

УДК 621.391

С.М. Бобрицкий¹, А.В. Колтыков², В.Н. Ткачѐв²¹Харьковский НИИ судебных экспертиз им. Н.С. Бокариуса, Харьков²Харьковский национальный университет радиоэлектроники, Харьков

ОБНАРУЖЕНИЯ ИНТЕРВАЛОВ ГРУППИРОВКИ ПОТЕРЬ РЕЧЕВЫХ ПАКЕТОВ ПО ПАКЕТНЫМ СЕТЯМ С ИСПОЛЬЗОВАНИЕМ КОДЕКОВ С ПОДАВЛЕНИЕМ ПАУЗ РЕЧИ И КОМПЕНСАЦИЕЙ ПОТЕРЬ

Усовершенствован метод обнаружения интервалов группировки потерь речевых пакетов по пакетным сетям с использованием кодеков с подавлением пауз речи и компенсацией потерь, основанный на анализе фрагментов речевого трафика в вейвлет-базисе Мейера, позволяющий определить расположение и длительности указанных группировок и исключить их из анализа речевого трафика на контентном и пакетном уровнях.

Ключевые слова: пакетные сети, форманты речевых сигналов, вейвлет-базис Мейера, пакетный уровень, контентный уровень.

Введение

В телекоммуникационных системах важное место занимают технологии передачи речевой информации [1]. При передаче речевых сообщений через сеть с коммутацией пакетов возможна ситуация, когда один или несколько пакетов теряются вследствие перегрузок в канале связи, что может существенно влиять на результат идентификации диктора.

Значимую роль играет разработка методов обнаружения интервалов потерь речевых пакетов для определения их расположения и длительности с дальнейшим исключением из анализа речевого трафика как на пакетном так и на контентном уровнях.

Постановка проблемы и анализ литературы.

Анализ качества идентификации диктора при передаче речевых данных через Интернет [2 – 4] показывает, что основным источником возникновения искажений, снижения качества и разборчивости синтезированной речи является прерывание потока речевых данных, вызванное потерями пакетов при передаче по сети связи, либо превышением допустимого времени доставки пакета с речевыми данными.

С учётом возможных потерь пакетов в сети для восстановления речевого потока на приёмной стороне используется протокол реального времени – Real Time Protocol (RTP) [5]. В заголовке данного протокола, в частности, передаются временная метка и номер пакета. Эти параметры позволяют при минимальных задержках определить порядок и момент декодирования каждого пакета, а также интерполировать потерянные пакеты.

Восстановленная последовательность, с возможными пропусками как одиночных пакетов, так и групп пакетов, поступает на декодер. Декодер обеспечивает восстановление речевой информации, заполнение пауз фоновым шумом, а также эхокомпенсацию кодируемого сигнала, обнаружение и детектирование телефонной сигнализации.

Кодеки G.723.1 и G.729 частично адаптированы для условий потерь пакетов [6]. Различия этих стандартов включают в себя определение голосовой активности (Voice Activity Detector) и элементы, ответственные за синтез речевого сигнала на фрагментах, соответствующих потерянным речевым данным.

Влияние потерь кадров на качество воспроизводимой речи зависит от используемого кодека. Если потерян кадр, состоящий из N речевых отсчетов кодека G.711, то на приемном конце будет отмечен пропуск звукового фрагмента длительностью 125 мкс. Если используется более совершенный узкополосный кодек, то потеря одного кадра может сказаться на воспроизведении нескольких следующих, так как декодеру потребуется время для того, чтобы достичь синхронизации с кодером – потеря кадра длительностью 20 мс может приводить к слышимому эффекту в течение 150 мс и более.

Кодеры типа G.723.1 разработаны так, что они функционируют без существенного ухудшения качества в условиях некоррелированных потерь до 3% кадров, однако при превышении этого порога качество ухудшается катастрофически [7].

Целью статьи является разработка метод обнаружения интервалов потерь речевых пакетов, передаваемых по пакетным сетям с использованием кодеков с подавлением пауз речи и компенсацией потерь, позволяющего определить расположение и длительности указанных интервалов и исключить их из анализа речевого трафика на контентном и пакетном уровнях.

