

13. Болотов, А. А. Алгоритмические основы эллиптической криптографии [Текст] / А. А. Болотов. - М. : Изд-во МЭИ, 2000. – 100 с.
14. Bierbrauer, J. Universal hashing and geometric codes, to appear in Designs [Electronic resource] / J. Bierbrauer // Codes and Cryptography. – 1994. – Access mode: <http://www.math.mtu.edu/~jbierbra/hashcol.ps>
15. Кларк, Дж.-мл. Кодирование с исправлением ошибок в системах цифровой связи [Текст] : пер. с англ. / Дж.-мл. Кларк, Дж. Кейн ; под ред. Б. С. Цыбакова. – М. : Радио и связь, 1987. – 392 с.
16. Кузнецов, А. А. Алгебраическое декодирование кодов по кривым Эрмита [Текст] / А. А. Кузнецов, А. В. Северинов, Д. А. Задворный, В. Н. Лысенко // Вісник ХПІ. – 2003. – № 26. – С 95-102.

УДК 621.395

ОБ УМЕНЬШЕНИИ ОБЪЕМА ТРАФИКА ПРИ ПАКЕТНОЙ ПЕРЕДАЧЕ РЕЧЕВЫХ СООБЩЕНИЙ ЗА СЧЕТ КОДИРОВАНИЯ ПАУЗ*

С. П. Белов, Е. И. Прохоренко

ВВЕДЕНИЕ

Речевые сообщения передаются посредством речевых сигналов, под которыми, в контексте данной работы, понимается результат кодирования звуков речи с помощью некоторых устройств. В ряде случаев, как, например, в традиционной телефонии, передаются электрические сигналы, возникающие на выходе микрофона при воздействии колебаний воздушного потока. В современных цифровых системах речевые сигналы представляют в дискретном виде и передают, в соответствии с некоторыми правилами, значения дискретных отсчетов, что приводит к необходимости кодирования сигнала.

Некоторые системы связи не могут работать непрерывно, например спутниковые системы передают сообщения в определенные моменты времени, находясь в определенных точках пространства, метеорные каналы связи работают при наличии определенных условий передачи. В этих случаях предварительно формируются пакеты битовых представлений сигналов, которые затем посылаются в канал связи, а на приемной стороне декодируются по известным правилам – такой принцип передачи информации принято называть пакетным режимом.

В процессе формирования пакетов исходные данные могут быть преобразованы таким образом, чтобы сократить объем их битового представления, что позволяет уменьшить трафик. Это можно считать основным преимуществом пакетной передачи. Во многом благодаря этому в настоящее время режим пакетной передачи получил широкое распространение в телекоммуникационных сетях, например в мобильных системах, IP-телефонии и др.

При передаче речевых сообщений сокращение трафика может быть достигнуто за счет кодирования пауз, так как при диалоге каждый из участников говорит, в среднем, только 35 процентов времени [1]. Кодирование пауз заключается в определении интервала, на котором отсутствуют звуки речи, фиксации начала этого интервала и его длительности. Кроме того, для воспроизведения речи с комфортным звучанием необходимо определить некоторые параметры этого интервала, например значения математического ожидания и наименьшего среднеквадратичного отклонения.

Известны различные методы обнаружения пауз в речевых сообщениях, среди которых достаточно широкое распространение получила технология VAD (детектор речевой активности) [3]. Технология VAD используется совместно с большим числом речевых кодеков. Обнаружение паузы основано на поиске сигнала продолжительностью в несколько сотен миллисекунд, со значениями отсчетов ниже заданного уровня (обычно –

* Работа финансировалась в рамках гранта Белгородского государственного университета.

20Дб) на всем отрезке, входящем в пакет. Такой подход не имеет строгого математического обоснования. Использование предположения, что продолжительность паузы должна быть не менее нескольких сотен миллисекунд, отражает стремление не допустить удаления частей сигнала соответствующим звукам речи, что гарантирует сохранность ее разборчивости. Вместе с тем, очевидно, что длительность пауз между словами, фразами зависит от дикторов. Кроме того, известно, что речь состоит из фонем, перемежающихся паузами гораздо меньшей, чем несколько сотен миллисекунд, длительности, которые также целесообразно кодировать.

