УДК 004.056

МЕТОДЫ ЗАЩИТЫ АУДИОФАЙЛОВ ОТ НЕСАНКЦИОНИРОВАННОГО КОПИРОВАНИЯ И РАСПРОСТРАНЕНИЯ

Борисова С.Н.

ФГБОУ ВПО «Пензенский государственный технологический университет», Пенза, e-mail: romi_s@list.ru

Настоящая статья посвящена вопросам защиты авторских прав музыкальных произведений, представленных в цифровом виде, а именно, защите от несанкционированного копирования и распространения во всемирной сети Интернет. Основная проблема при отслеживании подобного контента веб-страниц – множество копий аудиофайлов, умышленно модифицированных и имеющих неоригинальное название. В данной работе предложено два способа отслеживания несанкционированного размещения музыкальных файлов в сети Интернет. Первый способ – использование методов стеганографии, а именно встраивание цифрового водяного знака в защищаемый файл. Рассмотрены различные технологии встраивания, основанные на представлении файла во временной и частотной области. Удовлетворительными с точки зрения устойчивости к различным искажениям признаны методы встраивания, использующие спектральное представление сигнала. В качестве второго способа предложено формировать цифровой отпечаток файла (сигнатуру). Анализ англоязычной литературы позволил выделить два способа формирования цифровых отпечатков аудиофайлов, основанных на анализе частотно-временных признаков сигнала: измерение местоположения точек локальных максимумов амплитуды в спектрограммах и покадровое вычисление известных характеристик звука, устойчивых к искажениям.

Ключевые слова: аудиофайл, авторское право, цифровой водяной знак, цифровой отпечаток файла, стеганография, спектрограмма, сигнатура

METHODS OF PROTECTION FROM THE AUDIO-FILES FROM UNAUTHORIZED COPYING AND DISTRIBUTION

Borisova S.N.

Penza State Technological University, Penza, e-mail: romi_s@list.ru

This article is devoted to the protection of musical works copyright, presented in digital form, namely from unauthorized copying and distribution in the World Wide Web. The main problem of tracking such web-page content is multiple copies of audio files, deliberately modified and having unoriginal name. In this paper, we propose two ways of tracking the unauthorized placement of music files on the Internet. The first way is embedding the digital watermark into the protected file. Various embedding technologies, based on the representation of the file in the time and frequency domain. Methods of embedding are recognized satisfactory in terms of resistance to various distortions, using the spectral representation of the signal. The second method is to form a digital fingerprint (the signature) of the file. Analysis of English literature makes it possible to identify two ways of generating digital audio-fingerprint based on the analysis of time-frequency features of the signal: the measurement of locations of local maxima points in the amplitude spectrograms and shot-by-shot calculation of known sound characteristics, resistant to distortions.

Keywords: audio-file, copyright, digital watermark, digital fingerprint, steganography, the spectrogram, the signature

В эпоху быстро растущих интернет-технологий и доступности мультимедийных вычислительных средств защита прав интеллектуальной собственности стала жизненно важным вопросом. В частности, это актуально для изображений, аудио- и видеоинформации. Повсеместное использование глобальных сетей, а также распространение электронных средств массовой информации дают возможность художникам и фотохудожникам демонстрировать свои работы множеству людей по всему миру, фотокорреспондентам – оперативно размещать репортажи о происходящих событиях. С одной стороны, эти данные должны быть доступны любым пользователям, так как это основное их назначение, с другой стороны, подобный свободный доступ к информации делает ее уязвимой для угроз несанкционированного копирования и распространения от чужого имени. Несколько иначе дело обстоит с музыкальными и кинопроизведениями, представленными в цифровом виде. Собственники подобных произведений, как правило, нацелены на получение выгоды от их продаж и не публикуют их в открытом доступе. Однако многие владельцы авторских прав обеспокоены защитой от любого незаконного копирования своих работ. Поэтому в последнее время проводится серьезная работа по отслеживанию несанкционированного распространения музыкальных файлов и фильмов. В связи с этим необходимо упомянуть, что в нашей стране действует закон об авторских и смежных правах [2], который регулирует отношения,

