

## Оценка качества передачи речи в сетях IP

**Г.Г. ЯНОВСКИЙ, заведующий кафедрой сетей связи СПбГУТ, доктор технических наук**

**В статье дается характеристика субъективных и объективных подходов к оценке качества обслуживания при передаче речи. Определены недостатки модели MOS в сетях IP и описаны особенности E-модели, применяемой в пакетных сетях. Представлен анализ факторов, влияющих на качество передачи речи в сетях IP.**

При оценке качества услуг в сетях IP необходимо учитывать, что требования к сетевым характеристикам со стороны приложений данных и приложений, связанных с передачей голоса, существенно различаются.

Исходное требование при развертывании приложений VoIP: качество речи при использовании VoIP должно быть таким же, как и в ТфОП. Отметим, что уровень качества в сети ТфОП иногда называется уровнем качества междугородного соединения и является наивысшим уровнем качества доставки речи в сетях электросвязи. Как известно, качество обслуживания определяется набором сетевых параметров, в число которых входят пропускная способность сети, надежность сети/ сетевого оборудования, задержки, вариации задержки (джиттер) и потери пакетов.

До недавнего времени согласованные количественные оценки, определяющие качество передачи речи в сетях связи с учетом того, как это воспринимается пользователем, отсутствовали. Первоначально МСЭ предложил подход (Рекомендации МСЭ Р.800), в основе которого лежали субъективные оценки качества передачи речи (такие, как "отличное качество", "хорошее качество", "приемлемое качество" и т. д.). Субъективные оценки, к сожалению, не могут быть точно соотнесены с сетевыми характеристиками, которые используются при проектировании и эксплуатации сетей. Не могут быть они точно сопоставлены и с процессами, реализуемыми в терминальном оборудовании (т. е. вне сети). Речь идет об алгоритмах сжатия, схемах кодирования, механизмах защиты информации, восстановления данных и т. д. Тем не менее, субъективные оценки использовались в течение многих лет как единственный подход к оценке качества в телефонных сетях и в определенной степени сохраняют свое значение сегодня. В 1998 г. МСЭ стандартизировал подход, основанный на объективных оценках качества обслуживания, который позволяет описать показатели качества при передаче речи в пакетной форме (Рекомендация МСЭ G.107). В статье рассматриваются оба подхода, но основное внимание уделяется анализу Рекомендации G.107.

### Субъективная оценка качества обслуживания при передаче речи

Первичным критерием качества аудио- и видеoinформации является восприятие качества услуги пользователем. Определение качества услуг может базироваться как на субъективных, так и на объективных оценках. Наиболее широко используемая методика субъективной оценки качества описана в Рекомендации МСЭ Р.800 (первоначальная редакция относится к 1993 г.) и известна как методика MOS (Mean Opinion Score). В

## СЕТИ

соответствии с ней качество речи, получаемое при прохождении сигнала от говорящего (источник) через систему связи к слушающему (приемник), оценивается как арифметическое среднее от всех оценок, выставляемых экспертами после прослушивания тестируемого тракта передачи.

Экспертные оценки определяются в соответствии со следующей пятибалльной шкалой: 5 — отлично, 4 — хорошо, 3 — приемлемо, 2 — плохо, 1 — неприемлемо. Оценки 3,5 балла и выше соответствуют стандартному и высокому телефонному качеству, 3,0 — 3,5 — приемлемому качеству, 2,5 — 3,0 — синтезированному звуку. Для передачи речи с хорошим качеством целесообразно ориентироваться на значения MOS не ниже 3,5 баллов.

Хотя методика MOS, основанная на субъективных оценках, является достаточно надежным инструментом в телефонных сетях, в ней отсутствует возможность количественно учесть влияющие на качество речи факторы. В частности, не учитываются:

- сквозная (end-to-end) задержка между говорящим по телефону и слушающим;
- влияние вариации задержки (джиттера);
- влияние потерь пакетов.

