Expected Quality of Service (eQoS)
A Network Metric for Capturing End-User Experience

Gianluigi Pibiri, Ciarán Mc Goldrick, Meriel Huggard
School of Computer Science and Statistics
Trinity College Dublin, Ireland
Email: pibirig@tcd.ie, Ciaran.McGoldrick@tcd.ie, Meriel.Huggard@tcd.ie

Abstract—Quality-oriented network service provisioning can take place at the network level using estimates of Intrinsic Quality of Service (IQoS) parameters or at the user level through measurements of the end-user Quality of Experience (QoE). While IQoS parameters are quantitative and measurement based, QoE estimates are more difficult to obtain as they usually rely on subjective end-user reporting.

A new metric for the instantaneous estimation of QoE is proposed, expected Quality of Service (eQoS). This Perceived Quality of Service metric is calculated using IQoS parameters. eQoS estimates the QoE of common real time services for mobile devices (e.g. smartphones, tablets): Voice of IP via Constant Bit Rate, Audio and Video streaming via Variable Bit Rate.

The efficiency of the proposed eQoS metric is evaluated via a realisation of an infrastructure-based wireless network. Unlike existing QoE metrics, eQoS provides an instantaneous estimate of the perceived QoS thereby establishing eQoS as an essential parameter for inclusion in future traffic management algorithms.

I. INTRODUCTION

Internet Service Providers (ISPs) must be able to provision for the increase in real-time services engendered by modern mobile devices and in order to meet service guarantees made to their end-users, they must monitor and manage the network quality. Quality of Service (QoS) and Quality of Experience (QoE) [1] are two common metrics used to evaluate the quality of service provided in terms of physical parameters or final user opinion respectively.

Traditional QoS metrics may be used to estimate the user QoE but there are a number of key problems associated with this approach. Firstly, QoS measurements do not consider any direct relationship between network parameters and the quality as perceived, or assessed, by the end-user. Secondly, QoE measurements are not instantaneous and the time needed to collect these measurements is too long for them to be used for the adjustment of network parameters to the prevailing conditions.

The challenge is to obtain meaningful QoE measurements at the node and to then include these as parameters in the traffic management process. In this paper we define the Expected Quality of Service (eQoS), a perceived QoS metric obtained from intrinsic QoS parameters. In contrast to existing QoE metrics, eQoS is inferred from statistics gathered at the end-node and so it can be calculated instantaneously without the need for end-user input.

The definition of eQoS depends on the nature of the service (e.g. VoIP, Audio, Video-streaming) and the encoding used (e.g. G729, MPEG-4, AAC, MP3). eQoS is independent from the network architecture. In the evaluation of eQoS we mainly focus on the IEEE802.11[2] protocol, however it could also be considered for use in other wireless networks e.g. WiMax, 4G [2] and in future IEEE802.11 variants (e.g. 802.11ac [3] and IEEE802.11ae [4]).

In the following section we review existing QoS and QoE metrics. We then consider the common real-time services offered on networks: VoIP, audio and video. Expected Quality of Service (eQoS) is then defined and evaluated. Conclusions and directions for future research are reviewed in the final section of this paper.

II. QoS AND QoE

Three different sub QoS layers have been proposed [1]: Intrinsic QoS, Perceived QoS and Assessed QoS. Definitions of these are provided by the ITU [5] [6] [7] and the IETF [8].

Intrinsic QoS (IQoS) measures the physical network parameters and it indicates if a network can satisfy an arbitrary delivery criterion. IQoS metrics include packet delay, jitter, packet loss etc. IQoS does not provide any indication of the impact of packet loss on the end-user's QoE.

The Perceived QoS (PQoS) is related to the QoS perceived by the customer or end-user [1]. Four aspects of PQoS have been defined by the ITU [7]; two of these relate to the end-user and two relate to the ISP. End-users are interested in the required and perceived QoS; while ISP's focus on the offered and achieved QoS. The need to use PQoS in the evaluation of the quality of future networks is emphasised in [9] where an evaluation methodology to measure the four perceived QoSs is detailed.

