

УДК 621.39

ЛОЗИНСКАЯ В.Н., доцент кафедры «Автоматика и телекоммуникации» (ДОННТУ);
ЯРЕМКО И.Н., доцент кафедры «Автоматика и телекоммуникации» (ДОННТУ);
МОРОЗ Е.Ю., магистрант

Усовершенствование алгоритма низкоскоростного кодирования

Loznskaya V.N., assistant professor of "Automatics and telecommunication" department (DNTU);
Yaremko I.N., assistant professor of "Automatics and telecommunication" department (DNTU);
Moroz E.Yu., graduate student

Low-speed coding algorithm development

Постановка проблемы

Конечный пользователь IP-телефонии не только сохраняет явные преимущества телефонной сети общего пользования, но и экономит средства за счет использования единой сети (снижение стоимости международных и междугородних звонков). Основным отличием от традиционной телефонии является форма сигналов. Для этого голос кодируется в последовательность символов (нулей и единиц) и пересылается к вызываемому абоненту. На приемной стороне, у вызываемого абонента процесс происходит в обратном порядке, т.е. происходит декодирование. Протокол IP «разбивает» закодированный сигнал на пакеты, которые передаются к пункту назначения различными маршрутами.

Очевидно, что в зависимости от типа транспортного протокола (TCP, UDP, RTP) могут возникать потери пакетов. Как итог, на приемной стороне, у вызываемого абонента, может возникнуть понижение качества голоса. Другим источником ухудшения качества голоса, при передаче кодированного сигнала, является компрессия (сжатие) голоса. Последний

процесс необходим из-за ограниченности пропускной способности канала передачи данных, обычно она составляет 64 кБит/с.

Для VoIP телефонии наиболее критичными являются значения параметров задержки пакетов и ее джиттера. В ходе разработки алгоритма кодирования речевых сигналов, необходимо найти компромисс между качеством речи и пропускной способностью канала.

Цель работы

Целью работы является уменьшение объема использования канального ресурса за счет усовершенствования алгоритма кодирования с обеспечением требуемого параметра обслуживания.

Основная часть

Для успешной передачи речи в IP-сети, необходимо осуществить ряд физических преобразований в канале связи. Эти преобразования выполняют специальные элементы сети, среди которых особое место занимают кодеки. Кодек – устройство, а обычно

программа, преобразующее данные или сигнал. Кодеки часто используются при цифровой обработке видео и звука. Данное устройство необходимо по той причине, что передача речи в несжатом виде занимает слишком высокую пропускную способность сети и, кроме этого, не дает никаких непосредственных преимуществ по сравнению с телефонными сигналами, поэтому в функциональной схеме VoIP-шлюза есть так называемый вокодер или кодер речи. Он совершает компрессию речи перед передачей в сеть IP и ее декомпрессию на приемной стороне. В общем случае, при сжатии речи ее первичное звучание (аналоговая форма) может быть воссоздано лишь приблизительно, поэтому при сильном сжатии качество будет значительно хуже, независимо от других факторов. Аналогично ухудшается качество речи и при уменьшении пропускной способности канала. Однако для такого случая существуют различные алгоритмы сжатия, которые имеют возможность адаптироваться.

Кодеки, как часть системы IP-телефонии, участвуют в процессе обеспечения качества голосовой информации при ее передаче через IP-сеть. Для обеспечения соответствующего качества речи необходимо учитывать такие параметры:

- качество передачи речи после согласования кодеками изменяется в зависимости от пропускной способности канала. Если у разных абонентов разные типы кодеков, то при установлении соединения используется «наихудший» из них, т.е. с наименьшей скоростью кодирования;
- чем меньше интервал передачи пакетов, тем выше качество речевой информации;
- чем выше качество речи, обрабатываемой средствами конечного

устройства VoIP, тем больший объем пропускной способности необходим для передачи голосового трафика.

Кроме обеспечения качества речи за счет применения алгоритмов сжатия, особое внимание стоит уделить проблемам транспортировки по IP-сети. К основным параметрам, позволяющим оценить качество транспортировки речи по IP-сети, относят значения [1-4]:

- пропускной способности сети;
- задержки передачи пакетов;
- джиттера передачи пакетов;
- потери речевых пакетов;
- готовности сети;
- акустического эха.

