О ВОЗМОЖНОСТИ СКРЫТНОЙ ПЕРЕДАЧИ ИНФОРМАЦИИ ВНУТРИ ЦИФРОВЫХ ЗВУКОВЫХ СИГНАЛОВ НА ОСНОВЕ УЧЕТА ЭФФЕКТОВ МАСКИРОВАНИЯ

Сычев А.В., Александров Э.В.

Воронежский государственный университет
394693, Россия, Воронеж, Университетская пл., 1, каф. информационных систем,
т. (073)-2-789-724, e-mail: sav@is.main.vsu.ru

Реферат. Рассматривается возможность скрытной передачи информации внутри сигнала-контейнера с использованием методов стеганографии. Предлагаются два подхода, основанные на использовании психоакустической модели стандарта MPEG. Оба подхода основаны на замещении избыточной с точки зрения слухового восприятия информации конфиденциальными данными, подлежащими скрытной передаче. При первом подходе модификации подвергается непосредственно дискретный спектр фрагментов цифрового звукового сигнала. Во втором случае скрытая информация внедряется в узкополосные сигналы, образующие исходный широкополосный, в общем случае, сигнал. Обсуждаются некоторые вопросы, связанные с реализацией предложенных подходов.

1. Введение

Доминирующую роль среди средств логического контроля и разграничения доступа к каналам и линиям связи играет шифрование или кодирование сообщений с использованием методов криптографии, изменяющее их исходное представление и, как правило, делающее невозможным их чтение, если неизвестен ключ или алгоритм шифрования (кодирования). Однако значительные достижения, достигнутые в области криптоанализа, в том числе и за счет массированного применения средств вычислительной техники и сетевых технологий, а также все более растущее стремление государств многих стран регламентировать использование криптографических алгоритмов и распределение ключей вынуждают искать альтернативные способы защиты информации.

Методы *стеганографии* (буквальный перевод с греческого – “тайнопись”) позволяют передавать конфиденциальную информацию таким образом, что сам факт такой передачи скрыт от посторонних лиц. Обычно для передачи такой информации в качестве *контейнера* используется один из общедоступных носителей, например, если речь идет об электронных носителях, то это, очень часто, аудиосигнал или видеоизображение. Конфиденциальная информация внедряется в контейнер так чтобы не вызывать заметных для глаза или слуха изменений, способных вызвать какие-либо подозрения.

В зависимости от способа представления сигнала-контейнера существует две группы методов внедрения скрытой информации в контейнер.

Если для выполнения манипуляций с контейнером используется его представление во временной области, то внедрение осуществляется путем модификации наименее значимых битов контейнера либо путем преобразования шума.

В случае спектрального представления (полученного, например, с помощью быстрого косинусного преобразования или одного из вейвлетных преобразований) манипуляции над контейнером затрагивают более значимые области сигнала. Следует отметить, что эта группа методов более устойчива к преобразованиям, направленным на искажение скрытой в контейнере информации.

В данной работе рассматривается возможность использования психоакустических особенностей слухового восприятия человека для скрытной передачи информации внутри оцифрованного звукового сигнала.

На сегодняшний день широкое распространение в компьютерной индустрии получили форматы для представления аудио и видеоданных, разработанные группой MPEG. В частности, для аудио данных (частота дискретизации 44.1 кГц, разрядность отсчетов - 16 бит, стерео) в формате MPEG-I степень сжатия при сохранении хорошего качества звучания варьируется в пределах от 1:4 до 1:12.

Таким образом, учет психоакустических особенностей слухового восприятия позволяет сократить общий размер кодового ресурса, необходимого для представления аудиосигнала, до 10% без заметной потери субъективного качества звучания. Сжатие достигается за счет удаления из спектра звукового сигнала тех компонент, которые маскируются доминирующими тональными и нетональными компонентами спектра.

В данной работе предлагается два подхода к замещению избыточной информации, содержащейся в сигнале-контейнере, полезной информацией, которую необходимо скрытно передать.

2. Общая постановка задачи.

Пусть на входе имеется:

1. Исходный фрагмент оцифрованного звукового сигнала - файл-контейнер.
2. Исходное сообщение (стего-сообщение), которое необходимо поместить в файл-контейнер.