возникающие в связи с созданием и использованием произведений науки, литературы и искусства, фонограмм, исполнений, постановок, передач организаций эфирного или кабельного вещания и устанавливает ответственность за несанкционированное использование этих произведений, а также за плагиат (ст. 15 закона об авторском праве). Однако данный закон не всегда помогает правообладателям защитить свои произведения. И как показывает практика, можно найти сотни копий аудио- и видеофайлов на различных web-страницах. Кроме того, названия фильмов и музыкальных файлов при несанкционированном распространении часто меняются, что затрудняет их выявление правообладателями. Поэтому автору не только необходимо отыскать все копии, существующие в сети, но и, согласно закону [2], доказать первоочередность своего авторства.

Для защиты аудиофайлов от несанкционированного копирования и распространения, а также доказательства первоочередности своего авторства можно предложить разные способы, позволяющие идентифицировать файлы:

1. Встраивать в аудиофайлы идентификационную метку. Данные метки незаметны для человеческого слуха, но легко обнаруживаются специальными детекторами.

2. Вычислить «цифровой отпечаток» звукового файла и хранить его в базе данных (БД). Цифровой отпечаток будет занимать значительно меньше места в БД, нежели сам файл, что позволит создать большую базу отпечатков.

Цифровые водяные знаки

Первый способ относится к методам стеганографии, а именно к одному из направлений стеганографии – цифровым водяным знакам (ЦВЗ). Цель стеганографии – скрытие факта передачи защищаемой информации. Отправитель внедряет секретное сообщение в какой-либо объект (контейнер) и только принимающая сторона, знающая о факте передачи, может извлечь это сообщение. В отличие от обычной стеганографии, ЦВЗ не ставят целью скрыть факт встраивания, скорее наоборот. В ЦВЗ контейнерами служат мультимедийные файлы (изображения, видео- и аудиофайлы), а скрываемыми данными – различная информация, идентифицирующая автора объекта. Учитывая то, что нарушитель знает или может догадываться о наличии ЦВЗ и предпринять попытки модификации защищаемого файла, при внедрении информации в аудио-сигналы существует определенный ряд требований [3]:

– скрываемая информация должна быть стойкой к наличию различных окрашенных шумов, сжатию с потерями, фильтрованию, аналогово-цифровому и цифро-аналоговому преобразованиям;

– скрываемая информация не должна вносить в сигнал искажения, воспринимаемые системой слуха человека (ССЧ);

– попытка удаления скрываемой информации должна приводить к заметному повреждению контейнера (для ЦВЗ) или его непригодности для восприятия;

– ЦВЗ должен однозначно идентифицировать автора защищаемого файла;

– скрываемая информация не должна вносить заметных изменений в статистику контейнера.

Цифровые водяные знаки имеют небольшой объём, однако при их встраивании должны использоваться сложные методы встраивания. Для внедрения скрываемой информации в аудиосигналы можно использовать методы, применимые в других видах стеганографии [1, 3]. Например, можно внедрять информацию, замещая наименее значимые биты (все или некоторые). Или можно строить стегосистемы, основываясь на особенностях аудиосигналов и системы слуха человека.

На настоящий момент популярны следующие методы создания ЦВЗ для аудиофайлов (рис. 1), подробно описанные в [1, 3, 4, 6]:

1. *Метод замены наименьших значащих бит (НЗБ)* является простейшим способом внедрить конфиденциальные данные в иные структуры данных. Используя звуковой сигнал, путем замены наименьших значащих бит каждой точки осуществления выборки, представленной двоичной последовательностью, можно зашифровать значительный объем информации. Сам процесс встраивания информации аналогичен тому, который используется для изображения-контейнера [3], то есть в каждом значении амплитуды наименьший значащий бит заменяется на бит сообщения. Недостаток метода – слабая устойчивость к посторонним воздействиям на сигнал (сжатие, воздействие шумов).

