

ОБЕСПЕЧЕНИЕ ГАРАНТИРОВАННОГО КАЧЕСТВА ОБСЛУЖИВАНИЯ В СЕТЯХ NGN С ИСПОЛЬЗОВАНИЕМ ОЦЕНОК КОНЕЧНЫХ ПОЛЬЗОВАТЕЛЕЙ

Введение

Предоставление услуг гарантированного качества независимо от типа передаваемой информации является отличительной особенностью сетей следующего поколения (Next Generation Network, NGN). Качество обслуживания абонентов в NGN является ключевым вопросом и в той или иной форме присутствует при решении практически любой задачи, связанной с проектированием или эксплуатацией сетей данного типа. Как правило, качество предоставляемой пользователю услуги фигурирует в роли ограничений в рамках математической постановки какой-либо более общей задачи, например задачи выбора пропускных способностей каналов передачи на этапе проектирования или задачи распределения этих пропускных способностей между информационными потоками на этапе управления сетью в процессе ее эксплуатации. Зачастую эти ограничения формулируются в терминах сетевых показателей качества [1, 2], отражающих качество передачи трафика на участке транспортной сети и не учитывающих влияние других элементов соединения, например локальных сетей и оконечного оборудования. В общем случае формулировка требований и ограничений должна осуществляться в терминах оценок качества, воспринимаемого конечным пользователем, которые являются производными качества работы сети.

Рассматривая управление сетевыми ресурсами как основной инструмент достижения заданного качества обслуживания на этапе эксплуатации телекоммуникационной сети (ТКС), следует отметить, что наибольшая эффективность управления трафиком, канальными и буферными ресурсами достигается при условии реализации динамической многопутевой стратегии. При ориентировании именно на такой подход возникает следующая научно-техническая задача: сформулировать в аналитическом виде условия (ограничения) обеспечения требований конечных пользователей к воспринимаемому качеству услуги в условиях реализации динамической многопутевой стратегии управления трафиком и другими сетевыми ресурсами.

Поставленную задачу условно можно разделить на две последовательно решаемые подзадачи: первая связана с получением формализованной зависимости оценок качества конечного пользователя от текущего значения сетевых параметров, вторая – с формированием зависимости сетевых параметров качества от текущего распределения ресурсов при условии реализации динамической многопутевой стратегии.

Уровни качества обслуживания

Сквозное качество обслуживания пользователей определяется рядом факторов, среди которых качество сетевого соединения (на уровне IP), специфические для конкретного приложения действия, аспекты восприятия самого пользователя. В соответствии с перечисленным можно выделить три уровня качества обслуживания, для описания каждого из которых вводится свой термин (рис. 1 – сквозная оценка качества телекоммуникационной услуги, рис. 2 – точки измерения параметров качества; табл. 1) [3, 4]:

- качество работы сети (Network Performance, *NP*), соответственно на уровне сети;
- качество телекоммуникационных услуг (обслуживания) (Quality of Service, *QoS*) на уровне приложения;
- воспринимаемое качество телекоммуникационной услуги (Quality of Experience, *QoE*) на уровне пользователя.

Оценка *QoE* представляет собой количественно выраженную степень удовлетворенности конечного пользователя качеством полученной им телекоммуникационной услуги. На достижение хороших значений именно этих оценок направлены все усилия операторов связи. Это интегральная оценка качества, выражаемая не в технической терминологии, а в неко-

торых баллах. Рекомендациями ITU-T определены несколько таких оценок, среди которых в качестве основных следует выделить (табл 2):

- рейтинг качества R (Quality Rating), $0 \leq R \leq 100$ [5];
- среднюю экспертную оценку MOS (Mean Opinion Score), $1 \leq MOS \leq 5$ [6].

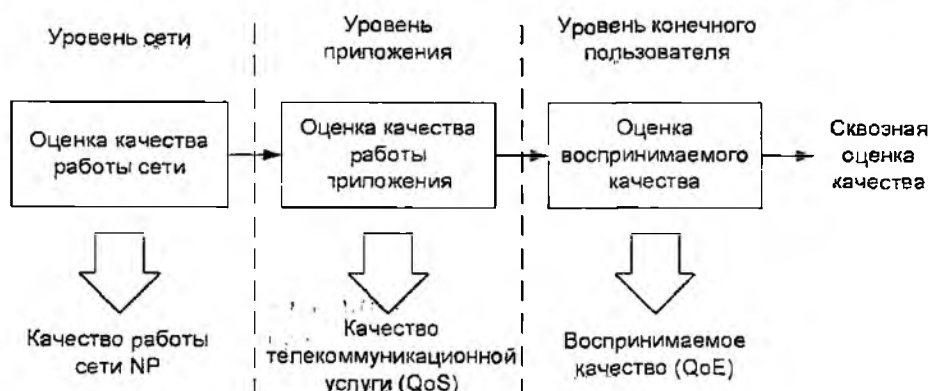


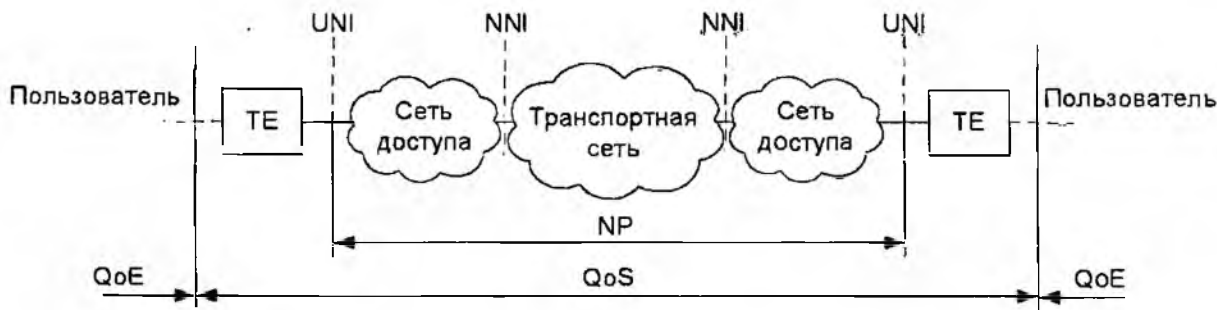
Рис. 1

Таблица 1

	Воспринимаемое качество QoE	Качество телекоммуникационной услуги QoS	Качество работы сети (сетевые характеристики NP)
Ориентирована	на пользователя		на оператора (провайдера)
Уровень	Конечного пользователя	Приложения	Сети
Указывает на	Соответствие ожидаемому пользователем качеству услуги, отражает потребительские свойства услуги	Получаемое пользователем качество услуги	«Слабые места» в работе сети, отражает текущее качество работы сети, ориентировано на разработку, проектирование, эксплуатацию и техническое обслуживание
Определяется	Пользователем с учетом аспектов восприятия	Путем измерения в трюках (между точками) доступа к услуге	Типом (возможностью) сетевых элементов, протоколов и соединений в целом
Описывается параметрами (атрибутами)	Восприятия пользователя (оценки MOS, R)	Услуги (суммарная с учетом терминального оборудования задержка, результирующий уровень потерь и принятых с ошибками пакетов, скорость с учетом TCP)	Элементы соединения (выделенная пропускная способность, сетевая задержка с учетом времени распространения и обслуживания в очереди, сетевой джиттер, вероятности потери пакета в сети, возникновения ошибок в пакете, нарушения порядка следования)
Характер оценки	Интегральная	Совокупность оценок	
Документация ITU-T	G 107, G 109, G 1070, P 800, P 830	G 1010, G 1030, Y 1541	Y 1540, Y 1541

Таблица 2

R (нижний предел)	MOS (нижний предел)	Удовлетворенность пользователя
90	4,34	Очень удовлетворен
80	4,03	Удовлетворен
70	3,60	Некоторые пользователи не удовлетворены
60	3,10	Многие пользователи не удовлетворены
50	2,58	Почти все пользователи не удовлетворены



TE – Terminal Equipment, оконечное оборудование

Рис 2

Оценки MOS и R связаны между собой нелинейной зависимостью [7], которую в диапазоне $2,5 < MOS < 4,4$ можно аппроксимировать выражением $MOS = R/20$.

Требование хорошего качества в терминах QoE можно записать как $R_{треб} > 70$ или $MOS_{треб} > 3,5$.

На практике QoE определяется качеством соединения, а также психологическими аспектами восприятия, где главную роль играет не столько получаемое качество, сколько его соответствие ожидаемому. Это указывает на важность предварительного соглашения о качестве предоставляемых услуг SLA (Service Level Agreement) между пользователем и провайдером.

Хотя QoS , так же, как QoE , определяется в точке потребления услуги (интерфейс пользователь-машина), однако, в отличие от QoE , качество телекоммуникационной услуги QoS выражается через совокупность специфических (технических) показателей и параметров услуги, которые, с одной стороны, определяются качеством работы телекоммуникационной сети, а с другой – определяют способность данной услуги удовлетворять установленные или прогнозируемые требования пользователя [3]. QoS отражает качество предоставляемой услуги на уровне приложения и в соответствии с делением приложений по типу передаваемой информации (речь, видео и данные) в качестве ключевых параметров использует задержку, джиттер и величину потерь для аудио- и видеоприложений (для видеотелефонии добавляется еще синхронизация между потоками звука и изображения) и задержку и величину потерь для приложений передачи данных [8]. Рекомендация ИТУ-T G 1010 содержит граничные значения указанных параметров, необходимые для приемлемого качества работы различных приложений. Следует отметить, что для приложений передачи данных фактором, определяющим результирующее качество услуги (QoE), является величина потерь, в то время как для передачи речи и видео критичными будут величины задержки и джиттера (рис 3) [8].

