

ОБНАРУЖЕНИЕ МАРКЕРА В ЦИФРОВОМ АУДИОСИГНАЛЕ АВТОРИЗОВАННЫМ ПОЛУЧАТЕЛЕМ¹

Гофман Максим Викторович

К.т.н., доцент, Петербургский государственный
университет путей сообщения Императора
Александра I
maxgof@gmail.com

DETECTION OF MARKER IN DIGITAL AUDIO SIGNAL BY AUTHORIZED RECIPIENT²

M. Gofman

Summary. A method for robust watermarking of digital audio signals is being developed, focused on the transmission of watermarked audio signals in an air audio channel. Special attention is paid to the development of a method for detecting the presence of a watermark in a digital audio signal by an authorized receiver. The article contains the results of field experiments to assess the noiseproof of the transmission of watermarked audio signals through an air audio channel.

Keywords: audio signal, steganography, audio signal marking, air audio channel, communication noiseproof.

Аннотация. Разрабатывается метод устойчивого маркирования цифровых аудиосигналов, ориентированный на передачу маркированных аудиосигналов в условиях воздушного аудиоканала. Особое уделяется разработке метода обнаружения наличия маркера в цифровом аудиосигнале авторизованным получателем. Статья содержит результаты натурных экспериментов по оценке помехоустойчивости передачи маркированных аудиосигналов через воздушный аудиоканал.

Ключевые слова: аудиосигнал, стеганография, маркирование аудиосигналов, воздушный аудиоканал, помехоустойчивость связи.

Введение

С ростом количества и доступности мобильных медиаустройств, смартфонов, цифровых диктофонов и других разнообразных инструментов вещания и записи акустических сигналов возникает задача создания методов маркирования аудиосигналов, способных помимо слуховой транспарентности (не слышимость) обеспечить устойчивость маркера, даже в условиях шумовой обстановки воздушного аудиоканала и возможных преднамеренных воздействий и преобразований, осуществляемых нарушителем над записанными, например, с помощью мобильных устройств, маркированными аудиосигналами.

Методы маркирования аудиоинформации и аудиосигналов в целом можно разделить на методы, работающие с временной областью сигнала и/или с областями преобразований [1, 2]. Методы цифрового маркирования аудиосигналов, ориентированные на передачу через воздушный аудиоканал, описаны в статьях [3, 4]. В них для передачи используется весь слышимый

диапазон, что обычно повышает устойчивость маркера. Однако, из-за того, что маркер внедряется путём полной замены фаз, то такой способ внедрения может оказывать значительное влияние на слышимость маркера. Кроме этого, так как внедрение выполняется во все коэффициенты преобразования, то это не позволяет использовать, например, их номера в качестве возможного ключа, который был бы известен только авторизованному приёмнику, и применялся им для выделения частей маркера только из тех коэффициентов, индексы которых составляют ключ.

Длительное время в системах связи используется ортогональное частотно-разделённое мультиплексирование или OFDM (от англ. orthogonal frequency division multiplexing). В работе [5] предлагается методика и система цифрового маркирования аудиосигналов, названная авторами системой «Dolphin», которая использует OFDM, и позволяет выполнять передачу информации маркированными аудиосигналами через воздушный аудиоканал. Система Dolphin выполняет внедрение маркера в диапазон частот от 8 кГц до 20 кГц. В системе Dolphin

¹ Исследование выполнено при финансовой поддержке Минобрнауки России (грант ИБ; Соглашение № 7/2020)

² The reported study was funded by Russian Ministry of Science (information security № 7/2020)

используется адаптивный метод маркирования, основанный на анализе распределения энергии в маркируемом аудиосигнале перед выбором метода модуляции аудиосигнала для каждого информационного пакета. Так как диапазон частот меньших 8 кГц не используется для внедрения маркера, то систему Dolphin нельзя отнести к тем, которые используют весь слышимый диапазон частот. Вследствие того, что внедрение выполняется по результатам энергетического анализа аудиосигнала, при этом сам исходный аудиосигнал обычно не известен получателю, то это затрудняет использование ключей, опираясь на которые авторизованный приёмник точно бы знал, какие части аудиосигнала будут содержать скрытую информацию.

