Стеганография для судебного исследователя. Краткий Обзор

Гари Кесслер (Gary C. Kessler)

Резюме

Стеганография - искусство тайного или скрытого письма. Цель стеганографии – тайная передача информации – скрыть существование сообщения от третьих лиц. Эта статья предназначена как углубленное техническое введение в стеганографию для тех, кто незнаком с этой областью. Она адресована судебным исследователям, которые нуждаются в практическом понимании стеганографии, не копаясь в математике, хотя в ней приведены ссылки на некоторые проводимые сейчас исследования для людей, которым нужны дополнительные подробности. Несмотря на то, что эта статья и предоставляет исторический контекст для стеганографии, основное внимание уделяется цифровым приложениям, сосредотачиваясь на скрытии информации в изображениях или звуковых файлах передаваемым по компьютерным сетям. Также, будут представлены примеры программных средств, которые используют стеганографию для скрытия данных в других файлах, и примеры программного обеспечения для обнаружения таких скрытых файлов.

Введение

Стеганография – искусство тайного или скрытого письма. Цель стеганографии – тайная передача информации с целью скрытия существования сообщения от третьих лиц. Этим она отличается от криптографии, искусства секретного письма, которое предназначено сделать сообщение недоступным для чтения третьими лицами, но не скрывает существование секретной коммуникации. Хотя стеганография отделяется и отличается от криптографии, между ними существует много аналогий, и некоторые авторы классифицируют стеганографию как форму криптографии, так как скрытая информация является формой секретного письма (Bauer 2002). Однако эта статья рассматривает стеганографию как отдельную область.

Хотя термин « стеганография» появился только в конце 15- того столетия, использовать стеганографию начали несколько тысячелетий тому назад. В древние времена, сообщения скрывали на задней стороне восковых табличек для письма, писали на желудках кроликов или наносили в виде татуировки на скальпе рабов. Невидимые чернила использовались в течение многих столетий - детьми и студентами для забавы и тайными агентами и террористами при серьёзном шпионаже. Микрофотоснимки и микрофильмы, основные элементы кинофильмов о войне и шпионах, появлялись после изобретения фотографии (Arnold и другие 2003; Johnson и другие 2001; Kahn 1996; Wayner 2002).

Стеганография скрывает тайное сообщение, но не факт того, что две стороны общаются друг с другом. Процесс стеганографии, как правило, включает в себя размещение скрытого сообщения в некотором носителе для транспортировки, который называется контейнер. Секретное сообщение вставляют в контейнер для образования стеганографического носителя. Стеганографический ключ может использоваться для шифрования скрытого сообщения и/или для рандомизации, передаваемых данных, в стеганографической схеме. Таким образом:

носитель_стеганографии = скрытое_сообщение + контейнер + стеганографический ключ

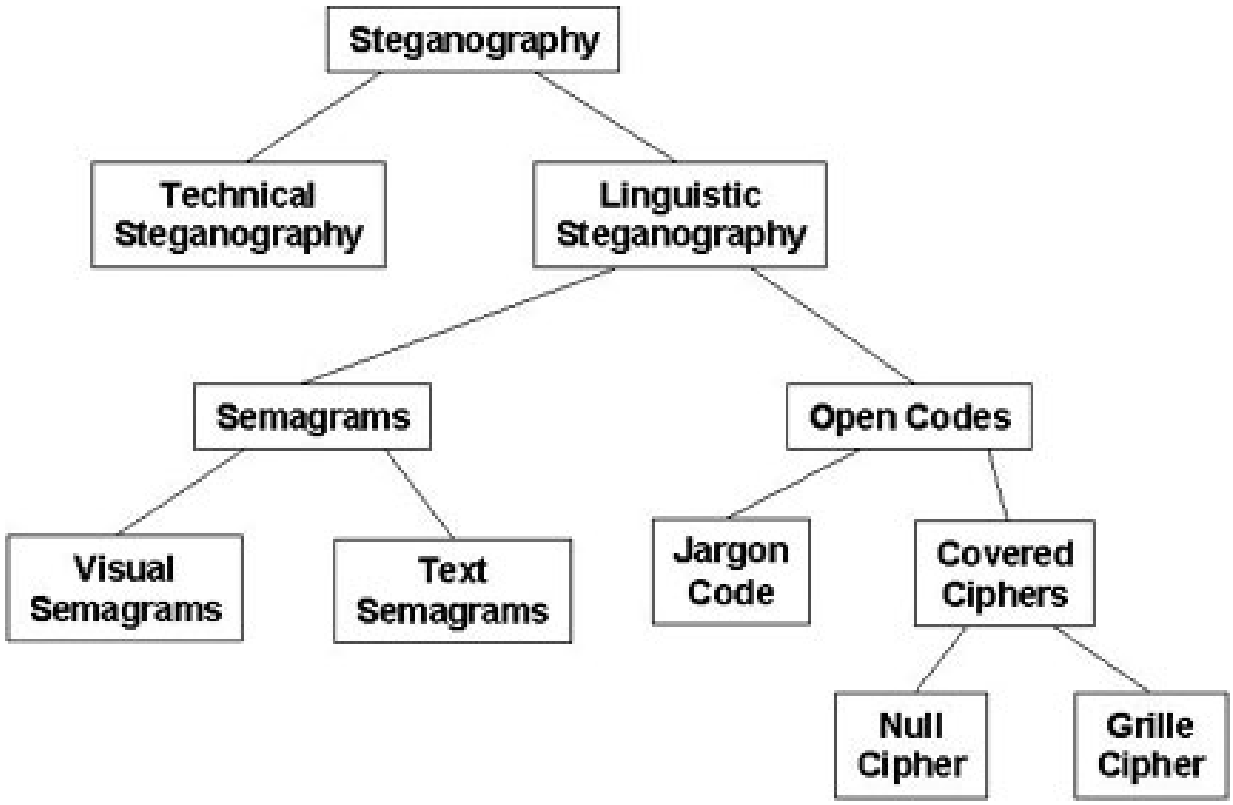

Иллюстрация 1. Классификация стеганографических методов (адаптировано из Bauer 2002).

Иллюстрация 1 показывает общую таксономию стеганографических методов (Arnold и другие 2003; Bauer 2002).

- Техническая стеганография (technical steganography) использует научные методы для скрытия сообщения, такие как, использования невидимых чернил или микрофотоснимков и другие методы сокращения размера;
- Лингвистическая стеганография (linguistic steganography) скрывает сообщение в контейнере некоторыми неочевидными способами и далее классифицируется как семаграммы (semagrams) или открытые коды (open codes);
- Семаграммы скрывают информацию при помощи символов или знаков. Визуальная семаграмма (visual semagrams) использует безобидные на первый взгляд или обычные физические объекты для передачи сообщения, например, каракули или расположения элементов на рабочем столе или веб-сайте. Текстовая семаграмма (text semagrams) скрывает сообщение, изменяя внешний вид текста-контейнера, например, едва различимые изменения в размере или типе шрифта, добавляя дополнительные пробелы или различные завитушки в буквах или рукописном тексте;
- Открытые коды скрывают сообщение в «законном» сообщении-контейнере такими способами, которые не видимы для неподозревающего наблюдателя. Такое сообщение-контейнер иногда называют открытой коммуникацией, тогда как скрытое сообщение – тайной коммуникацией. Эта категория подразделена на жаргонные коды и скрытые шифры;
- Жаргонный код (jargon code), как и предполагает название, использует язык, который понятен одной группе людей, но не имеет смысла для других. Жаргонные коды включают в себя нанесение пиктограмм (символы, используемые для указания присутствия и типа сигнала беспроводной сети [Warchalking 2003]), тайную терминологию, или невинный разговор, который передает особый смысл вследствие того, что факты известны только говорящим. Подкласс жаргонных кодов – коды условных знаков, когда значение передают некие заранее подготовленные фразы;
- Скрытые или замаскированные шифры (covered ciphers) скрывают сообщение в носителе-контейнере так, чтобы его мог восстановить любой, кто знает секрет того, как

оно было скрыто. Шифр “ решетка” (grille cipher) применяет шаблон, который используется, чтобы скрыть сообщение-контейнер. Слова, которые появляются в отверстиях шаблона, являются скрытым сообщением. Нулевой шифр (null cipher) скрывает сообщение согласно некоторому заранее подготовленному набору правил, например, «прочитайте каждое пятое слово» или «посмотрите на третью букву в каждом слове».

Поскольку увеличивающееся количество данных хранится в компьютерах и передаётся по сетям, не удивительно, что стеганография вошла в цифровой век. В компьютерах и сетях стеганографические приложения позволяют кому-нибудь скрыть любой тип бинарного файла в любом другом бинарном файле, хотя графические и звуковые файлы являются сегодня самыми распространенными контейнерами.

В мире компьютерных технологий стеганография предоставляет некоторые очень полезные и коммерчески важные функции, наиболее известные из которых – создание цифровых водяных знаков. В этом приложении, автор может вставить скрытое сообщение в файл, чтобы позднее можно было доказать право интеллектуальной собственности и/или гарантировать целостность содержимого. Художник, например, может поместить оригинал иллюстрации на веб-сайте. Если кто-нибудь другой украдёт файл и заявит, что эта работа является его собственностью, художник может позже доказать право собственности, потому что только он или она может восстановить водяной знак (Arnold и другие 2003; Barni и другие 2001; Kwok 2003). Хотя создание цифровых водяных знаков концептуально подобно стеганографии, оно обычно имеет другие технические цели. Обычно только небольшое количество повторяющейся информации вставляется в контейнер, нет необходимости скрывать информацию о водяных знаках, и водяной знак можно удалить, сохраняя при этом целостность контейнера.

В стеганографии есть ряд незаконных приложений; наиболее известны те, которые скрывают записи о незаконной деятельности, финансовом мошенничестве, индустриальном шпионаже и обмен информацией между членами преступных или террористических организаций (Hosmer и Hyde 2003).

Нулевые шифры

Исторически, нулевые шифры – это способ скрыть одно сообщение в другом без использования сложного алгоритма. Одни из самых простых нулевых шифров показаны в классических примерах ниже:

PRESIDENT'S EMBARGO RULING SHOULD HAVE IMMEDIATE NOTICE. GRAVE SITUATION AFFECTING INTERNATIONAL LAW. STATEMENT FORESHADOWS RUIN OF MANY NEUTRALS. YELLOW JOURNALS UNIFYING NATIONAL EXCITEMENT IMMENSELY.

APPARENTLY NEUTRAL'S PROTEST IS THOROUGHLY DISCOUNTED AND IGNORED. ISMAN HARD HIT. BLOCKADE ISSUE AFFECTS PRETEXT FOR EMBARGO ON BYPRODUCTS, EJECTING SUETS AND VEGETABLE OILS.

Немецкое Посольство в Вашингтоне, округ Колумбия, отправило эти сообщения в телеграммах в свои штаб-квартиры в Берлине во время первой мировой войны (Kahn 1996). В результате чтения первой буквы каждого слова в первом сообщении или второй буквы каждого слова во втором сообщении получится следующий скрытый текст:

PERSHING SAILS FROM N.Y. JUNE 1
(ПЕРШИНГ ОТПЛЫВАЕТ ИЗ НЬЮ-ЙОРКА ПЕРВОГО ИЮНЯ)

В Интернете спам является потенциальным носителем тайных сообщений. Рассмотрите следующий пример:

Dear Friend , This letter was specially selected to be sent to you ! We will comply with all removal requests ! This mail is being sent in compliance with Senate bill 1621 ; Title 5 ; Section 303 ! Do NOT confuse us with Internet scam artists . Why work for somebody else when you can become rich within 38 days ! Have you ever noticed the baby boomers are

more demanding than their parents & more people than ever are surfing the web ! Well, now is your chance to capitalize on this ! WE will help YOU sell more & SELL MORE . You can begin at absolutely no cost to you ! But don't believe us ! Ms Anderson who resides in Missouri tried us and says "My only problem now is where to park all my cars" . This offer is 100% legal . You will blame yourself forever if you don't order now ! Sign up a friend and your friend will be rich too . Cheers ! Dear Salaryman , Especially for you - this amazing news . If you are not interested in our publications and wish to be removed from our lists, simply do NOT respond and ignore this mail ! This mail is being sent in compliance with Senate bill 2116 , Title 3 ; Section 306 ! This is a ligitimate business proposal ! Why work for somebody else when you can become rich within 68 months ! Have you ever noticed more people than ever are surfing the web and nobody is getting any younger ! Well, now is your chance to capitalize on this . We will help you decrease perceived waiting time by 180% and SELL MORE . The best thing about our system is that it is absolutely risk free for you ! But don't believe us ! Mrs Ames of Alabama tried us and says "My only problem now is where to park all my cars" . We are licensed to operate in all states ! You will blame yourself forever if you don't order now ! Sign up a friend and you'll get a discount of 20% ! Thanks ! Dear Salaryman , Your email address has been submitted to us indicating your interest in our briefing ! If you no longer wish to receive our publications simply reply with a Subject: of "REMOVE" and you will immediately be removed from our mailing list . This mail is being sent in compliance with Senate bill 1618 , Title 6 , Section 307 . THIS IS NOT A GET RICH SCHEME . Why work for somebody else when you can become rich within 17 DAYS ! Have you ever noticed more people than ever are surfing the web and more people than ever are surfing the web ! Well, now is your chance to capitalize on this ! WE will help YOU turn your business into an E-BUSINESS and deliver goods right to the customer's doorstep ! You are guaranteed to succeed because we take all the risk ! But don't believe us . Ms Simpson of Wyoming tried us and says "Now I'm rich, Rich, RICH" ! We assure you that we operate within all applicable laws . We implore you - act now ! Sign up a friend and you'll get a discount of 50% . Thank-you for your serious consideration of our offer .