Показатели QoS являются функцией качества работы терминального (оконечного) оборудования и качества работы сети, включая сети доступа и транспортную сеть. В результате можно выделить два направления в достижении требуемого качества телекоммуникационной услуги QoS :

1) улучшающие качество настройки терминального оборудования (выбор и настройка кодеков, настройка TCP и операционной системы, выбор алгоритма и размера буфера компенсации джиттера, правильный выбор скорости передачи и частоты кадров, обеспечение при необходимости синхронизации между звуком и изображением и пр.),

2) приемлемое качество работы сети NP . Учитывая долю сети в суммарном значении перечисленных выше параметров QoS , качество ее работы должно находиться под тщательным контролем со стороны сетевой системы управления и само по себе представляет предмет для исследования.

Качество работы сети NP – это количественные характеристики сети, технические показатели и параметры, полученные в результате испытаний и измерений параметров телекоммуникационной сети, каналов электросвязи, технических средств телекоммуникаций [3].

В качестве основных параметров сетевого уровня, определяющих качество обслуживания абонентов, в соответствии с Y 1540 [9] можно выделить:

- задержку передачи IP-пакета T_{NP} , которая включает в себя задержки распространения и обработки в очереди,
- межконцевую вариацию задержки (джиттер) D_{NP} ,
- процент потерянных IP-пакетов P_{NP} ,
- процент пакетов, принятых с ошибками P_{ENP} ,
- процент пакетов с нарушенной очередностью P_{RONP} ,
- скорость передачи $B_{\backslash P}$, выражаемую в пакетах или в битах в секунду.

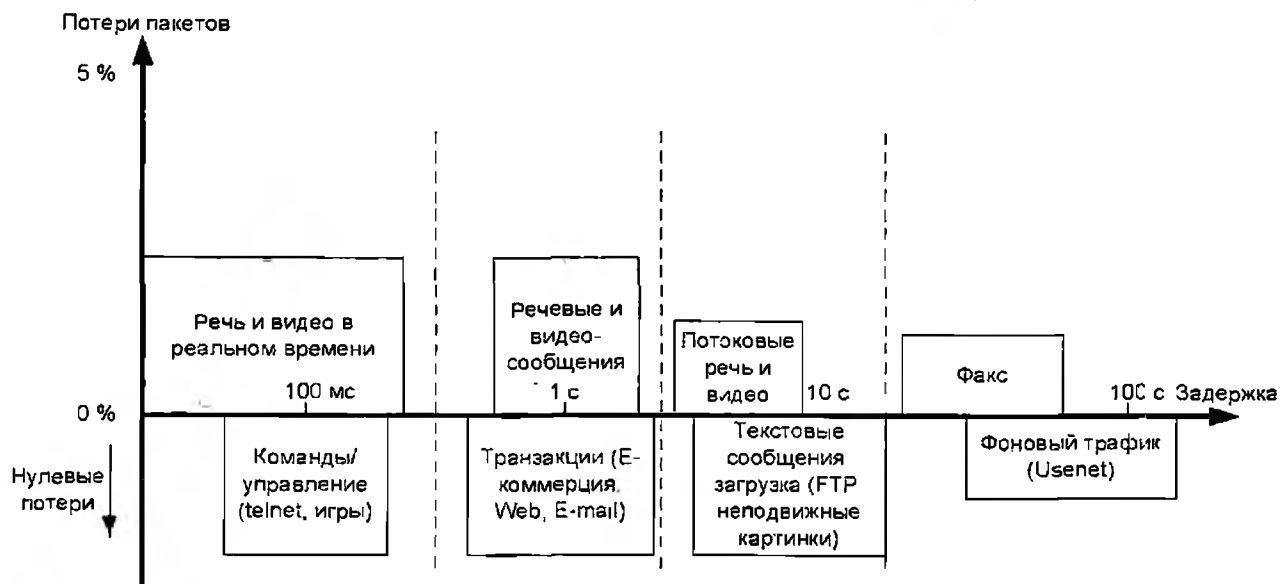


Рис 3

В рекомендации Y 1540 оговаривается, что параметры, связанные с потоком и пропускной способностью на уровне IP, не являются необходимыми, так как под влиянием протоколов более высоких уровней, прежде всего TCP, результирующие их значения могут существенно измениться.

Требования к параметрам различных приложений послужили основой для создания восьми классов сетевого качества обслуживания (табл. 3) [10]. Здесь класс 5 не содержит ни одного граничного значения, что подразумевает возможность их динамического формирования на основании текущего запроса.

Таблица 3

Параметр	Классы QoS								
	Класс 0	Класс 1	Класс 2	Класс 3	Класс 4	Класс 5	Класс 6	Класс 7	
T_{NP}	100 мс	400 мс	100 мс	400 мс	1 с	Н	100 мс	400 мс	
D_{NP}	50 мс		Н	Н	Н	Н	50 мс		
P_{NP}						Н	10^{-3}		
P_{ENP}						Н	10^{-6}		

Н – параметр не определен

Подходы к формулированию требований относительно качества обслуживания

На практике значения перечисленных параметров качества работы сети NP определяются текущим распределением сетевых ресурсов между принятыми к обслуживанию информационными потоками. При наличии системы динамического управления ресурсами сети параметры NP являются управляемыми и выступают в качестве инструмента достижения требуемого пользователем качества обслуживания (рис. 4).

Таким образом, используя введенную терминологию, можно рассматривать сквозное качество обслуживания как функциональную зависимость $QoE(QoS(NP))$ и в целях достижения качества обслуживания, удовлетворяющего пользователя, необходимо согласовать оценки потребителя QoE и оператора NP , что составляет основу системного подхода к качеству обслуживания [11]. Рассматривая управление параметрами NP как основной инструмент достижения запрашиваемого качества обслуживания, можно выделить два подхода к математической формулировке задачи его обеспечения. Оба они предполагают введение в существующую математическую модель ТКС дополнительных ограничений, однако в первом случае эти ограничения формулируются в терминах NP , а во втором – в терминах QoE (или QoS)

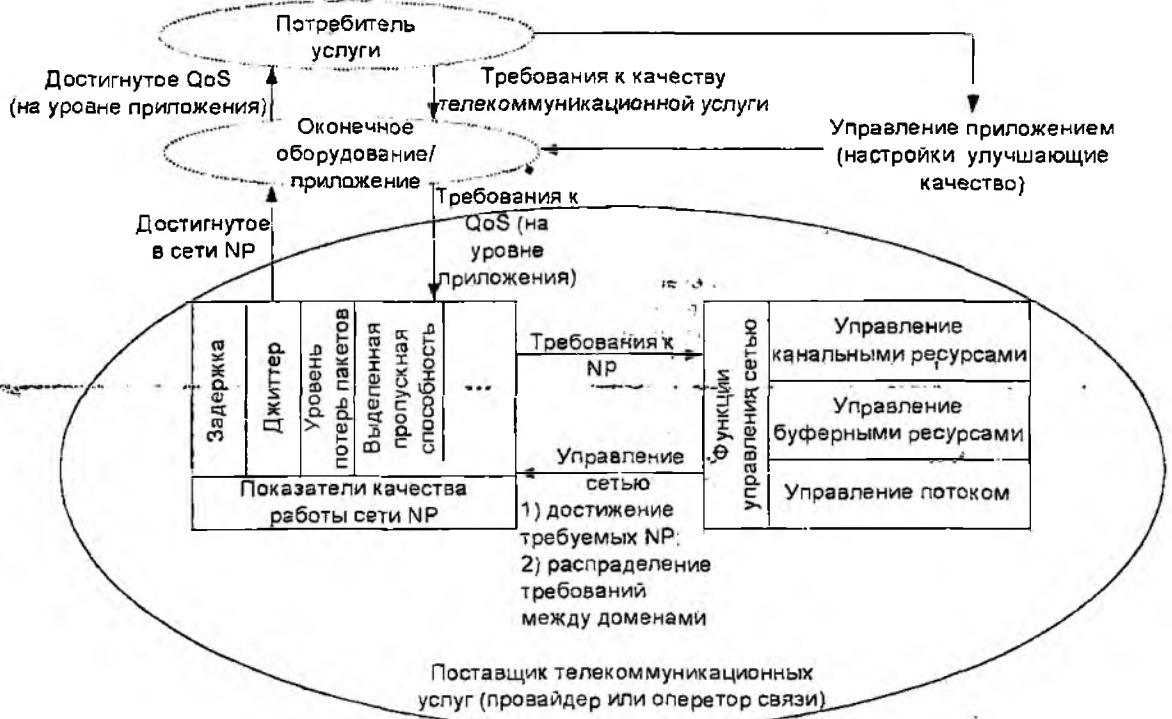


Рис 4

Ограничения на сетевые параметры качества в рамках первого, являющегося общепринятым, подхода могут быть записаны в общем виде как

$$\begin{cases} T_{NP} \leq T_{NP \text{ треб}}(QoE_{\text{треб}}), \\ D_{NP} \leq D_{NP \text{ треб}}(QoE_{\text{треб}}), \\ P_{NP} \leq P_{NP \text{ треб}}(QoE_{\text{треб}}), \\ B_{NP} \geq B_{NP \text{ треб}}(QoE_{\text{треб}}). \end{cases} \quad (1)$$

Данный подход предусматривает наличие предварительного этапа формирования требований к сетевым параметрам качества $NP_{\text{треб}}$ на основании известных требований на уровне пользователя $QoE_{\text{треб}}$ (или приложения $QoS_{\text{треб}}$), то есть как функцию $NP_{\text{треб}}(QoE_{\text{треб}})$ или $NP_{\text{треб}}(QoS_{\text{треб}})$.

