УДК [621.391 + 512.6] : 004.9

Точно обратимый метод встраивания данных в аудиофайл с сохранением гистограммы контейнера

А.А. Жарких, А.В. Гурин, В.Ю. Пластунов
Судоводительский факультет МА МГТУ, кафедра радиотехники и радиотелекоммуникационных систем

Аннотация. В работе представлен метод встраивания данных в аудиофайл-контейнер. Особенностью метода является его устойчивость к атакам, базирующимся на сравнении гистограмм. Достоинствами метода является его точная обратимость и возможность реализации в реальном масштабе времени. Платой за эти достоинства является увеличение объема данных в два раза и появление высокочастотных составляющих в аудиофайле-стего по сравнению с пустым. В работе также представлены результаты реализации алгоритмов метода и оценка относительной емкости контейнера. Метод также может быть использован в задаче встраивания водяных знаков в аудиофайлы.

Abstract. A new method of hiding information inside audio file-container has been described in the paper. This method is robust to steganalysis based on histogram comparison. Method's merits are its exact reversibility and capability of its realization in real time. But it doubles the size of the output file, and some high frequency components appear in audio file stego. The results of algorithm simulation and container capacity estimation have been given. The method can also be used for embedding watermarks in audio files.

Ключевые слова: стеганография, сохранение гистограмм, аудиофайл
Key words: information hiding, audiofile, histogram preserving

1. Введение

В последние годы наблюдается повышенный интерес к стеганографическим методам защиты информации. Об этом свидетельствуют публикации в ведущих мировых научных журналах и в материалах различных международных конференций, например (*Голубев, Емельянов*, 2009; *Грибунин и др*., 2002).

Под стеганографией в данной работе понимаем всю совокупность способов внедрения одного сообщения (внедряемое сообщение, далее просто сообщение) в другое (контейнер). Контейнер, содержащий сообщение, называется стего. Конкретная реализация многих методов внедрения тесно связана с физической природой сигнала-сообщения и сигнала-контейнера. Чаще всего в качестве контейнера выбираются аудиосигналы и изображения. Современная техника использует цифровую запись таких сигналов. В силу своей аналоговой природы, аудиосигналы и изображения, представленные в цифровой форме, содержат избыточную информацию, которую легко заменить на внедряемое сообщение.

В большинстве стеганографических алгоритмов происходит встраивание сообщения в контейнер либо в пустые, не используемые в контейнере места, либо заменой каких-то бит контейнера. При этом, и при встраивании, и при извлечении известен алгоритм, по которому можно найти конкретные координаты элементов встроенного сообщения, например (*Xiaojun Ki, Lewis*, 2009). Такие методы встраивания (по меньшей мере, простейшие из них) неустойчивы к статистическому анализу гистограмм, выявляющему факт встраивания сообщения. Дело заключается в том, что в контейнерах типа изображений и аудио множество отсчетов подчиняется Гауссовскому закону распределения. Встраивание путем подмены информации изменяет эту статистику, что можно увидеть при помощи сравнения гистограмм (*Bin Xia et al.*, 2009).

В качестве альтернативы рассматривается метод встраивания, основная идея которого заключается в том, что вместо перезаписи элементов контейнера происходит их перемещение. Для формирования стего все множество элементов контейнера анализируется на степень сходства. Находятся пары элементов, близкие по какому-то критерию. Стего формируется под управлением битов сообщения. Если передаваемый бит сообщения равен нулю, то 2 элемента некоторой пары не меняются, а если единице, то они переставляются местами. Таким образом, в процессе формирования стего просто переставляются местами некоторые элементы контейнера. Примеры методов встраивания данных в изображение с сохранением гистограммы описаны в (*Кучумов, Курахтенков*, 2009; *Xinpeng Zhang et al.*, 2009).

В данной работе предлагается точно обратимый метод встраивания сообщений в файл формата wav, содержащего несжатый звук PCM (pulse code modulation). Точная обратимость означает, что при извлечении из стего сообщения нет потерь в битовом составе как в сообщении, так и в контейнере.

