УДК 519.876.5:519.6:004.357 **DOI** 10.15622/sp.55.8

М.В. Гофман, А.А. Корниенко, Е.Т. Мирончиков, А.Б. Никитин

ЦИФРОВОЕ МАРКИРОВАНИЕ АУДИОСИГНАЛОВ ДЛЯ РОБАСТНОЙ СКРЫТОЙ АКУСТИЧЕСКОЙ СВЯЗИ ЧЕРЕЗ ВОЗДУШНЫЙ АУДИОКАНАЛ

Гофман М.В., Корниенко А.А., Мирончиков Е.Т., Никитин А.Б. **Цифровое маркирование аудиосигналов для робастной скрытой акустической связи через воздушный аудиоканал.**

Аннотация. В этой работе развивается методика цифрового маркирования аудиосигналов, ориентированная на передачу данных через воздушный аудиоканал. Внедряемый цифровой маркер занимает весь слышимый частотный диапазон. Цифровой маркер кодирует один бит информации. Решение о значении переданного бита выносится на основании знака центрального значения взаимно-корреляционной функции. Предлагаются две методики построения цифрового маркера. Обе методики специально ориентированы на частотные свойства обычных аудиосигналов. Невысокая вычислительная сложность предлагаемого метода маркирования позволяет использовать его для беспроводного обмена информацией между обычными смартфонами. Методика позволяет выполнять цифровое маркирования как речевых, так и музыкальных аудиосигналов без появления каких-либо заметно слышимых артефактов. Информация внедряется в виде маркера в частотную область аудиосигнала путем амплитудной модуляции его частотных составляющих. Эта работа снабжена результатами имитационного моделирования и натурных экспериментов, подтверждающих применимость методики для скрытой передачи данных через воздушный аудиоканал.

Ключевые слова: цифровой маркер, аудиосигнал, скрытая передача данных, воздушный аудиоканал, акустическая связь, стегоаудиосигнал.

1. Введение. Процедура цифрового маркирования аудиосигналов лежит в основе системы скрытой передачи данных с помощью таких сигналов. Она включает в себя три этапа: этап построения маркера, этап внедрения маркера в аудиосигнал и этап выделения маркера из маркированного аудиосигнала. Цифровое маркирование аудиосигнала предполагает создание из информации такого маркера и его внедрение в аудиосигнал таким образом, что становится возможным его выделить даже при условии, что маркированный аудиосигнал или стегоаудиосигнал подвергнется преднамеренной или ненамеренной атаке, то есть процедура цифрового маркирования должна быть робастной ко всякого рода воздействиям.

Существующие методы маркирования в общем случае можно разделить на работающие во временной области и в областях преобразований [1]. Далее их можно подразделить еще на несколько подкатегорий и, таким образом, методы из временной области включают методы с выравниваем по времени [2, 3] и методы [4-6], использующие эхо, тогда как методы из области преобразований также можно подразделить на методы расширения спектра [7, 8], методы модуляции с индексным

квантованием (от англ. quantization index modulation) [9, 10], методы «патчворка» или «лоскута» (от англ. patchwork) [11, 12].

Частым приложением процедуры цифрового маркирования является защита авторских прав при распространении аудио, видео и изображений в цифровых форматах. В этих случаях внедряемый сигнал либо сообщает получателю какие-то авторские данные или лицензионные ограничения, либо предотвращает или запрещает неавторизованное копирование. Другой пример приложения — внедряемый сигнал может либо разрешать, либо запрещать копирование некоторому копирующему устройству, которое проверяет внедряемый сигнал перед выполнением процедуры дупликации. Или же, например, согласованный с неким стандартом проигрыватель дисков может проверить наличие маркера, перед тем как решать проигрывать диск или нет. Еще одним приложением методов цифрового маркирования является так называемое гибридное полосовое канальное цифровое аудиовещание [13].

В зависимости от рабочей частотной полосы воздушная акустическая связь может использовать как обычный смартфон, который способен проигрывать/записывать аудио (до 22 кГц), так и ультразвуковой приемопередатчик (свыше 22 кГц). Поэтому любое устройство, оснащенное аудиоинтерфейсом может быть использовано как устройство аудиосвязи. Таким образом, появляется альтернативный интерфейс беспроводной связи между смартфонами, помимо имеющихся Wi-Fi и Bluetooth, которые часто выключены, для экономии батареи и/или предотвращения нежелательных подключений. Воздушная акустическая связь в частотном диапазоне до 22 кГц исследуется в работах [14-17].

Далее в этой статье развивается подход, описанный в статье [18], использующий троекратное кодирование передаваемых информационных битов перед передачей через воздушный аудиоканал. Предложенный в работе [18] подход показал свою устойчивость к искажающим влияниям воздушного аудиоканала. Однако он требовал использования двух различных последовательностей для кодирования значений бита информации. Особенностью предлагаемого далее подхода является кодирование битов в знаке корреляционной функции, что уменьшает требуемое количество последовательностей до одной и тем самым увеличивает потенциально возможную скорость передачи информации. Процедура передачи состоит из трех этапов. Первый этап заключается в получении кодового слова, соответствующего передаваемому биту — построение маркера. Второй — во внедрении этого кодового слова в аудиосигнал, в результате чего получается стегоаудиосигнал — марки-

рованный аудиосигнал. А третий этап состоит в восстановлении переданного бита из стегоаудиосигнала, полученного после передачи через воздушный аудиоканал.

2. Этап 1: построение маркера для аудиосигнала. Пусть требуется с помощью аудиосигнала осуществить скрытую передачу через воздушный аудиоканал двоичного символа $x \in \{0,1\}$. Для выполнения процедуры кодирования информации — первого этапа цифрового маркирования — требуется три специальные последовательности, составленные из элементов множества $\{+1,-1\}$. Первые две последовательности:

$$\boldsymbol{\alpha} = \left(\alpha(1) \quad \alpha(2) \quad \ldots \quad \alpha(N_\alpha)\right), \tag{1}$$

$$\boldsymbol{\beta} = \left(\beta(1) \quad \beta(2) \quad \ldots \quad \beta(N_\beta)\right) \tag{2}$$

должны обладать хорошими автокорреляционными свойствами.

С помощью третьей последовательности используется свойство обычных аудиосигналов, иметь близкие спектральные характеристики смежных последовательностей отсчетов. Поэтому третья последовательность:

$$\boldsymbol{\gamma} = \left(\gamma(1) \quad \gamma(2) \quad \ldots \quad \gamma(N_\gamma)\right) \tag{3}$$

должна обладать следующим свойством: ее подпоследовательности из +1-ц и –1-ц должны быть короткими, обычно чем короче, тем лучше.

Процесс получения кодового слова по двоичному символу x следующий. Если $x = 1$, то выбирается последовательность $\boldsymbol{\alpha}$, а если же $x = 0$, то последовательность $-\boldsymbol{\alpha}$. Затем каждый элемент выбранной последовательности заменяется либо последовательностью $\boldsymbol{\beta}$, если он равен 1, либо последовательностью $-\boldsymbol{\beta}$ в противном случае. И наконец, каждый элементы полученной последовательности также подвергаются замене: элемент, равный 1, заменяется на последовательность $\boldsymbol{\gamma}$, тогда как элемент, равный -1, заменяется на последовательность $-\boldsymbol{\gamma}$. Итоговую последовательность можно представить в виде вектора-строки:

$$\mathbf{y}(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},x) = \left(y(1,\alpha(1),\beta(1),\gamma(1),x) \quad y(2,\alpha(1),\beta(1),\gamma(2),x \right.$$
$$\left. \ldots \quad y(N_\alpha N_\beta N_\gamma, \alpha(N_\alpha), \beta(N_\beta), \gamma(N_\gamma), x)\right), \tag{4}$$

где

$$y\left(\left(i-1\right)N_{\beta}N_{\gamma}+\left(j-1\right)N_{\gamma}+k,\alpha\left(i\right),\beta\left(j\right),\gamma\left(k\right),x\right)= \\ =\left(2x-1\right)\alpha\left(i\right)\beta\left(j\right)\gamma\left(k\right), \tag{5}$$

где $i\in\left\{1,2,\ldots,N_{\alpha}\right\}$, $j\in\left\{1,2,\ldots,N_{\beta}\right\}$, $k\in\left\{1,2,\ldots,N_{\gamma}\right\}$. Если воспользоваться операцией Кронекерова произведения, то:

$$\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma},x\right)=\left(2x-1\right)\boldsymbol{\alpha}\otimes\boldsymbol{\beta}\otimes\boldsymbol{\gamma}, \tag{6}$$

где $\otimes$ — это оператор Кронекерова произведения.

Видно, что такой способ кодирования обладает небольшой скоростью, однако этот подход легко адаптировать для передачи не одного бита, а последовательностей бит. Например, когда требуется передавать последовательности из 4-х бит: $\mathbf{x}_0=\begin{pmatrix}0&0&0&0\end{pmatrix}$, $\mathbf{x}_1=\begin{pmatrix}0&0&0&1\end{pmatrix}$ и так далее до $\mathbf{x}_{15}=\begin{pmatrix}1&1&1&1\end{pmatrix}$. Тогда вместо одной последовательности $\boldsymbol{\alpha}$ используется несколько последовательностей $\boldsymbol{\alpha}_i$, где $i\in\left\{0,1,\ldots,7\right\}$, которые помимо хороших автокорреляционных свойств имеют слабую взаимную корреляцию. В таком случае можно последовательность $\mathbf{x}_i$ кодировать последовательностью $\mathbf{y}\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta},\boldsymbol{\gamma},1\right)$, а последовательность $\mathbf{x}_{15-i}$ кодировать последовательностью $\mathbf{y}\left(\boldsymbol{\alpha}_i,\boldsymbol{\beta},\boldsymbol{\gamma},0\right)$. Если же использовать еще и различные последовательности $\boldsymbol{\beta}$, обладающие хорошими автокорреляционными свойствами, но имеющими слабую взаимную корреляцию, то можно добиться еще большей скорости передачи, так как в таком случае такими же будут корреляционные свойства у итоговых последовательностей $\mathbf{y}$.