Целью исследования является поиск способов внедрения стего-информации обеспечивающих скрытность передачи данных в аудио-контейнере при максимальной сложности обнаружения и искажения сообщения. Кроме того, желательно обеспечить максимально возможную плотность записи сообщения в носитель.

3. Реализация

Алгоритм можно разбить на несколько последовательных этапов обработки, которые мы рассмотрим более подробно.

Построение психоакустической модели носителя и поиск характеристик носителя, которые можно использовать для передачи скрытой информации.

Психоакустической моделью контейнера будем называть некую обобщённую модель человеческого слуха, позволяющую пренебрегать некоторыми характеристиками звукового сигнала - контейнера и при этом сохранить для человека этот сигнал неотличимым от оригинала.

Разработки в этом направлении уже давно ведутся в исследовательском институте Fraunhofer-IIS. Ими была предложена и успешно использована в проекте MPEG - 1 психоакустическая модель. В данном исследовании за основу берутся эти разработки.

Отношение сигнал/маска (С/М) на выходе этой модели формируется на основе определения тональных (синусоподобных) и нетональных (шумоподобных) компонентов в спектре аудио сигнала.

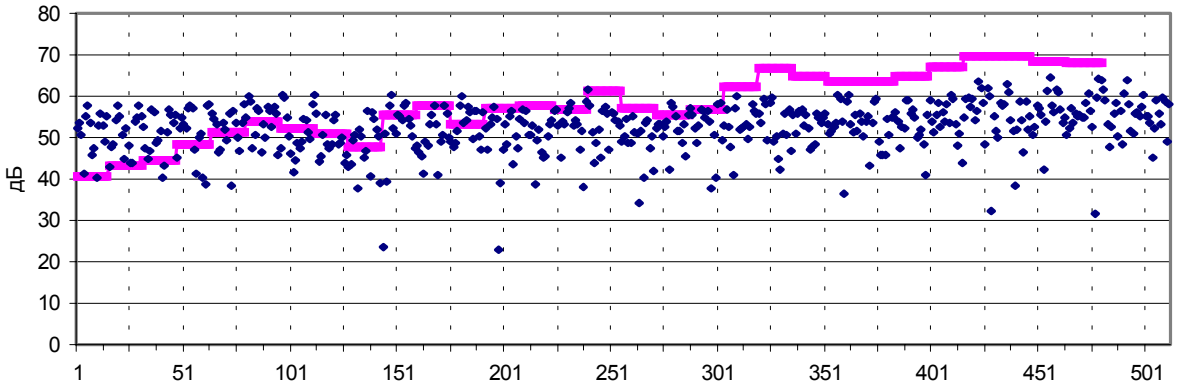

Рисунок 1. Энергетический спектр фрагмента цифрового звукового сигнала и пороги маскирования.

На рисунке 1 приводятся значения компонент энергетического спектра (в дБ) одного из фрагментов звукового сигнала и положение порогов маскирования (для 32 частотных поддиапазонов).

В результате должен получиться некоторый набор спектральных компонент сигнала-контейнера, наиболее существенных для точного воспроизведения сигнала в рамках этой модели. Остальные компоненты спектра, будучи избыточными в рамках этой модели, могут либо просто не учитываться при дальнейшей работе алгоритма либо для них может использоваться приближенное представление.

При первом подходе предлагается использовать оба этих набора для внесения скрываемого сообщения в спектр сигнала-контейнера таким образом, чтобы это не оказывало существенного влияния на субъективное восприятие человеком этого сигнала. Предполагается, что человек не сможет отличить оригинал от изменённого сигнала или, по крайней мере, не сможет заметить искажения, позволяющие обнаружить присутствие скрытой информации.

При втором подходе информация на выходе психоакустической модели используется для выбора числа уровней квантования *N* для каждого из 32 узкополосных сигналов, получаемых при прохождении исходного цифрового звукового сигнала через набор узкополосных фильтров. Очевидно, что чем меньше отношение сигнал-маска для заданного частотного диапазона, тем меньше двоичных разрядов (определяемых как $k = log_2 N$) требуется для представления значений отсчетов этих сигналов, и тем больше разрядов может быть использовано для размещения стего-информации. После внедрения последней узкополосные сигналы с помощью синтезирующего фильтра объединяются, образуя снова широкополосный сигнал, содержащий скрытую информацию.

Размещение стего-сообщения внутри контейнера

Схемы для двух вариантов внедрения стего-сообщений в сигнал-контейнер приведены на рис.2 и рис.3.