Метод обнаружения интервалов группировки потерь речевых пакетов

Для достижения поставленной в статье цели предлагается использовать свойства характерных искажений речевого сигнала при наличии однократных или групповых потерь. В качестве математиче-

ского аппарата, позволяющего описать и учесть указанные свойства был выбран вейвлет-анализ.

В случае учета подавления пауз речи (VAD) и отсутствия механизма компенсации потерь (PLC) (например, кодек G.711) для определения потерянных сегментов необходимо учитывать тот факт, что величина отсчетов в этих участках характеризуется минимальными значениями. Напротив, величины отсчетов в паузах речи и при речевой активности имеют большую энергию.

Предположим, что $x(i)$ – исходный сигнал, тогда сигнал, принятый на приемной стороне – $x'(i) = x(i) + e(i)$, где $e(i)$ – аддитивный шум (шум канала связи). Чтобы учесть влияние речевого сигнала на процесс обнаружения потерянных сегментов, необходимо выделить шумовую составляющую – остаточный шум, который определяется выражением

$$e'(i) = x'(i) - x''(i), \quad (1)$$

где $x''(i)$ – принятый сигнал после вейвлет-обработки.

Результат анализа данных зависит от выбора базиса. На практике с целью такого выбора часто пользуются критерием минимума энтропии. Чем меньше значение энтропии коэффициентов разложения суммируемых по номерам коэффициентов и уровням их размещения, тем адекватнее для исследуемых данных базис анализа

$$H = \exp \left(\left(- \sum_{j=1}^m \sum_{k=1}^n (d_{j,k})^2 \log d_{j,k}^2 \right) \right), \quad (2)$$

где d – коэффициенты разложения, j – текущий уровень разложения, k – номер коэффициента, m – конечный уровень разложения, n – количество коэффициентов.

Вейвлет-анализ речевых сигналов. В результате анализа речевых сигналов, было установлено, что наилучшим в условиях решения поставленной задачи является вейвлет Мейера [8]. Вейвлет Мейера и его масштабирующая функция в частотной области представлены на рис. 1.

Вейвлет-функция Мейера имеет вид:

$$\hat{\psi}(\omega) = \begin{cases} \sqrt{2\pi e^{-\frac{j\omega}{2}}} \sin\left(\frac{\pi}{2} v\left(\frac{3}{2\pi}|\omega|-1\right)\right) \text{ і } \delta \in |\omega| \in \left[\frac{2\pi}{3}; \frac{4\pi}{3}\right); \\ \sqrt{2\pi e^{-\frac{j\omega}{2}}} \cos\left(\frac{\pi}{2} v\left(\frac{3}{4\pi}|\omega|-1\right)\right) \text{ і } \delta \in \left[\frac{4\pi}{3}; \frac{8\pi}{3}\right); \\ 0 \text{ і } \delta \in \left[\frac{2\pi}{3}; \frac{8\pi}{3}\right), \end{cases} \quad (3)$$

где $v(a) = a^4(35 - 84a + 70a^2 - 20a^3)$, $a \in [0,1]$.

В свою очередь масштабирующая функция вейвлета Мейера имеет вид:

$$\hat{\phi}(\omega) = \begin{cases} \frac{1}{\sqrt{2\pi}} \text{ і } \delta \in |\omega| \leq \frac{2\pi}{3}; \\ \sqrt{2\pi e^{-\frac{j\omega}{2}}} \cos\left(\frac{\pi}{2} v\left(\frac{3}{2\pi}|\omega|-1\right)\right) \text{ і } \delta \in \left[\frac{2\pi}{3}; \frac{4\pi}{3}\right); \\ 0 \text{ і } \delta \in |\omega| > \frac{4\pi}{3}, \end{cases} \quad (4)$$

где ω – частота, t – время.

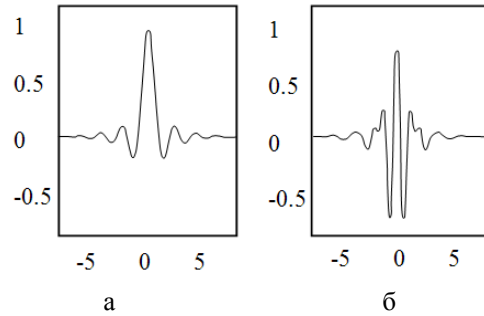


Рис. 1. Вейвлет Мейера: масштабирующая функция – а, вейвлет-функция – б

Разложение по вейвлет-базису, т.е. прямое преобразование вектора данных $x(t)$ осуществляется в соответствии со следующим выражением:

$$x(t) = \sum_{j,k=-\infty}^{\infty} c_{j,k} \Psi_{j,k}(t); \quad (5)$$

$$\Psi_{ab}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right); \quad a, b \in \mathbb{R}; \quad \psi \in L^2(\mathbb{R}). \quad (6)$$

