

УДК 621.394.396

## АНАЛИЗ МОДЕЛИ ОБЪЕКТИВНОЙ ОЦЕНКИ КАЧЕСТВА ПЕРЕДАЧИ ВИДЕО В IP-СЕТЯХ

М.А. Маколкина, старший преподаватель СПб ГУТ им. М.А. Бонч-Бруевича; makolkina@list.ru

**Ключевые слова:** качество передачи видео, пиковое отношение сигнал/шум (PSNR), искажение видео, сетевые потери.

**Введение.** Сегодня услуги IPTV завоевывают все большую популярность среди пользователей, и большинство операторов уже развернули или разворачивают площадки для их внедрения. Данный тип услуг мало изучен с точки зрения оценки качества. Существует ряд методов для оценки качества передачи видео (КПВ), но нет однозначного подхода, дающего четкое представление, что видит на экране телевизора пользователь. Поэтому сложно предотвратить возникающие перебои в вещании, предупредить клиента и повысить тем самым его лояльность.

В последние годы объем трафика IPTV существенно увеличивается и влияет на показатели качества обслуживания всех видов трафика, передаваемых в IP-сетях, в том числе и на качество доставки самого трафика IPTV. Возникает проблема оценки качества доставки телевизионного сигнала через сеть. Большой интерес вызывают методы объективной оценки, базирующиеся на сборе и анализе сетевых характеристик. Ниже рассмотрена модель объективной оценки качества IPTV, в которой КПВ оценивается на основе не только сетевых характеристик, но и специфических особенностей видеотрафика.

**Объективные методы.** На КПВ влияют различные факторы, зависящие как от сети, так и от специфики видеоприложений. Потери пакетов и джиттер задержки – наиболее значимые факторы при оценке качества, зависящие от функционирования сети. С другой стороны, тип кодека, техника маскировки ошибок, скорость кодирования, схема пакетизации и другие факторы, определяемые видом приложения, также оказывают воздействие на КПВ.

Применяемые сегодня методы оценки КПВ, основанные на анализе характеристик сети, фактически являются теми же средствами, которые используются для контроля качества передачи голосовой информации, т.е. измеряют параметры транспортной сети. В отличие от субъективных методов, объективные не дают представления о том, что на самом деле видит пользователь на экране телевизора. Объективные ме-

тоды не делают различия между пакетами, переносящими видео, и пакетами, переносящими другую информацию. Для того чтобы измерить в количественном отношении качество услуги IPTV, необходимо ввести определенную метрику.

Сегодня МСЭ-Т для оценки качества IPTV предлагает использовать метрику QoE (Quality of Experience) – качество восприятия. Вместе с тем существует ряд методов, имеющих значительные различия между собой и базирующихся на разных подходах – MDI, VQM, MPQM, NQM и т.д. Эти методы можно отнести к общим методам оценки работоспособности сети, в полной мере не учитывающим особенности приложения. Для более детальной оценки качества необходима разработка моделей, которые при расчете показателей ориентируются не только на параметры сети, но и на характеристики видеопотока. Ряд подходов основывается на моделях «потери–искажения», т.е. на моделях, отображающих влияние потерь пакетов на качество видео (в форме искажений).

При разработке модели оценки качества видео необходимо учесть два основных требования: пригодность модели, рассчитывающей различные сетевые параметры, параметры приложений и с высокой степенью точности проецирующей их на оценки КПВ; способность количественно определить эти параметры и применить оценку КПВ в реальном масштабе времени для потенциально большего количества видеопотоков.

Соблюдение обоих требований – достаточно сложная задача. С одной стороны, точные оценки КПВ требуют детального рассмотрения состояния сети и подробной информации о приложении. С другой – измерение всех параметров в реальном времени и в надлежащем объеме очень сложная задача. На практике необходимо найти компромисс между точностью и скоростью расчета или его сложностью. Поэтому на модель накладывается ряд ограничений, в частности модель не учитывает В-кадры, которые кодируются путем двунаправленного предсказания на основе предыдущего I-кадра и последующего Р-кадра. В них сокращается и временная, и пространственная избыточность, достигается самая большая степень компрессии. В целях упрощения расчетов

модель не работает с цветным изображением и не производит детального разбора видеопотока.