The Assessed QoS (AQoS) is the quality calculated through subjective end-user evaluation. AQoS has been identified with Quality of Experience (QoE) [10]. It should be noted that PQoS and AQoS are different concepts: QoE (AQoS) is a measure of customer satisfaction while PQoS relates to the networks ability to provide the service.

A relationship between QoS and QoE is explored in [11]. This is termed the Exponential Interdependency of QoE and QoS (IQX). IQX is specified for VoIP via an exponential relationship between the number of packets dropped and the QoE. There are three main differences between IQX and the eQoS proposed in this work: firstly eQoS is designed to be calculated instantaneously at the node, secondly it is inferred from a statistical evaluation

of QoE and IQoS parameters, thirdly it is independent of the service (VoIP, audio or video) and the protocol used. A logarithmic relationship between network parameters and QoE is proposed and evaluated in [12].

Three main techniques to improve the QoS provided to streaming services have also been explored in the literature: The first technique is a simple scheme for the prioritization of streaming traffic [13]. The second technique combines data gathered from different layers of the protocol stack [14] e.g. through network feedback [15] or via an application to gather data from the end-user. The third technique considers the use of network errors [16]. All these approaches demonstrate the need for a mechanism to measure eQoS at the node.

QoE is usually quantized for streaming video and audio via the Mean Opinion Score (MOS) [17] scale. MOS was initially created as a subjective evaluation of voice quality [5], where the user provides ratings on a scale from 1 (poor) to 5 (good). The pesq algorithm [18] can be used to measure QoE for VoIP traffic.

Peak Signal-to-Noise Ratio (PSNR) [19] is commonly used to evaluate video quality through image comparison techniques. The PSNR scale is not linearly related to the MOS, however a method for translating PSNR to the MOS scale has been suggested [20].

The ITU has defined the Perceptual Evaluation of Video Quality (PEVQ) [21] and Perceptual Evaluation of Audio Quality (PEAQ) [22]. Both algorithms try to reproduce human perception of video and audio quality respectively, comparing the original data with that obtained after encoding, transmission and decoding.

III. REAL TIME SERVICES

Mobile devices are heavy users of real time communications services that are sensitive to packet loss and packet delay. The transmission of media files over the network is called "streaming" [23]. The media stream is transmitted using the UDP and TCP protocols or dedicated streaming protocols, such as Real-time Streaming Protocol (RTSP) [24], the Real-time Transport Protocol (RTP) [25] and the Real-time Transport Control Protocol (RTCP) [25]. Non-specialist protocols at the application layer, like HTTP, can be used for streaming. In this work the UDP protocol is considered for two main reasons. Firstly feedback (RTCP) is not of concern to us and secondly the UDP packet header is much smaller than that of RTP. The most popular standards for audio and video encoding are MPEG-4 Part 10 [26] called also Advanced Video Coding (AVC) or H.264 and Advanced Audio Coding (AAC) [27]

Delay sources can be classed as "intrinsic" where the sources of delay are the sampling delay, compression or coding delay, and the packaging, or "extrinsic" where the sources of delay are links and nodes: queueing delay, transmission delay and the switching or routing delay.

A. VoIP

For VoIP encoding, the speech is first sampled and transformed into a digital signal using the Pulse Code Modulation (PCM) [28], using a sampling frequency of 8 KHz [29]. The samples are then compressed using algorithms such as G711 and G729 [30]. In G729 the VoIP packet is usually composed of two $10ms$ compressed speech frames; this operation gives rise to a small increase in the delay. ITU recommendation G.114 defines acceptable limits for a voice delay of between $0-150ms$. While a more consistent delay of between $150-400ms$ is still acceptable, some quality problems will ensue. When the delay exceeds $400ms$ the communication can not be considered conformant.