Среди наиболее используемых кодеков в IP-телефонии стоит отметить такие: G.711, G.722, G.722.2, G.723.1, G.726, G.727, G.728, G.729, RPE-LTP и iLBC. Каждый из перечисленных кодеков осуществляет преобразование информации на скоростях от 5,3 Кбит/с до 64 Кбит/с. Для перечисленных кодеков проведен эксперимент: при передаче голоса через канал IP-сети: при уменьшении полосы пропускания канала и проводили оценку стойкости кодека к ухудшению условий передачи информации. Результаты исследования [5] представлены в виде графической зависимости на рисунке 1.

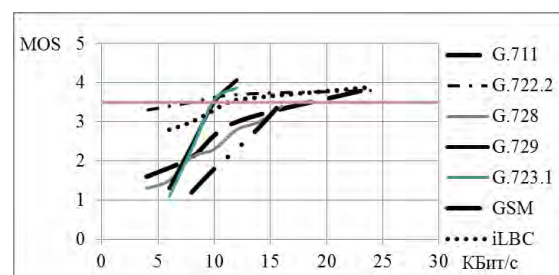


Рис. 1. – Результаты оценки MOS в зависимости от пропускной способности каналов

Исходя из представленной графической зависимости, можно

сделать вывод, что, начиная со скорости 10 Кбит/с, приемлемое качество кодирования (MOS>3,5) достигается моделями iLBC, G.729 и G.722.2. Эти модели кодеков относятся к гибридным кодерам. Алгоритмы данного класса занимают промежуточное значение между кодерами формы сигнала, в которых сохраняется форма колебания речевого сигнала в процессе его дискретизации и квантования, и параметрическими вокодерами, основанными на процедурах оценки и кодирования небольшого числа параметров речи, соединяют преимущества каждого из них.

В зависимости от способа описания поступающего на вход фильтра сигнала можно выделить следующие гибридные алгоритмы кодирования:

- с нарушением прореженной последовательности импульсов (MPLP);
- с самовозбуждением (SELP);
- с кодовым возбуждением (CELP).

Несмотря на то, что CELP кодеки имеют только приемлемые уровни оценок MOS (3.9) и DRT (не менее 95%), данные кодеки являются оптимальными для усовершенствования. Для этого рассмотрим основные этапы кодирования для данного класса алгоритмов (рисунок 2).



Рис. 2. Структурная схема кодера

Отметим, что достоинством CELP кодеков является отсутствие кодирования сигналов отсчет за отсчетом. Для кодирования разностного сигнала применяется «кодовая книга

возбуждений»: каждому разностному сигналу ставится строка (шаблон) сигналов. Эта строка содержит набор отсчетов, соответствующих переданной избыточной последовательности на основе полученного значения ошибки. На приемном конце вместо декодера разностного сигнала также используется «кодовая книга возбуждений». В качестве линейного предсказания (LPC) обычно используется авторегрессионная модель n -го порядка, а в качестве квантования - скалярное.

Для получения передаточной функции ФФ, которая наиболее точно отображает частотную характеристику голосового тракта для звука, стоит определить коэффициенты передаточной функции $\{a_k\}$ (LSF-коэффициенты), исходя из условий наименьшей ошибки линейного предсказания речевого сигнала (из условия минимума квадрата средней ошибки). Выражение для оценки дисперсии сигнала ошибки, которую необходимо свести к минимуму (1):

$$\hat{\sigma}_e^2 = \frac{1}{N} \cdot \sum_{n=1}^N (x_n - \hat{x}_n)^2 \rightarrow \min, \quad (1)$$

$$s^2 = \sum_{n=1}^N (x_n - \hat{x}_n)^2 = \sum_{n=1}^N \left(x_n - \sum_{k=1}^M a_k \cdot x_{n-k} \right)^2, \\ s^2 \rightarrow \min,$$

где, x_n – текущий отсчет источника;

\hat{x}_n – предсказанное значение x_n ;

a_k – коэффициенты предсказания.