3. Этап 2: внедрение маркера в аудиосигнал. Внедрение маркера в аудиосигнал составляет второй этап предлагаемого процесса скрытой передачи информации. Пусть задан цифровой аудиосигнал:

$$\mathbf{z}=\begin{pmatrix}z\left(1\right)&z\left(2\right)&\ldots&z\left(N_{\mathrm{Z}}\right)\end{pmatrix}, \tag{7}$$

где $z\left(i\right)$ — это *i*-й отсчет цифрового аудиосигнала, принимающий значения из диапазона $\left[-1,1\right]$; а N_{Z} — количество отсчетов, при этом:

$$N_{\text{Z}} = K_{\text{блок}} N_{\text{блок}}, \tag{8}$$

где

$$K_{\text{блок}} = N_{\alpha} N_{\gamma}, \tag{9}$$

а $N_{\text{блок}}$ — четное число, удовлетворяющее следующему неравенству:

$$N_{\text{блок}} \geq 2\left(N_{\beta} + 1\right). \tag{10}$$

Элементы вектора (4) встраиваются в цифровой аудиосигнал (7), в результате этого получается такого же размера стегоаудиосигнал:

$$\mathbf{z}' = \left(z'\left(1\right) \quad z'\left(2\right) \quad \ldots \quad z'\left(N_{\text{Z}}\right)\right), \tag{11}$$

где $z'\left(i\right) \in \left[-1, +1\right]$ — это *i*-й отсчет цифрового стегоаудиосигнала. Процесс встраивания начинается с разбиения вектора (7) на блоки, которые затем подвергаются прямому дискретному преобразованию Фурье, что позволяет перейти в частотную область и получить спектральные линии соответствующих блоков отсчетов. Путем амплитудной модуляции получаемых спектральных линий элементы вектора (4) встраиваются в скрывающий сигнала (7). Завершением процесса построения стегоаудиосигнала (11) является применение обратного дискретного преобразования Фурье к модифицированным спектральным линиям, а также нормировка, в результате и получается временной сигнал (11). Далее приводится детальное описание этого процесса.

Блоки, на которые разбивается вектор (7), представляют собой векторы из отсчетов:

$$\mathbf{z}_{\text{блок}}\left(j\right) = \left(z\left(N_{\text{блок}}\left(j-1\right)+1\right) \quad z\left(N_{\text{блок}}\left(j-1\right)+2\right) \quad \ldots \quad z\left(N_{\text{блок}} j\right)\right), \tag{12}$$

где $j \in \left\{1, 2, \ldots, K_{\text{блок}}\right\}$. Блок подвергается прямому дискретному преобразованию Фурье, что дает вектор спектральных линий:

$$\mathbf{Z}_{\text{блок}}\left(j\right) = \left(Z_{\text{блок}}\left(j,1\right) \quad Z_{\text{блок}}\left(j,2\right) \quad \ldots \quad Z_{\text{блок}}\left(j, N_{\text{блок}}\right)\right), \tag{13}$$

где

$$Z_{\text{блок}}\left(j,k\right)=\sum_{k=1}^{N_{\text{блок}}} z\left(N_{\text{блок}}\left(j-1\right)+k\right)\exp\left(-\frac{\iota 2\pi\left(j-1\right)\left(k-1\right)}{N_{\text{блок}}}\right), \qquad (14)$$

где $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, $k\in\left\{1,2,\ldots,N_{\text{блок}}\right\}$, $\iota=\sqrt{-1}$ — мнимая единица.

Прежде чем описать процесс применяемой амплитудной модуляции введем дополнительные объекты, которые нам потребуются. Пусть для *j*-го ($j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$) блока спектральных линий множества:

$$\mathbf{A}_{\text{НЛ}}\left(j\right)=\left\{A_{\text{НЛ}}\left(j,1\right),A_{\text{НЛ}}\left(j,2\right),\ldots,A_{\text{НЛ}}\left(j,N_{\beta}\right)\right\} \qquad (15)$$

и

$$\mathbf{A}_{\text{СВ}}\left(j\right)=\left\{A_{\text{СВ}}\left(j,1\right),A_{\text{СВ}}\left(j,2\right),\ldots,A_{\text{СВ}}\left(j,N_{\beta}\right)\right\} \qquad (16)$$

обозначают номера модифицируемых спектральных линий и величины сил встраивания соответственно. Номера линий $A_{\text{НЛ}}\left(i,j\right)$ удовлетворяют следующим неравенствам:

$$1<A_{\text{НЛ}}\left(j,1\right)<A_{\text{НЛ}}\left(j,2\right)<\ldots<A_{\text{НЛ}}\left(j,N_{\beta}\right)\le\frac{N_{\text{блок}}}{2}. \qquad (17)$$

Величины сил встраивания $A_{\text{СВ}}\left(j,i\right)$ представляют собой положительные вещественные числа обычно значительно меньшие единицы $A_{\text{СВ}}\left(j,i\right)\ll 1$.

Встраивание происходит в частотную область, а именно в амплитуду спектральных линий. Количество спектральных линий $N_{\text{блок}}$ в каждом блоке и число блоков $K_{\text{блок}}$ в получаемом стегоаудиосигнале (11) не отличаются от таких же величин в исходном аудиосигнале. Обозначим *k*-ю спектральную линию *j*-го блока отсчетов стегоаудиосигнала (11) символом $Z'_{\text{блок}}\left(j,k\right)$, тогда для первой половины блока спектральных линий будет выполняться следующее равенство:

$$Z'_{\text{блок}}\left(j,k\right)= \\ =\begin{cases} Z_{\text{блок}}\left(j,k\right)\left(1+A_{\text{СВ}}\left(j,i\right)y\left(m\left(j,i\right)\right)\right), & \text{если } k=A_{\text{НЛ}}\left(j,i\right), \\ Z_{\text{блок}}\left(j,k\right), & \text{иначе,} \end{cases} \quad (18)$$

где $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, $k\in\left\{1,2,\ldots,N_{\text{блок}}/2\right\}$, $i\in\left\{1,2,\ldots,N_{\beta}\right\}$,

$$m\left(j,i\right)=\left(\left(j-1\right)\bmod N_{\gamma}\right)+1+\left(i-1\right)N_{\gamma}+\left\lfloor\frac{j-1}{N_{\gamma}}\right\rfloor N_{\gamma}N_{\beta}, \quad (19)$$

где $\left\lfloor a\right\rfloor$ — целая часть вещественного числа a. Спектральные линии j-го блока $Z_{\text{блок}}\left(j,k\right)$ с номерами k из диапазона от $\left(N_{\text{блок}}/2\right)+1$ до $N_{\text{блок}}$ для сохранения свойства сопряженной симметричности также подвергаются изменениям в соответствии со следующим равенством:

$$Z'_{\text{блок}}\left(j,k\right)=\begin{cases} \left(Z'_{\text{блок}}\left(j,N_{\text{блок}}-k+2\right)\right)^{*}, & \text{если } j\neq\left(N_{\text{блок}}/2\right)+1, \\ Z_{\text{блок}}\left(j,k\right), & \text{иначе,} \end{cases} \quad (20)$$

где $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, $k\in\left\{\left(N_{\text{блок}}/2\right)+1,\left(N_{\text{блок}}/2\right)+2,\ldots,N_{\text{блок}}\right\}$, $\left(a\right)^{*}$ — число, комплексно сопряженное числу a.

Таким образом, вместо исходных блоков спектральных линий получаются блоки модифицированных спектральные линии. Однако, чтобы перейти во временную область требуется выполнить обратное дискретное преобразование Фурье над каждым таким модифицированным блоком. Выполняя такие преобразования и объединяя их результаты, получим вещественный цифровой стегоаудиосигнал (11), элементы которого удовлетворяются равенству:

$$z'\left(\left(j-1\right)N_{\text{блок}}+i\right)=\frac{1}{N_{\text{блок}}}\sum_{k=1}^{N_{\text{блок}}}Z'_{\text{блок}}\left(j,k\right)\exp\left\{\frac{\iota 2\pi\left(i-1\right)\left(k-1\right)}{N_{\text{блок}}}\right\}, \quad (21)$$

где $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, $i\in\left\{1,2,\ldots,N_{\textit{блок}}\right\}$, $\iota=\sqrt{-1}$ — мнимая единица.

Значения величин (21) могут оказаться вне диапазона $\left[-1,1\right]$. Поэтому завершающим этапом построения стегоаудиосигнала (11) явля-

ется нормировка значений его отсчетов. Мультипликативный нормирующий коэффициент, например, может быть получен на основании всех величин (21), таким образом:

$$\theta = \frac{1}{\max\left\{\left|z'\left(1\right)\right|,\left|z'\left(2\right)\right|,\ldots,\left|z'\left(N_{\text{Z}}\right)\right|\right\}}, \tag{22}$$

где $\left|a\right|$ — абсолютное значение числа a. Или же нормировка может выполняться поблоково, в таком случае задержка перед отправкой стегоаудиосигнала в канал уменьшается.

4. Этап 3: выделение скрытой информации из стегоаудиосигнала после его передачи через воздушный аудиоканал. Третий этап предлагаемого метода скрытой передачи информации состоит в выделении из стегоаудиосигнала, принятого микрофоном, внедренной информации. Стегоаудиосигнал (11) передается через воздушный аудиоканал. Пусть на выходе аудиоканала выполняется дискретизация с частотой F_{s}, равной частоте отправки отсчетов стегоаудиосигнала (11) в аудиоканал. Таким образом, из-за шумов на выходе канала, а также из-за отсутствия синхронизации между передатчиком и приемником получится последовательность отсчетов:

$$r\left(1\right), r\left(2\right), \ldots, r\left(N_{\text{Z}}\right), r\left(N_{\text{Z}}+1\right), \ldots\ , \tag{23}$$

в общем случае отличающихся от переданных.