B. Audio encoding

Advanced Audio Coding (AAC) [27] is a lossy encoding procedure that involves 3 important steps [27]: firstly the audio is subsampled and transformed from the time domain into the frequency domain, the signal is then filtered to remove perceptual components, finally the signal is quantised and encoded.

In our discussion of eQoS below, we consider audio encoded at 44KHz 16 bit stereo, however it should be noted that AAC allows for encoding of up to 48 channels at 96KHz. The number of frames (the smallest independent sample of encoded audio) per second is constant but the packet size varies. Therefore the number of frames per second is considered as the most appropriate unit in the calculation of eQoS for audio.

C. Video encoding

The most popular standard for video streaming is MPEG-4 part 10 known as H.264 or MPEG-4 Advanced Video Coding (AVC) [26], proposed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG).

A sequence of images is encoded in three kind of frames: I frame, P frames, and B frames [31]. I frames contain a reference image, completely independent of other frames, while P frames are coded and decoded using information from the last I frame. They are encoded with the Motion Compensator Algorithm [31]. P frames are used when the differences with the previous I frame are small. B frames, where B is for Bidirectional, are coded using information from the last and next I or P frames. The sequence of frames is called a Group of Pictures, labelled $GoP(N, M)$ [31], where N is the number of frames between two full image I frames and M is the number of frames between two anchor (I or P) frames. A high encoding profile includes more detail and the quality level is high while a baseline profile has low detail and does not include B frames. From the network perspective, the number of packets per frame and the GOP have to be considered carefully.

IV. EQOS

Definition 1. *eQoS is an instantaneous measure of the Perceived QoS and reflects the loss in quality per unit time.*

It is calculated from physical parameters and is shown to be reflective of the QoE. eQoS takes on values between 0 and 1 and can also be expressed as a percentile. A near zero value for eQoS means no perceived loss in quality while a value close to one indicates substantial reduced quality.

A key feature of eQoS is that it can be used to evaluate the quality by considering only delayed packets and lost packets at the access point It is assumed that packets arrive regularly at the AP queue, with a delay close to the theoretical minimum value for the network.

The sampling time for the IQoS parameters is calculated by considering the kind of service provided e.g. for audio service (G729) with PCM (8KHz), the sampling time is $1s$, while for video the encoding and sampling are based on a $1s$ interval (MP4, FPS).

eQoS is easy to calculate, an essential requirement on wireless networks with an high throughput (e.g. IEEE802.11ac). eQoS is calculated as sum of two contributions:

$$eQoS = F(P_D, P_d) + G(P_D + P_d), \qquad (1)$$

where: $F(P_D, P_d)$: is the loss contribution, $G(P_D, P_d)$: is the contribution due to loss of immediately adjacent packets, P_D: is the number of packets lost through dropping and P_d: is the number of packets lost due to excessive delay. Both P_D and P_d are calculated at the AP queue and the loss contribution $F(P_D, P_d)$ is:

$$F(P_D, P_d) = \alpha \frac{P_D + P_d}{P_T}. \qquad (2)$$

P_T is the number of packets each service needs to transmit (for G729 coding, P_T is 50 packets per second). α is a proportionality factor dependent on the service configuration and it will be discussed in more detail below. P_d is the number of packets lost when the max delay threshold is exceeded.

$G(P_D, P_d)$ is the loss in quality due to the dropping of immediately adjacent packets. It depends on the codec and captured by a parameter, β ; the choice of which will be discussed below.

If we consider a sequence of n packets transmitted in one second where k of these packets are dropped and all packets are equally likely to be dropped, then the probability that the k packets dropped are in sequence is:

$$p(n, k) = \frac{\binom{n}{k} - \binom{n-k+1}{k}}{\binom{n}{k}}. \qquad (3)$$

For G729 we iterate this process to calculate the probability up to $k = 25$ that at least a minimum of 3 packets in sequence will be dropped.

