Исследование возможности обнаружения стеганографических вложений при передаче медиафайлов

Кольцов Андрей Сергеевич,
доцент, профессор кафедры информационной безопасности телекоммуникационных систем (ВИ ФСИН России), к.т.н.,
старший лейтенант внутренней службы

В настоящее время широкое применение средств вычислительной техники и многообразие способов обработки информации характерны для большинства учреждений, организаций и предприятий.

Существующие меры защиты используют большое количество средств обнаружения несанкционированных воздействий и реагирования на угрозы. В распоряжении персонала, обеспечивающего безопасную обработку информации, как правило, имеется широкий набор программного обеспечения, которое даже при недостаточном функционале позволяет провести предварительный (предшествующий выбору профессиональных программ) мониторинг, анализ и принять решение о наличие угроз в обрабатываемой информации. Однако применительно к области стеганографии необходимо отметить, что, как правило, отсутствует информация о характере воздействий и происходящих процессах. Что, в свою очередь, не даёт возможности определения степени опасности и приводит к неадекватному реагированию.

Усиление тенденций гуманизации содержания заключённых в местах лишения свободы приводит к тому, что спецконтингент получает возможность персональной работы с компьютером и всемирной паутиной. Ни для кого не секрет, что сегодня огромное количество информации находится в открытом доступе, и сотрудники учреждений УИС не всегда в состоянии отследить трафик.

Так, в соответствии с Концепцией развития уголовно-исполнительной системы до 2020 года [1] между УФСИН России и негосударственными образовательными учреждениями высшего профессионального образования было заключено соглашение об осуществлении образовательной деятельности для получения высшего и дополнительного профессионального образования осужденными, находящимися в местах лишения свободы, а также сотрудниками учреждений ГУФСИН и членами их семей по заочной форме с применением электронного обучения и дистанционных образовательных технологий (в ред. распоряжения Правительства РФ от 23.09.2015 № 1877 р).

В настоящее время организовано обучение осужденных в Нижегородской области (ИК-5, ИК-16), а В Ульяновской области организовано обучение осужденных в Современной гуманитарной академии и ряде других областей.

Таким образом, у осужденных появляется возможность создания нелегальных каналов передачи информации. Например, с помощью программы S-Tools (Steganography Tools), имеющей статус freeware, можно спрятать информацию в графическом или звуковом файле (только WAV). Причем графический файл после этого можно просмотреть, а звуковой — прослушать. Утилита не требует инсталляции, достаточно распаковать архив и запустить файл s-tools.exe. Архив программы занимает всего лишь порядка 280 KB.

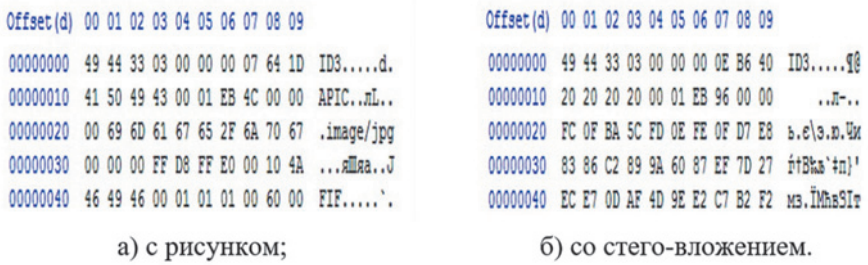
Таким образом, заключенные могут получить доступ к методикам скрытой передачи информации, одной из которых является использование стеганографии.

Методы стеганографии позволяют не только скрытно передавать данные (так называемая классическая стеганография), но и успешно решать задачи помехоустойчивой аутентификации, защиты информации от несанкционированного копирования, отслеживания распространения информации сетями связи, поиска информации в мультимедийных базах данных, и, что не менее важно, осуществлять несанкционированный обмен информацией, спрятанной в файлах электронного документооборота и передающейся по цифровым сетям связи в обход существующей политики информационной безопасности, применяемой в УИС и аппаратнопрограммных средств контроля информации.

В связи с этим становится актуальной проблема контроля наличия нелегальных скрытых вложений в файлах, разрешённых для использования спецконтингентом учреждений ФСИН России.

В реальной среде у заключённых мало времени для организации передачи информации, поэтому рассматривались только mp3 файлы, так как в статических изображениях стеговложения обнаруживаются наиболее просто (видеоизображение не будем рассматривать, так как они занимают большой объём памяти и снижают скорость передачи). Остальные звуковые форматы либо задействуют большее количество времени и памяти, как WAV, либо имеют плохое качество и малую вместимость файловых вложений типа WMA.

MP3 получил наибольшее распространение среди форматов аудио



а) с рисунком; б) со стего-вложением.