## Объективная оценка качества обслуживания при передаче речи в пакетных сетях

Для преодоления указанных недостатков в 1998 г. МСЭ принял Рекомендацию G.107, в которой был описан подход к объективной оценке качества услуг в телекоммуникациях. В его основу положена так называемая *E*-модель, которая открыла новое направление в оценке качества услуг, связанное с измерением характеристик терминалов и сетей. После создания *E*-модели было проведено большое число испытаний, в которых менялся уровень воздействия искажающих сетевых факторов. Данные этих тестов были использованы в *E*-модели для вычисления объективных оценок. Результатом вычислений в соответствии с *E*-моделью является число, называемое *R*-фактором ("коэффициентом рейтинга"). Значения *R*-фактора однозначно сопоставляются с оценками MOS (см. табл. 1 и рис. 1).

Таблица 1

Оценка QoS на основе R-фактора и оценок MOS		
Значение R-фактора	Категория качества и оценка пользователя	Значение оценки MOS
$90 < R < 100$	Самая высокая (отлично)	4,34 — 4,50
$80 < R < 90$	Высокая (хорошо)	4,03 — 4,34
$70 < R < 80$	Средняя (приемлемо: часть пользователей оценивает качество как неудовлетворительное)	3,60 — 4,03
$60 < R < 70$	Низкая (плохо: большинство пользователей оценивает качество как неудовлетворительное)	3,10 — 3,60
$50 < R < 60$	Неприемлемая (не рекомендуется)	2,58 — 3,10

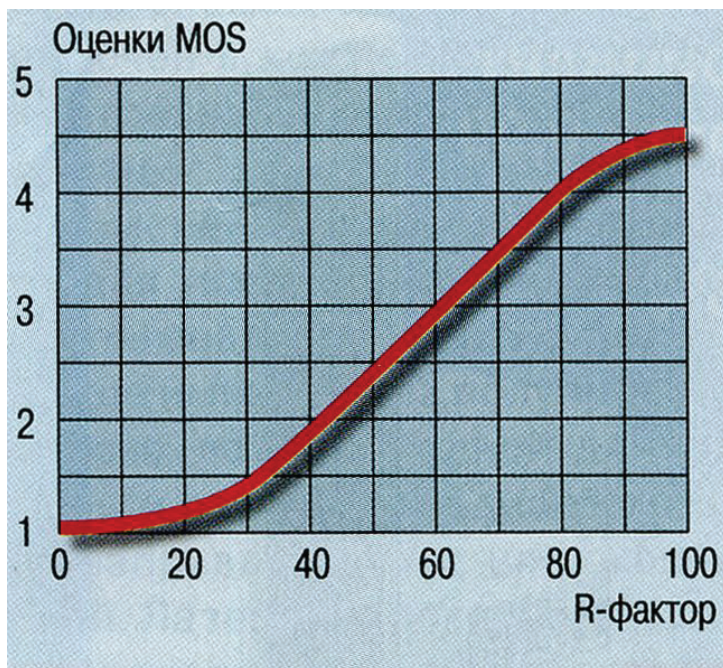


Рис. 1. Зависимость между оценками MOS и R-фактором

В соответствии с *E*-моделью *R*-фактор определяется в диапазоне значений от 0 до 100, где 100 соответствует самому высокому уровню качества. При расчете *R*-фактора учитываются 20 параметров, в числе которых:

- однонаправленная задержка;
- коэффициент потери пакетов;
- потери данных из-за переполнения буфера джиттера;
- искажения, вносимые при преобразовании аналогового сигнала в цифровой и последующем сжатии (обработка сигнала в кодеках);
- влияние эхо и др.