Воспринимаемое качество обслуживания QoE зависит как от качества работы сети, так от ряда действий, производимых над трафиком в конечном оборудовании в соответствии с типом приложения, генерирующего данный трафик. В результате функциональная зависимость $QoE(QoS(NP))$ определяется типом приложения. Как показали результаты анализа, формализованная зависимость $QoE(NP)$ (а значит $NP_{\text{треб}}(QoE_{\text{треб}})$) существует только для трех типов услуг: передача речи, видеотелефония и просмотр информации в Web. Для некоторых типов услуг (различные приложения передачи данных, цифровое телевизионное ве-

щание) существует формализованная зависимость $QoS(NP)$, что позволяет сформировать требования к сетевым параметрам как $NP_{треб}(QoS_{треб})$, где $QoS_{треб}$ — характерные для данного приложения требования к качеству в терминах QoS , содержащиеся в G 1010. В случае полного отсутствия формализованной зависимости показателей различных уровней при формировании требований к сетевым параметрам следует воспользоваться рекомендацией Y 1541 и выбрать класс обслуживания (табл. 3) с учетом изложенного там же руководства по применению классов QoS протокола IP (рис. 5).

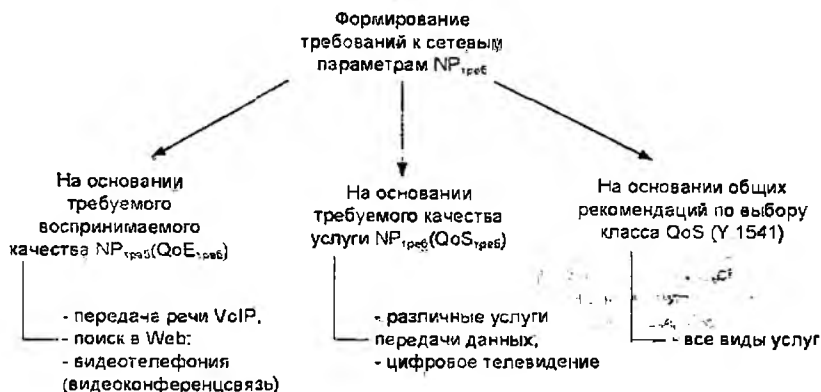


Рис. 5

Другой подход к обеспечению гарантированного качества связан с введением зависимости $QoSE(QoS(NP))$ в математическую модель сети. Тогда проверка выполнения требований относительно предоставляемого качества будет формализована в виде ограничений на уровне $QoSE$:

$$MOS \geq MOS_{треб} \text{ или } R \geq R_{треб} \quad (2)$$

Для тех видов услуг, для которых существует только формализованная зависимость $QoS(NP)$ ограничение (2) можно записать как

$$T_{QoS} \leq T_{QoS_{треб}}, D_{QoS} \leq D_{QoS_{треб}}, P_{QoS} \leq P_{QoS_{треб}}, B_{QoS} \geq B_{QoS_{треб}}, \quad (3)$$

где $T_{QoS}, D_{QoS}, P_{QoS}, B_{QoS}$ — соответствующие параметры качества на уровне приложения

Тогда в общем виде задача гарантированного качества обслуживания может быть сформулирована как оптимизационная (например, [1, 2]), связанная с поиском экстремума некоторого, как правило, стоимостного функционала при наличии ряда ограничений, среди которых вида (1) или (2), (3). Учитывая достаточно сложную взаимосвязь отдельных сетевых параметров в их влиянии на сквозную оценку качества и возможность ее более полного учета только в рамках второго подхода (2), (3), именно он заслуживает дальнейшего развития и предполагает следующие два этапа: формализацию зависимости $QoSE(NP)$ или $QoS(NP)$ и ее введение в математическую модель сети

Взаимосвязь различных уровней качества передачи речи

В случае передачи речевой информации через сети IP (VoIP) результирующее качество обслуживания $QoSE$ может быть оценено с использованием E-модели [7], в наиболее полной степени на данный момент отражающей сложную взаимосвязь факторов качества передачи речи (рис. 6).

В рамках E-модели для оценки воспринимаемого качества передачи речи используется рейтинг качества R , который представляет собой зависимость

$$R = R_0 - I_s - I_d - I_{e-эф} + A, \quad (4)$$

где R_0 — фактор, связанный со значением отношения сигнал/шум, учитывает влияние шумов, возникающих при передаче, таких как шумы оборудования, шумы в помещении, при значе-

нии всех входящих в его состав параметров, рекомендуемых по умолчанию, $R_{01} = 93,2$, I_x – коэффициент снижения качества, вызванного всеми одновременными искажениями, которым подвергается речевой сигнал, включая процессы преобразования в ЦАП/АЦП, а также вызванного неоптимальным местным эффектом, I_d – коэффициент снижения качества, обусловленного влиянием задержек и наличие эха сигнала, $I_{e-эф}$ – коэффициент снижения качества, вызванного искажениями, вносимыми в процессе кодирования/декодирования, а также потерями речевых пакетов; A – коэффициент выигрыша.

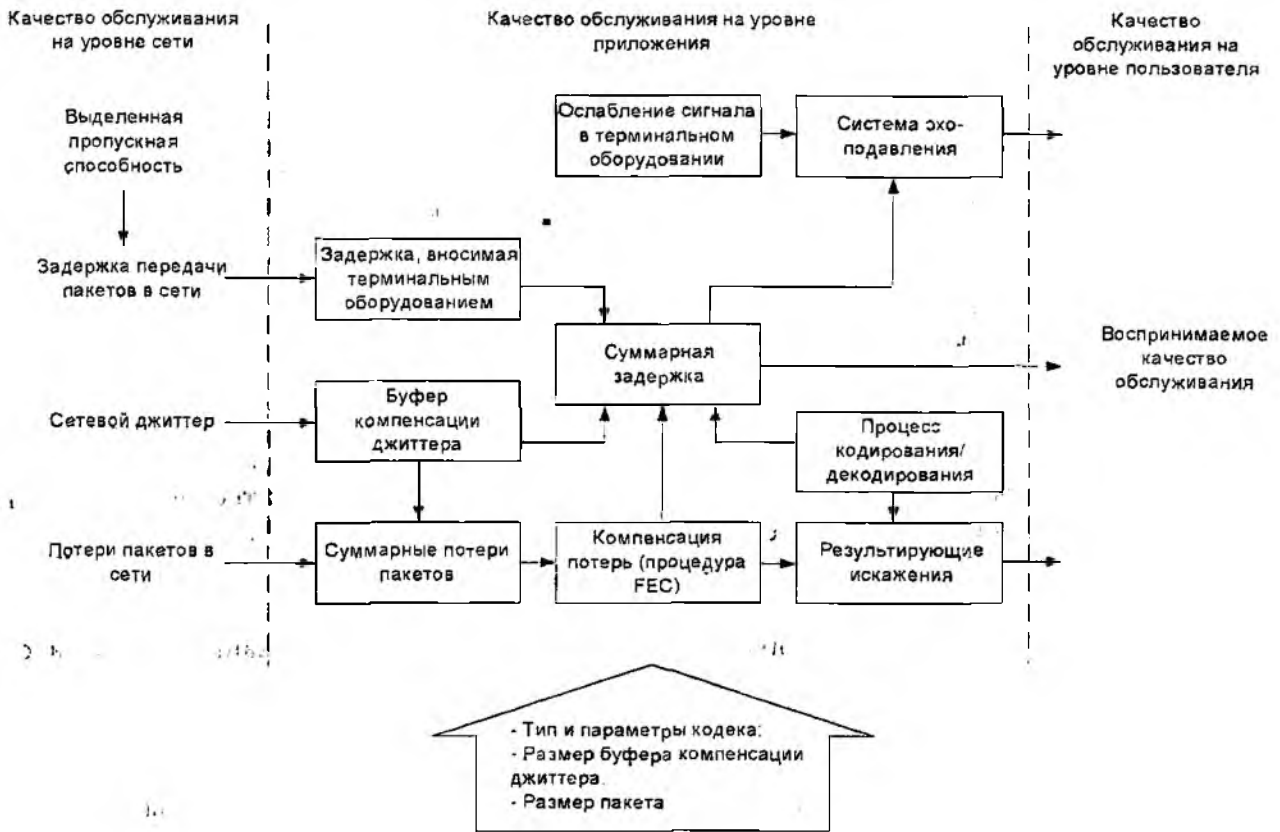


Рис. 6

В целом E-модель позволяет объединить в едином показателе качества как индивидуальные характеристики сигналов, так и сетевые параметры передачи (задержку и величину потерь пакетов в сети) и представляет собой совокупность ряда математических выражений, зависящих от множества входных параметров. Положив все входные параметры, не связанные с передачей по сети, равными значениям по умолчанию [7], можно сформировать зависимость $QoE(NP)$ для передачи речевого трафика следующим образом [12]

$$R = R_0 - I_{del}(T_{NP}) - I_{e-эф}(P_{NP}), \quad (5)$$

где $I_{del}(T_{NP})$ – коэффициент снижения качества, обусловленного длительной задержкой, как функция задержки в сети

