End-to-End QoS Issues of MPEG4-FGS Video Streaming Traffic Delivery in an IP/UMTS Network

Thomas Pliakas[1], George Kormentzas[1] and Charalabos Skianis [1,2]

[1]Universiy of Aegean,
Department of Information and Communication Systems Engineering,
GR-83200, Karlovassi, GREECE
{tpliakas, gkorm}@aegean.gr

[2]National Centre for Scientific Research 'Demokritos',
Institute of Informatics & Telecommunications,
skianis@iit.demokritos.gr

Abstract. The paper addresses the end-to-end QoS problem of MPEG4-FGS video streaming traffic delivery over a heterogeneous IP/UMTS network. It proposes and validates an architecture that explores the joint use of packet prioritization and scalable video coding together with the appropriate mapping of UMTS traffic classes to the DiffServ traffic classes. A set of simulation scenarios, involving eight different video sequences, demonstrates the quality gains of both scalable video coding and prioritized packetization.

Keywords: DiffServ, End-to-End QoS, MPEG4-FGS, Packet Prioritization, UMTS.

1 Introduction

Fixed and wireless/mobile network operators face a twin challenge: to create and deliver attractive IP-based multimedia services quickly in response to fast-changing business and customer demands; and to evolve their current underlying networking infrastructures to an architecture that can deliver such services in a highly adaptable and guaranteed end-to-end Quality of Service (QoS) way both from network and application perspectives.

Simultaneously, on the customer side, for the next few years at least, there will be a wide variety of mobile/wireless access technologies supporting IP connectivity. These technologies include: mobile communication networks, such as GPRS [1] and UMTS [2], the family of broadband radio access networks, like IEEE 802.11 [3] and HIPERLAN [4], and wireless broadcasting technologies, like digital video broadcasting (DVB—satellite and terrestrial [5]).

IP technology seems to be able to resolve the interworking amongst the diverse fixed core and wireless/mobile access technologies at the network level. In this all-IP network, the end-to-end QoS provision concerning the network perspective could be established through the appropriate mapping amongst the QoS traffic classes/services supported by the contributing underlying networking technologies. Building on this context, this work concerns a DiffServ-aware IP core network and a UMTS access network and examines end-to-end QoS issues regarding MPEG4-FGS video streaming traffic delivery over such a network.

The Differentiated Services (DiffServ) model proposed by IETF support (based on the DSCP field of the IP header) two different services, the Expedited Forwarding (EF) that offers low packet loss and low delay/jitter and the Assured Forwarding (AF), which provides QoS guarantees better than the best-effort service. Differences amongst AF services imply that a higher QoS AF class will give a better performance (faster deliver, lower loss probability) than a lower AF class [6].

The QoS provision in Universal Mobile Telecommunications System (UMTS) is achieved through the concept of "bearers". A bearer is a service providing a particular QoS level between two defined points invoking the appropriate schemes for either the creation of QoS guaranteed circuits, or the enforcement of special QoS treatments for specific packets. The selection of bearers with the appropriate characteristics constitutes the basis for the UMTS QoS provision. Each UMTS bearer is characterized by a number of quality and performance factors. The most important factor is the bearer's Traffic Class; four traffic classes have been defined in the scope of the UMTS framework (i.e., Conversational, Streaming, Interactive and Background). The appropriate mapping of UMTS traffic classes to the aforementioned DiffServ service classes could offer a vehicle for the end-to-end QoS provision over a heterogeneous DiffServ/UMTS network. In our work, we employ and evaluate the three different mapping approaches presented in [7]-[9] respectively.

The Fine Grain Scalability (FGS) [10] feature of MPEG4 is a promising scalable video solution to address the problem of guaranteed end-to-end QoS provision concerning the application perspective. According to MPEG4-FGS, the Base Layer (BL) provides the basic video quality to meet the minimum user bandwidth, while the Enhancement Layer (EL) can be truncated to meet the heterogeneous network characteristics, such as available bandwidth, packet loss, and delay/jitter [11].

To address the end-to-end QoS problem of MPEG4-FGS video streaming traffic delivery over a heterogeneous IP/UMTS network, the paper proposes and validates through a number of NS2-based simulation scenarios an architecture that explores the joint use of packet prioritization and scalable video coding together with the appropriate mapping of UMTS traffic classes to the DiffServ traffic classes.

