Труды научно-технической конференции кластера пензенских предприятий,
обеспечивающих **БЕЗОПАСНОСТЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ.**
Том 10, Пенза-2016, с. 66-73 (http://пниэи.рф/activity/science/BIT/T10-p66.pdf)

УДК: 004.032.26

Афанасьев А.А. (г. Орел)

Непрерывная аутентификация легитимного пользователя сети связи, опирающаяся на использование личностных особенностей низкоскоростного кодирования речевых данных

В настоящее время при организации доступа к ресурсу сети связи используются различные механизмы подтверждения легитимности абонента, основанные на биометрической аутентификации [1]. Одним из самых простых и доступных является речевая аутентификация, которая совместно с другими методами позволяет гарантированно осуществлять использование ресурсов сети связи ограниченному кругу лиц [2, 3]. Основным недостатком использования подобных методов является однократная проверка легитимности пользователя с дальнейшим предоставлением ему прав использования ресурсов сети на весь сеанс связи.

Для его устранения предлагается метод непрерывной аутентификации диктора, при этом процедура аутентификации становится связанной непосредственно с процессом ведения телефонных переговоров по сети связи. В основе данного метода лежит использование низкоскоростного кодирования речевого сигнала (РС) с применением линейного предсказания с возбуждением от кода при использовании алгоритмов векторного квантования параметров липредера и процедуры анализа через синтез [4]. Подавляющее большинство алгоритмов низкоскоростного кодирования РС в своей основе используют метод линейного предсказания [5].

Предлагаемый в данной статье подход базируется на использовании взаимозависимостей элементов декомпозиции РС возникающих при применении процедуры линейного предсказания с возбуждением от кода при векторном квантовании данных, данный класс алгоритмов относится к CELP-подобным. Таким образом, предлагаемые технические решения базируются на существующих технологиях обработки РС и могут быть использованы без изменения элементной базы построения подобных устройств.

При изучении распределения информационных ресурсов в кадрах соответствующих липредеров (табл.1), можно заметить, что основная информационная составляющая передачи расходуется на представление сигнала возбуждения.

Таблица 1

Соотношение информационных ресурсов в кодеках РС

Тип липредера	CS-ACELP	RPE-LTP-LPC	Усредненные значения
Стандарт	G.729	GSM	
$V_K, кбит/с$	8	13	
Коэффициенты модели	0,23	0,14	0,19
Сигнал возбуждения	0,77	0,86	0,81

В первом приближении можно рассматривать любой вокодер как некоторый архиватор речи, работающий в реальном масштабе времени "на лету". Низкоскоростные вокодеры в этом контексте являются архиваторами с высоким уровнем уплотнения информации. Идентификация сигнала возбуждения для них является примерно в 4 раза более ресурсоемкой задачей в сравнении с более простой задачей идентификации и передачи параметров модели. Чем выше коэффициент сжатия речевой информации, тем больше внимания разработчики вокодеров вынуждены уделять идентификации сигнала возбуждения и синтезу его архива в виде векторов кодовой книги.

Для уменьшения информационной избыточности представления данных при их передаче по каналу связи используются статистические зависимости, возникающие между элементами декомпозиции РС при его обработке на основе алгоритмов CELP и использовании положений кластерного анализа при обработке данных.

Так, теоретическим обоснованием метода линейного предсказания является авторегрессионная модель, используемая при цифровом спектральном анализе и предполагающая бесконечный порядок формирующей системы при возбуждении ее сигналом в виде дискретного белого гауссовского шума. Решение системы алгебраических матричных уравнений Юла-Уокера дает возможность определить ее параметры и структуру [6]. Таким образом, возбуждение формирующего фильтра авторегрессионной модели цифрового спектрального анализа осуществляется сигналом $u(n)$ в виде дискретных квантованных значений белого шума с математическим ожиданием равным нулю и единичной дисперсией.

$$\begin{cases} M\{u(n)\} = 0, \\ D\{u(n)\} = \sigma^2\{u(n)\} = 1. \end{cases} \qquad (1)$$

Повышение порядка модели M улучшает точность описания исследуемого процесса. Критерием вычисления параметров модели в предположении о сходимости к гауссовскому закону распределения исходного процесса является взвешенная среднеквадратическая ошибка [6].

$$e^2(n) = d_2(\vec{S}, \vec{S}') = \frac{1}{N}(\vec{S} - \vec{S}')^T(\vec{S} - \vec{S}') == \frac{1}{N}\sum_{i=1}^{N}(\vec{S} - \vec{S}')^2, \qquad (2)$$

где $\vec{S}$ - вектор оригинального РС, $\vec{S}'$- вектор синтезированного РС, N- количество отсчетов на сегменте анализа.