Figure 1 shows the probability two or more packets are dropped in sequence when 50 packets are transmitted and k packets are dropped. Some consideration needs to be given to specific cases: Firstly, if the sampling window is not a sliding window then the packets dropped or delayed
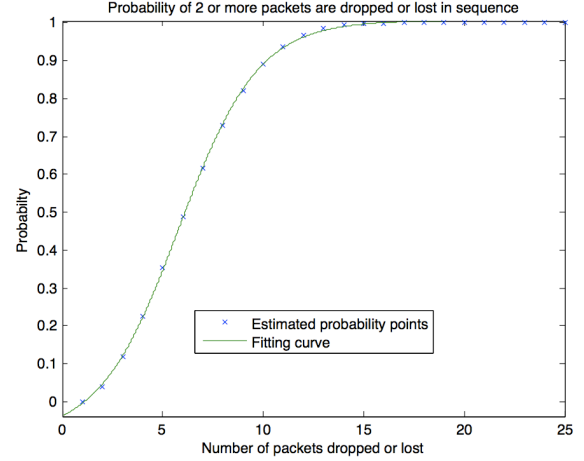

Figure 1. Estimated probability fitted by sigmodal (Boltzman) equation

in sequence between adjacent sampling windows need not be considered. Secondly, it is likely to be more probable that when a packet is dropped the next packet is dropped too. Thirdly, the probability of two packets being dropped in sequence is higher than that of finding larger packet drop groupings in sequence.

A sigmodal (Boltzman) equation has been established as an appropriate fit to the probability of sequential packet drops. The fitted curve for 50 packet transmissions per second (i.e. for G729) is shown in figure 1. The sigmoidal fits the (3) distribution for G729 (50 packets per second) with parameters $a \cong 1$, $b \cong -0.1$, $c \cong 5.8$ and $d \cong 2$. Combining this with our estimate of the loss contribution gives the following equation for the evaluation of eQoS for VoIP:

$$eQoS = \alpha \frac{P_D + P_d}{P_T} + \beta \left(-.1 + \frac{1.1}{1 + \exp(-\frac{P_D + P_d - 5.8}{2})} \right). \qquad (4)$$

A. The α and β parameters

eQoS is an instantaneous measure of the perceived QoS and it is calculated as the sum of two contributions, α and β, where $\alpha + \beta = 1$. If we look at the behaviour of eQoS as we vary α and β we obtain the family of curves shown in figure 2. We see that these vary from a straight line when the contribution from $G(P_D + P_d)$ is ignored (i.e. when $\beta = 0$) to the pure sigmodal curve when the contribution of $F(P_D, P_d)$ is zero ($\alpha = 0$).

A curve shape similar to the logarithmic one obtained by Reichl et al. [12] is obtained with the quantization:

$$\beta > \alpha \qquad (5)$$

α and β must be consistent with the desired and achievable quality, and they are easily set once the application details are known.

B. eQoS for audio streaming

Audio streaming does not give rise to CBR traffic due to the variable packet sizes it produces. However, the
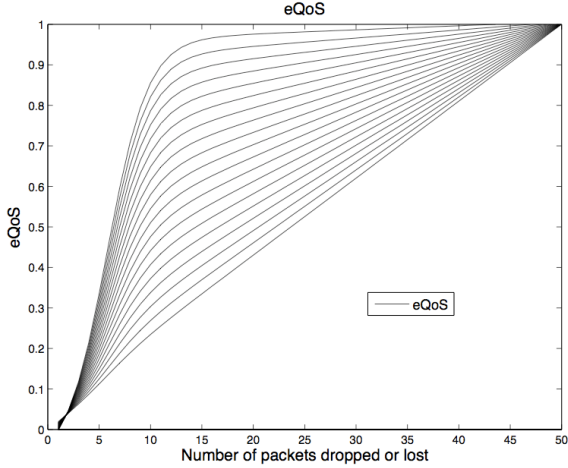
Figure 2.   eQoS family curves

smallest informational unit is the frame and each frame is encapsulated in a packet. It is possible to exploit the constant frame rate in each time unit (typically 43 frames per second(fps)) to define eQoS for audio streaming using an analysis similar to that used for VoIP traffic above.