The rest of the paper is organized as follows. In Section 2, the proposed video coding and prioritization framework for providing QoS guarantees for MPEG4-FGS video streaming traffic delivery over a heterogeneous IP/UMTS network is presented. In Section 3, we demonstrate how video-streaming applications can benefit from the use of the proposed architecure. Finally, Section 4 draws the conclusions of this work.

2 Overview of the Proposed Architecture

Our architecture integrates the concepts of MPEG4-FGS video streaming, prioritized packetization based on content and DiffServ/UMTS classes coupling. The proposed architecture is depicted in Figure 1. It consists of three key components: (1)

MPEG4-FGS scalable video encoding, (2) simple prioritized packetization according to the type of content (I, P, B frame type), and (3) DiffServ/UMTS classes coupling in order to achieve QoS continuity of MPEG4-FGS video streaming traffic delivery over DiffServ and UMTS network domains. Each one of these components is discussed in detail in the following subsections.

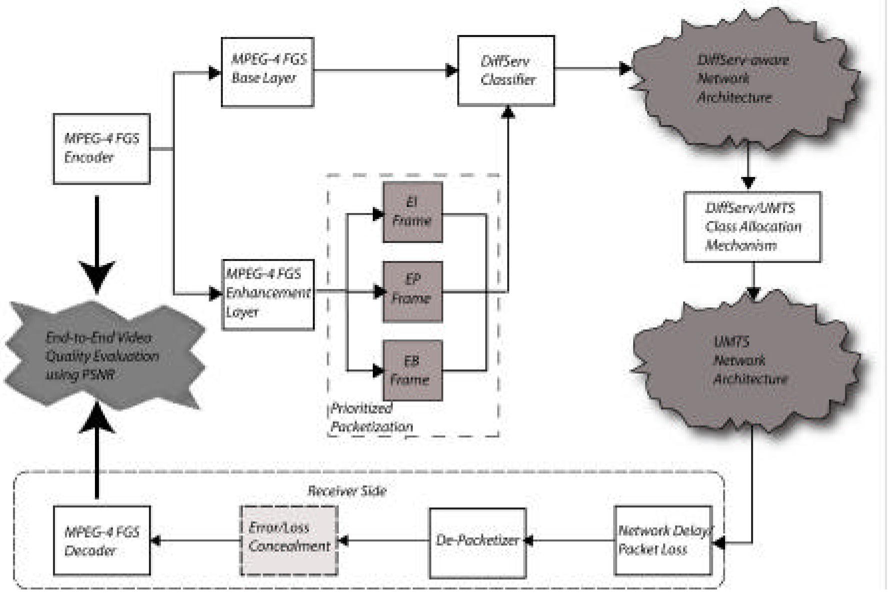

Fig. 1. Overview of proposed architecture

2.1 MPEG4 -FGS Scalable Video Coding

MPEG4-FGS scalable video coding constitutes a new video coding technology that increases the flexibility of video streaming. Similar to the conventional scalable encoding, the video is encoded into a BL and one or more ELs. For MPEG4-FGS, the EL can be efficient truncated in order to adapt transmission rate according to underlying network conditions. This feature can be used by the video servers to adapt the streamed video to the available bandwidth in real-time (without requiring any computationally demanding re-encoding). In addition, the fine granularity property can be exploited by the intermediate network nodes (including base stations, in case of wireless networks) in order to adapt the video stream to the currently available downstream bandwidth.

In contrast to conventional scalable methods, the complete reception of the EL for successful decoding is not required [13]. The received part can be decoded, increasing the overall video quality according to the rate-distortion curve of the EL as it described [14]. The overall video quality can be also improved from the error concealment method that is used. In our architecture, when a frame is lost, the decoder inserts a successfully previous decoded frame in the place of each lost frame.

A packet is also considered as lost, if the delay of packet is more than the time of the play-out buffer. (For the experiments discussed in the next Section III, this time is set to 1sec).

In order to measure the improvements in video quality by employing MPEG4-FGS, we adopt the Peak Signal-to-Noise Ratio (PSNR) and Structural SIMilarity (SSIM) [12] metrics. PSNR is one of the most commonly used objective metric to assess the application-level QoS of video transmissions and SSIM is a novel metric for measuring the structural similarity between two image sequences, exploiting the general principle that the main function of the human visual system is the extraction of structural information from the viewing field.