Таким образом, представляется актуальной разработка метода определения в речевых сигналах отрезков соответствующим паузам между словами, фразами и фонемами.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ МЕТОДА

Пауза в речевом сигнале – отрезок, не имеющий регулярных периодических циклов, а более или менее нерегулярные случайные изменения, т.е. процесс, который можно отнести к классу случайных. Если принять, что вероятностная структура паузы не изменяется со временем, то тогда сигнал можно считать **случайным стационарным процессом**.

Понятие стационарности отражает предположение о том, что с течением времени вероятностные свойства отсчетов последовательности $y(n)$ и статистические связи между ними не изменяются [7]. Для стационарных случайных последовательностей должны быть справедливы неравенства:

$$\begin{aligned} m(k) &= const, \\ \sigma^2(k) &= \sigma^2 = const, \\ R(n, k) &= R(|n - k|). \end{aligned} \quad (1)$$

Здесь

$m(k) = Ey(k)$ – математическое ожидание,

$\sigma^2(k) = E(y(k) - m(k))^2$ – дисперсия,

$R(n, k) = E[(y(k) - m(k)) \cdot (y(n) - m(n))]$ – автоковариационная функция (АКФ), используемая для описания статистической связи между последовательностями, где E – оператор математического ожидания случайной величины [7].

Сигнал, соответствующий звукам, не является стационарным, так как формируется при активном воздействии речевого аппарата. Процедура обнаружения пауз может быть основана на принципе обнаружения отличий характеристик сигналов на данном интервале по сравнению с характеристиками сигнала в паузе.

Для описания характеристик сигнала в паузе, которые определяются заранее, в данной работе используется модель авторегрессии [4], или стохастическое разностное уравнение, которое можно представить следующим образом:

$$y(n) - m = \sum_{k=1}^p \alpha_k (y(n-k) - m) + \sigma_0 u_t, \quad (2)$$

где m – математическое ожидание $y(n)$, u_t – некоррелированная последовательность с единичной дисперсией и нулевым математическим ожиданием, σ_0 – параметр, определяющий уровень среднеквадратичной погрешности предсказания на основе линейной комбинации вида:

$$\hat{y}(n) = \sum_{k=1}^p \alpha_k (y(n-k) - m). \quad (3)$$

Отрезки речевого сигнала, которые не относятся к паузе, будут отличаться по структуре от сигнала в паузе, т.е. для каждого звука возможна своя модель авторегрессии. Таким образом, можно сформулировать следующие гипотезы:

H_0 : отрезок сигнала генерируется по схеме (2) с известными параметрами α_k, m, σ_0, p .

H_1 : отрезок сигнала генерируется по схеме отличной от (2).

Для проверки этих гипотез необходимо ввести решающую функцию (РФ). В данной работе будут использованы результаты, полученные в [5], где найдена не имеющая мертвых зон векторная решающая функция $S_m(\tau)$, использующая в качестве компонент статистики максимальной чувствительности:

$$S_1(\tau) = \sum_{k=1}^{\tau} z(k), \quad (4)$$

$$S_2(\tau) = \frac{1}{2} \sum_{k=1}^{\tau} z^2(k) - \frac{\tau}{2}, \quad (5)$$

$$S_i(\tau) = \sum_{k=1}^{\tau+2-i} z(k)z(k+i-2), \quad i = 3, \dots, m, \quad (6)$$

где $z(i)$ последовательность вида:

$$z(i) = \sum_{k=0}^p \alpha_k (y(i-k) - m) / \sigma_0 \quad (7)$$

Соотношения для границ доверительных интервалов РФ имеют вид

$$a_i(\tau) = -b_i(\tau), \quad b_i(\tau) = k\sqrt{\tau - i + 2}, \quad (8)$$

где k – некоторый коэффициент.

Реально АКФ и ее параметры не известны, и их можно только оценить по значениям реализаций случайных последовательностей. В данной работе рассматривается один из возможных подходов к решению этой задачи с применением принципа адаптации к реализациям пауз речевых сигналов, используемых при формировании пакета.

Для оценки работоспособности предложенного метода определения в речевых сигналах отрезков, соответствующих паузам между словами, фразами и фонемами были проведены вычислительные эксперименты для установления возможности использования отрезков шумов микрофона для оценивания параметров модели генерации сигнала в паузах вида (2).