Таким образом, *E*-модель и *R*-фактор могут быть использованы для объективной оценки качества передачи речи в технологии VoIP. Как только *R*-фактор получен, могут быть вычислены соответствующие оценки MOS. Вычисление *R*-фактора начинается для случая, когда искажения сигнала в канале не учитываются, а принимаются во внимание искажения, которые имеют место при преобразовании реальной речи в электрический сигнал (и обратно). Теоретическое значение *R*-фактора уменьшается от 100 до 93,2, что соответствует оценке MOS, равной 4,4. Таким образом, при использовании *E*-модели оценка 4,4 в системе MOS является максимально возможной оценкой качества речи в сети без искажений. Величина *R*-фактора меняется от 0 до 93,2, что соответствует изменению оценок MOS от 1 до 4,4. Значение *R*-фактора определяется по следующей формуле:

$$R = R_o - I_s - I_d - I_e + A,$$

где:  $R_o = 93,2$  — исходное значение *R*-фактора;

$I_s$  — искажения, вносимые кодеками и шумами в канале;

$I_d$  — искажения за счет суммарной сквозной задержки ("из конца в конец") в сети;

$I_e$  — искажения, вносимые оборудованием, включая и потери пакетов;

$A$  — так называемый фактор преимущества.

Например, мобильные пользователи могут соглашаться с низким уровнем качества, получая дополнительные удобства. В большинстве случаев расчета  $R$ -фактора параметр  $A$  принимается равным нулю.

## Анализ факторов, влияющих на качество речи в пакетных сетях

**Влияние кодеков на качество пакетизированной речи.** При расчете  $R$ -фактора одна из составляющих —  $ls$ , уменьшающая значение  $R$ -фактора, определяется искажениями, возникающими в кодеке при пакетизации речевого сигнала. Качество передачи речи в сетях с коммутацией пакетов в последние годы было значительно улучшено путем создания эффективных кодеков, обеспечивающих хорошую разборчивость речевого сигнала на приемном конце. В состав этих методов входят:

- методы эффективного кодирования речи (рекомендации МСЭ-Т серии G.7xx);
- механизмы подавления пауз (механизм кодирования речи при прерывистой передаче, известный как Voice Activity Detection, VAD);
- механизмы эхоподавления (рекомендация МСЭ G.164) и эхокомпенсации (рекомендации МСЭ G.165 и G.168);
- механизмы маскирования ошибок (packet loss concealment), обеспечивающие компенсацию пробелов в речевом потоке, вызванных потерей отдельных пакетов.

**Характеристики речевых кодеков.** Сегодня имеется большой набор эффективных кодеков с различными характеристиками. В табл. 2 представлены характеристики кодеков, соответствующих стандартам МСЭ-Т. Исторически первый тип кодека, известный как G.711 (версии G.711a и G.711u, скорость выходного сигнала 64 кбит/с), преобразует аналоговый сигнал в цифровой с очень высоким качеством без применения операции сжатия. Однако, при этом требуется значительная пропускная способность по сравнению с кодеками, в которых осуществляется сжатие информации. При создании первых кодеков (70-е гг.) технология современных цифровых сигнальных процессоров (DSP) была недоступна. Сегодня на базе DSP возможно построить весьма эффективные кодеки со значительно меньшими требованиями к пропускной способности тракта передачи.

Таблица 2

Типы речевых кодеков и их характеристики

Кодек	Скорость передачи, кбит/с	Длительность датаграммы, мс	Задержка пакетизации, мс	Полоса пропускания для двунаправленного соединения, кГц	Задержка в буфере джиттера	Теоретическая максимальная оценка MOS
G.711u	64	20	1	174,4	2 дейтаграммы, 40 мс	4,4
G.711a	64	20	1	174,4	2 дейтаграммы, 40 мс	4,4
G.726-32	32	20	1	110,4	2 дейтаграммы, 40 мс	4,22
G.729	8	20	25	62,4	2 дейтаграммы, 40 мс	4,07
G.723m	6,3	30	67,5	43,73	2 дейтаграммы, 60 мс	3,87
G.723a	5,3	30	67,5	41,6	2 дейтаграммы, 60 мс	3,69

## СЕТИ

Низкоскоростные кодеки требуют существенно меньших значений пропускных способностей, однако оказывают значительно большее влияние на качество речевого сигнала по сравнению с высокоскоростными кодеками.