2.2 Prioritized Packetization

We define two groups of priority policies, one for BL and one for EL. These policies are used from Edge Router of the DiffServ-aware underlying network to mark the packets to the appropriate traffic classes. The packetization process can affect the efficiency as well as the error resiliency of video streaming. Fixed length packetization scheme is adopted for both BL and EL streams as proposed by the MPEG4 specification. Based on the content of each packet, we assign priorities according to the anticipated loss impact of each packet on the end-to-end video quality (considering the loss impact to itself and to dependencies). Each layer has a priority range, and each packet has different priority according to its payload. The packets which contain data of an I Frame are marked with lowest drop probability, the packets which contain data of a P Frame are marked with medium drop probability and the packets which contain data of a B Frame are marked with high drop probability.

Note that MPEG4-FGS specification assumes guaranteed delivery to BL and best-effort one to EL. In our framework, we use EF for transmitting BL and AF with different priorities for the EL based on the frame type. With assigned priorities, the packets are sent to underlying network and receive different forwarding treatments. Table 1 depicts the relation between the type of the EL content and the corresponding DiffServ classes. The first digit of the AF class indicates forwarding priority and the second indicates the packet drop precedence.

Table 1. DiffServ classes allocation for EL

Frame Type	DiffServ Classes
I Frame	AF11
P Frame	AF12
B Frame	AF13

2.3 DiffServ/UMTS Classes Coupling

The proposed MPEG4-FGS video streaming traffic delivery framework adopts three different coupling of DiffServ/UMTS classes approaches depicted in Table 2. Note that the actual QoS that can be obtained heavily depends on the traffic engineering for both UMTS and DiffServ networks.

Table 2. DiffServ/UMTS classes coupling

DiffServ Classes	UMTS Traffic Classes (Setting I) [7]	UMTS Traffic Classes (Setting II) [8]	UMTS Traffic Classes (Setting III) [9]
EF	Streaming	Conversational	Conversational
AF11	Interactive 1	Streaming	Streaming
AF12	Interactive 2	Streaming	Streaming
AF13	Interactive 3	Streaming	Interactive
BE	Background	Background	Background

3 Framework Evaluation

This section evaluates the performance of the proposed architectural framework through a set of experimental cases. A NS2- based simulation environment with the appropriate EURANE package extensions for simulating a UMTS network is adopted. We study the performance of our framework by enabling or disabling scalable video coding and/or by enabling or disabling prioritized transmission. The quality gains of scalable video coding in comparison with non-scalable video coding and the quality gains of prioritized transmission in comparison with non-prioritized transmission applying three different DiffServ/UMTS traffic classes mapping approaches are discussed in detail.

Fig. 2 depicts our simulation setup, which includes a DiffServ-aware autonomous system of a single 512Kbps wired link and a single UMTS cell of 1Mbps with the following rate allocation for the supported traffic classes: 200Kbps for the Conversional class, 300Kbps for the Streaming class, 200kbps for the Interactive 1 class, 100kbps for both Interactive 2 and 3 classes, and 200Kbps for the Background class. For the DiffServ-aware network the buffer management is considered to be WRED. The qualitative remarks being the outcome of our experiments can be also applied over more complex heterogeneous IP/UMTS infrastructures.

