СТЕГАНОГРАФИЯ В ЗВУКОВЫХ ФАЙЛАХ НА ПРИМЕРЕ ПРОГРАММЫ MP3STEGO

Кесельман М.М.
Степанова И.П.
Томский политехнический университет
srv1@tpu.ru

Введение

Ученые университета Калифорниподсчитали объём медийного контента, которое потребляет каждый американец, включая телевидение, веб-сайты, радио, газеты и пр.

Средний американец затрачивает около 12 часов в день на получение информации. За это время он просматривает 34 гигабайта аудиовизуальной информации и 100 тысяч слов.

В 2010 году в день по статистике на популярный видеоресурс YouTube загружалось 50 тысяч часов видео. Чтобы следить за авторскими правами на YouTubeиспользуется программаContentID, в которой на данный момент хранится уже более 50 млн. цифровых отпечатков.

Технологии, созданные для внедрения в файлы цифровых водяных знаков или цифровых отпечатков, собирают под общее понятие стеганография (от греч. стегано – скрытый, графос – пишу, буквально «тайнопись»).

Помимо внедрения цифровых отпечатков, данные технологии используются для внедрения любой информации в файлы, скрывая своё присутствие от пользователей.

Виды аудиостеганографии

Аудиостеганография – это вид стеганографии, который использует аудиофайлы разных расширений в виде контейнеров для скрываемой информации.

На данный момент времени распространены 3 вида аудиостеганографии, которые используют различные типы сокрытия информации в файлах:

1) Фазовая стеганография.
2) Сокрытие информации в mp3-тегах.
3) Сокрытие информации в mp3-фреймах.

Сокрытие информации в mp3-фреймах

Существует метод, который преобразовывает файл срасширением wmv в файл с расширением Mp3 с добавлением скрытой информации. Методреализован в программе MP3Stego. Идея реализации заключается в том, что данные сначала шифруются, а затем в процессе кодирования MP3-файла (из WAV) подмешиваются в конечный результат. В итоге получается обычный MP3-файл без заметных для слуха искажений, но хранящий в себе закодированные данные.

Алгоритм сжатия Mp3Stego.

Биты встраиваемого сообщения кодируются значениями числа бит, необходимых для кодирования коэффициентов дискретно-косинусного преобразования (ДКП) и масштабирования кодом Хаффмана (см. рис.1). Данные алгоритмы созданы для сжатия информации, например, они используются для сжатия файлов MPEG, JPEG.

Так, если необходимо закодировать нулевой бит, то значение этого числа должно быть четным, а если единичный бит, то нечетным. Необходимую чётность получают путемуменьшения шага квантования. Затем осуществляется квантование коэффициентов ДКП с новым шагом. После чего подсчитывается число бит, необходимых для кодирования коэффициентов ДКП, и суммируется с числом бит, необходимых для кодирования коэффициентов масштабирования. В случае если результат равен необходимой чётности и шум квантования оказался ниже порогового значения, то осуществляется переход к кодированию следующего бита, в отрицательном случае уменьшается шаг квантования, и процедура повторяется.

Пороговое значение вычисляется психоакустической моделью. Психоакустические модели слуха помогают производить компрессию сигнала с потерей информации в высоком качестве. Они помогают за счёт того, что позволяют точно описать, как можно безопасно удалить частоты из исходного сигнала — то есть без значительного ухудшения качества звука.

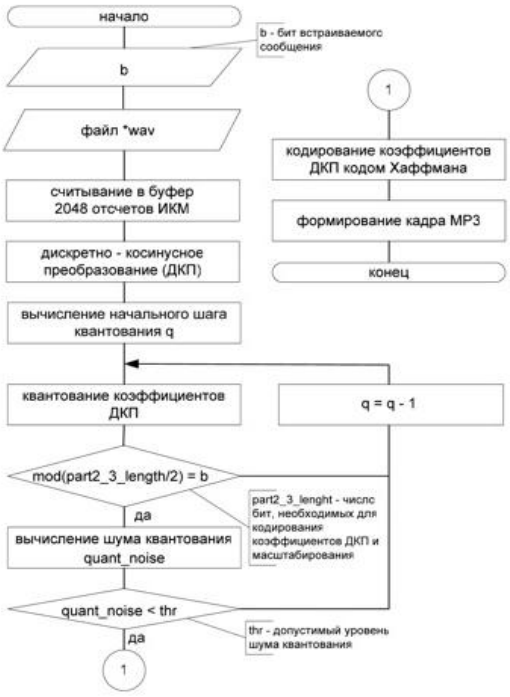


Рис. 1. Алгоритм сжатия информации

Оценка объёма внедряемой информации

Для оценки объёма внедряемой информации введём коэффициент использования контейнера, определяемый выражением $K_{исп} = \frac{V_c}{V_к}$,

где $V_с$ - объём внедренного сообщения в байтах; $V_к$ -объём контейнера с сообщением в байтах.Статистика на ресурсе программы гласит: в проведённых экспериментах по внедрению информации в звуковые файлы эффективность программы MP3Stego показала коэффициент использования контейнера, равный 0.001

Использование и помехи в использовании программы Mp3Stego.

