http://dx.doi.org/10.1109/EUROCON.2011.5929312

Semi-reliable Transport Protocol for IPTV over Mobile WiMAX

Ismail Ali, Salah Al-Majeed, Martin Fleury, and Mohammed Ghanbari
School of Computer Science and Electronic Eng.
University of Essex, Colchester CO4 3SQ, UK
{iaali, ssaleha, fleum, ghan}@essex.ac.uk

ABSTRACT
As Internet IPTV extends to mobile devices, suitable transport protocols are sought that can adapt streaming to wireless access networks. This paper proposes a semi-reliable video-rate protocol that provides selective retransmission of scalable video layers should channel packet loss occur. The semi-reliable protocol leads to good video quality and reduces end-to-end delay and start-up delay.

Keywords- error resiliency; H.264/SVC; mobile TV; transport protocols

I. INTRODUCTION

Internet Protocol Television (IPTV) is under active consideration for IEEE 802.16e (mobile WiMAX) [1]. A typical delivery chain [2], refer to Fig. 1, is from a Video Hub Office (VHO), perhaps receiving DVB from a satellite, then over an IP-based network before delivery over an access network. Rather than ADSL to a set-top box or PC, mobile devices receive video from a broadband wireless mast. Notice that to facilitate fast channel swapping [3] it is likely that through intelligent management, content will be placed relatively close to the receiver. This implies that a feedback channel for acknowledgments will not experience long delays. A key issue then becomes selection of a protocol that will provide end-to-end transport of a video stream.

In commercial systems, TCP acts in conjunction with Adobe Flash Player technology for progressive download, which is also known as pseudo-streaming. However, in [4], it was reported that 10% of viewers of YouTube may have interrupted downloads due to the jerky nature of the video download rather than due to poor content. Given TCP already has difficulty distinguishing congestion packet drops from packet losses through channel noise [5], reliable TCP-based delivery is unlikely to be satisfactory in the long term for delivery to mobile device. TCP-friendly Rate Control (TFRC) [6] employs unreliable UDP transport but imposes a version of TCP congestion control without the aggressive 'sawtooth' shaped sending rates, which are unsatisfactory for video streaming. Unfortunately, TFRC can lead to poor wireless channel utilization [7], as TFRC changes its sending rate by increasing the inter-packet gap and it does this whether there is packet loss from: buffer overflow through congestion; or through channel-corrupted packets.

TCP is too reliable for video streaming, as some packet losses can be tolerated. Instead, UDP video streaming has been used for fixed WiMAX [8]. However, too many UDP packet losses can seriously harm a compressed video stream. This is due to the predictive nature of video coding, which operates through motion compensation and entropy coding. Bell Labs introduced a reliable form of UDP, R-UDP, see [9], and there is also an R-UDP protocol employed by Microsoft in their MediaRoom product for IPTV service delivery over multicast networks. However, this paper's contribution is not to make UDP completely reliable but to provide reliability for prioritized video data. That is it provides a type of a semi-reliability. This occurs through a selective negative acknowledgment (NACK) facility. In the proposed scheme, a NACK is only sent once per lost packet in order to avoid introducing excessive delay, which would impact upon live IPTV channels. Selective NACK is used in conjunction with the Scalable Video Coding (SVC) extension of the H.264/AVC (Advanced Video Coding) codec [10]. This allows only base-layer or selected quality enhancement layer packets to be acknowledged for retransmission. Taking UDP transport of a video stream as the baseline, the proposed scheme can result in over 1.5 dB improvement in video quality (PSNR), depending on WiMAX cannel conditions. This is not as much an improvement in quality as unselective NACKs, again with only single retransmission. However, in compensation, selective NACKs trade quality against reduced delay.

Importantly, the overall sending period is reduced compared to unselective NACKs. Consequently, there will be a reduced start-up delay at the receiver to avoid the risk of stream interruptions. This in turn will lead to a smaller playout buffer at the mobile receiver, contributing to a reduced energy budget through passive and active energy consumption from memory access and refresh. The proposed transport for ease of reference is called Broadband Video Streaming (BVS) with the scalable variety being BVS-s. The paper now develops the BVS algorithm.

