

**Леонова Антонина Валерьевна** – Федеральное государственное автономное образовательное учреждение высшего профессионального образования «Южный федеральный университет»; e-mail: antonina\_tsure@mail.ru; 347928, г. Таганрог, пер. Некрасовский, 44; тел.: 89043467991; кафедра электрогидроакустической и медицинской техники; ассистент; к.т.н.

**Лебедева Юлия Сергеевна** – e-mail: lebedeva\_yulia@list.ru; тел.: 89287547999; кафедра автоматизированных систем научных исследований и экспериментов; лаборант НИЧ; студентка.

**Leonova Antonina Valer'evna** – Federal State-Owned Autonomy Educational Establishment of Higher Vocational Education “Southern Federal University”; e-mail: antonina\_tsure@mail.ru; 44, Nekrasovskiy, Taganrog, 347928, Russia; phone: +79043467991; the department of hydroacoustic and medical engineering; assistant; cand. of eng. sc.

**Lebedeva Yulia Sergeevna** – e-mail: lebedeva\_yulia@list.ru; phone: 89287547999; the department of automated research systems; SRU assistant; student.

УДК 004.93

**В.В. Копытов, Д.В. Якушев**

#### **МЕТОДЫ КОДИРОВАНИЯ РЕЧЕВЫХ СИГНАЛОВ С ПОМОЩЬЮ РЕКОНСТРУИРОВАННОЙ МОДЕЛИ РЕЧЕВОГО ПРОЦЕССА**

*Представлена реконструированная модель речевого процесса. Рассмотрена возможность использования различных методов численного интегрирования для восстановления речевого сигнала с целью снижения ошибки реконструкции. Представлены результаты численного эксперимента. Разборчивость речи – лучшая для методов трапеций и Симпсона. Метод Симпсона наиболее трудоемкий в плане вычислительных мощностей и занимает большее количество времени для реконструкции речевого сигнала, в отличие от других методов. Поэтому на практике целесообразно введение дополнительной метрики для оценки разборчивости речи и сочетание представленных методов для получения максимально качественного речевого сигнала. Результаты исследования могут быть использованы для кодирования речевой информации.*

*Кодирование речевой информации; речевой сигнал; реконструкция речевого процесса; ошибки реконструкции; методы численного интегрирования; метод прямоугольников; метод трапеций; метод Симпсона.*

**V.V. Kopytov, D.V. Yakushev**

#### **METHODS OF CODING OF SPEECH SIGNALS WITH THE HELP OF THE RECONSTRUCTED MODEL OF SPEECH PROCESS**

*The article presents renovated model of the speech process. The possibility of using various methods of numerical integration for the restoration of the speech signal in order to reduce errors of reconstruction. The results of the numerical experiment. Speech intelligibility is the best method for trapezoids and Simpson. Method Simpson most time-consuming in terms of computing power, and takes more time for the reconstruction of the speech signal, in contrast to other methods. Therefore, in practice, it is expedient introduction of an additional metrics for evaluation of speech intelligibility, and a combination of the presented methods to obtain the maximum quality of the speech signal. The results of the study can be used for the encoding of the speech information.*

*Coding of speech information; the voice signal; the reconstruction of the speech process; the errors of the reconstruction; methods of numerical integration method of rectangles; the trapezoid method; method Simpson.*

В современных условиях автоматизации информационных процессов возрастает актуальность обеспечения безопасности конфиденциальной речевой информации от несанкционированного доступа. Для обеспечения безопасности передачи информации по каналам связи используются различные технологии преобразования сигналов. Реализация данных подходов связана с кодированием исходного речевого сигнала, т.е. преобразованием каждого сегмента сигнала в набор параметров, предназначенных для последующей обработки. В настоящее время применяются различные методы и алгоритмы преобразования речевых сигналов, основанные на применении цифровых фильтров, принципах обработки сигналов Хургина–Яковлева, технологиях вейвлет-преобразования и технологиях преобразования графических образов узкополосных спектрограмм с последующим синтезом по ним новых речеподобных сигналов.

Таким образом, существующие методы обработки речевых сигналов основываются на преобразовании их спектральных характеристик в частотно-временной области. Задача спектрального оценивания подразумевает оценивание некоторой функции частоты, точность которой зависит от объема выборки отсчетов речевого сигнала. Чем выше объем, тем точнее оценивание. Поэтому использование спектрального преобразования находится в прямой зависимости от предположения стационарности речевого процесса. В этом предположении заключается основной недостаток подобных методов, так как речевой процесс является нестационарным и его параметры неустойчивы относительно условий среды образования и распространения речевых сигналов.