2. Описание метода и результаты моделирования

Существует два варианта аудиофайла формата wav, содержащего несжатый звук PCM – одноканальный (моно) и двухканальный (стерео). Для описания метода ограничимся одноканальным вариантом. В случае необходимости встраивания в два канала можно расширить базовый одноканальный метод до двухканального двумя способами. Первый способ заключается во встраивании фрагментов битового потока независимо в оба канала. Второй способ базируется на объединении двух каналов в единый контейнер, формировании стего по одноканальному алгоритму и разъединении стего на два канала. В этом варианте каждый отсчет первого и второго каналов считаются соответственно нечетным и четным в объединенном контейнере. После встраивания сообщения в объединенный контейнер стего разделяется по двум каналам по следующему правилу: каждая нечетная пара записывается в первый канал, а каждая четная – во второй канал.

Как для одноканального варианта, так и для двуканального, при любом способе реализации частота дискретизации удваивается.

Вернемся к описанию одноканального варианта.

Обозначим через $x(k)$, k=1, 2 ...n, последовательность отсчетов исходного контейнера. Это последовательность отсчетов аудиосигнала, записанных согласно требованиям теоремы Котельникова. В предлагаемом алгоритме независимо от параметров аудиосигнала и параметров встраивания объем стего строго в два раза больше объема исходного контейнера. Обозначим стего через $y(k)$, k=1, 2 ...2n.

Если частота дискретизации контейнера равна f_c, то частота дискретизации стего равна $f_s = 2f_c$. Это изменение частоты необходимо зафиксировать в выходном файле.

Работа алгоритма встраивания опирается на некоторое "условие встраивания", которое проверяется для каждой текущей пары контейнера ($x_{2k\text{-}1}$ x_{2k}), k=1, 2 ...n. Встраивание одного бита b сообщения в фрагмент расширенного контейнера ($y_{4k\text{-}3}$ $y_{4k\text{-}2}$ $y_{4k\text{-}1}$ y_{4k}) осуществляется при условии $0 < |x_{2k\text{-}1} - x_{2k}| < L_{off}$, где L_{off} – некоторый порог.

Для упрощения анализа и описания алгоритмов введем следующие матрицы:

– вектор-столбец из двух последовательных отсчетов контейнера $(x_{2k\text{-}1}, x_{2k})^T$;

– вектор-столбец из четырех последовательных отсчетов стего $(y_{4k\text{-}3}, y_{4k\text{-}2}, y_{4k\text{-}1}, y_{4k})^T$;

– матрица встраивания бита в контейнер

$$T_{emb}(b) = \begin{pmatrix} 1 & 0 & 1-b & b \\ 0 & 1 & b & 1-b \end{pmatrix}^T;$$

– матрица формирования четырех отсчетов стего без встраивания

$$T_{noemb} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}^T.$$

Действие матрицы встраивания определяется значением текущего бита встраиваемого сообщения b. Для формирования выходной четверки отсчетов необходимо матрицу $T_{emb}(b)$ умножить на вектор-столбец из двух отсчетов контейнера:

$$\begin{pmatrix} y_{4k-3} \\ y_{4k-2} \\ y_{4k-1} \\ y_{4k} \end{pmatrix} = T_{emb}\begin{pmatrix} x_{2k-1} \\ x_{2k} \end{pmatrix} = \begin{pmatrix} x_{2k-1} \\ x_{2k} \\ (1-b)x_{2k-1} + bx_{2k} \\ bx_{2k-1} + (1-b)x_{2k} \end{pmatrix}.$$

Действие матрицы формирования четверки отсчетов стего без встраивания

$$(y_{4k\text{-}3}, y_{4k\text{-}2}, y_{4k\text{-}1}, y_{4k})^T = T_{noemb}\,(x_{2k\text{-}1}, x_{2k})^T = (\,x_{2k\text{-}1}, x_{2k\text{-}1}, x_{2k}, x_{2k})^T.$$

Пусть e=1, если выполняется условие встраивания, и e=0 в противном случае. Тогда на текущем шаге

$$T_{stego} = eT_{emb}(b) + (1 - e)T_{noemb}.$$

Если контейнер содержит нечетное число бит, необходимо сформировать последние биты стего по правилу $y_{2n\text{-}1} = x_n$, $y_{2n} = x_n$, т.е. последний отсчет контейнера дважды повторить и записать в конец выходной последовательности (стего).