**Базовая модель [1].** Для оценки качества видео необходимо исследовать отношение между потерями пакетов и искажениями в декодированном видео. Рассмотрим видеопоследовательность с кадрами размером  $N_1 \times N_2$  пикселей. Обозначим  $f[k]$  – вектор **2-D** (размером  $N_1 \times N_2$ ), полученный через строчную развертку кадра  $k$ ;  $f[k]$  – соответствующий кадр, восстановленный на декодере. Для упрощения выражения **2-D** вектор ( $N=N_1 \times N_2$  пикселя в каждом кадре  $k$ ) представим в виде **1-D** вектора  $f[k]$  размера  $N$ . Тогда сигнал ошибки в кадре  $k$ , представляющий ухудшение сигнала, возникшее из-за потери пакетов, определяется как:

$$e[k] = \hat{f}[k] - f[k]. \tag{1}$$

Квадрат ошибки в кадре  $k$ :

$$\sigma^2[k] = (e^T[k]e[k]) / N. \tag{2}$$

Значение  $\sigma^2[k]$  подвержено влиянию некоторых сетевых и зависимых от приложения факторов. В данной модели «потери—искажения» реализована попытка найти взаимосвязь между субъективными и объективными методами оценки. Для простоты оценки искажения измеряются в разгах.

Важная проблема при моделировании искажений – определение пределов, в которых ошибка размножается среди кадров. Так как временное предсказание предполагает зависимость между соседними кадрами, одиночная потеря пакета затрагивает не только кадр с данными, которые переносились в потерянном пакете, но и другие кадры с кодовыми зависимостями от потерянного кадра. Если ошибка происходит в кадре  $k$ , то значение размноженной ошибки в кадре  $[k+i]$  можно записать так:

$$\sigma^2[k+i] = \sigma^2[k]\gamma^i. \tag{3}$$

Фактор затухания  $\gamma$  ( $\gamma < 1$ ) учитывает влияние пространственной фильтрации, зависит от спектра сигнала ошибки и типа пространственной фильтрации, применяемой в декодере, т.е. меняется как функция характеристик видео и процесса декодирования.

Если  $(T-1)$  кадров закодировано с однонаправленным предсказанием (P-кадры) между двумя последовательностями кадров с внутренним кодированием (I-кадры), то общее искажение, возникающее по причине потерь в кадре  $k$ , составит

$$D = \sum_{i=0}^{x-1} \sigma^2[k+i], \tag{4}$$

где  $x$  – число кадров (от кадра, в котором произошла первичная ошибка – кадр  $k$ , до следующего I-кадра);  $T$  – количество P-кадров в группе кадров, закодированных с предсказанием на основе предыдущего I-кадра.

Элементарный поток видео имеет иерархическую структуру, состоящую из шести типов элементов: видеопоследовательности, группы кадров (изображений), кадра, среза, макроблока и блока. В зависимости от факторов, специфических для видеоприложений, например, формата дискретизации, IP-пакет может содержать как целый кадр, так один или несколько блоков, относящихся к одному кадру. Предполагая, что ожидаемое исходное искажение, возникшее из-за потери блока, равно  $\sigma_s^2$  и что внутри группы кадров, закодированной с предсказанием, местоположение  $x$  кадра с потерянным блоком равномерно распределено на интервале

$[0, T-1]$ , общее среднее искажение, возникшее из-за потери одного блока, определяется выражением

$$D_1 = \sum_{i=0}^{T-1} \sigma_s^2 \gamma^i \left(1 - \frac{i}{T}\right) = \frac{\gamma^{T+1} - (T+1)\gamma + T}{T(1-\gamma)^2} \sigma_s^2 = \alpha \sigma_s^2, \tag{5}$$

где  $\alpha$  – функция  $\gamma$  и  $T$ , учитывающая общее влияние размножения сигнала ошибки.

