



# ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫЕ ТЕХНОЛОГИИ

УДК 004.415.24

## ИССЛЕДОВАНИЕ ЧУВСТВИТЕЛЬНОСТИ НЕКОТОРЫХ МЕР КАЧЕСТВА СКРЫТИЯ ИНФОРМАЦИИ В РЕЧЕВЫХ СИГНАЛАХ

## RESEARCH OF SENSITIVITY OF SOME MEASURES QUALITY ASSESSMENT HIDDEN INFORMATION IN THE SPEECH SIGNAL

**Е.Г. Жиляков, П.Г. Лихолоб, А.А. Медведева, Е.И. Прохоренко**  
**E.G. Zhilyakov, P.G. Likholob, A.A. Medvedeva, E.I. Prokhorenko**

*Белгородский государственный национальный исследовательский университет,  
308015 Россия, г. Белгород, ул. Победы, 85*

*Belgorod State National Research University, 308015 Russia, Belgorod, Pobedy, 85*

*Zhilyakov@bsu.edu.ru, Likholob@bsu.edu.ru, Medvedeva\_aa@bsu.edu.ru, Prokhorenko@bsu.edu.ru*

*Аннотация.* В работе представлено сравнение некоторых мер степени различия между исходным сигналом и результатом внедрения дополнительной информации. Проведено сравнение чувствительности различных мер. Сравнение проводилось на основе анализа результатов внедрения на основе стеганографического метода расширения спектра. В работе представлены результаты сравнения некоторых мер различия на основе анализа речевых сигналов при разбиении на отрезки одинаковой длительности.

*Resume.* The paper presents a comparison of some measures the degree of difference between the original signal and the result of the introduction of additional information. A comparison of the sensitivity of the various measures. A comparison was based on the analysis of the implementation of results-based steganographic method of spreading. The paper presents the results of a comparison of some measures the difference based on the analysis of speech signals in the division into segments of equal length.

*Ключевые слова:* речевые сигналы, стеганография, меры различия, коэффициент корреляции, среднеквадратическая ошибка, относительная погрешность, отношение сигнал-шум, мера расстояния Итакуры-Санто, метод расширения спектра.

*Keywords:* speech signals, steganography, measures the differences, the correlation coefficient, the mean square error, relative error, the signal-to-noise ratio, a measure of distance Itakura Santo method of spreading.

Развитие современных информационно-телекоммуникационных систем направлено на обеспечение возможности предоставления естественных для человека форм информационного обмена. Одной из таких форм, наиболее часто используемых удобных для человека, является речь. Современные информационные системы позволяют осуществлять хранение и передачу речевых сообщений на расстояние. Обеспечение такой возможности привело к бурному развитию технологий, обеспечивающих внедрение в аудиозаписи дополнительной информации, которая не будет восприниматься органами чувств человека. Это могут быть метки даты и времени, метки, подтверждающие авторское право и т.д. Внедрение дополнительной информации таким образом, чтобы сам факт внедрения не был обнаружен, занимается стеганография. Именно этот аспект и описывает основной принцип стеганографии [Грибунин и др., 2002; Жиляков, 2015; Конахович, Пузыренко, 2006; Крыжевич, Белобородов, 2014; Fugui, 2000; Zhilyakov E.G. 2015].

В случае использования в качестве объекта, в который будет внедряться информация (контейнера), речевого сигнала, результат внедрения, т.е. стегоконтейнер (контейнер вместе с внедренной информацией), «на слух» не должен отличаться от исходного контейнера.

Очевидно, что наиболее эффективными способами обнаружения изменения (выявления степени изменения) являются субъективные оценки. Однако рост спроса на стегоалгоритмы и,

как следствие, увеличение объемов обрабатываемых речевых данных приводит к необходимости автоматизации процесса оценки результатов внедрения дополнительной информации.

Для этого необходимо использование объективных методов, позволяющих в некоторой числовой форме оценить степень различия речевых сигналов до и после внедрения дополнительной информации.