В классической постановке задачи параметрического цифрового спектрального анализа на основе авторегрессионной модели – линейное разностное уравнение формирующего фильтра выглядит следующим образом:

$$y(nT) = \sum_{m=1}^{M} a_m y(nT - mT) + u(nT), \qquad (3)$$

где $y(nT)$ – выходной сигнал, T – интервал дискретизации, a_m- коэффициенты фильтра.

Линейное предсказание РС также использует подобную модель, повышение порядка передаточных функций фильтров анализа и синтеза в ней приводит к "обелению" сигнала остатка предсказания, который является наилучшим сигналом возбуждения.

Работа фильтра синтеза модели линейного предсказания в идеальном случае описывается следующим выражением

$$S'(n) = \sum_{i=1}^{M} a_i S(n-i) + e(n), \qquad (4)$$

где *S'(n)* – предсказанное значение РС; a_i – весовой коэффициент или коэффициент линейного предсказания; *M* – число коэффициентов или порядок линейного предсказания, *e*(*n*) – ошибка предсказания.

Повышение порядка модели в выражениях (4) приводит к получению более точных оценок относительно анализируемого сигнала *S*(*n*) . В идеале $e(n) \to u(n)$ при $M \to \infty$.

Сравнивая выражения (3) и (4) видно, что при фиксированном порядке *M* сигнал возбуждения *e*(*n*) не является реализацией белого шума и определяется как разность между реальным и предсказанным значениями РС.

Ограниченный порядок модели линейного предсказания предопределяет появление данных зависимостей при дальнейшей кластеризации сигнальных пространств сигнала возбуждения и параметров формирующей системы на основе векторного квантования, более углубленное обоснование появления данных зависимостей представлено в [7].

Векторное квантование является одним из методов эффективного кодирования РС с потерями, при котором подразумевается использование кодовой книги. Кодовой книгой называется хранилище конечного количества опорных векторов (опорными назовем те вектора, которые будут записаны в кодовой книге после ее обучения и непосредственно примут участие в кодировании сигналов при использовании ее по назначению), каждый из которых имеет свой порядковый номер. Процедуры создания кодовых книг и векторного квантования более подробно раскрыты в [8].

Таким образом, с каждым из векторов кодовой книги параметров формирующей системы голосового тракта становится связана определенная область (кластерное подпространство) кодовой книги векторов сигналов возбуждения. После процедуры обучения, в результате которой определяются данные подпространства сигналов возбуждения для всех векторов кодовой книги параметров формирующей системы голосового тракта, осуществляется запись в память системы обработки соответствующих зависимостей.

Блок-схема алгоритма нахождения взаимозависимостей элементов декомпозиции речи при линейном предсказании в CELP-подобных алгоритмах представлена на рис. 1

Рис.1 . Блок-схема алгоритма нахождения взаимозависимостей элементов декомпозиции речевого сигнала при линейном предсказании

При функционировании системы обработки РС вычисленный, вектор кодовой книги параметров формирующей системы голосового тракта определяет подпространство векторов сигналов возбуждения синтезирующей модели. Использование такой реструктуризации кодовых книг дает возможность значительно сократить среднюю скорость передачи данных в канале связи при низкоскоростном кодировании РС. Если же производить индивидуализацию кодовых книг, построение которых осуществлялось на основе элементов декомпозиции РС, то такой подход дает возможность учитывать индивидуальные особенности диктора и осуществлять процедуру непрерывной речевой аутентификации в процессе функционирования системы обработки РС. Обобщенная блок-схема алгоритма ее функционирования на передающей стороне представлена на рис.2.