$$eQoS = F(f_D, f_d) + G(f_D + f_d) \qquad (6)$$

Where $F(f_D, f_d)$ is the loss contribution, $G(f_D, f_d)$ is the contribution due to loss of immediately adjacent frames, f_D is the number of frames lost through dropping and f_d: is the number of frames lost due to excessive delay. The resulting family of curves obtained when α and β vary from the $F(f_D, f_d)$ contribution only (a straight line) to the $G(f_D, f_d)$ contribution only (sigmodal) are practically the same as shown in the VoIP case.

C. eQoS for video streaming

To define eQoS for VBR video streaming the most popular H.264 encoding used on mobile devices is considered with the basic profile configuration for a video movie with 30 fps. The file used in these simulations contains information for I and P frames only, with a ratio of around 1 I frame every 23 P frames. The loss contribution to eQoS is expressed as:

$$F(B_D, B_d) = \frac{B_D + B_d}{B_T}, \qquad (7)$$

where B_D is the number of bytes lost due to dropping, B_d is the number of bytes lost due to excessive delay and B_T is the total number bytes transmitted in the sampling time. The byte is the smallest information unit in terms of quality perceived. The most important frames, I frames, are usually larger than the less important P frames

The contribution due to the loss of adjacent frames is particularly important in the case of video streaming: Firstly because I frames are very important for video streaming quality, the loss of sequential packets that damage the I frame have a concomitant impact on of the P frames decoded using that I frame and this may affect the video for a number of seconds. Secondly the human

eye is sufficiently sensitive to be able to detect the loss of sequential P frames.

The adjacency contribution in case of I frames is obviously most significant than the adjacency contribution in case of P frames. For these reasons we need to define two adjacency contributions one for I fames and one for P frames. The resulting eQoS for video streaming is:

$$eQoS = \alpha F(B_D, B_d) + \beta G_I(P_{ID}, P_{Id}) + \gamma G_P(P_{PD}, P_{Pd}) \qquad (8)$$

where G_I and G_P are the adjacency contributions for I and P frames respectively. P_{ID}, P_{Id}, P_{PD} and P_{Pd} refer to packets dropped and delayed for I and P frames respectively. The sigmodal curves for I an P frames are similar to the VoIP case and are obtained by calculating the average number of packets and the number of frames per second.

D. eQoS parameterisation for Audio and Video

In the case of audio steaming the α and β parameters do not differ dramatically from the VoIP parameters chosen. Audio streaming is sampled at higher frequency than VoIP, also more tracks are usually available and the decoding applications are designed to reproduce the sound with the best quality. All these features result in the need to give more weight to the adjacency contribution (β).

When finding the eQoS for video streaming three parameters need to be set. The content (e.g. detail in the frames, difference between the frames, changes on the scene, colours) and the physical characteristics of the video (e.g. resolution, frames per second, packet size) must both be considered. Like the parameters used for VoIP and audio streaming α, β and γ must satisfy two equations:

$$\alpha + \beta + \gamma = 1, \qquad (9)$$

and

$$\beta + \gamma > \alpha. \qquad (10)$$

The latter gives more weight to the adjacency contribution for I and P frames than to the loss contribution. The values used in the simulations are α equal to 0.4, β equal to 0.3 and γ equal to 0.3. These three parameters values are conformant with the QoE estimation for videos with resolution close to 480x320 and a basic H.264 encoding.