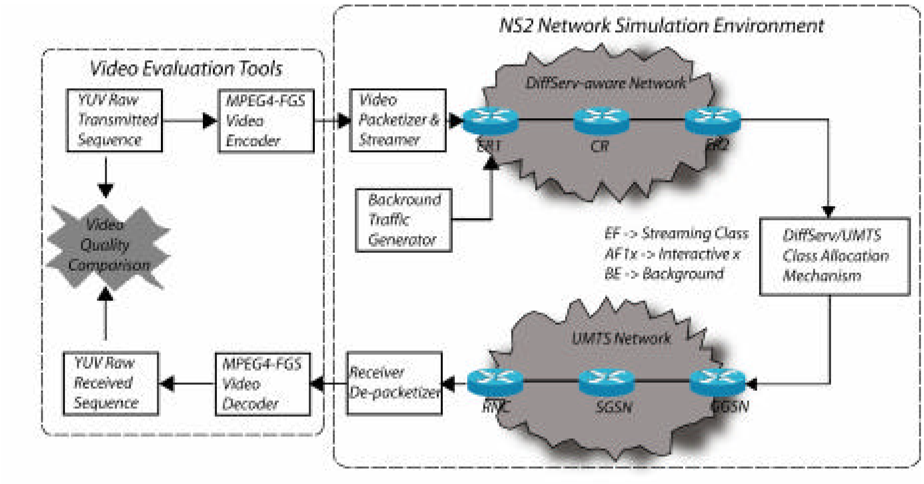

Fig. 2. Simulation Setup

Several YUY QCIF (176x144) video sequences consisting of 300 to 2000 frames are used as video sources. A number of background flows are also transmitted in the simulated network in order to fill in the respective DiffServ/UMTS class capacity in the link. The background traffic is increased from 210Kbps to 540Kbps leading the system in congestion.

The validation of the quality gains offered by the proposed framework concerns four simulation cases consisting in a number of experiments referring to eight different source video sequences transmissions over an all-IP network consisting of a DiffServ-aware IP core network and a UMTS access one.

The first simulation case refers to a single layer MPEG4 stream transmission. The video frames are sent every 33ms for 30fps video. For this simulation scenario, we use EF for transmitting I frames and AF12 and AF13 for transmitting P and B frames respectively. The mapping of DiffServ classes to the UMTS ones is performed through Table 2.

The second simulation case concerns a scalable MPEG4 stream transmission consisting in two layers. The BL packets are encoded using the MPEG4-FGS codec with MPEG2 TM5 rate control at 128kbps and the EL ones are encoded at 256kbps. For this case, we have direct application of Tables 1 and 2.

The third simulation case concerns a scalable MPEG4 stream transmission consisting in one BL and two ELs, i.e., EL1 and EL2. The encoding of BL packets remains at 128kbps as in the second simulation case, while the encoding of packets of both ELs is at 128kbps. For this simulation scenario, we use EF for transmitting BL, AF11 for transmitting EL1, and Best Effort (BE) for transmitting EL2. The mapping of DiffServ classes to the UMTS ones is performed through Table 2.

The fourth simulation case adopts the setup of the third case, while it applies the prioritized packetization scheme of the second case to the packets of the first EL, i.e., for this simulation scenario, we use EF for transmitting BL, Table 1 for transmitting EL1, and Best Effort (BE) for transmitting EL2.

Tables 3 to 5 depict the simulation results in terms of PSNR and SSIM video quality metrics for eight different YUV video sequences for all simulation cases (1 to 4) for the three settings (I to III) concerning Diffserv/UMTS classes coupling. For Setting I, each configuration case increases the video quality and the gain increment that offers each case is around 2db in terms of PSNR. For Setting II, the Cases 3 and 4 produce the same results.

Table 3. Quality Results for all Simulation Cases for Diffserv/UMTS classes coupling of Setting I

Video Sequences	***Number of Frames***	***Case 1***		***Case 2***		***Case 3***		***Case 4***	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bridge-Close	2001	25.455	0.0673	27.025	0.0772	29.565	0.0815	31.026	0.0896
Highway	2000	28.321	0.0761	30.658	0.0874	31.875	0.0937	33.451	0.0986
Grandma	871	28.365	0.0761	29.982	0.0832	31.453	0.0905	32.821	0.0949
Claire	494	27.981	0.0731	30.025	0.0896	31.751	0.0936	32.973	0.0978
Salesman	444	28.456	0.0762	31.563	0.0912	32.961	0.0957	34.361	0.0985
Foreman	400	29.012	0.0816	31.454	0.0905	33.568	0.0982	34.816	0.0993
Carphone	382	25.565	0.0675	28.234	0.0796	31.028	0.0896	32.564	0.0942
Container	300	24.545	0.0684	27.194	0.0784	29.729	0.0829	31.581	0.0912

Table 4. Quality Results for all Simulation Cases for Diffserv/UMTS classes coupling of Setting II