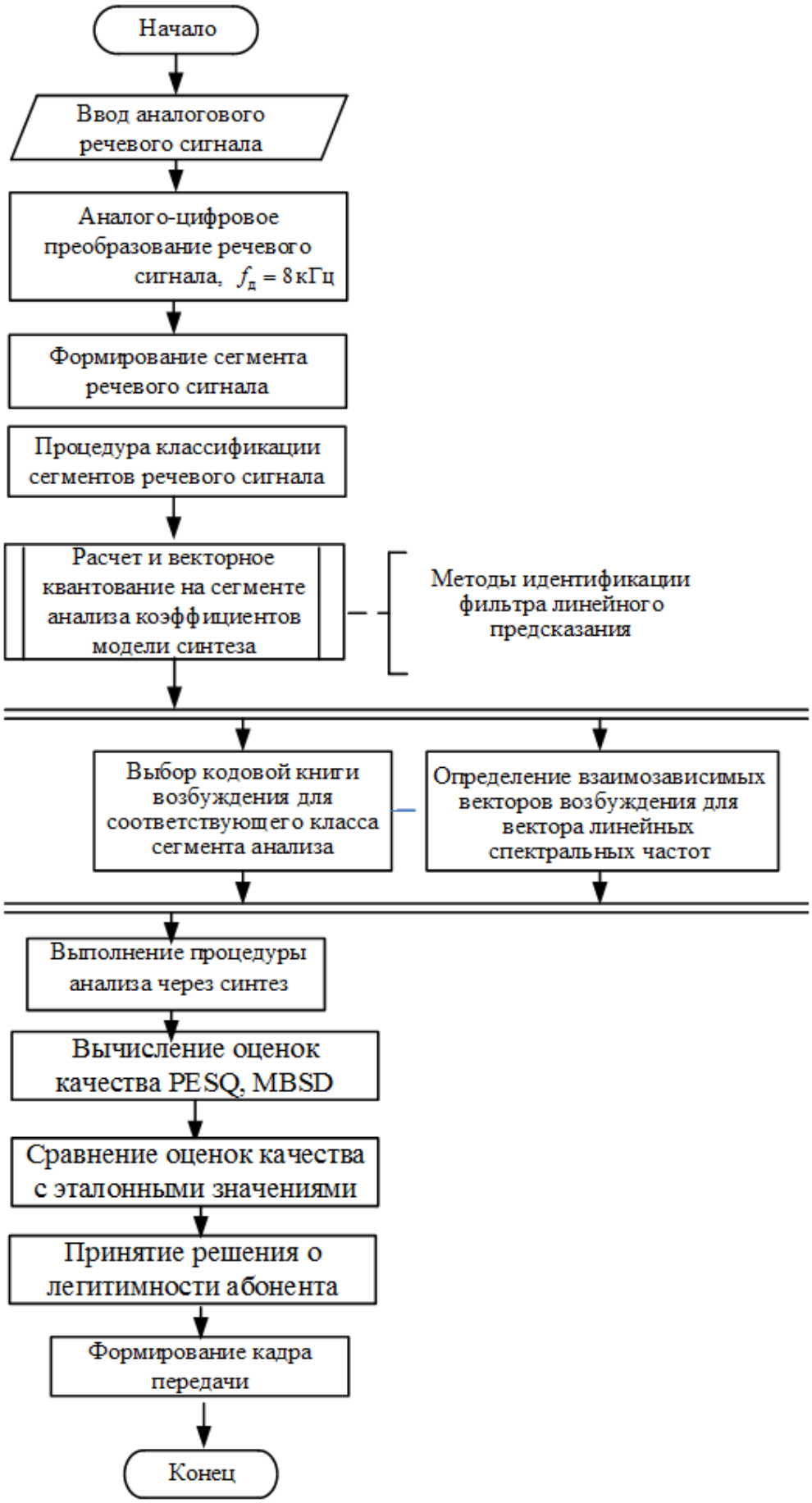

Рис.2. Обобщенная блок-схема функционирования системы обработки речевого сигнала на передающей стороне

Вычисленный вектор параметров голосового тракта определяет индивидуальность диктора, отражая его формантные особенности, а их совокупность определяет динамику изменения формантных областей в процессе ведения разговора. Таким образом, по распределению значений составляющих нескольких подобных векторов можно сделать заключение с высокой степенью вероятности о принадлежности диктора легитимной группе пользователей сети связи, при этом априорные вероятности появления объектов аутентификации с заданными свойствами являются неизвестными. Важными показателями при этом являются вычисленные ошибки первого и второго рода [9]. Для понижения вероятности возникновения ошибки второго рода P_2 (пропуска "чужого") необходимо увеличение длины анализируемой выборки активной речи. При частоте дискретизации 8 кГц и использовании линейного 8-битного квантователя снижение ошибки второго рода до приемлемых значений (P_2<0.05) достигается при увеличении длины анализируемого отрезка активной речи до 20с. Однако такая процедура, как правило, выполняется одноразово на этапе активации абонентского устройства, при многократном ее повторении будет значительно возрастать вычислительная сложность используемых алгоритмов аутентификации и соответственно возрастать задержка алгоритма по принятию решения.