2. *Метод внедрения информации с использованием эхо-сигнала*. Данный метод предполагает встраивание ЦВЗ в контейнер путем изменения параметров эхо-сигнала. К параметрам эхо, несущим внедряемую информацию, относятся: начальная амплитуда, время спада и сдвиг (время задержки между исходным сигналом и его эхо). Оригинальный сигнал смешивается с одной или несколькими точными копиями, которые слегка отстают во времени. Когда сдвиг между оригинальным сигналом и его эхо уменьшается, ССЧ человека воспринимает эхо-сигнал как добавочный резонанс. Метод слабо устойчив к сжатию.

Рис. 1. Методы встраивания ЦВЗ

3. *Метод фазового кодирования*. Фазовый метод кодирования заключается в разбиении сигнала на сегменты и замещении фазы начального сегмента аудиосигнала на фазу, которая характеризует данные. Фазы последующих сегментов регулируются в целях сохранения разности фаз между сегментами. Метод фазового кодирования является одним из методов, который устойчив к сжатию и воздействию шумов.

4. *Метод растяжения спектра*. В данном методе псевдослучайная последовательность, представляющая собой «белый шум», модулируется сигналом несущей, представляющей ЦВЗ и затем добавляется к аудиосигналу-контейнеру. Метод является устойчивым к некоторым посторонним воздействиям.

5. *Time base modulation (изменение масштаба временной оси)* [5]. Временная ось делится на части, которые незначительно растягиваются и сжимаются. Местоположение и степень сжатия/растяжения являются величинами, которые используются для кодирования информации. Метод устойчив к сжатию, искажениям.

Некоторые из методов были реализованы. На рис. 2 представлены временные диаграммы для исходного *C* и модифицированного *S* контейнера, а также разницы между контейнерами (*C*-*S*) для метода НЗБ.

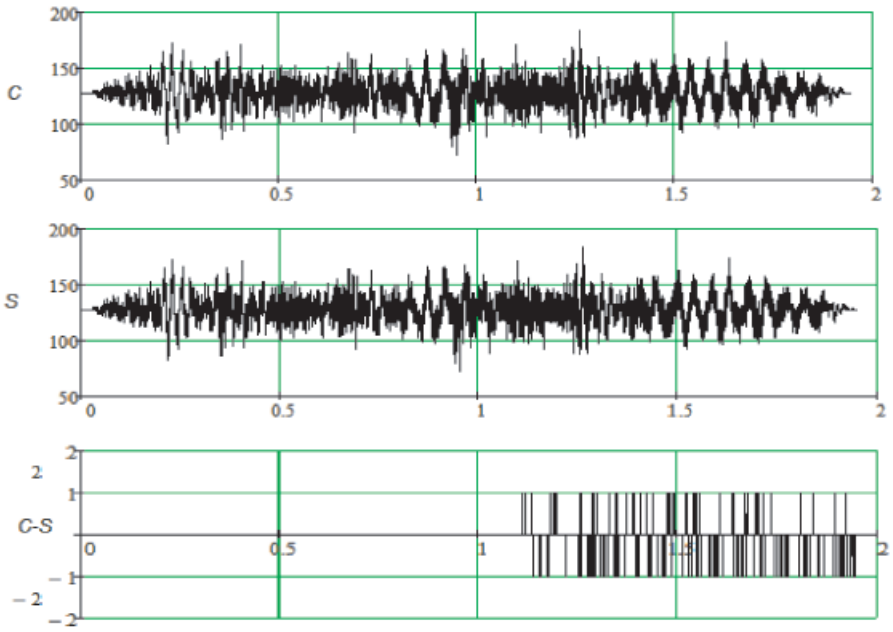

Рис. 2. Результаты реализации метода встраивания в НЗБ

Показатели звукового искажения PSNR (максимальное отношение сигнал – шум), AF (качество звучания), SC (индекс структурного подобия) для четырех методов встраивания (НЗБ, фазовое кодирование, расширение спектра, использование эхо-сигнала) представлены на рис. 3. Значения показателей звукового искажения оценивались в соответствии с выражениями, представленными в источнике [3]. Индекс структурного содержания является наиболее близким параметром для оценки искажений, воспринимаемых ССЧ. Соответственно полученным результатам лучшие показатели наблюдаются у методов НЗБ, фазового кодирования и расширения спектра. Однако метод НЗБ является самым неустойчивым к любым искажениям сигнала, его преобразованиям и зашумлению. Метод использования эхо-сигнала показывает самые неудовлетворительные оценки звукового искажения. Поэтому можно сделать вывод, что для ЦВЗ предпочтительно использовать методы, производящие встраивание в частотную область сигнала.