2. ЭКСПЕРИМЕНТАЛЬНОЕ ИССЛЕДОВАНИЕ МЕТОДА

Один из экспериментов заключался в следующем. Были записаны сигнал $y_1(n)$, не имеющий информационной нагрузки, основной составляющей которого являлись шумы микрофона, сигнал $y_2(n)$, являющийся паузой, расположенной между двумя словами записанной ранее фразы, и сигнал $y_3(n)$, являющийся паузой расположенной между двумя фонемами слова «аппроксимация».

Из рисунка 1 видно, что согласно критерию Акаике [7], для моделей авторегрессии всех трех сигналов можно принять порядок $p = 14$.

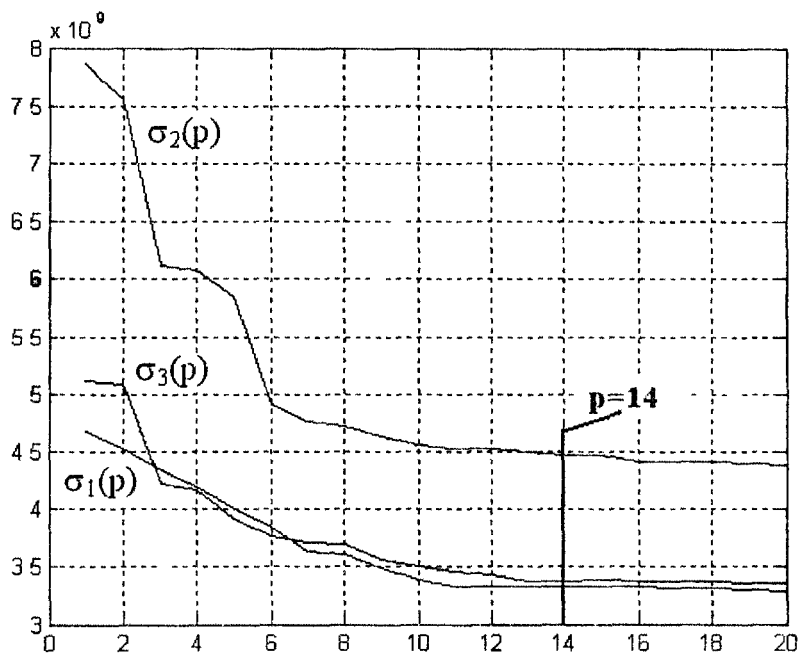


Рис 1. График определения порядка модели авторегрессии

Была проведена оценка параметров сигналов m , α_k , σ_0^2 (таблица 1), с использованием метода наименьших квадратов [7]. Численные значения оцениваемых параметров достаточно близки. Это обстоятельство определило возможность принять в качестве параметров гипотезы H_0 параметры сигнала $y_1(n)$ для дальнейшего исследования.

Таблица 1
Параметры моделей сигналов вида (2)

	$m \times 10^{-3}$	α_k , при $k=1, \dots, p$														$\sigma_0^2 \times 10^{-9}$
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	
$y_1(n)$	42.95	0.84	-0.77	0.88	-0.89	1.00	-0.90	0.83	-0.68	0.58	-0.47	0.31	-0.16	0.16	-0.04	6.26
$y_2(n)$	42.99	1.04	-0.97	1.18	-0.99	1.07	-1.03	0.72	-0.59	0.50	-0.38	0.26	-0.17	0.14	-0.05	4.47
$y_3(n)$	42.91	0.92	-0.70	0.83	-0.62	0.66	-0.52	0.38	-0.34	0.29	-0.19	0.04	0.02	0.08	0.04	3.38

При вычислении решающей функции $S_2(\tau)$ (5) для сигналов $y_1(n)$, $y_2(n)$, $y_3(n)$ были выбраны интервалы $\tau = 3, 5, 10, 15$ отсчетов как достаточные для быстрого обнаружения разладки при сохранении высокого качества воспроизведения речи. В качестве иллюстрации на рисунке 2 приведены графики рассчитанных значений РФ для сигнала $y_1(n)$ с параметрами, приведенными в таблице 1, при различных τ и при $k = 4$.