Чтобы из этой последовательности выделить переданный сигнал (4), требуется определить ту подпоследовательность отсчетов, в которой он скрыт. Поиск начинается с того, что последовательность (23) разбивается на перекрывающиеся подпоследовательности одинаковой длины N_{Z}. Обозначим отдельную подпоследовательность вектором:

$$\mathbf{r}\left(i_{\text{шаг}}\right) = \left(r\left(i_{\text{шаг}}\right) \quad r\left(i_{\text{шаг}}+1\right) \quad \ldots \quad r\left(i_{\text{шаг}}+N_{\text{Z}}-1\right)\right), \tag{24}$$

где $i_{\text{шаг}} \in \left\{1,2,\ldots\right\}$.

Далее выполняются действия, некоторые из которых обратны, выполненным процедурой встраивания. Так, осуществляются дискретные преобразования Фурье над смежными блоками элементов вектора (24), при этом длина блока также, как и процедуре встраивания, равна $N_{\text{блок}}$; например, первый блок можно описать вектором:

$$\left(r\left(i_{\text{шаг}}\right) \quad r\left(i_{\text{шаг}}+1\right) \quad \ldots \quad r\left(i_{\text{шаг}}+N_{\text{блок}}-1\right)\right), \tag{25}$$

а второй блок — вектором:

$$\left(r\left(i_{\text{шаг}}+N_{\text{блок}}\right) \quad r\left(i_{\text{шаг}}+N_{\text{блок}}+1\right) \quad \ldots \quad r\left(i_{\text{шаг}}+2N_{\text{блок}}-1\right)\right), \tag{26}$$

и так далее. В результате этих преобразований каждый блок преобразуется в такой же размерности комплексный вектор — вектор спектральных линий. Обозначим *i*-ю спектральную линию *j*-го блока через

$$R\left(j,i,i_{\text{шаг}}\right)=\sum_{k=0}^{N_{\text{блок}}-1} r\left(i_{\text{шаг}}+\left(j-1\right)N_{\text{блок}}+k\right)\exp\left(-\frac{\iota 2\pi k\left(i-1\right)}{N_{\text{блок}}}\right), \tag{27}$$

где $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, $i\in\left\{1,2,\ldots,N_{\text{блок}}\right\}$, $\iota=\sqrt{-1}$ — мнимая единица.

Вычислим абсолютные значения тех спектральных величин (27), в которых $i\in\mathbf{A}_{\text{НЛ}}\left(j\right)$, а потом вычислим натуральные логарифмы от этих абсолютных значений. Затем сформируем из результатов этих вычислений, выполненных для всех $j\in\left\{1,2,\ldots,K_{\text{блок}}\right\}$, вектор

$$\mathbf{R}_{\text{НЛ}}\left(i_{\text{шаг}}\right)=\left(R_{\text{НЛ}}\left(1,i_{\text{шаг}}\right) \quad R_{\textit{НЛ}}\left(2,i_{\text{шаг}}\right) \quad \ldots \quad R_{\textit{НЛ}}\left(K_{\text{блок}}N_{\beta},i_{\text{шаг}}\right)\right) \tag{28}$$

где

$$R_{\text{НЛ}}\left(k,i_{\text{шаг}}\right)=\ln\left|R\left(k_1\left(k\right),A_{\text{НЛ}}\left(k_1\left(k\right),k_2\left(k\right)\right),i_{\text{шаг}}\right)\right|, \tag{29}$$

где $k\in\left\{1,2,\ldots,K_{\text{блок}}N_{\beta}\right\}$,

$$k_1\left(k\right)=\left\lceil\frac{k-1}{N_{\beta}}\right\rceil+1, \tag{30}$$

$$k_2\left(k\right)=\left(\left(k-1\right)\bmod N_{\beta}\right)+1. \tag{31}$$

Если же скрывающий сигнал (7) известен, то обычно после его вычитания:

$$R_{\text{НЛ}}\left(k,i_{\text{шаг}}\right)=\ln\left|R\left(k_1\left(k\right),A_{\text{НЛ}}\left(k_1\left(k\right),k_2\left(k\right)\right),i_{\text{шаг}}\right)\right|-\\-\ln\left|Z\left(k_1\left(k\right),A_{\text{НЛ}}\left(k_1\left(k\right),k_2\left(k\right)\right)\right)\right|,\tag{32}$$

вероятность успешного детектирования значительно увеличивается.

Учитывая, что длина вектора (28) кратна $K_{\text{блок}}=N_\alpha N_\gamma$, этот вектор можно переформатировать в матрицу $\mathbf{D}\left(i_{\text{шаг}}\right)$ размера $N_\gamma\times N_\alpha N_\beta$:

$$\mathbf{D}\left(i_{\text{шаг}}\right)=\begin{pmatrix}D\left(1,1,i_{\text{шаг}}\right) & D\left(1,2,i_{\text{шаг}}\right) & \cdots & D\left(1,N_\alpha N_\beta,i_{\text{шаг}}\right)\\ D\left(2,1,i_{\text{шаг}}\right) & D\left(2,2,i_{\text{шаг}}\right) & \cdots & D\left(2,N_\alpha N_\beta,i_{\text{шаг}}\right)\\ \vdots & \vdots & \ddots & \vdots\\ D\left(N_\gamma,1,i_{\text{шаг}}\right) & D\left(N_\gamma,2,i_{\text{шаг}}\right) & \cdots & D\left(N_\gamma,N_\alpha N_\beta,i_{\text{шаг}}\right)\end{pmatrix},\tag{33}$$

где

$$D\left(i,j,i_{\text{шаг}}\right)=R_{\text{НЛ}}\left(k,i_{\text{шаг}}\right),\tag{34}$$

где $k\in\left\{1,2,\ldots,K_{\text{блок}}N_\beta\right\}$, а зависимость между целыми числами i,j,k является следующей:

$$i=\left(\left[\frac{k-1}{N_\beta}\right]\bmod N_\gamma\right)+1,\tag{35}$$

$$j=\left[\frac{k-1}{N_\beta N_\gamma}\right]N_\beta+\left(\left(k-1\right)\bmod N_\beta\right)+1.\tag{36}$$

Цифровое корреляционное детектирование скрытого сигнала. Вычислим линейную комбинацию строк матрицы (33) путем умножения этой матрицы слева на вектор (3). Применим к элементам вектора, получаемого в результате этого умножения, функцию:

$$\operatorname{sign}\left(a\right)=\begin{cases}1, & \text{если } a\geq 0,\\ -1, & \text{если } a<0,\end{cases}\tag{37}$$

и получим вектор:

$$\mathbf{s}\left(i_{\text{шаг}}\right)=\left(s\left(1,i_{\text{шаг}}\right)\quad s\left(2,i_{\text{шаг}}\right)\quad\ldots\quad s\left(N_{\alpha}N_{\beta},i_{\text{шаг}}\right)\right),\tag{38}$$

где

$$s\left(i,i_{\text{шаг}}\right)=\operatorname{sign}\left(\sum_{k=1}^{N_{\gamma}}\gamma\left(k\right)D\left(k,i,i_{\text{шаг}}\right)\right),\tag{39}$$

где $i\in\left\{1,2,\ldots,N_{\alpha}N_{\beta}\right\}$. Результаты натурных экспериментов показали, что в некоторых случаях для повышения вероятности успешности передачи выгодно отказаться от применения функции (37) в равенстве (39), то есть использовать такой вариант:

$$s\left(i,i_{\text{шаг}}\right)=\sum_{k=1}^{N_{\gamma}}\gamma\left(k\right)D\left(k,i,i_{\text{шаг}}\right).\tag{39a}$$

Переформатируем вектор (38) в матрицу:

$$\mathbf{S}\left(i_{\text{шаг}}\right)=\begin{pmatrix}S\left(1,1,i_{\text{шаг}}\right) & S\left(1,2,i_{\text{шаг}}\right) & \cdots & S\left(1,N_{\beta},i_{\text{шаг}}\right)\\ S\left(2,1,i_{\text{шаг}}\right) & S\left(2,2,i_{\text{шаг}}\right) & \cdots & S\left(2,N_{\beta},i_{\text{шаг}}\right)\\ \vdots & \vdots & \ddots & \vdots\\ S\left(N_{\alpha},1,i_{\text{шаг}}\right) & S\left(N_{\alpha},2,i_{\text{шаг}}\right) & \cdots & S\left(N_{\alpha},N_{\beta},i_{\text{шаг}}\right)\end{pmatrix},\tag{40}$$

где

$$S\left(i,j,i_{\text{шаг}}\right)=s\left(k,i_{\text{шаг}}\right),\tag{41}$$

где $k\in\left\{1,2,\ldots,N_{\alpha}N_{\beta}\right\}$, а зависимость между i,j,k определяется следующими равенствами:

$$i=\left\lfloor\frac{k-1}{N_{\beta}}\right\rfloor+1,\tag{42}$$

$$j=\left(\left(k-1\right)\bmod N_{\beta}\right)+1.\tag{43}$$

Теперь вычислим скалярные произведения, между вектором (2) и теми векторами, которые извлечены из (28) и которые представляют

собой строки матрицы (40), скалярно умножив эту матрицу на вектор (2). К элементам вектора, полученного в результате этого умножения, применим функцию (37). И наконец, после применения этой функции вычислим скалярное произведение полученного вектора и вектора (1), что даст целое число:

$$\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)=\sum_{i=1}^{N_\alpha}\alpha\left(i\right)\operatorname{sign}\left(\sum_{k=1}^{N_\beta}S\left(i,k,i_{\text{шаг}}\right)\beta\left(k\right)\right), \tag{44}$$

лежащее в диапазоне от $-N_\alpha$ до N_α.