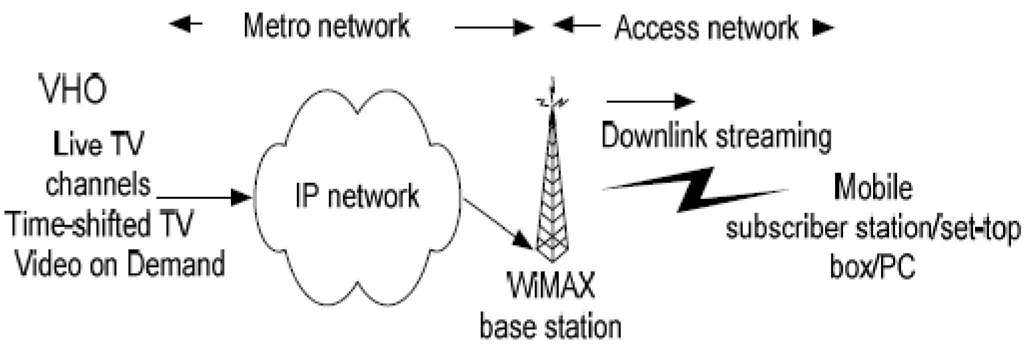

Figure 1. Typical IPTV delivery chain

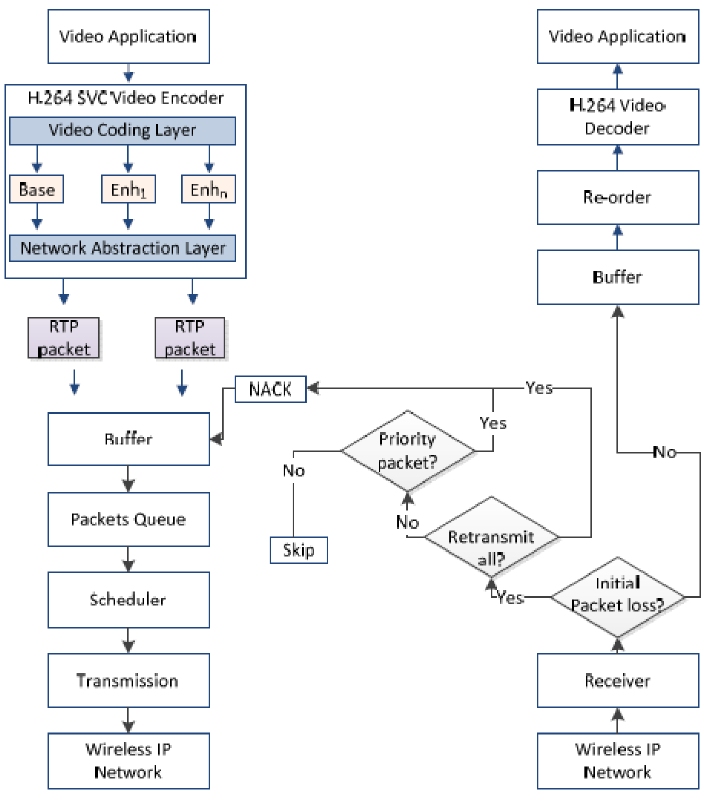

Figure 2. Operation of BVS.

II. BVS ALGORITHM

Fig. 2 is a general representation of the processing involved, showing the NACK response of the receiver. The following describes the operation assuming downlink streaming from a VHO to a mobile subscriber station (SS). At a mobile SS a record is kept of packet sequence numbers available through the RTP header. If an out-of-sequence packet arrives, a NACK may be transmitted to the BS in the next sub-frame for forwarding to the VHO. The SS only transmits a NACK if this is the first time that particular packet has been lost. If it is the first time and the non-selective NACK version of BVS (BVS-a in subsequent Figures) is in operation then a NACK is sent. However, I prioritized (BVSs in subsequent Figures) operation is in use then a decision is made according to the type of SVC packet that has been lost, reflecting the importance of that packet to the reconstruction of the video. Further discussion of how packets are selected is reserved for Section III.