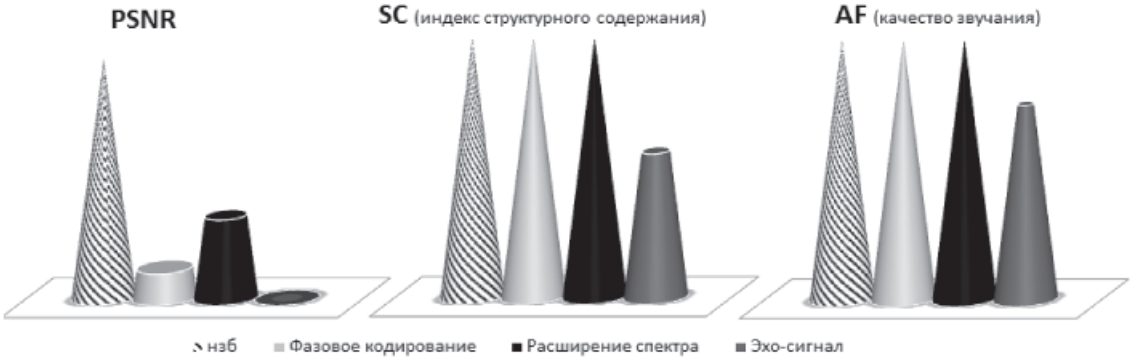

Рис. 3. Сравнение показателей звукового искажения для четырех методов встраивания

Цифровые отпечатки файлов

Второй способ защиты звуковых файлов от несанкционированного копирования и распространения – использование цифровых отпечатков файлов, назовем их «звуковые отпечатки», так как анализируются звуковые файлы. Технология цифровых отпечатков используется в настоящее время для защиты различного рода информации, большей частью для защиты текстовых документов от утечек. Однако данная технология может быть распространена и на другие типы файлов, например на изображения, аудио- и видеофайлы. При формировании цифровых отпечатков нужно учитывать 3 важных момента: во-первых, отпечаток должен быть максимально компактным, во-вторых, отпечаток должен учитывать характерные особенности файлов (звуковые или визуальные), в-третьих, для использования отпечатков необходимо создать базу отпечатков всех защищаемых файлов. Минкомсвязи РФ летом 2014 года заявило о планах создания общероссийского реестра аудио-видео- и другого контента, где можно будет регистрировать все цифровые отпечатки. Это в свою очередь приведет к тому, что на web-страницах нельзя будет размещать копии музыкальных файлов даже в измененном виде (в другом размере или качестве, с вырезанными фрагментами и т.п.). Такая технология защиты авторских прав используется на западе, а с июня 2014 года и в РФ, в частности, в социальной сети «В Контакте». Однако детали формирования звуковых отпечатков и их проверки не публикуются. С учетом всего вышесказанного возникает вопрос: какими способами можно формировать звуковой отпечаток и какие особенности звукового файла необходимо использовать для успешной его идентификации?

Как уже было отмечено выше, звуковой отпечаток (сигнатура) представляет собой краткое описание аудиофайла, учитывающее его звуковые особенности. Причем должны учитываться такие особенности, которые являются не только стойкими к различным модификациям звукового файла (сжатию, аналого-цифровому и цифро-аналоговому преобразованию, фильтрованию и т.п.), но и имеющими отношение к его восприятию. Например, при использовании криптографических хеш-функций в качестве цифрового отпечатка два звуковых трека одного музыкального произведения, но исполненного в разных интерпретациях, будут соответствовать разным цифровым отпечаткам. То есть функцию вычисления сигнатуры аудиофайла необходимо построить таким образом, чтобы воспринимаемые одинаково звуковые объекты приводили к одинаковым

отпечаткам. Вторая проблема использования цифровых отпечатков – выбор эффективного алгоритма сравнения отпечатков и алгоритма отсечения ложных срабатываний. В то же время получение и поиск отпечатков должно быть быстрым и простым.