Поскольку в IP-сетях потери видеоданных происходят из-за потери пакетов, а не блоков, следующий шаг приводит к отображению потери IP-пакетов на потери блоков видеокadra. Когда теряется  $n$  ( $n \geq 1$ ) последовательных пакетов за одно событие, будет затронуто  $f(n)$  блоков, где  $f(n)$  – отражение числа потерянных пакетов на число потерянных блоков. Это отношение зависит от реализации кодека и техники восстановления потерь. Например, если каждый пакет содержит один блок и декодер просто пропускает декодирование блоков, содержащихся в потерянном пакете, то  $f(n)=n$ . Однако, если декодер отбрасывает видеокادر всякий раз, когда один из IP-пакетов потерян, отношение  $f(n)$  принимает другую форму. Тем не менее, для любого кодека  $f(n)$  может быть определено, и общие искажения, возникающие при  $n$  последовательно потерянных пакетов, могут быть смоделированы пропорционально искажениям, возникающим при потере блоков, т.е.

$$D_n = f(n)D_1. \tag{6}$$

Далее в процессе моделирования выразим средние искажения как функцию модели потерь, учитывая размер пачки ошибок и место, затронутое ошибкой. Используем  $P_n$  для определения вероятности потери  $n$  последовательных пакетов и  $P_m$  для вычисления вероятности того, что два последовательных события потери затронули  $m$  пакетов по отдельности (с начала пакета, попавшего в первое событие потери, до пакета, попавшего во второе событие).

Предположим, что каждый кадр передается  $L$  пакетами и что  $n$  и  $m$  являются независимыми случайными величинами. Тогда общее среднее искажение восстановленного видео может быть рассчитано по формуле:

$$\bar{D} = \frac{\sum_n P_n D_n}{\sum_m P_m (n-r)} = \frac{\overline{f(n)}}{\bar{m}} L D_1 \tag{7}$$

или

$$\bar{D} = P_e f(n) L D_1, \tag{8}$$

где  $P_e$  – вероятность потери любой длины в видеопотоке;  $f(n)$  – среднее число блоков, на которых повлияли потери. Величины  $P_e$  и  $f(n)$  выражают характеристики процесса потерь ( $f(n)$  также зависит от схемы пакетизации и техники восстановления после потерь, используемой в декодере), в то время как  $L$  и  $D_1$  являются параметрами, учитывающими особенности видеоприложения.

**Моделирование влияния различных кодеков.** Несмотря на то, что большинство стандартов сжатия видео поддерживает разделение видеокadров на блоки, кодеки реализуют разные схемы восстановления видеокadra при потере блоков. Для иллюстрации и сравнения результатов рассмотрим кодеки, реализованные в соответствии со стандартами MPEG-2, H.264 и имеющие разные реакции на произошедшую ошибку.

В кодеке MPEG-2 при потере IP-пакетов происходит следующее: если декодер обнаруживает, что какое-то число IP-пакетов, переносящих информацию одного видеокadra, потеряно, то он отбрасывает весь поврежденный кадр и заменяет его предыдущим декодированным кадром. В кодеке

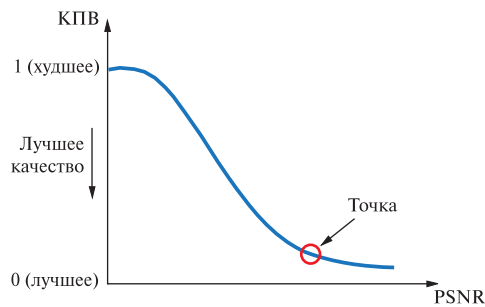


Рис. 1

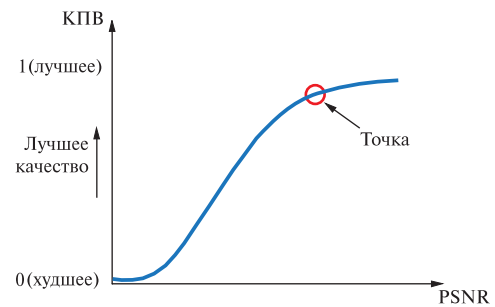


Рис. 2

H.264 реализована другая техника маскировки ошибок: все полученные блоки декодируются, в то время как блоки, переносимые потерянными IP-пакетами, восстанавливаются, используя блоки предыдущего кадра и информацию компенсации движения других блоков этого же кадра.

В кодеке MPEG-2 потеря затрагивает не только блоки, находящиеся в потерянных IP-пакетах, но и другие блоки этого кадра, в то время как в кодеке H.264 затрагиваются только блоки, входящие в потерянные пакеты. Заметим, что вышеизложенное описание MPEG-2 и H.264 представляет конкретную реализацию кодека.