Данное сообщение похоже на типичный спам, который обычно игнорируется и удаляется. Это сообщение было создано на веб-сайте “Spam Mimic” («Имитатор Спама»), который превращает короткое текстовое сообщение в текстовый блок, похожий на спам, используя основанную на грамматике идею имитации, впервые предложенную Питером Вэйнером (spam mimic 2003; Wayner 2002). Читатель ничего не узнает, смотря на пробелы между словами или орфографические ошибки в сообщении. Нули и единицы закодированы посредством выбора слов. В скрытом сообщении в спам-контейнере выше говорится:

Meet at Main and Willard at 8:30
(Встречаемся у Майна и Уилларда в 8:30)

Не нужны специальные инструменты или навыки скрытия сообщений в цифровых файлах, используя изменения нулевого шифра. Изображение или текстовый блок могут быть скрыты под другим изображением в файле программы **PowerPoint**, например. Сообщения могут быть скрыты в свойствах файла текстового редактора **Word**. Сообщения могут быть скрыты в комментариях к веб-страницам или в других элементах форматирования, которые игнорируются браузерами (Artz 2001). Текст может быть скрыт как штриховой рисунок в документе, при помощи раскрашивания текста тем же цветом, что и фон, и помещая другой рисунок на передний план. Получатель может восстановить скрытый текст, изменив его цвет (Seward 2004). Всё это, несомненно, низко-технологичные алгоритмы, но они могут быть очень эффективными.

Цифровые графические и аудио файлы

Многие распространённые методы цифровой стеганографии используют графические или звуковые файлы в качестве носителей-контейнеров. Следовательно, будет полезно сделать обзор кодирования графических и звуковых файлов перед обсуждением того, как стеганография и стегоанализ работают с этими контейнерами.

Иллюстрация 2 показывает цветовой куб модели RGB, обычное средство для представления конкретного цвета по относительной интенсивности его трех составляющих цветов – красного, зеленого и синего – каждый со своей собственной осью (moreCrayons 2003). Отсутствие всех цветов даёт черный цвет, показанный как пересечение трех цветных осей в нулевой точке. Смесь 100 процентов красного, 100 процентов синего, и отсутствия зеленого цвета образует пурпурный цвет; голубой цвет – это 100 процентов зеленого и 100 процентов синего цвета без красного; а соединение 100 процентов зеленого и 100 процентов красного цвета без синего образуют желтый. Белый цвет – это наличие всех трех цветов.

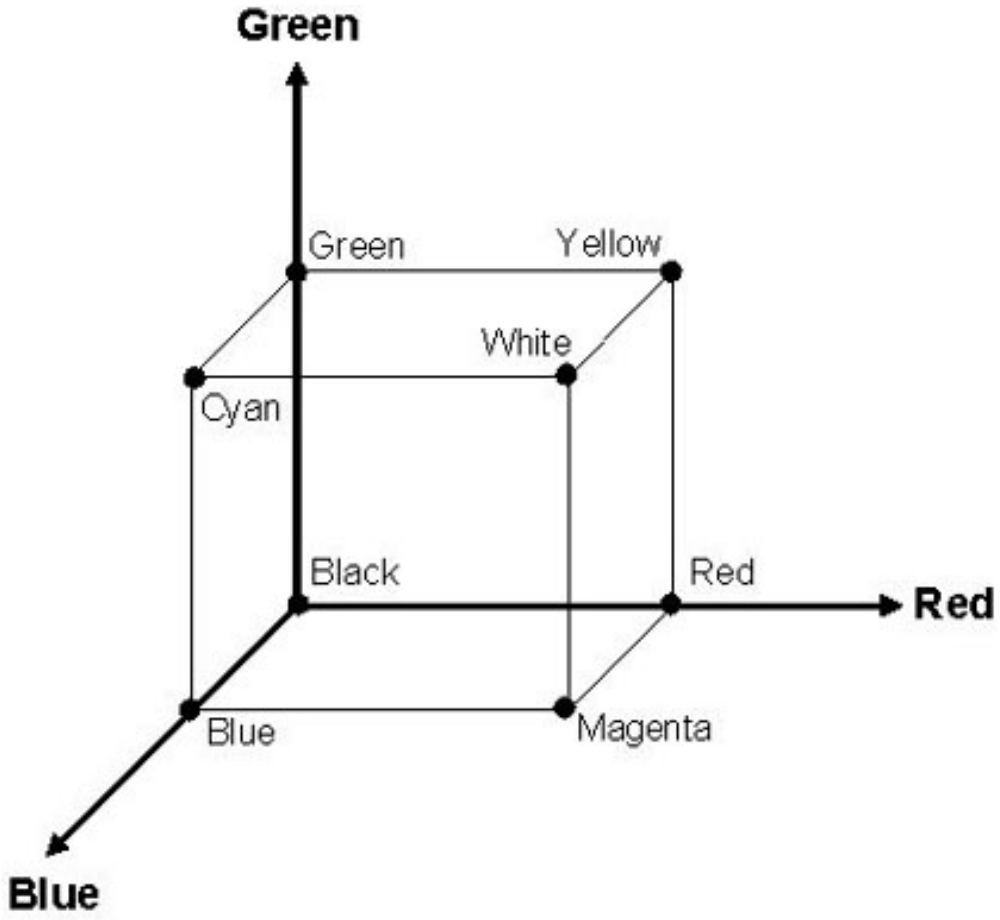

Иллюстрация 2. Цветовой куб модели RGB.

Иллюстрация 3 показывает уровни интенсивности (яркости) модели RGB нескольких случайных цветов. Каждый компонент модели RGB задан одним байтом, поэтому значения каждой интенсивности цвета могут меняться от 0 до 255. Данный конкретный оттенок обозначен уровнем красного 191 (BF в hex-формате), уровнем зеленого 29 (1D в hex-формате) и уровнем синего 152 (98 в hex-формате). Затем, один пиксель пурпурного цвета будет закодирован, используя 24 бита, как 0xBF1D98. Эта 24- битовая схема кодирования поддерживает 16 777 216 (2^24) уникальных цветов (Curran и Bailey 2003; Johnson и Jajodia 1998A).

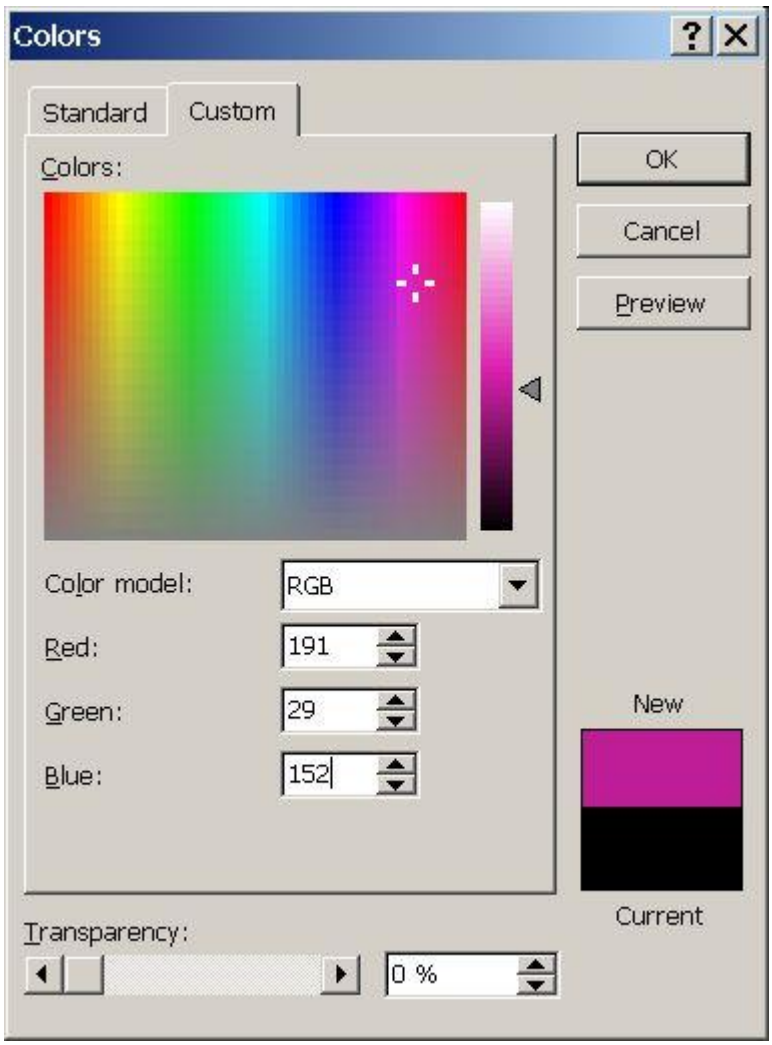

Иллюстрация 3. Это диалоговое окно выбора цвета показывает уровни красного, зеленого и синего (RGB) данного выбранного цвета.

Большинство цифровых графических приложений сегодня поддерживает 24-битовую реалистичную цветопередачу, где каждый элемент изображения (пиксел) закодирован в 24 битах, включая три байта модели RGB, как описано выше. Другие приложения кодируют цвет, используя восемь битов/пикселей. Эти схемы также используют 24- битовую реалистичную цветопередачу, но применяют палитру, которая определяет, какие цвета используются в изображении. Каждый пиксел закодирован в восьми битах, где значение указывает на 24- битовую запись цвета в палитре. Этот метод ограничивает уникальное число цветов в данном изображении до 256 (2^8).

Кодирование выбора цвета заметно влияет на размер изображения. Изображение размером 640 X 480 пиксела, использующее восьмибитовую кодировку цвета, займёт приблизительно 307 Кб (640 X от 480 до 307 200 байтов), тогда как изображению размером 1400 X 1050 пикселей, использующему 24- битовую кодировку цвета, потребуется 4,4 Мб (1400 X 1050 X 3 = 4 410 000 байтов).

Цветные палитры и восьмибитовый цвет обычно используются с такими форматами изображений, как формат графического обмена (GIF) и формат растровых графических файлов (BMP). Обычно полагают, что форматы GIF и BMP предлагают сжатие без потерь, так как изображение, восстановленное после кодирование и сжатия – до единого бита идентично оригинальному изображению (Johnson and Jajodia 1998A).

Графический формат Объединенной группы экспертов по машинной обработке фотографических изображений (JPEG), использует дискретное косинусное преобразование, а не попиксельное кодирование. В формате JPEG, изображение разделено на 8X8 блоков для каждого отдельного компонента цвета. Цель состоит в том, чтобы

найти блоки, где количество изменения в значении цвета пикселя низко. Если уровень цветов слишком различен, блок делится на 8X8 подблоков, пока уровень цвета не будет достаточно низок. Каждые 8X8 блоков (или подблоков) преобразуются в 64 коэффициента дискретных косинусных преобразований, которые приблизительно равны светимости (яркость, оттенок и контраст) и хроматическим данным (цвету) той части изображения. Обычно считают, что формат JPEG является сжатием с потерей данных, так как изображение, восстановленное из сжатого файла JPEG, является хорошим приближением, но не идентично оригиналу (Johnson и Jajodia 1998A; Monash University 2004; Provos и Honeyman 2003).