Thus, an SS only transmits a NACK in BVS-s if a prioritized packet has been lost for the first time. The reorder buffer at the SS is a 'playout' buffer which may change the sending order of video data to fit the decode order. Upon receiving a NACK, the server prevents transmission from its input buffer until a single retransmission of the missing packet in the sequence has taken place. Not shown in Fig. 2, is a holding buffer that retains sent packets in the case of the need for a retransmission.

III. SVC STRUCTURE

Fig. 3 shows an illustrative size four SVC Group-of- Pictures (GOP) structure with hierarchical B-pictures (or P-pictures) between the key frames. In tests, a GOP size of 16 was employed. We have also utilized key pictures [10] which form the coarsest temporal layer and which serve to delimit the extent of drift within GOP borders. Partitioning the enhancement layer transform coefficients into multiple slices, medium-grained scalability (MGS), increases the granularity of quality scalability and enables packet-based quality scalable coding. Unfortunately, this arrangement introduces interpacket dependency, which, together with the predictive structure, introduces complex dependencies. If not preserved, these dependencies can result in video quality reduction when the decoder drops packets that depend on non-received or erroneous ones. Our proposed scheme negatively acknowledges the non-received or erroneous important packets shown shaded in Figure 4. These are the base SNR (quality scalability) layer packets together with selected packets from the first MGS sub-layer. Limitation of space in this paper does not permit further justification of the choice of packet type for selective NACK.

IV. EVALUATION

In Fig. 5's scenario, showing link delays and capacities, the VHO at position C streams video over the IP network. Various sources of congestion exist: node A sources to node B constant bit-rate (CBR) data at 1.5 Mbps with packet size 1 kB and sinks a continuous TCP FTP flow sourced at node B. Node B also sources an FTP flow to the BS and CBR data at 1.5 Mbps with packet size 1 kB. Video is transferred to the mobile SS via the WiMAX BS across the downlink (DL). NACKs are sent from the SS to the VHO at position C.

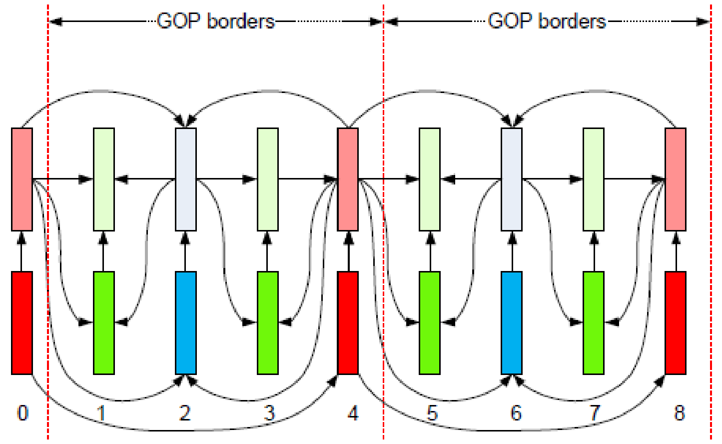

Figure 3. Illustrative coding structure for GOP size of four with base and single enhancement layer, showing the prediction structure. Pictures appearing at 0, 4, and 8 are coded as key pictures