Другим подходом к моделированию речевого сигнала является линейное предсказание. Коэффициенты модели линейного предсказания оцениваются с помощью автокорреляционного метода по предыдущему сигналу. Передаточную функцию  $A(z)$  можно представить как [1]:

$$A(z) = \frac{1}{1 + \sum_{k=1}^p a_k z^{-k}}, \quad (1)$$

где  $a_i$  определяются с помощью анализа речевого сигнала моделью линейного предсказания порядка  $p$ . Сигнал пропускается через фильтр с передаточной функцией  $1/A(z)$ . Коэффициенты линейного предсказания содержат информацию о форме речевого тракта. Основной проблемой данного подхода является неустойчивость модели линейного предсказания к типу микрофона и внешним шумам. Искажения проявляются в виде всплесков на графике производной объемной скорости и объемной скорости после обратной фильтрации, по амплитуде превосходящих пики исследуемых сигналов.

Задачей настоящего исследования является использование альтернативного подхода к преобразованию речевого сигнала с помощью процедуры реконструкции математической модели динамической системы по порождаемому временному ряду [2, 3, 4, 5]. В этом случае коэффициенты  $a_i$  аппроксимации передаточной функции речевого тракта (1) определяются с помощью реконструкции речевого процесса по речевому сигналу. Такой подход учитывает нестационарность и внутреннюю нелинейную динамику речевого тракта и, соответственно, более устойчив к внешним искажениям.

Применим к исследуемому сегменту речевого сигнала процедуру реконструкции математической модели динамической системы по порождаемому временному ряду. Под динамической системой будем понимать речевой процесс, под временным рядом будем понимать речевой сигнал, являющийся реализацией динамической системы.

Реконструируемая модель речевого процесса имеет вид

$$\int_0^t x dt = x_1; \int_0^t x_1 dt = x_2; \int_0^t x_2 dt = x_3; \quad (2)$$

$$x = f(x_1, x_2, x_3) = c_0 + c_1 x_1 + c_2 x_2 + c_3 x_1 x_2 + c_4 x_1^2 + c_5 x_2^2 + c_6 x_1^2 x_2 + c_7 x_1 x_2^2 + c_8 x_1^2 x_2^2.$$

Здесь  $x$  – речевой сигнал, зависящий от времени  $t$ ;  $x_1, x_2$  – возбуждение в голосовой щели (входной процесс),  $c_0, \dots, c_8$  – коэффициенты, содержащие информацию о форме речевого тракта. Выбор порядка и размерности многочлена в (2) продиктован адекватностью представленной модели исследуемой динамической системе.

Точность представленного подхода на практике зависит от используемых в (2) методов численного интегрирования, из которых находят применение следующие:

- ◆ метод прямоугольников;
- ◆ метод трапеций;
- ◆ метод Симпсона.

Для проведения численного эксперимента использовалось произнесение одним и тем же диктором слов: «ДОМ» и «КОТОРОМ». Выбор слов определялся наличием наиболее распространенных в русском языке фонем. Полученный речевой сигнал сглаживался с помощью скользящего среднего и разбивался на сегменты по пересечениям нулевой линии. К каждому из полученных сегментов применялась процедура реконструкции математической модели речевого процесса, для этого оценивались коэффициенты модели (2), которые предназначены для передачи по каналам связи. Затем по значениям коэффициентов восстанавливался исходный речевой сигнал. В качестве ошибки реконструкции использовалась формула среднего квадратического отклонения реконструированного речевого сигнала

$$e = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (z_i - x_i)^2}, \quad (3)$$

где  $n$  – окно реконструкции речевого процесса,  $x$  – исходный речевой сигнал,  $z$  – реконструированный речевой сигнал,  $e$  – ошибка реконструкции.

В табл. 1 представлены результаты эксперимента.

Таблица 1

**Значения ошибок реконструкции речевого процесса для различных методов численного интегрирования**

Методы численного интегрирования, используемые для реконструкции речевого процесса	ДОМ			КОТОРОМ		
	Количество переходов через ноль			Количество переходов через ноль		
	3	5	9	3	5	9
	$e$			$e$		
Метод прямоугольников	0,33	0,29	0,30	0,22	0,14	0,14
Метод трапеций	0,42	0,35	0,34	0,18	0,20	0,18
Метод Симпсона	0,40	0,39	0,34	0,19	0,18	0,19

На рис. 1 и 2 представлены графики исходных и реконструированных речевых сигналов «ДОМ» и «КОТОРОМ» для метода трапеций.