Рис. 1. Шестнадцатеричный код звукового файла

и в достаточном количестве имеется на любом компьютере, как в составе программного обеспечения, так и в личных и служебных библиотеках пользователей.

В работе рассматривается способ обнаружения внедренной информации, основанный на записи метаданных файлов формата MP3 — в теги ID3v2.

Структурно тег файла MP3 состоит из общего заголовка и вложенных кадров с собственными заголовками. В спецификации тегов описаны несколько десятков типов кадров, в которые может быть записана текстовая, цифровая и графическая информация. Такую информацию можно считать легальной (или официальной). Как правило, она считывается стандартным программным обеспечением: отображается в свойствах файлов или в окне проигрывателя, доступна программам чтения записи тегов.

Для проведения статистического анализа были выбраны 15 файлов с пустыми заголовками — теги ID3v2.3 (объемом около 1000 байт заполненных нулями). Продолжительность воспроизведения — минимальная, так как анализу подлежат только заголовки.

Предварительные исследования показали, что для получения достоверных данных о порогах принятия решений необходимо провести анализ не менее 10–15 Кбайт вложений. Размер вложений выбран с избытком с учётом снятия информации реализуемого стегопрограммами. Анализу подлежат 20 Кбайт файла с вложениями. Собственные заголовки файлов-вложений проверены и имеют только минимум служебной информации (необходимой для корректного чтения этих файлов официальными программами).

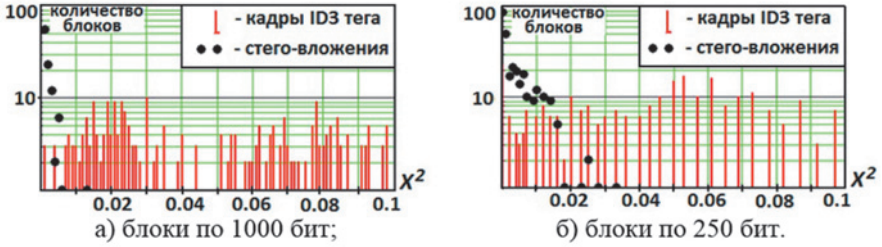
С помощью программ Mp3TagTools v1.2 проведена запись информации в стандартные кадры тегов заголовков и сформированы две группы по 15 файлов:

1) файлы с рисунком, который отображается при воспроизведении стандартным плеером, как обложка альбома (кадр APIC);
2) файлы с текстом, полностью скопированным из файлов. txt и внесенным в стандартный кадр COMM (кадр комментария, содержание которого отображается при просмотре свойств файла) [2].

С помощью программ реализации стеганографических алгоритмов Fox-Secret также сформированы две группы по 15 файлов mp3 с той же информацией, что в предыдущих группах, только вложенной в кадр с неизвестным заголовком. Эти кадры игнорируются программами просмотра ID3-тегов и стандартными плеерами. Все четыре группы файлов mp3 воспроизведены стандартным плеером без ошибок.

Для примера (с помощью программы HxD-Hex-редактор v1.7.7.0), в 16-ричном коде представлены байты файла с рисунком, вложенным в стандартный кадр APIC (рис. 1а). Как заголовок кадра, так и заголовок тега ID3 считываются в правой текстовой колонке (кодировка ANSI). На рисунке 1б представлен 16-ричный код такого же файла со стего-вложением: заголовок тега также отображается в текстовой колонке, а заголовок кадра (в 16-ричном формате — 20202020) — не описан в спецификации тега IDv2.3 и игнорируется всеми программами.

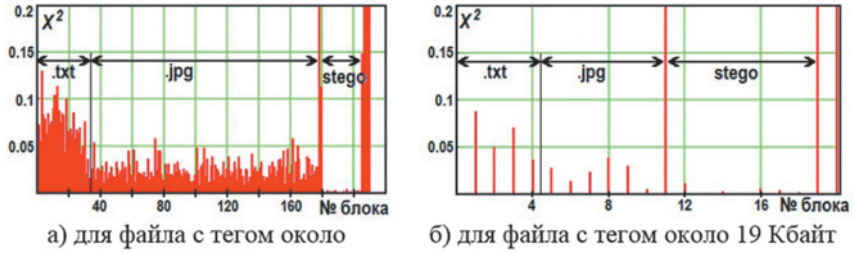
При анализе аудиофайлов следы открытия проявляются при анализе статистических характеристик наименьших значащих битов отсчетов и результаты вычислений по критерию χ^2.



а) блоки по 1000 бит; б) блоки по 250 бит.