При этом для методов, оценивающих качество вложения можно предъявить следующие требования:

- метод должен позволить выразить качество звучания количественной мерой;
- метод должен учитывать свойства слухового восприятия;
- при использовании методов не должно возникать необходимости использования экспертов,

но при этом необходимо, чтобы он обеспечивал лучшую корреляцию с субъективными оценками;

– метод должен позволять определять критическую отметку (порог обнаружения) по достижении которой, изменения, вызываемые стеганографическим методом кодирования, будут заметны на слух;

– метод не должен зависеть от параметров анализируемого сигнала (частоты дискретизации, разрядности и т.д.), он должен одинаково реагировать на изменения во временной и частотных областях.

В настоящее время наиболее широкое использование получили методы оценки различия сравниваемых сигналов, основанных на анализе отрезков речевых сигналов во временной области. При этом используются такие оценки различия, как среднеквадратическая ошибка (*CKO*), относительная погрешность (*HCKO*), отношение сигнал-шум (*ОСШ*), коэффициент корреляции (*cor*), мера расстояния Итакуры-Санто (расстояние наибольшего правдоподобия, *ISD*). Каждая из этих оценок позволяет выявить различия в сравниваемых сигналах. Однако они имеют разную чувствительность.

В частности, среднеквадратическая ошибка (*CKO*) отражает абсолютное различие энергии отрезков сигналов во временной области [Ozer H., 2000; Hicsonmez S., 2013; Iser B., 2008]:

$$CKO = \sum_{n=1}^N (x_n - \tilde{x}_n)^2, \quad (1)$$

где  $x_n$  – значение амплитуды исходного отрезка данных;  $\tilde{x}_n$  – значение амплитуды отрезка данных содержащего дополнительную информацию,  $N$  – количество отсчетов сравниваемых отрезков сигналов.

Данная мера позволяет выявить различия в огибающих амплитуд отрезков речевых сигналов. Чем меньше изменений вносится при внедрении дополнительной информации, тем ближе значение этой оценки к нулю.

Однако данная оценка не учитывает энергию самого сигнала, а это значит, что при выборе данной оценки возникают сложности с выбором порогового значения. Поэтому чаще используют нормированную оценку *CKO* к норме исходного сигнала [Жилияков Е.Г., 2010]:

$$HCKO = \sum_{n=1}^N (x_n - \tilde{x}_n)^2 / \sum_{n=1}^N x_n^2. \quad (2)$$

Реакция данной оценки аналогична реакции *CKO*.

Также для учета степени отличия исходного сигнала и результата внедрения дополнительной информации используют оценку, чувствительную ко времени выравнивания сравниваемых отрезков сигналов [Ozer H., 2000; Hicsonmez S., 2013; Iser B., 2008]:

$$ОСШ = 10 \cdot \lg \frac{\sum_{n=1}^N x_n^2}{\sum_{n=1}^N (x_n - \tilde{x}_n)^2}. \quad (3)$$

Чем выше оценка *ОСШ*, тем меньше изменений было внесено. В случае равенства двух отрезков (исходного и подвергнутого изменениям при кодировании) оценка будет равна бесконечности ( $\infty$ ).

Для оценки степени схожести двух отрезков данных, часто используют оценку взаимной энергии этих сигналов, определяемую коэффициентом корреляции [Ozer H., 2000; Iser B., 2008]:

$$cor = \frac{\sum_{n=1}^N \left( x_n - \frac{1}{N} \sum_{n=1}^N x_n \right) \cdot \left( \tilde{x}_n - \frac{1}{N} \sum_{n=1}^N \tilde{x}_n \right)}{\sqrt{\sum_{n=1}^N \left( x_n - \frac{1}{N} \sum_{n=1}^N x_n \right)^2 \cdot \left( \tilde{x}_n - \frac{1}{N} \sum_{n=1}^N \tilde{x}_n \right)^2}}. \quad (4)$$

Чем ближе значение корреляции к единице, тем выше схожесть отрезка данных содержащего контрольную информацию и исходного.