Аудио кодирование включает в себя преобразование аналогового сигнала в поток битов. Аналоговый звук – голос и музыка – представлен гармоническими волнами различных частот. Человеческое ухо может номинально слышать частоты в диапазоне 20 - 20 000 циклов/секунду (Герц или Гц). Звук – аналоговый, это означает, что он является непрерывным сигналом. Хранение звука в цифровой форме требует, чтобы непрерывная звуковая волна была преобразована в ряд фрагментов, которые могут быть представлены последовательностью нулей и единиц.

Аналого-цифровое преобразование достигается выборкой аналогового сигнала (при помощи микрофона или другого звукового детектора) и преобразованием этих фрагментов в уровни напряжения. Затем уровень напряжения или сигнала преобразовывается в цифровое значение, используя схему, которая называется кодово-импульсная модуляция. Устройство, которое выполняет это преобразование, называется кодер-декодер или кодек.

Кодово-импульсная модуляция обеспечивает только приближение оригинального аналогового сигнала, как показано на Иллюстрации 4. Если аналоговый уровень звукового давления измерить, например, на уровне 4,86, он будет преобразован в пять в кодово-импульсных модуляций. Это называют ошибкой дискретизации. Различные аудио приложения определяют различное число уровней кодово-импульсных модуляций, чтобы эта « ошибка» была почти необнаружима человеческим ухом. Телефонная сеть преобразовывает каждый голос абонента в восьмибитовое значение (0-255), тогда как музыкальные приложения обычно используют 16- битовые значения (0-65535) (Fries и Fries 2000; Rey 1983).

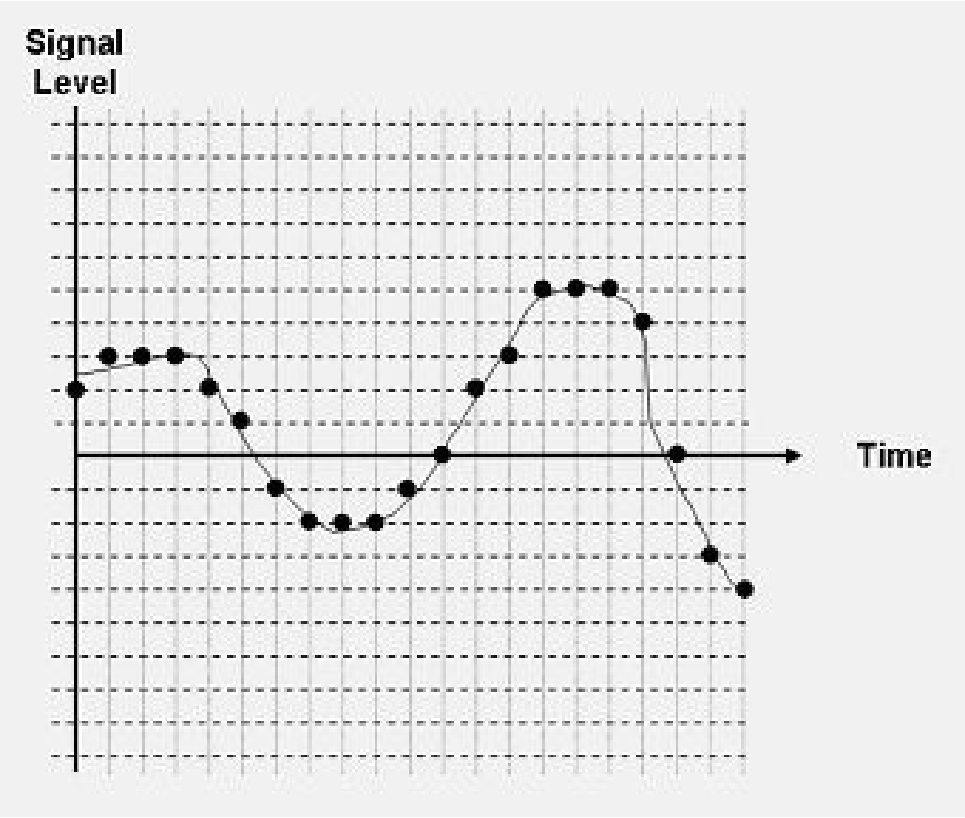

Иллюстрация 4. Простая кодово-импульсная модуляция
где: Signal Level –уровень сигнала, Time – время.

Аналоговые сигналы должны быть записаны с частотой в два раза большей самой высокой частотной составляющей сигнала, чтобы можно было правильно воспроизвести оригинал. В телефонной сети человеческий голос передаётся в полосе частот 0-4000 Гц (хотя для передачи голоса фактически используются только приблизительно 400-3400 Гц); поэтому, выборка голоса производится 8000 раз в секунду (частота выборки 8 кГц). Музыкальные аудио приложения предполагают полный диапазон частот доступных для человеческого уха и поэтому обычно используют частоту выборки 44.1 кГц (Fries и Fries 2000; Rey 1983).

Из частоты выборки (44,1 кГц), разрешающей способности кодово-импульсной модуляции (16 битов) и числа звуковых каналов (два) можно легко вычислить, что скорость передачи битов несжатой музыки равна 1411200 битам в секунду. Это предполагает, что одноминутный аудио файл (несжатый) займет 10,6 Мбайт (1411200*60/8 = 10584000). В действительности, звуковые файлы уменьшаются при помощи разнообразных методов сжатия. Один очевидный метод – это уменьшить число каналов до одного или уменьшать частоту выборки, в некоторых случаях до 11 кГц. Другие кодеки используют собственные схемы сжатия. Все эти решения снижают качество звука.

Таблица 1: Некоторые распространённые цифровые аудио форматы (Fries и Fries 2000)

Тип Аудио	Расширение Файла	Кодек
AIFF (Mac)	.aif, .aiff	Кодово-импульсная модуляция (или другой)
AU (Sun/Next)	.au	μ-law (или другой)
CD audio (CDDA)	нет данных	Кодово-импульсная модуляция

MP3	.mp3	MPEG Audio Layer III
Windows Media Audio	.wma	Разработанный компанией Microsoft
QuickTime	.qt	Разработанный компанией Apple Computer
RealAudio	.ra, .ram	Разработанный компанией Real Networks
WAV	.wav	Кодово-импульсная модуляция (или другой)

Цифровые контейнеры

Существует много способов скрытия сообщений в цифровых носителях. Компьютерные судебные исследователи знакомы с данными, которые остаются в зазорах файловой системы или свободной области, как остатки предыдущих файлов, и можно написать программы для получения доступа непосредственно к зазорам и свободным областям файловой системы. Небольшие количества данных также могут быть скрыты в неиспользуемых частях заголовков файлов (Curran и Bailey 2003).

Информация может также быть скрыта на накопителе на жестких магнитных дисках (далее, НЖМД) в скрытом разделе. Скрытый раздел не будет виден при нормальных условиях, хотя программы конфигурирования НЖМД и другие утилиты могут дать полный доступ к скрытому разделу (Johnson и другие 2001). Эта теория была реализована в стеганографической файловой системе **ext2fs** для **Linux**. Скрытая файловая система представляет особый интерес, так как она защищает пользователя от того, чтобы быть неразрывно привязанным к определенной информации на его НЖМД. Такая форма возможности отрицать, что было совершено какое-либо действие, позволяет пользователю утверждать, что он не владеет определенной информацией или утверждать, что определенные события никогда не происходили. В этой системе пользователи могут скрыть ряд файлов на НЖМД, гарантировать тайну содержания файлов, и не разрушить нескрытые файлы удалением стеганографического драйвера файлов (Anderson и другие 1998; Artz 2001; McDonald и Kuhn 2000).

Сетевые протоколы могут быть ещё одним цифровым контейнером. Например, тайный протокол управления передачей Крэйга Роуланда формирует скрытые каналы связи, используя поле идентификации в пакетах протокола IP или поле порядкового номера в сегментах протокола управления передачей (TCP) (Johnson и другие 2001; Rowland 1996).

Существует несколько звуковых характеристик, которые могут быть изменены таким образом, что будут неразличимыми человеческими чувствами, и такие изменения, например, небольшие изменения фазового угла, модуляции речи и частоты, могут переносить скрытую информацию (Curran и Bailey 2003).

Тем не менее, из-за изобилия уже существующих потенциальных файлов-контейнеров, их способности создавать бесчисленное количество новых файлов-контейнеров, и свободного доступа к стеганографическому программному обеспечению, которое будет работать с этими контейнерами, графические и аудио файлы остаются самыми легкими и распространёнными носителями-контейнерами в Интернете. По этой причине, внимание в данной работе уделяется графическим и аудио файлам.

Самый распространенный метод скрытия стеганографических данных в аудио и графических файлах использует некоторый тип замены или переписывания самого младшего двоичного разряда. Термин «самый младший двоичный разряд» происходит от числовой значимости битов в байте. Старший или самый старший двоичный разряд – это

тот, который имеет самое высокое арифметическое значение (то есть, 2^7=128), тогда как младший или самый младший двоичный разряд – это тот, который имеет самое низкое арифметическое значение (то есть, 2^0=1).

В качестве простого примера замены младшего двоичного разряда, представьте «скрытие» буквы “G” в следующих восьми байтах файла-контейнера (младшие двоичные разряды подчеркнуты):

10010101 00001101 11001001 10010110

00001111 11001011 10011111 00010000

Буква “G” представлена в американском стандартном коде для обмена информацией (ASCII) в виде двоичной цепочки 01000111. Эти восемь битов могут быть «записаны» в самый младший двоичный разряд каждого из восьми байтов контейнера следующим образом:

1001010*0* 00001101 1100100*0* 10010110

0000111*0* 11001011 10011111 0001000*1*

В вышеизложенном примере, фактически, была изменена только половина младших двоичных разрядов (выделены выше курсивом). Это имеет некоторый смысл, когда один набор нулей и единиц заменяется другим набором нулей и единиц.

Замена младшего двоичного разряда может использоваться для перезаписи законных кодировок цвета модели RGB или указателей палитры в файлах GIF и BMP, коэффициентов в файлах JPEG и уровней кодово-импульсной модуляции в аудио файлах. Переписыванием младшего двоичного разряда, числовое значение байта изменяется очень незначительно и маловероятно, что это будет обнаружено человеком визуально или на слух.

Замена младшего двоичного разряда – это простой и распространённый, метод стеганографии. Однако, использование этого метода, не обязательно так просто, как это кажется. Только самое простое стеганографическое программное обеспечение всего лишь перезапишет каждый младший двоичный разряд со скрытыми данными. Почти все используют какие-нибудь средства для расположения фактических битов в случайном порядке в файле-контейнере, которые различны. Это один из факторов, который делает обнаружение стеганографии таким трудным.

Ещё один способ скрытия информации в изображении с палитрой состоит в том, чтобы изменить порядок цветов в палитре или использовать кодирование младшего двоичного разряда на цветах палитры, а не на данных изображения. Однако, эти методы потенциально слабы. Многие графические программные средства упорядочивают цвета палитры по частоте, светимости или другим параметрам, а случайно упорядоченная палитра выделяется при статистическом анализе (Fridrich и Du 2000).

Продолжают появляться более новые, более сложные стеганографические методы. Стеганографические методы передачи сигналов с расширенным спектром аналогичны методам радиосвязи с передачей сигналов в широком спектре (разработанные во время второй мировой войны и обычно используемые в системах передачи данных сегодня), где уровень сигнала распространяется по спектру широкой частоты, а не сосредотачивается на единственной частоте, чтобы усложнить обнаружение и глушение сигнала. Стеганография с широким диапазоном имеет ту же самую функцию - избежать обнаружения. Эти методы используют в своих интересах тот факт, что небольшие искажения в графических и аудио файлах являются наименее обнаружимыми в «высокоэнергетических» частях контейнера (то есть, высокая интенсивность в аудио файлах или яркие цвета в графических файлах). Даже когда их специально сравнивают, человеческие чувства легче обмануть, когда маленькие изменения сделаны в громких звуках и/или ярких цветах (Wayner 2002).

