

АНАЛИЗ АЛГОРИТМОВ НИЗКОСКОРОСТНОГО КОДИРОВАНИЯ

Мороз Е.Ю., магистрант

(ГОУ ВПО « Донецкий национальный технический университет », г. Донецк, ДНР)

Технология передачи голосовой информации по цифровым каналам с использованием протокола IP (Voice over IP) позволяет конечному пользователю получать не только качественную голосовую связь по каналам передачи данных, а и возможность получать дополнительные услуги. При этом конечный пользователь IP-телефонии не только сохраняет явные преимущества телефонной сети общего пользования, а и получает следующие:

- IP-телефония поддерживает голос и данные одновременно, из-за чего получают дополнительную экономию средств за счет использования единой сети;
- снижение стоимости международных и междугородних звонков;
- доступ к новым услугам.

Основным отличием от традиционной телефонии является форма сигналов, которые передаются по сети, для VoIP телефонии эти сигналы цифровые. 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. Кроме этого, 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 понижения качества голоса.

Еще одним 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. Обычно, в ходе нахождения оптимального алгоритма кодирования речевых сигналов, необходимо найти компромисс между качеством речи и пропускной способностью канала. Кроме искажений, вносимых в ходе преобразования речи, при транспортировке последней через сеть передачи данных на качество голоса также влияют показатели качества обслуживания. Для VoIP телефонии наиболее критичными являются значения параметров задержки пакетов и ее джиттера [1]. Целью представленной работы является выбор алгоритма кодирования среди существующих усовершенствование которого не изменит значение показателей качества обслуживания.

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 [2].

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:

- пропускной способности сети;

- задержки передачи пакетов;

- джиттера передачи пакетов;

- потери речевых пакетов;

- готовности сети;

- акустического эха.

Кодеры формы [3] характеризуются способностью сохранять основную форму речевого сигнала. Сохраняя огибающую форму сигнала, подобные кодеры работают по принципу выборки. Их характеристики эффективно измеряются соотношением сигнал/шум, потому что квантование является основным источником искажений формы выходного сигнала.

Импульсно-кодовая модуляция (PCM) используется в рекомендации G.711.

При построении систем цифровой передачи принципиальным моментом является определение полосы частот, необходимой для обеспечения заданного качества воспроизведения переданного сообщения. В общем, для высококачественной передачи речевого сообщения необходима полоса не менее 10 кГц. Однако, для достижения удовлетворения уровня разборчивости при передаче речи достаточно передать спектр в полосе 300..3400 Гц. Вместе с ИКМ используется дифференциальная импульсно-кодовая модуляция (ДИКМ). Например, с целью снижения требований к пропускной способности можно использовать наличие корреляций между отчетными значениями переданного сообщения. При этом последовательность значений $\lambda(h)$ поступает на один вход разностного устройства, в то самое время, когда на другой вход поступает спрогнозированное значение $\bar{\lambda}(h)$, полученное тем или иным методом в устройстве предсказания на основе анализа как предыдущих, так и текущих переданных значений на входе разностного устройства.

Адаптивная дифференциальная импульсно-кодовая модуляция (АДИКМ), рекомендации G.721 и G.726 имеет практически такое же качество восстановления речи, как и ИКМ. Однако для передачи информации при его использовании необходимо всего лишь 32 кбит/с. Метод основан на том, что в аналоговом сигнале, передающим речь, невозможны резкие изменения интенсивности. В АДИКМ изменение уровня сигнала кодируется четырехразрядным числом, при этом частота измерения амплитуды сигнала сохраняется неизменной.

Другой тип алгоритмов сжатия – вокодеры [4]. Они представляют собой устройство, осуществляющее параметрическую компрессию голосовых сигналов. Компрессия речевых сигналов на передаче проводится в анализаторе, который выделяет из речевого сигнала соответствующие составляющие, и передает их в канал связи в виде кодовых сообщений. На приеме с помощью местных источников сигналов, управляемых принятыми параметрами,

синтезируется речевой сигнал. По принципу определения параметров речи различают вокодеры:

- полосовые канальные;
- формантные;
- ортогональные;
- вокодеры с линейным предсказанием;
- гомоморфные.