Снижение качества, вызванное длительной суммарной (абсолютной) задержкой передачи T_d , определяется в единицах рейтинга в соответствии с E-моделью как (рис. 7) [12]:

$$I_{del} = \begin{cases} 0, & T_d \leq 100 \text{ мс}, \\ 25 \left[(1 + X^6)^{\frac{1}{6}} - 3 \left(1 + \left[\frac{X^6}{3} \right]^{\frac{1}{6}} \right) + 2 \right], & T_d > 100 \text{ мс}, \end{cases} \quad (6)$$

$$\text{где } X = \log \frac{\left(\frac{T_a}{100}\right)^{1.4}}{\log 2}$$

В случае передачи речи между двумя терминалами *VoIP*, взаимодействующими через сеть передачи данных, абсолютная задержка T_a представляет собой сумму одноконцевой сетевой задержки, отражающей качество работы сети, T_{cp} и задержек обработки в каждом терминальном оборудовании T_{TE} , возникающих вследствие кодирования t_{enc} , пакетизации t_{pack} , декодирования t_{dec} , компенсации джиттера t_{jb} . В среднем задержка T_{TE} составляет около 80 мс для устройства категории В Р 1010 и 50 мс для устройства категории А Р 1010 [10]. Наличие в соединении участка LAN увеличивает задержку T_a в среднем на 5 мс

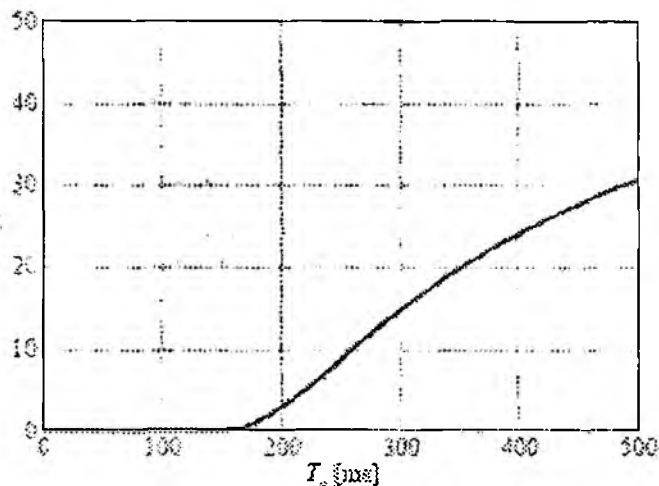


Рис. 7

Влияние потерь пакетов на качество передачи речи формализовано в виде коэффициента (рис. 8) [12]

$$I_{e-eff} = I_e - (95 - I_e) \frac{P_{pl}}{\frac{P_{pl}}{BurstR} + B_{pl}}, \quad (7)$$

где I_e – коэффициент снижения качества из-за использования низкоскоростных кодеков, который зависит только от типа кодека и определяется в Дополнении I/G 113 [13], P_{pl} – суммарная вероятность потери пакета, учитывающая потери пакетов в сети P_{NP} и в терминальном оборудовании P_{TE} , предполагая потери в терминальном оборудовании незначительными, можно считать $P_{pl} = P_{NP}$, B_{pl} – фактор, учитывающий устойчивость кодека к потерям (Дополнение I/G 113), $BurstR$ – коэффициент «всплеска» потерь, который равен 1 при независимой потере пакетов и превышает 1 при наличии групповых потерь

Выражения (6) и (7) в совокупности представляют собой формализацию зависимости $QoE(NP)$ для услуги передачи речи *VoIP*. Их влияние на результирующее воспринимаемое качество (для примера для кодека G 711 со случайными потерями и PLC) отражено на рис. 9 [12].

В результате требования пользователя к сквозному качеству предоставляемой услуги передачи речи при заданном типе терминального оборудования и используемом кодеке могут быть записаны так

$$R_c - I_{del}(T_{NP}) - I_{e-eff}(P_{NP}) \geq R_{треб} \quad (8)$$

Выражение (8) представляет собой ограничение типа (2) для приложений передачи речи которое в дальнейшем подлежит введению в математическую модель сети с целью ее расширения на предоставление услуг гарантированного качества

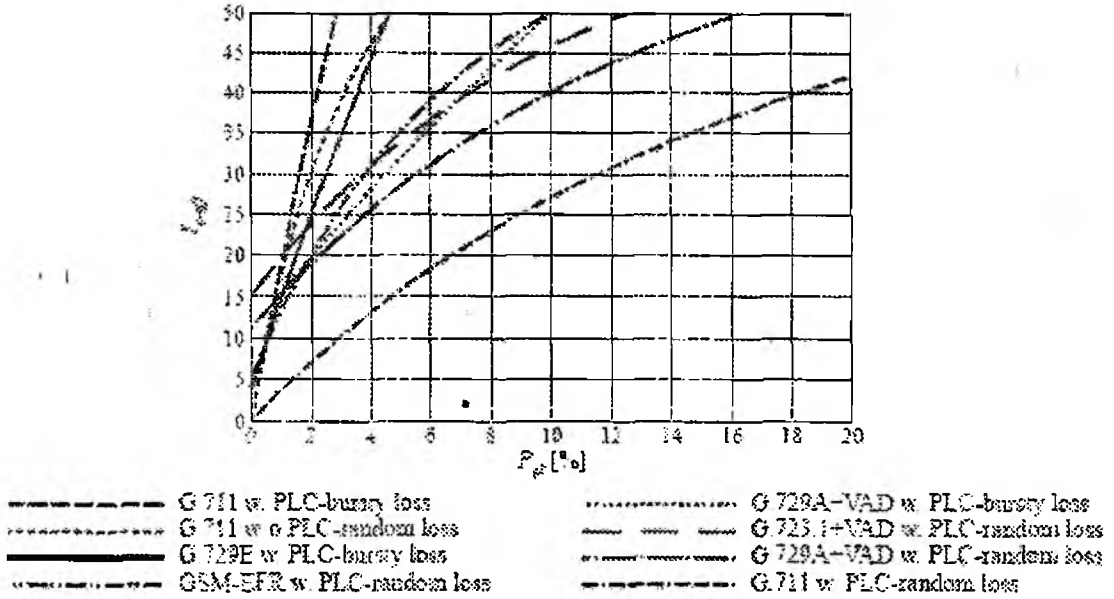


Рис. 8

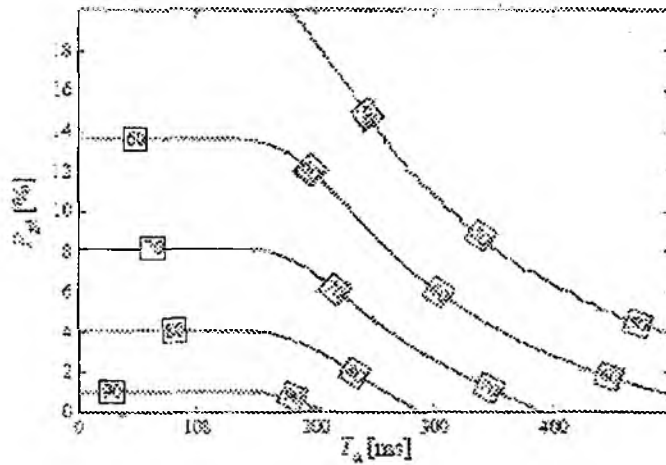


Рис. 9

Взаимосвязь различных уровней качества передачи видеoinформации

Методика оценки качества передачи видеoinформации через сеть IP $QoE(NP)$ применительно в услуге видеотелефонии подробно изложена в рекомендации ITU-T G 1070. Взаимосвязь факторов, влияющих на качество данной услуги, приведена на рис. 10.

В соответствии с G 1070 воспринимаемое качество мультимедийной информации определяется как

$$MM_q = m_1 MM_{st} + m_2 MM_T + m_3 MM_{st} MM_T + m_4, 1 \leq MM_q \leq 5, \quad (9)$$

где MM_{st} – качество передачи аудиовизуальной информации, MM_T – ухудшение качества вследствие наличия задержек и рассинхронизации сигналов звука и видео, m_i – коэффициенты, зависящие от размеров дисплея и целей общения (приведены в G 1070).