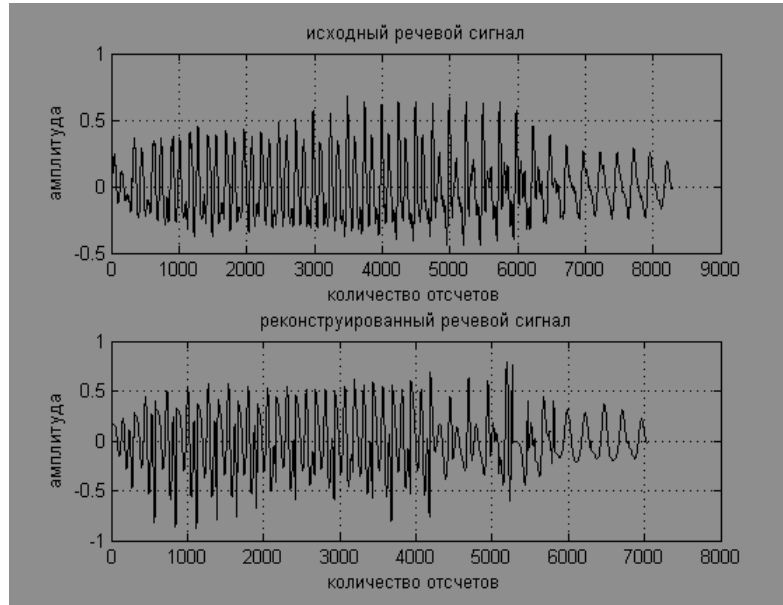


Рис. 1. Исходный и реконструированный речевой сигнал «ДОМ» для метода трапеций

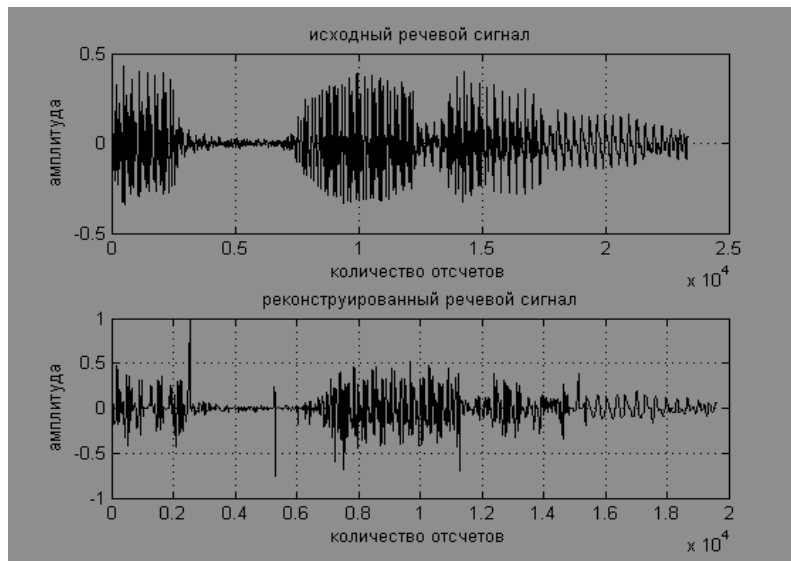


Рис. 2. Исходный и реконструированный речевой сигнал «КОТОРОМ» для метода трапеций

На рис. 3, 4 и 5 представлен пример сегмента исходного и реконструированного речевого сигнала «ДОМ» с оцененными параметрами модели (2) и фазовым пространством для случая использования методов: прямоугольников, метода трапеций и метода Симпсона.

