УДК 004.415.24

ОПРЕДЕЛЕНИЕ ВОЗМОЖНОГО ОБЪЁМА ВНЕДРЯЕМОЙ ИНФОРМАЦИИ ПРИ СКРЫТОЙ ПЕРЕДАЧЕ МЕТОК В РЕЧЕВЫХ ДАННЫХ

Е. Г. ЖИЛЯКОВ
С.Н. ДЕВИЦИНА
П.Г. ЛИХОЛОБ

Белгородский государственный национальный исследовательский университет

e-mail:
Zhilyakov@bsu.edu.ru
Devitsyna@bsu.edu.ru
Likholob@bsu.edu.ru

В работе оценено доступное место для скрытой передачи информации в речевых данных. Предложен алгоритм автоматизированного поиска пространства внедрения меток, основанный на закономерностях распределения энергии в речевых данных. Поиск пространства внедрения меток осуществляется среди узких частотных интервалов, энергия которых не входит в 90-99% энергии. Эмпирически оценен для каждого звука русской речи максимально возможный объем внедряемой информации при заданных условиях.

Ключевые слова: ЦВЗ, метка, доля энергии, частотный интервал, стеганография, речевые данные, скрытая передача информации.

Объединение телекоммуникационных каналов в единое целое позволяет получить удалённый доступ к многочисленным ресурсам. Основной ресурс – это информация, представленная преимущественно в цифровом виде, хранимая, обрабатываемая и передаваемая в виде данных. Наиболее широко используемым и удобным способом обмена информации между людьми является речь. Для передачи и хранения устную речь кодируют, преобразуя её в цифровой код. Устную речь, представленную в цифровом коде, состоящую из последовательности отсчетов или выборок мгновенных значений ($\vec{X} = [x_1, x_2, \ldots, x_l, \ldots, x_L]^T$, $L \in \mathbf{N}$) аналогового сигнала, отстоящих один от другого на одинаковый интервал времени ($T_д = 1/f_д$ – период дискретизации, $f_д$ – частота дискретизации), принято называть речевыми данными. В некоторых случаях речевые данные приобретают особую ценность. Примером может послужить результат обмена служебной информацией, такой как аудиопротокол совещания, переговоры в диспетчерских, а также при передаче речевых команд управления. Существует множество способов защиты данных, позволяющих автоматизировать процессы обработки и накопления информации, осуществить защищенную передачу информации, идентифицировать владельца данных. Нередко для этих задач используют криптографические методы. Криптографические методы перекодируют информацию, зачастую увеличивая объем передаваемых и хранимых данных, усложняя процесс накопления и обработки информации. При этом стойкость и эффективность криптографических методов ограничена вычислительными возможностями владельца информации и правонарушителя. Альтернативой криптографическим методам являются методы, применяемые в стеганографии. Использование средств и методов, применяемых в цифровой стеганографии, позволяет преобразовывать информацию, представленную в виде данных, внедряя в данные дополнительную информацию. Совокупность вносимых незначительных изменений (меток $w_m \in \{1,0\}$), в которых закодирована информация, используемая для маркировки и защиты данных, называется цифровым водяным знаком (ЦВЗ). Искажения, возникающие при преобразовании и последующем воспроизведении, должны быть не ощутимы органами чувств человека. Разнообразие стегоалгоритмов, кодирующих ЦВЗ в данные, показывает стремление разработчиков повысить эффективность работы этих алгоритмов. Эффективность работы стегоалгоритмов определяется скрытностью и стойкостью внедряемой информации. Известно, что стего- система является разрушенной, если обнаружен факт её существования, либо разрушена внедряемая информация. Обнаружение меток зависит от объема внедряемой информации и позиций её внедрения.