В работе [6] предлагается методика цифрового маркирования аудиосигналов, которая также ориентирована на связь через воздушный аудиоканал. Её особенностью является то, что она предполагает передачу маркированных аудиосигналов между смарт-устройствами, обычно смартфонами. При передаче используется диапазон частот от 19.5 кГц до 22 кГц. Из-за того, что передача ведётся только в области высоких частот, близкой к ультразвуковому и отчасти ультразвуковому диапазону частот, то предъявляются особые требования к микрофонам и динамикам. По утверждениям авторов, эта методика позволяет выполнять передачу на расстоянии до 25 м.

В работе [7] предлагается система цифрового маркирования аудиосигналов, ориентированная на организацию связи на малых расстояниях до 10–15 см. Передача ведётся в частотном диапазоне от 6 кГц до 7 кГц. При этом окружающий шум используется в качестве скрывающего сигнала. Если уровень окружающего шума падает ниже определённого уровня, то передача прекращается в целях безопасности.

Таким образом, среди нерешённых задач остается следующая: допустимость внедрения в произвольные частотные составляющие при сохранении возможности слепого приема маркированных аудиосигналов. В этой статье разрабатывается метод маркирования цифровых аудиосигналов, предложенный в работах [8, 9]. Разрабатываемый метод ориентирован на обеспечение достаточной устойчивости маркирования при передаче маркированного аудиосигнала через воздушный аудиоканал. Особенностью разрабатываемого метода является возможность внедрения маркера в требуемые частотные составляющие маркируемого аудиосигнала, а также отсутствие необходимости приёмнику знать аудиосигнал, который был подвергнут маркированию. Таким образом появляется возможность реализации слепого приёма. Разработке подвергается та часть метода, с помощью которой вы-

полняется обнаружение маркера в сигнале, принятом из акустического канала.

Обнаружение маркера

Предположим, что отправитель передал в воздушный акустический аудиоканал маркированный аудиосигнал, полученный в результате цифро-аналогового преобразования отсчетов цифрового маркированного аудиосигнала. Также будем предполагать, что маркирование выполнялось по методу, представленному в работе [8, 9], авторизованному получателю известны векторы α, β, γ , использованные при создании маркера, множество номеров $A_{НД}$, использованных при внедрении маркера, частота дискретизации F_s , использованная отправителем, а также известна длина $N_{блок}$ блока отсчетов, на которые выполнялись разбиение исходного цифрового аудиосигнала в процессе внедрения маркера.

С частотой дискретизации F_s выполним аналого-цифровое преобразование (ЦАП) сигналов, поступающих с выходов микрофона авторизованного получателя. Сформируем вектор из последовательности отсчетов цифрового сигнала, получаемого после ЦАП

$$\mathbf{r} = (r_1 \quad r_2 \quad \dots \quad r_{N_\alpha N_{блок} N_\gamma} \quad \dots),$$

где r_j — это j -й отсчет цифрового аудиосигнала, получаемого на основании выхода акустического микрофона. Значения отсчетов r_j лежат в диапазоне $[-1, 1]$. Далее описан предлагаемый процесс обнаружения наличия маркера в аудиосигнале, принятом акустическим микрофоном.

Процесс обнаружения маркера организован таким образом, что в нём выполняется считывание и анализ последовательностей из $N_\alpha N_{блок} N_\gamma$ элементов цифрового сигнала \mathbf{r} . Обозначим величиной $i_{шаг}(m)$ количественное значение «шага», на который смещено начало считанной последовательности относительно первого элемента вектора \mathbf{r} . В таком случае последовательность из $N_\alpha N_{блок} N_\gamma$ элементов, смещённая относительно первого элемента строки на $i_{шаг}(m)$, может быть описана вектором

$$\hat{\mathbf{r}}(i_{шаг}(m)) = (r_{i_{шаг}(m)+1} \quad r_{i_{шаг}(m)+2} \quad \dots \quad r_{i_{шаг}(m)+N_\alpha N_{блок} N_\gamma}),$$

при этом $i_{шаг}(m) \in \{0, 1, \dots\}$.