Для упрощения анализа предположим, что каждый видеопакет содержит  $s$  блоков и каждый видеокادر передается  $L$  пакетами [2, 3]. Тогда общее искажение, влияющее на видеопоследовательность при  $n$  последовательно потерянных пакетах, может быть представлено как:

$$\begin{aligned} \bar{D} &= s(\bar{n} + L - 1)P_eLD_1 - \text{ для кодека MPEG-2;} \\ \bar{D} &= s\bar{n}P_eLD_1 - \text{ для кодека H.264.} \end{aligned} \quad (9)$$

Описанная выше модель позволяет учесть:

- модель потерь пакетов (величины  $n$  и  $P_e$ );
- битовую скорость передачи, выражаемую через число IP-пакетов, передающих видеокادر, и число блоков в видеокadre ( $s$  и  $L$ );
- схему пакетизации, выражаемую через  $L$  – число IP-пакетов, передающих видеокادر;
- механизмы восстановления от потерь (значение  $f(n)$  для кодеков MPEG-2 и H.264);
- чувствительность видеопоследовательности к ошибкам (параметр  $D_1$ ).

Поскольку искажение  $\bar{D}$  определено, результирующее качество видео может быть выражено с помощью пикового отношения сигнал/шум (PSNR – Peak Signal-to-Noise Ratio), т.е.

$$PSNR = 10 \log \left( \frac{255^2}{\bar{D}} \right). \quad (10)$$

Однако PSNR не дает четкого представления о том, как зритель оценит КПВ, т.е. PSNR показывает, что качество приемлемое, а на самом деле происходит систематическая потеря I-кадров видеопоследовательности и пользователь наблюдает картинку низкого качества с замираниями и артефактами. Из-за особенностей формирования видеопотока объективные и субъективные методы могут давать разные оценки КПВ.

Группа экспертов VQEG (Video Quality Expert Group) в сотрудничестве с различными университетами США, Европы, Азии и исследовательскими лабораториями проводит ряд экспериментов, нацеленных на сопоставление результатов объективных методов с субъективными оценками. Воспользуемся результатами VQEG и выразим КПВ через PSNR [4], как показано на рис. 1.

Отношение между этими двумя параметрами имеет вид:

$$КПВ = \frac{1}{1 + \exp(b_1(PSNR - b_2))}, \quad (11)$$

где  $b_1$  и  $b_2$  – параметры, учитывающие характеристики видео. Выражение (11) предполагает, что PSNR может оценить субъективное качество видео. Например, когда его значение меньше соответствующей «точки насыщения» (рис. 1), PSNR имеет более линейную зависимость с КПВ. Когда значение PSNR достигает «точки насыщения», субъективное качество видео по существу «насыщается до предела» и дальнейшее увеличение PSNR не приводит к заметному для человеческого глаза улучшению качества видео.

Перестроим график, предложенный VQEG. Для этого введем коэффициент  $C = 1 + \exp(b_1(PSNR - b_2))$ , учитывающий характеристики видео и отображающий искажения видеопоследовательности. Тогда формула (11) примет вид:

$$КВП = 1 - \frac{1}{C}. \quad (12)$$

Взаимосвязь между объективными и субъективными оценками КПВ показана на рис. 2. При достижении объективных оценок «точки насыщения» улучшать КПВ далее нецелесообразно, поскольку особенности человеческого зрения таковы, что пользователь уже не заметит разницы.

**Оценка модели.** Для оценки точности модели было проведено имитационное моделирование, учитывающее влияние различных приложений и сетевых параметров на качество видео. Моделирование осуществляли для двух кодеков MPEG-2 и H.264. Для проверки чувствительности модели к этим факторам изменяли размер кадра и формат. Исследовали два формата кадров QCIF и CIF, два вида потерь (независимые, назовем их Бернулли, и пачечные), уровень потерь – 2%.

Определим КПВ для кодека MPEG-2. Для этого рассчитаем общее искажение, влияющее на видеопоследовательность через  $n$  последовательных потерь, по формулам (9)–(11). Возьмем следующие значения параметров:

$s=2$  – количество блоков в пакете;

$L=2$  – количество пакетов в кадре;

$N=3$  – количество последовательно потерянных пакетов за одно событие;

$D_1=150$  – общее среднее искажение, возникшее по причине потери одного блока видеокadre.