Все рассмотренные выше оценки вычисляют меру различия, используя для сравнения значения отсчетов во временной области. Однако на ряду с изменениями во временной области необходимо также учитывать различия в частотной области. Для этого используется мера, основанная на расстоянии Итакуры-Саито [Ozer H., 2000; Hicsonmez S., 2013; Iser B., 2008]:



$$ISD = \int_{-\pi}^{\pi} \left( \frac{|\tilde{X}(v)|^2}{|X(v)|^2} + \ln \frac{|X(v)|^2}{|\tilde{X}(v)|^2} - 1 \right) dv / 2\pi. \quad (5)$$

Известно, что энергия отрезка сигнала может быть выражена следующим образом [Жиляков, 2010; Жиляков, 2015]:

$$\|\tilde{x}\|^2 = \sum_{n=1}^N x_n^2 = \sum_{r=1}^R \left( \int_{v \in \Omega_r} |X(v)|^2 dv / 2\pi \right) = \sum_{r=1}^R P_r, \quad (6)$$

где  $P_r$  – значение энергии частотной компоненты отрезка сигнала.

Тогда мера, основанная на расстоянии Итакуры-Саито, может быть представлена в виде:

$$ISD = \sum_{r=1}^R \Delta \omega_r \cdot \left( \frac{\tilde{P}_r}{P_r} + \ln \frac{P_r}{\tilde{P}_r} - 1 \right) / \pi, \quad (7)$$

где  $\tilde{P}_r$  – значение энергии частотной компоненты исходного отрезка данных;  $P_r$  – значение энергии частотной компоненты отрезка данных содержащего дополнительную информацию.

Мера имеет смысл расстояния между спектрами двух сигналов и оценивает несоответствие между энергией измененного и исходного отрезка данных. При равенстве отрезков данных мера обращается в ноль.

Сравнение чувствительности представленных оценок осуществлялось на основе использования одного из наиболее распространенных стеганографических методов [Hicsonmez S., 2013], учитывающих частотные характеристики речевых сигналов, – метода расширения спектра. Суть метода заключается в добавлении к отрезку исходного речевого сигнала псевдослучайной последовательности (ПСП) в соответствии с выражением [Жарких и др., 2009; Fridrich, 2012; Nedeljko, Tapio, 2004; Stanković et al, 2012; Белов С.П., 2015]:

$$\tilde{x} = \bar{x} + \alpha_m \cdot e_m \cdot \bar{u}, \quad (8)$$

где  $\bar{x}$  – исходный отрезок данных;  $\bar{u}$  – отрезок, соответствующий псевдослучайной последовательности;  $\alpha_m$  – весовой коэффициент;  $e_m$  – кодовое отображение двоичного бита контрольной информации, определяемое по формуле:

$$e_m = 2e_m - 1, \quad m = 1, \dots, M, \quad (9)$$

где  $e_m$  – бит контрольной информации в двоичной системе счисления,  $e_m \in \{0, 1\}$ ;  $M$  – объем скрытно кодируемой контрольной информации;  $e_m$  – кодовое отображение двоичного бита контрольной информации,  $e_m \in \{-1, 1\}$ ;  $m$  – порядковый номер бита контрольной информации.

Весовой коэффициент  $\alpha_m$  определяет скрытность системы. В работах [Thierry, Ferran, 2009; Vercoe, 1995] его предлагается выбирать равным:

$$\alpha_m = \frac{\langle \bar{x}, \bar{u} \rangle}{\|\bar{u}\|^2}. \quad (10)$$

Стоит отметить, что использование в качестве шума сигнальной конструкции  $\bar{u}$ , не обладающей взаимной энергией с данными  $\bar{x}$ , позволяет повысить помехоустойчивость стеганографически закодированной контрольной информации  $e_m$ , а использование коэффициента проекции  $\alpha_m$  повышает скрытность контрольной информации.

Декодирование бита контрольной информации из данных происходит путем определения знака скалярного произведения отрезка данных и псевдослучайной последовательности:

$$\tilde{e}_m = \text{sign}(\langle \tilde{x}, \bar{u} \rangle), \quad (11)$$

где  $\text{sign}(\ )$  – операция выделения знака.

Для исследования чувствительности рассмотренных мер качества скрытия информации, были проведены вычислительные эксперименты с использованием различных звуков русской речи. На рисунках 1-3 представлены отрезки речевых сигналов, соответствующих некоторым звукам русской речи, а также распределение их энергии по частотным интервалам.