Решение о том, что в векторе (24) скрыт вектор (4), будет положительным, если будут выполняться три условия. Первое условие заключается в том, чтобы выполнялось неравенство:

$$\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right|\geq\rho_{\text{порог}}, \tag{45}$$

где $\rho_{\text{порог}}$ — это заданное положительное целое число, удовлетворяющее неравенству $\rho_{\text{порог}}\leq N_\alpha$.

Второе условие учитывает не только рассматриваемый вектор (24), но и $W_{\text{справа}}-1$ последующих векторов, а формулируется оно так: абсолютное значение $\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right|$ должно быть наибольшим среди всех абсолютных значений $\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}+j\right)\right|$, где $j\in\left\{0,1,\ldots,W_{\text{справа}}-1\right\}$, то есть должно выполнятся равенство:

$$\begin{aligned}\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right|=\max\Big\{&\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right|,\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}+1\right)\right|,\ldots\\&\ldots,\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}+W_{\text{справа}}-1\right)\right|\Big\}.\end{aligned} \tag{46}$$

Обычно в качестве величины $W_{\text{справа}}$ следует брать положительное целое число, не меньшее произведения $N_{\text{блок}}N_\gamma$.

Для упрощения формулирования третьего условия введем сумму:

$$\mu\left(i_{\text{шаг}}\right)=\sum_{i=1}^{N_\alpha}\left(\text{sign}\left(\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right)\alpha\left(i\right)N_\beta-\sum_{k=1}^{N_\beta}S\left(i,k,i_{\text{шаг}}\right)\beta\left(k\right)\right)^2. \quad (47)$$

Итак, третье условие: для последовательности (24) $\mathbf{r}\left(i_{\text{шаг}}\right)$ должно выполняться равенство:

$$\mu\left(i_{\text{шаг}}\right)=\min\left\{\mu\left(i_{\text{шаг}}+j_1\right),\ldots,\mu\left(i_{\text{шаг}}+j_{\text{К}}\right)\right\}, \quad (48)$$

где $j_k\in\left\{j_1,\ldots,j_{\text{К}}\right\}\subseteq\left\{0,1,\ldots,W_{\textit{справа}}-1\right\}$ — это такие целые числа, для которых выполняется равенство $\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}\right)\right|=\left|\rho_{\text{цифр}}\left(i_{\text{шаг}}+j_k\right)\right|$.

Цифро-аналоговое корреляционное детектирование скрытого сигнала. Помимо описанного выше цифрового метода результаты натурных экспериментов показали полезность гибридного — цифро-аналогового метода. Процедура цифро-аналогового корреляционного детектировании выполняется следующим образом. Вначале вычисляется величина (44). Дальнейшие действия повторяют те, которые выполняются в цифровом подходе, однако в них не используется функция (37). Так, в цифро-аналоговом подходе элементы вектора (38) будут вычисляться по формуле (39а) и в общем случае окажутся вещественными числами. Матрица (40) будет составляться именно из вещественных значений (39а). В конце будет вычисляться величина:

$$\rho_{\text{аналог}}\left(i_{\text{шаг}}\right)=\sum_{i=1}^{N_\alpha}\alpha\left(i\right)\sum_{k=1}^{N_\beta}S\left(i,k,i_{\text{шаг}}\right)\beta\left(k\right). \quad (49)$$

Этот подход называется цифро-аналоговым, так как решение о наличии скрытого сигнала принимается не только на основании величины (49), но также используется величина (44). Есть два требования, которым должен удовлетворять вектор (24), чтобы было вынесено положительное решение о наличии в нем скрытого сигнала. Первое требование совпадает с первым требованием цифрового корреляционного детектирования, определенном неравенством (45). Второе требование состоит в том, что величина (49) должна удовлетворять равенству:

$$\begin{aligned}\left|\rho_{\text{аналог}}\left(i_{\text{шаг}}\right)\right|=\max\left\{\left|\rho_{\text{аналог}}\left(i_{\text{шаг}}\right)\right|,\left|\rho_{\text{аналог}}\left(i_{\text{шаг}}+1\right)\right|,\ldots\right.\\ \left.\ldots,\left|\rho_{\text{аналог}}\left(i_{\text{шаг}}+W_{\text{справа}}-1\right)\right|\right\}.\end{aligned} \quad (50)$$

Иногда оказывается полезным предъявлять еще и третье требование, заключающее в том, чтобы величина (49) превосходила заданный вещественный порог $\rho_{\text{апорог}}\left(i_{\text{шаг}}\right)$:

$$\rho_{\text{аналог}}\left(i_{\text{шаг}}\right) \geq \rho_{\text{апорог}}\left(i_{\text{шаг}}\right). \tag{51}$$

Выбор величины порога $\rho_{\text{апорог}}\left(i_{\text{шаг}}\right)$ зависит от условий распространения передаваемого стегоаудиосигнала. Также полезно сделать его зависимым от характеристик самого исходного аудиосигнала, однако в таком случае приемнику потребуются сведения об исходном аудиосигнале.

Следует отметить, что в цифро-аналоговом подходе при проверке первого требования может использоваться числовое значение $\rho_{\text{порог}}$, меньшее, чем значение этой же величины в цифровом подходе, когда выполняется обработка одной и той же принятой последовательности отсчетов, однако успешность детектирования при этом обычно оказывается выше.

Восстановление переданного информационного бита. Когда последовательность (24) удовлетворяет всем требованиям, указанным в выбранном способе корреляционного детектирования, тогда она рассматривается как содержащая скрытый сигнал (4). В таком случае значение скрытого бита вычисляется следующим образом:

$$x' = \begin{cases} 1, & \text{если} \quad \operatorname{sign}\left(\rho\left(i_{\text{шаг}}\right)\right) = 1, \\ 0, & \text{если} \quad \operatorname{sign}\left(\rho\left(i_{\text{шаг}}\right)\right) = -1, \end{cases} \tag{52}$$

где $\rho\left(i_{\text{шаг}}\right)$ — это либо величина (44), либо величина (49), в зависимости от выбранного способа детектирования. Следует отметить, что искажения в канале могут повлиять так, что значение бита x' может не совпасть с значением переданного бита x.

5. Две методики построения последовательностей $\boldsymbol{\gamma}$. В обычных аудиосигналах смежные блоки отсчетов:

$$\left(r\left(i_{\text{шаг}}\right) \quad r\left(i_{\text{шаг}}+1\right) \quad \ldots \quad r\left(i_{\text{шаг}}+jN_{\text{блок}}-1\right)\right), \tag{53}$$

где $j \in \left\{1,2,\ldots,K_{\text{блок}}\right\}$ обладают очень близкими спектральными характеристиками. Поэтому полезно использовать такую последовательность (3), которая при вычислении элементов вектора (38) учитывает это свойство смежных блоков отсчетов. Опираясь на это свойство, элементы вектора (38) при небольшом шуме в канале и наличии скрытого сигнала (4) в рассматриваемом векторе (24) будут представлять собой слабо зашумленные элементы вектора (2).

Далее представлены две методики построения последовательностей (3). Первая методика предполагает, что последовательность (3) выбирается такая, которая обладает хорошими автокорреляционными свойствами, но при этом в ней подпоследовательности из +1-ц или –1-ц имеют небольшую длину. Однако в процессе вычисления элементов вектора (38) (см. равенства (39) или (39а)) первая методика не использует саму последовательность (3). Вместо нее используется другая последовательность $\boldsymbol{\gamma}'$, вычисляемая следующим образом. Вначале вычисляется вектор $\boldsymbol{\delta}$, представляющий собой разность между первыми $N_\gamma - 1$ элементами вектора (3) и последними $N_\gamma - 1$ элементами вектора (3):

$$\boldsymbol{\delta} = \left(\delta\left(1\right) \quad \delta\left(2\right) \quad \cdots \quad \delta\left(N_\gamma - 1\right)\right), \tag{54}$$

где

$$\delta\left(i\right) = \gamma\left(i\right) - \gamma\left(i+1\right), \tag{55}$$

где $i \in \left\{1,2,\ldots,N_\gamma - 1\right\}$; смежные элементы вектора (3), которые равны между собой, приведут к появлению нулевого элемента в векторе (54). Далее пусть возрастающая последовательность положительных целых $i_1,\ldots,i_K$ составлена из порядковых номеров тех элементов вектора (54), которые не равны 0; при этом для простоты изложения будем считать, что $i_0 = 0$ и $i_{K+1} = N_\gamma$. Используя эти целые числа и элементы вектора (3), сформируем вектор:

$$\boldsymbol{\gamma}' = \left(\gamma'\left(1\right) \quad \gamma'\left(2\right) \quad \ldots \quad \gamma'\left(N_\gamma\right)\right), \tag{56}$$

где

$$\gamma'\left(i\right) = \frac{\gamma\left(i\right)}{i_k - i_{k-1}}, \tag{57}$$

для всякого $i_{k-1} < i \le i_k$ при $k \in \left\{1,2,\ldots,K+1\right\}$. Еще раз следует отметить, что в при использовании этой методики именно вектор (56) используется вместо вектора (3) в процедуре выделения. Использование в приемнике вектора (56) позволяет частично учесть в процессе детектирования спектральное сходство смежных блоков отсчетов. При выборе последовательности (3) среди последовательностей одинаковой длины для первой методики следует отдавать предпочтение той, которая порождает такую последовательность (56), у которой абсолютное значение суммы элементов меньше; например, между двумя последовательностями Баркера $\left(1 \quad -1 \quad 1 \quad 1\right)$ и $\left(1 \quad -1 \quad -1 \quad -1\right)$ следует выбирать вторую.