После определенного преобразования получается выражение, которое имеет структуру кратковременной ненормированной автокорреляционной функции (АКФ), но зависимой не только от относительного смещения последовательности i , но и от

положения этих последовательностей в сегменте сигнала, которые обозначаются индексом k , который входит в границы суммирования (2).

$$\varphi(k, m) = \sum_{j=M+1-k}^{N-k} x_j \cdot x_{j+i}, \quad (2)$$

где N – количество отсчетов в сегменте сигнала;

M – количество отсчетов, необходимых для расчета коэффициентов предсказания $(M + 1)$ -го отсчета;

x_j – текущий отсчет источника.

При использовании этого метода получаем смещенные оценки коэффициентов линейного предсказания. После дальнейшего преобразования выражение (2) получается в виде системы линейных уравнений, которые можно решить с помощью алгоритма Левинсона-Дарбина, который требует меньших затрат, чем стандартные алгоритмы.

Далее рассмотрим случай не скалярного, а векторного преобразования, его алгоритм выглядит следующим образом: входной вектор \bar{s}_i сравнивается с каждым вектором из кодовой книги. Входной вектор \bar{s}_i (речевой сигнал на входе кодера) представляет собой вектор признаков речевого сигнала (например, спектральных):

$$\bar{s}_i = [s_i(0), s_i(1), \dots, s_i(N-1)]^T. \quad (3)$$

Кодер отображает входной вектор \bar{s}_i в выходной символ $u_n, n=1,2,\dots,L$ с помощью кодовой книги. В результате из кодовой книги выбирается вектор \hat{s}_i , ближайший к вектору \bar{s}_i и в канал

передается u_n , представляющий адрес найденного кодового вектора. В такой интерпретации векторное квантование, по сути, это распознавание образов, где вектор \bar{s}_i представляет собой входной образ, кодовая книга соответствует базе эталонов. Метод поиска по кодовой книге использует ограничение шума, полученное от LPC-фильтров, где основным критерием выбора является минимизация квадратичной ошибки (2) между входным вектором и кодовыми векторами. Каждый вектор в кодовой книге берется из одной секции. Первая секция заполнена уже закодированными разностными векторами. Векторы в другой секции кодовой книги собираются из определенных линейных комбинаций векторов, хранящихся в первой секции.

В кодовой книге находится L векторов

$$\hat{s}_n = [\hat{s}_n(0), \hat{s}_n(1), \dots, \hat{s}_n(N-1)]^T, \quad (4)$$

где $n=1,2,\dots,L$.

Кодовая книга создается с помощью обучающего алгоритма, в соответствии с которым положение центроидов на каждом шаге уточняется в соответствии рекуррентной формуле:

$$\bar{s}_n(t+1) = \bar{s}_n(t) + \alpha(t) \cdot [\bar{s}_i(t) - \bar{s}_n(t)], \quad (5)$$

где t – номер шага;

α – коэффициент обучения, $\alpha \approx 1/\sqrt{t}$.

Выражение (5) уточняет местоположение только того центроида, для которого входной вектор \bar{s}_i является ближайшим. И соответствует правилу обучения нейронных сетей, в частности, правилу Кохонена [6].

После этого происходит формирование пакета для передачи по IP-сетям. Из-за добавления соответствующих заголовков размер пакета существенно увеличивается. При этом в канал связи поступает следующая информация; LSF-коэффициенты, описание сегмента основного состояния и индексы кодовой книги.

Использование принципа векторного квантования позволяет уменьшить количество бит для описания текущего состояния. Итак, вектор текущего состояния разбивается на 19 подвекторов, каждый из которых квантуется кодовой книгой, в результате чего, вместо 171 бита описания текущего состояния необходимо передать $19 \cdot 6 = 114$ бит. Таким образом, использование векторного квантования вместо скалярного снижает суммарную скорость цифрового потока вокодера до 247 бит для каждого кадра длиной 20 мс, что составляет фиксированную битовую скорость 12,35 кбит/с, что ниже, чем для режима с кадром, длиной в 30 мс.