Проведенный анализ англоязычной литературы по выбранной тематике позволил предложить использовать для отслеживания нелегального аудиоконтента web-страниц технологии, которые применяются для поиска интересующих пользователя аудиофайлов в различных базах данных, например в сервисе *Shazam*. Из всех рассмотренных было выделено два эффективных алгоритма распознавания звука, которые успешно применяются для поиска аудиоконтента в базах данных.

В работе [7] музыкальное произведение рассматривается как частотно-временной график, называемый спектрограммой. На одной оси откладывают время, на другой – частоту, на третьей – интенсивность пиков спектрограммы, в качестве которых используют точки локального максимума амплитуды. Каждая точка на графике представляет интенсивность конкретной частоты в данный момент времени. Местоположение данных точек на сетке «частота – время» мало меняется при зашумлении. Для улучшения поиска для каждого пика спектрограммы находится соседняя точка вперед по оси времени, называемая опорной. Затем эти точки связываются в пары, называемые созвездиями. Цифровой отпечаток представляет собой хеш-таблицу, в которой роль ключа исполняет значение частоты для пика интенсивности и для его опорной точки, напротив которых ставится время в секундах от начала трека. Процесс формирования одного элемента хеш-таблицы представлен на рис. 4.

В работе [4] для формирования отпечатков звукового файла используются хорошо известные характеристики звука, устойчивые к искажениям, такие как коэффициенты Фурье, коэффициенты косинусного преобразования Фурье (MFFC), неравномерность спектра (spectral flatness) коэффициенты линейного кодирования с предсказанием (LPC) и другие. Предложенная схема получения отпечатков базируется на этом подходе обработки потока. Обзор схемы показан на рис. 5 (а). Сначала звуковой сигнал разделяется на кадры (фреймирование). Для каждого кадра вычисляется набор характеристик звукового файла, в качестве которых выбраны коэффициенты преобразования Фурье (ПФ). Для вычисленных параметров ПФ сохраняется только абсолютная величина спектра, то есть спектральная плотность мощности (ABS). Для одного кадра формируется один суб-отпечаток длиной 32 бита.

