоритмы пространственной области, стеганоалгоритмы области преобразования, мультимедиа, медиа-пространство, авторское право.

Введение

Использование цифровых форматов мультимедиа в настоящее стало повсеместным [1]. Но наряду с этим в современном информационном обществе, исследования и разработки в области стеганографии

становятся все более популярными. Это связано с тем, что существуют проблемы управления цифровыми ресурсами и контроля использования прав собственности на компьютерные файлы. Отсюда возникает актуальнейшая задача сокрытия информации в условиях развитой инфраструктуры сетевого общения пользователей – интернет-участников открытого и неконтролируемого взаимодействия в медиа-пространстве.

Сокрытие информации в медиа-пространстве обычно производят при помощи стеганографических алгоритмов. Существует несколько задач, для решения которых используют такие алгоритмы, например:

1. обеспечение тайны переписки (postal privacy);
2. общение удаленных абонентов, обменивающихся цифровыми массивами информации;
3. общение удаленных абонентов в открытых сетевых структурах;
4. достижение скрытности хранимой информации большого объема.

Одним из наиболее эффективных методов защиты мультимедийной информации является встраивание в защищаемый объект невидимых меток – цифровых водяных знаков (ЦВЗ). Название этот метод получил известного способа защиты ценных бумаг, в том числе и денег, от подделки. Термин «digital watermarking» был впервые применен в работе [2]. В отличие от обычных водяных знаков ЦВЗ могут быть не только видимыми, но и (как правило) невидимыми. Невидимые ЦВЗ анализируются специальным декодером, который выносит решение об их корректности.

Стегосистемы ЦВЗ, в частности, должны выполнять задачу защиты авторских и имущественных прав на электронные сообщения при различных попытках активного нарушителя искажения или стирания встроенной в них аутентифицирующей информации. Формально говоря, системы ЦВЗ должны обеспечить аутентификацию отправителей электронных сообщений. Подобная задача может быть возложена на криптографические системы электронной цифровой подписи (ЭЦП) данных, но в отличие от стегосистем ЦВЗ, известные системы ЭЦП не обеспечивают защиту авторства не только цифровых, но и аналоговых сообщений в условиях, когда активный нарушитель вносит искажения в защищаемое сообщение и аутентифицирующую информацию. Иные требования по безопасности предъявляются к стегосистемам, предназначенным для скрытия факта передачи конфиденциальных сообщений от пассивного нарушителя. Также имеет свои особенности обеспечение имитостойкости стегосистем к вводу в скрытый канал передачи ложной информации [3, 4].

Основные требования, предъявляемые к ЦВЗ для графических файлов, представлены в [5, 6].

Математическая модель генерации ЦВЗ

При формальном представлении генерации ЦВЗ в виде математической модели воспользуемся общепринятой записью:

$$\varphi : X \to Y,$$

где φ – отображение (функция); X – область определения; Y – область значений. Введем следующие обозначения: $Y_{\text{ЦВЗ}}$ – множество ЦВЗ; $X_{\text{КЛЮЧ}}$ – множество ключей, $X_{\text{КОНТЕЙНЕР}}$ – множество контейнеров; $X_{\text{СООБЩЕНИЙ}}$ – множество скрываемых сообщений. Тогда генерация ЦВЗ может быть представлена в виде

$$F : X_{\text{КОНТЕЙНЕР}} \times X_{\text{КЛЮЧ}} \times X_{\text{СООБЩЕНИЙ}} \to Y_{\text{ЦВЗ}},$$

или

$$y_{\text{ЦВЗ}} = F\left(x_{\text{КОНТЕЙНЕР}}, x_{\text{КЛЮЧ}}, x_{\text{СООБЩЕНИЙ}}\right),$$

где $y_{\text{ЦВЗ}} \in Y_{\text{ЦВЗ}}$, $x_{\text{СООБЩЕНИЙ}} \in X_{\text{СООБЩЕНИЙ}}$, $x_{\text{КЛЮЧ}} \in X_{\text{КЛЮЧ}}$, $x_{\text{КОНТЕЙНЕР}} \in X_{\text{КОНТЕЙНЕР}}$. Функция F (отображение) может быть произвольной, но на практике требования робастности ЦВЗ накладывают на нее определенные ограничения. Формально это можно записать так:

$$y_{\text{ЦВЗ}} = F\left(x_{\text{КОНТЕЙНЕР}}, x_{\text{КЛЮЧ}}, x_{\text{СООБЩЕНИЙ}},\right) \approx F\left(x_{\text{КОНТЕЙНЕР}} + \varepsilon, x_{\text{КЛЮЧ}}, x_{\text{СООБЩЕНИЙ}}\right),$$

