УДК 003.26+347.78

Н. П. Шутько, магистрант (БГТУ)

ЗАЩИТА АВТОРСКИХ ПРАВ НА ЭЛЕКТРОННЫЕ ТЕКСТОВЫЕ ДОКУМЕНТЫ МЕТОДАМИ СТЕГАНОГРАФИИ

Объект исследования данной статьи – текстовые документы, обрабатываемые с помощью известных текстовых процессоров и редакторов, а также коды компьютерных программ и файлы баз данных. Раскрывается сущность созданного метода текстовой стеганографии, предназначенного для защиты прав интеллектуальной собственности и основанного на изменении цветовых координат символов текста; описывается созданное программное средство для анализа эффективности предложенного метода. Анализируются некоторые аспекты, относящиеся к проблеме стойкости данного класса стегосистем к атакам.

The object of study of this article – text documents that are processed with the help of some word processors and editors, as well as codes of software and database files. The essence of the devised method of text steganography, intended to protect the copyright of intellectual property and based on the basis of changing the color of coordinates of text characters, is covered; moreover, the author gives a description of software tool created to analyze the effectiveness of the proposed method. Some of the aspects related to the problem of resistance this class stegosystems to attacks is analyzed.

Введение. Защита авторских прав становится менее надежной по мере того, как компьютерные сети все чаще используются для передачи электронных документов. Рассылка документов по сети предполагает, что их может получить большое число адресатов. Это также дает возможность недобросовестным пользователям адаптировать или перерабатывать информацию с целью извлечения коммерческой выгоды. Угроза информационного пиратства стала реальностью.

Одним из направлений решения указанной проблемы является применение современной стеганографии. Стеганография – это искусство передачи скрытого сообщения. Название взято из работы Тритемиуса (1462–1516) – «Стеганография» и происходит от греческого στεγανό-ς,γραφ-ειν, что в переводе означает «тайнопись». Большая часть примеров, которые приводят современные ученые, чтобы объяснить стеганографию, на самом деле пришли из трудов Геродота [1].

Компьютерные технологии придали новый импульс развитию и совершенствованию стеганографии, появилось новое направление в области защиты информации – компьютерная стеганография. Причем, в отличие от криптографии, данные методы скрывают сам факт передачи информации [2].

В то время, как большая часть современной стеганографии сфокусирована на изображениях, аудиосигналах и других цифровых данных, существует также множество письменных источников, в которых можно скрыть информацию. В данной работе акцент делается на применение методов стеганографии в тексте.

Многие существующие методы текстовой стеганографии [3] недостаточно эффективно скрывают сообщения, факт наличия секретной информации (ключ) является очевидным либо почти очевидным. Этот вывод сделан нами на основе сравнительного анализа известных методов текстовой стеганографии с помощью специально разработанного программного средства [4, 5]. Таким образом, актуальной является задача разработки новых методов, повышающих устойчивость к атакам, т. е. снижающих вероятность извлечения сообщения из контейнера («контейнером» называют в данном случае текстовый файл, в который осаждается секретная информация). Указанным осаждаемым сообщением может быть любая информация, известная только автору документа, с помощью которой автор может защитить права интеллектуальной собственности.

Основная часть. Компьютерная стеганография базируется на двух принципах. Первый заключается в том, что файлы, содержащие оцифрованное изображение или звук, могут быть до некоторой степени видоизменены без потери функциональности, в отличие от других типов данных, требующих абсолютной точности.

Второй принцип состоит в неспособности органов чувств человека различить незначительные изменения в цвете изображения или качестве звука, что особенно легко использовать применительно к объекту, несущему избыточную информацию, будь то 16-битный звук, 8-битное или, еще лучше, 24-битное изображение. Если речь идет об изображении, то изменение значений наименее важных битов, отвечающих за цвет пиксела, не приводит к сколь-нибудь заметному для человека изменению цвета.