При использовании для низкоскоростного кодирования речи выявленных взаимозависимостей между векторами соответствующих кодовых книг элементов декомпозиции РС обученных на уникальной речи диктора, возникающие связи будут характеризовать индивидуальные особенности пользователя подобной системы. При этом время оценки факта несанкционированного доступа на фразе активной речи абонента значительно снижается (до 2-5 с), что связано с наличием структурно организованных связей между подпространствами кодовой книги векторов сигналов возбуждения и векторами кодовой книги параметров голосового тракта легитимного пользователя, которые будут нарушены при попытке использования системы связи.

Подобные зависимости взаимного отношения частот обращения к одинаковым областям кодовой книги векторов возбуждения для отдельного легитимного пользователя можно выделить, используя контроль потока обращений к этим областям. Такой подход несколько усложняет процедуру обучения, однако дает возможность использовать в абонентских устройствах обработки РС, основанных на векторном квантовании параметров линейного предсказания, универсальные кодовые книги параметров декомпозиции. Таким образом, нет необходимости в коренной модернизации уже работающего абонентского оборудования.

Поддержка решение о получении доступа к услугам сети связи с персонального абонентского терминала будет осуществляться постоянно (непрерывно). Если микрофон оказывается в «чужих» руках, изменится поток обращения к областям кодовой книги. Изменение потока векторов кодовой книги для параметров голосового тракта «чужого» голоса повлечет за собой иной выбор подпространства кодовой книги. Это приведет к разрушению выявленных статистических связей между элементами декомпозиции РС и позволяет выявить факт несанкционированного доступа к сети связи. Если при этом отслеживать обращение к не свойственным пользователю фрагментам кодовой книги и запрещать их, то следствием будет являться резкое ухудшения объективных показателей качества синтезированной речи.

На рис.3 показана упрощенная схема аутентификации легитимного диктора в сети связи.