т.е. незначительно измененный контейнер не приводит к изменению ЦВЗ. Кроме того, функция F часто является составной:

$$F = T \mathbf{o} G,$$

где $G : X_{\text{КЛЮЧ}} \times X_{\text{СООБЩЕНИЙ}} \to X_{\text{КОД}}$ и $T : X_{\text{КОНТЕЙНЕР}} \times X_{\text{КОД}} \to Y_{\text{ЦВЗ}}$; $\mathbf{o}$ – суперпозиция.

Функция G может быть реализована при помощи криптографически безопасного генератора псевдослучайных последовательностей с $x_{\text{КЛЮЧ}}$, в качестве начального значения.

Отсчеты ЦВЗ принимают обычно значения из множества {–1, 1}, при этом для отображения {0,1} → {–1, 1} можно применить двоичную относительную фазовую модуляцию ФМн-2 (Binary Phase Shift Keying – BPSK) [7]. Данный вид модуляции нашел очень широкое применение ввиду высокой помехоустойчивости и простоты модулятора и демодулятора. Оператор T модифицирует кодовые слова $X_{\text{КОД}}$, в результате чего получается ЦВЗ – $Y_{\text{ЦВЗ}}$. На этот оператор не накладывают условие существования у

него обратного, так как соответствующий выбор G уже гарантирует необратимость F. Функция T должна быть выбрана так, чтобы незаполненный контейнер $x_{\text{КОНТЕЙНЕР}_0} \in X_{\text{КОНТЕЙНЕР}}$, заполненный контейнер $x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \in X_{\text{КОНТЕЙНЕР}}$ и незначительно модифицированный заполненный контейнер $x'_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \in X_{\text{КОНТЕЙНЕР}}$ порождали бы один и тот же ЦВЗ:

$$T\left(x_{\text{КОНТЕЙНЕР}_0}, x_{\text{КОД}}\right) = T\left(x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}, x_{\text{КОД}}\right) = T\left(x'_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}, x_{\text{КОД}}\right),$$

т.е. она должна быть устойчивой к малым изменениям контейнера.

Процесс встраивания ЦВЗ $y_{\text{ЦВЗ}}(i,j)$ в исходное изображение $x_{\text{КОНТЕЙНЕР}_0}(i,j)$ можно описать как суперпозицию двух сигналов:

$$\Psi : X_{\text{КОНТЕЙНЕР}} \times Y_{\text{ЦВЗ}} \times X_{\text{МАСКА}} \to X_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}},$$

или

$$x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}(i,j) = x_{\text{КОНТЕЙНЕР}_0}(i,j) \oplus x_{\text{МАСКА}}\, y_{\text{ЦВЗ}}(i,j)\, p(i,j),$$

где $X_{\text{МАСКА}}$ – маска встраивания ЦВЗ, учитывающая характеристики зрительной системы человека, служит для уменьшения заметности ЦВЗ; $p(i,j)$ – проектирующая функция, зависящая от ключа; знаком $\oplus$ – обозначен оператор суперпозиции, включающий в себя, помимо сложения, усечение и квантование.

Проектирующая функция осуществляет «распределение» ЦВЗ по области изображения. Ее использование может рассматриваться, как реализация разнесения информации по параллельным каналам. Кроме того, эта функция имеет определенную пространственную структуру и корреляционные свойства, использующиеся для противодействия геометрическим атакам.

Одним из наиболее важных устройств в стегосистеме является стегодетектор. В зависимости от типа он может выдавать двоичные либо М-ичные решения о наличии/отсутствии ЦВЗ (в случае детектора с мягкими решениями). Рассмотрим вначале более простой случай «жесткого» детектора стего. Обозначим операцию детектирования через D. Тогда

$$D : X_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \times Y_{\text{ЦВЗ}} \to \{0,1\},$$

$$D\left(x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}, y_{\text{ЦВЗ}}\right) = D\left(x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}, F\left(x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}, x_{\text{КЛЮЧ}}, x_{\text{СООБЩЕНИЙ}}\right)\right) =$$

$$= \begin{cases} 1, \text{ если } y_{\text{ЦВЗ}} \text{ есть} \\ 0, \text{ если } y_{\text{ЦВЗ}} \text{ нет} \end{cases}.$$

В качестве детектора ЦВЗ обычно используют корреляционный приемник.