1. Использование спектрального представления сигнала-контейнера.

II. На основе БПФ вычисляется спектр фрагмента сигнала-контейнера *S(t)*.

III. С помощью блока "*психоакустическая модель*" определяется положение маскируемых $f_m(i_1)$ и маскирующих частотных компонент. Последние разделяются на тональные $f_t(i_2)$ и нетональные $f_n(i_3)$ компоненты.

IV. Значения частотных компонент представляются в логарифмическом виде. Маскируемые компоненты подвергаются квантованию. Двоичные разряды, которые оказываются в пределах шага квантования, используются для размещения блоков стего-сообщения.

V. После выполнения обратного БПФ восстанавливается исходный фрагмент сигнала-контейнера *S'(t)*, но уже содержащий скрытую информацию.

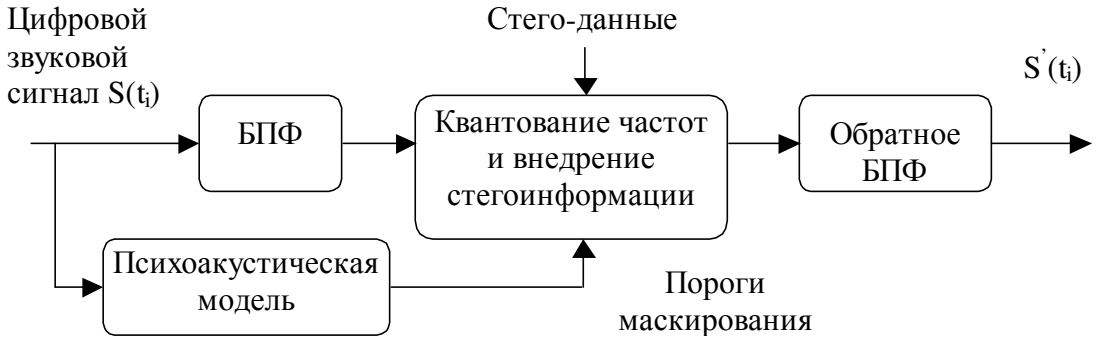

Рисунок 2. Схема внедрения скрытой информации в спектр сигнала-контейнера.

1. Представления сигнала-контейнера в виде суммы узкополосных сигналов.

VI. С помощью набора фильтров формируется набор из 32 узкополосных сигналов $S_k(t_i)$, k = 1,2,…,32, перекрывающих весь спектр исходного сигнала.

VII. На выходе блока "психоакустическая модель" формируются значения отношения сигнал-маска для каждого из 32 частотных поддиапазонов.

VIII. В зависимости от соответствующей величины отношения сигнал-маска выбирается число уровней квантования приведенных к логарифмическому представлению отсчетов узкополосного сигнала $S_k(t_i)$. Двоичные разряды, оказавшиеся в пределах шага квантования, используются для размещения стего-сообщения.