Значения РФ практически не превышают границ доверительных интервалов $a_i(\tau)$ (8) при всех рассмотренных значениях τ , но при $\tau=5$ почти все значения РФ близки пороговому значению, но не превышают его (исключение составляет лишь одно значение РФ, т.е. всего 5 отчетов сигнала $y_1(n)$ не были идентифицированы как пауза). Это свойство представляется полезным, так как, с одной стороны, достигается приемлемый уровень ложных тревог, а с другой – обеспечивается наибольшая чувствительность в обнаружении отрезка сигнала, соответствующего отсутствию паузы.

При анализе сигналов $y_2(n)$ и $y_3(n)$ наблюдалась схожая картина, т.е. значения РФ рассчитанные для $y_2(n)$ и $y_3(n)$, также не превышают границ доверительных интервалов на заданных τ , при оптимальном значении $\tau=5$. Поэтому представляется обоснованным использовать именно это значение для интервала вычисления решающей функции.

Для исследования поведения РФ при наличии разладки (моменты окончания пауз) было записано одно и то же слово «аппроксимация» двумя различными дикторами $x_1(n)$ и $x_2(n)$.

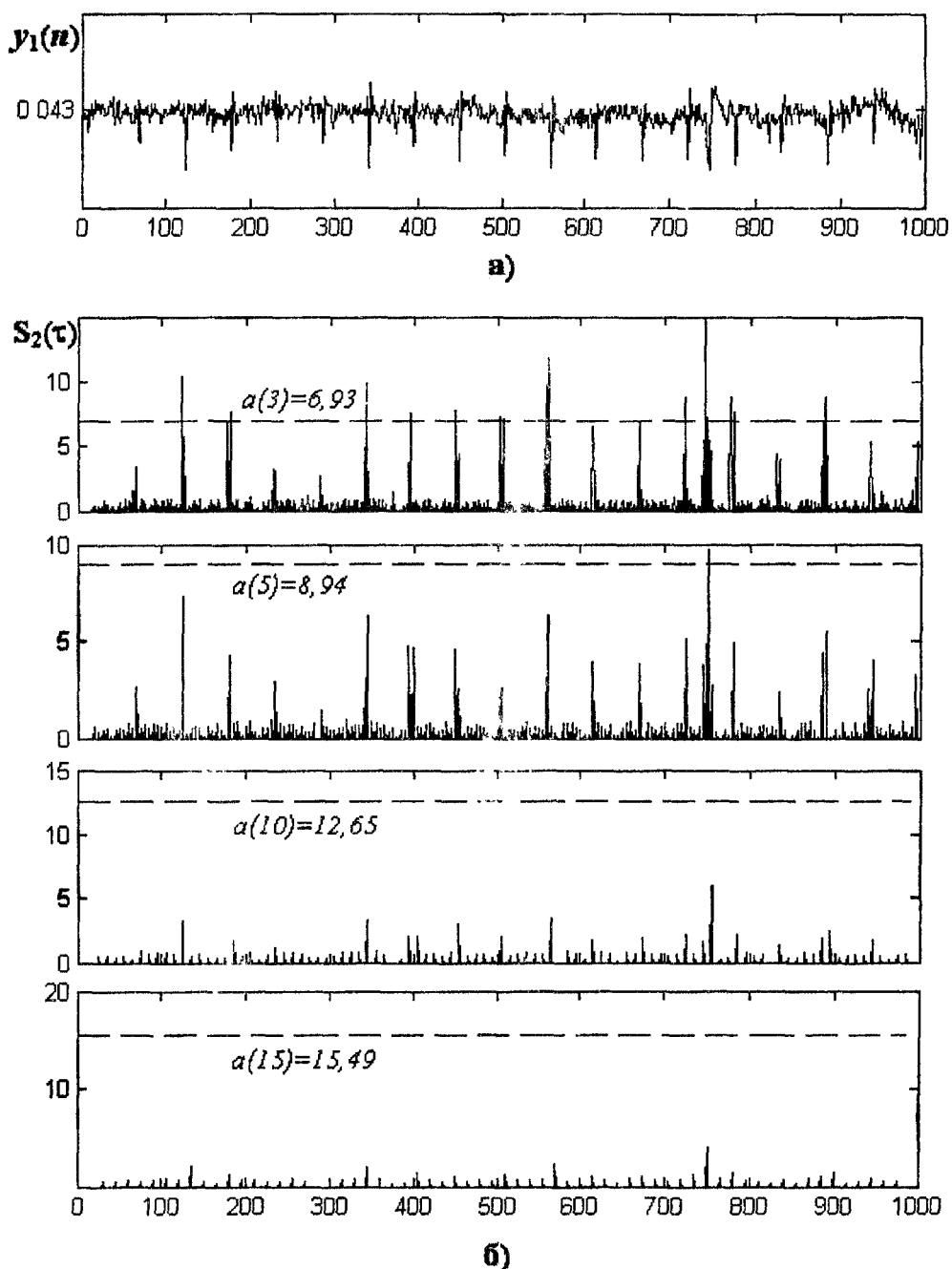


Рис. 2. Графики рассчитанных значений РФ для сигнала $y_1(n)$:
 а) отрезок сигнала $y_1(n)$; б) решающая функция второго рода $S_2(\tau)$ при различных τ

Для сигнала $x_1(n)$ рассчитана РФ $S_2(\tau)$ с использованием параметров, полученных при анализе сигнала $y_1(n)$ (таблица 1). На рисунке 3а показаны отрезок речевого сигнала $x_1(n)$ и значения $S_2(\tau)$ на всем интервале его анализа, а на рисунке 3б поведение РФ на участке сигнала включающего как паузу, так и звук (отсчеты с 2000 по 3000 сигнала $x_1(n)$). Значения РФ значительно превышают границу доверительного интервала, показанную на рисунке пунктирной линией, на тех отсчетах сигнала, которые порождены звуками речи.