При полностью дискретном процессе декомпозиции выражения для коэффициентов имеют вид:

$$c_{j,k} = 2^{1/2} \sum_n c_{j-1,n} h_{n+2k}; \quad (7)$$

$$d_{j,k} = 2^{1/2} \sum_n c_{j-1,n} g_{n+2k}, \quad (8)$$

где $c_{j,k}$ – аппроксимирующие коэффициенты, $d_{j,k}$ – детализирующие коэффициенты, а последовательности h_n и g_n являются соответственно низкочастотным и высокочастотным фильтрами декомпозиции [3].

Обратный процесс заключается в получении c_{j-1} из c_j и d_j :

$$c_{j-1,n} = 2^{1/2} \sum_k c_{j,k} h_{n+2k} + 2^{1/2} \sum_k d_{j,k} g_{n+2k}. \quad (9)$$

Определение значимости детализирующих коэффициентов. Для определения значимых детализирующих коэффициентов при проведении вейвлет-анализа голосовых сигналов обычно используется пороговая обработка [9].

Выбор типа пороговой обработки осуществляется следующим образом.

Пусть Z значение порога. Жесткая пороговая обработка сигнала выглядит следующим образом:

$$x = \begin{cases} x, & |x| > Z, \\ 0, & |x| \leq Z. \end{cases} \quad (10)$$

Мягкая пороговая обработка:

$$x = \begin{cases} \text{sign}(x)(|x| - Z), & |x| > Z, \\ 0, & |x| \leq Z. \end{cases} \quad (11)$$

Жесткая процедура создает сосредоточенные неоднородности на интервале $x = \pm Z$, тогда как при мягкой процедуре таких неоднородностей нет. В связи с этим в разработанном методе обнаружения интервалов потерь речевых пакетов выбрана мягкая пороговая обработка.

При этом выбор порога осуществляется по минимаксному алгоритму, который использует фиксированную форму порога, выбранную на основании среднеквадратической ошибки [10].

Выделение экстраполированных участков речи. В системах связи, в которых используются кодеки с механизмом компенсации потерь (PLC) (например, кодек G.729) для определения потерянных сегментов необходимо учитывать, что потеря одного кадра может сказаться на воспроизведении нескольких следующих, так как декодеру потребуется время для того, чтобы достичь синхронизации с кодером – потеря кадра длительностью 20 мс может приводить к слышимому эффекту в течение 150 мс и более.

При потере кадра кодек восстанавливает информацию на основании уже полученных данных, что в основном приводит к появлению экстраполированных предыдущих сегментов речи. Определить такой участок речи корреляционным методом не всегда возможно (рис. 5).

Особенность разработанного метода обнаружения интервалов потерь речевых пакетов заключается в том, что исследуемый сигнал раскладывается на вейвлет-коэффициенты, которые позволяют выявлять экстраполированные участки речи, устраняя при этом крупномасштабные составляющие.

Вычислительная схема разработанного метода. Разработанный метод обнаружения интервалов потерь речевых пакетов состоит из следующих этапов: удаление постоянной составляющей; очистка сигнала от шума с помощью вейвлета Мейера; разбиение сигнала на участки по размеру окна анализа; построение автокорреляционной функции детализирующих коэффициентов; вычисление энергии каждого участка сигнала; вычисление энергии отсчетов автокорреляционной функции каждого участка сигнала; вычисление порога обнаружения потерянных сегментов; определение местоположения потерянных сегментов.

1. Определение и удаление постоянной составляющей принятого сигнала.

$$x_{cp} = \frac{1}{N} \sum_{i=0}^{N-1} x'(i), \quad (12)$$

где $x'(i)$ – значение i -го отсчета сигнала; N – количество исследуемых отсчетов сигнала.

Затем из каждого элемента сигнала вычитается постоянная составляющая и результат записывается в следующем виде:

$$x''(i) = x'(i) - x_{cp}, \quad (13)$$

где $i = 1..N$.