При восстановлении контейнера и извлечении сообщения используется то же условие встраивания. Введем матрицы

$$T_{emptycont} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad T_{fullcont} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Тогда

$$T_{cont} = (1 - e) \cdot T_{emptycont} + e \cdot T_{fullcont}.$$

Если после анализа всех четверок стего остались два одинаковых отсчета, один из них дописывается в конец восстановленного контейнера.

Нетрудно показать, что

$$T_{cont}T_{stego} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Это говорит о сохранении битового состава исходного контейнера при восстановлении.

Для вычисления абсолютной емкости контейнера p для данного порога L_{off} необходимо анализировать каждую пару отсчетов на выполнение "условия встраивания". Если условие выполняется, p увеличивается на один бит. Абсолютная емкость p – это число бит сообщения, которое можно записать, относительная емкость η – это отношение числа бит сообщения к общему числу бит контейнера. Если исходный аудиосигнал содержит n отсчетов и записан с точностью q бит на отсчет, то относительная емкость расширенного контейнера равна $\eta = p/(2nq)$. p зависит как от структуры контейнера, так и от порога, и удовлетворяет неравенству $0 \le p \le [n/2]$. Таким образом, в данный контейнер при заданном пороге L_{off} можно встроить $p = p(L_{off})$ бит, т.е. некоторое сообщение bit (t), $t = 1,2, \ldots p$.

Ниже приведен пример использования данного алгоритма к аудиофайлу формата PCM, содержащему запись произнесения числа девятнадцать. Файл моно, частота дискретизации 22 кГц, разрядность 16 бит, длительность 2 сек.

На рис. 2 показаны спектры сигнала контейнера и стего. Видно, что из-за удвоения частоты дискретизации в стего проявляются составляющие с частотами, отсутствующими в контейнере. Эти составляющие находятся в верхней части спектра и прослушиваются при воспроизведении стего. Их заметность существенно зависит от контейнера. При использовании в качестве контейнера записи речи с шумами эти высокочастотные составляющие не воспринимались на слух.

На рис. 3 показана зависимость емкости контейнера от величины выбранного порога. Исходя из того, что используется 16-битный контейнер, величина порога в данном случае может меняться от единицы до 2^{16}-1. График построен до порога, значение которого составляет 2^{13}, так как дальнейшее увеличение порога встраивания не практически увеличивает емкости контейнера. Максимально достижимая емкость для тестируемого контейнера составляет 0,71 % от его размера. Максимально возможная емкость контейнера $\eta = 100\%/2N_{bit}$, N_{bit} – разрядность исходного аудиофайла, для 16-битного файла η = 3,125 %. Как правило, для реальных контейнеров эта величина емкости не достигается из-за присутствия в контейнере одинаковых соседних отсчетов.

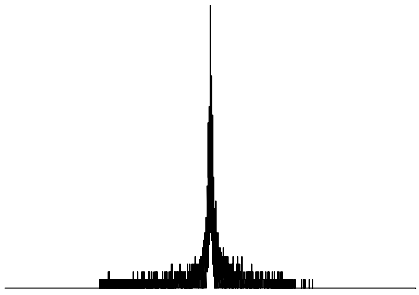

Рис. 1. Нормированная на максимальное значение одинаковых отсчетов гистограмма файла контейнера.

По оси абсцисс значения отсчетов в интервале от -2^{15} до 2^{15}.
По оси ординат отложена частота появления отсчетов, нормированная на число n

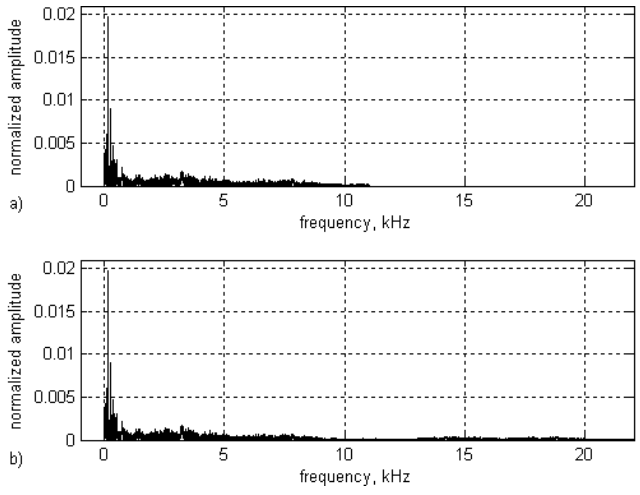

Рис. 2. Спектр контейнера (a) и стего (b)