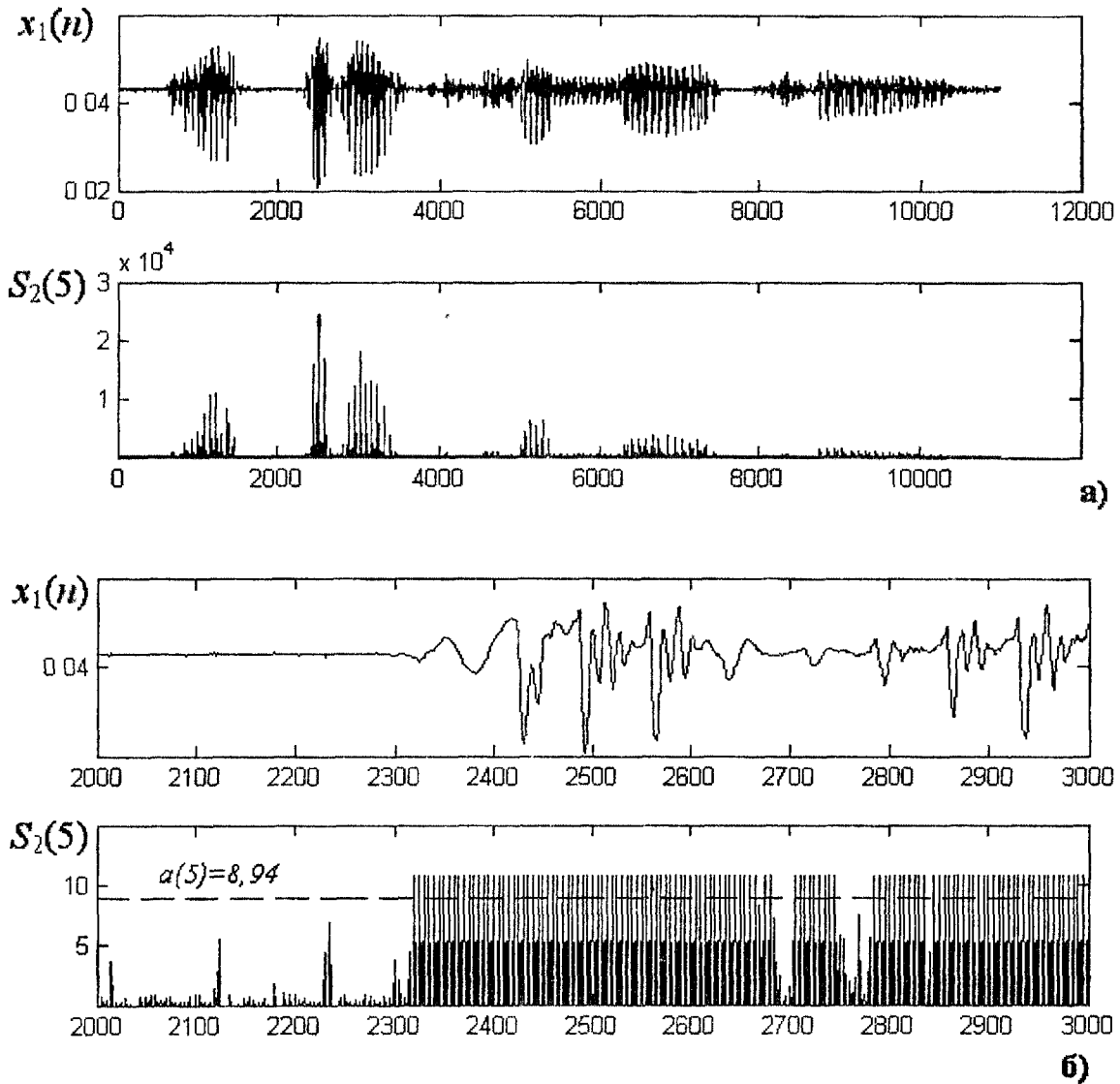


Рис. 3. Отрезок речевого сигнала $x_1(n)$ и значения $S_2(\tau)$:

- а) на всем интервале его анализа;
- б) на интервале $n=2000:3000$

Таким образом, применение данной РФ действительно позволяет быстро обнаружить разладку при малых размерах интервалов запаздываний.

Аналогичные испытания были проведены с использованием различных пороговых значений границ доверительных интервалов для сигналов $x_1(n)$ и $x_2(n)$ (таблица 2). Степень сжатия определялась как отношение количества отсчетов исходного сигнала к количеству отсчетов сигнала с удаленными отрезками сигнала – паузами.

Для подтверждения правильности определения границ пауза/звук было оценено качество воспроизведения речи при удалении и последующем восстановлении обнаруженных пауз.

Если значения РФ на i -ом интервале не превышали пороговых значений, то отсчеты речевого сигнала $x_1(n)$ на данном интервале удалялись и кодировались (значения номеров начальных отсчетов пауз и длительность паузы), таким образом формировалась последовательность $x'_1(n)$, представляющая собой «сжатый» речевой сигнал.

После восстановления пауз, качество воспроизведения было оценено несколькими экспертами.

Существует множество подходов к проблеме определения качества [3]. Наиболее широко используемый подход оперирует оценкой MOS (Mean Opinion Score), которая определяется как средняя оценка качества большой группой слушателей по пятибалльной шкале. Для прослушивания экспертам предъявляются разные звуковые фрагменты – речь, музыка, речь на фоне различного шума и т.д. Оценки интерпретируются следующим образом: 4-5 – высокое качество;