Предположим, что у половины пикселей изображения значение яркости увеличено на 1, а у остальных – осталось неизменным, или уменьшено на 1. Тогда:

$$x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} = x_{\text{КОНТЕЙНЕР}_0} + y_{\text{ЦВЗ}},$$

где $y_{\text{ЦВЗ}} = F\left(x_{\text{КОНТЕЙНЕР}_0}, x_{\text{КЛЮЧ}}, x_{\text{СООБЩЕНИЙ}}\right)$. Коррелятор детектора ЦВЗ вычисляет величину

$$x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \cdot y_{\text{ЦВЗ}} = \left(x_{\text{КОНТЕЙНЕР}_0} + y_{\text{ЦВЗ}}\right) \cdot y_{\text{ЦВЗ}} = x_{\text{КОНТЕЙНЕР}_0} \cdot y_{\text{ЦВЗ}} + y_{\text{ЦВЗ}} \cdot y_{\text{ЦВЗ}}.$$

Так как $y_{\text{ЦВЗ}}$ может принимать значения ±1, то $x_{\text{КОНТЕЙНЕР}_0} \cdot y_{\text{ЦВЗ}}$ будет мало, а $y_{\text{ЦВЗ}} \cdot y_{\text{ЦВЗ}}$ будет всегда положительно. По этой причине величина $x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \cdot y_{\text{ЦВЗ}}$ будет очень близка к $y_{\text{ЦВЗ}} \cdot y_{\text{ЦВЗ}}$. Следовательно, можно определить вероятность неверного обнаружения стего, как дополнительную (комплементарную) функцию ошибок от корня квадратного из отношения $y_{\text{ЦВЗ}} \cdot y_{\text{ЦВЗ}}$ («энергии сигнала») к дисперсии значений пикселей яркости («энергия шума»).

Для случая мягкого детектора и закрытой стегосистемы имеем две основные меры похожести:

$\delta = \dfrac{x_{\text{КОНТЕЙНЕР}_0} \cdot x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}}{\left\| X_{\text{КОНТЕЙНЕР}_0} \right\| \left\| X_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}} \right\|}$ – нормированный коэффициент взаимной корреляции и

$\delta = N - \sum_i^N x_{\text{КОНТЕЙНЕР}_0}(i) \cdot x_{\text{КОНТЕЙНЕР}_{\text{ЗАПОЛНЕННЫЙ}}}(i)$ – расстояние по Хэммингу.

В детекторе возможно возникновение двух типов ошибок. А именно, существует вероятность того, что детектор не обнаружит имеющийся ЦВЗ и есть вероятность ложного нахождения ЦВЗ в пустом контейнере (вероятность ложной тревоги). Снижение одной вероятности приводит к увеличению другой. Надежность работы детектора характеризуют вероятностью ложного обнаружения. Система ЦВЗ должна быть построена таким образом, чтобы минимизировать вероятности возникновения обеих ошибок, так как каждая из них может привести к отказу от обслуживания.

Алгоритм внедрения сообщений

При разработке системы скрытой передачи и стеганоалгоритма, представленной авторами в [8], обнаружились трудности, касающиеся межформатных преобразований. Поскольку JPEG – формат сжатия с потерями, то они (потери) в общем случае не позволяют восстановить встроенное сообщение, поскольку восстановление происходит после процедур межформатных преобразований JPEG – RGB BMP – JPEG. В разработанной системе указанная проблема была решена. Процесс внедрения сообщения включает ряд превентивных мер. Непосредственно внедрение происходит по алгоритму, представленному на рис. 1.