Первые стеганографические методы предполагали работу в пространственной области, обеспечивая малый объем и низкую скрытность внедряемого ЦВЗ. С разработкой методов вейвлет-преобразования большинство методов цифровой стеганографии внедряют ЦВЗ в данные, используя разбиение на частотные интервалы в соответствии с параметрами окна данных во временной области. Следовательно, эффективность работы стегоалгоритмов зависит от выбора частотного интервала, в конкретном отрезке (окне) речевых данных. Анализ и использование закономер-

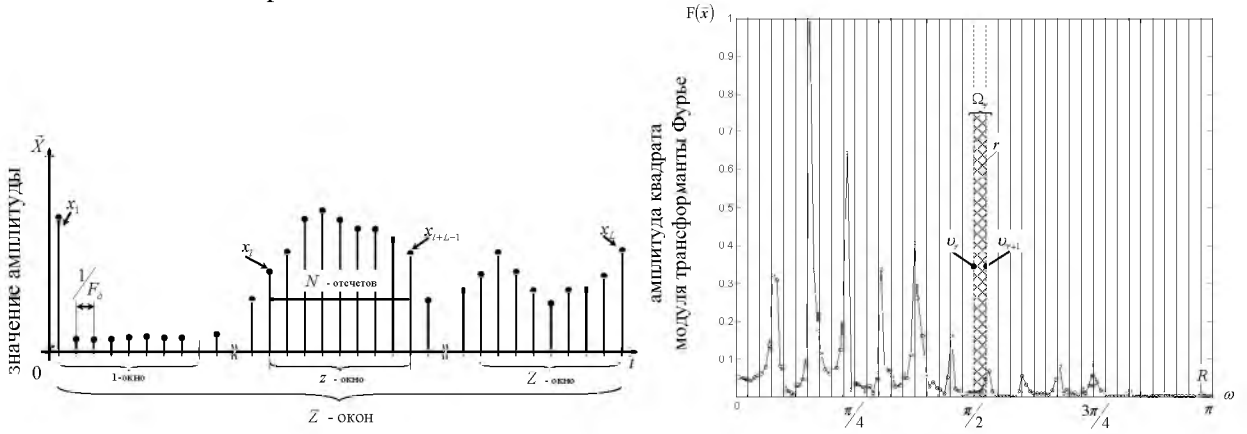
ностей распределения энергии по частотным интервалам отрезка (окна) речевых данных определит доступный объём внедряемой информации, не вызывающий заметных искажений.

Математические основы.

Для анализа речевые данные во временной области условно разбивают на z око, равной длины (рис.1 а).

$$x_n = X_{(z-1)\cdot N+n-1},\ n=1\ldots N,\ z=1\ldots Z, \quad (1)$$

где z – порядковый номер окна во временной области; n – порядковый номер отчета речевых данных; N – длительность окна (в отчетах) анализируемого окна речевых данных $\vec{x}$; Z – количество окон во временной области.



а) б)

Рис. 1. Отрезок речевых данных: а) значения амплитуд во временной области; б) огибающая амплитуд трансформанты Фурье, в частотной области

В рамках данной работы предложено разбивать частотную ось на R частотных интервалов равной ширины Ω_r, с выполнением условия:

$$f_\partial = \frac{1}{2\pi}\cdot\sum_{r=1}^{R}\Omega_r,\ \forall\Omega_r=\Omega \quad (2)$$

где $r=1,\ldots,R$ – номер частотного интервала; R – количество частотных интервалов, на которые разбивается частотная ось.

Часть энергии P_r, для каждого частотного интервала z -окна речевых данных, может быть рассчитана с использованием выражения [1],:

$$P_{z,r} = \frac{1}{2\pi}\int_{\omega\in\Omega_r}|X(\omega)|^2\,d\omega = \vec{x}^{\,\prime}A_r\vec{x}, \quad (3)$$

где $X(\omega)$ - трансформанта Фурье; $A_r=\left\{a_{ik}^r\right\}$ – субполосная матрица, определяемая для каждого из R частотных интервалов с элементами вида:

$$a_{ik}^r = (\sin(\upsilon_{r+1}(i-k))-\sin(\upsilon_r(i-k)))/(\pi(i-k)),\quad i,k=1,\ldots,N \quad (4)$$