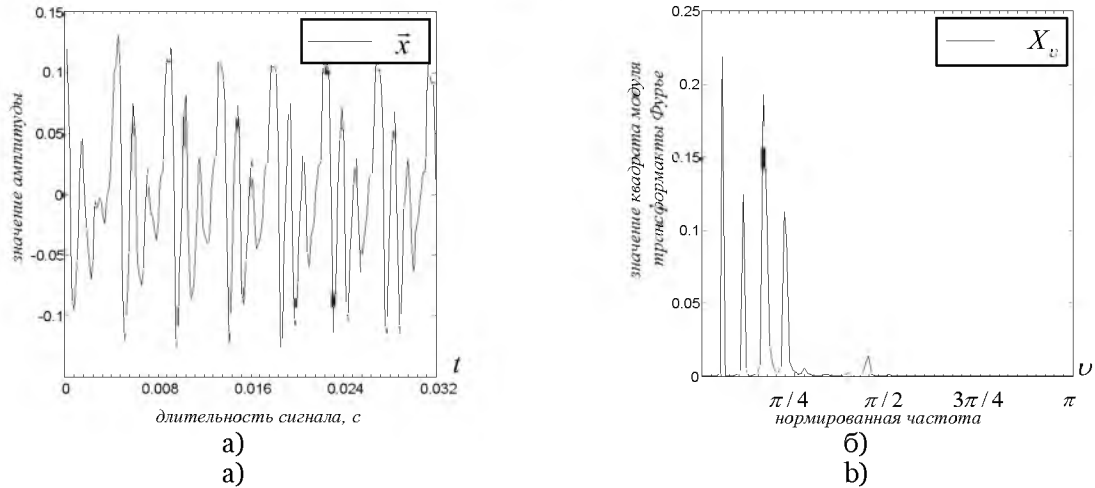


Рис. 1. Речевой сигнал, соответствующий звуку «а»: а) отрезок речевого сигнала; б) распределение энергии по частотным интервалам  
 Fig. 1. A speech signal corresponding to the sound «a»: a) the segment of the speech signal; b) the distribution of energy over the frequency interval

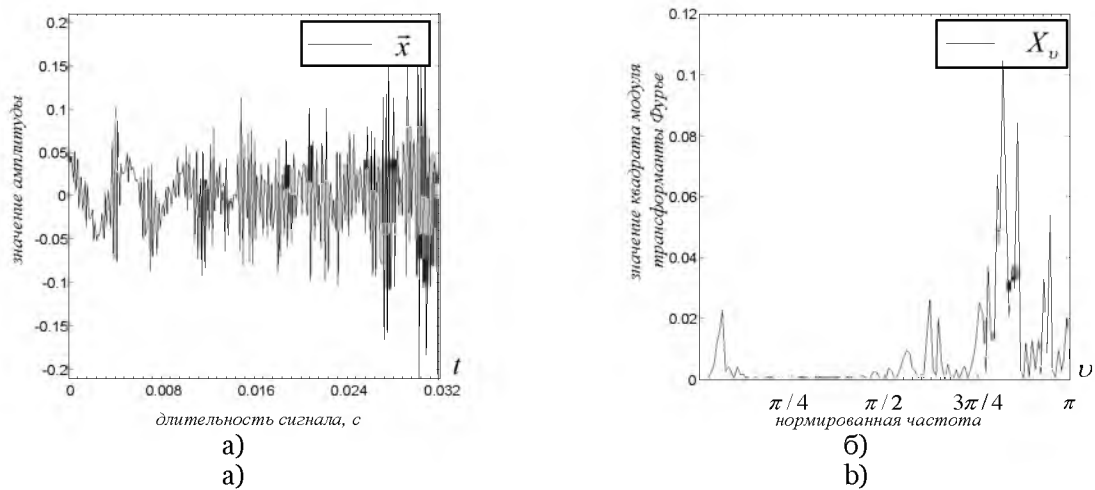


Рис. 2. Речевой сигнал, соответствующий звуку «ch»: а) отрезок речевого сигнала; б) распределение энергии по частотным интервалам  
 Fig. 2. A speech signal corresponding to the sound «ch»: a) the segment of the speech signal; b) the distribution of energy over the frequency interval