Вторая методика отличается от первой тем, что в ней используется некоторая последовательность:

$$\boldsymbol{\varphi} = \left(\varphi\left(1\right) \quad \varphi\left(2\right) \quad \ldots \quad \varphi\left(N_{\varphi}\right)\right), \tag{58}$$

где $\varphi\left(i\right) \in \left\{+1,-1\right\}$, которая дублируется $N_{\text{дубль}}$ раз. Дублирование при выборе подходящего вектора (58) будет увеличивать отношение максимального значения к минимальному значению автокорреляционной функции последовательности (3). После дублирования вектора (58) и объединения копий, каждый элемент объединенной последовательности заменяется на последовательность $\boldsymbol{\eta}$ четной длины N_{η}, первая половина которой состоит из одних единиц, а вторая — из минус единиц:

$$\boldsymbol{\eta} = \left(\eta\left(1\right) \quad \eta\left(2\right) \quad \ldots \quad \eta\left(N_{\eta}\right)\right), \tag{59}$$

где

$$\eta\left(i\right) = \begin{cases} 1, & \text{если} \quad i \le \dfrac{N_{\eta}}{2}, \\ -1, & \text{если} \quad i > \dfrac{N_{\eta}}{2}, \end{cases} \tag{60}$$

где $i \in \left\{1,2,\ldots,N_{\eta}\right\}$. При этом, когда выполняется замена, тогда заменяющая последовательность (59) масштабируется с коэффициентом,

представляющим собой заменяемый элемент. В итоге после замены получается последовательность (3), длина которой равна $N_{\gamma} = N_{\text{дубль}} N_{\varphi} N_{\eta}$, а элементы удовлетворяют равенству:

$$\gamma\left(j\right) = \eta\left(\left(\left(j-1\right) \bmod N_{\eta}\right)+1\right)\varphi\left(\left(\left[\frac{j-1}{2}\right] \bmod N_{\varphi}\right)+1\right), \qquad (61)$$

где $j \in \left\{1,2,\ldots,N_{\gamma}\right\}$, $\left[a\right]$ — целая часть числа a. Если использовать оператор произведения Кронекера, то последовательность (3) можно получить таким образом:

$$\boldsymbol{\gamma} = \begin{pmatrix}1 & 1 & \ldots & 1\end{pmatrix} \otimes \boldsymbol{\varphi} \otimes \boldsymbol{\eta}, \qquad (62)$$

где первый вектор, составленный из одних единиц, имеет длину $N_{\text{дубль}}$, а $\otimes$ — это оператор произведения Кронекера. Во второй методике, в отличие от первой, эта же последовательность (3) используется в процедуре вычислении элементов вектора (38). Выбор значения N_{η} следует делать с учетом того, сколько смежных блоков отсчетов аудиосигнала имеют приблизительно одинаковые амплитудные спектры. Вторая методика уже при передаче учитывает спектральное сходство смежных блоков отсчетов.

Анализ свойств последовательностей $\boldsymbol{\gamma}$. Определим взаимную корреляционную функцию последовательностей $\mathbf{a} = \begin{pmatrix}a\left(1\right) & a\left(2\right) & \ldots & a\left(N\right)\end{pmatrix}$, $\mathbf{b} = \begin{pmatrix}b\left(1\right) & b\left(2\right) & \ldots & b\left(N\right)\end{pmatrix}$ из вещественных чисел следующим образом:

$$\mathrm{f}_{\text{кор}}\left(\mathbf{a},\mathbf{b},i\right) \triangleq \sum_{j=1}^{i} a\left(j\right) b\left(N-i+j\right),\ i \geq 1, \qquad (63)$$

при этом будем считать, что $a\left(k\right) = b\left(k\right) = 0$, когда $k > N$ или $k \leq 0$. Обозначим через $\mathrm{F}\left(\mathbf{a},\mathbf{b}\right)$ абсолютное значение отношения максимального значения функции (63) к минимальному значению этой же функции, то есть:

$$\mathrm{F}\left(\mathbf{a},\mathbf{b}\right) \triangleq \left|\frac{\max\limits_{\forall i}\left\{\mathrm{f}_{\text{кор}}\left(\mathbf{a},\mathbf{b},i\right)\right\}}{\min\limits_{\forall i}\left\{\mathrm{f}_{\text{кор}}\left(\mathbf{a},\mathbf{b},i\right)\right\}}\right|. \qquad (64)$$

А также для краткости изложения обозначим через $\mathrm{F}\left(\mathbf{a}\right)$ величину $\left|\max\limits_{\forall i}\left\{\mathrm{f}_{\text{кор}}\left(\mathbf{a},\mathbf{a},i\right)\right\} / \min\limits_{\forall i}\left\{\mathrm{f}_{\text{кор}}\left(\mathbf{a},\mathbf{a},i\right)\right\}\right|$.

Путем выбора подходящих последовательностей (1) и (2) можно добиться того, что будет выполняться как минимум приближенное равенство:

$$\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},1,1,x\right)\right) \approx \mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},1,x\right)\right). \tag{65}$$

В качестве примера приведем следующий случай. Пусть последовательности (1) и (2) — это, соответственно, кодовые слова кодов Касами и Голда, длины которых 255 и 127. При этом в качестве кодового слова кода Касами будет использоваться последовательность, порождаемая регистром с обратной связью, описываемой полиномом $x^8 + x^4 + x^3 + x^2 + 1$. Тогда так в качестве кодового слова кода Голда будет использоваться последовательность, порождаемая регистром с обратной связью, описываемой полиномом $x^7 + x^3 + x^2 + x + 1$. При таких последовательностях (1) и (2) будет выполнятся равенство (65), то есть замена элементов последовательности (1) на элементы последовательности (2) не приведет к изменению абсолютного значения отношения максимального значения автокорреляционной функции к ее минимальному значению. Действительно, когда информационный бит $x = 1$, тогда $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},1,1,x\right)\right) = \left|-255 \,/\, 30\right| = 8.5$, но при этом также выполняется равенство $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},1,x\right)\right) = \left|-32385 \,/\, 3810\right| = 8.5$.

Когда последовательности (1) и (2) таковы, что выполняется равенство (65), тогда особую важность приобретает выбор подходящей последовательности (3). Для простоты будем обозначать последовательности (3), получаемые первой или второй методикой построения, описанным выше, символами $\boldsymbol{\gamma}^{(1)}$ и $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}, N_{\text{дубль}}, N_{\eta}\right)$ соответственно. В таком случае именно последовательность (3) будет определять величину $\mathrm{F}\left\{\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(1)},x\right),\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}',x\right)\right\}$, когда используется первый способ построения последовательностей (3), и величину $\mathrm{F}\left\{\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}, N_{\text{дубль}}, N_{\eta}\right),x\right)\right\}$, когда используется второй способ. В таблице 1 приведены значения $\mathrm{F}\left\{\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(1)},1\right),\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}',1\right)\right\}$ для некоторых последовательностей (3). В таблице 2 приведены некоторые из тех

последовательностей (58), которые при увеличении величины $N_{\text{дубль}}$ увеличивают величину $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$, а также указаны такие, которые не приводят к такому увеличению. При этом в качестве последовательностей (1) и (2) использовались упомянутые ранее последовательности Касами и Голда, длины которых 255 и 127 соответственно.

Таблица 1. Примеры зависимости абсолютного значения отношения максимальной величины корреляционной функции к ее минимальному значению от последовательностей (3) при использовании первого метода построения этих последовательностей

Последовательность $\boldsymbol{\gamma}$	Величина $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(1)},1\right),\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}',1\right)\right)$
$1,1,-1$	2
$1,1,1,-1,-1,1,-1$	~3.01
$1,1,1,-1,-1,-1,1,-1,-1,1,-1$	~3.01

Таблица 2. Примеры зависимости абсолютного значения отношения максимальной величины корреляционной функции к ее минимальному значению от последовательностей (3) при использовании второго метода построения этих последовательностей

		$\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},2\right),1\right)\right)$		
Десятичный эквивалент	Последовательность $\boldsymbol{\varphi}$	$N_{\text{дубль}}=1$	$N_{\text{дубль}}=10$	$N_{\text{дубль}}=100$
1	1	2.033	1.0535	1.0051
1	1,-1	2.033	1.0535	1.0051
3	1,1,-1	1.9892	2.8549	2.9848
13	1,-1,1,1	2.6812	2.0521	2.0051
1	1,-1,-1,-1	1.5948	1.9504	1.9949
3	1,1,-1,-1,-1	1.4253	1.6389	1.6638
5	1,-1,1,-1,-1,-1	2.3922	2.9257	2.9949

В целом о последовательностях (58), длины которых равны 3, 4 или 5 можно сказать следующее:

– среди последовательностей (58), длины 3, только последовательности (1,1,1) и (–1, –1, –1), которые соответствуют числам 7 и 0 соответственно, не приводят к увеличению

$\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ при увеличении $N_{\text{дубль}}$ при любом N_{η}

Остальные с ростом величины $N_{\text{дубль}}$ приближают значение $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ к 3.

– среди последовательностей (58) длины 4 только двоичные последовательности, которые соответствуют числам из множества $\{1,7,8,14\}$, увеличивают величину $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ при увеличении $N_{\text{дубль}}$ при любом N_{η}. Вторая половина этого множества может быть получена по первой путем обращения двоичных разрядов соответствующих чисел первой половины, например, числу 7 соответствует последовательность (1,1,1, –1), а числу 8 — последовательность (–1, –1, –1,1). Выбирая элементы из этого множества и увеличивая $N_{\text{дубль}}$ при любом допустимом N_{η}, значение $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ приближается к 2.

– среди последовательностей (58) длины 5 только двоичные последовательности, которым соответствуют числа из множества $\{1,3,7,10,15,16,21,24,28,30\}$, увеличивают величину $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ при увеличении $N_{\text{дубль}}$, при любом допустимом N_{η}. Выбирая элементы из этого множества и увеличивая $N_{\text{дубль}}$ при любом N_{η}, значение $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ приближается к $1\frac{2}{3}$. Следует отметить, что при выборе последовательности (58), соответствующей числу из множества $\{6,12,19,25\}$, величина $\mathrm{F}\left(\mathbf{y}\left(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right),1\right)\right)$ равна $1\frac{2}{3}$ и вовсе не меняется при увеличении $N_{\text{дубль}}$ при любом допустимом N_{η}.