Входящие в состав (9) показатели определяются в соответствии со следующими выражениями

$$MM_{st} = m_5 S_q + m_6 I'_q + m_7 S_q V_q + m_8, 1 \leq MM_{st} \leq 5; \quad (10)$$

$$MM_T = \max \{AD + MS, 1\}, 1 \leq MM_T \leq 5, \quad (11)$$

$$AD = m_9(T_S + T_V) + m_{10}, \quad (12)$$

$$MS = \begin{cases} \min[m_{11}(T_S - T_V) + m_{12}, 0], & \text{если } T_S \geq T_V, \\ \min[m_{13}(T_V - T_S) - m_{14}, 0], & \text{если } T_S < T_V, \end{cases} \quad (13)$$

где S_q и V_q – качество передачи звука и изображения соответственно, AD – отражает влияние абсолютных задержек сигналов звука T_S и видео T_V ; MS – отражает рассинхронизацию между звуком и изображением.

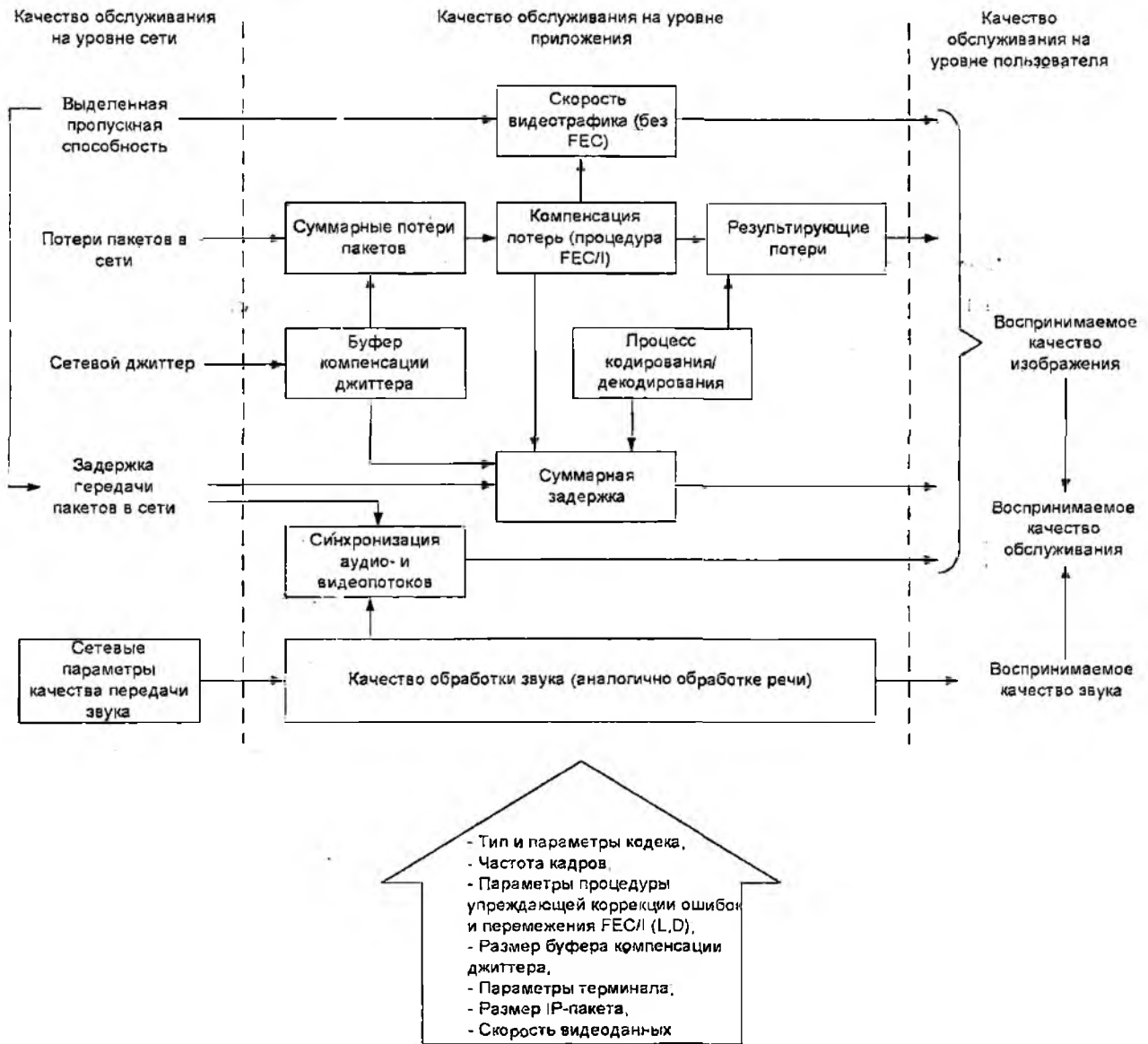


Рис 10

Здесь, как и в случае передачи речи, T_S и T_V представляют собой суммарные задержки звука и видео, в которых одним из компонентов является зависящая от работы сети T_{AP} при постоянных других компонентах.

Качество передачи звука S_q и изображения V_q при отсутствии эхо-сигнала является функцией величины потерь пакетов P_{pl} , и, предполагая $P_{pl} = P_{NP}$, можно считать $S_q(P_{NP})$ и $V_q(P_{NP})$. В этом случае качество передачи звука оценивается по аналогичной (5) формуле

$$Q = 93,193 - I_{\text{эфф}}(P_{NP}), \quad Q \leq 100, \quad (14)$$

которая затем пересчитывается в S_q [14]:

$$S_q = \begin{cases} 1, & \text{если } Q < 0, \\ 1 + 0,035Q + 7 \cdot 10^{-6} \cdot Q(Q-60)(100-Q), & \text{если } 0 < Q < 100, \\ 4,5, & \text{если } Q > 100. \end{cases} \quad (15)$$

Параметр качества передачи изображения V_q обусловлен тремя факторами: битовой видеоскоростью Br_f , частотой кадров Fr_f и вероятностью потерь пакетов P_{pl} и определяется следующим образом [14]:

$$V_q = 1 + I_{\text{кодир}} \exp\left(-\frac{P_{pl}}{D_{rpl}}\right), \quad (16)$$

где $I_{\text{кодир}}$ – коэффициент снижения качества видео, обусловленного использованием кодеков; D_{rpl} – фактор устойчивости кодека в потере пакетов.

При оптимальной частоте кадров $Fr_f = v_1 + v_2 Br_f$, $1 \leq Fr_f \leq 30$, максимизирующей качество передачи изображения для данной битовой скорости,

$$I_{\text{кодир}} = v_3 - \frac{v_3}{1 + \left(\frac{Br_f}{v_4}\right)^{v_5}}, \quad 0 \leq I_{\text{кодир}} \leq 4, \quad (17)$$

где v – коэффициенты, зависящие от типа кодека, формата видео, временного интервала между кадрами и размера дисплея [14].

Фактор устойчивости кодека в потере пакетов определяется как [14]

$$D_{rpl} = v_{10} + v_{11} \exp\left(-\frac{Fr_f}{v_8}\right) + v_{12} \exp\left(-\frac{Br_f}{v_9}\right), \quad 0 < D_{rpl} < \infty \quad (18)$$

Выражения (10) – (18) представляют собой зависимость $QoE(NP)$ для услуги видеотелефонии и могут быть использованы для формирования ограничений вида (2) на основе известного требуемого уровня воспринимаемого качества $MM_{q \text{ треб}}$.

Для других услуг передачи видео, например цифрового телевидения, формализованной зависимости $QoE(NP)$ нет, однако форумом видеослужб сформированы требования к качеству предоставляемых услуг на уровне приложения, т. е. в терминах QoS (табл. 4) [10]. Определяющим фактором для услуги цифрового телевидения является величина потерь пакетов P_{pl} .

Требования к сетевому параметру качества P_{NP} , отражающему вероятность потери пакета при передаче через сеть, будут значительно ниже при условии использования механизма упреждающей коррекции ошибок и перемежения FEC/I. Так, в [10] показано, что использование FEC/I позволяет снизить вероятность потери с 10^{-4} до $1,5 \times 10^{-8}$ (для уровня доступа) и с 10^{-5} до 2×10^{-10} (для всех уровней). Однако реализация механизма FEC/I требует увеличения пропускной способности соединения в среднем на 5 – 20 % от скорости видеопотока.

Таким образом, для услуги цифрового телевидения требования к параметрам качества могут быть сформулированы на уровне QoS как

$$\begin{cases} P_{pl} = k_{FEC/I}^p P_{NP} \leq P_{pl \text{ треб}}; \\ B_{QoS} = k_{FEC/I}^b B_{NF} \geq B_{QoS \text{ треб}}; \end{cases} \quad (19)$$

где $k_{ГКСИ}^p$, $k_{ГКСИ}^h$ – коэффициенты влияния применения процедуры FEC/I на параметры потерь и скорости.