Оценим объем полезной нагрузки в канале связи. Он зависит от размера голосового семпла (звукового файла), который непосредственно влияет на необходимую полосу пропускания. Стандартный размер кадра представленного кодека при скорости кодирования 12,35 кбит/с составляет 20 мс.

Размер семпла в байтах рассчитаем по формуле:

$$B_{sample} = \frac{t_{sample} \cdot B_{codek}}{8}, \quad (6)$$

где B_{sample} – размер семпла в байтах;

t_{sample} – размер семпла в секундах;

B_{codek} – битрейт кодека.

Для вычисления полосы пропускания, занимаемой одним звонком, канала, при использовании разработанного кодека:

$$B_1 = \frac{B_{sample} + B_{l2} + B_{ip,rtп,udp}}{8} \cdot B_{codek}, \quad (7)$$

где B_1 – полоса пропускания, необходимая для одного звонка;

B_{l2} – объем служебной информации протокола канального уровня в байтах. Размер служебной информации при использовании Ethernet составляет 18 байт.

$B_{ip,rtп,udp}$ – заголовки IP, UDP, RTP без компрессии в байтах (20,8,12) соответственно.

При подстановки всех параметров в (7) получаем, что размер семпла в байтах составляет 31, максимальная полоса пропускания канала, необходимая для осуществления одного звонка, протокол IPv4 – 35,5 кбит/с, для версии IPv6 – 40,2 кбит/с. Таким образом, представленный кодек позволяет уменьшить размер IP пакетов, что способствует уменьшению необходимой полосы пропускания канала для осуществления одного звонка.

Выводы

В связи с необходимостью преобразования голоса для его дальнейшей передачи по IP-сетям разработка и усовершенствование существующих алгоритмов кодирования становится актуальной проблемой. Анализ кодеков IP-телефонии показала, что наиболее эффективными для этих целей являются гибридные кодеры: CELP. Для усовершенствования существующего

алгоритма предложено использовать следующие модификации: использование определенных моделей для линейного предсказания, использование векторного квантования и обучаемой кодовой книги. Вышеперечисленные меры позволяют уменьшить размер формируемых пакетов, что способствует уменьшению необходимого объема канального ресурса для совершения одного телефонного звонка.

Список литературы:

1. Liao, J.-C. L. Wanjiun. Voip mobility in ip/cellular network interworking/ J.-C. L. Wanjiun Liao. IEEE Communications Magazine, – 2000.
2. Klepec, A. K. Borut. Performance of voip applications in a simple differentiated services network architecture/ A. K. Borut Klepec. IEEE, – 2001.
3. Goode, B. Voice over internet protocol/ B. Goode. Proc.of IEEE. – 2002.
4. M. X. Z. Wang, Mai. Implementation of qos provisioning system/Mai M. X. Z. Wang. Proc.of IEEE Real-time and Embedded Technology and Applications Symposium, – 2002.
5. Максимов, М.И. Проектирование низкоскоростных речепреобразующих устройств для каналов с высоким процентом ошибок/

Максимов М.И., Н.А. Сидорова, О.В. Чернойров // Электросвязь. – 2008. – №7. – с. 48-50.

6. Горбаченко, В.И. Сети и карты Кохонена [Электронный ресурс]/ В.И. Горбаченко. – Режим доступа: http://gorbachenko.self-organization.ru/articles/Self-organizing_map.pdf.

Аннотации:

В работе представлен анализ алгоритмов низкоскоростных кодеков, используемых в IP-телефонии. Оптимальными, с точки зрения возможности усовершенствования, являются CELP кодеки. Предложенные рекомендации по усовершенствованию кодека позволяют уменьшить размер формируемых пакетов, что способствует уменьшению необходимого объема канального ресурса.

Ключевые слова: вокодер, CELP, автокорреляционная функция, векторное квантование, обучаемая кодовая книга, полезная нагрузка в канале связи.

The article is coding algorithms analyzes used in IP-phone is given in the work. With ability of modernization, the optimal are CELP coders. Given recommendations with coder's modernization let decrease the size of made packets. This contribute decreasing of needed bandwidth.

Keywords: vocoder, CELP, autocorrelation function, vector quantization, learning codebook, communication channel payload.