6. Экспериментальная оценка вероятности успешной передачи. Основными параметрами предлагаемой методики являются по-

следовательности (1), (2), (3), сам исходный скрывающий (или несущий) сигнал (7), количество отсчетов $N_{\text{блок}}$ в каждом блоке, в который будет выполняться встраивание, а также множество порядковых номеров спектральных линий $\mathbf{A}_{\text{НЛ}}\left(j\right)$, которые будут подвергнуты модификации, и множества величин сил встраивания $\mathbf{A}_{\text{СВ}}\left(j\right)$, где $j \in \left\{1,2,\ldots,K_{\text{блок}}\right\}$. При выполнении выделения скрытого сигнала важным параметром является размер окна $W_{\text{справа}}$ считывания.

При выполнении имитационного моделирования, а также натурных экспериментов в качестве последовательностей (1) и (2) будем использовать коды Касами и Голда, длины которых $N_{\alpha} = 255$ и $N_{\beta} = 127$ соответственно. Следовательно, количество блоков отсчетов, требуемых для встраивания скрытого сигнала будет равно $K_{\text{блок}} = 255N_{\gamma}$.

Размер блока отсчетов сделаем равным $N_{\text{блок}} = 256,$ при этом в каждом из $K_{\text{блок}}$ блоков для встраивания будут использоваться спектральные линии с одними и теми же порядковыми номерами от 2 до 128. Таким образом, для всех $j \in \left\{1,2,\ldots,K_{\text{блок}}\right\}$ будут использовать одинаковые множества $\mathbf{A}_{\text{НЛ}}\left(j\right) = \left\{2,3,\ldots,128\right\}$.

Силу встраивания сделаем равной 0.1 для любой спектральной линии; значит, для всех блоков $j \in \left\{1,2,\ldots,K_{\text{блок}}\right\}$ и порядковых номеров $k \in \left\{1,2,\ldots,N_{\textit{блок}}\right\}$ будет $A_{\text{СВ}}\left(j,k\right) = 0.1$; такая величина силы встраивания обычно обеспечивает слуховую транспарентность (неслышимость) скрываемого сигнала. Размер окна $W_{\text{справа}} = 256N_{\gamma}$.

Результаты имитационного моделирования. На вероятность успешной передачи влияют как условия распространения аудиосигнала, так и характеристики приемника и передатчика. Элемент матрицы $\mathbf{D}\left(i_{\text{шаг}}\right)$ (см. формулу (34)), расположенный в *i*-й строке и *j*-м столбце, будем моделировать как удовлетворяющий следующему равенству:

$$D\left(i,j,i_{\text{шаг}}\right) = \ln\left|h\left(i,j,i_{\text{шаг}}\right)Z'_{\text{блок}}\left(k_1\left(i,j\right),A_{\text{НЛ}}\left(k_1\left(i,j\right),k_2\left(j\right)\right)\right) + n_{\text{фон}}\left(i,j,i_{\text{шаг}}\right)\right|, \tag{67}$$

где $i \in \left\{1,2,\ldots,N_{\gamma}\right\}$, $j \in \left\{1,2,\ldots,N_{\alpha}N_{\beta}\right\}$,

$$k_1\left(i,j\right)=\left\lfloor\frac{j-1}{N_\beta}\right\rfloor N_\gamma+i, \tag{68}$$

$$k_2\left(j\right)=\left(\left(j-1\right)\bmod N_\beta\right)+1, \tag{69}$$

$h\left(i,j,i_{\text{шаг}}\right)$ — это комплексный коэффициент передачи стегоаудиосигнала по воздушному аудиоканалу; $n_{\text{фон}}\left(i,j,i_{\text{шаг}}\right)$ — комплексный аддитивный шум, который представляет собой фоновый акустический шум, присутствующий в среде распространения стегоаудиосигнала. Будем предполагать, что величины $h\left(i,j,i_{\text{шаг}}\right)$ будут сохранять свои значения в течение временных интервалов передачи $N_\gamma N_{\text{блок}}$ отсчетов стегоаудиосигнала, то есть интервалов времени, равных $N_\gamma N_{\text{блок}}\,/\,F_{\text{s}}$ секундам, где F_{s} — частота дискретизации. Значит, будет выполняться следующая последовательность равенств $h\left(1,j,i_{\text{шаг}}\right)=h\left(2,j,i_{\text{шаг}}\right)=\ldots=h\left(N_\gamma,j,i_{\text{шаг}}\right)$; такие модели каналов называются моделями с блоковыми замираниями, их разновидности рассмотрены в работах [19, 20].

В процедуре выделения скрытого сигнала не предполагалось знание приёмником скрывающего (несущего) аудиосигнала, вследствие этого его можно отнести к коэффициенту передачи h. Кроме этого, если учесть, что в процедуре встраивания применялась амплитудная модуляция частотных составляющих скрывающего аудиосигнала, то, внеся соответствующие изменения в приведенную выше модель, элемент матрицы $\mathbf{D}\left(i_{\text{шаг}}\right)$, расположенный в *i*-й строке и *j*-м столбце, можно описать следующим равенством:

$$\begin{aligned}D\left(i,j,i_{\text{шаг}}\right)=\ln\Big|h\left(i,j,i_{\text{шаг}}\right)\Big(1+A_{\text{СВ}}\left(k_1\left(i,j\right),k_2\left(j\right)\right)\times\\ \times y\Big(k_3\left(i,j\right),\alpha\Big(k_4\left(j\right)\Big),\beta\Big(k_2\left(j\right)\Big),\gamma\left(i\right),x\Big)\Big)+n_{\text{фон}}\left(i,j,i_{\text{шаг}}\right)\Big|,\end{aligned} \tag{70}$$

где

$$k_3\left(i,j\right)=\left(j-1\right)N_\gamma+i, \tag{71}$$

$$k_4\left(j\right)=\left\lfloor\frac{j-1}{N_\beta}\right\rfloor+1, \tag{72}$$

тогда как величины k_1 и k_2 были определены равенствами (68) и (69). В качестве коэффициентов передачи и фонового шума будем использовать случайные гауссовы комплексные величины, а именно:

$$h\left(i,j,i_{\text{шаг}}\right)=\frac{1}{\sqrt{2}}\left(h_1\left(j,i_{\text{шаг}}\right)+\iota h_2\left(j,i_{\text{шаг}}\right)\right), \tag{73}$$

$$n_{\text{фон}}\left(i,j,i_{\text{шаг}}\right)=\frac{\sigma_{\text{фон}}}{\sqrt{2}}\left(n_{\text{фон},1}\left(i,j,i_{\text{шаг}}\right)+\iota n_{\text{фон},2}\left(i,j,i_{\text{шаг}}\right)\right), \tag{74}$$

где $\iota=\sqrt{-1}$ — мнимая единица, а $h_1\left(j,i_{\text{шаг}}\right)$, $h_2\left(j,i_{\text{шаг}}\right)$, $n_{\text{фон},1}\left(i,j,i_{\text{шаг}}\right)$ и $n_{\text{фон},2}\left(i,j,i_{\text{шаг}}\right)$ — это статистически независимые одинаково распределенные гауссовы случайные величины, имеющие нулевое математическое ожидание и единичную дисперсию. Меняя вещественную величину $\sigma_{\text{фон}}$, мы будем изменять дисперсию амплитуды аддитивного шума, которая в таком случае будет равна $\sigma_{\text{фон}}^2$, тогда как дисперсия амплитуды мультипликативного шума будет оставаться неизменной и равной 1.

Так же, как и выше, обозначим через $\boldsymbol{\gamma}^{(1)}$ последовательность $\boldsymbol{\gamma}$, полученную первым способом построения, а через $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi},N_{\text{дубль}},N_{\eta}\right)$ обозначим последовательность $\boldsymbol{\gamma}$, полученную вторым способом. При выполнении имитационного моделирования использовались последовательности $\boldsymbol{\gamma}_1^{(1)}=\boldsymbol{\varphi}_1$, $\boldsymbol{\gamma}_2^{(1)}=\boldsymbol{\varphi}_2$, а также $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1,1,2\right)$, $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_2,1,2\right)$, $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1,2,2\right)$, $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_2,2,2\right)$, где $\boldsymbol{\varphi}_1=\begin{pmatrix}1 & 1 & -1\end{pmatrix}$ и $\boldsymbol{\varphi}_2=\begin{pmatrix}1 & 1 & 1 & -1 & -1 & 1 & -1\end{pmatrix}$. Так, например:

$$\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1,1,2\right)=\begin{pmatrix}1 & -1 & 1 & -1 & -1 & 1\end{pmatrix},$$

$$\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1,2,2\right)=\begin{pmatrix}1 & -1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 & 1\end{pmatrix}.$$

Так как в приемнике при использовании первой методики вместо последовательности (3) используется (56), то, например, вместо $\boldsymbol{\gamma}_2^{(1)}$ в приемнике использовалась последовательность:

$$\boldsymbol{\gamma}_2'=\begin{pmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3} & -\frac{1}{2} & -\frac{1}{2} & 1 & -1\end{pmatrix}.$$

Вероятность успешной передачи $P_{\text{успех}}$ информационного бита $x\in\{0,1\}$ зависит от двух других: вероятности $P_{\text{детек}}$ детектировать в

принятом сигнале скрытый сигнал и вероятности восстановления $P_{\text{восст}}$ переданного бита. Первая вероятность зависит от того, найдется ли последовательность отсчетов (24) такая, которая будет удовлетворять всем требованиям, указанным ранее в описании процесса выделения скрытого сигнала. Вторая вероятность зависит от того, совпадет ли знак числа $\rho\left(i_{\text{шаг}}\right)$ в равенстве (52) со знаком числа $2x-1$. Таким образом, $P_{\text{успех}}$ определяется следующим равенством:

$$P_{\text{успех}} = P_{\text{детек}} P_{\text{восст}}. \tag{75}$$

Выполняя имитационное моделирования предполагалось, что для любой реализации матрицы (33) решение о наличии скрытого сигнала было положительным, это позволило сосредоточиться только на оценке вероятности восстановления $P_{\text{восст}}$ переданного бита x при различных значениях дисперсии фонового шума $\sigma^2_{\text{фон}}$. Таким образом, выполняя имитационное моделирование, предполагалось, $P_{\text{детек}} = 1$, следовательно, выполнялось равенство $P_{\text{успех}} = P_{\text{восст}}$.