Тогда, при вероятности события потери (любой длины) в видеопоследовательности  $P_e=0,04$ ,  $\bar{D}=96$  и  $PSNR=28,38$  дБ получим  $КВП=0,013$ . При  $P_e=0,4$ ,  $\bar{D}=960$  и  $PSNR=18,35$  дБ имеем  $КВП=0,28$ .

В ходе анализа модели «потери–искажения» было рассмотрено несколько случаев с различными параметрами. На рис. 3 представлен один из результатов имитационного моделирования, в ходе которого ставилась задача определить

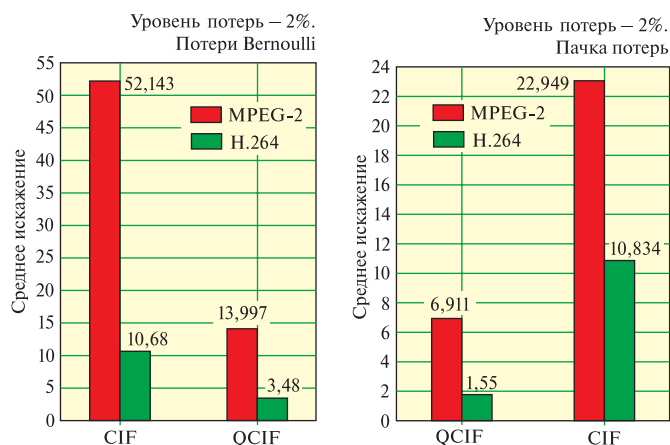


Рис. 3

среднее искажение для различных типов кодеков (MPEG-2, H.264) и форматов (CIF, QCIF).

**Результаты моделирования.** Количество потерянных пакетов – доминирующий фактор, влияющий на качество видео в каждой конфигурации видео. В частности, для данной схемы пакетизации (представленной  $s$  и  $L$ ), размера кадра, выбора кодека и пачечности  $n$  искажения пропорциональны вероятности события потерь  $P_e$ .

Влияние пакетизации зависит от типа кодека. Для кодека MPEG-2 искажения увеличиваются с возрастанием  $L$ , поскольку одиночная потеря пакета затрагивает не только блок в этом пакете, но и все другие блоки кадра. Ситуация совершенно иная в случае с H.264 кодеком, для которого (рис. 3) изменение  $L$  имеет меньшее воздействие на качество видео.

Схема маскировки ошибок определяет чувствительность кодека к моделям потерь. Как видно из рис. 3, более сложная схема для кодека H.264 чувствительна к пачкам потерь, обычно снижающим способность кодека к маскировке ошибок. Это происходит потому, что пачки потерь затрагивают большое количество блоков в кадре. Отсюда для заданной скорости потери пакетов предпочтительнее группировать потери пакетов в одном кадре, чем распространять их на несколько кадров.

**Заключение.** Рассмотренная модель «потери–искажения» позволяет контролировать КПВ в IP-сетях. Модель учитывает влияние различных факторов, зависящих от параметров сети и специфики приложения, на качество видео. Модель не требует детальных знаний о характеристиках видео или глубокого разбора видеопотока. Тем не менее она дает точную оценку КПВ, которая может быть скоррелирована с субъективными методами оценки.

#### ЛИТЕРАТУРА

1. Liang Y.J., Apostolopoulos J.G., Girod B. Analysis of packet loss for compressed video: Does burst-length matter?/ ICASSP, Hong Kong, China. – Apr. 2003.
2. Perkins C., Hodson O., Hardman V. Survey of packet loss recovery techniques for streaming audio// IEEE Network. – Sep. 1998. – Vol. 12. – P. 40–48.
3. Feamster N., Balakrishnan H. Packet loss recovery for streaming video / 2<sup>nd</sup> Int. Packet Video Workshop, Pittsburgh, PA. – Apr. 2002.
4. VQEG. Final Report on the Validation of Objective Models of Video Quality Assessment. – Aug. 2003[Online]. Available: <http://www.vqeg.org>.

Получено 28.09.11