

АНАЛИЗ ВЛИЯНИЯ КОДЕКОВ НА КАЧЕСТВО РЕЧИ ПРИ ПЕРЕДАЧЕ В IP-СЕТЯХ

Заводов С.А., Агеев Д.В.

Харьковский национальный университет ра-
диоэлектроники, Украина.

E-mail: zavodov2007@mail.ru

Abstract

Peculiarities of modern stage of development of the society is to strengthen the role of media in all major areas of human activity. Further social development, of course, will be primarily due to the computerization. A number of devices integrated telecommunication networks increases continuously, which leads to an increase in the transmitted data and its diversity. Verbal communication (telephony) remains one of the most popular ways of direct exchange of information between people. In recent years, widely developed telecommunications network, in which all types of information, including voice, are transmitted in digital form. In recent years extensive development received packet network, in which all types of information, including voice, are transmitted in digital form. In implementing the digital transformation of speech signals there are specific distortions affecting speech quality. The process of converting an analog voice signal into digital form is called an analysis or digital encoding of speech, and the reverse recovery process of analog voice signal form - synthesis or decoding of speech. The purpose of any coding scheme - is to obtain a digital sequence that requires a minimum transmission rate and from which a decoder can recover the original speech signal with minimal distortion.

Особенностью современного этапа развития общества является усиление роли информации во всех основных сферах человеческой деятельности. Дальнейшее социальное развитие, безусловно, будет в первую очередь связано с информатизацией.

Количество устройств, объединенных телекоммуникационными сетями, непрерывно возрастает, что приводит к увеличению передаваемой информации и ее разнообразию. Речевая коммуникация (телефония) остается одним из наиболее оперативных и востребованных способов непосредственного обмена информацией между людьми.

В последние годы широкое развитие получили пакетные сети, в которых все виды информации, в том числе речь, передаются в цифровом виде. При реализации цифровых преобразований речевых сигналов возникают специфические искажения, влияющие на качество речи.

Процесс преобразования аналогового речевого сигнала в цифровую форму называется анализом или цифровым кодированием речи, а обратный процесс восстановления аналоговой формы речевого сигнала – синтезом или декодированием речи.

Цель любой схемы кодирования – получить такую цифровую последовательность, которая требует минимальной скорости передачи и из которой декодер может восстановить исходный речевой сигнал с минимальными искажениями.

Скорость передачи, которую предусматривают имеющиеся сегодня узкополосные кодеки, лежит в пределах 1,2 – 64 Кбит/с. Естественно, что от этого параметра прямо зависит качество воспроизводимой речи. Существует множество подходов к проблеме определения качества. Наиболее широко используемый подход оперирует оценкой MOS (Mean Opinion Score), которая определяется для конкретного кодека как средняя оценка качества большой группой слушателей по пятибалльной шкале.

Большинство кодеков, используемых в телекоммуникационных сетях, (G.711, G.723.1, G.726, G.728, G.729) имеют оценку выше 3,5, что соответствует качеству ТфОП. Тем не менее, самая высокая оценка у кодека G.711, который требует самую высокую скорость передачи.

Для снижения требуемой пропускной способности следует применять дополнительные алгоритмы, подавляющие периоды молчания (Voice Activity Detector – VAD, Comfort Noise Generator – CNG, Discontinuous Transmission – DTX). При диалоге один его участник говорит, в среднем, только 35 процентов времени. Таким образом, если применить алгоритмы, которые позволяют уменьшить объем информации, передаваемой в периоды молчания, то можно значительно сузить необходимую полосу пропускания. В двустороннем разговоре такие меры позволяют достичь сокращения объема передаваемой информации до 50%, а в децентрализованных многоадресных конференциях (за счет большего количества говорящих) – и более.