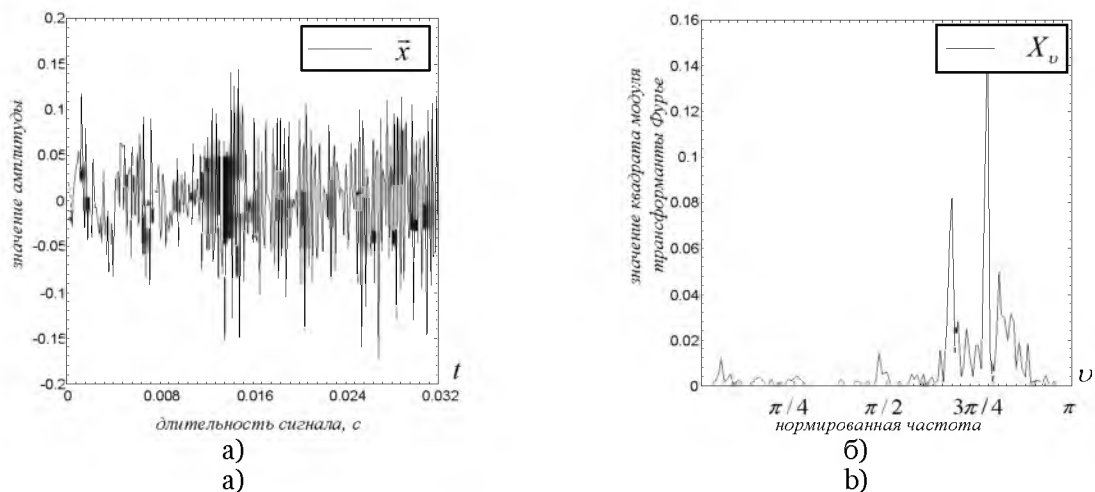


Рис. 3. Речевой сигнал, соответствующий звуку «sh»: а) отрезок речевого сигнала; б) распределение энергии по частотным интервалам  
 Fig. 3. A speech signal corresponding to the sound «sh»: a) the segment of the speech signal; b) the distribution of energy over the frequency interval



Как видно из рисунков, различные звуки русской речи имеют разное распределение энергии. Так, энергия речевого сигнала, соответствующего звуку «а», сосредоточена в полосе нижних частот, а энергия отрезков речевых сигналов, соответствующих звукам «ч» и «ш», сосредоточена в области высоких частот. Кроме того, энергия отрезка сигнала, соответствующего звуку «ш», распределена в более широкой полосе частот. Так как разные звуки речи имеют различное распределение энергии в частотной области, то внедрение дополнительной информации в различные звуки может приводить к различным искажениям. Таким образом, используемые меры качества скрытия информации должны учитывать особенности частотного распределения отрезков речевых сигналов.

В таблице 1 представлены результаты оценки рассмотренных мер различия для всех звуков русской речи. При этом для анализа использовались отрезки речевых сигналов, записанных с частотой дискретизации 8 кГц и разрядностью кода 16 бит. Для реализации метода расширения спектра речевые сигналы разбивались на отрезки одинаковой длительности по  $T=32$ мс. Важно также отметить, что исследование рассматриваемых мер осуществлялось при реализации наложения шума на сигнал при отсутствии взаимной корреляции и использовании весового коэффициента вида:

$$\alpha_m = \frac{K_m}{\|z\|^2} \cdot \quad (12)$$

При этом параметр  $K_m$  изменялся в диапазоне от 0,0001 до 0,2000.