IX. С помощью синтезирующего фильтра восстанавливается исходный сигнал, содержащий скрытые данные.

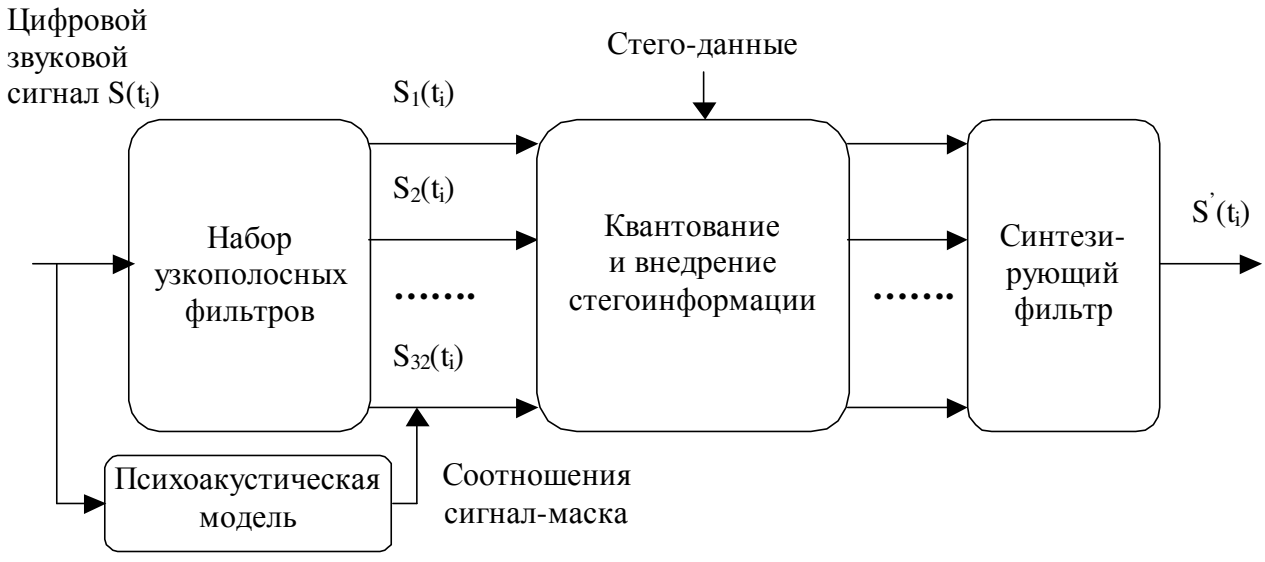

Рисунок 3. Схема внедрения скрытой информации в узкополосных составляющих сигнала-контейнера.

Проблема извлечения стего-сообщения из контейнера

Сжатие информации по стандарту MPEG изначально является сжатием с потерями, поэтому перед ним не ставится требование обратимости преобразования. Поэтому необходимо предусмотреть меры, обеспечивающие возможность восстановления исходного формы представления сигналов. Если в качестве сиганала-контейнера используется оцифрованный звуковой сигнал (чаще всего 16-ти разрядный), ошибка операции округления, используемой при преобразованиях "вещественное число - целое" число составляет порядка 10^{-5} – 10^{-6}, что фактически значительно сокращает двоичных число разрядов, доступных для размещения стего-сообщений. Данная проблема снимается в случае использования вещественного представления отсчетов сигнала.

В случае когда стего-сообщение внедряется непосредственно в спектр сигнала-контейнера, появляется еще одна проблема, обусловленная тем, что энергия *нетональных* маскирующих компонент определяется как сумма *маскируемых* компонент текущего частотного поддиапазона, и поэтому любые изменения в величине последних приводят к изменению величины порога маскирования.

Заключение

Применение стеганографических методов, основанных на учете психоакустических свойств слухового анализатора человека, позволяет добиться высокой степени скрытности передачи конфиденциальной информации внутри звуковых цифровых сигналов-контейнеров, что особенно эффективно может быть использовано в компьютерных сетях и в Internet. Однако вопросы реализации предложенных в докладе подходов требуют более тщательной проработки и исследований.

Библиография.

I. M.D. Swanson, B. Zhu, A.H. Tewk, L. Boney "Robust audio watermarking using perceptual masking" / "Signal Processing", N66, Elsevier, 1998.

II. ISO/IEC "Information technology – Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbits/s". – Part3:Audio. ISO/IEC 11172-3 International Standard, 1993.

ON THE POSSIBILITIES OF THE INFORMATION HIDING INSIDE DIGITAL AUDIO SIGNALS BASED ON AUDIO MASKING EFFECTS

Sytchov A.V., Aleksandrov E.V.

Voronezh State University

394693, Russia, Voronezh, Universitetskaya sq.., 1, VSU, Information Systems Dept.,

phone: (073)-2-789-724, e-mail: sav@is.main.vsu.ru

Abstract. An approach for information hiding inside innocent looking cover medium using steganography methods is proposed. The approach is based on psychoacoustic model used in MPEG specification (audio part). Instead of redundant frequencies in the spectrum of audio signal (working like cover medium) to be removed futher one proposes to allocate blocks of useful information which is required to be hiddenly transmitted. Some possible problems arising from software implementation are discussed.