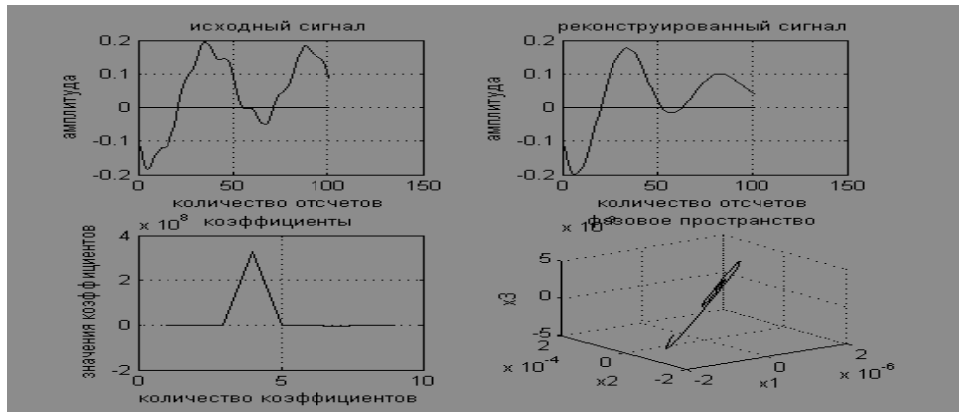


Рис. 3. Графики сегментов исходного и реконструированного речевого сигнала слова «ДОМ» для метода прямоугольников

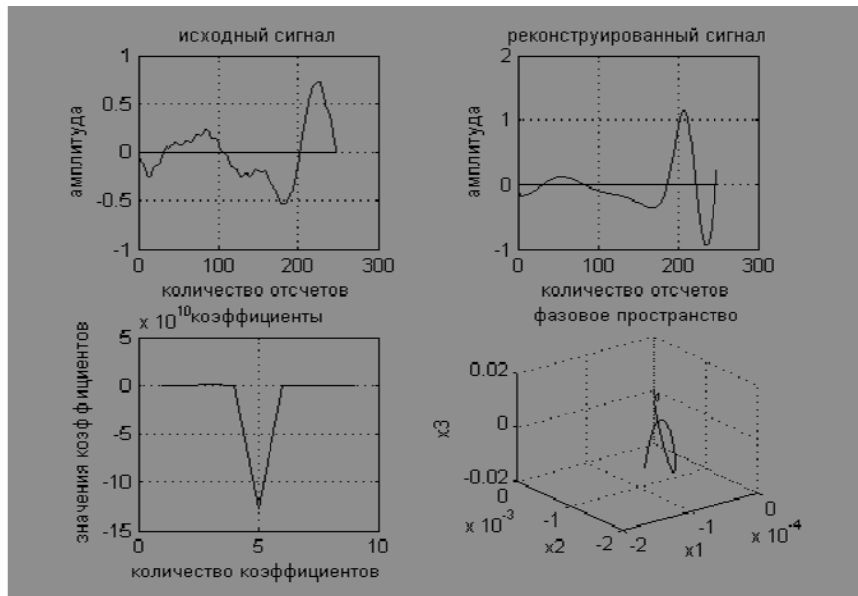


Рис. 4. Графики сегментов исходного и реконструированного речевого сигнала слова «ДОМ» для метода трапеций

На рис. 6, 7 и 8 представлен пример сегмента исходного и реконструированного речевого сигнала «КОТОРОМ» с оцененными параметрами модели (2) и фазовым пространством для случая использования методов: прямоугольников, метода трапеций и метода Симпсона.

Можно сделать вывод, что по результатам численного эксперимента каждый из представленных методов численного интегрирования обладает преимуществами и недостатками. Следует отметить, что, несмотря на величину ошибки реконструкции, разборчивость речи лучше для методов трапеций и Симпсона. Метод Симпсона наиболее трудоемкий в плане вычислительных мощностей и занимает большее количество времени для оценки коэффициентов модели (2) и последующей реконструкции речевого сигнала, в отличие от других методов. Метод прямо-

угольников наиболее перспективен с точки зрения скорости, но резко падает разборчивость восстановленного речевого сигнала. Поэтому на практике целесообразно введение дополнительной метрики для оценки разборчивости речи и сочетание представленных методов для получения максимально качественного речевого сигнала. Также следует отметить «сглаживающий эффект» реконструированных сегментов речевого сигнала, связанный с использованием интегрирования в модели (2), что обеспечивает независимость качества речевого сигнала от условий среды распространения.