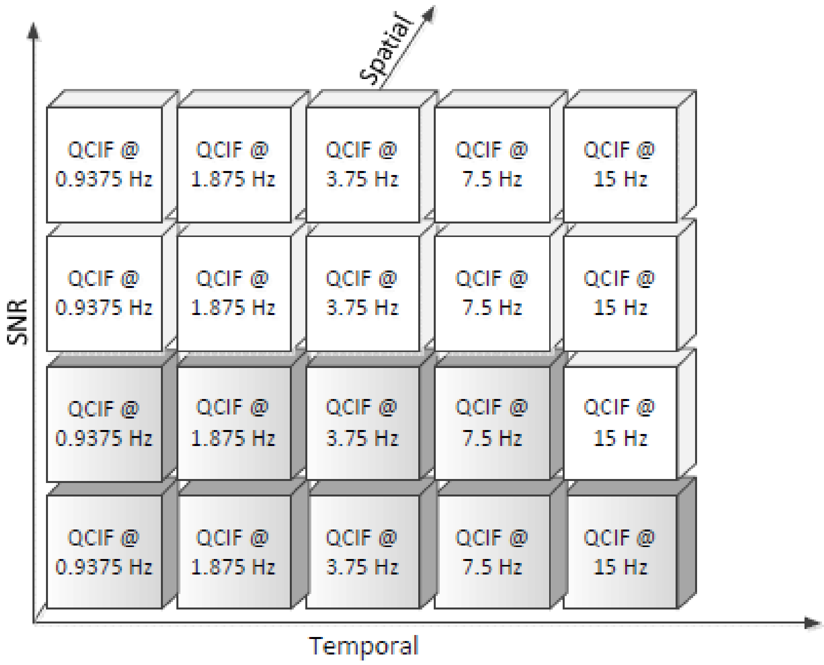

Figure 4. H.264 SVC mapping to BVS-s. Packets belonging to darkly shaded layers are those that are negatively acknowledged

The PHYsical layer settings selected for WiMAX simulation are given in Table I. The antenna is modelled for comparison purposes as a single half-wavelength dipole, whereas a sectored set of antenna on a mast might be used in practice. The antenna heights are typical ones taken from the standard [1]. The Time Division Duplex (TDD) frame length was set to 20 ms, as this value from the standard [1] reduces polling access delay for real-time services. The data rate results from the use of one of the mandatory coding modes [1] for a TDD downlink/uplink sub-frame ratio of 3:1. The BS is assigned more bandwidth capacity than the uplink to allow the BS to respond to multiple mobile devices. Thus, the parameter settings in Table I such as the modulation type and physical layer coding rate are required to achieve a data-rate of 10.67 Mbps over the downlink.

To judge the impact of BVS-S during streaming the well-known ns-2 simulator was augmented with a WiMAX module [11]. Data points are the mean of 25 runs, with 95% confidence intervals plotted in the results' graphs. As an input, 1065 frames of the *Paris* sequence at 30 Hz (display time 35.5 s) were encoded in Quarter Common Intermediate Format (QCIF) with the JSVM v. 9.19.1 encoder for H.264 SVC.

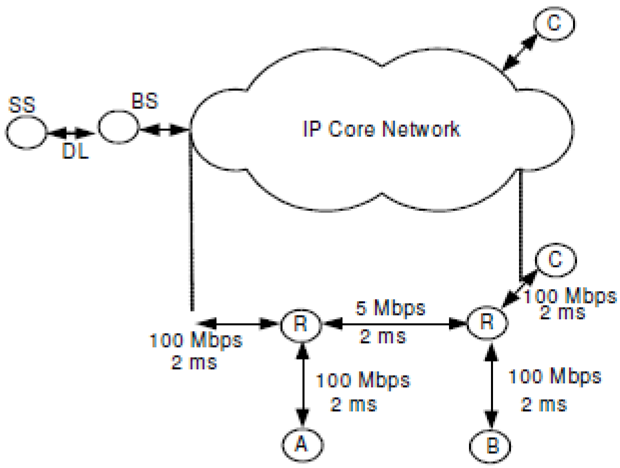

Figure 5. Video streaming scenario

TABLE I. SIMULATED WIMAX SETTINGS

Parameter	*Value*
PHY	OFDMA
Frequency band	5 GHz
Duplexing mode	TDD
Frame length	20 ms
Max. packet length	1024 B
Raw data rate	10.67 Mbps
IFFT size	1024
Modulation	16-QAM 1/2
Guard band ratio	1/16
DL/UL ratio	3:1
Channel model	Gilbert-Elliott
MS transmit power	250 mW
BS transmit power	20 W
Approx. range to MS	0.7 km
Antenna type	Omni-directional
Antenna gains	0 dBD
MS antenna height	1.5 m
BS antenna height	32 m