где υ_r, υ_{r+1} – границы r-ого частотного интервала, причем:

$$0\le\upsilon_r<\upsilon_{r+1}\le\frac{\pi f_\partial}{2},\ \ (\upsilon_{r+1}-\upsilon_r)=\pi f_\partial/R=\Omega, \quad (5)$$

Матрица A_r обладает полной системой ортонормальных собственных векторов, соответствующих неотрицательным собственным числам и удовлетворяющих соотношениям [2]:

$$\lambda_{kr}\vec{q}_{kr}=A_r\vec{q}_{kr}. \quad (6)$$

где λ_{kr} – собственные числа субполосной матрицы A_r; $\vec{q}_{kr}$ – собственные вектора субполосной матрицы A_r; k – порядковый номер собственного числа и соответствующего ему собственного вектора.

Это свойство можно использовать при анализе речевых данных, т.к. близость к «единице» максимального собственного числа $\lambda_{\max r}$ указывает на степень влияния части энергии сосредоточенной вне частотного интервала на энергию, сосредоточенную внутри частотного интервала Ω_r.

Для анализа речевых данных удобно рассматривать распределение долей энергии по частотным интервалам, рассчитанное для каждого окна речевых данных

$$Pd_r \cong \frac{\vec{x}^{'} A_r \vec{x}}{\left\|\vec{x}\right\|^2} \cdot 100\%, \qquad (7)$$

где Pd_r – значение доли энергии сигнала; $\left\|\vec{x}\right\|^2$ – энергия анализируемого отрезка сигнала.

Распределение энергии по частотным интервалам для каждого звука русской речи разнообразно и области скрытия нецелесообразно привязывать к определенной сетке частот. Это несложно увидеть из рис.2. Следовательно, необходимо делать выбор частотного интервала доступного для скрытия в каждом окне. При этом количество частотных интервалов которые можно использовать для внедрения ЦВЗ будет изменяться от окна к окну.

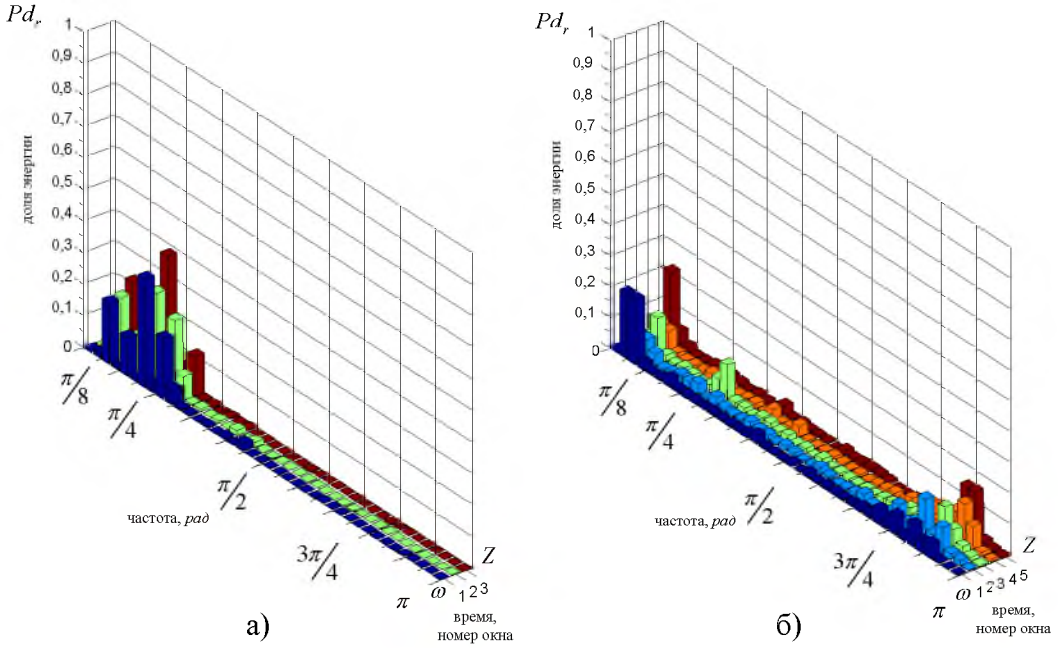


Рис. 2. Результат вычисления долей энергии для отрезков речевых данных соответствующих звукам русской речи: а) – звуку «а» в слове «с**а**вА»; б) – звуку «з» в слове «арбУ**з**»

Алгоритм поиска маски внедрения меток.