Таблица 4

Уровень вещания	Скорость видеопотока	Качество вещания		
		Одно кратковременное прекращение работы за 10 дней	Одно кратковременное прекращение работы в день	10 кратковременных прекращений работы в день
Доставка	270 Мбит/с	4×10^{-11}	4×10^{-10}	4×10^{-9}
Первичное распространение	40 Мбит/с	3×10^{-10}	3×10^{-9}	3×10^{-8}
Доступ к распространению	3 Мбит/с	4×10^{-9}	4×10^{-8}	4×10^{-7}

Учитывая вид (19), для цифрового телевидения без потери информации можно перейти к ограничениям вида (1):

$$P_{NR} \leq (k_{ГКСИ}^p)^{-1} P_{pl_пред}, B_{NR} \geq (k_{ГКСИ}^h)^{-1} P_{QoS_треб}. \quad (20)$$

Требования к временным сетевым параметрам качества T_{NR} и D_{NR} для услуги цифрового телевидения находятся в рамках классов QoS 6 и 7 Y.1541 (табл. 3) [10].

Взаимосвязь параметров качества передачи данных

На данный момент модель для получения оценки сквозного качества предоставления услуг передачи данных в терминах QoE предложена только для случая поиска и просмотра информации в Web. Взаимосвязь различных факторов, оказывающих влияние на воспринимаемое качество передачи данных, показана на рис. 11.

Воспринимаемое пользователем качество просмотра информации в Web в соответствии с G.1030 является функцией времени сеанса связи и вычисляется по одной из следующих формул:

$$MOS_{2-стр} = \frac{4}{\ln\left(\frac{0,011T_{max}}{T_{max}} + 0,47\right)} (\ln(T_{сеанс}) - \ln(0,011T_{max} + 0,47)) + 5, \quad (21)$$

$$MOS_{1-стр} = \frac{4}{\ln\left(\frac{0,005T_{max}}{T_{max}} + 0,24\right)} (\ln(T_{сеанс}) - \ln(0,005T_{max} + 0,24)) + 5, \quad (22)$$

$$MOS_{одн} = \frac{4}{\ln\left(\frac{0,003T_{max}}{T_{max}} + 0,12\right)} (\ln(T_{сеанс}) - \ln(0,003T_{max} + 0,12)) + 5, \quad (23)$$

где $MOS_{2-стр}$, $MOS_{1-стр}$, $MOS_{одн}$ – оценки MOS для двухстраничного сеанса связи поиска и просмотра информации, для одностраничного сеанса связи и для произвольных одностраничных сеансов (одиночное событие) соответственно; $T_{сеанс}$ – время сеанса связи; T_{max} – максимальное ожидаемое время сеанса связи.

Время сеанса связи $T_{сеанс}$ представляет собой время, в течение которого будет полностью загружена нужная Web-страница. В случае двухстраничного поиска и просмотра информации это сумма времени загрузки страницы поиска, самого поиска и загрузки запрошенной страницы. Время сеанса связи является параметром, отражающим совокупное качество работы сети, и его можно отнести к уровню QoS.

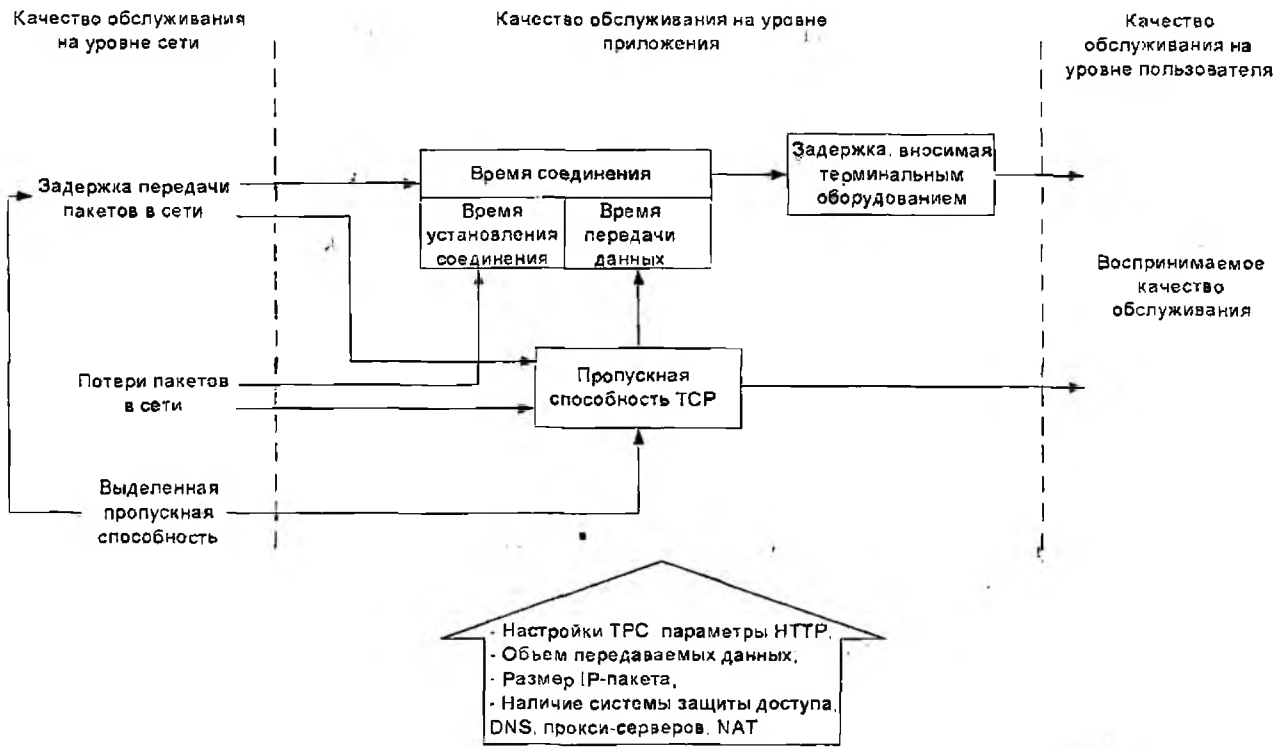


Рис. 11

При использовании протокола TCP для передачи данных время $T_{сетей}$ является функцией выделенной пропускной способности B_{NP} , задержки в сети T_{NP} и вероятности потери пакетов в сети P_{NP} . Основой для формирования зависимости $T_{сетей}(B_{NP}, T_{NP}, P_{NP})$ является формула пропускной способности TCP [4, 10]

$$B_{TCP} \approx \min \left(\frac{W_{max}}{RTT}, \frac{1}{RTT \sqrt{\frac{2bP_{pl}}{3}} + T_0 \min \left(1, 3 \sqrt{\frac{3bP_{pl}}{8}} \right) P_{pl} (1 + 32P_{pl}^2)} \right), \quad (24)$$

где B_{TCP} – приближительная модель пропускной способности TCP, пак/с; W_{max} – максимальный размер буферного окна получателя, пакеты; RTT – период кругового обращения, с, $RTT = 2T_{NP}$, b – количество пакетов, запрошенных для подтверждения приема полученным ACK; T_0 – время ожидания для повторной передачи не подтвержденного (потерянного) пакета

Для эталонной конечной точки TCP (TCP Рино) [10] приняты следующие параметры: максимальное окно 16, 64 или 256 *кбайт*, время ожидания $T_0 = 1$ с, $b = 2$ пакета (одно подтверждение ACK на два пакета). Характерной особенностью TCP-соединения является его способность предотвращать перегрузки, однако в результате таких действий появляются дополнительные потери пакетов (потери вследствие проверки протоколом TCP возможности передачи с недопустимыми параметрами) с вероятностью $P_{прое}$. Для оценки $P_{прое}$ можно воспользоваться упрощенной формулой пропускной способности TCP [4]

$$B_{TCP} < \frac{MSS}{RTT} \cdot \frac{C}{\sqrt{P_{прое}}}, \quad (25)$$

где C – постоянная, которая учитывает влияние случайных/периодических потерь и стратегии ACK, как правило, $C = 0,866$, MSS – максимальный размер сегмента TCP

Учитывая, что TCP ограничивает пропускную способность приблизительно до 75 %, то есть $B_{TCP} \approx 0,75B_{NP}$, имеем [4]:

$$P_{прое} \approx \frac{MSS}{RTT} \cdot \frac{C}{0,75B_{NP}} \quad (26)$$

Тогда суммарная вероятность потерь [4]

$$P_{пл} = 1 - (1 - P_{\text{пр}})(1 - P_{\text{прое}}) \quad (27)$$

При просмотре одной Web-страницы время загрузки приблизительно можно оценить как

$$T_{\text{заг}} = T_{\text{уст. соед.}} + T_{\text{пер}} \quad (28)$$

где $T_{\text{уст. соед.}}$ – время установления соединения TCP, $T_{\text{пер}}$ – время передачи содержимого Web-страницы.

Время передачи определяется объемом загружаемой страницы L (для страницы со сложной графикой $L \approx 1$ Мбит) и пропускной способностью TCP

$$T_{\text{пер}} = \frac{L}{B_{TCP}} \quad (29)$$

При незначительных потерях в сети, когда $P_{NP} < P_{\text{прое}}$, можно считать $T_{\text{пер}} = \frac{L}{0,75B_{NP}}$.

Время установления соединения представляет собой сумму всех временных задержек (RTT), которые имеют место на данном этапе. Для соединения, предложенного в G 1030 в качестве эталонного (рис. 12), это время DNS-опроса (26 мс) и время открытия TCP-соединения ($39 \text{ мс} + T_{NP}$) (табл. 5).