V. EVALUATION

The first scenario considered was VoIP streaming where packet drops occur. In this simulation, using a modified version of [32], the instantaneous eQoS values were compared to the pesq (QoE) [18] measurements. The simulations were carried out by introducing packet drops into the standard VoIP data streams used in the ITU Software [18]. Computer simulations included packet drops due to buffer overflow or excessive delay for an infrastructure wireless network. A number of simulations were performed, each with a different number of packet drops per second, where packet drops occur randomly.

Figure 3 shows the scatter plot between eQoS and pesq with the linear regression line and the 95% Predictive
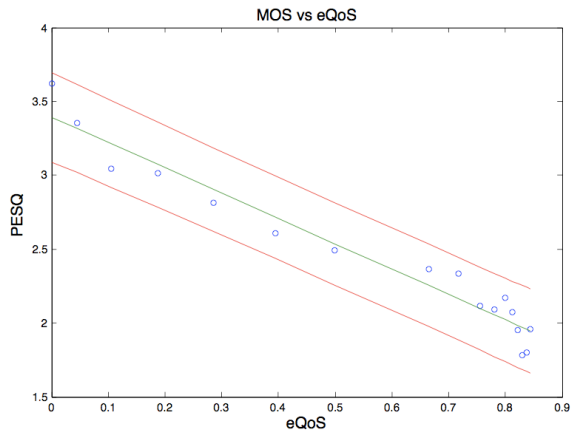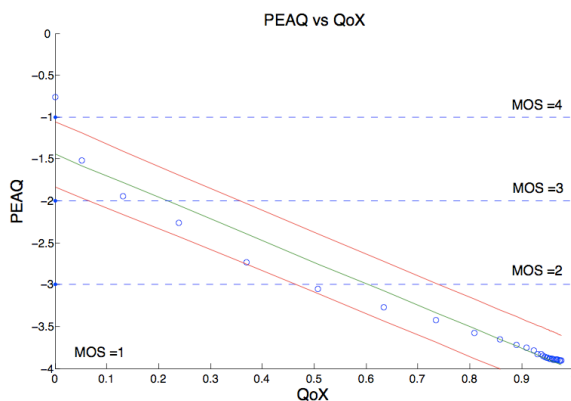
Figure 3. PESQ MOS vs eQoS for VoIP traffic

Figure 4. Regression graph of PEAQ and eQoS for audio streaming
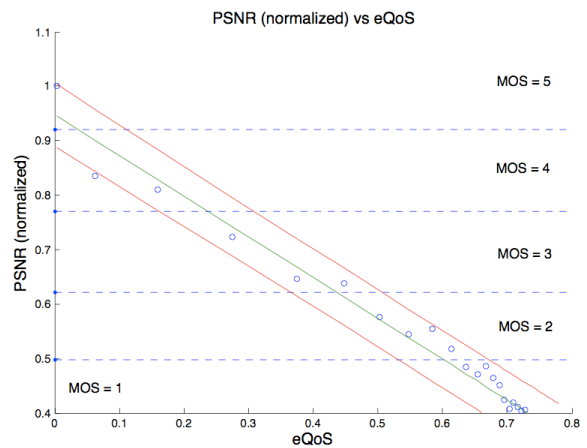

Figure 5. Regression graph of PSNR (normalized) and eQoS for video

the y-axis gives the normalised PSNR (where the PSNR values are divided by the maximum value). The horizontal dashed lines reports the MOS levels for PSNR.

The regression line and relative confidence interval suggest there is a strong linear relationship between PSNR and eQoS. The resulting coefficient of determination is:
`R-square=0.0.8894 adjR-square=0.8855`
The intercepts between the regression line and the MOS horizontal lines give the MOS points corresponding to the eQoS values.

The results show a strong relationship exists between eQoS and PSNR, therefore eQoS is well suited for instantaneous QoE estimation for video streaming at the AP, where packet loss is the only intrinsic QoS metric available.

VI. Conclusions and Future work

eQoS is a compelling metric for the estimation of QoE in an access point. It is obtained via straight forward calculations and is intended to be reflective of end-user stream quality. This makes it a powerful tool for network provisioning and traffic management. It provides a measurable link between well-established network parameters and the perceived end-user experience.