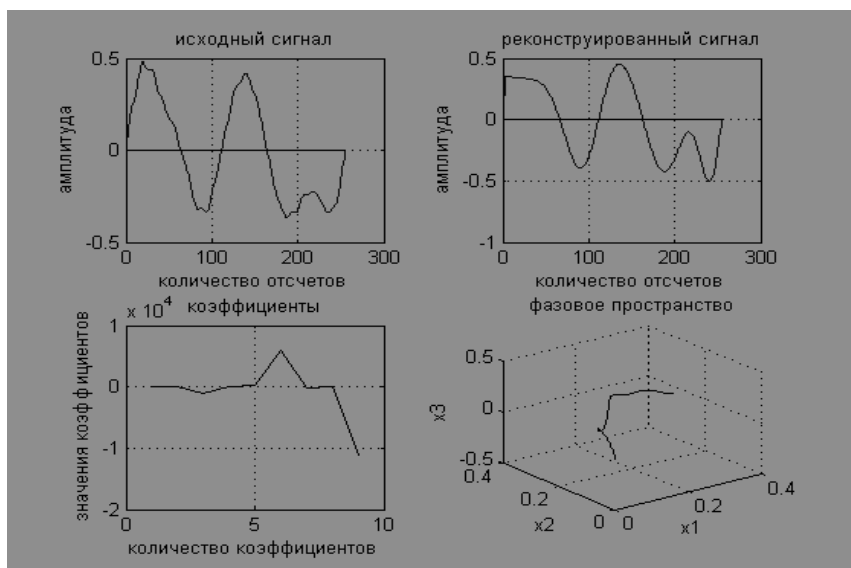


Рис. 5. Графики сегментов исходного и реконструированного речевого сигнала слова «ДОМ» для метода Симпсона

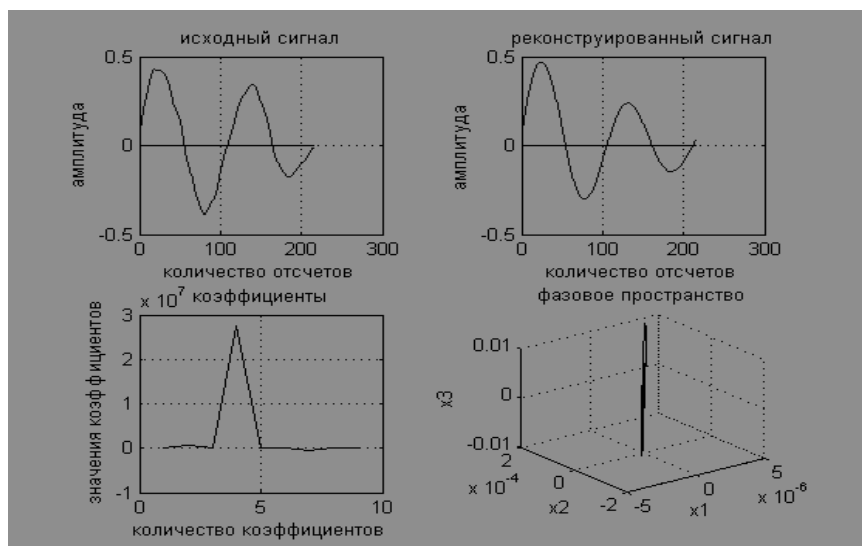


Рис. 6. Графики сегментов исходного и реконструированного речевого сигнала слова «КОТОРОМ» для метода прямоугольников

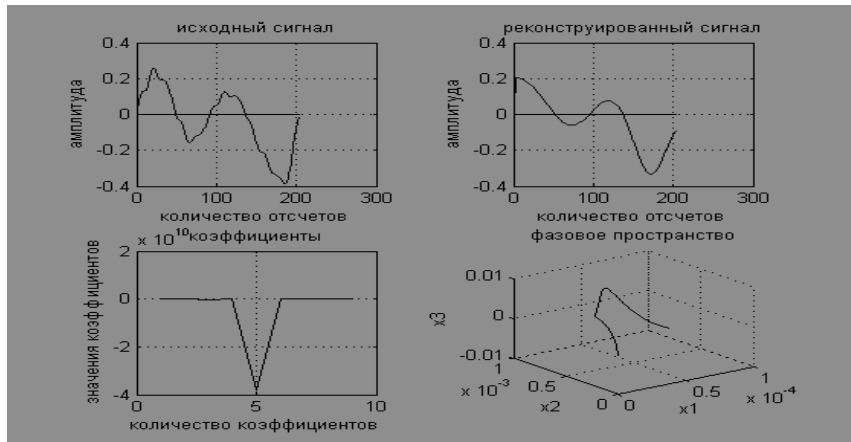


Рис. 7. Графики сегментов исходного и реконструированного речевого сигнала слова «КОТОРОМ» для метода трапеций