OFDMA = Orthogonal Frequency Division Multiple Access, QAM = Quadrature Amplitude Modulation, TDD = Time Division Duplex

Paris consists of two figures seated around a table in a TV studio setting, with moderate to high spatial-coding complexity. For comparison with other results, previous frame replacement was used as a simple form of error concealment. The video packet size was limited to a maximum of 1 kB. One

SNR enhancement layer was used with transform coefficients written to three MGS layers and, hence, a finer scalable granularity was achieved. The Quantization Parameters (QP) was set to 35 and 31 for the base and enhancement layer respectively. The motion estimation and mode decision QP for the base layer was set to 33 and to 29 for the enhancement layer. Consequently, the composite SVC average bit-rate was 69.5 kbps, with up to layer 2, 1, and base layer amounting to 58.6, 47.9, and 33.1 kbps respectively, corresponding (before packet drops) to mean PSNRs of 36.2, 35.3, 34.6 and 33.6 dB.

A Gilbert-Elliott two-state, discrete-time, ergodic Markov chain [12] modeled the wireless channel error characteristics at the ns-2 physical layer. Depending on burst length the impact of small-scale or large-scale fading is modelled. The probability of remaining in the good state was set to 0.95 and of remaining in the bad state was 0.94, with both states modelled by a Uniform distribution. The packet loss probability in the good state was fixed at 0.01 and the bad state default was 0.05. However, the bad state packet loss probability, P_B, was also varied as [0.01, 0.05 , 0.1 , …, 0.2]. In this way, we were able to judge the effect of worsening burst error channel conditions. We are aware that the channel could also be modelled more closely for propagation effects but it is the presence of burst errors [13] that mostly affect’s compressed video quality.

Fig. 6 compares the video quality of UDP video stream transport [8], reliable BVS-a (non-selective NACK with single retransmission) and BVS-s (semi-reliable with NACKs for selected SVC packets). As the WiMAX channel conditions deteriorate, then the two BVS schemes improve further upon UDP transport of the scalable video stream. As one might expect, BVS-a results in superior video quality but considering the relative gap in dB, BVS-s is closer to BVS-a than it is to UDP.

Against any reduction in video quality, should be judged the relative reduction in end-to-end delay, Fig. 7 and average overall sending time, Fig. 8. Mean end-to-end delay increases for BVS-a and BVS-s compared to UDP, because a packet’s delay includes the time to retransmit it if it is lost and all lost packets are retransmitted in BVS-a. However, not only does BVS-s achieve a lower mean packet end-to-end delay but its overall sending time is very much closer to the 35.5 s display time. Thus, the start-up time is significantly reduced. BVS-s also succeeds in reducing throughput as measured at the sender. In Fig. 9, compared to the constant UDP’s throughput, retransmissions result in a limited increase in throughput for both BVS-a and BVS-s, especially when channel conditions deteriorate. However, BVS-s’s share of the restricted WiMAX capacity is reduced.