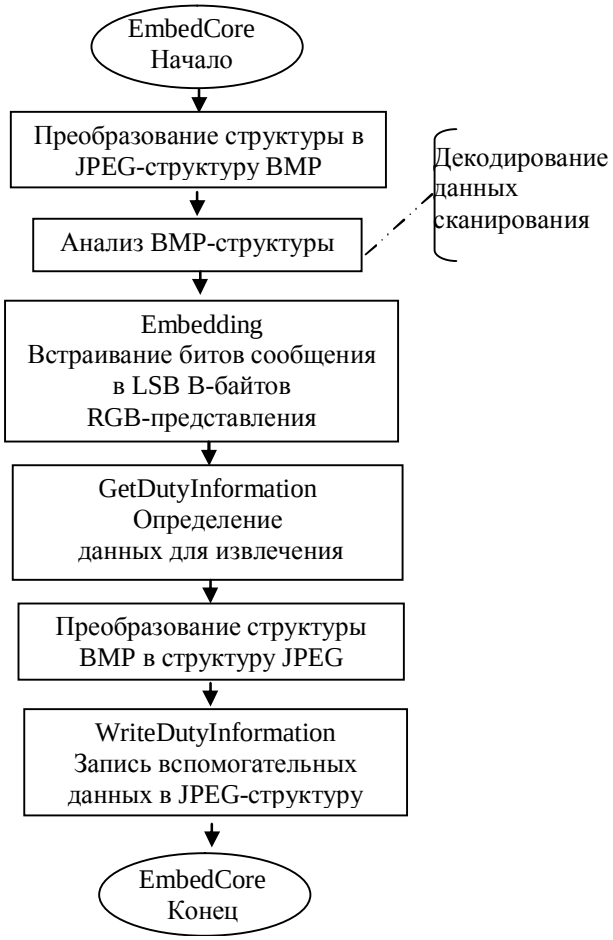

Рис. 1. Блок-схема алгоритма внедрения сообщения

На первом этапе производится преобразование потока данных JPEG в поток данных BMP. При этом увеличивается размер потока за счет изменения принципа кодирования информации об 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 в себе.

Известно, что система человеческого зрения обладает особенностью слабой чувствительности к изменениям в оттенках синего цвета, поэтому для встраивания используются B-составляющие RGB-структур [9]. На самом деле, человеческий глаз также редко может отследить изменения в наименьшем значащем бите красной и зеленой компоненты RGB-структуры. Для минимизации объема изменения пространственной области в режиме работы по умолчанию используется только младший бит такого байта, что до минимума снижает вероятность обнаружения изменения даже на изображениях с большой площадью заливки синего цвета. Простейший способ замены битов – последовательная замена в каждом *b*-байте – представлен на рис. 2.

Поскольку JPEG – формат сжатия с потерями, необходимо учесть этот факт для возможности извлечения сообщения на принимающей стороне. Разработанный механизм компенсирования потерь при межформатных преобразованиях был представлен авторами в [8].

Для определения факта наличия скрытого сообщения пять выходных файлов JPEG с внедренными сообщениями различной длины были проверены хорошо известной программой Stegdetect. Эта програм-

ма детектирования факта встраивания ориентирована на поиск байтовых сигнатур, выдающих стеганографическое вмешательство.