2. Далее с помощью вейвлет-анализа определяется значение остаточного шума.

Для этого реализуется процедура вейвлет-фильтрации сигнала, включающая три шага:

- вычисление вейвлет-декомпозиции сигнала (7)-(8);

- пороговая обработка детализирующих коэффициентов (11);

- вычисление вейвлет-реконструкции сигнала (9).

На рис. 2 изображен фрагмент принятого сигнала, а на рис. 3 форма его остаточного шума.

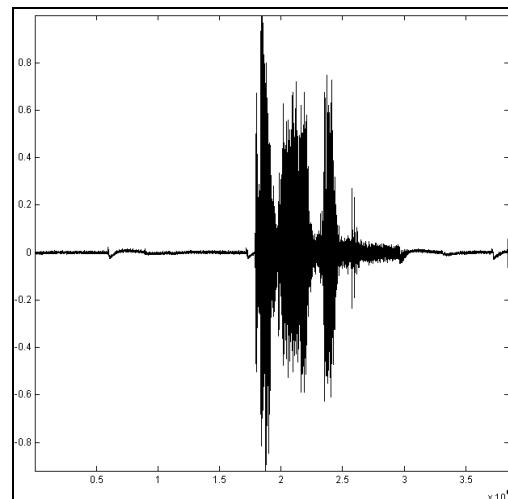


Рис. 2. Фрагмент принятого сигнала

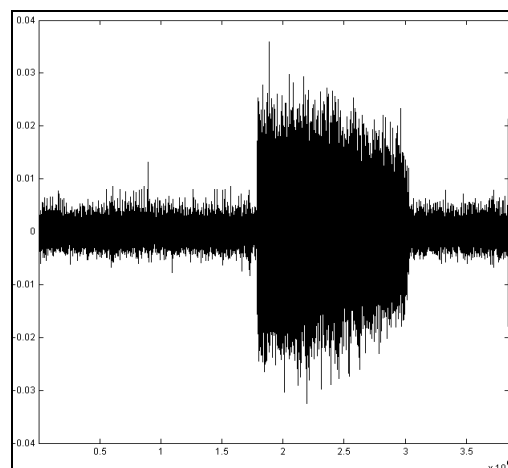


Рис. 3. Форма остаточного шума принятого сигнала

3. Производится разбиение остаточного шума на участки по размеру окна анализа.

Согласно данной процедуре сигнал $e'(i)$ продолжительностью, определяемой максимальным числом отсчетов $i = 1..N$, разбивается на участки (окна), количество которых определяется предварительно заданным числом отсчетов N_{window} в окне и равно

$$N_{window} = \frac{t_{window} \cdot f_{\text{дискретизации}}}{1000}, \quad (14)$$

где N_{window} – размер окна в отсчетах; t_{window} – размер окна в миллисекундах; $f_{\text{дискретизации}}$ – частота дискретизации сигнала в герцах.

4. В каждом окне определяется значение нормированной энергии сигнала:

$$S_{window} = \frac{1}{N_{window}} \sum_{i=0}^{N_{window}-1} X^2(i). \quad (15)$$

и шума:

$$S_e = \frac{1}{N} \sum_{i=0}^{N-1} e'^2(i). \quad (16)$$

5. Вычисляются значения автокорреляционной функции детализирующих коэффициентов:

$$r(i) = \frac{1}{N_{window}} \sum_{i=0}^{N_{window}-1} d(i)d(i+j-N_{window}), \quad (17)$$

где d – детализирующие коэффициенты вейвлет разложения, i – номер отсчета, j – номер окна анализа, а N_{window} – размер окна анализа.

6. Вычисление энергии отсчетов автокорреляционной функции каждого участка сигнала. При этом нормированная энергия корреляционных коэффициентов j -го окна производится согласно нахождению значений нормированной энергии остаточного шума (14) и определяется:

$$S_r = \frac{1}{N} \sum_{i=0}^{N-1} r^2(i). \quad (18)$$

7. Принятие решения об обнаружении интервалов потерь речевых пакетов осуществляется на основе сравнения нормированной энергии шума с пороговым значением, которое определяется выражением:

$$Z_{\text{порог}} = S_e / 2. \quad (19)$$

При этом местоположение начала обнаруженного интервала определяется первым отсчетом, в котором порог (19) превышен нормированной энергией шума.