В полосовых вокодерах спектр речи делится на 7-20 полос (каналов) аналоговыми или цифровыми полосовыми фильтрами. Большая часть каналов в вокодере дает большую разборчивость речи. С каждого полосового фильтра сигнал поступает на детектор и фильтр низких частот с частотой среза F_{cp} . Таким образом, сигналы на выходе каждого канала изменяются с частотой не менее F_{cp} . Их передача возможна в аналоговом или цифровом виде.

В формантных вокодерах огибающая спектра раскладывается в ряд по выбранной системе ортогональных базисных функций. Вычисленные коэффициенты этого разложения передаются на приемную сторону. Расширение получили гармонические вокодеры, которые используют разложение в ряд Фурье.

Энергия речи может концентрироваться в трех-четыре пиках – формантах. Формантный вокодер определяет и передает положение пика энергии в частотном диапазоне, амплитуду спектральных пиков. В следствие этого снижается объем переданной информации. Качество восстановленной речи зависит от точности определения этих параметров. Принцип устройства формантного вокодера основаны схоже с принципом устройства полосового вокодера. Но при этом для передачи отбираются только данные про возбужденные спектры. Это снижает требования к объему переданной информации. Декодер восстанавливает сигнал также с помощью генерации основного тона и сигналов разного типа (шумовых и импульсных). Такой тип вокодера обеспечивает передачу речи со скоростью 1 Кбит/с, но применяется достаточно редко из-за трудностей, связанных с точным вычисление формант.

Ортогональные вокодеры отличаются от полосовых тем, что функции фильтров с помощью цифровых методов. Это или быстрое преобразование Фурье или функции Уолша.

Гомоморфная обработка позволяет разделить генераторную и фильтровую функции, которые составляют речевой сигнал. Такая обработка речевого сигнала приводит к довольно удобному описанию, где основные параметры сигналу отдалены друг от друга, т.е. информация о сигнале возбуждения расположена в области больших времен, а информация о речевом тракте и форме импульса возбуждения – в области малых времен. Гомоморфный вокодер, как и любые вокодерные системы, в которых пытаются разделить параметры речи на сигнал возбуждения и параметры речевого тракта, позволяет достичь малую скорость передачи и дополнительную громкость при обработке речи, ценой усложнения в описании и потерь в качестве.

Вокодеры с линейным предсказанием (LPC) основаны на оригинальном математическом аппарате. Они получили наибольшее распространение и будут рассмотрены в дальнейшем. Рекомендации G.728, G.729, G.723.

LPC работает с блоками отсчетов, для каждого из которых вычисляется и передается частота основного тона, его амплитуда и информация о типе возбуждающего влияния.

Каждый из вышеперечисленных 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.