В основе некоторых методов стеганографии лежат именно данные положения для скрытия сообщения. В графической стеганографии таким методом является, например, метод LSB (Least Significant Bit, метод наименьшего значащего бита), в текстовой – метод пробелов различной длины, метод изменения межстрочных интерва-

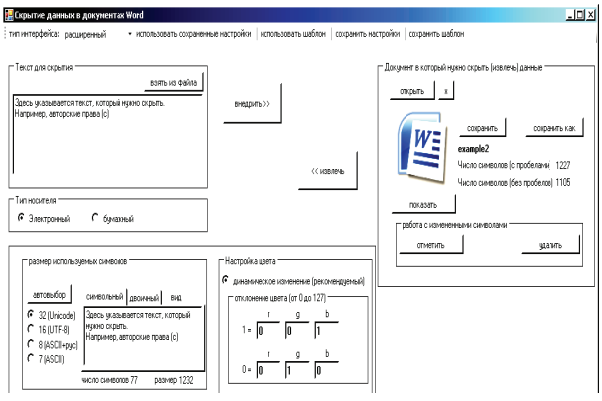
лов и др. [6]. Мы предлагаем новый метод текстовой стеганографии для скрытия секретной (авторской) информации, который использует такой атрибут символа, как цвет.

Сущность предлагаемого метода заключается в следующем. Вся информация в цифровых устройствах представлена в двоичной системе, т. е. записывается с помощью «0» и «1». Мы предлагаем шифровать данные с помощью изменения исходного цвета символа в текстовом документе. То есть, отклоняя значение базового цвета в ту или иную сторону, можно скрыть необходимую информацию (например, увеличение значения цвета будет шифровать «0», а уменьшение – «1»).

Цвет символов в документах текстового процессора Microsoft Word представлен в цветовой модели RGB, поэтому можно изменять значения лишь трех цветовых каналов – красного, зеленого и синего. Отклонение цвета должно быть незначительным для эффективного скрытия секретного сообщения. Это обосновано ограниченными возможностями человеческих органов чувств, в силу чего люди не способны различать незначительные вариации цветов. Что является очень важным, так как стеганография призвана незаметно скрывать сообщение.

Такое добавление своеобразных меток в документ может служить многим целям. Для такого применения, как защита авторских прав, знаки, помещенные в тексте, могут быть использованы как «идентификатор» законного получателя документа. Через анализ любой восстановленный помеченный документ может быть соединен или связан с оригиналом, предназначенным получателю. К тому же, личная информация, которую получатель документа может не обнародовать (например, номер телефона или кредитной карты), также может содержаться в скрываемом тексте.

Для реализации предлагаемого метода изменения цветовых координат было создано программное средство Sword [7]. На рисунке представлено главное окно программного средства.



Главное окно программного средства Sword

Нами были проанализированы программные средства для выполнения подобных стеганографических операций. Далее кратко охарактеризуем некоторые существующие программы.

Один из лучших и самых распространенных продуктов в этой области для платформы Windows 95/NT – программа S-Tools (имеет статус freeware). Программа позволяет прятать любые файлы как в изображениях формата GIF и BMP, так и в аудиофайлах формата WAV. При этом S-Tools – это стеганография и криптография вместе, потому что файл, подлежащий сокрытию, еще и шифруется с помощью одного из криптографических алгоритмов с симметричным ключом: DES (времена которого прошли), тройной DES или IDEA – два последних на сегодня вполне заслуживают доверия. Файл-носитель перетаскивается в окно программы, затем в этот файл перетаскивается файл с данными любого формата, вводится пароль, выбирается алгоритм шифрования. Разглядеть отличия можно только при очень внимательном сопоставлении исходного изображения и изображения, содержащего дополнительную информацию, а если рассматривать только результат, то он вообще никаких подозрений вызвать не может. Причем, не зная пароля, сам факт использования S-Tools установить и доказать нельзя.

Другая распространенная стеганографическая программа – Steganos for Windows 95 (shareware). Она обладает практически теми же возможностями, что и S-Tools, но использует другой криптографический алгоритм (HWY1) и, кроме того, способна прятать данные не только в файлах формата BMP и WAV, но и в обычных текстовых и HTML-файлах, причем весьма оригинальным способом – в конце каждой строки добавляется определенное число пробелов. А еще Steganos добавляет в меню «Отправить» (то, которое появляется при правом щелчке мышью на файле) опцию отправки в шредер, что позволяет удалить файл с диска без возможности его последующего восстановления.