Примеры стеганографии

В настоящее время существует более 100 доступных стеганографических программ, начиная с бесплатного ПО, заканчивая коммерческими продуктами. В данном разделе будет показано некоторые простые примеры стеганографии: скрытие 11067-битовой GIF-карты аэропорта г. Берлингтона, Штат Вермонт, (Иллюстрация 5) в GIF, JPEG и WAV-файлах.

Иллюстрация 5. Эта карта скрыта в различных контейнерах, описанных в этой статье.

В первом примере применяется **Gif-It-Up**, программа компании **Nelsonsoft**, которая скрывает информацию в GIF-файлах, используя замену младшего двоичного разряда (и включает в себя опцию шифрования). Иллюстрация 6 показывает GIF- изображение ночной аллеи г. Вашингтон, округ Колумбия, которое использовалось программой **Gif-It-Up**, для помещения в нее карты аэропорта, показанную на Иллюстрации 5. Размер оригинального контейнера - 632 778 байт и он использует 249 уникальных цветов, тогда как размер файла, содержащего стеганографические данные, 677 733 байт и он использует 256 уникальных цветов. Размер файла больше в файле, содержащем стеганографические данные, из-за опции расширения цвета, которая используется для минимизации искажения в изображении файла-контейнера. Если расширение цвета не используется, различия в размерах файла чуть менее заметны.

Иллюстрация 6. Файл-контейнер в формате GIF, содержащий карту аэропорта.

Иллюстрация 7. Палитра файла-контейнера с изображением Вашингтонской аллеи до (слева) и после (справа) скрытия файла с изображением карты.

Иллюстрация 7 показывает палитры файла-контейнера до, и после вложения сообщения. Как и все программы для вложения самого младшего двоичного разряда, которые работают с восьмибитовыми цветными изображениями, **Gif-It-Up** изменяет цветную палитру и обычно заканчивает многими двойными парами цвета.

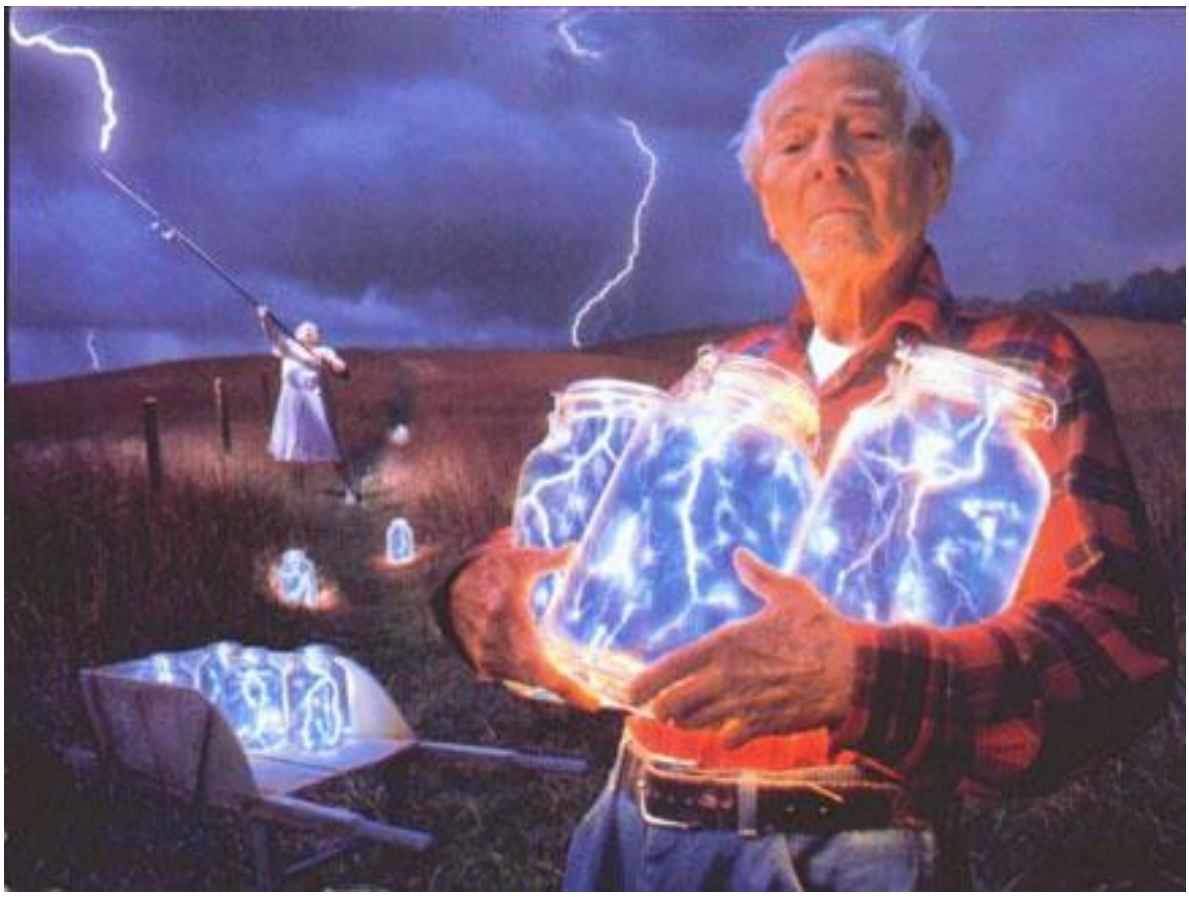

Иллюстрация 8. Файл-контейнер в формате JPEG, содержащий карту аэропорта.

JP Hide-&-Seek (JPHS), программа Аллана Латама, предназначена для работы с файлами JPEG и сжатием с потерями. Программа **JPHS** использует перезапись самого младшего двоичного разряда в коэффициентах дискретного косинусного преобразования, которые использует алгоритм JPEG. Алгоритм шифрования BlowFish используются для рандомизации и шифрования младшего двоичного разряда (Johnson and Jajodia 1998B). Иллюстрация 8 показывает пример JPEG- файла, с вложенной в него картой аэропорта. Оригинальный файл-контейнер имеет размер 207244 байт и содержит 224274 уникальных цвета. Этот же файл, содержащий стеганографические данные, имеет размер 207275 байт и содержит 227870 уникальных цветов. Здесь нет цветной палитры используемой для просмотра, потому что JPEG использует 24- битовое кодирование цвета, и дискретное косинусное преобразование.

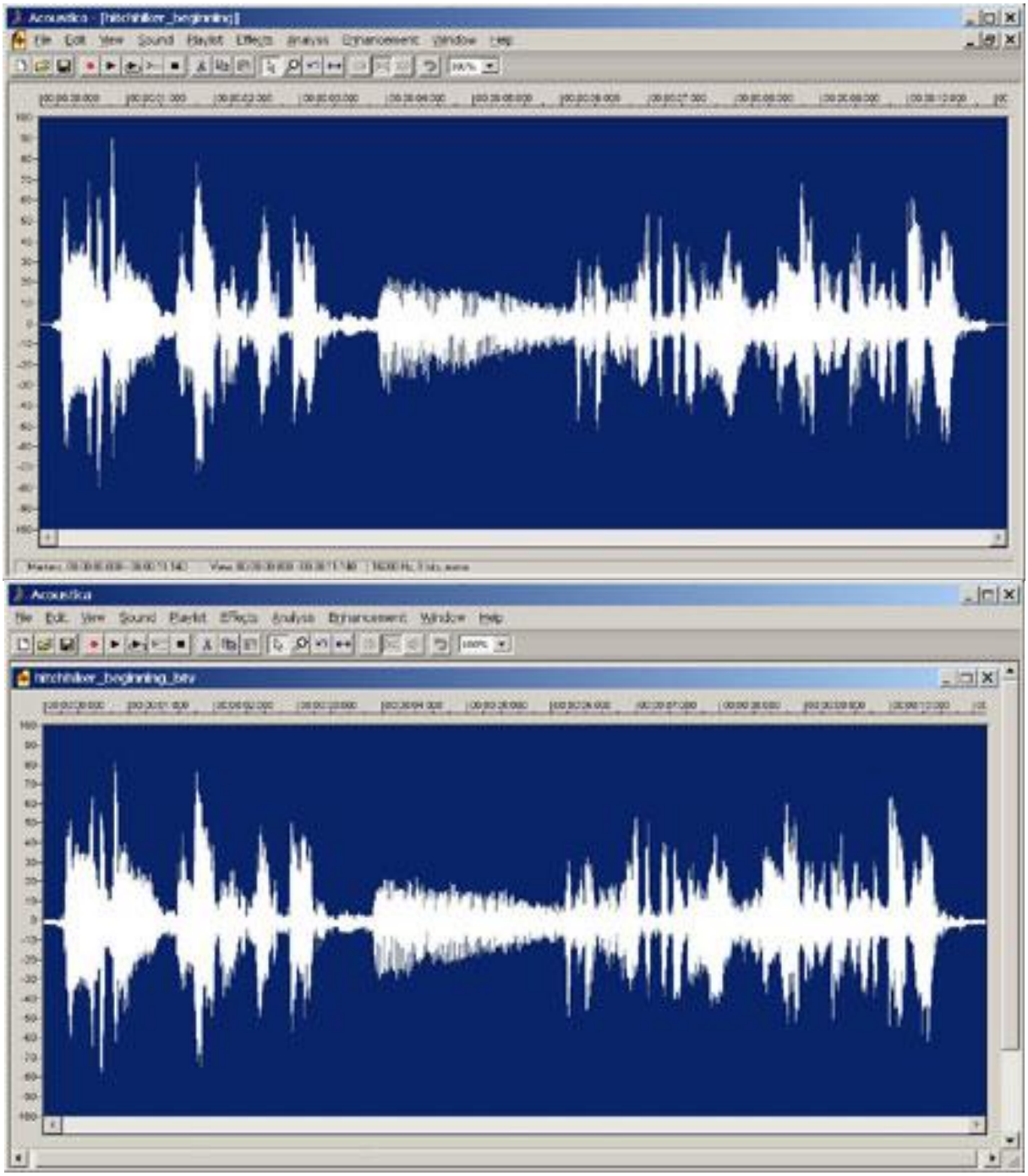

Иллюстрация 9. Сравнения уровня сигнала между файлом-контейнером в формате WAV до (выше) и после (ниже) вложения карты аэропорта.

В заключительном примере используется **S-Tools**, программа Энди Брауна, которая может скрывать информацию в файлах формата GIF, BMP и WAV. Программа **S-Tools** использует замену самого младшего двоичного разряда в файлах, которые применяют сжатие без потерь, например, восьми или 24- битовую цветовую и кодово-импульсную модуляцию. Программа **S-Tools** использует пароль при рандомизации самого младшего двоичного разряда в файлах и может шифровать данные, используя Стандарт шифрования данных США (DES), Международный алгоритм шифрования данных (IDEA), Шифр профиля сообщения (MDC), или Тройной-DES (Johnson и Jajodia 1998A; Johnson и Jajodia 1998B; Wayner 2002). Иллюстрация 9 показывает сравнение уровня сигнала между файлом-контейнером в формате WAV до и после скрытия карты аэропорта. Оригинальный WAV-файл имеет размер 178544 байт, тогда как этот же файл, содержащий стеганографические данные, имеет размер 178298 байт. Хотя из-за относительно небольшого размера рисунка трудно увидеть детали, заметны некоторые различия в

начале и конце звукового образца (то есть, в течение периодов тишины). (Некоторые стеганографические инструменты имеют встроенные вычислительные средства, чтобы избегать низко-интенсивные части сигнала.) Звуковые файлы очень хорошо подходят для скрытия информации, потому что они обычно имеют относительно большой размер, что затрудняет нахождение небольших скрытых элементов.

Программы **Gif-It-Up**, **JPHS**, и **S-Tools** использовались только в качестве примера. Они бесплатны, легки в использовании и хорошо выполняют свои задачи. Есть много других программ, которые могут использоваться для скрытия информацию в файлах формата BMP, GIF, JPEG, MP3, программы Paintbrush (PCX), формата переносимой сетевой графики (PNG), тегированного формата файлов изображений (TIFF), WAV и других типов файлов-контейнеров. Веб-сайт **StegoArchive.Com** имеет очень хороший список свободно распространяемого, условно-бесплатного и коммерческого стеганографического программного обеспечения для операционных систем DOS, Linux/Unix, MacOS, Windows и других (StegoArchive.com 2003).