Шаг 1. Задать длину окна отрезка речевых данных N.

Шаг 2. Задать значение ширины частотного интервала Ω_r, необходимое для скрытия метки.

Шаг 3. Задать значения доли энергии окна речевых данных m. Примем, что частотные интервалы, доля энергии которых входит в значения доли энергии m, условно считаются информационным.

Шаг 4. Вычислить для всех $r \in R$ частотных интервалов субполосных матриц $A_r = \left\{a_{ik}^r\right\}$ по (3). Контролируя близость первого собственного числа субполосной матрицы к единице, можно оценить степень принадлежности доли энергии анализируемому частотному интервалу Ω_r.

Шаг 5. Вычислить для всех окон речевых данных $\vec{x}$ значения долей энергии, распределенных по частотным интервалам $r \in R$.

Шаг 6. Упорядочить значение долей энергии Pd_r от максимального значению к минимальному.

Шаг 7. Определить информационные частотные интервалы, суммируя упорядоченные доли энергии до превышения порога h. Частотные интервалы, доля энергии которых не вошла в значение доли энергии m, доступны для скрытия в них меток и определяют маску вложения.

Вычислительные эксперименты.

Для исследования использовались речевые данные, представляющие собой отрезки аудиосигнала фраз русской речи, записанных с частотой дискретизации $f_\partial = 8$кГц в 16-битовом представлении в режиме моно. В анализируемых фразах речевых данных выделялись отрезки во временной области, соответствующие звукам русской речи. Звуки (речевые данные) условно разделялись на равные окна длиной $N = 256$ отсчетов (для первого эксперимента) и $N = 512$ отсчетов (для второго эксперимента), количество частотных интервалов $R = 40$ отсчетов (для первого эксперимента) и $R = 80$ (для второго эксперимента).

Для повышения эффективности работы стего- алгоритмов предлагается кодировать (скрывать) метки в узких частотных интервалах речевых данных с долей энергии не входящей в 90-99% энергии отрезка (окна) речевых данных. Количество частотных интервалов выбрано из эмпирически полученного соотношения [4]:

$$J = N / R, \quad (8)$$

где величина $J \geq 6$, что позволяет получить высокую избирательность при вычислении долей энергии в частотных интервалах (при этом максимальное собственное число близко к единице и обеспечивается высокая избирательность энергии внутри частотного интервала).

Таблица 1

Усредненная доля частотных интервалов доступных для внедрения

Гласные										
звук	а	е	ё	и	О	у	ы	э	ю	я
N=256	70%	75%	75%	80%	82,5%	75%	80%	67,5%	52,5%	72,5%
N=512	75%	86,5%	82,5%	91,2%	83,7%	87,5%	87,5%	73,75%	-	81,25%

сонорные согласные					
звук	й	л	м	н	р
N=256	82,5%	80%	85%	87,5%	57,5%
N=512	88,75%	88,75%	93,75%	93,75%	52,5%

звонкие согласные						
звук	б	в	Г	д	ж	з
N=256	77,5%	75%	70%	75%	65%	62,5%
N=512	91,25%	87,5%	81,25%	82,5%	61,25%	72,5%

глухие согласные										
звук	к	п	с	т	Ф	х	ц	ч	ш	щ
N=256	75%	72,5%	47,5%	45%	35%	45%	35%	27,5%	42,5%	32,5%
N=512	27,5%	72,5%	37,5%	40%	38,75%	46,25%	33,75%	27,5%	50%	37,5%

Как видно из табл. 1 и рис. 1, все звуки отличаются долей узких частотных интервалов, которые можно использовать для внедрения:

– глухие согласные: «с», «т», «ф», «х», «ч», «ц», «ш», «щ», гласный – «ю», сонорная согласная – «р», звонкие согласные – «ж», «з» менее сконцентрированы в частотной области следовательно, частотных интервалов, которые можно использовать для внедрения у них меньше;

– сонорные согласные: «й», «л», «м», «г», «д» и гласные: «а», «е», «ё», «и», «о», «у», «ы», «э», «я», звонкие согласные – «б», «в», «г», глухие согласные: «к», «п» обладают наибольшей концентрацией энергии, следовательно, в частотные интервалы речевых данных, соответствующих этим звукам, можно поместить больше информации.