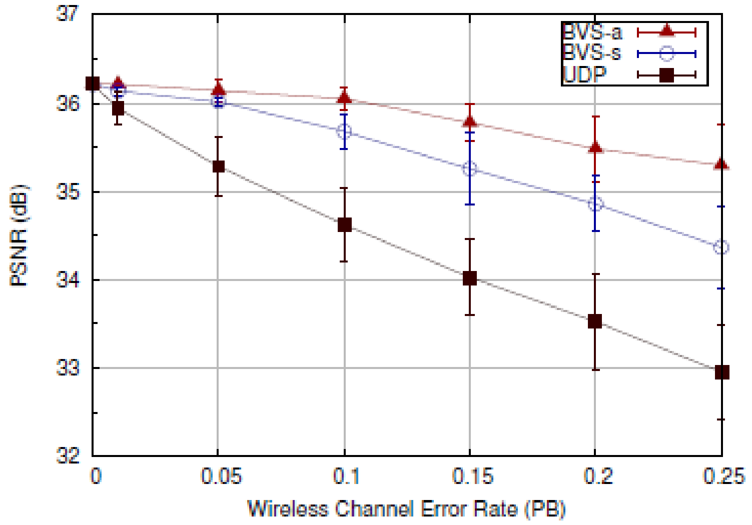

Figure 6. Received video quality of the various UDP-based schemes

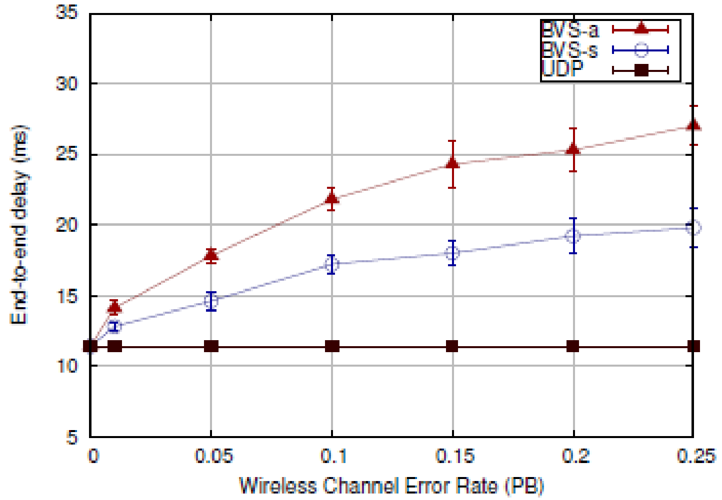

Figure 7. Mean packet end-to-end delay of the various UDP-based schemes

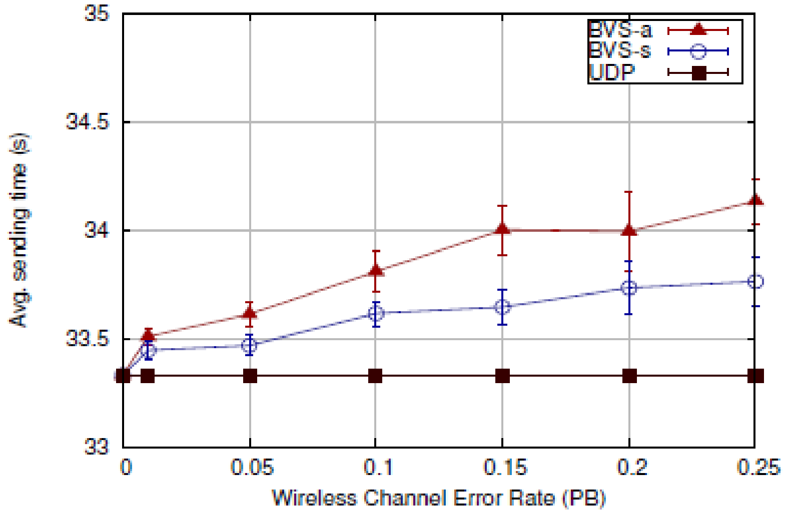

Figure 8. Avg. sending time for the various UDP-based schemes

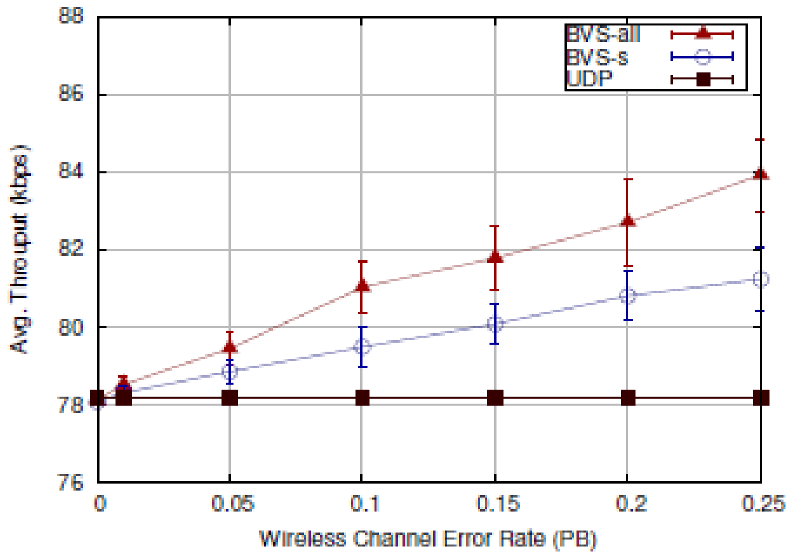

Figure 9. Throughput at the sender for the various UDP-based schemes

V. CONCLUSION

Clearly there is a trade-off between BVS-a and BVS-s in terms of balancing video quality (which is good in both cases for the *Paris* video sequence). In cases where delay is important, especially in proposed interactive varieties of live IPTV (e.g. the viewer responding to quiz shows) then BVS-s should be preferred. Start-up delay can also impact upon a consumer's preference for an IPTV scheme and BVS-s is preferable. Equally, if there is competition for WiMAX capacity then BVS-s should also be preferred.

REFERENCES

[1] IEEE, 802.16e-2005. IEEE Standard for Local and Metropolitan Area Networks. Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems, 2005.

[2] I. V. Uilecan, C. Zhou, and G. E. Atkin, "Framework for delivering IPTV services over WiMAX wireless networks," *IEEE Electro/Info.Technol. Conf.*, pp. 470–475, 2007.

[3] N. Degrande, K. Laevens, and D. De Vleeschauwer, "Increasing the user perceived quality for IPTV services," *IEEE Communs. Mag.*, vol. 46, no. 2, pp. 94–100, Feb. 2008.