Хотя обсуждение выше сосредоточилось только на графических и звуковых файлах, стеганографические носители не ограничены этими типами файлов. Другие типы файлов также имеют характеристики, которые могут быть использованы для скрытия информации. Например, программа **Hydan** может скрывать текстовые сообщения в исполняемых файлах операционных систем OpenBSD, FreeBSD, NetBSD, Red Hat Linux и Windows XP. Программа **Hydan**, разработанная Раканом Эль-Халилом (Rakan El-Khalil), использует в своих интересах избыточность в наборе команд i386 и вставляет скрытую информацию, определяя наборы функционально эквивалентных команд, концептуально подобно мимикрии на основе грамматики (например, когда команды ADD – это нулевой бит, а команды SUB - один бит). Программа может скрыть приблизительно один байт сообщения в каждых 110 байтах кода программы и сохранить оригинальный размер файла приложения. Также можно применять шифрование алгоритмом BlowFish (El-Khalil 2003).

Обнаружение стеганографии

Так называемая « Проблема Заключенных» (Simmons 1983) часто используется, чтобы описать стеганографию, хотя первоначально её вводили для описания сценария криптографии.

Проблема включает в себя двух заключенных, Элиса и Боба, которые заперты в отдельных тюремных камерах и желают сообщить друг другу некоторый секретный план. Элису и Бобу разрешают обмениваться сообщениями друг с другом, но Уильям, тюремный надзиратель, может читать все сообщения. Элис и Боб знают, что Уильям прекратит их связь, если он обнаружит секретный канал (Chandramouli 2002; Fridrich и другие 2003B).

Уильям может действовать пассивно или активно. В пассивной модели надзирателя, Уильям исследует каждое сообщение и решает, отправить сообщение или нет, основываясь на своей способности обнаружить скрытое сообщение. В активной модели надзирателя, Уильям может изменить сообщения, если желает. Консервативный или злой надзиратель мог бы фактически изменять все сообщения в попытке разрушить любой тайный канал так, чтобы Элису и Бобу нужно было бы использовать очень надёжный метод стеганографии (Chandramouli 2002; Fridrich и другие 2003B).

Трудность задачи начальника будет в значительной степени зависеть от сложности алгоритма стеганографии и количества предшествующих знаний Уильяма (Chandramouli 2002; Fridrich и другие 2003B; Provos and Honeyman 2003).

В чистой модели стеганографии, Уильям ничего не знает о методе стеганографии, используемом Элисом и Бобом. Со стороны Элиса и Боба будет неправильным полагаться на это, так как безопасность редко работает в условиях неясности и особенно катастрофична применимо к криптографии. Однако, часто это является моделью

проведения компьютерно-техническим аналитиком поиска возможного использования стеганографии на веб-сайте или жестком диске.

Стеганография, с использованием секретного ключа, предполагает, что Уильям знает алгоритм стеганографии, но не знает секретный стеганографический/криптографический ключ, применяемый Элисом и Бобом. Это последовательно при условии, что пользователь криптографии будет действовать по Принципу Керкхоффа (то есть, « безопасность криптографической схемы лежит в управлении ключом, а не в секретности алгоритма».) (Kahn 1996). Возможно, это является слишком сильным предположением для применения на практике, потому что полная информация включала бы в себя доступ к источнику файла-контейнера.

Стегоанализ, обнаружение стеганографии третьим лицом, является относительно молодой дисциплиной исследований, несколько статей по этой теме появились перед концом 1990-ых. Искусство и наука стегоанализа предназначены, чтобы обнаружить или оценить скрытую информацию, основываясь на наблюдении некоторой передачи данных и не делая никаких предположений об алгоритме стеганографии (Chandramouli 2002). Обнаружение скрытой информации может быть не достаточным. Возможно, стегоаналитик также захочет извлечь скрытое сообщение, повредить скрытое сообщение, чтобы получатель не мог извлечь его, и/или изменить скрытое сообщение, чтобы послать дезинформацию получателю (Jackson и другие 2003). Обычно достаточным является обнаружение и извлечение стеганографии, если цель – это сбор доказательств, связанных с прошлым преступлением, хотя разрушение и/или изменение скрытой информации также могли бы быть законными целями правоприменительных органов в ходе расследования деятельности преступных или террористических групп.

Методы стегоанализа можно классифицировать подобно методам криптоанализа, которые в значительной степени основаны на том, какое количество предшествующей информации известно (Curran и Bailey 2003; Johnson и Jajodia 1998B).

- Атака только стеганографии: Стеганографический носитель - единственный элемент, доступный для анализа;
- Атака известного контейнера: Для анализа доступен и контейнер (исходный файл в который позднее были помещены стеганографические данные) и стеганографический носитель;
- Атака известного сообщения: Известно скрытое сообщение;
- Атака выбранной стеганографии: Известен как стеганографический носитель, так и стеганографический алгоритм;
- Атака выбранного сообщения: Известное сообщение и алгоритм стеганографии используются, чтобы создать стеганографический носитель для будущего анализа и сравнения;
- Атака известной стеганографии: Известны контейнер и стеганографический носитель, а так же алгоритм стеганографии.

Методы стеганографии для цифровых носителей можно широко классифицировать как работающие в области изображений или области преобразований. Инструменты области изображений скрывают сообщение в контейнере неким видом побитовой манипуляции, например, изменением самого младшего бита. Инструменты области преобразований управляют алгоритмом стеганографии и фактическими преобразованиями, используемыми в скрытии информации, например, коэффициентами дискретного косинусного преобразования в изображениях формата JPEG (Johnson и Jajodia 1998B).

В таком случае, отсюда вытекает, что стегоанализ следует по пути, в котором развивается стеганография. Первый простой подход состоит в том, чтобы визуально осмотреть контейнер и носитель стеганографии. Многие простые стеганографические инструменты работают в области графических изображений и используют для стеганографического скрытия сообщения биты, в контейнере, независимо от содержания

контейнера. Хотя легче скрыть сообщение в области более яркого цвета или более громкого звука, программа не может отыскать эти области. Таким образом, визуальный осмотр может быть достаточным, чтобы появилось подозрение что исследуемый носитель содержит стеганографические данные (Wayner 2002).

Второй подход состоит в том, чтобы искать странности в структуре, которые предполагают манипуляцию. Вложение самого младшего бита в изображение на основе палитры часто является причиной большого количества двойных цветов, где идентичные (или почти идентичные) цвета дважды появляются в палитре и имеют отличия только в самом младшем двоичном разряде. Стеганографические программы, которые скрывают информацию просто, управляя порядком цветов в палитре, также вызывают структурные изменения. Структурные изменения часто создают сигнатуру алгоритма стеганографии, который использовался (Jackson и другие 2003; Wayner 2002).

Стеганографические методы обычно изменяют статистические данные контейнера и, очевидно, что более длинные скрытые сообщения изменят контейнер больше чем короткие (Farid 2001; Fridrich и Du 2000; Fridrich и Goljan 2002; Ozer и другие 2003). Статистический анализ, как правило, используется для обнаружения скрытых сообщений, особенно когда аналитик работает в слепую (Jackson и другие 2003). В области статистического стегоанализа имеется большое количество работ.

Статистический анализ графических и звуковых файлов может показать, отклоняются ли статистические свойства файлов от ожидаемой нормы (Farid 2001; Ozer и другие 2003; Provos и Honeyman 2001). Эти так называемые статистические данные первого порядка – средства, дисперсия, критерий хи-квадрат (X^2) - могут определить количество избыточной информации и/или искажения в носителе. Хотя эти измерения могут спрогнозировать, было ли содержание изменено или кажется подозрительным, они не являются окончательными (Wayner 2002).

Процесс статистического стегоанализа осложняется, так как некоторые алгоритмы стеганографии стараются сохранить статистические данные первого порядка файла-контейнера, чтобы избежать только этого типа обнаружения. Шифрование скрытого сообщения также затрудняет его обнаружение, потому что зашифрованные данные обычно имеют высокую степень случайности, и единицы и нули появляются с равным правдоподобием (Farid 2001; Provos и Honeyman2001).

Восстановление скрытого сообщения добавляет ещё один уровень сложности по сравнению с простым обнаружением присутствия скрытого сообщения. Восстановление сообщение требует знания или оценки длины сообщения и, возможно, ключа шифрования и знания криптографического алгоритма (Fridrich и др. 2003B).

Алгоритмы для определенных типов файлов-контейнеров могут сделать анализ более простым. Формат JPEG, в частности, получил большое внимание исследователей из-за способа, с помощью которого различные алгоритмы работают с этим типом файла. Формат JPEG является плохим контейнером при использовании для сокрытия вложения стеганографического метода изменения самого младшего бита, так как модификация в файле, вызванная сжатием JPEG, упрощает задачу обнаружения скрытой информации (Fridrich и Du2000). Существует несколько алгоритмов, которые скрывают информацию в файлах формата JPEG, и все они работают по-разному. Алгоритм **JSteg** последовательно вкладывает скрытые данные в самые младшие биты, алгоритм **JP Hide&Seek** использует вероятностный процесс для выбора самых младших битов, алгоритм **F5** использует матричное кодирование, основанное на коде Хемминга, а алгоритм **OutGuess** сохраняет статистические данные первого порядка (Fridich и др. 2001; Fridich и др. 2002A; Fridrich и др. 2002B; Fridich и др. 2003A; Provos и Honeyman 2001; Provos и Honeyman 2003).

Были описаны более расширенные статистические тесты графических и звуковых файлов, использующие статистику высшего уровня, линейный анализ, случайные поля Маркова, статистику волны малой амплитуды и другие (Farid 2001; Farid и Lyu2003; Fridrich и Goljan2002; Ozer и др. 2003). Детальное обсуждение выходит за пределы данной

статьи, но результаты этого исследования можно увидеть в некоторых инструментах детектирования стеганографии.

Сегодня большинство стегоанализов основаны на сигнатурах, подобно антивирусным приложениям и системам обнаружения вторжения. Системы стегоанализа на основе детектирования аномалий только начинают появляться. Хотя прежние системы являются точными и надёжными, последние будут более гибкими и приспособленными быстро отвечать требованиям детектирования новых методов стеганографии. Одна форма так называемого « слепого поиска стеганографии» различает между собой « чистые» изображения и изображения, содержащие стеганографические данные, используя статистику, основанную на разложении волны малой амплитуды, или исследовании пространства, ориентации, и масштаба через подмножества большего изображения (Farid 2001; Jackson и др. 2003).

Этот тип статистического стегоанализа не ограничивается графическими и звуковые файлами. Программа **Hydan** сохраняет размер оригинального контейнера, но, используя набор « функционально эквивалентных» команд, изменяет некоторые команды, которые обычно не используются. Это обнаруживает программу **Hydan** при исследовании статистического распределения команд программы. Будущие версии программы **Hydan** будут поддерживать целостность статистического профиля оригинального приложения, чтобы защититься от этого анализа (El-Khalil 2003).

Система правоохранительных органов не всегда может узнать, когда и где использовалась стеганография или алгоритм, который применялся. Общие инструменты, которые могут обнаружить и классифицировать стеганографию, все еще находится в начальной стадии исследования, но уже становятся доступным в программных средствах, некоторые из которых описаны в следующем разделе (McCullagh 2001).

И повторяется тот же самый цикл, как и в мире криптографии – стегоанализ помогает находить вложенную стеганографию, но также и показывает разработчикам новых алгоритмов стеганографии, как избежать обнаружения.

Инструменты для поиска стеганографии

Данная статья больше предназначена практикующим судебным экспертам, чем исследователям. Следовательно, в данном разделе будет приведено несколько примеров доступного в настоящее время программного обеспечения, которое может обнаружить присутствие программ стеганографии, обнаружить подозрительные файлы-контейнеры, и разрушить сообщения, скрытые при помощи стеганографии. Это ни в коем случае не является обзором всех доступных инструментов, а лишь примером доступных возможностей. Сайт StegoArchive.com содержит список многих программ стегоанализа (StegoArchive.com 2003).