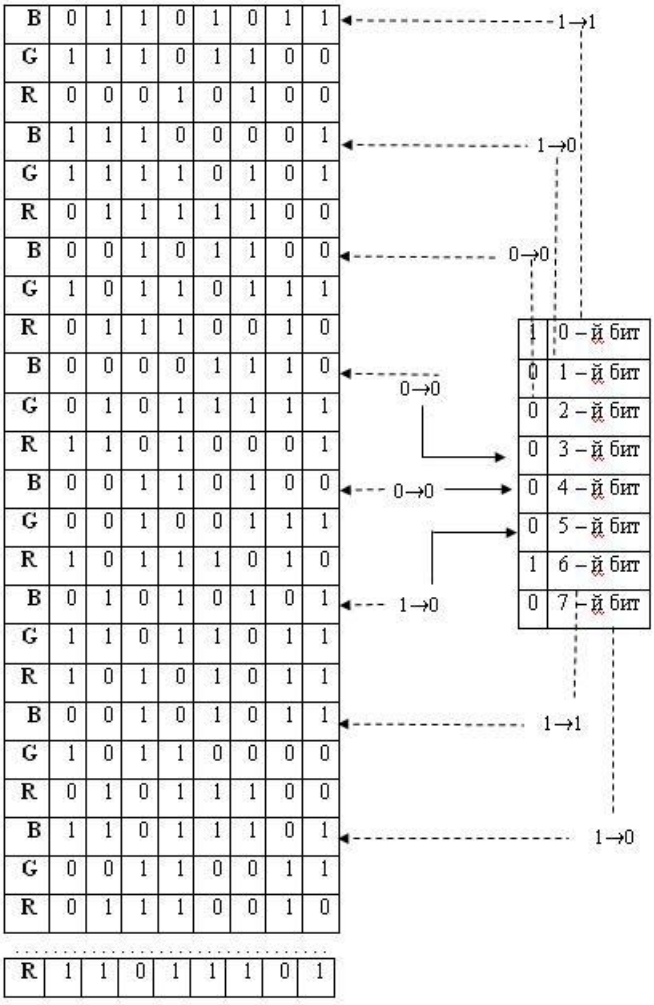

Рис. 2. Последовательная замена битов

В результате проведенного эксперимента программа Stegdetect не смогла корректно указать на факт внедрения данных с помощью разработанного алгоритма. Визуально определить этот факт также не представляется возможным. Визуальный анализ проводился группой людей при наличии оригинального файла JPEG без внедренного сообщения. Определить, в какой из двух файлов встроены данные, им не удалось.

Заключение

Для графических изображений с точки зрения защиты авторского права на их файлы принципиально необходимо реализовывать автоматическое подписывание файлов с целью опубликования информации об авторе. Это может быть текст или иная графическая информация, размещенная в какой-либо (например, нижней) части изображения, однозначно ассоциирующаяся с личностью автора-правообладателя. Такие «метки» служат неопровержимой ссылкой на источник, предоставивший конкретный графический файл. Внедрение в изображения цифровых водяных знаков, позволяющих подтвердить и проверить права разработчика на данный файл мультимедиа, является также эффективной защитной мерой для соблюдения прав интеллектуальной собственности. Такие метки могут быть различным образом расположены в мультимедийном файле и служить противодействием для таких неправомерных деяний, например, как подмена авторства и отказ от авторства. Представленный алгоритм позволяет решить эту задачу. Но необходимо отметить тот факт, что данный алгоритм не проверялся на устойчивость к различного рода искажениям изображения. Эта задача входит в план ближайших научных исследований авторов данной работы.

Литература

1. Сидоркина И.Г., Коробейников А.Г., Кудрин П.А. Алгоритм распознавания трехмерных изображений с высокой детализацией // Вестник Марийского государственного технического университета. – 2010. – № 2 (9). – С. 91–99.

2. Osborne C., van Schyndel R., Tirkel A. A Digital Watermark // IEEE In-tern. Conf. on Image Processing. – 1994. – P. 86–90.
3. Ramkumar M. Data Hiding in Multimedia. – PhD Thesis. – New Jersey Institute of Technology, 1999. – 72 p.
4. Simmons G. The prisoner`s problem and the subliminal channel // Proc. Workshop on Communications Security (Crypto′83). – 1984. – P. 51–67.
5. Барсуков В.С. Романцов А.П. Компьютерная стеганография: вчера, сегодня, завтра. Технологии информационной безопасности XXI века // Специальная техника. – 1998. – № 4 [Электронный ресурс]. – Режим доступа: http://st.ess.ru, свободный. Яз. рус. (дата обращения – 10.10.2012).
6. Bender W., Gruhl D., Morimoto N., Lu A. Tehniques for data Hiding // IBM Systems Journal. – 1996. – V. 35. – № 3&4. – P. 313–336.
7. Скляр Б. Цифровая связь. Теоретические основы и практическое применение: Пер. с англ. – М.: Вильямс, 2003. – 1104 с.
8. Коробейников А.Г., Кувшинов С.С., Блинов С.Ю., Лейман А.В., Нестеров С.И. Разработка стеганоалгоритма на базе форматных и пространственных принципов сокрытия данных // Научно-технический вестник информационных технологий, механики и оптики. – 2012. – № 1 (77). – С. 116–119.
9. Марр Д. Зрение: информационный подход к изучению представления и обработки зрительных образов: Пер. с англ. – М.: Радио и связь, 1987. – 400 с.

Коробейников Анатолий Григорьевич – Санкт-Петербургский филиал Института земного магнетизма, ионосферы и распространения радиоволн им. Н.В.Пушкова РАН, 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

Кутузов Илья Михайлович – Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики, студент, formalizator@gmail.com