Далее, вектор $\hat{\mathbf{r}}(i_{шаг}(m))$ подвергается ряду преобразований. Вектор разделяется на непересекающиеся блоки по $N_{блок}$ элементов в каждой. Каждый блок подвер-

гается дискретному преобразованию Фурье, в результате получаются блоки частотных составляющих спектра Фурье. Из каждого блока частотных составляющих выделяются те составляющие, номера которых совпадают с номерами из известного авторизованному получателю множества номеров A_{HJL} . У каждой выделенной частотной составляющей вычисляются амплитуды, а после этого натуральный логарифм от амплитуды. Таким образом, получается

$$D(i_{uaz}(m)) = \begin{pmatrix} D_{1,1}(i_{uaz}(m)) & \dots & D_{1,N_\alpha N_\beta}(i_{uaz}(m)) \\ D_{2,1}(i_{uaz}(m)) & \dots & D_{2,N_\alpha N_\beta}(i_{uaz}(m)) \\ \vdots & \ddots & \vdots \\ D_{N_\gamma,1}(i_{uaz}(m)) & \dots & D_{N_\gamma,N_\alpha N_\beta}(i_{uaz}(m)) \end{pmatrix},$$

в которой

$$D_{i,j}(i_{uaz}(m)) = \ln |f(i, j, \hat{r}(i_{uaz}(m)), A_{HJL})|,$$

где $k \in \{1, 2, \dots, N_\alpha N_\gamma\}$, $l \in \{1, 2, \dots, N_\beta\}$, $\ln|a|$ — натуральный логарифм от абсолютного значения числа a . Здесь f обозначает необходимые преобразования, более детально представленные в статьях [8, 9].

Теперь вычислим коэффициент корреляции между принятыми данными и маркером, кодирующим значение 1. Для этого сначала вычислим вектор из скалярных произведений между столбцами матриц $D(i_{uaz}(m))$ и вектором γ , получим вектор из $N_\alpha N_\beta$ элементов. Потом, вычислим скалярные произведения между непересекающимися последовательностями из N_β элементов полученного вектора и вектором β , получим вектор

$$\hat{d}(i_{uaz}(m)) = (\hat{d}_1(i_{uaz}(m)) \quad \hat{d}_2(i_{uaz}(m)) \quad \dots \quad \hat{d}_{N_\alpha}(i_{uaz}(m))).$$

Теперь выполним статистическое нормирование. Для этого определим сначала среднее арифметическое элементов вектора $\hat{d}(i_{uaz}(m))$, обозначенное как $M(\hat{d}(i_{uaz}(m)))$. Затем, вычислим скалярное произведение между вектором $\hat{d}(i_{uaz}(m))$ и α , при этом будут использоваться значения элементов вектора $\hat{d}(i_{uaz}(m))$, смещённые на среднее арифметическое

$$M(\hat{d}(i_{uaz}(m))).$$

Поделим это скалярное произведение на нормирующий коэффициент и получим вещественное число

$$\tilde{d}(i_{uaz}(m)) = \frac{\sum_{i=1}^{N_\alpha} \alpha_i (\hat{d}_i(i_{uaz}(m)) - M(\hat{d}(i_{uaz}(m))))}{\sqrt{\sum_{i=1}^{N_\alpha} (\alpha_i (\hat{d}_i(i_{uaz}(m)) - M(\hat{d}(i_{uaz}(m))))^2}}$$

Величина $\tilde{d}(i_{uaz}(m))$ является нормированным коэффициентом корреляции между последовательностью отсчетов $\hat{r}(i_{uaz}(m))$ и маркером, кодирующим число 1. Сама по себе случайная величина $\tilde{d}(i_{uaz}(m))$ представляет собой сумму большого количества случайных величины, которые можно рассматривать как статистически слабо зависимые. Поэтому в соответствии с центральной предельной теоремой можно ожидать, что, когда в принятом сигнале маркера нет, случайная величина $\tilde{d}(i_{uaz}(m))$ будет иметь стандартное нормальное распределение с нулевым математическим ожиданием и единичной дисперсией. Эти предположения были подтверждены результатами натуральных экспериментов.