8. Для сетей с использованием кодеков с подавлением пауз речи и компенсацией потерь допол-

нительно выявляются экстраполированные участки речи. Для этого с порогом сравнивается значение нормированной энергии корреляционных коэффициентов j -го окна (18).

Пример нахождения потерянных сегментов приведен на рис. 4, 5.

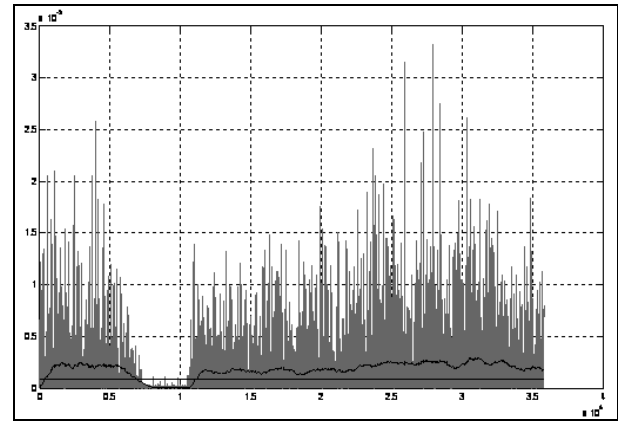


Рис. 4. Нахождение потерянных сегментов по энергии остаточного шума

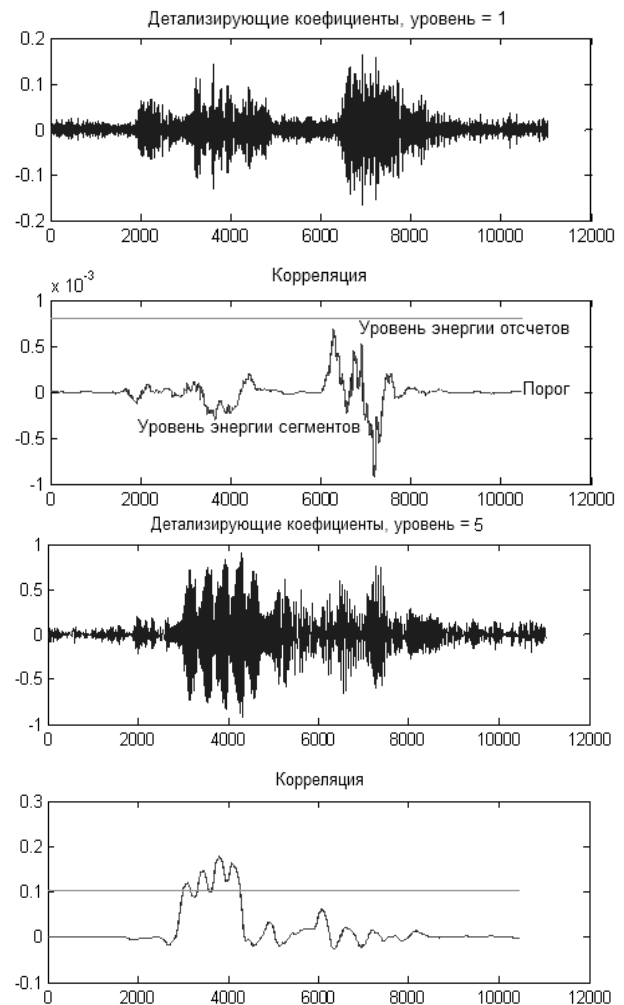


Рис. 5. Нахождение потерянных сегментов по детализирующим коэффициентам вейвлет-разложения

Выводы

В статье разработан метод обнаружения интервалов потерь речевых пакетов, передаваемых по пакетным сетям с использованием кодеков с подавлением пауз речи (VAD) и компенсацией потерь (PLC), позволяющего определить расположение и длительности указанных интервалов.

Разработанный метод включает следующие операции: удаление постоянной составляющей; деконпозиция речевого сигнала с помощью вейвлета Мейера; построение автокорреляционной функции детализирующих коэффициентов; вычисление порога обнаружения потерянных сегментов; разбиение отсчетов автокорреляционной функции на участки по размеру окна анализа; вычисление энергии отсчетов автокорреляционной функции каждого участка сигнала; определение местоположения потерянных сегментов.