Обнаружение стеганографического программного обеспечения на исследуемом компьютере является важной частью последующего судебного анализа. Как показывает практика, многие программы для поиска стеганографии работают лучше, когда есть ключевая информация относительно типа использованных стеганографических алгоритмов. Обнаружение стеганографического программного обеспечения на компьютере даёт повод подозревать, что на исследуемом компьютере фактически имеются стеганографические файлы со скрытыми сообщениями. Кроме того, тип найденного стеганографического программного обеспечения непосредственно влияет на любой последующий стегоанализ (например, наличие на анализируемом носителе программы **S-Tools** может указать на необходимость исследовать, на содержание стеганографических данных, файлы форматов GIF, BMP и WAV, тогда как программа **JP Hide-&-Seek** может побудить аналитика более пристальнее осмотреть JPEG-файлы).

Программное обеспечение **Gargoyle** (ранее StegoDetect) компании WetStone Technologies (WetStone Technologies 2004) может использоваться для обнаружения присутствия стеганографического программного обеспечения. Программа **Gargoyle**

применяет собственный набор данных (или хэш набор) всех файлов в известном стеганографическом программном обеспечении, сравнивая их с хэшами файлов, подлежащих поиску. Иллюстрация 10 показывает результаты поиска программы **Gargoyle** в каталоге, где хранились стеганографические программы. Базы хэшей программы Gargoyle также могут использоваться для обнаружения присутствия криптографических программ, программ мгновенной передачи сообщений, клавиатурных шпионов, троянских программ, программ взлома пароля, и другого вредоносного программного обеспечения.

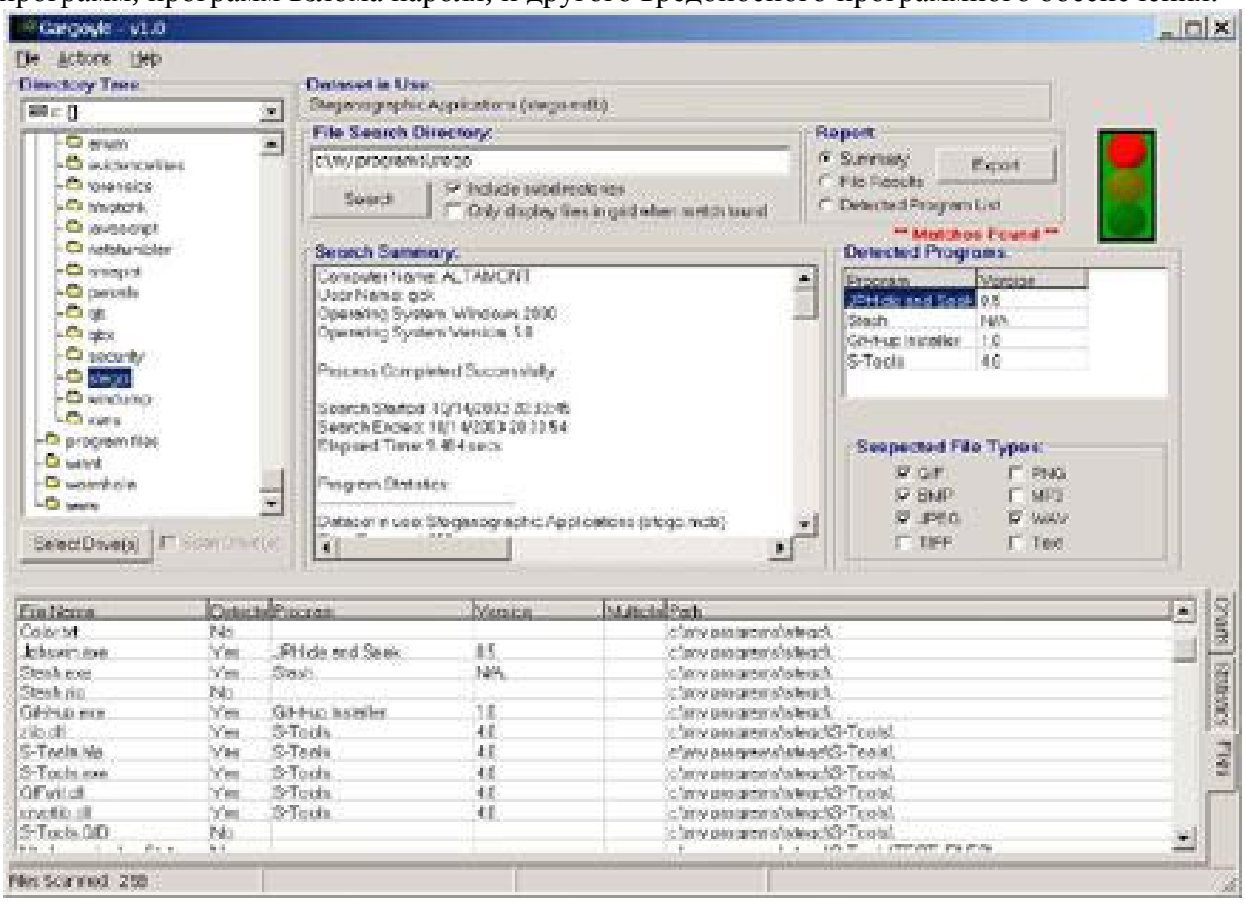

Иллюстрация 10. Результаты поиска программы Gargoyle в одном из каталогов на НЖМД.

Программа **Forensic Toolkit** компании AccessData (AccessData 2003) и программа **EnCase** компании Guidance Software (Guidance Software 2003) может использовать наборы хэшей программ **HashKeeper** (Hashkeeper 2003), Maresware (Maresware 2003) и Национальной справочной библиотеки программного обеспечения (National Software Reference Library 2003) для поиска большого разнообразия программ. В общем, их наборы хэшей предназначены исключить хэш известных «хороших» файлов из индексов поиска в течение судебного анализа. Программа **Gargoyle** также может импортировать эти наборы хэшей.

Обнаружение стеганографического программного обеспечения продолжает становиться более трудным процессом по ещё одной причине - небольшой размер программного обеспечения вместе с увеличивающейся емкостью сменных носителей. Например, программа **S-Tools** требуют менее 600 Кб дискового пространства, и может выполняться непосредственно, без дополнительной установки, с накопителя на гибких магнитных дисках или USB- накопителя. При данных обстоятельствах, на НЖМД подозреваемого не будет найдено никаких артифактов этой программы.

Вторая важная функция программного обеспечения для обнаружения стеганографии – найти возможные файлы-контейнеры. В идеале, программное обеспечение для детектирования стеганографии также должно предоставлять некоторые ключи относительно стеганографического алгоритма, который использовался для скрытия

информации в исследуемом файле, чтобы аналитик мог попытаться восстановить скрытую информацию.

Одна из самых распространённых программ для детектирования – программа **stegdetect** Нильса Провоса (Niels Provos). Программа **stegdetect** может детектировать информацию в изображениях формата JPEG, закодированную такими стеганографическими методами как: **F5**, **Invisible Secrets** , **JPHide** и **JSteg** (OutGuess 2003). Иллюстрация 11 показывает результаты из **xsteg**, графического интерфейса для программы **stegdetect**, при исследовании двух файлов на НЖМД – оригинального контейнера и графического файла, в формате JPEG содержащего стеганографические данные, показанного на Иллюстрации 8. Обратите внимание, что стеганографический файл не только помечен как содержащий скрытую информацию, но программа также предполагает (правильно) используемый стеганографический метод - **JPHide**.

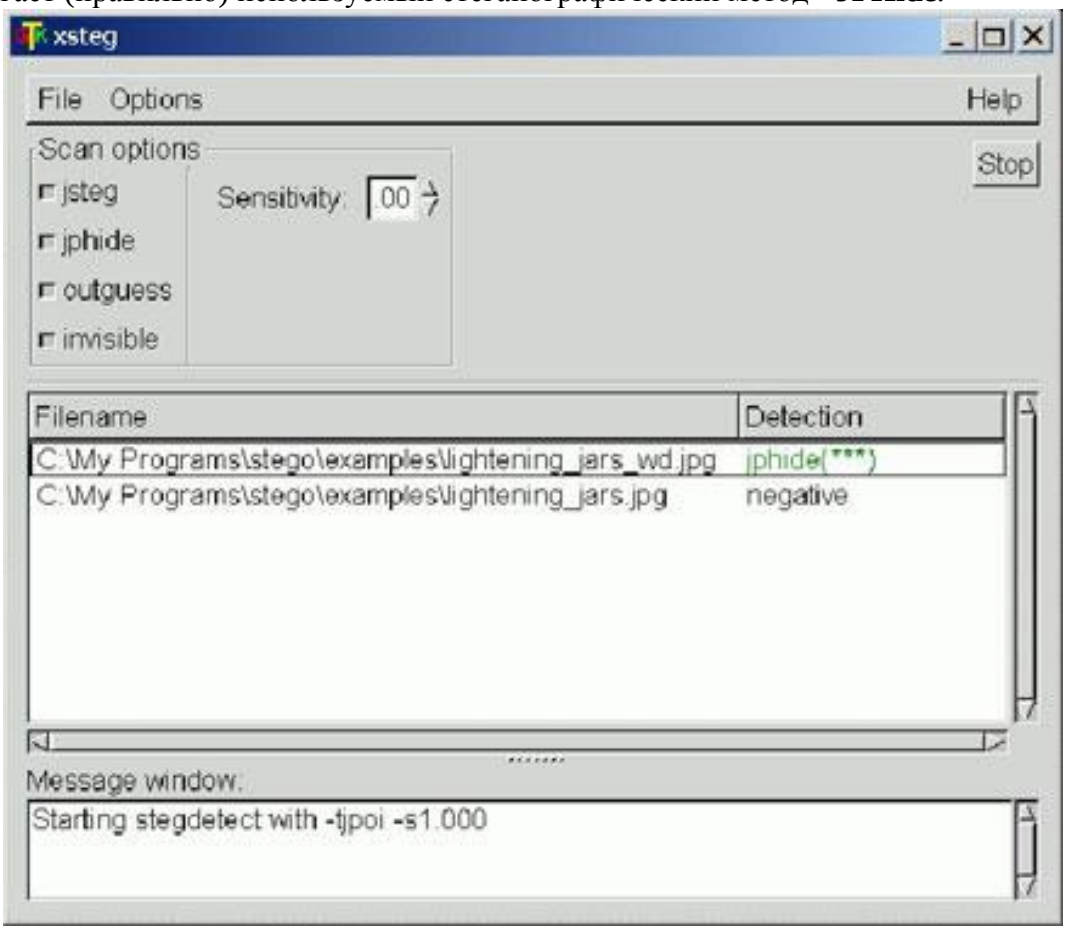

Иллюстрация 11. Результаты из Xsteg при экспертизе двух исследуемых JPEG – файлов.

Программа **Stego Watch** компании WetStone Technologies (WetStone Technologies 2004) анализирует файлы и даёт вероятностный ответ относительно того, какие из них являются стеганографическими носителями, а также, указывает вероятный алгоритм, использованный для скрытия стеганографических данных (который, в свою очередь, предоставляет сведения о наиболее вероятностных методах для извлечения стеганографических данных). Анализ использует разнообразные выбираемые пользователем статистические тесты, основанные на характеристиках файлов-контейнеров, которые могли быть изменены различными стеганографическими методами. Знание о наличии стеганографического программного обеспечения находящегося на исследуемом компьютере поможет аналитику выбрать наиболее подходящие статистические тесты.

Иллюстрация 12 показывает результаты программы **Stego Watch** после исследования файла-контейнера в формате JPEG, показанного на Иллюстрации 8. Секция «Алгоритмы

обнаружения стеганографии» (“ Steganography Detection Algorithms”) показывает статистические алгоритмы, используемые для анализа и те, которые дали положительные результаты для этого изображения. Как указано выше, программа **Stego Watch** правильно распознаёт стеганографическое программное обеспечение для работы с форматом JPEG, которое было использовано для сокрытия данных.

Институт изучения защитных технологий (Institute for Security Technology Studies) при Дартмутском Колледже разработал программное обеспечение, способное обнаружить скрытые данные в графических файлах, используя статистические модели, которые не зависят от формата изображения или метода стеганографии. Данный продукт пока не доступен для приобретения. Эта программа была проверена на 1800 изображениях и четырех различных стеганографических алгоритмах и смогла обнаружить присутствие скрытых сообщений с 65- процентной точностью, с ошибочно-положительной оценкой меньше чем 0,001 процента (Dartmouth College 2003).