Сформируем из величин $\tilde{d}(i_{uaz}(m))$ вектор

$$\rho(i_{uaz}(m)) = (\rho_1(i_{uaz}(m)) \quad \rho_2(i_{uaz}(m)) \quad \dots \quad \rho_{2W+1}(i_{uaz}(m))),$$

в котором

$$\rho_i(i_{uaz}(m)) = \tilde{d}(w(i_{uaz}(m), i)),$$

$$\text{где } w(i_{uaz}(m), i) = i_{uaz}(m) + i - 1,$$

при этом W — это заранее определённое положительное целое число. В качестве значения W полезно использовать достаточно большое значение, кратное $N_{\text{блок}} N_\gamma$. Таким образом, вектор $\rho(i_{uaz}(m))$ представляет собой результат обработки последовательности векторов

$$\hat{r}(i_{uaz}(m)), \dots, \hat{r}(i_{uaz}(m) + 2W).$$

Процесс обнаружения состоит из двух этапов. На первом этапе определяется множество смещений

$$\{i_{uaz}(m), i_{uaz}(m) + 1, \dots, i_{uaz}(m) + 6W\}$$

относительно момента начала записи, среди которых имеется начало сигнала, содержащего маркер. На втором этапе в этом множестве выполняется поиск начала сигнала, содержащего маркер.

По каждому вектору из последовательности векторов $\rho(i_{uaz}(m)), \dots, \rho(i_{uaz}(m) + 2W)$ вычислим коэффициент эксцесса и получим вектор

$$\mu(i_{шаг}(m)) = (\mu_1(i_{шаг}(m)) \quad \mu_2(i_{шаг}(m)) \quad \dots \\ \dots \quad \mu_{2W+1}(i_{шаг}(m))).$$

У стандартного нормального распределения коэффициент эксцесса $\mu_j(i_{шаг}(m)) = 3$. Таким образом, чем сильнее элементы вектора $\rho(w(i_{шаг}(m), j))$ похожи на элементы выборки значений случайной величины, имеющей стандартное нормальное распределение, тем ближе значение $\rho(w(i_{шаг}(m), j))$ к числу 3.

Теперь вычислим пик-факторы по каждому вектору из последовательности векторов

$$\mu(i_{шаг}(m)), \dots, \mu(i_{шаг}(m) + 2W)$$

и сформируем из этих пик-факторов вектор

$$\zeta(i_{шаг}(m)) = (\zeta_1(i_{шаг}(m)) \quad \zeta_2(i_{шаг}(m)) \quad \dots \\ \dots \quad \zeta_{2W+1}(i_{шаг}(m))),$$

в котором

$$\zeta_k(i_{шаг}(m)) = \frac{\max\{\mu(i_{шаг}(m))\}}{\sqrt{\frac{\sum_{i=1}^{2W+1} (\mu_i(w(i_{шаг}(m), k)))^2}{2W+1}}},$$

где

$$k \in \{1, 2, \dots, 2W+1\}, w(i_{шаг}(m), k) = i_{шаг}(m) + k - 1.$$

Для стандартного нормального распределения пик-фактор равен 1, так как для такого распределения и числители, и знаменатели этих дробей будут равны 3.

Первый этап завершается успешно, если выполняются два условия:

должно выполняться равенство

$$\zeta_{W+1}(i_{шаг}(m)) = \max\{\zeta_1(i_{шаг}(m)), \dots, \zeta_{2W+1}(i_{шаг}(m))\},$$

то-есть центральный элемент вектора $\zeta(i_{шаг}(m))$ должен быть максимальным среди всех элементов этого вектора. Отметим, что все элементы вектора $\zeta(i_{шаг}(m))$ — это положительные числа.