В таблицах 3 и 4 приведены результаты имитационного моделирования при различных значениях дисперсии фонового шума $\sigma^2_{\text{фон}}$.

Таблица 3. Таблица вероятностей успешного восстановления переданного бита для имитационного моделирования при использовании первой методики построения последовательностей (3)

Дисперсия $\sigma^2_{\text{фон}}$ амплитуды фонового шума, выраженная в дБ	Вероятность восстановления $P_{\text{восст}}$ переданного бита x (цифровое детектирование)	Вероятность восстановления $P_{\text{восст}}$ переданного бита x (цифро-аналоговое детектирование)
результаты для последовательности $\boldsymbol{\gamma}_1^{(1)}$		
30	0.5143	0.5221
25	0.5385	0.5547
20	0.6316	0.6712
15	0.8389	0.9179
10	0.9988	0.9999
результаты для последовательности $\boldsymbol{\gamma}_2^{(1)}$		
30	0.5139	0.5257
25	0.5539	0.5827
20	0.6714	0.7478
15	0.9196	0.9801

Таблица 4. Таблица вероятностей успешного восстановления переданного бита для имитационного моделирования при использовании второй методики построения последовательностей (3)

Дисперсия $\sigma^2_{\text{фон}}$ амплитуды фонового шума, выраженная в дБ	Вероятность восстановления $P_{\text{восст}}$ переданного бита x (цифровое детектирование)	Вероятность восстановления $P_{\text{восст}}$ переданного бита x (цифро-аналоговое детектирование)
результаты для последовательности $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1, 1, 2\right)$		
30	0.5217	0.5226
25	0.5570	0.5844
20	0.6768	0.7494
15	0.9196	0.9831
результаты для последовательности $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1, 2, 2\right)$		
30	0.5264	0.5373
25	0.5872	0.6245
20	0.7291	0.8368
15	0.9715	0.9987
результаты для последовательности $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_2, 1, 2\right)$		
30	0.5250	0.5416
25	0.5751	0.6240
20	0.7489	0.8477
15	0.9821	0.9991
результаты для последовательности $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_2, 2, 2\right)$		
30	0.5431	0.5606
25	0.6217	0.6848
20	0.8296	0.9272
15	0.9988	0.9999

По результатам видно, что предлагаемая методика скрытой передачи сигналов является очень устойчивой к аддитивному шуму. Это можно заключить на основании следующего. При дисперсии амплитуды фонового шума 15 дБ уже практически невозможно на слух различить сам стегоаудиосигнал, однако даже при такой дисперсии обеспечивается вероятность восстановления больше 0.8.

Результаты натурных экспериментов. Длительность скрытой передачи одного бита равна отношению $N_\alpha N_\gamma N_{\text{блок}} / F_{\text{s}}$, где F_{s} — ча-

стота выборки. Для примера рассчитаем эту длительность при следующих значениях параметров. Пусть последовательности (1) и (2) имеют длины $N_{\alpha} = 255$ и $N_{\beta} = 127$ двоичных символов соответственно. Значит, если $N_{\text{блок}} = 256,$ то при использовании последовательности (3), длина которой $N_{\gamma} = 3,$ и частоте выборки $F_{\text{s}} = 32$ кГц окажется, что длительность скрытой передачи одного бита равна $195840 / 32000 = 6.12$ секунды.

В реальных условиях, помимо характеристик аудиосистемы, включающей микрофон и динамик, на вероятность успешной передачи влияют: окружающий акустический шум, характеристики пути, который проходит стегоаудиосигнал от динамика до микрофона, характеристики самого исходного аудиосигнала, в который внедряется информация. Натурные эксперименты были выполнены в лабораторных условиях: тихое помещение, отсутствие препятствий между микрофоном и динамиком, а сами они неподвижны. Менялось только расстояние между динамиком и микрофоном. В качестве аппаратных средств передачи и приема аудиосигналов использовались динамики Sennheiser MX170, микрофон Philips SBC ME570. Мощность динамика в операционной системе Windows была установлена на 50% при проведении всех натурных экспериментов. Цифровая обработка аудиосигналов выполнялась с помощью программной среды MATLAB. Пороговое значение $\rho_{\text{порог}}$ было задано равным 71.

При проведении экспериментов в качестве исходных аудиосигналов использовались следующие два их вида: речевая запись и музыкальная композиция. Речевой аудиосигнал характеризовался слабостью высоких частотных составляющих, так как она была записана в условиях, когда высокочастотные шумы имели очень малую мощность. В качестве музыкальной композиции использовалась песня «This Moment» музыкальной группы «In This Moment»; эта композиция характеризуется помимо низкочастотных составляющих наличием еще и мощных высокочастотных составляющих. Оба выбранных аудиосигнала позволяют выполнять внедрение с силой встраивания $A_{\text{СВ}}\left(j,i\right),$ равной 0.1 для всех допустимых пар $\left(j,i\right)$, при этом не появляется каких-либо слышимых звуковых артефактов, вызванных внедрением скрытого сигнала.

По результатам натурных экспериментов (таблица 5) передачи 100 бит информации можно сделать следующие выводы:

Таблица 5. Таблица вероятности успешной передачи, построенная по результатам натурных экспериментов, при использовании цифрового детектирования

Расстояние между динамиком и микрофоном (см) / Используемая последовательность $\boldsymbol{\gamma}$	50	60	70
$\boldsymbol{\gamma}_1^{(1)} = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$	0.95	0.98	0.90
$\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1, 1, 2\right) = \begin{pmatrix} 1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}$	0.99	0.99	0.99

– использование речевой записи в качестве исходного аудиосигнала позволило выполнить успешную передачу на расстояние, не превышающее 5 см между динамиком и микрофоном; такое небольшое расстояние вызвано небольшой мощностью исходного аудиосигнала.

– использование указанной музыкальной композиции из-за своих частотных характеристик позволило значительно увеличить расстояние между динамиком и микрофоном (таблица 5); так, вероятность успешной передачи на расстояние 70 см оказалась равной 0.99 при использовании последовательности $\boldsymbol{\gamma}^{(2)}\left(\boldsymbol{\varphi}_1, 1, 2\right) = \begin{pmatrix} 1 & -1 & 1 & -1 & -1 & 1 \end{pmatrix}$, и равной 0.9 при использовании последовательности $\boldsymbol{\gamma}_1^{(1)} = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix}$.

7. Заключение. В статье сформулированы свойства, которыми следует наделить скрываемую последовательность, чтобы повысить ее робастность к искажающим влияниям воздушного аудиоканала, когда используется предлагаемый метод цифрового маркирования аудиосигналов. Предложены способы построения последовательностей, обладающих подходящими свойствами. Способ детектирования наличия скрытого сигнала (цифрового маркера), предложенный в статье, является слепым, так как не требует знания приемником скрывающего сигнала. Простота предлагаемого в этой статье метода скрытой передачи данных позволяет применять его в режиме реального времени без необходимости требовать от передатчика и приемника значительной вычислительной мощности.

Литература

1. *Hua G. et al.* Twenty years of digital audio watermarking — a comprehensive review // Signal Processing. 2016. vol. 128. pp. 222–242.
2. *Bassia P., Pitas I., Nikolaidis N.* Robust audio watermarking in the time domain // IEEE Transactions on multimedia. 2001. vol. 3. no. 2. pp. 232–241.
3. *Lemma A.N., Aprea J., Oomen W., Kerkhof L.V.D.* A temporal domain audio watermarking technique // IEEE transactions on signal processing. 2003. vol. 51. no. 4. pp. 1088–1097.
4. *Gruhl D., Lu A., Bender W.* Echo hiding // International Workshop on Information Hiding. 1996. pp. 295–315.
5. *Hu P., Peng D., Yi Z., Xiang Y.* Robust time-spread echo watermarking using characteristics of host signals // Electronics Letters. 2016. vol. 52. no. 1. pp. 5–6.
6. *Xiang Y. et al.* A dual-channel time-spread echo method for audio watermarking // IEEE Transactions on Information Forensics and Security. 2012. vol. 7. no. 2. pp. 383–392.
7. *Kirovski D., Malvar H.S.* Spread-spectrum watermarking of audio signals // IEEE transactions on signal processing. 2003. vol. 51. no. 4. pp. 1020–1033.
8. *Verma C., Tarar S.* Secure Random Sequence based Frequency Hoping Spread Spectrum Audio Watermarking // International Journal of Engineering Science. 2016. vol. 6. no. 5. pp. 5026–5032.
9. *Chen B., Wornell G.W.* Quantization index modulation: A class of provably good methods for digital watermarking and information embedding // IEEE Transactions on Information Theory. 2001. vol. 47. no. 4. pp. 1423–1443.
10. *Tian H., Liu J., Li S.* Improving security of quantization-index-modulation steganography in low bit-rate speech streams // Multimedia systems. 2014. vol. 20. no. 2. pp. 143–154.
11. *Yeo I.K., Kim H.J.* Modified patchwork algorithm: A novel audio watermarking scheme // IEEE Transactions on speech and audio processing. 2003. vol. 11. no. 4. pp. 381–386.
12. *Xiang Y. et al.* Patchwork-based audio watermarking method robust to de-synchronization attacks // IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP). 2014. vol. 22. no. 9. pp. 1413–1423.
13. *Wang F. et al.* Simultaneous Broadcasting of Analog FM and Digital Signals by Separating Co-Channel FM Signals // IEEE Communications Letters. 2016. vol. 20. no. 11. pp. 2197–2200.
14. *Lee H., Kim T.H., Choi J.W., Choi S.* Chirp signal-based aerial acoustic communication for smart devices // IEEE Conference on Computer Communications (INFOCOM). 2015. pp. 2407–2415.
15. *Nandakumar R., Chintalapudi K. K., Padmanabhan V., Venkatesan R.* Dhwani: secure peer-to-peer acoustic NFC // ACM SIGCOMM Computer Communication Review. 2013. vol. 43. no. 4. pp. 63–74.
16. *Wang Q. et al.* Messages behind the sound: real-time hidden acoustic signal capture with smartphones // Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking. 2016. pp. 29–41.
17. *Hanspach M., Goetz M.* On covert acoustical mesh networks in air // Journal of Communications. 2013. vol. 8. no. 11. pp. 758–767.
18. *Гофман М.В.* Методика скрытой передачи данных при связи через воздушный аудиоканал // Труды СПИИРАН. 2017. Вып. 51. С. 97–122.
19. *Collins A., Polyanskiy Y.* Dispersion of the coherent MIMO block-fading channel // IEEE International Symposium on Information Theory (ISIT). 2016. pp. 1068–1072.
20. *Гофман М.В.* Помехоустойчивое пространственное блоковое кодирование // LAP Lambert Academic Publishing. 2013. 176 с.