[4] P. Gill, M. Arlitt, Z. Li and A. Mahanti, "YouTube traffic characterization: A view from the edge," in *7th ACM SIGCOMM Conf. on Internet Measurement*, 2007, pp. 15–18.

[5] H. Balkrishnan, V. Padmanabhan, S. Seshan, and R. Katz, "A comparison of mechanisms for improving TCP performance over wireless links," *IEEE/ACM Trans. Networking*, vol. 5, no. 6, pp. 756– 769, 2007

[6] M. Handley, J. Pahdye, S. Floyd, and J. Widmer, "TCP-Friendly Rate Control (TFRC): Protocol specification," RFC 3448, 2003.

[7] M. Chen, and A. Zakhor, "Rate control for streaming video over wireless," *IEEE Wireless Comms.*, vol. 12, no. 4, pp. 32–14, 2005.

[8] O. Issa, W. Li, and H. Liu, "Performance evaluation of TV over broadband wireless access networks," *IEEE Trans. roadcasting*, vol. 56, no. 2, pp. 201–210, 2010.

[9] C. Partridge, and R. Hinden, "Version 2 of the Reliable Data Protocol," RFC 1151, 1990.

[10] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1119, 2007.

[11] F. C.-D.Tsai et al. "The design and implementation of WiMAX module for ns-2 simulator," in *Workshop on NS2: The IP Network Simulator*, 2006, article no. 5.

[12] G. Haßlinger and O. Hohlfeld, “The Gilbert-Elliott model for packet loss in real time services on the Internet,” in *14th GI/ITG Conf. on Measurement, Modelling, and Evaluation of Computer and Commun. Systs.,* 2008, pp. 269–283.

[13] Y.J. Liang, J.G. Apostolopoulos, and B. Girod, “Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames,” *IEEE Trans. Circ. Syst. Video Technol.*, vol. 18, no. 7, pp. 861–874, 2008.