УДК 658.012:004.42

Х. Апьхрейсат, П. П. Кокорин

Санкт-Петербургский институт информатики и автоматизации РАН

МЕТОД ВСТРАИВАНИЯ ЦИФРОВЫХ ВОДЯНЫХ ЗНАКОВ В ЗВУКОВЫЕ ФАЙЛЫ ФОРМАТА MP3

Представлен метод встраивания цифровых водяных знаков (ЦВЗ) в файлы формата MP3. Экспериментально показано, что ЦВЗ сохраняются и успешно восстанавливаются даже после модификации аудиоданных.

К настоящему времени формат MP3 приобрел большую популярность как средство передачи звуковых файлов в сетях, в частности Интернет, при этом происходит копирование и распространение MP3-файлов в нарушение авторских прав. С целью предотвращения незаконного копирования представляет интерес изучение возможности встраивания цифровых водяных знаков (ЦВЗ, watermarking) [1, *Генне О. В.* <www.confident.ru/magazine>] во время кодирования MP3-файла.

Подходы к стеганографическому сокрытию информации в аудиоданных. В последние годы наметился значительный прогресс в создании ЦВЗ, устойчивых к операциям кодирования/декодирования сигнала в формате MP3. Известно несколько алгоритмов встраивания ЦВЗ в поток аудиоданных и их извлечения [2, 3].

Методы ЦВЗ схожи с методами стеганографии, но в плане постановки задачи они принципиально различны. Задача стеганографии — скрыть факт передачи информации и максимально усложнить несанкционированное ее извлечение, в то время как ЦВЗ должен легко обнаруживаться в сигнале, обладать высокой устойчивостью к фальсификации и удалению, а также минимально искажать исходный сигнал.

Существует несколько методов сокрытия информации в потоке аудиоданных.

1) Сокрытие в служебных областях звукового файла (поля ID3-тегов, поля контрольной суммы CRC16, биты заполнения (padding bits) и др.).

2) Встраивание информации в область звуковых данных.

3) Сокрытие в частотно-фазовой области (модификация спектра).

4) Сокрытие информации во временной области (амплитудная модуляция).

Для формата MP3 проще всего использовать методы первой группы. В настоящее время уже создано несколько программ, скрывающих данные в служебных областях MP3-файла, например Rohos, Camouflage, MP3Stego и Steganos. Эти методы позволяют вмещать в файлы значительно больше информации по сравнению с любым другим методом. Как правило, они не ухудшают качества звука, но в то же время их стойкость к атакам активного злоумышленника крайне низка. Также они нестойки по отношению к модификациям звуковых данных.

Методы второй и третьей групп значительно более сложны в реализации, но и более эффективны. Модификация частотно-фазовой области является достаточно популярной областью исследований. Было предложено несколько способов сокрытия данных в частотной и фазовой областях [4]. Данные методы более стойки к атакам. С другой стороны, они в некоторой степени снижают качество звука.

Обнаружение стеганографических вставок в файлах MP3. Рассмотрим некоторые особенности применения атак пассивного рода к методам стеганографии.

Факт использования методов первой группы в основном может быть относительно легко обнаружен (детектирован) простой проверкой значений служебных полей по спецификации

формата файла. Детектирование методов, основанных на сокрытии данных в промежутках между фреймами (элементами файла формата MP3), не вызывает сложностей и сводится к вычислению их размеров.

Обнаружение стеговставок в частотно-фазовой, а также во временной областях представляет гораздо больший интерес для исследований. Данный метод позволяет использовать некоторые статистические характеристики распределений параметров потока аудиоданных, при этом устанавливаются возможные места закладки, полученные данные проверяются по критериям отклонения от ожидаемых значений.

При использовании методов, основанных на изменении фазовой области, как правило, изменяется абсолютное значение фазы некоторых гармоник. Существуют варианты методов, в которых восстанавливаются значения разностей фаз гармоник, смежных фреймов. Ухо человека при восприятии звука чувствительно к изменению разности фаз, но не их абсолютного значения. Вследствие применения других методов фаза выбранной гармоники изменяется на фиксированное значение.

Возможны комбинации нескольких подходов, например, использование модели психоакустического восприятия звука [5], различных методов модуляции и др., при этом возрастает устойчивость таких методов к атакам.

Создание и встраивание ЦВЗ. Представленный в настоящей работе метод встраивания и извлечения ЦВЗ устойчив к различным процедурам сжатия и кодирования сигнала, он состоит из двух основных этапов:

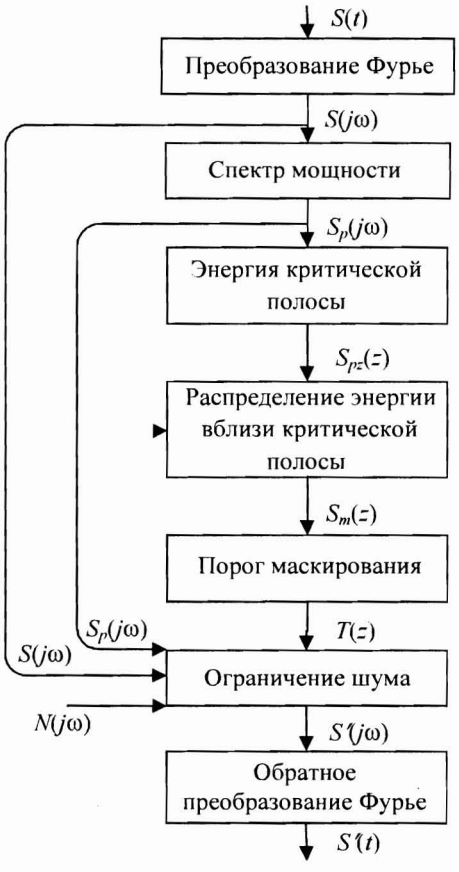

Рис. 1

— создание и внедрение ЦВЗ;

— извлечение ЦВЗ из аудиопотока.

В нем применяются комбинация методов модуляции с быстрым переключением несущей частоты и психоаустическая модель восприятия, полученный сигнал обладает хорошей устойчивостью к атакам. Эффект маскирования частот, присутствующий в психоакустической модели, используется для формирования и сокрытия ЦВЗ в MP3-файле.

Для установления получателем ЦВЗ не требуется исходный MP3-файл. Извлечение ЦВЗ происходит так называемым слепым детектированием (blind detection).

Основные этапы метода представлены на рис. 1 [5, 6]. Для оценки порога маскирования $T(z)$ следует определить характеристики сигнала: спектр мощности $S(j\omega)$, энергию критической полосы $S_{pz}(z)$ и распределение энергии вблизи критической полосы $S_m(z)$. Значение $T(z)$ используется в процедуре встраивания ЦВЗ в „свободные места“, найденные на этапе психоакустического анализа.

Создание ЦВЗ сводится к преобразованию битового потока в поток аудиоданных $x(t)$. Выходной сигнал должен быть устойчивым к удалениям некоторых фрагментов потока. В теории передачи информации эта проблема решатся путем расширения спектра сигнала и коррекции ошибок. В настоящей работе для создания сигнала ЦВЗ $x(t)$ применяется метод коррекции ошибок кодом Боуза—Чоудхури—Хоквенгема (БЧХ) и многопозиционная частотная модуляция (MFSK) [7, 8].

Алгоритм создания сигнала ЦВЗ приведен на рис. 2, где $\{w\}$ — исходная последовательность битов сигнала, $\{w_R\}$ — ЦВЗ после кодирования его методом БЧХ, I, H — размерность матрицы смешивания, $\{заголовок\}$ — заголовок последовательности и $\{d\} = \{заголовок\} + \{w_1\}$ — последовательность битов для последующего расширения частоты и передачи.

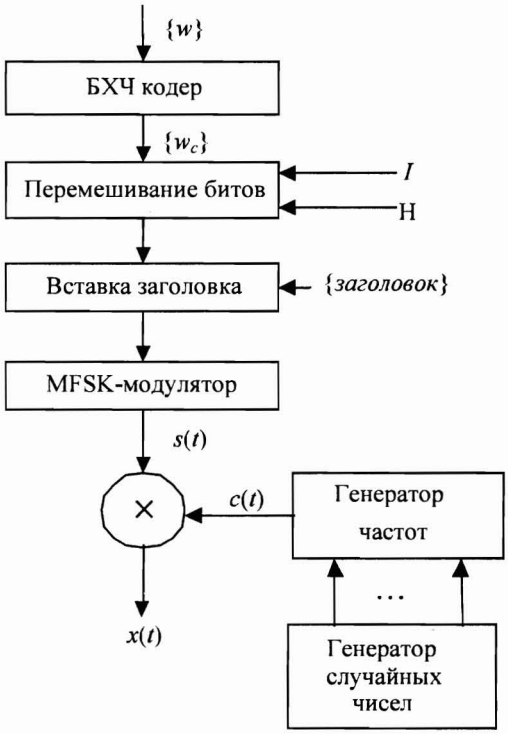

Рис. 2

В данной работе применяется метод, основанный на БЧХ-кодировании для коррекции множественных случайных ошибок [8; см. также *Bowman J. C.* Coding Theory & Cryptography <http://www.math.ualberta.ca/~bowman/m422/m422.pdf>]. Также используется техника перемешивания битов исходного сигнала. Шифратор перемешивает биты данных, поэтому компоненты входного сигнала становятся практически статистически независимыми. Если закодированный символ последовательности представить в виде $\{w_c\} = x_1, x_2, x_3 \cdots x_n$, то эта же последовательность после перемешивания битов, становится следующей: $\{w_I\} = x_1, x_6, x_{11}, \ldots$ для $I = 3$ и $H = 5$. Затем перемешанный сигнал поступает на вход MFSK-модулятора. Модулированный сигнал смешивается с сигналом, поступившим с генератора частот, для последующей модуляции с разбросом по частоте [7]. Модулированный сигнал представляется следующей формулой:

$$x(t) = c(t)S(t).$$

Порог маскирования $T(z)$ используется для встраивания ЦВЗ в „свободные места“ в спектре частот сигнала, а также для формирования самого сигнала ЦВЗ. Для этого $T(z)$ сравнивается со значениями спектра мощности сигнала $S_p(j\omega)$. Компоненты спектра с уровнем ниже порога $T(z)$ считаются маскированными для человеческого слуха [5]. Такие элементы $S_w(j\omega)$ спектра $S_p(j\omega)$ будут удалены, а на их место будет встроен сигнал ЦВЗ. Обозначим вновь вставляемые компоненты как $X_w(j\omega)$.