Меньшая пропускная способность означает, что можно организовать большее число телефонных соединений по одному и тому же тракту, но при этом уменьшается разборчивость речи, возрастают задержки и качество речи становится более чувствительным к потере пакетов.

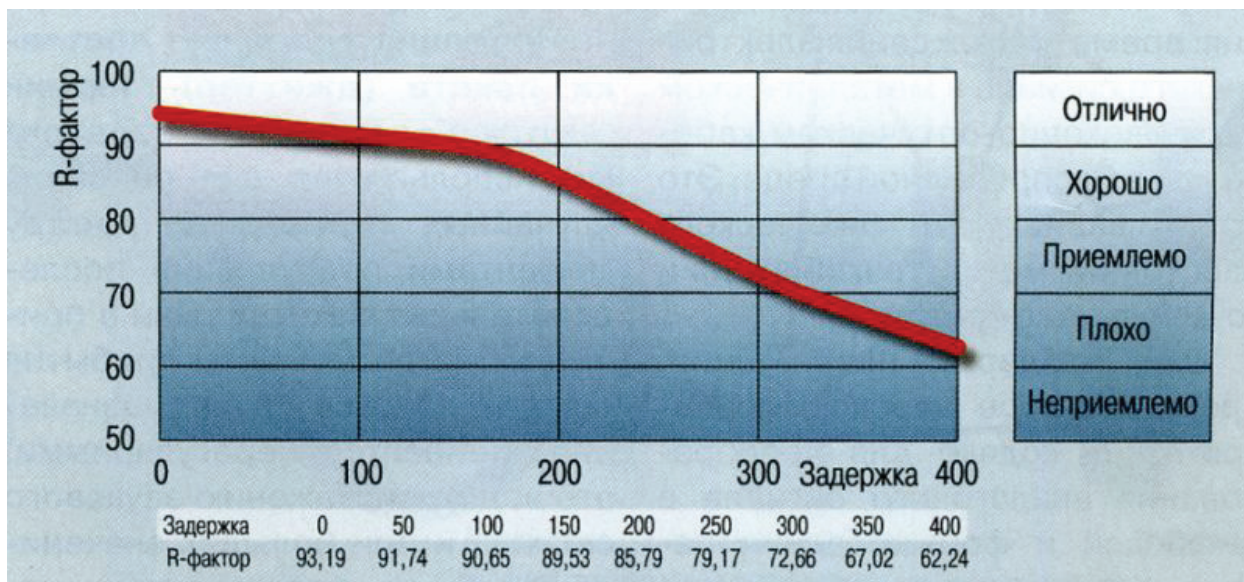


Рис. 2. Влияние суммарной задержки на оценки МОС и величину R-фактора

**Задержки и джиттер в сетях IP.** Задержка доставки пакета определяется временем переноса пакета от источника до получателя. Время задержки меняется в зависимости от трафика в сети и доступных сетевых ресурсов, в частности, пропускной способности, во время доставки. Речь представляет собой трафик, чувствительный к задержке, тогда как большинство приложений данных относительно устойчиво к задержке. Если задержка доставки пакета превышает определенное значение, пакет отбрасывается. В результате, при большом числе отброшенных пакетов качество речи ухудшается, что и отражено в приведенной выше формуле для R-фактора, где влияние задержки учтено через составляющую ID.

Естественным является вопрос, какая задержка допустима при пакетной передаче речи? В результате исследований качества речевого сигнала еще в 60-х годах прошлого столетия было установлено, что человек начинает чувствовать задержки речевого сигнала, превышающие 150 мс, и ощущает заметный дискомфорт, если задержка превышает 250 мс. Позднее, при поддержке МСЭ, были проведены масштабные исследования влияния сетевой задержки на качество телефонного разговора. Эти результаты нашли отражение в Рекомендации МСЭ G.114, в соответствии с которой рекомендуемый порог задержки при передаче речи равен 150 мс. При задержке 300 мс разговор распадается на фрагменты, которые невозможно связать в слитную речь.