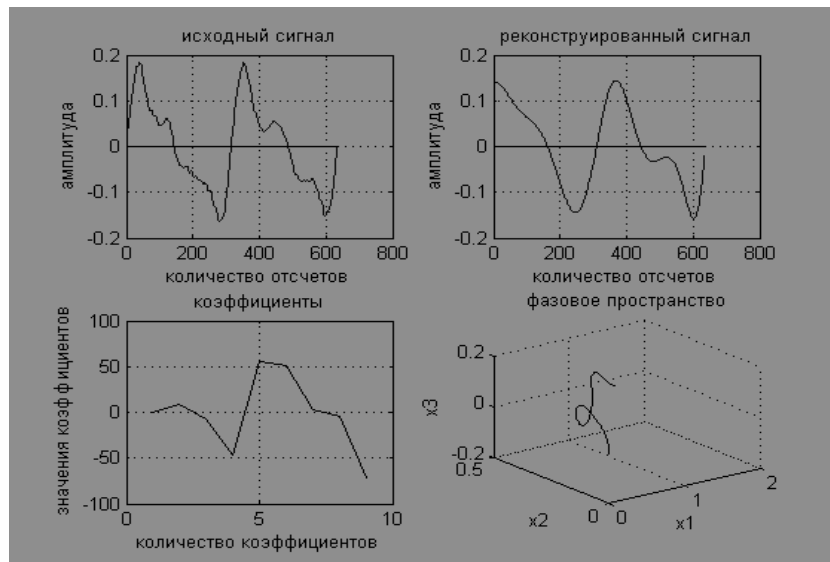


Рис. 8. Графики сегментов исходного и реконструированного речевого сигнала слова «КОТОРОМ» для метода Симпсона

Следует отметить, что для улучшения показателей реконструкции речевого сигнала целесообразно использовать модель с известной структурой, описывающей геометрию речевого тракта и голосового источника, но неизвестными параметрами. Это связано с тем, что многочлен в реконструированной модели (2) частично воспроизводит динамику процесса речеобразования, не учитывая особенностей спектральных характеристик речи на участках «выбросов», связанных с появлением высоких формант голоса по частотной составляющей.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Сорокин В.Н., Макаров И.С. Обратная задача для голосового источника // Информационные процессы. – 2006. – Т. 6, № 4. – С. 375-395.

2. *Анищенко В.С., Астахов В.В., Вадивасова Т.Е., Нейман А.Б., Стрелкова Г.И., Шиманский-Гайер Л.* Нелинейные эффекты в хаотических и стохастических системах. – М.,–Ижевск: Институт компьютерных исследований, 2003. – 544 с.
3. *Безручко Б.П., Смирнов Д.А.* Математическое моделирование и хаотические временные ряды. – Саратов: ГосУНЦ «Колледж», 2005. – 320 с.
4. *Анищенко В.С., Астахов В.В., Вадивасова Т.Е., Нейман А.Б., Стрелкова Г.И., Шиманский-Гайер Л.* Нелинейные эффекты в хаотических и стохастических системах. – М.,–Ижевск: Институт компьютерных исследований, 2003. – 529 с.
5. *Малинецкий Г.Г., Потапов А.Б.* Современные проблемы нелинейной динамики. – 2-е изд., исправл. и доп. – М.: Едиториал УРСС, 2002. – 360 с.

Статью рекомендовал к опубликованию д.т.н., профессор В.Д. Ковалев.

**Копытов Владимир Вячеславович** – ФГАОУ ВПО «Северо-Кавказский федеральный университет»; e-mail: kopytov@stavsu.ru; 355004, г. Ставрополь, ул. Лермонтова, 103, кв. 80; тел.: 89187513343; кафедра организации и технологии защиты информации; д.т.н.; профессор.

**Якушев Дмитрий Владимирович** – e-mail: botashev32@mail.ru; 369204, КЧР, г. Карачаевск, ул. Ленина, 56, корпус 1, кв. 141; тел.: 89187609005; кафедра организации и технологии защиты информации; доцент; к.т.н.

**Kopytov Vladimir Vyacheslavovich** – North-Caucasian Federal University; e-mail: kopytov@stavsu.ru; 103, Lermontov street, kv. 80, 355004, Stavropol, Russia; phone: +79187513343; the department of organization and technology of information protection; dr. of eng. sc.; professor.

**Yakushev Dmitry Vladimirovich** – North-Caucasian Federal University; e-mail: botashev32@mail.ru; 56, Lenina street, building 1, kv. 141, Karachaevsk, KChR, 369204, Russia; phone: +79187609005; the department of organization and technology of information protection; cand. of eng. sc.; associate professor.