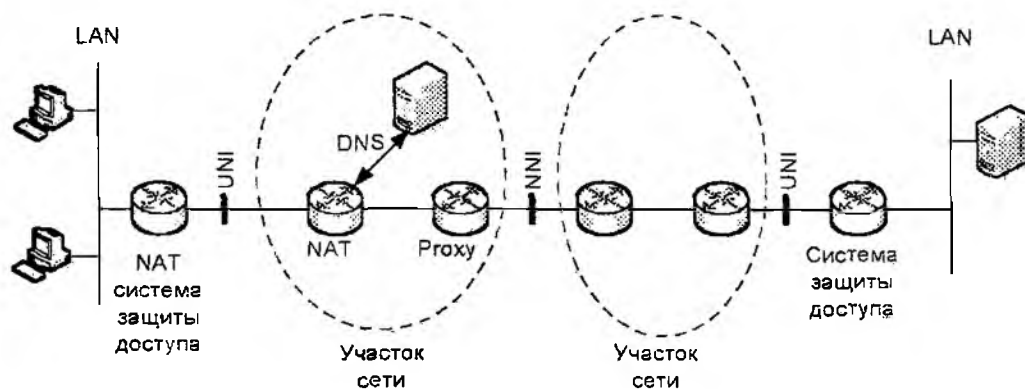


Рис. 12

Таким образом, формулы (21) – (29) представляют собой зависимость $QoS(NP)$ для просмотра Web-информации, которая может быть использована для формирования ограничений вида (2) исходя из требуемого уровня MOS . Для остальных видов услуг передачи данных (транзакции, e-mail, telnet, передача неподвижных изображений, интерактивные игры и пр.) формулы (24) – (29) отражают зависимость $QoS(NP)$ и могут быть использованы для формирования ограничений вида (3) исходя из требуемого QoS [4].

Математическая модель ТКС с гарантированным качеством обслуживания

Возьмем в качестве основы тензорную математическую модель ТКС [15 – 19], которая уже ориентирована на предоставление услуг гарантированного качества, однако реализуется это посредством ограничений в терминах NP (1). В рамках тензорного описания ТКС представляется возможной формулировка ограничений на качество предоставляемых услуг в виде (2) (или (3)) с последующим применением выражений (5) – (29) для оценки возможного уровня качества (в терминах MOS , R или QoS) в ходе поиска решения задачи управления и распределения сетевых ресурсов.

Время установления соединения TCP					
Время DNS-опроса			Время открытия TCP-соединения		
Источник задержки		Значение, мс	Источник задержки		Значение, мс
передачи (RTT)	обработки		передачи (RTT)	обработки	
LAN		2	LAN		2
	NAT/система защиты доступа	4		NAT/система защиты доступа	4
Линия		10	Линия UNI		10
NAT, линия к DNS		4	NAT, линия к прокси		4
	DNS	6		Прокси-сервер	4
			Участок сети		$T_{\text{сн}}$
			Линия UNI		10
				Система защиты доступа	2
			LAN		2
				Сервер (обработка TCP)	1

Математическая модель ТКС с гарантированным качеством обслуживания

Возьмем в качестве основы тензорную математическую модель ТКС [15 – 19], которая уже ориентирована на предоставление услуг гарантированного качества, однако реализуется это посредством ограничений в терминах NP (1). В рамках тензорного описания ТКС представляется возможной формулировка ограничений на качество предоставляемых услуг в виде (2) (или (3)) с последующим применением выражений (5) – (29) для оценки возможного уровня качества (в терминах MOS , R или QoS) в ходе поиска решения задачи управления и распределения сетевых ресурсов.

Как показано в [16 – 19], в рамках тензорного описания ТКС в условиях реализации динамических многопутевых стратегий управления трафиком, канальными и буферными ресурсами сетевые параметры качества применительно к обслуживанию θ -го класса трафика, передаваемого между z -й парой узлов сети, $z = \overline{1, Z}$, могут быть выражены следующим образом:

$$T_{NP(z)}^{(\theta)} = \left(\begin{matrix} \langle 4,1 \rangle E_{(-)z\pi\eta}^{(\theta)} & - & \left[\langle 4,2 \rangle E_{(-)z\pi\eta}^{(\theta)} \right] \left[\langle 4,4 \rangle E_{(-)z\pi\eta}^{(\theta)} \right]^{-1} \left[\langle 4,3 \rangle E_{(-)z\pi\eta}^{(\theta)} \right] \end{matrix} \right)^{-1} B_{NP(z)}^{(\theta)}, \quad (30)$$

$$P_{NP(z)}^{(\theta)} = 1 - \left(\begin{matrix} \langle 4,1 \rangle X_{(-)z\pi\eta}^{(\theta)} & - & \left[\langle 4,2 \rangle X_{(-)z\pi\eta}^{(\theta)} \right] \left[\langle 4,4 \rangle X_{(-)z\pi\eta}^{(\theta)} \right]^{-1} \left[\langle 4,3 \rangle X_{(-)z\pi\eta}^{(\theta)} \right] \end{matrix} \right)^{-1} B_{NP(z)}^{(\theta)}, \quad (31)$$

$$\text{где } \left\| \begin{matrix} \langle 1 \rangle E_{(-)z\pi\eta}^{(\theta)} & | & \langle 2 \rangle E_{(-)z\pi\eta}^{(\theta)} \\ \text{---} & | & \text{---} \\ \langle 3 \rangle E_{(-)z\pi\eta}^{(\theta)} & | & \langle 4 \rangle E_{(-)z\pi\eta}^{(\theta)} \end{matrix} \right\| = E_{(-)z\pi\eta}^{(\theta)}, \quad \left\| \begin{matrix} \langle 1,1 \rangle E_{(-)z\pi\eta}^{(\theta)} & | & \langle 4,2 \rangle E_{(-)z\pi\eta}^{(\theta)} \\ \text{---} & + & \text{---} \\ \langle 4,3 \rangle E_{(-)z\pi\eta}^{(\theta)} & | & \langle 4,4 \rangle E_{(-)z\pi\eta}^{(\theta)} \end{matrix} \right\| = \langle 4 \rangle E_{(-)z\pi\eta}^{(\theta)},$$

$$\left\| \begin{matrix} \langle 1 \rangle X_{(-)z\pi\eta}^{(\theta)} & | & \langle 2 \rangle X_{(-)z\pi\eta}^{(\theta)} \\ \text{---} & | & \text{---} \\ \langle 3 \rangle X_{(-)z\pi\eta}^{(\theta)} & | & \langle 4 \rangle X_{(-)z\pi\eta}^{(\theta)} \end{matrix} \right\| = X_{(-)z\pi\eta}^{(\theta)}, \quad \left\| \begin{matrix} \langle 4,1 \rangle X_{(-)z\pi\eta}^{(\theta)} & | & \langle 4,2 \rangle X_{(-)z\pi\eta}^{(\theta)} \\ \text{---} & + & \text{---} \\ \langle 4,3 \rangle X_{(-)z\pi\eta}^{(\theta)} & | & \langle 4,4 \rangle X_{(-)z\pi\eta}^{(\theta)} \end{matrix} \right\| = \langle 4 \rangle X_{(-)z\pi\eta}^{(\theta)}; \quad E_{(-)z\pi\eta}^{(\theta)}, X_{(-)z\pi\eta}^{(\theta)} - \text{проекции}$$

метрических тензоров E и X применительно к обслуживанию трафика θ -го класса, передаваемого между z -й парой узлов сети, в системе координат базисных контуров и узловых пар; $T_{NP(z)}^{(\theta)}$, $P_{NP(z)}^{(\theta)}$ – проекции в пространстве узловых пар тензоров задержки и вероятности

потерь пакета применительно к обслуживанию трафика θ -го класса, передаваемого между z -й парой узлов сети (их элементами являются межконцевые сетевые параметры качества). $B_{\lambda, \pi(z)}^{(\theta)}$ – пропускная способность, выделенная для трафика θ -го класса, передаваемого между z -й парой узлов сети.

В рамках тензорного представления ТКС важную роль играют двухвалентные тензоры E и X , которые зависят от используемой модели обслуживания в том или ином тракте передачи. Данные тензоры связывают тензоры сетевых параметров качества следующим образом:

$$\Lambda = ET, \Lambda = XP_n, \quad (32)$$

где Λ – одновалентный контравариантный тензор величин пакетной интенсивности трафика; T – одновалентный ковариантный тензор временных задержек передачи; P_n – одновалентный ковариантный тензор вероятностей своевременной доставки пакетов.

Используемые в (30) и (31) проекции тензоров E и X в пространстве контуров и узловых пар $E_{(z)\pi\eta}^{(0)}$ и $X_{(z)\pi\eta}^{(0)}$ могут быть получены координатным преобразованием при переходе в данное пространство из более простого пространства ветвей, представляющего собой совокупность отдельно взятых трактов передачи сети. Законы преобразования соответствующих проекций имеют вид [16–19]:

$$E_{(z)\pi\eta}^{(0)} = A'_{(z)v} E_{(z)v}^{(\theta)} A_{(z)}, \quad X_{(z)\pi\eta}^{(0)} = A'_{(z)v} X_{(z)v}^{(\theta)} A_{(z)},$$

где $E_{(z)v}^{(\theta)}$ и $X_{(z)v}^{(\theta)}$ – проекции соответствующих тензоров применительно к обслуживанию трафика θ -го класса, передаваемого между z -й парой узлов сети, в системе координат отдельных ветвей; $A_{(z)}$ – матрица ковариантного преобразования при переходе от системы координат отдельных ветвей к системе координат независимых контуров и узловых пар (определяет структурными свойствами сети).