- 3.5-4 – качество ТфОП; такое качество обычно обеспечивается в большинстве телефонных разговоров;
- 3-3.5 – качество речи по-прежнему удовлетворительно, однако его ухудшение явно заметно на слух;
- 2.5-3 – речь разборчива, однако требует концентрации внимания для понимания.

Результаты оценки качества воспроизведения речи также приведены в таблице 2.

Таблица 2

Результаты сжатия и оценки качества воспроизведения речи при различных k

$x_1(n)$						
Пороговый коэффициент, k	4	5	10	20	40	60
Степень сжатия, Ксж	1,52	1,57	1,796	2,14	2,74	3,32
Оценка качества воспроизведения речи, балл	5	4,6	4	3,5	3,1	2,8
$x_2(n)$						
Пороговый коэффициент, k	4	5	10	20	40	60
Степень сжатия, Ксж	1,3	1,4	1,53	1,7	2,03	2,34
Оценка качества воспроизведения речи, балл	5	4,8	4,3	3,9	3,5	3,2

Как видно из табл. 2, степень сжатия определяется выбранным значением порогового коэффициента, оптимальное значение которого зависит от нескольких факторов. В данной работе значения коэффициентов выбирались экспериментально, а задача их оптимизации не решалась.

ЗАКЛЮЧЕНИЕ

Результаты эксперимента свидетельствуют о том, что предлагаемый метод позволяет достаточно быстро обнаруживать границу пауза/звук с высокой степенью точности, т.е. обнаруживать паузы между фразами, словами и фонемами. При этом можно обеспечить изменение коэффициента сжатия в широких пределах, что позволяет оптимизировать трафик в системах связи, работающих в режиме пакетной передачи, по критерию качества воспроизведения речи во всем диапазоне экспертных оценок.