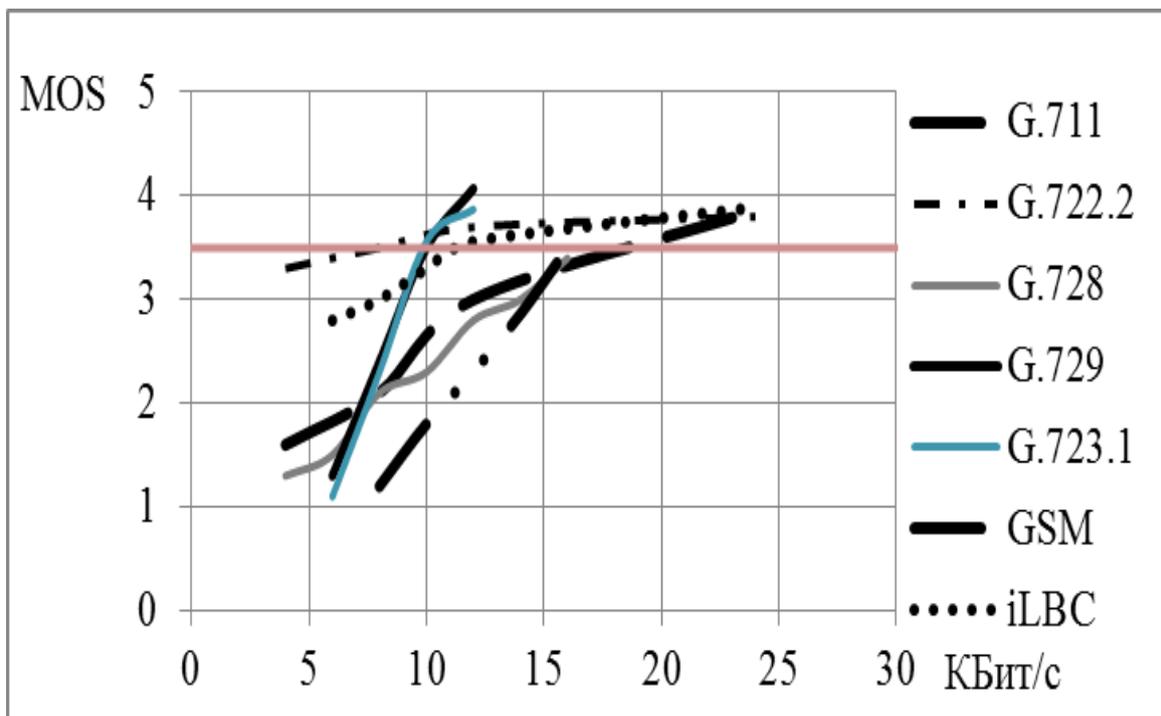


Рисунок 1 – Результаты оценки MOS в зависимости от пропускной способности каналов

Анализ алгоритмов кодирования показал, что наилучшее качество кодирования голоса имеют гибридные кодеки и кодеки формы, хуже данные для вокодеров. Сигнал на выходе кодеков формы занимает достаточно большую полосу пропускания и является мало эффективным для IP-телефонии. Лучший из гибридных кодеков – CELP, однако он требует больше вычислительных ресурсов, а соответственно – увеличивает задержку кодирования информации. Лучший из вокодеров – LPC. Таким образом, оптимальным для разработки между скоростью передачи с предсказанием – LPC.

Анализ методов представления сигналов показал, что оптимальным методом является метод линейного предсказания. Именно такой метод предсказания используется LPC-алгоритмом.

Таким образом, именно этот алгоритм является оптимальным между скоростью передачи и задержкой обработки информации, обусловленной сложностью алгоритма кодирования. Он может обеспечить низкоскоростное кодирование при сохранении хорошего качества голосового трафика.

Перечень ссылок

1. Объективная оценка качества обслуживания при передаче речи в пакетных сетях / [Электронный ресурс]. — Режим доступа : <http://helpiks.org/7-25961.html>.
2. Гольдштейн, Б. С. IP-телефония / Б. С. Гольдштейн. – Москва : Радио и связь, 2001. – 335 с.
3. Кудряшов, Б. Д. Основы теории кодирования: учебное пособие / Б. Д. Кудряшов. – Санкт-Петербург : БХВ-Петербург, 2016. – 400 с.
4. Штарьков, Ю. М. Универсальное кодирование. Теория и алгоритмы / Ю. М. Штарьков. – Москва : ФИЗМАТЛИТ, 2013. – 288 с.
5. Максимов, М. И. Проектирование низкоскоростных речепреобразующих устройств для каналов с высоким процентом ошибок / М. И. Максимов, Н. А. Сидорова, О. В. Черноярков // Электросвязь. – 2008. – №7. – С. 48-50.