Video Sequences	***Number of Frames***	***Case 1***		***Case 2***		***Case 3***		***Case 4***	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bridge-Close	2001	32.123	0.0938	32.783	0.0942	29.213	0.0817	29.218	0.0817
Highway	2000	34.342	0.0987	34.632	0.0989	31.321	0.0908	31.341	0.0908
Grandma	871	34.943	0.0991	34.232	0.0984	31.763	0.0919	31.768	0.0919
Claire	494	33.231	0.0979	33.683	0.0977	32.497	0.0926	31.591	0.0927
Salesman	444	35.039	0.0996	35.913	0,0999	31.938	0.0937	31.942	0.0937
Foreman	400	35.725	0.0998	35.281	0.0997	32.321	0.0944	31.327	0.0943
Carphone	382	33.184	0.0983	33.432	0.0987	31.293	0.0915	31.284	0.0913
Container	300	32.718	0.0948	32.782	0.0948	29.123	0.0817	29.128	0.0817

Table 5. Quality Results for all Simulation Cases for Diffserv/UMTS classes coupling of Setting III

Video Sequences	***Number of Frames***	***Case 1***		***Case 2***		***Case 3***		***Case 4***	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Bridge-Close	2001	32.118	0.0936	33.562	0.0968	29.218	0.0818	29.217	0.0817
Highway	2000	34.212	0.0985	34.432	0.0987	31.319	0.0909	31.314	0.0907
Grandma	871	34.679	0.0988	34.782	0.0989	31.764	0.0917	31.763	0.0917
Claire	494	33.235	0.0979	33.783	0.0978	31.497	0.0925	31.489	0.0923
Salesman	444	34.671	0.0988	34.732	0.0990	31.942	0.0937	31.936	0.0937
Foreman	400	34.983	0.0995	35. 243	0.0997	32.316	0.0329	32.297	0.0328

Carphone	382	32.928	0.0953	33.421	0.0973	31.292	0.0982	31.286	0.0979
Container	300	32.594	0.0941	33.783	0.0978	29.432	0.0821	29.425	0.0821

For the Highway video sequence, we measure the packet/frame losses for I, P, and B frames for the four simulation cases for the three settings (I to III) concerning Diffserv/UMTS classes coupling. For Cases 3 and 4 the depicted measurements concern EL1. The results presented in Tables 6-8 are in accordance with the ones depicted in Tables 3-5. For Setting I, each case improves the previous one and Case 4 offers the best video quality gain as it experiences the lower packet/frame losses. For Settings II and III, Case 2 offers the best video quality.

Table 6. Packet/Frame losses for the Highway video sequence for Diffserv/UMTS classes coupling of Setting I

Frame Type	***Case 1 Frame Loss***	***Case 1 Packet Loss***	***Case 2 - EL Frame Loss***	***Case 2 - EL Packet Loss***	***Case 3 - EL1 Frame Loss***	***Case 3 - EL1 Packet Loss***	***Case 4 - EL1 Frame Loss***	***Case 4 - EL1 Packet Loss***
I	0.1%	3,4%	0.1%	3,1%	0.1%	2,1%	0.1%	0.1%
P	11.4%	12.6%	11.1%	11.9%	10.7%	11.8%	5.7%	6,3%
B	47,3%	47.7%	43,6%	43.9%	42,6%	42.8%	23.9%	27.8%

Table 7. Packet/Frame losses for the Highway video sequence for Diffserv/UMTS classes coupling of Setting II

Frame Type	***Case 1 Frame Loss***	***Case 1 Packet Loss***	***Case 2 - EL Frame Loss***	***Case 2 - EL Packet Loss***	***Case 3 - EL1 Frame Loss***	***Case 3 - EL1 Packet Loss***	***Case 4 - EL1 Frame Loss***	***Case 4 - EL1 Packet Loss***
I	0.1%	3.2%	0.1%	2.4%	0.1%	2.7%	0.1%	3.1%
P	6.3%	7.5%	5.7%	6.2%	5.6%	6.8%	5.5%	6.1%
B	19.7%	12.7%	16.7%	9.8%	15.6%	8.7%	15.4%	8.9%

Table 8. Packet/Frame losses for the Highway video sequence for Diffserv/UMTS classes coupling of Setting III