Гофман Максим Викторович — к-т техн. наук, доцент кафедры информатики и информационной безопасности, Петербургский государственный университет путей сообщения Императора Александра I. Область научных интересов: системы связи, системы передачи данных. Число научных публикаций — 17. maxgof@gmail.com; Московский пр., 9, Санкт-Петербург, 190031; р.т.: +7(812)310-34-72, Факс: +7(812)570-76-68.

Корниенко Анатолий Адамович — д-р техн. наук, профессор, заведующий кафедрой информатики и информационной безопасности, Петербургский государственный университет путей сообщения Императора Александра I. Область научных интересов: информатика, информационная безопасность. Число научных публикаций — 200. kaa.pgups@yandex.ru; Московский пр., 9, Санкт-Петербург, 190031; р.т.: +7(812)310-34-72, Факс: +7(812)570-76-68.

Мирончиков Евгений Тимофеевич — д-р техн. наук, профессор, заведующий кафедрой информатики и информационной безопасности, Петербургский государственный университет путей сообщения Императора Александра I. Область научных интересов: теория информации. Число научных публикаций — 200. etm937@yandex.ru; Московский пр., 9, Санкт-Петербург, 190031; р.т.: +7(812)310-34-72, Факс: +7(812)570-76-68.

Никитин Александр Борисович — д-р техн. наук, профессор, заведующий кафедрой автоматики и телемеханики на железных дорогах, Петербургский государственный университет путей сообщения Императора Александра I. Область научных интересов: автоматика и телемеханика на железных дорогах. Число научных публикаций — 200. nikitin@crtc.spb.ru; Московский пр., 9, Санкт-Петербург, 190031; р.т.: +7(812)310-07-88, Факс: +7(812)457-82-92.

DOI 10.15622/sp.55.8

M.V. GOFMAN, A.A. KORNIENKO, E.T. MIRONCHIKOV, A.B. NIKITIN

DIGITAL WATERMARKING OF AUDIO SIGNALS FOR ROBUST HIDDEN AUDIO COMMUNICATION VIA AIR AUDIO CHANNEL

Gofman M.V., Kornienko A.A., Mironchikov E.T., Nikitin A.B. **Digital Watermarking of Audio Signals for Robust Hidden Audio Communication via Air Audio Channel.**

Abstract: The article presents a digital audio watermarking method for air audio data transmission. The digital watermark occupies the whole frequency range of the audio signal. The digital watermark encodes one bit of information. A decision about transmitted bit is based on a sign of the center value of the mutual correlation function. Two methods to design the digital audio watermark are presented. Low complexity of presented methods of digital audio watermarking makes them suitable for use in smartphones. The method can be used in digital audio watermarking of both speech and music. Information is embedded via the digital watermark in the frequency domain of a host audio signal via amplitude modulation of its frequency constituents. The paper includes results of simulation modeling and natural experiments.

Keywords: digital watermark, audio signal, covert data transmission, air audio channel, acoustic telecommunication, steganography audio signal

Gofman Maksim Viktorovich — Ph.D., associate professor of informatics and information security department, Emperor Alexander I st. St. Petersburg State Transport University. Research interests: communication systems, systems of data transmission. The number of publications — 17. maxgof@gmail.com; 9, Moskovsky pr., Saint Petersburg, 190031; office phone: +7(812)310-34-72, Fax: +7(812)570-76-68.

Kornienko Anatoliy Adamovich — Ph.D., Dr. Sci., professor, head of informatics and information security department, Emperor Alexander I st. St. Petersburg State Transport University. Research interests: informatics, information security. The number of publications — 200. kaa.pgups@yandex.ru; 9, Moskovsky pr., Saint Petersburg, 190031; office phone: +7(812)310-34-72, Fax: +7(812)570-76-68.

Mironchikov Evgenij Timofeevich — Ph.D., Dr. Sci., professor, head of informatics and information security department, Emperor Alexander I st. St. Petersburg State Transport University. Research interests: information theory. The number of publications — 200. etm937@yandex.ru; 9, Moskovsky pr., Saint Petersburg, 190031; office phone: +7(812)310-34-72, Fax: +7(812)570-76-68.

Nikitin Aleksandr Borisovich — Ph.D., Dr. Sci., professor, head of automation and telemechanics on the railways department, Emperor Alexander I st. St. Petersburg State Transport University. Research interests: automation and telemechanics on railways. The number of publications — 200. nikitin@crtc.spb.ru; 9, Moskovsky pr., Saint Petersburg, 190031; office phone: +7(812)310-07-88, Fax: +7(812)457-82-92.

References

1. Hua G. et al. Twenty years of digital audio watermarking — a comprehensive review. *Signal Processing*. 2016. vol. 128. pp. 222–242.
2. Bassia P., Pitas I., Nikolaidis N. Robust audio watermarking in the time domain. *IEEE Transactions on multimedia*. 2001. vol. 3. no. 2. pp. 232–241.

3. Lemma A.N., Aprea J., Oomen W., Kerkhof L.V.D. A temporal domain audio watermarking technique. *IEEE transactions on signal processing*. 2003. vol. 51. no. 4. pp. 1088–1097.
4. Gruhl D., Lu A., Bender W. Echo hiding. International Workshop on Information Hiding. 1996. pp. 295–315.
5. Hu P., Peng D., Yi Z., Xiang Y. Robust time-spread echo watermarking using characteristics of host signals. *Electronics Letters*. 2016. vol. 52. no. 1. pp. 5–6.
6. Xiang Y. et al. A dual-channel time-spread echo method for audio watermarking. *IEEE Transactions on Information Forensics and Security*. 2012. vol. 7. no. 2. pp. 383–392.
7. Kirovski D., Malvar H.S. Spread-spectrum watermarking of audio signals. *IEEE transactions on signal processing*. 2003. vol. 51. no. 4. pp. 1020–1033.
8. Verma C., Tarar S. Secure Random Sequence based Frequency Hoping Spread Spectrum Audio Watermarking. *International Journal of Engineering Science*. 2016. vol. 6. no. 5. pp. 5026–5032.
9. Chen B., Wornell G.W. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. *IEEE Transactions on Information Theory*. 2001. vol. 47. no. 4. pp. 1423–1443.
10. Tian H., Liu J., Li S. Improving security of quantization-index-modulation steganography in low bit-rate speech streams. *Multimedia systems*. 2014. vol. 20. no. 2. pp. 143–154.
11. Yeo I.K., Kim H.J. Modified patchwork algorithm: A novel audio watermarking scheme. *IEEE Transactions on speech and audio processing*. 2003. vol. 11. no. 4. pp. 381–386.
12. Xiang Y. et al. Patchwork-based audio watermarking method robust to de-synchronization attacks. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*. 2014. vol. 22. no. 9. pp. 1413–1423.
13. Wang F. et al. Simultaneous Broadcasting of Analog FM and Digital Signals by Separating Co-Channel FM Signals. *IEEE Communications Letters*. 2016. vol. 20. no. 11. pp. 2197–2200.
14. Lee H., Kim T.H., Choi J.W., Choi S. Chirp signal-based aerial acoustic communication for smart devices. IEEE Conference on Computer Communications (INFOCOM). 2015. pp. 2407–2415.
15. Nandakumar R., Chintalapudi K.K., Padmanabhan V., Venkatesan R. Dhwani: secure peer-to-peer acoustic NFC. *ACM SIGCOMM Computer Communication Review*. 2013. vol. 43. no. 4. pp. 63–74.
16. Wang Q. et al. Messages behind the sound: real-time hidden acoustic signal capture with smartphones. Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking. 2016. pp. 29–41.
17. Hanspach M., Goetz M. On covert acoustical mesh networks in air. *Journal of Communications*. 2013. vol. 8. no. 11. pp. 758–767.
18. Gofman M.V. [A Method of Hidden Data in Communication via Air Audio Channel] *Trudy SPIIRAN – SPIIRAS Proceedings*. 2017. vol. 2(51). pp. 97–122. (In Russ.).
19. Collins A., Polyanskiy Y. Dispersion of the coherent MIMO block-fading channel. IEEE International Symposium on Information Theory (ISIT). 2016. pp. 1068–1072.
20. Gofman M.V. Pomehoustojchivoe prostranstvennoe blokovoe kodirovanie [Noiseproof space-time block coding]. LAP Lambert Academic Publishing. 2013. 176 p. (In Russ.).