Анализ доступной информации показал, что наше стеганографическое средство Sword не имеет аналогов. Оно позволяет работать с документами любых форматов, способных хранить цвет символов, и с теми, которые можно открыть с помощью OfficeWord: MS Word 93–2010 (*.doc, *.docx), *.rtf (межплатформенный формат хранения размеченных текстовых документов), *.odt (открытый формат документов для офисных приложений). Поддержка данных форматов должна быть доступна из установленного на компьютере пользователя приложения MS Word.

Как видно из рисунка, программа имеет несколько блоков установки. В блоке «*Текст для скрытия*» пользователь может вписать сообщение, которое он хочет скрыть. Программа также позволяет встраивать сообщение, которое уже имеется (функция «*Взять из файла*»).

Далее выбираем тип носителя – электронный или бумажный. Разница заключается в том, что при выборе типа «бумажный» информация будет встраиваться лишь в символы, тогда как в типе «электронный» осаждение будет осуществляться и в пробельных элементах.

Блок «*Размер используемых символов*». С помощью переключателей можно посмотреть, как выбранный текст будет выглядеть в двоичном виде и в текущей кодировке (например, если используются китайские иероглифы, то они будут видны только в Unicode). Изменяя и выбирая этот режим, можно заранее понять, в каком именно виде будет скрыт текст. При смене кодировки исходный текст не теряется. Поэтому если в одной кодировке символы поменялись, можно выбрать другую – и они вернутся. Также в данном блоке отображаются сведения о количестве встраиваемых символов, о необходимом минимальном размере файла, в который будет происходить скрытие.

С помощью «*Настроек цвета*» можно выбрать значение отклонения исходного цвета символа. При этом необходимо учитывать, что чем больше отклонение, тем больше вероятность обнаружения секретной (скрытой) информации.

И, наконец, выбираем документ, в который необходимо произвести скрытие (контейнер).

Первоначально сообщение, которое необходимо спрятать, переводилось в двоичный вид, как указывалось выше. Однако мы решили использовать для сравнения 16-ричную систему представления или даже 32-ричную. Это обусловлено тем, что при такой кодировке стегосообщение можно скрыть в меньшем объеме текста. Например, нам необходимо зашифровать «Слово» в текстовом документе. При этом число символов в документе, необходимых для скрытия, в двоичной системе составляет 79, а в 32-ричной – 15.

Для оценки эффективности предлагаемого метода рассчитаем количество скрываемых бит на каждый бит генерируемого текста.

Для этого возьмем для примера первый абзац введения из электронного варианта данной статьи. Он содержит 52 слова, 387 знаков без пробелов и 438 с пробелами. Рассчитаем среднее количество знаков в слове – $387/52 \approx 7$. Тогда $7/(438 \times 8) = 0{,}0019$ скрываемых бит на каждый бит текста. Чем меньше данное значение, тем меньше вероятность привлечения внимания злоумышленника к тексту, в котором скрыта секретная информация.

В качестве другой оценки эффективности стегосистемы можно, по аналогии с криптографическими системами, использовать понятие и оценку стойкости системы ко взлому – *стеганографическую стойкость*. Однако в нашем случае, в отличие от классических криптосистем, количественная оценка стойкости или безопасности представляется более сложной, поскольку может допускать различные толкования. Возможно, в силу последнего обстоятельства оценка стеганографической стойкости системы, в которой контейнером выступает компьютерный файл, а контентом является текст, практически не изучена.

Под стойкостью стегосистемы (в общем виде) здесь понимается ее способность скрывать от квалифицированного нарушителя факт передачи (внедрения) сообщений, способность противостоять его попыткам разрушить, модифицировать или удалить тайную информацию, а также способность подтвердить или опровергнуть подлинность осажденной (авторской) информации.