Frame Type	***Case 1 Frame Loss***	***Case 1 Packet Loss***	***Case 2 - EL Frame Loss***	***Case 2 - EL Packet Loss***	***Case 3 - EL1 Frame Loss***	***Case 3 - EL1 Packet Loss***	***Case 4 - EL1 Frame Loss***	***Case 4 - EL1 Packet Loss***
I	0.0%	0.0%	0.1%	1.8%	0.1%	1.2%	0.1%	1.7%
P	5.2%	7,8%	6.8%	11.3%	6.4%	7.2%	6.7%	7.1%
B	22.7%	23,8%	21.9%	23.5%	15.3%	17.1%	24.3%	26.8%

As an overall remark of the above results, we could note that Case 4 of Setting I could offer almost the same video quality as Case 2 of Settings II and III, without however employing conversational class.

4 Conclusions

Nowadays, continuous media applications over heterogeneous all-IP networks, such as video streaming and videoconferencing, are become very popular. Several approaches have been proposed in order to address the end-to-end QoS both from network perspective, like DiffServ and UMTS QoS traffic classes, and from application perspective, like scalable video coding and packetized prioritization mechanisms. The paper addresses the end-to-end QoS problem of MPEG4-FGS video streaming traffic delivery over a heterogeneous IP/UMTS network. It proposes and validates through a number of NS2-based simulation scenarios a framework that explores the joint use of packet prioritization and scalable video coding together with the appropriate mapping of UMTS traffic classes to the DiffServ traffic classes.

Acknowledgment

The work reported in this paper is carried out within the project "Study and Development of Interactive Broadband Services based on DVB-T/DVB-H Technologies" in the context of framework 2.2 of ‘Pythagoras II – Research Group Support of the University of the Aegean’ jointly funded by the European Union and the Hellenic Ministry of Education.

References

1. 3GPP, General Packet Radio Service (GPRS); Service Description, Tech. Spec. 3GPP TS 23.060. v3.12.0, June 2002; http://www.3gpp.org
2. Richardson KW. UMTS overview. *Journal of Electronics and Communication Engineering* 2000; 12(3): 93–100.
3. IEEE Std 802.11, 1999 Edition (ISO/IEC 8802-11: 1999).
4. Khun-Jush J, Schramm P, Malmgren G, Torsner J. ‘HiperLAN2: broadband wireless communications at 5 GHz’. *IEEE Communication Magazine* 2002; 40(6): 130–136.
5. ETSI, Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television, EN300-744, January 2001.
6. IETF Web site: http://www.ietf.org, DiffServ and IntServ Working Groups.
7. S. Maniatis, E. Nikolouzou and I. Venieris, “QoS Issues in the Converged 3G Wireless and Wired Networks”, *IEEE Communication Magazine*, Aug. 2002.
8, H. Wang, D. Prasad, O. Teyeb and H-P. Schwefel, “Performance Enhancements of UMTS networks using end-to-end QoS provisioning”, IWS 2005/WPMC’05, Aalborg, Denmark. Sept. 2005
9. R. Chakravorty, I. Pratt and J. Crowcroft, “A Framework for Dynamic SLA-based QoS Control for UMTS”, *IEEE Communication Magazine,* Oct. 2003
10. W. Li, "Overview of fine granular scalability in mpeg-4 video standard", *IEEE Transaction on Circuits and Systems for Video Technology,* vol. 11, no. 3, pp. 301–317, 2001.
11. H. M. Radha, M. van de Schaar, and Y. Chen, "The MPEG-4 fine grained scalable video coding method for multimedia streaming over IP", I*EEE Transactions on Multimedia*, vol. 3, no. 1, pp. 53–68, 2001.

12. Zhou Wang, Ligang Lu, and Alan C. Bovik, “Video Quality Assessment Based on Structural Distortion Measurement”, *Signal Processing: Image Communication*, special issue on “Objective Video Quality Metrics”, vol. 19, no. 2, pp. 121-132, Feb. 2004.
13. P. de Cuetos, M. Reisslein, and K. W. Ross, "Evaluating the streaming of fgs-encoded video with rate-distortion traces", Institute Eurecom Technical Report, Tech Report RR-03-078, 2003
14. P. Seeling, P. de Cuetos, and M. Reisslein, "Fine granularity scalable video: Implications for streaming and a trace based evaluation methodology", IEEE Communication Magazine, vol. 43, no. 4, pp. 138–142, 2005.