Таблица 1  
Table 1

**Оценка мер различия исходного сигнала и результатов внедрения при использовании стеганографического метода расширения спектра ( $T=32$ мс)**  
**Evaluation of the differences of the original signal and implementation results using steganographic technique spreading ( $T=32$ ms)**

Тип звука	$K_m$	<i>СКО</i>	<i>НСКО</i>	<i>ОСШ</i>	<i>cor</i>	<i>ISD</i>
А	0,0001	0,0001	0,0001	80,0000	1,0000	0,0021
	0,0002	0,0002	0,0002	73,9794	0,9999	0,0045
	0,0100	0,0100	0,0100	40,0000	0,9950	0,4529
	0,1000	0,1000	0,1000	20,0000	0,9524	6,3492
	0,2000	0,2000	0,2000	13,9794	0,9091	13,3037
Ч	0,0001	0,0001	0,0001	80,0000	1,0000	0,0002
	0,0002	0,0002	0,0002	73,9794	0,9999	0,0005
	0,0100	0,0100	0,0100	40,0000	0,9950	0,0182
	0,1000	0,1000	0,1000	20,0000	0,9524	0,3009
	0,2000	0,2000	0,2000	13,9794	0,9091	0,8142
Ш	0,0001	0,0001	0,0001	80,0000	1,0000	0,0007
	0,0002	0,0002	0,0002	73,9794	0,9999	0,0014
	0,0100	0,0100	0,0100	40,0000	0,9950	0,0523
	0,1000	0,1000	0,1000	20,0000	0,9524	0,6402
	0,2000	0,2000	0,2000	13,9794	0,9091	1,5429

Из представленных в таблице данных видно, что значения всех рассмотренных мер, кроме меры, основанной на расстоянии Итакуры-Саито, зависят только от коэффициента  $K_m$ . В свою очередь, значение меры, основанной на расстоянии Итакуры-Саито, зависит как от коэффициента  $K_m$ , так и от типа звука. Так для вокализованных звуков русской речи добавление широкополосного шума приводит к более существенному увеличению меры, основанной на расстоянии Итакуры-Саито, чем при добавлении того же фрагмента шума к шипящим звукам. Таким образом, мера, основанная на расстоянии Итакуры-Саито, позволяет учитывать особенности распределения энергии звуков русской речи.

### Выводы

Таким образом, использование меры, основанной на расстоянии Итакуры-Саито, целесообразно для оценки меры качества скрытия информации в речевых сигналах.

Работа выполнена при поддержке грантов РФФИ № 15-07-01463 «Разработка методов и алгоритмов автоматического распознавания устной речи с использованием субполосного анализа речевых сигналов».