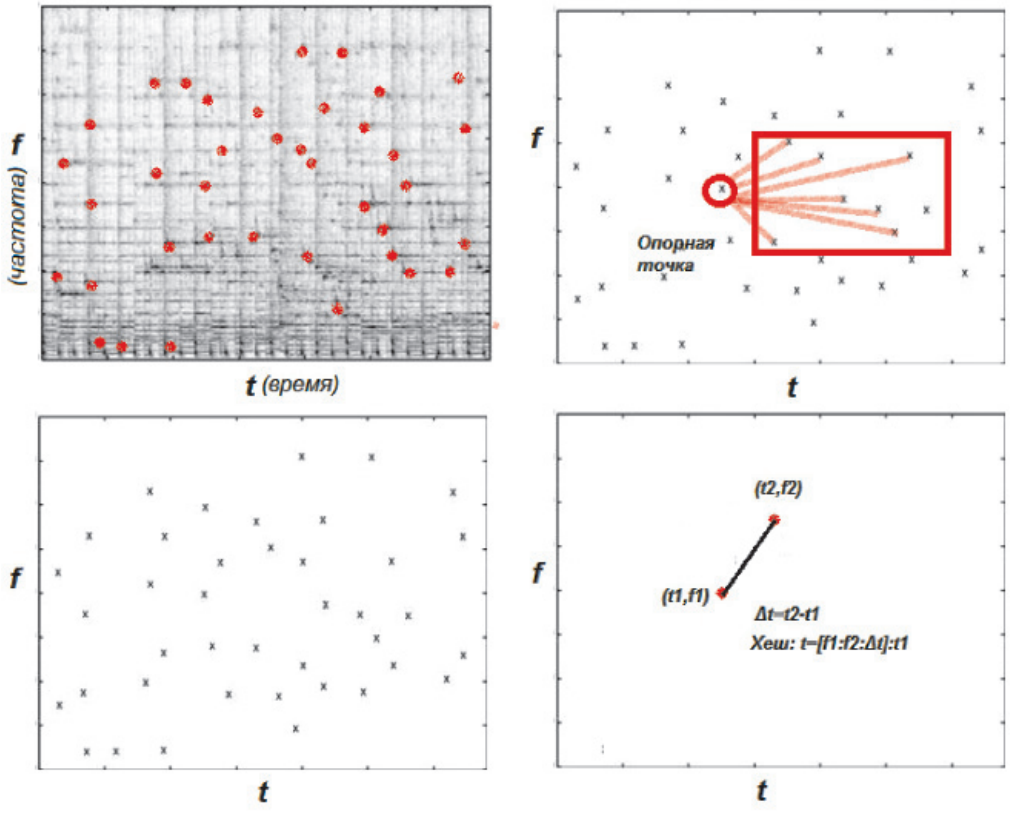

Рис. 4. Схема формирования отпечатка звукового файла по его спектрограмме

Для получения суб-отпечатка для каждого кадра выбираются 33 неперекрывающихся диапазона частот от 300 до 2000 Гц. Каждый бит *m* суб-отпечатка для кадра *n* определяется по формуле

$$F(n,m) = \begin{cases} 1, \text{ если } E(n,m) - E(n,m+1) - (E(n-1,m) - E(n-1,m+1)) > 0, \\ 0, \text{ если } E(n,m) - E(n,m+1) - (E(n-1,m) - E(n-1,m+1)) \le 0, \end{cases} \quad (*)$$

где $E(n,m)$ – энергия диапазона частот *m* для *n*-го кадра. Бит «1» кодирует белый пиксель, бит «0» – черный пиксель. Разность энергий (одновременно по осям времени и частоты) – свойство, которое очень устойчиво ко многим видам обработки. Суб-отпечатки объединяются в блоки по 256 (рис. 5 (б)). Такой способ формирования отпечатков описывает лишь 3 секунды звукового файла, что, по мнению авторов метода, является достаточным для точной идентификации звукового файла.

отпечатков для идентификации файлов заключается в том, что решение о сходстве или различии аудиофайлов выдается с некоторой долей вероятности, то есть результат распознавания файлов оценивается двумя ошибками: ошибкой первого рода (файл не найден при его наличии) и ошибкой второго рода (ошибочно найден другой файл). Ошибка второго рода у метода спектрограмм примерно в 2,5 раза выше, чем у метода покадрового вычисления отпечатков.

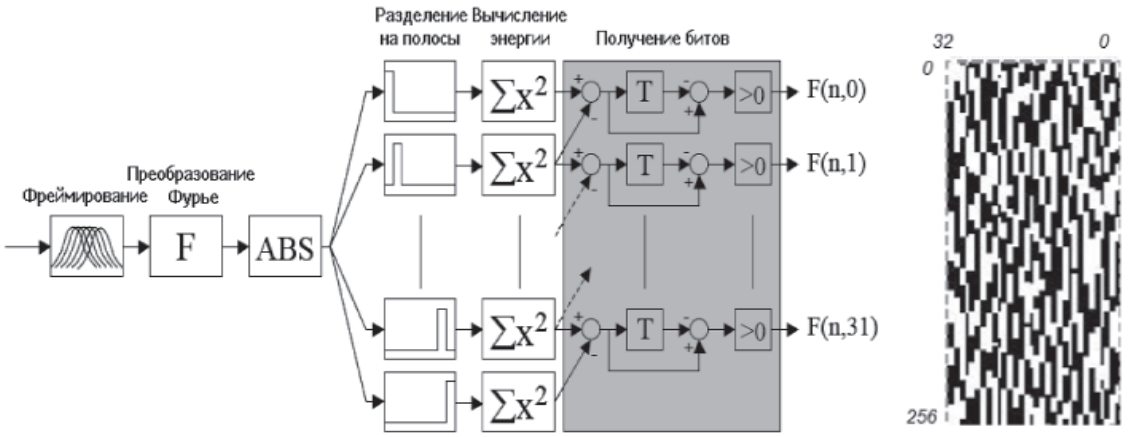

Рис. 5. Схема формирования отпечатка звукового файла по его частотным характеристикам