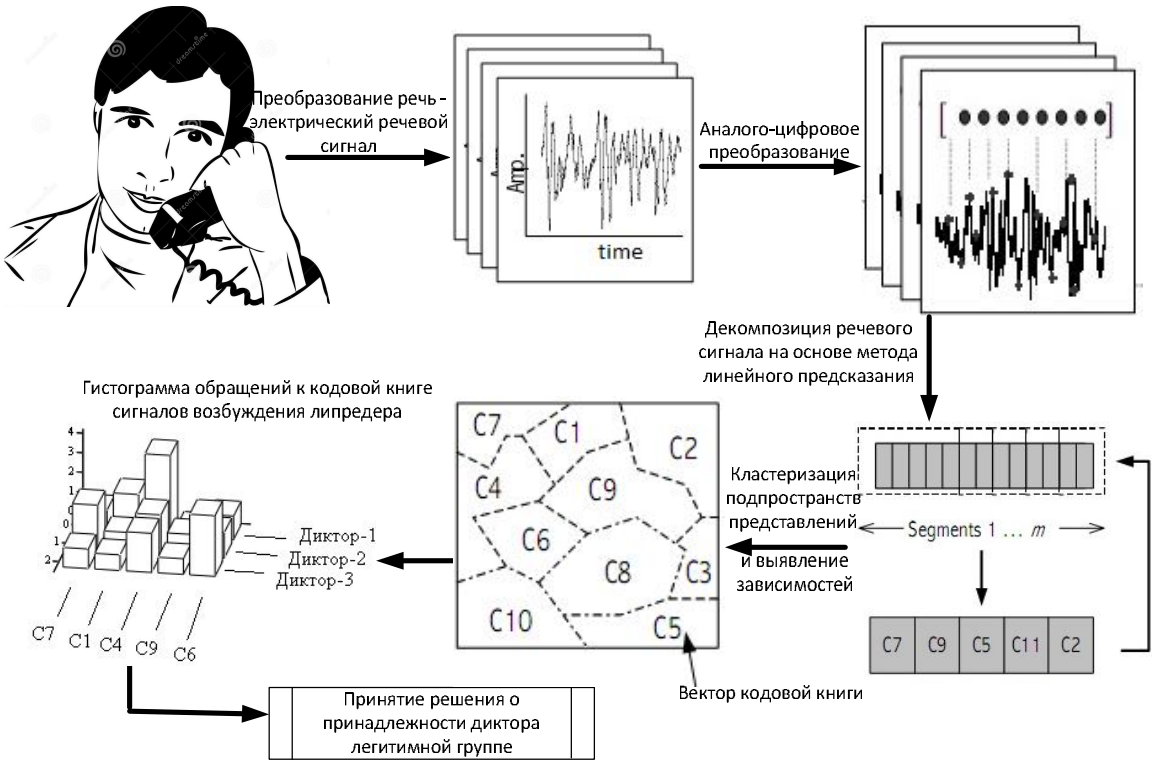

Рис.3 Аутентификации легитимного диктора в сети связи

Применение взаимозависимости элементов декомпозиции РС, возникающее при линейном предсказании на основе CELP-подобных алгоритмов в процессе его обработки и низкоскоростного кодирования, позволяет по-новому рассмотреть задачу создания индивидуально ориентированных систем абонентского речевого кодирования. Подобная система обработки РС позволяет достигать значительного коэффициента сжатия по сравнению с существующими классическими вокодерами на основе линейного предсказания. При этом, индивидуализация абонентского устройства осуществляется в процессе его использования путем поэтапного создания индивидуальных кодовых книг элементов декомпозиции РС на речи пользователя на основе самообучения системы, использующей стандартный полностью автоматический алгоритм обучения больших искусственных нейронных сетей ГОСТ Р 52633.5 [10]. Следует подчеркнуть, что наличие полностью автоматизированного обучения – это принципиально важное требование к средствам биометрической аутентификации личности по базовому стандарту ГОСТ Р 52633.0 [11].

Таким образом, использование положений многомерного статистического кластерного анализа и методов классического линейного предсказания РС на основе CELP-моделей позволяют получить одновременно значительное снижение средней скорости передачи при персонификации абонентского доступа к сети связи.

Литература

1. Руд, Б. Руководство по биометрии/ Б. Руд. – М. :Техносфера, 2007. – 368 с.

2. Волчихин В.И., Иванов А.И., Назаров И.Г., Фунтиков В.А., Язов Ю.К. Нейросетевая защита персональных биометрических данных/ Под ред. Ю.К. Язова. М.: Радиотехника, 2012.– 160с.

3. Monrose F., Reiter M., Li Q., Wetzel S. Cryptographic key generation from voice. //Proc. IEEE Symp. on Security and Privacy, pp. 202-213, 2001.

4. Быков, С. Ф. Цифровая телефония : учеб. пособие для вузов / С. Ф. Быков, В. И. Журавлев, И. А. Шалимов.–М.: Радио и связь, 2003. – 144 с.

5. Маркел, Дж. Д. Линейное предсказание речи / Д. Маркел, А. Х. Грей.пер. с англ./под ред. Ю. Н. Прохорова и В. С. Звездина. – М. : Связь, 1980. – 308 с.

6. Марпл – мл. С. Л. Цифровой спектральный анализ и его приложения. – М.: Мир, 1990. – 584 с.

7. Афанасьев А.А. Непрерывная аутентификация диктора при ведении телефонных переговоров по низкоскоростным цифровым каналам. Научно-практический журнал "Вопросы кибербезопасности", №3(16), с. 60-68.

8. Макхоул, Д. Векторное квантование при кодировании речи / Д. Макхоул, С. Рукос, Г. Гиш. ТИИЭР. – 1985. – Т.73. – №11. – С. 19–61.

9. А.Л. Горелик, В.А. Скрипкин Методы распознавания: уч. Пособие для вузов. – 4-е изд., М.: Высш. шк., 2004. – 261с.

10. ГОСТ Р 52633.5-2011 «Защита информации. Техника защиты информации. Автоматическое обучение нейросетевых преобразователей биометрия-код доступа».

11. ГОСТ Р 52633.0-2006 «Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации».

Статья поступила 07.12.2016, опубликована 17.12.2016
по положительной рецензии д.т.н. Иванова А.И.