**Список литературы  
References**

- Грибунин В.Г., Оков И.Н., Туринцев И.В. 2002. Цифровая стеганография. Аспекты защиты. М., Солон-Пресс. С.261.
- Белов С.П., Жилияков Е.Г., Лихолоб П.Г., Пашинцев В.П. 2015. О методе скрытного кодирования контрольной информации в речевые данные. Инфокоммуникационные технологии. Саратов. Поволжский государственный университет телекоммуникаций и информатики, 13(3): 325-333.
- Belov S.P., Zhilyakov E.G., Liholob P.G. Pashincev V.P. 2015. O metode skrytnogo kodirovaniya kontrolnoj informacii v rechevye dannye. Infokommunikacionnye tekhnologii. Saratov. Povolzhskij gosudarstvennyj universitet telekommunikacij i informatiki. 13(3): 325-333.
- Gribunin V.G., Okov I.N., Turintsev I.V. 2002. Tsifrovaya steganografiya. Aspekty zaschity. M., Solon-Press: 1-261.
- Жарких А.А., Гурин А.В., Пластунов В.Ю. 2009. Метод стеганографии на основе прямого расширения спектра сигнала. Материалы VII Международной научно-технической конференции, 7 – 11 декабря 2009 г. INTERMATIC. - М.: МИРЭА часть 4: 78-83.
- Zharkih A.A., Gurin A.V., Plastunov V.Yu. 2009. Metod steganografii na osnove pryamogo rasshireniya spektra signala. Materialy VII Mezhdunarodnoy nauchno-tehnicheskoy konferentsii, 7 – 11 dekabrya 2009 g. INTERMATIC. - M.: MIREA chast 4: 78-83.
- Жилияков Е.Г., Белов С.П., Черноморец А.А. 2010. Вариационные методы анализа сигналов на основе частотных представлений. Вопросы радиоэлектроники, сер. ЭВТ, вып.1. – Москва: Изд-во ОАО «ЦНИИ «Электроника»: 10-26.
- Zhilyakov E.G., Belov S.P., Chernomorets A.A. 2010. Variatsionnyie metody analiza signalov na osnove chastotnyih predstavleniy. Voprosy radioelektroniki, ser. EVT, vyip.1. – Moskva: Izd-vo ОАО «TsNII «Elektronika»: 10-26.
- Жилияков Е.Г. 2015. Оптимальные субполосные методы анализа и синтеза сигналов конечной длительности. Автоматика и телемеханика. – М.: Академический научно-издательский, производственно-полиграфический и книгораспространительский центр Российской академии наук "Издательство "Наука". 4: 51-66
- Zhilyakov E.G. 2015. Optimalnyie subpolosnyie metody analiza i sinteza signalov konechnoy dlitelnosti. Avtomatika i telemehanika. Moskov. Akademicheskij nauchno-izdatelskiy, proizvodstvenno-poligraficheskij i knigorasprostranitel'skiy tsentr Rossiyskoy akademii nauk "Izdatel'stvo "Nauka"" № 4: 51-66.
- Hicsonmez S., Uzun E., Sencar H.T. 2013. Methods for identifying traces of compression in audio. In Communications, Signal Processing, and their Applications (ICCSIPA), 2013 1st International Conference on IEEE: 1-6.
- Конахович Г.Ф., Пузыренко А.Ю. 2006. Компьютерная стеганография. Теория и практика. Киев, «МК-Пресс»: 1-288.
- Konahovich G.F., Puzyrenko A.Yu. 2006. Kompyuternaya steganografiya. Teoriya i praktika. Kiev, "MK-Press.": 1-288.
- Крыжевич Л.С., Белобородов Д.А. 2014. Стеганографические методы сокрытия данных в звуковых файлах на основе всплесковых преобразований. Auditorium: электронный научный журнал Курского государственного университета. – Курск: № 2, 2014г. «Аудиториум» электронный научный журнал <http://auditorium.kursksu.ru>. URL: <http://auditorium.kursksu.ru/index.php?page=6&new=2>
- Kryzhevich L.S., Beloborodov D.A. Steganograficheskie metody sokryitiya dannyih v zvukovyih faylah na osnove vspleskovyih preobrazovaniy. Auditorium: elektronnyiy nauchnyiy zhurnal Kurskogo gosudarstvennogo universiteta. – Kursk: # 2, 2014g. «Auditorium» elektronnyiy nauchnyiy zhurnal <http://auditorium.kursksu.ru>. URL: <http://auditorium.kursksu.ru/index.php?page=6&new=2>
- Iser B., Schmidt G., Minker, W. 2008. Bandwidth extension of speech signals. Springer Science & Business Media. 13: 1-187
- Fridrich, J. 2012. Steganography in digital media: Principles, algorithms, and applications. Steganography in Digital Media: 1-441.
- Furui, Sadaoki. 2000. Digital speech processing, synthesis, and recognition. 2nd ed., rev. and expanded
- Nedeljko Cvejic, Tapio Seppanen. 2004. Spread spectrum audio watermarking using frequency hopping and attack characterization. Signal Processing 84: 207 – 213.
- Ozer H., Avcibas, I., Sankur, B. and Memon, N.D. 2003. Steganalysis of audio based on audio quality metrics. In Electronic Imaging 2003 International Society for Optics and Photonics: 55-66
- Stanković, S., Orović, I., Sejdić, E. 2012. Multimedia signals and systems. Multimedia Signals and Systems: 1-349.
- Thierry Dutoit, Ferran Marques. 2009. Applied Signal Processing A MATLAB TM-Based Proof of Concept.
- Vercoe B.L. 1995. Csound: A Manual for the Audio-Processing System, MIT Media Lab, Cambridge.
- Zhilyakov E.G. 2015. Optimal subband methods of analysis and synthesis of signals of finite duration / Automation and Remote Control. - M .: Academic Scientific Publishing, Production and Publishing and Bookselling Center of the Russian Academy of Science "Publishing House" Science "№ 4, 2015: 51-66.