Новые значения $S_w(j\omega)$ и $X_w(j\omega)$ вычисляются по формулам

$$S'_{wi}(j\omega) = \begin{Bmatrix} Sw_{wi}(j\omega) & S_{pi}(j\omega) \ge T(z) \\ 0 & S_{pi}(j\omega) < T(z) \end{Bmatrix},$$

$$X_{wi}(j\omega) = \begin{Bmatrix} 0 & S_{pi}(j\omega) \ge T(z) \\ X_{wi}(j\omega) & S_{pi}(j\omega) < T(z) \end{Bmatrix},$$

где $i = 1, 2\ldots$ — номер компоненты спектра; z, ω — значения, соответствующие компоненте i.

Для формирования ЦВЗ $X''(j\omega)$ вычисленное новое значение фактора F_z умножается на соответствующий компонент $X'_w(j\omega)$. Значение F_z вычисляется по формуле

$$F_z = \frac{\sqrt{T(z)}}{\max(|X'_w(j\omega)|)}.$$

Окончательные значения заменяемых элементов спектра $output(j\omega)$ вычисляются как сумма $S'_w(j\omega)$ и X''.

Восстановление ЦВЗ. Разработанная система способна восстанавливать ЦВЗ в отсутствие исходного MP3-файла. Для этого вычисляется значение $T(z)$, которое сравнивается со значениями спектра мощности входного сигнала. Все частотные компоненты спектра с уровнем ниже порога $T(z)$ удаляются из спектра. Оставшиеся компоненты преобразуются во временную область $r(t)$. Полученные сигналы анализируются на предмет наличия в них ЦВЗ.

Получатель по имеющемуся у него заголовку битовой последовательности определяет начало сигнала ЦВЗ. Далее на основе сигнала ЦВЗ методом MFSK генерируется сигнал заголовка. Для его обнаружения в потоке $r(t)$ используются адаптивные фильтры высокого разрешения. После этого производится синхронизация и извлечение ЦВЗ [7, 8].

Оценка полученных результатов. Для проверки и оценки устойчивости алгоритма к различного рода сжатиям, он был реализован в среде MatLab. Созданный ЦВЗ встраивался в MP3-файл, который затем редактировался в программе CoolEdit pro v2.1, после чего проводилось извлечение ЦВЗ из сигнала.

После сжатия MP3-файла с различными интенсивностями и амплитудами (–2, –4, –6 и –8 дБ) вычислялся процент успешно восстановленных битов каждого ЦВЗ.

В настоящей работе представлен алгоритм встраивания и извлечения ЦВЗ в звуковые файлы формата MP3, разработанный на основе модели психоакустического восприятия, модуляции с быстрым переключением несущей по случайной последовательности, многопозиционной частотной модуляции и с применением БЧХ кода для коррекции ошибок. Формирование и встраивание ЦВЗ выполняется в соответствии с принципом маскирования частот. Алгоритм показал высокую точность обнаружения ЦВЗ (порядка 87 %) даже после сжатия MP3-файла с различной интенсивностью и применения основных способов искажения сигнала.

СПИСОК ЛИТЕРАТУРЫ

1. *Лысенко А. В.* Построение цифрового водяного знака, устойчивого к потере синхронизации // Изв. УрГУ. 2006. № 43. С. 155—166.

2. *Bassia P., Pitas I.* Robust audio watermarking in the time domain // IEEE Transaction on Multimedia. 2001. Vol. 3, N 2. P. 232—241.

3. *Kirovski D., Malvar H.* Robust covert communication over a public audio channel using spread spectrum // 4th Int. Inf. Hidding Workshop. Pittsburg, USA, 2001.

4. *Чваркова И. Л.* Стеганографические методы скрытия информации в аудиоданных // Электроника. 2003. № 11. С. 54—56.

5. *Zwicker E., Fastl H.* Psychoacoustics: Facts and Models. N. Y.: Springer, 1999.

6. *Roederer J. G.* The Physics and Psychophysics of Music. N. Y.: Springer, 2001.

7. *Torrieri D.* Principles of Spread-Spectrum Communication Systems. N. Y.: Springer, 2004. P. 129—179.

8. *Trappe W., Washington L. C.* Introduction to Cryptography with Coding Theory. Prentice Hall, 2006. P. 392—449.

Рекомендована институтом

Поступила в редакцию
17.05.07 г.