В пространстве ветвей проекции тензоров $E_{(z)v}^{(\theta)}$ и $X_{(z)v}^{(\theta)}$ имеют вид диагональных $n \times n$ матриц, где n – количество трактов передачи. Выражение для элементов главной диагонали этих матриц зависит от используемой модели процессов обслуживания пакетов при рассмотрении тракта передачи как системы массового обслуживания.

Исходя из ориентации на сети NGN одним из требований к выбору модели процессов обслуживания пакетов является учет мультисервисного характера трафика, а также наличие различных классов обслуживания трафика. Как показано в [20], мультисервисный трафик является самоподобным (фрактальным) случайным процессом, для моделирования которого может использоваться фрактальное броуновское движение (fB) или распределение Парето (P).

В первом случае, при моделировании тракта передачи системой массового обслуживания вида $fB_0 / D / 1 / K$ или $fB_0 / M / 1 / K$, $K \rightarrow \infty$ имеем [20]:

$$e_{(z)v}^{(0)ii} = \frac{\lambda_{(z)v}^{(\theta)'}}{\frac{\rho_{\theta}}{m(1-\rho_{\theta-1})(1-\rho_{\theta})} + \frac{1}{\mu}};$$

$$x_{(z)v}^{(0)ii} = \lambda_{(z)v}^{(0)ii} \left[\log_2 \left(1 - \rho_{\theta} \Phi \left\{ \frac{1}{\sqrt{a\rho_{\theta}}} \left(\frac{1-\rho_{\theta}}{H_{\theta}} \right)^{H_{\theta}} \left(\frac{mK}{1-H_{\theta}} \right)^{1-H_{\theta}} \right\} \right) \right]^{-1},$$

где $i = \overline{1, n}$, $e_{(z)v}^{(0)ii}$, $x_{(z)v}^{(0)ii}$ – элементы главной диагонали проекций тензоров $E_{(z)v}^{(0)}$ и $X_{(z)v}^{(0)}$; $\lambda_{(z)v}^{(\theta)'}$ – интенсивность трафика θ -го класса, передаваемого между z -й парой узлов сети, в тракте, моделируемом ветвью i ; μ – интенсивность обслуживания пакетов в тракте i ;

$\rho_{\theta} = \sum_{k=1}^z \sum_{j=1}^{\theta} \frac{\lambda_{(k)v}^{(j)'}}{\mu}$ – относительная интенсивность суммарной входящей для тракта i нагрузки с

приоритетами $\overline{1, \theta}$; ρ_R – относительная интенсивность суммарной входящей для тракта i нагрузки (с учетом трафика всех приоритетов); m – параметр, характеризующий распределение длительности обслуживания пакетов в тракте i , $m=1$ для экспоненциального распределения и $m=2$ для фиксированного; H_θ – показатель самоподобности (Херста) для объединенного (в рамках i -го тракта) трафика θ -го класса; $\overline{\Phi\{\bullet\}}$ – дополнительная функция распределения стандартного гауссовского распределения; $a \in (-\infty, \infty)$ – параметр, отвечающий за положение гауссовского распределения (соизмерим со средним значением).

Во втором случае при моделировании тракта передачи системой массового обслуживания вида $\overline{P_\theta} / D / 1 / K$ или $\overline{P_\theta} / M / 1 / K$, $K \rightarrow \infty$ имеем [20]

$$e_{(z,r)}^{(\theta,m)} = \frac{\lambda_{(z,r)}^{(\theta)} \mu}{1 + \frac{\rho_R}{(1-\rho_{r-1})(1-\rho_r)} C'_z};$$

$$x_{(z,r)}^{(\theta,m)} = \lambda_{(z,r)}^{(\theta)} \left[\log_2 \left(1 - \frac{\rho_R - \rho_R^{k+2}}{1 - \rho_R^{k+2}} \frac{1 - \rho_r}{1 - \rho_r^{k+1}} \rho_r^{\frac{k}{C'_z}} \right) \right]^{-1},$$

где $i = \overline{1, n}$, $C'_z = \frac{C_u^2 + C_b^2}{2}$.

В результате выражения для оценок NP (30) – (31) в совокупности с выражениями (5) – (29) представляют собой расширение тензорной модели ТКС для возможности формирования и выполнения требований к качеству предоставляемых услуг в терминах его восприятия конечными пользователями. Наследуя от исходной модели динамический и многопутевой характер управления трафиком и ресурсами, полученное расширение обладает гораздо большей степенью гибкости по отношению к предоставляемому качеству.

Выводы

Таким образом, наличие формализованной взаимосвязи оценок качества обслуживания различных уровней (5) – (29) позволяет синтезировать управление сетевыми параметрами на основании требований конечных пользователей, сформулированных в виде интегральных оценок воспринимаемого качества MOS или R . Именно такой способ расширения математических моделей ТКС на предоставление услуг гарантированного качества является более гибким и позволит в полной степени учесть всю сложность взаимосвязи сетевых параметров NP в рамках QoS . Так, например, как следует из E-модели, воспринимаемое качество передачи речи определяется не столько абсолютными значениями задержки и вероятности потерь, сколько их взаимным расположением на плоскости $T_u - P_{rt}$ (см. рис. 9). Формулировка ограничений в виде (2) (или (3)) позволяет это учесть, не ставя жестких ограничений на значения отдельных параметров (вида (1)).

Список литературы: 1. Симоненко Д. В. Динамическая модель распределения сетевых ресурсов с поддержкой гарантированного качества обслуживания в территориально-распределенных ТКС // Наукові записки УНДІЗ. 2008. №3 (5). С. 39 – 46. 2. Дробот О. А. Комплексная модель обеспечения гарантированного качества обслуживания с реализацией динамических стратегий распределения сетевых ресурсов // Радиотехника: Всеукр. міжвед. науч.-техн. сб. 2007. № 148. С.43-54. 3. Рекомендації Державного департаменту з питань зв'язку та інформатизації Міністерства транспорту та зв'язку України Р 45-020-2007 Визначення вимог до показників та норм якості телекомунікаційних послуг та послуг поштового зв'язку. Загальні положення. 4. ITU-T Recommendation G.1030 Estimating end-to-end performance in IP networks for data applications. 5. ITU-T Recommendation G.109 Definition of categories of speech transmission quality. 6. ITU-T Recommendation P.800 Methods for subjective determination of transmission quality. 7. ITU-T Recommendation G.107: The E-model. a computational model for

use in transmission planning. 8. *ITU-T Recommendation G.1010 End-user multimedia QoS categories*. 9. *ITU-T Recommendation Y.1540 Internet protocol data communication service – IP packet transfer and availability performance parameters*. 10. *ITU-T Recommendation Y.1541 Network performance objectives for IP-based services*. 11. *ETSI Document ETR 003, October 1994, Network Aspects (NA). General aspect of Quality of Service (QoS) and Network Performance (NP)*. 12. *ITU-T Recommendation G.109 Amendment 1 New Appendix I – The E-model-based quality contours for predicting speech transmission quality and user satisfaction from time-varying transmission impairments*. 13. *ITU-T Recommendation G.113 Appendix I (2002). Provisional planning values for the equipment impairment factor I_e and packet-loss robustness factor B_{pl}* . 14. *ITU-T Recommendation G.1070 Opinion model for video-telephony applications*. 15. *Крон Г. Тензорный анализ сетей: Пер. с англ. М.: Сов. радио, 1978. 719 с*. 16. *Лемешко А. В. Особенности моделирования двухполюсной сети связи ортогональной сетью в рамках тензорного анализа // Радиотехника. Всеукр. межвед. науч.-техн. сб. 2002. Вып. 128. С. 16-25*. 17. *Лемешко А. В., Дробот О. А. Модель многопутевой QoS-маршрутизации в мультисервисной телекоммуникационной сети // Радиотехника: Всеукр. межвед. науч.-техн. сб. 2006. Вып. 144. С. 16-22*. 18. *Лемешко А. В. Вероятностно-временная модель QoS маршрутизации с предвычислением путей в условиях неидеальной надежности элементов телекоммуникационной сети // Радиотехника. 2005. Вып. 142. С. 11-20*. 19. *Лемешко А. В. Мультитензорная интерпретация решения маршрутных задач в телекоммуникационных сетях, представленных мнопродуктовыми многополюсными моделями евклидова пространства // Радиотехника и компьютерные системы. 2003. Вып. 3. С. 115-126*. 20. *Сычев К. И. Математические модели процессов формирования и обслуживания мультисервисного (самоподобного) трафика // Телекоммуникации. 2008. № 8. С. 19 – 25*.

*Харьковский национальный
университет радиозлектроники*

Поступила в редакцию 17.11.2008