Иллюстрация 12. Информация из программы Stego Watch о JPEG-файле, который, как подозревается, является стеганографическим контейнером.

Обнаружение стеганографии в файле, который, как подозревается, содержит её – относительно легко по сравнению с извлечением скрытых данных. Большая часть стеганографического программного обеспечения использует пароли для обеспечения секретности, рандомизацию и/или шифрование. **Stegbreak**, программа сопутствующая приложению **stegdetect**, использует для дешифрации скрытых данных, зашифрованных алгоритмами **JSteg-Shell**, **JPHide** и **OutGuess,** метод подбора по словарю, но, опять же, она применима только к JPEG- файлам (OutGuess 2003). Подобно ей, программа **Stego Break**, сопутствующая программе **Stego Watch** компании WetStone, использует метод восстановление пароля исследуемых файлов по словарю (WetStone Technologies 2004). Схемы обнаружения стеганографии не помогают непосредственно в восстановлении пароля. Процесс обнаружения соответствующих ключей – вот, где начинает действовать остальная часть судебных исследователей и экспертов.

Работая с уликами криминального дела, у судебного эксперта, вероятно, не будет никакой причины изменять какие-нибудь исследуемые файлы. Однако, при исследовании, которое является составной частью наблюдения за террористами, возможно, будет необходимо разрушить скрытую информацию, даже если её нельзя восстановить. Скрытое содержание, такое как стеганографические и цифровые водяные знаки, может быть атаковано несколькими способами, чтобы его можно быть удалить или изменить (Hernandez Martin и Kutter 2001; Voloshynovskiy и другие 2001), и существует программное обеспечение, специально предназначенное для атаки на цифровые водяные знаки. Такие атаки имеют один из двух возможных эффектов - они либо уменьшают стеганографическую пропускную способность контейнера (необходимую для избежания атаки) или полностью блокируют способность контейнера служить стеганографическим носителем.

Хотя этот тема также выходит за рамки этой статьи, чтобы закрыть это обсуждение можно использовать один интересный пример разрушения файлов содержащих стеганографические данные. Программа **2Mosaic** Фабьена Петиткола (Fabien Petitcolas) использует так называемую « презентационную атаку», направленную прежде всего на изображения на веб-сайте. Программа **2Mosaic** атакует систему создания цифровых водяных знаков, разбивая изображение на меньшие части изображения. На веб-сайте, несколько маленьких изображений располагаются друг за другом и выглядят так же, как и оригинальное большое изображение (Petitcolas 2003).

Иллюстрация 13. Часть JPEG-изображения, созданного программой 2Mosaic, со скрытой картой аэропорта.

Иллюстрация 13 показывает пример работы программы **2Mosaic** с JPEG-изображением показанном Иллюстрации 8. В этом случае, файл-контейнер разбит на 165 частей, как показано выше (11 рядов по 15 частей исходного изображения). Метод программы **2Mosaic** очевиден при использовании. Человек, который будет просматривать измененное изображение, немедленно узнает, что здесь что-то не в порядке.

Резюме и заключение

Рассмотрите следующий гипотетический сценарий. По предварительному соглашению с членами террористической организации, лидер террористической ячейки помещает какой-то предмет для продажи на аукционе eBay каждый понедельник и размещает фотографию того предмета. Предмет для продажи является законным. Предложения принимаются, деньги получаются, и предметы поставляются с уплатой всех налогов. Но в некоторое заранее согласованное время в течение недели, размещается версия фотографии, которая содержит скрытое сообщение. Члены ячейки знают это время и еженедельно загружают файлы содержащие скрытые сообщения. Если эти люди не находятся под активным следствием, вряд ли, что кто-нибудь обратит внимание на эту деятельность.

Этот сценарий, или подобный, является жизнеспособным методом для коммуникации террористов или преступников. Но, реален ли он? После 11 сентября 2001, появился ряд статей, предполагающих, что террористы Аль Каиды применяют стеганографию (Kelly 2001; Kolata 2001; Manoo 2002; McCullagh 2001). В частичном ответе на эти сообщения, было сделано несколько попыток, чтобы детектировать присутствие стеганографических изображений в Интернете. В течение одного известного исследования был произведён поиск в более чем трёх миллионах изображений формата JPEG в архивах eBay и USENET. Использовалась программа **stegdetect**, один - два

процента изображений были признаны подозрительными, но никаких скрытых сообщений, с использованием программы **stegbreak,** восстановлено не было (Provos and Honeyman 2001; Provos and Honeyman 2003). Другое исследование провело анализ нескольких сотен тысяч изображений из случайного набора веб-сайтов и, также с использованием программ **stegdetect** и **stegbreak,** получило подобные результаты (Callinan and Kemick 2003).

Хотя эти проекты предоставляют основу для поиска стеганографических изображений на веб-сайтах, на основании только их результатов нельзя делать никаких выводов относительно наличия изображений, содержащих стеганографические данные, в Интернете. Первое и самое главное, программа **stegdetect** работает только с изображениями в формате JPEG. Другие типы изображений не исследовались. Во вторых, было исследовано ограниченное число веб-сайтов - слишком малое, чтобы сделать любые категорические утверждения об Интернете в целом. Также, интересно отметить, что некоторые исследователи стеганографии специально не публикуют информацию о том, какие Интернет сайты они исследуют или что они находят (Kolata 2001; McCullagh 2001).

Существует мало неопровержимых статистических данных о частоте, с которой сотрудники правоохранительных органов обнаруживают стеганографическое программное обеспечение или носители в ходе производства расследований и судебных экспертиз. Однако, отдельные примеры из жизни наводят на мысль, что многие судебные эксперты не ищут как положено стеганографическое программное обеспечение, а многие не смогли бы распознать такие инструментальные средства, если бы они их нашли. Кроме того, инструментальные средства, которые используются для обнаружения стеганографического программного обеспечения, часто не отвечают требованиям, и эксперты часто полагаются исключительно на наборы хэшей или непосредственно на стеганографические инструменты (Kruse and Heiser 2001; Nelson и другие 2003; Security Focus 2003). Полный поиск улик стеганографии на исследуемом НЖМД, который, возможно, содержит тысячи изображений, аудио файлов и видеоклипов может занять несколько дней (Hosmer and Hyde2003).

Несомненно, многие судебные эксперты считают поиск стеганографических инструментальных средств и/или стеганографических носителей стандартной частью каждой экспертизы (Security Focus 2003). Но чего, как оказывается, недостает - системы ориентиров, предоставляющей систематический подход к обнаружению стеганографии. Даже рекомендации Министерства юстиции США по поиску и сбору компьютерных улик, едва упоминают стеганографию (U.S. Department of Justice 2001; U.S. Department of Justice 2002). Стегоанализ будет только одной частью исследования, однако, и следователю, возможно, понадобятся ключи из других аспектов дела, чтобы указать судебным экспертам правильное направление. Судебный эксперт может подозревать использование стеганографии из-за характера преступления, книг в библиотеке подозреваемого, обнаруженных аппаратных или программных средств, больших наборов изображений, с виду похожих на дубликаты, утверждений, сделанных подозреваемым или свидетелями, или другими факторами. Веб-сайт может быть подозреваемым из-за характера его содержания или людей, которых он обслуживает. Также, эти же самые элементы могут дать эксперту ключи к паролям. А поиск стеганографии необходим не только в уголовных расследованиях или в процессе предварительного сбора информации. Судебные следователи, занимающиеся делами бухгалтерского учета, понимают необходимость поиска стеганографии, поскольку это становится типичным способом скрытия финансовых отчетов (Hosmer and Hyde 2003; Seward 2003).

Невозможно узнать, насколько сегодня широко распространено использование стеганографии преступниками и террористами (Hosmer and Hyde 2003). Однако, точные сведения сегодня, возможно, даже не имеют значения. Использование стеганографии наверняка увеличивается и будет растущим препятствием для правоохранительной и анти-террористической деятельности. Игнорирование значения стеганографии из-за

нехватки статистики – это « безопасность через отрицание» и не является хорошей стратегией.

Стеганографию не найдут, если её не будут искать. Существует несколько отчётов о том, что террористы Аль-Каиды использовали порнографические графические файлы в качестве стеганографических носителей (Kelly 2001; Manoo 2002). Использование данным конкретным противником стеганографии и порнографии может быть неожиданным с технологической и культурной точки зрения, но это демонстрирует способность работать по-новому. В области компьютерных расследований, мы также должны думать и расследовать творчески.

Литература

AccessData. Forensic Toolkit product page [Online]. (December 29, 2003). Available: http://www.accessdata.com/Product04_Overview.htm.

Anderson, R., Needham, R., and Shamir, A. Steganographic file system. In: *Proceedings of the Second International Workshop on Information Hiding* (IH '98), Lecture Notes in Computer Science, vol. 1525. D. Aucsmith, ed., Portland, Oregon, April 14-17, 1998. Springer-Verlag, Berlin, Germany, 1998, pp. 73-82. Also available: http://www.cl.cam.ac.uk/ftp/users/rja14/sfs3.pdf.

Arnold, M., Schmucker, M., and Wolthusen, S. D. *Techniques and Applications of Digital Watermarking and Content Protection.* Artech House, Norwood, Massachusetts, 2003.

Artz, D. Digital Steganography: Hiding data within data. *IEEE Internet Computing* (2001) 5(3):75-80. Also available: http://www.cc.gatech.edu/classes/AY2003/cs6262_fall/digital_steganography.pdf.

Barni, M., Podilchuk, C. I., Bartolini, F., and Delp, E. J. Watermark embedding: Hiding a signal within a cover image, *IEEE Communications* (2001) 39(8):102-108.

Bauer, F. L. *Decrypted Secrets: Methods and Maxims of Cryptology* , 3rd ed. Springer-Verlag, New York, 2002.

Callinan, J. and Kemick, D. Detecting steganographic content in images found on the Internet. Department of Business Management, University of Pittsburgh at Bradford [Online]. (December 11, 2003). Available: http://www.chromesplash.com/jcallinan.com/publications/steg.pdf.

Chandramouli, R. Mathematical approach to steganalysis. In: *Proceedings of the SPIE Security and Watermarking of Multimedia Contents IV* , vol. 4675. International Society for Optical Engineering, San Jose, California, January 21-24, 2002, pp. 14-25. Also available: http://www.ece.stevens-tech.edu/~mouli/spiesteg02.pdf.

Curran, K. and Bailey, K. An evaluation of image-based steganography methods. *International Journal of Digital Evidence* [Online]. (Fall 2003). Available: http://www.ijde.org/docs/03_fall_steganography.pdf.

Dartmouth College, Institute for Security Technology Studies. A Novel Software for Detection of Hidden Messages within Digital Images [Online]. (December 29, 2003). Available: http://www.ists.dartmouth.edu/text/steganography.php.

El-Khalil, R. Hydan [Online]. (December 30, 2003). Available: http://www.crazyboy.com/hydan/.

Farid, H. Detecting Steganographic Messages in Digital Images. Technical Report TR2001-412, Dartmouth College, Computer Science Department, 2001. Also available: http://www.cs.dartmouth.edu/~farid/publications/tr01.pdf.

Farid, H. and Lyu, S. Higher-order wavelet statistics and their application to digital forensics. *IEEE Workshop on Statistical Analysis in Computer Vision*, Madison, Wisconsin, June 2003. Also available: http://www.cs.dartmouth.edu/~farid/publications/sacv03.pdf.

Fridrich, J. and Du, R. Secure steganographic methods for palette images. In: *Proceedings of the 3rd Information Hiding Workshop* , Lecture Notes in Computer Science, vol. 1768.

Dresden, Germany, September 1999. Springer-Verlag, Berlin, Germany, 2000, pp. 47-60. Also available: http://www.ws.binghamton.edu/fridrich/Research/ihw99_paper1.dot.

Fridrich, J. and Goljan, M. Practical steganalysis of digital images: State of the art. In: *Proceedings of the SPIE Security and Watermarking of Multimedia Contents IV* , vol. 4675. International Society for Optical Engineering, San Jose, California, January 21-24, 2002, pp. 1-13. Also available: http://www.ws.binghamton.edu/fridrich/Research/steganalysis01.pdf.