1. Introduction

Steganography encompasses methods of transmitting secret messages through innocuous cover carriers in such a manner that the very existence of the embedded messages is undetectable. Steganography literally means "covered writing" and is the art of hiding the very existence of a message. The possible cover carriers are innocent looking carriers (images, audio, video, text, or some other digitally representative code) which will hold the hidden information. A message is the information hidden and may be plaintext, ciphertext, images, or anything that can be embedded into a bit stream. Together the cover carrier and the embedded message create a stego-carrier. Hiding information may require a stegokey which is additional secret information, such as a password, required for embedding the information. For example, when a secret message is hidden within a cover image, the resulting product is a stego-image. A possible formula of the process may be represented as:

cover medium + *embedded message* + *stegokey* = *stego-medium*

Steganographic tools can be categorized into two groups: those in the *Image Domain* and those in the *Transform Domain*. *Image Domain* tools encompass bit-wise methods that apply least significant bit (LSB) insertion and noise manipulation. These approaches are common to steganography and are characterized as "simple systems. The *transform domain* grouping of tools include those that involve manipulation of algorithms and image transforms such as discrete cosine transformation (DCT) and wavelet transformation. These methods hide messages in more significant areas of the cover and may manipulate image properties such as luminance. These techniques are typically far more robust than bit-wise techniques; however a tradeoff exists between the about of information added to the image and the robustness obtained.

Two approaches which directly exploits temporal and frequency perceptual masking to guarantee that the stego message, embedded inside digital audio, is inaudible are considered in this report.

Nowadays digital audio and video formats developed by MPEG are very popular in computer applications. It should be noted that for audio (sampling frequency 44.1 kHz, samples digit capacity – 16 bits) in the MPEG-I format the compression ratio obtained is varied from 1:4 to 1:12 while preserving well enough sound quality. So, it is theoretically possible to reduces the bit allocation down to 10% by removing redundant components from the spectrum of digital audio medium.

2. Problem definition.

Let us given:
Initial frame of digital audio – cover medium.
Embedded message to be inserted into cover medium.
The purpose of the research is to find the ways for obtaining the hidden message transmission inside stego-medium while providing high level of security against detection and corruption.

3. Implementation.

An algorithm may include the steps as follows.
The Psychoacustic model of the medium cover reconstuction and the search for components which could be used for data hiding.

The *Psychoacoustic Model* is supposed to be a generalized model of human auditory system, which provides the opportunity of reducing some digital audio features while preserving the identity to original sound. On the output of the psychoacoustic model the signal-to-mask ratio (SMR) is obtained. The SMR calculation is based on determination of tonal (sinusoid like) and nontonal (noise like) masking components inside the spectrum of the cover medium. Other (masked) components are the subject of reduction or replacing by blocks of stego message because these components are not essential for correct (from the subjective point of view) sound reproduction.
Data embedding into cover medium.

Two approaches for stego message embedding into cover medium are available.

Transform Domain approach.

Digital audio frame considered to be cover medium is processed in parallel by FFT and Psychoacoustic Model blocks.

As result masked and masking (tonal and nontonal) frequency components are determined.

The masked components are subjected to reduction or quantization. Binary digits which occur inside the quantization step are used for stego message allocation.

The modified spectrum is subjected to inverse FFT restoring digital audio frame containing hidden data.

Time Domain approach.

Using special filter bank the digital audio frame is splited into 32 narrow band signals overlapping completely the spectrum of the initial audio signal.

On the output of the Psychoacoustic Model 32 values of SMR individually for each of frequency subbands are calculated.

I n dependence of SMR value necessary number of quantization levels is defined. After the samples of narrow band signals have been quantizated the remaining binary digits which are inside of quantization step are used for stego message blocks allocation.

After passing synthesis filter the stego-cover is obtained.

Some problems related to stego message extraction from stego-medium.
MPEG specification is does not provide lossless compression of data. That is why special techniques for uniquely inversable transformation should be developed. Because in the most of multimedia applications the 16 bit audio samples are used for digital audio encoding the “float-to-integer” conversion cause to the roundoff error of the order of $10^{-5} – 10^{-6}$. In fact, this reduces the number of useful binary digits for hidden data allocation.

Conclusion

Steganographic applications based on features of the human auditory system provide the high level of security. This feature may be effectively used especially in computer networks and Internet. But the implementation of the approaches proposed in the report should be carefully elaborated and one requires special research.

Bibliography.

1. M.D. Swanson, B. Zhu, A.H. Tewk, L. Boney “Robust audio watermarking using perceptual masking” / “Signal Processing”, N66, Elsevier, 1998.

II. ISO/IEC “Information technology – Coding of moving pictures and associated audio for digital storage media up to about 1,5 Mbits/s”. – Part3:Audio. ISO/IEC 11172-3 International Standard, 1993.