Рассмотрим, какие факторы определяют суммарную величину задержки доставки пакета. Сквозная задержка доставки пакета  $D_d$  ("из конца в конец") определяется как сумма четырех составляющих:

## СЕТИ

$$D_{\text{д}} = D_p + D_{\text{нк}} + D_{\text{нн}} + D_{\text{бд}},$$

где:

$D_p$  — задержка распространения: время прохождения электрического сигнала в металлическом или волоконно-оптическом кабеле или в беспроводной среде. Это время зависит от физического расстояния между точкой входа и точкой выхода из сети.

$D_{\text{нк}}$  — задержка пакетизации: время, которое необходимо затратить в кодеке для преобразования аналогового сигнала в цифровой и формирования пакета. Как видно из табл. 2, чем ниже скорость сигнала на выходе кодека, тем выше задержка пакетизации, поскольку кодек тратит больше времени на процессы компрессии и декомпрессии сигнала;

$D_{\text{нн}}$  — задержка переноса пакета: время прохождения пакета через все устройства сети, расположенные вдоль пути передачи пакета, включая маршрутизаторы, шлюзы, сетевые экраны, обработчики трафика, сегменты сети с относительно малой пропускной способностью в условиях перегрузки и т. д. Для некоторых устройств, например, синхронных мультиплексоров, эта величина постоянна, для других, таких, как маршрутизаторы, задержка переноса меняется с изменением нагрузки в сети;

$D_{\text{бд}}$  — задержка на приемной стороне в буфере джиттера: буфер джиттера используется для уменьшения вариаций между моментами поступления пакетов на вход приемного устройства. Буфер может накапливать от одной до нескольких датаграмм. В соответствии с данными табл. 2 типичный буфер джиттера накапливает две датаграммы и задержка  $D_{\text{бд}}$  составляет от 20 до 30 мс в зависимости от типа кодека.

Очевидно, что задержка распространения, задержки в кодеке и буфере джиттера являются постоянными величинами для выбранного пути передачи пакета, тогда как задержка переноса — случайная величина, зависящая от условий в сети в конкретный момент времени. На рис. 2 показано, как задержки влияют на R-фактор и показатели МОС.

**Вариация задержки доставки пакета (джиттер).** Термин "джиттер" (вариация задержки) используется для описания случайных изменений между моментами поступления последовательных пакетов речи в приемник. Когда моменты прибытия речевых пакетов в пункт назначения становятся нерегулярными, это ведет к искажению звукового сигнала и при больших значениях джиттера, превышающих несколько десятков мс, речь становится неразборчивой.

Потери пакетов определяются как процент пакетов, не доставленных к месту назначения. В сети имеется ряд причин потери пакетов, среди которых отметим следующие:

- при перегрузке сети очереди в коммутаторах и маршрутизаторах быстро растут. Если перегрузка сохраняется в течение длительного времени, происходит переполнение буферов, и пакеты теряются;
- при наличии ошибок в пакетах данных пакеты могут быть переданы повторно в соответствии с запросом приемной стороны, тогда как повторная передача пакетов речи увеличивает задержку пакетов. Поэтому при пакетной передаче речи пакеты с ошибками сбрасываются. Потери речевых пакетов не должны превышать 1 % в среднем на достаточно большом временном интервале. При больших значениях коэффициента потерь в восстановленной на приемной стороне речи возникают разрывы.

### Заключение

Рекомендации МСЭ Р.800 определен метод, основанный на субъективных оценках, выносимых группой экспертов. Этот метод, используемый в телефонных сетях, к сожалению, не может учитывать влияние вероятностно-временных характеристик сетей на качество передачи речи в пакетных сетях.

Второй рассмотренный в статье метод вычисления R-фактора, основанный на E-модели, может рассматриваться как общая модель МСЭ-Т для объективной оценки качества передачи речи. Главной особенностью E-модели является учет большого набора факторов, отражающих воздействие оконечных устройств и транспортной среды на качество обслуживания. Показана связь между субъективными оценками, основанными на модели MOS, и R-фактором. Эта вычислительная модель может быть полезна как разработчикам сетей, так и поставщикам услуг.