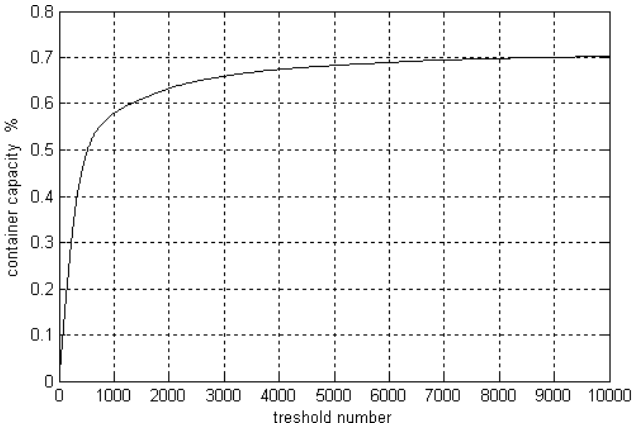

Рис. 3. Зависимость емкости контейнера η от значения выбранного порога, нормированная на размер контейнера

3. Заключение

В работе предложен метод встраивания данных в аудиофайл. Этот метод сохраняет гистограмму исходного файла, но объем записанных данных и частота дискретизации увеличиваются в два раза. Кроме того, возможно появление дополнительных, воспринимаемых на слух, спектральных составляющих. Максимальный объем встраиваемой информации зависит от выбранного порога.

На этапе извлечения полностью восстанавливаются как контейнер, так и сообщение без потерь.

Эту особенность метода можно использовать при распространении записей со встроенной информацией. Пользователь, знающий настройки (величину порога), сможет восстановить контейнер и сообщение. Если настройки неизвестны пользователю, придется мириться с потерей качества звука и невозможностью воспроизвести скрытое сообщение.

При использовании в качестве контейнера записи речи вместе с шумом высокочастотные составляющие стего, появляющиеся вследствие встраивания и увеличения частоты дискретизации, на слух практически неразличимы.

Можно отметить два приложения, в которых предложенный метод может быть использован:

1) при встраивании водяных знаков в аудиофайлы. Внедрение сообщения, являющегося водяным знаком, приведет к появлению воспринимаемых на слух частотных составляющих, простая фильтрация которых приведет к потере встроенного сообщения.

2) встраивание сообщения в аудиофайл среднего качества с сокрытием факта самого встраивания. В этом случае паразитные высокочастотные составляющие не воспринимаются на слух.

Литература

Bin Xia, Xingming Sun, Jiaohua Qin. Steganalysis based on neighbourhood node degree histogram for LSB matching steganography. *MINES 2009 – The 2009 International Conference on Multimedia Information NEtworking and Security (MINES 2009)*, v.2, p.79-82, 2009.

Xiaojun Ki, Daniel Lewis. QIM-based histogram preserving and high capacity steganography for JPEG images, 2009. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.116.5492

Xinpeng Zhang, Shuozhong Wang, Weiming Zhang. Steganography with histogram-preserving and distortion-constraining properties. *MINES 2009 – The 2009 International Conference on Multimedia Information NEtworking and Security (MINES 2009)*, v.1, p.30-34, 2009.

Голубев Е.А., Емельянов Г.В. Стеганография как одно из направлений обеспечения информационной безопасности. *T-Comm Телекоммуникации и транспорт. Спецвыпуск "Технологии информационного общества"*, ч. III, Август, с.185-186, 2009.

Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. *М., "Солон-Пресс"*, 272 с., 2002.

Кучумов А.А., Курахтенков Л.В. Алгоритм стеганографического сокрытия информации с помощью графов. *T-Comm Телекоммуникации и транспорт. Спецвыпуск "Технологии информационного общества"*, ч. III, Август, с.196-198, 2009.