Отметим здесь некоторые основные, с нашей точки зрения, моменты, относящиеся к оценке стегостойкости систем. Анализируемый нами класс стегосистем, в частности, должен выполнять задачу защиты авторских или иных имущественных прав на электронные сообщения при различных попытках активного нарушителя исказить или стереть встроенную в них аутентифицирующую информацию (стегосообщение). Формально говоря, наша система должна обеспечить аутентификацию отправителей (создателей, авторов) электронных текстовых сообщений. Подобная задача в криптографических системах возложена на электронную цифровую подпись (ЭЦП). Однако, в отличие от стегосистем, известные системы ЭЦП не обеспечивают защиту авторства в условиях, когда нарушитель вносит искажения в защищаемое сообщение и аутентифицирующую информацию. Вместе с тем, требования по безопасности, предъявляемые к стегосистемам, предназначенным для скрытия факта передачи конфиденциальных сообщений от нарушителя, являются также иными.

Определение 1. Стегосистема является стойкой, если нарушитель не способен обнаружить и тем более читать скрываемое в контейнере сообщение.

Здесь под нарушителем понимается объект системы (человек, устройство или процесс), цель

которого направлена как раз на присвоение или нарушение прав интеллектуальной собственности.

Проводя дальнейшие параллели между стеганографией и криптографией, в нашем случае целесообразно ввести модели нарушителя. Пусть и для систем нашего класса выполняются основные принципы Кергоффа: нарушитель знает полное описание стегосистемы, нарушителю известны вероятностные характеристики авторского сообщения, контейнера, ключа, формируемых стегограмм. Нарушитель обладает необходимыми вычислительными ресурсами, бесконечно большим временем для стегоанализа, и ему известно произвольно большое множество стегограмм. Но нарушителю в нашем случае могут быть не известны две вещи: используемый ключ для данной конкретной стегосистемы (конкретного текстового документа) и то, использует ли данный конкретный автор одинаковый ключ для различных документов.

Определение 2. Если нарушитель не в состоянии установить, содержится или нет скрываемое сообщение в текстовом документе, то назовем такую стегосистему стойкой к атакам пассивного нарушителя или совершенной.

Формальное описание модели требует выполнения новых исследований.

Заключение. В качестве метода для защиты и доказательства прав собственности на документы предложен стеганографический метод, основанный на специальном изменении цветовых координат символов текста (объектом-контейнером является текстовый документ). Разработано программное средство, которое зарегистрировано в Национальном центре интеллектуальной собственности. Описаны некоторые особенности моделирования и оценки стойкости стегосистем к попыткам обнаружения скрытой информации в текстах-контейнерах на основе аналогий между стеганографическими и криптосистемами.

Литература

1. Bennett, K. Linguistic steganography: Survey, analysis, and robustness concerns for hiding information in text. – Purdue Univ., CERIAS Tech. Rep., 2004.

2. Конахович, Г. Ф. Компьютерная стеганография. Теория и практика / Г. Ф. Конахович, А. Ю. Пузыренко. – К.: МК-Пресс, 2006. – 288 с.

3. Пласковицкий, В. А. Управление защитой информации на основе стеганографических методов / В. А. Пласковицкий, Н. П. Шутько // Международная научно-техническая конференция: Автоматический контроль и автоматизация производственных процессов. – Минск: УО «БГТУ», 2012. – С. 283–285.

4. Urbanovich, N. Development, analysis of efficiency and performance in an electronic textbook methods of text steganography / N. Urbanovich // Printing future days: 4th International Scientific Conference on Printing and Media Technology. – Chemnitz, 2011. – P. 189–193.

5. Урбанович, Н. П. Исследование эффективности стеганографических методов скрытия информации в тексте / Н. П. Урбанович // XIII Республиканская научная конференция студентов и аспирантов: Новые математические методы и компьютерные технологии в проектировании, производстве и научных исследованиях. – Ч. 2. – Гомель: УО «ГГУ им. Ф. Скорины», 2011. – С. 27–28.

6. Urbanovich, N. The use of steganographic techniques for protection of intellectual property rights / N. Urbanovich, V. Plaskovitsky // New Electrical and Electronic Technologies and their Industrial Implementation. – 2011. – P. 147–148.

7. Свидетельство о регистрации компьютерной программы Sword v.1.0 / В. А. Пласковицкий, Н. П. Шутько // Реестр Национального Центра интеллектуальной собственности РБ. – 2011. – Запись № 383 от 04.01.2012.

Поступила 07.03.2013