Также на оценку MOS существенно влияют потери пакетов, которые являются неотъемлемым атрибутом IP-сетей. Так как пакеты содержат кадры, сформированные кодеком, то это вызывает потери кадров. Существуют подходы (такие как применение кодов с исправлением ошибок – forward error correction), позволяющие уменьшить число потерянных кадров при данном числе потерянных пакетов. Требуемая для этого дополнительная служебная информация распределяется между несколькими пакетами, так что при потере некоторого числа пакетов кадры могут быть восстановлены. Однако введение избыточности отрицательно сказывается на задержке воспроизведения сигнала. Так как, чтобы обеспечить возможность восстановления информации при потере трех подряд переданных пакетов, то декодер вынужден поддерживать буфер из четырех пакетов, что вносит значительную дополнительную задержку воспроизведения.

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 $N \cdot 125$ 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.

При потере пакетов деградация качества существенно зависит от длины пакета, так как алгоритмы компенсации потерь пакетов (PLC) обычно имеют верхний предел компенсируемой потери в 60 мсек. График ниже иллюстрирует то, что пакетизация 40 мсек существенно начинает ухудшать качество речи при потере более 5% пакетов.

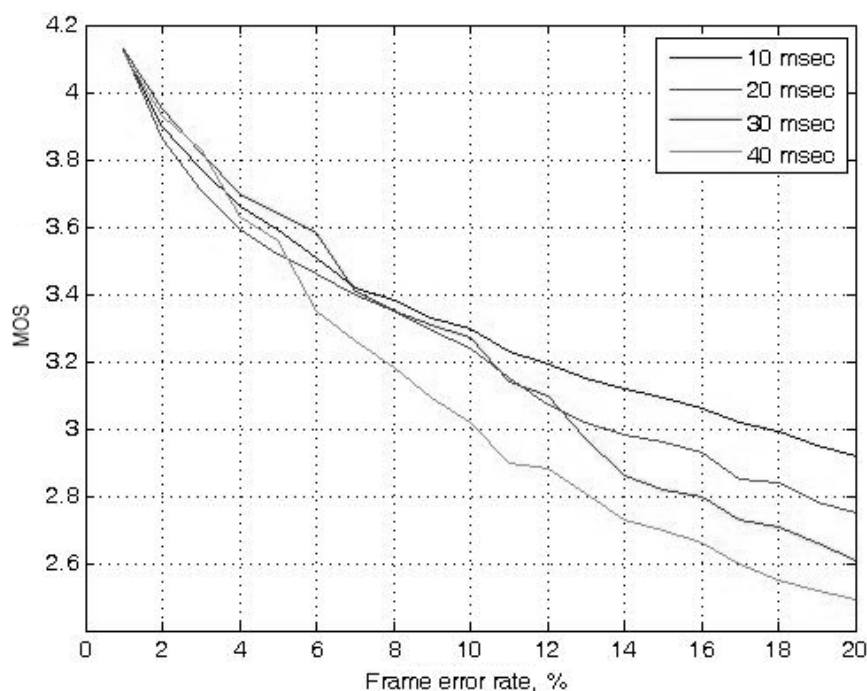


Рис.1. Зависимость оценки качества речи (MOS) от уровня потерь для G.711 PLC при разном уровне пакетизации

Таким образом, для минимального снижения качества передаваемой речи, необходимо исходя из пропускной способности сети правильно подобрать кодек, а так же размер пакета. При этом следует применять дополнительные алгоритмы, подавляющие периоды молчания.

Литература:

1. Гольдштейн Б.С., Пинчук А.В., Суховицкий А.Л. IP+Телефония. — М.: Радио и связь, 2006. — 336 с.
2. ITU-T Recommendation G.113 Transmission impairments due to speech processing — 2007.
3. Бобрицкий С.М. Алгоритм оценки качества передачи речи с анализом характеристик эффективных потерь пакетов на основе вейвлет-преобразования [Электронный ресурс] // Проблемы телекоммуникаций. — 2011. — № 2 (4). — С. 95 – 107. — Режим доступа: http://pt.journal.kh.ua/2011/2/1/112_bobrickij_voice.pdf.