Acknowledgment

References

[1] R. Stankiewicz, P. Cholda, and A. Jajszczyk, "Qox: What is it really?" *Communications Magazine, IEEE*, vol. 49, no. 4, pp. 148–158, 2011.

[2] G. J. Mullet, *Wireless Telecommunications Systems and Networks*. Thomson, Delmar Learning, 2006.

[3] "Status of Project IEEE 802.11 Task Group AC." [Online]. Available: https://mentor.ieee.org/802.11/documents

[4] "Status of Project IEEE 802.11 Task Group AE." [Online]. Available: https://mentor.ieee.org/802.11/documents

Interval. It is evident that the scatter plot data is well inside the 95% PI . The high values of R-square and adjR-square also confirm the linear relationship between eQoS and pesq when no extrinsic delay is affecting the transmission.
`R-square=0.9507 adjR-square=0.9474`
The comparision of QoE and eQoS for audio and video was carried out using Evalvid [33] [34]. The Evalvid for NS-2 software [35] was modified to simulate audio streaming together with video streaming. The results comparing the calculated eQoS and the PEAQ [36], [37] are shown in figure 4. The regression shows a linear relationship exists between the two metrics where
`R-square=0.9772 adjR-square=0.9764`
The efficiency of eQoS for video streaming is examined through a comparison of PSNR/MOS and eQoS evaluations. PSNR is calculated by comparing the original video with the "damaged" video where loss has occurred. While it is recognised that PSNR is an imperfect metric for the estimation QoE, nevertheless it is widely used in the literature as a valid estimation measure [20]. PSNR values can be easily converted to the MOS scale [20], therefore it is suited to our proposes.

The values used for the parameters α, β and γ were as discussed above: 0.4, 0.3 and 0.3 respectively. Figure 5 shows the linear regression between PSNR/MOS and eQoS. The x-axis gives the reported eQoS values and

[5] "ITU-T recommendation E.800: Definitions of terms related to quality of service," Sep. 2008.

[6] "ITU-T recommendation E.802: Framework and methodologies for the determination and application of QoS parameters," Feb. 2007.

[7] "ITU-T recommendation G.1000: Communications Quality of Service: A framework and definitions," Nov. 2001.

[8] E. Crawley and R. Nair, "Rfc2386-1998," *A framework for QoS-based routing in the internet. USA: Network Working Group*, vol. 5, 1998.

[9] E. Ibarrola, J. Xiao, F. Liberal, and A. Ferro, "Internet qos regulation in future networks: a user-centric approach," *Communications Magazine, IEEE*, vol. 49, no. 10, pp. 148 –155, Oct. 2011.

[10] "ITU-T recommendation P.10/G.100: Vocabulary for performance and quality of service," Jul. 2006.

[11] T. Hoßfeld, D. Hock, P. Tran-Gia, K. Tutschku, and M. Fiedler, "Testing the IQX Hypothesis for Exponential Interdependency between QoS and QoE of Voice Codecs iLBC and G.711," University of Würzburg, Tech. Rep. 442, Mar. 2008.

[12] P. Reichl, S. Egger, R. Schatz, and A. D'Alconzo, "The logarithmic nature of QoE and the role of the Weber-Fechner law in QoE assessment," in *Communications (ICC), 2010 IEEE International Conference on*. IEEE, 2010, pp. 1–5.

[13] I. Ali, M. Fleury, S. Moiron, and M. Ghanbari, "Enhanced prioritization for video streaming over qos-enabled wireless networks," *Wireless Advanced (WiAd), 2011*, pp. 268–272, 20-22 June 2011.

[14] Q. Zhang, W. Zhu, and Y. Zhang, "Qos-adaptive multimedia streaming over 3g wireless channels," *MMSA2000) Nov*, pp. 9–10, 2000.