Fridrich, J., Goljan, M., and Du, R. Steganalysis based on JPEG compatibility. In: *Proceedings of the SPIE Multimedia Systems and Applications IV* , Special Session on Theoretical and Practical Issues in Digital Watermarking and Data Hiding, vol. 4518. International Society for Optical Engineering, Denver, Colorado, August 21-22, 2001, pp. 275-280. Also available: http://www.ws.binghamton.edu/fridrich/Research/jpgstego01.pdf.

Fridrich, J., Goljan, M., and Hogea, D. Attacking the OutGuess. In: *Proceedings of the ACM Workshop on Multimedia and Security 2002* , Juan-les-Pins, France, December 2002A. Also available: http://www.ws.binghamton.edu/fridrich/Research/acm_outguess.pdf.

Fridrich, J., Goljan, M., and Hogea, D. New methodology for breaking steganographic techniques for JPEGs. In: Proceedings of the SPIE Security and Watermarking of Multimedia Contents V, vol. 5020. International Society for Optical Engineering, Santa Clara, California, January 21-24, 2003A, pp. 143-155. Also available: http://www.ws.binghamton.edu/fridrich/Research/jpeg01.pdf.

Fridrich, J., Goljan, M., and Hogea, D. Steganalysis of JPEG images: Breaking the F5 algorithm. *Proceedings of the 5th International Workshop on Information Hiding* (IH 2002). F. A. P. Petitcolas, ed., Noordwijkerhout, The Netherlands, October 7-9, 2002B. Springer-Verlag, Berlin, Germany, pp. 310-323. Also available: http://www.ws.binghamton.edu/fridrich/Research/f5.pdf.

Fridrich, J., Goljan, M., Hogea, D., and Soukal, D. Quantitative steganalysis of digital images: Estimating the secret message length, *Multimedia Systems* (2003B) 9(3):288-302. Also available: http://www.ws.binghamton.edu/fridrich/Research/mms100.pdf.

Fries, B. and Fries, M. *MP3 and Internet Audio Handbook*. TeamCom Books, Burtonsville, Maryland, 2000.

Guidance Software. EnCase [Online]. (December 29, 2003). Available: http://www.guidancesoftware.com/.

Hashkeeper. Hashkeeper Files [Online]. (December 29, 2003) Available: http://www.hashkeeper.org/files/.

Hernandez Martin, J. R. and Kutter, M. Information retrieval in digital watermarking, *IEEE Communications* (2001) 39(8):110-116.

Hosmer, C. and Hyde, C. Discovering covert digital evidence. *Digital Forensic Research Workshop* (DFRWS) 2003, August 2003 [Online]. (January 4, 2004). Available: http://www.dfrws.org/dfrws2003/presentations/Paper-Hosmer-digitalevidence.pdf.

Jackson, J. T., Gregg, H., Gunsch, G. H., Claypoole, R. L., and Lamont, G. B. Blind Steganography detection using a computational immune system: A work in progress. *International Journal of Digital Evidence* [Online]. (Winter 2003) (December 21, 2003). Available: http://www.ijde.org/docs/02_winter_art4.pdf.

Johnson, N. F., Duric, Z. and Jajodia, S. *Information Hiding: Steganography and Watermarking: Attacks and Countermeasures*. Kluwer Academic, Norwell, Massachusetts, 2001.

Johnson, N. F. and Jajodia, S. Exploring steganography: Seeing the unseen, *Computer* (1998A) 31(2):26-34. Also available: http://www.jjtc.com/pub/r2026.pdf.

Johnson, N. F. and Jajodia, S. Steganalysis of images created using current steganography software. In: *Proceedings of the Second International Workshop on Information Hiding* (IH '98), Lecture Notes in Computer Science, vol. 1525. D. Aucsmith, ed. Portland, Oregon, April 14-17, 1998. Springer-Verlag, Berlin, Germany, 1998B, pp.273-289. Also available: http://www.jjtc.com/ihws98/jjgmu.html.

Kahn, D. Codebreakers: *The Story of Secret Writing* . Revised ed., Scribner, New York, 1996.

Kelly, J. Terror groups hide behind Web encryption. *USA Today*, February 5, 2001. Also available: http://www.usatoday.com/tech/news/2001-02-05-binladen.htm.

Kolata, G. Veiled messages of terror may lurk in cyberspace, *New York Times*, October 30, 2001, p. 1.

Kruse, W. G. and Heiser, J. G. *Computer Forensics: Incident Response Essentials* . Addison-Wesley, Boston, Massachusetts, 2001.

Kwok, S. H. Watermark-based copyright protection system security, *Communications of the ACM* (2003) 46(10):98-101.

Manoo, F. Case of the missing code, *Salon.com*, July 17, 2002 [Online}. (December 29, 2003). Available: http://www.salon.com/tech/feature/2002/07/17/steganography/.

Maresware. Hash Set CD [Online]. (December 29, 2003). Available: http://www.dmares.com/maresware/hash_cd.htm.

McCullagh, D. Secret messages come in .Wavs. *WIRED News* , February 20, 2001 [Online]. (December 11, 2003). Available: http://www.wired.com/news/politics/0,1283,41861,00.html.

McDonald, A. D. and Kuhn, M. G. StegFS: A steganographic file system for Linux. In: *Proceedings of the Third International Workshop on Information Hiding* (IH '99), Lecture Notes in Computer Science, vol. 1768. A. Pfitzmann, ed., Dresden, Germany, September 29-October 1, 1999. Springer-Verlag, Berlin, Germany, 2000, pp. 462-477. Also available: http://www.cl.cam.ac.uk/~mgk25/ih99-stegfs.pdf.

Monash University. JPEG Image Coding Standard [Online]. (January 10, 2004). Available: http://www.ctie.monash.edu.au/emerge/multimedia/jpeg/.

moreCrayons. color cube [Online]. (December 12, 2003). Available: http://www.morecrayons.com/palettes/webSmart/colorcube.php.

National Software Reference Library. NSRL Project Web Site [Online]. (December 29, 2003). Available: http://www.nsrl.nist.gov/.

Nelson, B., Phillips, A., Enfinger, F., and Steuart, C. *Guide to Computer Forensics and Investigations*. Course Technology, Boston, Massachusetts, 2003.

OutGuess. Steganography Detection with Stegdetect [Online]. (December 29, 2003). Available: http://www.outguess.org/detection.php.

Ozer, H., Avcibas, I., Sankur, B., and Memon N. Steganalysis of audio based on audio quality metrics. In: *Proceedings of the SPIE, Security and Watermarking of Multimedia Contents V*, vol. 5020, SPIE, Santa Clara, California, 2003, pp. 55-66. Also available: www.busim.ee.boun.edu.tr/~sankur/SankurFolder/Audio_Steganalysis_16.doc.

Petitcolas, F. A. P. 'mosaic' attack [Online]. (December 29, 2003). Available: http://www.petitcolas.net/fabien/watermarking/2mosaic/index.html.

Provos, N. and Honeyman, P. Detecting Steganographic Content on the Internet. Center for Information Technology Integration, University of Michigan, CITI Technical Report 01-11 [Online]. (August 2001). Available: http://www.citi.umich.edu/techreports/reports/citi-tr-01-11.pdf.

Provos, N. and Honeyman, P. Hide and seek: An introduction to steganography. *IEEE Security & Privacy* (2003) 1(3):32-44. Also available: http://niels.xtdnet.nl/papers/practical.pdf.

Rey, R. F. (ed.). *Engineering and Operations in the Bell System* , 2nd. ed., AT&T Bell Laboratories, Murray Hill, New Jersey, 1983.

Rowland, C. H. Covert Channels in the TCP/IP Protocol Suite. *First Monday* , 1996 [Online]. (January 10, 2004). Available: http://www.firstmonday.dk/issues/issue2_5/rowland/ or http://www.guides.sk/psionic/covert/covert.tcp.txt.

Security Focus. Forensics mailing list, personal communication, December 1-26, 2003.

Seward, J. Debtor's digital reckonings. *International Journal of Digital Evidence*, Fall 2003 [Online]. (January 3, 2004). Available: http://www.ijde.org/docs/03_fall_seward.pdf.

Seward, J. Personal communication, January 2004.

Simmons, G. J. Prisoners' problem and the subliminal channel. In: *Advances in Cryptology: Proceedings of CRYPTO 83*. D. Chaum, ed. Plenum, New York, 1983, pp. 51-67.

spam mimic [Online]. (December 29, 2003). Available: http://www.spammimic.com/.

StegoArchive.com [Online]. (December 30, 2003). Available: http://www.stegoarchive.com/.

U.S. Department of Justice. *Electronic Crime Scene Investigation: A Guide for First Responders.* Office of Justice Programs, National Institute of Justice, Technical Working Group for Electronic Crime Scene Investigation, NCJ 187736, July 2001. Also available: http://www.ncjrs.org/pdffiles1/nij/187736.pdf.

U.S. Department of Justice. *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations* . Criminal Division, Computer Crime and Intellectual Property Section, July 2002. Also available: http://www.cybercrime.gov/s&smanual2002.pdf.

Voloshynovskiy, S., Pereira, S., Pun, T., Eggers, J. J., and Su, J. K. Attacks on digital watermarks: Classification, estimation-based attacks, and benchmarks, *IEEE Communications* (2001) 39(8):118-126.

Warchalking. Warchalking: Collaboratively creating a hobo-language for free wireless networking [Online]. (December 21, 2003). Available: http://www.warchalking.org/.

Wayner, P. *Disappearing Cryptography: Information Hiding: Steganography & Watermarking.* 2nd. ed., Morgan Kaufmann, San Francisco, California, 2002.

WetStone Technologies. Stego Suite [Online]. (May 24, 2004). Available: http://www.wetstonetech.com/f/Stego_Suite_Datasheet_for_web.pdf.

Приложение А: Дополнительные веб-сайты

- Computer Forensics, Cybercrime and Steganography Resources Website, Steganography & Data Hiding - Articles, Links, and Whitepapers page (http://www.forensics.nl/steganography)
- GCK's steganography links (www.garykessler.net/library/securityurl.html#crypto)
- Neil Johnson's Steganography and Digital Watermarking page (http://www.jjtc.com/Steganography/)

Приложение Б: Загружаемые дополнения к статье

Исходный файл (который был скрыт стеганографическими методами), « чистые» файлы-контейнеры, и файлы (содержащие стеганографические данные), упомянутые в данной статье, могут быть разгружены с веб-адреса http://digitalforensics.champlain.edu/fsc/. Используйте пароль "tyui", чтобы дешифровать и извлечь скрытый файл из файлов содержащих стеганографические данные.

- Иллюстрация 5 карта аэропорта: btv_map.gif
- Иллюстрация 6 «чистый» файл-контейнер: mall_at_night.gif
- Иллюстрация 6 файл, содержащий стеганографические данные: mall_at_night_btv2.gif
- Иллюстрация 8 «чистый» файл-контейнер: lightening_jars.jpg
- Иллюстрация 8 файл, содержащий стеганографические данные: lightening_jars_btv.jpg
- Иллюстрация 9 «чистый» файл-контейнер: hitchhiker_beginning.wav
- Иллюстрация 9 файл, содержащий стеганографические данные: hitchhiker_beginning_btv.wav
- Иллюстрация 13 поврежденный файл, содержащий стеганографические данные: disrupt/lighte~1.html

Некоммерческое программное обеспечение, используемое в примерах в этой статье, может быть разгружено со следующих веб-адресов:

- 2Mosaic (http://digitalforensics.champlain.edu/download/2Mosaic_0_2_2.zip)
- Gif-It-Up (http://digitalforensics.champlain.edu/download/Gif-it-up.exe)
- JPHS for Windows (http://digitalforensics.champlain.edu/download/jphs_05.zip)
- Stegdetect (http://digitalforensics.champlain.edu/download/stegdetect-0.4.zip)
- S-Tools (http://digitalforensics.champlain.edu/download/s-tools4.zip)

Приложение B: Коммерческие программные продукты упомянутые в статье

AccessData Corp. Orem, Utah www.accessdata.com
Guidance Software Pasadena, California www.guidancesoftware.com
WetStone Technologies Cortland, New York www.wetstonetech.com

Источник: http://www.fbi.gov/hq/lab/fsc/backissu/july2004/research/2004_03_research01.htm

Перевод
Капинус О.В. (info@computer-forensics-lab.org)
Михайлов И.Ю. (info@computer-forensics-lab.org)
Бочков Д.С.