должно выполняться неравенство

$$\zeta_{W+1}(i_{шаг}(m)) > \zeta_{порог},$$

то-есть величина этого центрального элемента должна превосходить некоторый порог, который сам превосходит 1.

Если эти условия не выполняются для текущего значения $i_{шаг}(m)$, то это значение увеличивается на 1 и первый этап начинается заново.

Второй этап начинается, если успешно завершён первый этап. Начиная второй этап, найдём в векторе $\rho(i_{шаг}(m) + 4W)$ порядковые номера элементов, для которых один или оба смежных элемента будут иметь противоположные знаки. Знание этих порядковых номеров укажет на границы положительных и отрицательных последовательностей значений элементов вектора $\rho(i_{шаг}(m) + 4W)$. Как только будут определены «границы» положительных и отрицательных последовательностей, найдём отношения «площадей» под/над кривыми (по методу трапеций), образуемыми элементами этих последовательностей, к числу слагаемых «площадей». Выберем ту последовательность, «площадь» которой окажется максимальной. Позиция максимального по абсолютному значению элемента в этой последовательности будет использоваться в качестве указателя на начало сигнала, содержащего маркер. Обозначим номер этой позиции — так $i_{начало}(m)$. Это значит, что маркированный сигнал в цифровом сигнале, принятом с помощью акустического микрофона авторизованного получателя, будет начинаться с отсчета, имеющего порядковый номер $i_{начало}(m)$, и занимать $N_\alpha N_{блок} N_\gamma$ отсчетов цифрового сигнала.

Результаты натуральных экспериментов

В качестве акустического микрофона использовался 4-х капсульный микрофон AKG Lyra в режиме Tight Stereo (фронтальная стереозапись), а в качестве акустического динамика использовался один динамик из пары Edifier R1280DB.

В качестве вектора α использовалась последовательность Касами длиной $N_\alpha = 255$, в качестве вектора β использовалась последовательность Голда длиной $N_\beta = 127$. Вектор γ имел длину $N_\gamma = 8$ и удовлетворял следующему равенству:

$$\gamma = (1 \quad -1 \quad 1 \quad 1) \otimes (1 \quad -1),$$

где \otimes — это оператор кронекерова произведения. Размер блока отсчетов, на

$$N_{блок} = 2(N_\beta + 1) = 256.$$

При таких значениях параметров маркера и частоте дискретизации $F_s = 44100$ Гц передача маркера займёт

$$\frac{N_{\alpha} N_{\text{блок}} N_{\gamma}}{F_s} \cong 11.8(c),$$

Значение для величины W было взято равным 4096.

В качестве музыкальной композиции для внедрения маркера использовалось песня «This Moment» музыкальной рок-группы «In This Moment»; эта композиция характеризуется помимо низкочастотных составляющих наличием еще и мощных высокочастотных составляющих.

Множество номеров частотных составляющих, использованных для внедрения элементов маркера в блок, было одним и тем же для всех блоков отсчетов маркируемого цифрового аудиосигнала, при этом использовались все частотные составляющие блоков, кроме постоянной составляющей. Силы встраивания элементов маркера были одинаковыми для всех частотных составляющих и равнялись 0.1. Такое значение силы встраивания обеспечивало отсутствие слышимых искажений аудиосигнала после внедрения маркера в сравнении с аудиосигналом без маркера. В качестве порогового значения пик-фактора использовалось значение 1.15. Усиление сигналов производилось с помощью усилителей динамика и микрофона. Каждый из них оказывал влияние, помимо усиления полезного сигнала, и на усиление шума. Для оценки передачи в целом использовалось отношение среднего значения амплитуд отсчетов принятого цифрового маркированного аудиосигнала к среднему значению амплитуд отсчетов переданного цифрового маркированного аудиосигнала. Для оценки шума использовалось среднее значение амплитуд постороннего акустического сигнала, записанного вне времени вещания динамиком маркированного аудиосигнала, но в тех же условиях передачи, в которых осуществлялось вещание маркированного аудиосигнала.