Основным отличием данного метода от известных является возможность получения оценки в реальном масштабе времени, без использования передачи заранее сохраненного (тестового) речевого фрагмента.

Для этого предлагается использовать свойства характерного искажения речевого сигнала при наличии однократных или групповых потерь. В качестве базовой, был выбран метод, основанный на вейвлет-анализе речевых сигналов.

Проведенные экспериментальные исследования показали работоспособность разработанного метода в условиях характерных для современных телекоммуникационных сетей.

Дальнейшие исследования целесообразно сконцентрировать на повышении адекватности

оценки качества передачи речевых сигналов в условиях ненулевой вероятности наличия группировки потерь пакетов.

Список литературы

1. Олифер В.Г. Компьютерные сети. Принципы, технологии, протоколы / В.Г. Олифер, Н.А. Олифер. – М.: Питер, 2007. – 520 с.
2. Назаров М.В. Методы цифровой обработки и передачи речевых сигналов / М.В. Назаров, Ю.Н. Прохоров – М.: Радио и связь, 2005. – 333 с.
3. Шелухин О.И. Цифровая обработка и передача речи / О.И. Шелухин, Н.Ф. Лукьянцев. – М.: Радио и связь, 2000. – 251 с.
4. Коллинс Д. IP telephony / Д. Коллинс. – К.: Издательство «Мак Гроу Хилл», 2007. – 736 с.
5. Hoffman D. RTP Payload Format for MPEG1/MPEG2 Video, RFC 2250 / D. Hoffman, G. Fernando, V. Goyal, M. Civanlar. – January, 1998.
6. Горохин В.Н. Передача речи по ip-сетям. / В.Н. Горохин, Д.К. Шакуров. – М.: Питер, 2009. – 220 с.
7. ITU Recommendation G.723.1. Dual Rate Speech Coder for Multimedia Communications Transmitting at 5.3 & 6.3 kbit/s, March 1996.
8. Meyer, Y. (1993), Wavelets: Algorithms and Applications, Society for Industrial and Applied Mathematics, Philadelphia.
9. Короновский А.А. Непрерывный вейвлетный анализ и его приложения. / А.А. Короновский, А.Е. Храмов – М.: Физматлит, 2003. – 350 с.
10. Брандт З. Анализ данных: Статистические и вычислительные методы для научных работников и инженеров / пер. с англ. – М.: Мир, АСТ, 2003. – 686 с.

Поступила в редколлегию 25.05.2011

Рецензент: д-р техн. наук, проф. В.А. Поповский, Харьковский национальный университет радиоэлектроники, Харьков.

ВИЯВЛЕННЯ ІНТЕРВАЛІВ УГРУПОВАННЯ ВТРАТ МОВНИХ ПАКЕТІВ ЗА ПАКЕТНИМИ МЕРЕЖАМИ З ВИКОРИСТАННЯМ КОДЕКА ІЗ ПРИДУШЕННЯМ ПАУЗ МОВИ ТА КОМПЕНСАЦІЄЮ ВТРАТ

С.М. Бобрицький, А.В. Колтиков, В.Н. Ткачов

Удосконалено метод виявлення інтервалів угруповання втрат мовних пакетів по пакетним мережам з використанням кодеків із придушенням пауз мови і компенсацією втрат, заснований на аналізі фрагментів мовного трафіку в вейвлет-базисі Мейєра, що дозволяє визначити розташування і тривалість зазначених угруповань та виключити їх з аналізу мовного трафіку на контентному і пакетному рівнях.

Ключові слова: пакетні мережі, форманти мовних сигналів, вейвлет-базис Мейєра, пакетний рівень, контентний рівень.

DETECTION OF GROUPING INTERVALS LOSSES VOICE PACKETS OVER PACKET NETWORKS USING CODECS WITH SILENCE SUPPRESSION OF SPEECH AND LOSS COMPENSATION

S.M. Bobritskiy, A.V. Kolytkov, V.N. Tkachov

Improved method for detecting grouping intervals of loss of speech packets in a packet cross-tyam using codecs with silence suppression of speech and loss compensation based on the analysis of fragments of voice traffic in the wavelet basis Meier, allowing to determine the location and duration of these groups and exclude them from the analysis of voice traffic for content and batch levels.

Keywords: package networks, formant of linguistic signals, vejvlet-basis of Meyera, package level, contextniy level.