Рис. 2. Сравнение χ^2 для официальных вложений и стего

Расчеты и наглядное представление результатов реализовано средствами MathCAD, что потребовало дополнительной подготовки анализируемых файлов: считывание бинарных кодов файла MP3 и их запись в текстовый файл [3, 4]. Такое преобразование реализовано в среде Visual Studio на языке C#. Результатом работы программы является файл с последовательно записанными в один столбец двоичными значениями всех байт исходного файла MP3. Очевидны особенности распределения стего, которые показаны на рис. 2, в сравнении с суммарной статистикой официальных вложений. Для блоков длиной 1000 бит (рис. 2а) характерны малые значения χ^2 до 0,015 (ось абсцисс), более 95% блоков не превышает значений



а) для файла с тегом около 205 Кбайт; б) для файла с тегом около 19 Кбайт

Рис. 3. Распределение значений χ^2 с вложениями в заголовке ID3 (текст, фото и стего), анализируемые блоки по 1000 бит

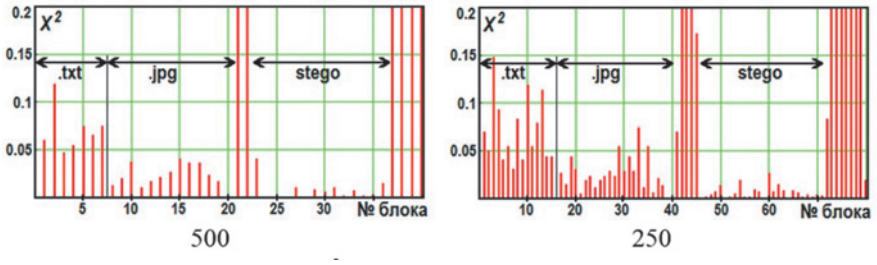


Рис. 4. Распределение значений χ^2 в коротком заголовке ID3 (около 19 Кбайт) при малых длинах блоков

$\chi^2 = 0{,}01$. При сокращении размера блока до 250 бит (рисунок 2б) аналогичным 95 процентным порогом можно считать значение 0,02 [5].

Подобный характер распределения обусловлен использованием криптографических алгоритмов, которые применяются совместно со стего для дополнительного шифрования и имеют близкое к равномерному распределение значений младшего бита [6].

Анализ одного тега с достаточно объемными кадрами (общий размер около 205 Кбайт) блоками по 1000 байт позволяет выделить три интервала с распределением, характерным для стеганографического и официальных вложений (рис. 3). Для тегов с вложениями по 5–8 Кбайт (размер тега 19 Кбайт) подобный анализ возможен, но в общем случае количество отсчетов может быть недостаточным для принятия решения о наличии и характере вложений (рис. 3б).

Для сравнительно коротких заголовков результаты приведены на рис. 4. Границы кадров выражены не так четко, как в предыдущем случае, однако области малых значений χ^2, характерных для стего, могут быть определены. При этом проведение анализа возможно для достаточно коротких участков заголовков до 1500 бит (по блокам 125–500 бит). В этом случае следует говорить не об обнаружении стего, а о «подозрительных» участках файла.

Полученные характеристики распределения для разрешенных способов внедрения (вложения) информации могут быть использованы в качестве математической модели части пустого контейнера. Примерно 10-кратные различия статистики распределений значений χ^2 для официальных и стеганографических вложений, характерные для алгоритма встраивания в ID3-теги, позволяют реализовать первичную сортировку файлов и выделение предполагаемого стеганографического вложения, использующего дополнительное криптографическое шифрование.

Ответственность за обеспечение безопасности информации лежит, как правило, на администраторе сети, который будет определять значения порога обнаружения (в зависимости от важности информации, уровней доступа пользователей и т. д.) и возможные меры по пресечению нарушений: стирание и перезапись подозрительных участков файлов или их сохранение для последующего детального анализа.

Список используемых источников

1. Концепции развития уголовно-исполнительной системы Российской Федерации до 2020 года: распоряжение Правительства Российской Федерации от 14 октября 2010 г № 1772 р. // Собрание законодательства РФ.— 25.10.2010.— № 43 ст. 5544.
2. Nilsson M. ID3v2.3 [Электронный ресурс]: Informal Standard Document. / M. Nilsson. 1999. Режим доступа: http://id3.org/id3v2.3.0, свободный.
3. Барсуков, В. С. Стеганографический камуфляж в джунглях интернета [Текст] / В. С. Барсуков // Специальная техника.— 2005.— N 5.— С. 31–37.
4. Fridrich J., Du R., Long M. Steganalysis of LSB encoding in color images // ICME, 2000.
5. Конахович Г. Ф., Пузыренко А. Ю. Компьютерная стеганография. Теория и практика.— К.: «МК-Пресс», 2006.— 288 с, ил.
6. Душкин А. В., Кравченко А. С., Новосельцев В. И., Смоленцева Т. Е., Сумин В. И. Математические модели и информационные процессы управления сложным объектом: Монография // Воронеж: Научная книга, 2014.— 125 с.— ISBN 978 5 4446 0512 7.