Выводы

В результате проведенных в данной работе размышлений можно сделать следующий вывод: предложенные методы защиты аудиофайлов от несанкционированного копирования и распространения вполне могут быть использованы для поставленной задачи. Однако, чтобы отследить звуковые файлы по встроенной идентификационной метке (ЦВЗ), необходимо быть убежденным, что эта метка там присутствует изначально. Но невозможно встроить идентификационную метку во все копии файлов. При использовании второго способа отслеживания аудиофайла по его цифровому отпечатку можно отследить сколь угодно модифицированные копии файла-оригинала. Особенность использования цифровых

Для улучшения вероятностных показателей обнаружения предполагается введение некоторой избыточности, что может увеличить время поиска. Для доказательства представленных выше предположений и получения более детальных характеристик работы алгоритмов необходима их практическая реализация, что является темой для дальнейшего исследования.

Список литературы

1. Грибунин В.Г., Оков И.Н., Туринцев И.В. Цифровая стеганография. – М.: Солон-Пресс, 2002. – 272 с.

2. Закон Российской Федерации «Об авторском праве и смежных правах» от 09.07.1993 г. № 5351-1 // Официальный сайт компании «Консультант плюс» [Электронный ресурс] – Режим доступа: http://www.consultant.ru/popular/avtorpravo/ (дата обращения 22.05.2015).

3. Конахович Г.Ф., Пузыренко А.Ю. Компьютерная стеганография. Теория и практика. – Киев: МК-Пресс, 2006. – 288 с.

4. Avery Li-chun Wang. An Industrial-Strength Audio Search Algorithm. [Электронный ресурс] – Режим доступа: http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf (дата обращения 22.05.2015).

5. Bender W., Gruhl D., Morimoto N., Lu A. Techniques for Data Hiding // IBM Systems Journal. 1996. № 35.

6. J. Foote, J. Adco, and A. Girgensohn. Time base modulation: a new approach to watermarking audio. [Электронный ресурс] – Режим доступа: http://www.fxpal.com/publications/FXPAL-PR-03-212.pdf (дата обращения 22.05.2015).

7. Jaap Haitsma, Ton Kalker. A Highly Robust Audio Fingerprinting System. [Электронный ресурс] – Режим доступа: http://ismir2002.ismir.net/proceedings/02-FP04-2.pdf (дата обращения 22.05.2015).

References

1. Gribunin V.G., Okov I.N., Turincev I.V. Cifrovaja steganografija. Moscow: Solon-Press, 2002. 272 p.

2. *Zakon Rossijskoj Federacii «Ob avtorskom prave i smezhnyh pravah» ot 09.07.1993 g* (Law No. 5351-I of July 9, 1993 on Copyright and Related Rights) // Oficialnyj sajt kompanii «Konsultant pljus». [Jelektronnyj resurs] – Rezhim dostupa:http://www.consultant.ru/popular/avtorpravo (accessed 22 May 2015).

3. Konahovich G.F., Puzyrenko A. Ju. Kompjuternaja steganografija. Teorija i praktika. Kiev: MK-Press, 2006. 288 p.

4. Avery Li-chun Wang. An Industrial-Strength Audio Search Algorithm. Available at: http://www.ee.columbia.edu/~dpwe/papers/Wang03-shazam.pdf (accessed 22 May 2015).

5. Bender W., Gruhl D., Morimoto N., Lu A. Techniques for Data Hiding // IBM Systems Journal. 1996. no. 35.

6. J. Foote, J. Adco, and A. Girgensohn. Time base modulation: a new approach to watermarking audio. Available at: http://www.fxpal.com/publications/FXPAL-PR-03-212.pdf (accessed 22 May 2015).

7. Jaap Haitsma, Ton Kalker. A Highly Robust Audio Fingerprinting System. Available at: http://ismir2002.ismir.net/proceedings/02-FP04-2.pdf (accessed 22 May 2015).

Рецензенты:

Сальников И.И., д.т.н., профессор, заведующий кафедрой «Вычислительные машины и системы», Пензенский государственный технологический университет», г. Пенза;

Султанов Б.В., д.т.н., профессор кафедры «Информационная безопасность систем и технологий», Пензенский государственный университет, г. Пенза.