Библиографический список

1. Росляков, А. В. IP-телефония [Текст] / А. В. Росляков М. Ю. Самсонов, И. В. Шibaева. – М. : Экотредз, 2001. – 250 с.
2. Кораблин, М. А. Оценка эффективности использования технологий VoIP и VAD в корпоративных IP-сетях [Текст] / М. А. Кораблин, А. В. Мороз. – Электросвязь. – 2004. – № 8.
3. Гольдштейн, В. С. IP-Телефония [Текст] / В. С. Гольдштейн, А. В. Пинчук, А. Л. Суховицкий. – М. : Радио и связь, 2001. – 336с.: ил.

4. Андерсон, Т. Статистический анализ временных рядов [Текст] / Т. Андерсен ; пер. с англ. И. Г. Журбенко, В. П. Носко ; под ред. Ю. К. Беляева. – М. : Мир, 1976. – 755с.: ил.

5. Жилияков, Е. Г. Статистики максимальной чувствительности в задаче обнаружения изменений параметров процессов авторегрессии [Текст] / Е. Г. Жилияков, Э. К. Шпилевский // Заводская лаборатория. – 1992 – № 7.

6. Васюков, В. Н. Цифровая обработка сигналов и сигнальные процессоры в системах подвижной радиосвязи [Текст] : учебник / В. Н. Васюков. – Новосибирск : Изд-во НГТУ, 2003. – 292 с.: ил.

7. Жилияков, Е. Г. Методы и алгоритмы обработки экспериментальных данных в атомно-абсорбционной спектроскопии [Текст] / Е. Г. Жилияков, Н. И. Корсунов, Д. П. Лагода ; Харьков. политехн. ин-т ; под ред. Н. И. Корсунова. – Киев : Наук. думка, 1992. – 122 с.: ил.

УДК 621.391

СИСТЕМНЫЕ АСПЕКТЫ ПОСТРОЕНИЯ РЕГИОНАЛЬНОЙ МУЛЬТИСЕРВИСНОЙ СЕТИ СВЯЗИ И МЕТОДЫ ОЦЕНКИ НАДЕЖНОСТИ

Г. Н. Кузьменко

В современном обществе существенно возрастает роль и значимость информации под влиянием изменений структуры и масштабов производства, научно-технического прогресса, повышения деловой активности, роста образовательного и культурного уровня населения, расширения связей во всех сферах жизнедеятельности как в национальном, так и в интернациональном масштабах. Информацию справедливо рассматривают в качестве важнейшего национального ресурса и составной части национального богатства любой страны. Происходит постепенный переход от индустриального общества к информационному. В связи с этим возрастают и потребности в обмене информацией между различными ее источниками и потребителями. Телекоммуникации и связь являются неотъемлемой частью производственной и социальной инфраструктуры экономики России. В соответствии с основными положениями Концепции развития электросвязи Российской Федерации главной целью развития на период до 2010-2015 годов является достижение таких удельных показателей инфокоммуникаций, которые обеспечили бы эффективное развитие экономики страны в условиях рыночных отношений, удовлетворение возрастающих информационных потребностей общества, органов государственной власти (управления), обороны страны, правопорядка.

Развитие телекоммуникационных технологий привело к изменениям принципов и методов построения, а также путей развития современных цифровых сетей связи. Без современной высокотехнологичной среды передачи и распределения информации невозможен научно-технический прогресс и социальное развитие регионов страны. Электронный обмен информацией охватывает новые сферы жизнедеятельности человека. Наблюдается не только увеличение объема передаваемых данных, но и задаются более жесткие требования к пропускной способности сетевых структур и надежности их функционирования с целью обеспечения высокого качества новых услуг.

При этом в настоящее время отмечается тенденция роста влияния регионов на все стороны экономической и общественно-политической жизни страны.

Основополагающая роль в процессе обмена информацией отводится сетям электросвязи. В соответствии с законом «О связи», который вступил в действие в 2003 году, единая сеть электросвязи РФ (ЕСЭ) представляет собой совокупность сопряженных сетей электросвязи общего пользования, технологических (ранее корпоративных и ведомственных) сетей и других сетей на территории России, независимо от ведомственной принадлежности и форм собственности. Процессы конвергенции и интеграции сетей пакетной коммутации и традиционных сетей коммутации каналов