Программа MP3Stego получила широкую популярность за счет того, что была первой программой, которая производила mp3-файлы со скрытыми данными внутри. Также преимуществом этой программы было внедрение информации в Mp3-файл во время сборки самого файла, что гарантировало большую скрытность в отличие от других программ.

Шифратор Mp3Stego требует аудиофайл с расширением wav и текст в формате документа txt. На выходе же программа выводит уже зашифрованный файл в формате mp3.

ДешифраторпрограммыMp3Stego требует только файл со стегоинформацией в формате mp3.

Код программы Mp3Stego открыт для любых экспериментов и улучшений. Автор программы Mp3Stego не обновлял её с 2006 года, следовательно, последней официальной версии Mp3Stego10 лет.

К сожалению,программа Mp3Stego устаревает, её отладкой никто не занимается, следовательно, на более поздних версиях Windowsона показывает нулевой результат.

Также для работы Mp3Stegoтребуется не любой wav-файл, т.к. программа избирательна и можетне запускаться из-за отсутствия в файлекаких-либо данных. Результаты экспериментов с разными аудиофайлами представлены на рис. 2.

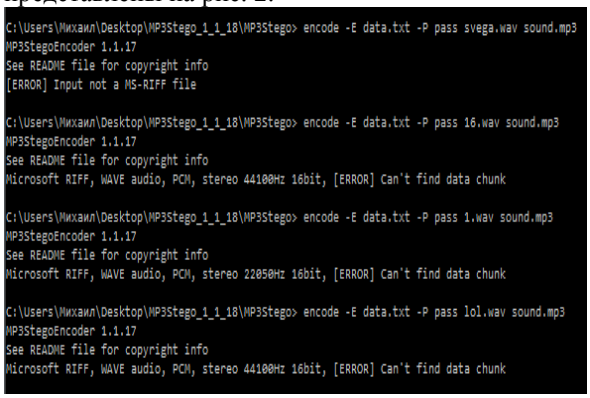
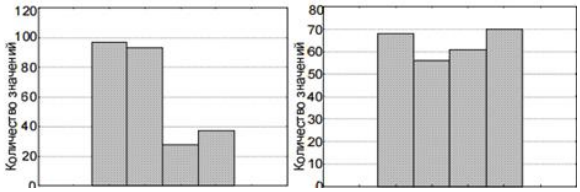


Рис. 2. Скриншот консоли результатов экспериментов с различными аудиофайлами

Детектирование стеговставок в файлах Mp3

Для детектирования стегоставок необходимо проанализировать распределение значений квадрантов фазгармоник[1]. Если во фрейм было встроено сообщение, будет заметнопреобладание некоторых квадрантов. Для пустого фрейма такогопреобладания не будет (см. рис. 3).



Квадрант а) Квадрант б)

Рис. 3. Гистограммы распределений значений фаз по квадрантам для фреймов со вставкой (а) и без вставки (б)

Как видно из рис. 3-а, значительное преобладание 1 и 2 квадрантов свидетельствует о том, что во фрейм был встроен единичный бит.

Заключение

Большое количество медиаинформации, а также плохое регулирование авторского права дают большое поле для использования стеганографии. Информационная достаточность форматов хранения аудиоданных даёт большое количество мест для сокрытия информации в аудиофайлах.

Помимо большого распространения медиафайлов, несомненным плюсом является высокая трудоёмкость нахождения и определение стегосистемы, внедренной с помощью программы Mp3Stego.

В данное время актуальность проектирования качественных и поддерживаемых стегопрограмм растёт, что даёт повышенный интерес к данной теме.

Литература

1. Кокорин П.П. О методах стегоанализа в аудиофайлах // Труды СПИИРАН. Вып. 4. — СПб: Наука, 2007.

2. Википедия [Электронный ресурс] - URL: https://ru.wikipedia.org (дата обращения 11.10.2016).

3. YouTube [Электронный ресурс] - URL: https://www.youtube.com (дата обращения 11.10.2016).