Микрофонный усилитель в AKG Lyra был настроен на максимальное значение усиления поступающего сигнала. Усиление в динамике менялось от 4 до 9 делений из 50 возможных делений по шкале усилителя мощности, встроенного в динамики Edifier R1280DB. Динамик и микрофон находились в области прямой видимости на расстоянии 3 метров друг от друга. Натурные эксперименты показали, что при таком расстоянии между динамиком и микрофоном, когда отношение среднего значения амплитуд отсчетов принятого цифрового маркированного аудиосигнала к среднему значению амплитуд отсчетов переданного цифрового маркированного аудиосигнала становилось больше -12 дБ, тогда вероятность ошибки обнаружения маркера в принятом аудиосигнале была меньше 10^{-6} , а вероятность восстановления информации при отсутствии ошибки обнаружения маркера была меньше 10^{-3} . Таким образом, можно ожидать, что разработанный метод обеспечивает вероятность удачной передачи информации маркированными аудиосигналами не меньше $1 - 10^{-9}$. Величина вероятности ошибки обнаружения оценивалась, исходя из общего числа отсчетов в принятом маркированном цифровом аудиосигнале, а вероятность восстановления информации после удачного обнаружения маркера оценивалась на основании величины W .

Заключение

В статье была продолжена разработка метода маркирования аудиосигналов, начатая в работах [8, 9]. Предложен метод обнаружения маркера в аудиосигнале и метод выделения информации, внедренной в маркированный аудиосигнал. Практическая работоспособность этого метода была подтверждена результатами натурных экспериментов передачи маркированных аудиосигналов через воздушный акустический канал в условиях значительного постороннего акустического шума.

ЛИТЕРАТУРА

1. Шелухин О. И., Канаев С. Д. Стеганография. Алгоритмы и программная реализация: Под редакцией проф. Шелухина О. И. — Москва: Горячая линия-Телеком. — 2017. — 592 с.
2. Грибунин В. Г., Оков И. Н., Туринцев И. В. Цифровая стеганография. — Москва: СОЛОН-ПРЕСС. — 2009. — 272 с.
3. Yun H. S., Cho K., Kim N. S. Acoustic data transmission based on modulated complex lapped transform // IEEE Signal Processing Letters. — 2010. — Т. 17. — №. 1. — С. 67–70.
4. Cho K., Baek S., Moon H. G., Kim N. S. Multi-microphone approach for reliable acoustic data transmission // 2016 IEEE International Conference on Consumer Electronics (ICCE). — IEEE. — 2016. — С. 557–560.
5. Zhou M., Wang Q., Ren K., Koutsonikolas D., Su L., Chen Y. Dolphin: Real-time hidden acoustic signal capture with smartphones // IEEE Transactions on Mobile Computing. — 2018. — Т. 18. — Вып 3. — С. 560–573.
6. Lee H., Kim T. H., Choi J. W., Choi S. Chirp signal-based aerial acoustic communication for smart devices // 2015 IEEE Conference on Computer Communications (INFOCOM). — IEEE. — 2015. — С. 2407–2415.
7. Nandakumar R., Chintalapudi K. K., Padmanabhan V., Venkatesan R. Dhvani: secure peer-to-peer acoustic NFC // ACM SIGCOMM Computer Communication Review. — 2013. — Т. 43. — №. 4. — С. 63–74.

8. Гофман М. В. Методика скрытой передачи данных при связи через воздушный аудиоканал // Труды СПИИРАН. — 2017. — Вып. 2. — С. 97–122.
9. Гофман М. В., Корниенко А. А., Мирончиков Е. Т., Никитин А. Б. Цифровое маркирование аудиосигналов для робастной скрытой акустической связи через воздушный аудиоканал // Труды СПИИРАН. — 2017. — Вып. 6. — С. 185–215.

© Гофман Максим Викторович (maxfog@gmail.com).

Журнал «Современная наука: актуальные проблемы теории и практики»