Среднее значение доступных узких частотных интервалов для звуков русской речи равно 67%-70%, если принять вероятность появления звуков одинаковой. В случае если использовать вероятности появления фонем приведенных в табл.2 [3], тогда среднее значение доступных узких частотных интервалов для звуков русской речи составит 67%-76%.

Таблица 2

Вероятности появления звуков русской речи (p_i)

Гласные										
звук	а	Е	ё	и	о	у	ы	э	ю	я
	0,183	0,096	0,01	0,023	0,061	0,014	0,016	0,032	0,011	0,01
сонорные согласные										
звук	Й	л		м		н		р		
	0,043	0,035		0,023		0,06		0,035		
звонкие согласные										
звук	Б	в		Г		д		ж	з	
	0,005	0,032		0,009		0,03		0,009	0,012	
глухие согласные										
звук	к	П	с	т	ф	х	ц	ч	ш	щ
	0,034	0,025	0,04	0,073	0,009	0,011	0,02	0,005	0,021	0,023

Выводы.

Сравнение результатов в табл.1 показывает, что увеличение ширины окна анализа речевых данных с $N = 256$ отсчетов до $N = 512$ отсчетов (при учете выражения8) дает не существенный прирост в количестве частотных интервалов которые можно использовать для внедрения. Но при этом изменяется маска согласно которой происходит внедрение. Хотелось бы отметить, что ширина частотного интервала уменьшается при изменении ширины окна с $N = 256$ отсчетов до $N = 512$ отсчетов. Следовательно, использование более широких окон во временной области и узких в частотных интервалов позволяют лучше учитывать свойства распределения долей энергии по частотным интервалам для различных окон речевых данных.

Работа выполнена при поддержке программы РФФИ проект №12-07-00514А.

Список литературы

1. Жиляков, Е.Г. Вариационные методы анализа сигналов на основе частотных представлений [Текст] / Е.Г. Жиляков, С.П. Белов, А.А. Черноморец// Вопросы радиоэлектроники, сер. ЭВТ, вып.1. – Москва: Изд-во ОАО «ЦНИИ «Электроника», 2010. – 185 с.

2. Гантмахер, Ф.Р. Теория матриц [Текст] / Ф.Р. Гантмахер. – М.: Физматлит, 2004. – 560 с.

3. Величкин А.И. Передача аналоговых сообщений по цифровым каналам связи. – М.: Радио и связь, 1983. – 240 с.

4. Жиляков, Е.Г. Вариационные методы анализа и построения функций по эмпирическим данным [Текст]/ Е.Г. Жиляков. – Белгород: Изд-во БелГУ, 2007. – 160 с.

DEFINITION POSSIBLE VOLUME INTRODUCES INFORMATION IN SECURE COMMUNICATION TAGS IN VOIGE DATA

E.G. ZHILYAKOV
S.N. DEVITSINA
P.G. LIKHOLOB

Belgorod National Research University

e-mail:
Zhilyakov@bsu.edu.ru
Devitsyna@bsu.edu.ru
Likholob@bsu.edu.ru

The paper evaluated the available space for the secure communication of information in the speech data. Algorithms for automated search of space-consistent implementation of tags based on the regularities of the distribution of energy in the speech data. Search space by introducing labels of frequency intervals, the energy of which is not included in the 0,90-0,99 energy. Empirically for each sound of Russian speech maximum amount of information being implemented under given conditions.

Keywords: digital watermark, the label, the label of the energy, the frequency interval, steganography, voice data, secure communication of information.