[15] P. Papadimitriou and V. Tsaoussidis, "End-to-end congestion management for real-time streaming video over the internet," in *Proc. 49th IEEE GLOBECOM*, 2006.

[16] A. Nafaa, T. Taleb, and L. Murphy, "Forward error correction strategies for media streaming over wireless networks," *Communications Magazine, IEEE*, vol. 46, no. 1, pp. 72–79, January 2008.

[17] "ITU-T recommendation P.910(04/08): Subjective video quality assessment methods for multimedia applications."

[18] "ITU-T recommendation P.862 (02/01): Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs; P.862 Amedment 2 (11/05): Revised Annex A – Reference implementations and conformance testing for ITU-T Recs P.862, P.862.1 and P.862.2; P.862 Corrigendum 1 (10/07)," Oct. 2007.

[19] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electronics letters*, vol. 44, no. 13, pp. 800–801, 2008.

[20] J. Gross, J. Klaue, H. Karl, and A. Wolisz, "Cross-layer optimization of OFDM transmission systems for MPEG-4 video streaming," *Computer Communications*, vol. 27, no. 11, pp. 1044–1055, 2004.

[21] "ITU-T recommendation J.247: Objective perceptual multimedia video quality measurement in the presence of a full reference," 08 2008.

[22] "ITU-T recommendation BS.1387: Method for objective measurements of perceived audio quality," 11 2001.

[23] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the h. 264/avc video coding standard," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 7, pp. 560–576, 2003.

[24] H. Schulzrinne, A. Rao, and R. Lanphier, "Real Time Streaming Protocol (RTSP)," RFC 2326 (Proposed Standard), Internet Engineering Task Force, Apr. 1998. [Online]. Available: http://www.ietf.org/rfc/rfc2326.txt

[25] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, RFC 3550 (Standard), Internet Engineering Task Force, Jul., updated by RFCs 5506, 5761, 6051, 6222.

[26] S. Kwon, A. Tamhankar, and K. Rao, "Overview of h.264/mpeg-4 part 10," *Journal of Visual Communication and Image Representation*, vol. 17, no. 2, pp. 186 – 216, 2006.

[27] K. Brandenburg, "Mp3 and AAC Explained," 1999.

[28] J. R. A. Bruce Carlson, Paul Crilly, *Communication Systems (4TH Edition)*. McGraw-Hill Science/Engineering/Math, June 25, 2001.

[29] (2003, Sep.) Understanding delay in packet voice networks. white paper. Available in http://www.cisco.com/warp/public/788/voip/delay-details.pdf, last visit in July 2012. Cisco Systems.

[30] "ITU-T recommendation G.729 : Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)," Jan 2007.

[31] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable h. 264/mpeg4-avc extension," in *Image Processing, 2006 IEEE International Conference on*. IEEE, 2006, pp. 161–164.

[32] C. Hoene, "Simulating playout schedulers for voip-software package," *URL: http://www. tkn. tuberlin. de/research/qofis*, 2004.

[33] J. Klaue, B. Rathke, and A. Wolisz, "Evalvid–a framework for video transmission and quality evaluation," *Computer Performance Evaluation. Modelling Techniques and Tools*, pp. 255–272, 2003.

[34] "http://www.tkn.tu-berlin.de/menue/research/evalvid/," *EvalVid - A Video Quality Evaluation Tool-set*.

[35] C. Ke, C. Shieh, W. Hwang, and A. Ziviani, "An evaluation framework for more realistic simulations of mpeg video transmission," *Journal of information science and engineering*, vol. 24, no. 2, pp. 425–440, 2008.

[36] "http://www-mmsp.ece.mcgill.ca/documents/downloads/pqevalaudio/," *PQevalAudio-v1r0*.

[37] P. Kabal, "An examination and interpretation of itu-r bs. 1387: Perceptual evaluation of audio quality," *McGill University*, 2002.