TUNING WIMAX FOR MULTI-CONNECTION VIDEO STREAMING

Salah Saleh Al-Majeed
London School of Commerce
White Hart Yard, London SE1 1NX

Martin Fleury
University of Essex
Wivenhoe Park, Colchester, UK

Uma Mohan
London School of Commerce
White Hart Yard, London SE1 1NX

ABSTRACT

Streaming a single video over multiple TCP-Friendly Rate Control (TFRC) connections is a promising way of separately coping with both wireless channel losses and traffic congestion, without the need for cross-layer intervention or retransmission delay at the data-link layer. At the same time, the wireless channel is properly utilized, as throughput improves with an increasing number of connections. Nevertheless over IEEE 802.16e (mobile WiMAX), tuning is needed to select the number of connections and the Time Division Duplex (TDD) frame size. The paper assesses the impact on video quality of packet drops due both to channel loss over a WiMAX access link and router buffer overflow across an all-IP network, consisting of broadband wireless access and core network. The paper also considers end-to-end delay and start-up delay when employing several connections. Results show that provided the TDD frame size is selected appropriately then using multiple connections preserves video quality and improves wireless channel utilization, with a minimal impact on end-to-end delay. As a trade-off, there is an increase in start-up delay arising from the need to avoid possible buffer underflow.

KEYWORDS

Multi-connections, TFRC, video streaming; WiMAX

1. INTRODUCTION

In multi-connection TCP-friendly Rate Control (TFRC) (Handley et al, 2003) video streaming, a *single* video source is multiplexed onto several connections across the wireless link in order to increase the throughput, thereby improving wireless channel utilization. By multiplexing a video stream across multiple connections it is hoped that the impact of packet loss on one or more of these connections will be mitigated by the datarate across the remaining connections. TFRC's main role when congestion occurs across the network path is to reduce the video streaming data rate across the wired portion of the concatenated network. It does this in response to packet drops at intermediate routers, which signal the presence of contending traffic. Unfortunately, TFRC can misinterpret as congestion packet losses due to wireless interference and noise, leading to a reduction in wireless channel utilization. Though cross-layer approaches to avoid misinterpretation are possible, these are complex to implement and inflexible. In fact, cross-layer approaches are most appropriate when a network has a fixed application, not one in which multimedia streaming is mixed in with other types of traffic

In pioneering work on multi-connection TFRC, in MULTTFRC (Chen and Zakhor, 2006) improved video quality comes about by increasing the quantity of video data that can be sent over the multiple connections. Of course, increased video data implies a lower compression ratio and, hence, higher quality delivered video, provided the rise in packet losses across the wireless channel does not degrade the quality. If burst errors occur then during the time that they occur all connections are affected, leading to a rise in packet losses, which was countered in Chen and Zakhor (2006) by means of application Forward Error Control (FEC). Unfortunately, if the number of connections varies, as it does in Chen and Zakhor (2006), then sending rate oscillations can occur. If the compression rate was varied at the source (either by changing the quantization parameter at the codec if live video or through a bit-rate transcoder) then oscillations in bitrate run the risk of disconcerting changes in displayed video quality.

However, we show that delivered video quality is maintained without the need to dynamically change the compression ratio by keeping the number of connections constant. This is because, with multiple TFRC connections, TFRC is better able to control its sending rate. In fact, TFRC (Handley et al, 2003) was originally designed with a high number of streams in mind, as may arise from a Video-on-Demand server, and special measures are recommended if the number of contending flows is *not* large enough. We consider an IEEE 802.16e (mobile WiMAX) (IEEE, 2005) uplink, which is the access network stage of an all-IP network (Lin and Pang, 2005). There is interest in uplink media services as a complement to IPTV video broadcasting. In this environment, the paper assesses the impact on video quality of packet drops due both to channel loss and router buffer overflow. It should be remarked that in Chen and Zakhor (2006) there was no investigation of actual video quality beyond the packet loss statistics.

The paper also considers end-to-end delay and start-up delay when employing several connections. Results show that provided the TDD frame size is selected appropriately then using multiple connections preserves video quality, as a result of the differential effect of packet loss patterns. Wireless channel utilization is considerably improved, with a minimal impact on end-to-end delay. As a trade-off, there is an increase in start-up delay arising from the need to avoid possible buffer underflow, though this is smaller compared to TCP-based streaming, when large buffers are normally employed (Hsiao, Kung, and Tan, 2003) to balance the possibility of repeated retransmissions. For example in Shen et al (2009) a buffer size of 10 s was required to counter wireless burst errors. However, for a WiMAX uplink this only becomes apparent if the Time Division Duplex (TDD) frame length is tuned to avoid queue servicing scheduling delays. The frame length is significant as a longer frame reduces delay at a WiMAX subscriber station, thus permitting more data to be removed from queues when the subscriber station's queues are polled.

It has been suggested in Tappayuthpijam et al (2009) that in Long Term Evolution (LTE) (Ekstrom, 2006) packet loss can be virtually eradicated by retransmission at the data-link layer. However, that approach has the potential to introduce unbounded delay across the wireless link, apart from the drop in throughput that results,. This reintroduces the problems that led to the search for an alternative to TCP transport for multimedia streaming. There is also the overhead of maintaining state at the evolved node B (an LTE radio head) and the delay arising if retransmissions are still occurring when a handoff occurs. Therefore, we consider that further investigation of multiple TFRC connections is a way forward in these networks, especially if there is a further wired network present beyond the wireless access link. This paper now details the network over which multi-connections video streaming takes place.

2. SCENARIO INVESTIGATED

The scenario tested in this paper is shown in Figure 1. The following describes the WiMAX part and this description is followed by a description of the inset, showing traffic sources and sinks within the core IP network.

2.1 WiMAX system

In Figure 1, once a Base Station (BS) has allocated bandwidth to each subscriber station (SS), each SS must manage its queue according to the data arrival rate from user applications. WiMAX networks support multiple service classes to accommodate heterogeneous traffic with varying requirements. WiMAX's rtPS is most suitable for real-time video services, particularly for Variable Bitrate Video (VBR), which is employed to maintain delivered video quality but may lead to 'bursty' arrival rates. Other congesting traffic is assumed to enter the non-real-time Polling Service (nrtPS) queue at the SS. In our experiments for both queues, a drop-tail queuing discipline was simulated. Queue sizes in tests were all set to fifty packets. This value was selected as it seems appropriate to mobile, real-time applications for which larger buffer sizes might lead both to increased delay and also greater active and passive energy consumption at the buffer's memory. Access to the SS service class queues was round-robin.

The physical layer (PHY) settings selected for WiMAX simulation are given in Table 1. The antenna is modeled for comparison purposes as a half-wavelength dipole. The Gilbert-Elliott 'bursty' channel model is further explained in Section 2.5. The TDD frame length was varied in experiments, because, as mentioned in Section 1, it has an important effect on the service rate at an SS. Current implementations have apparently

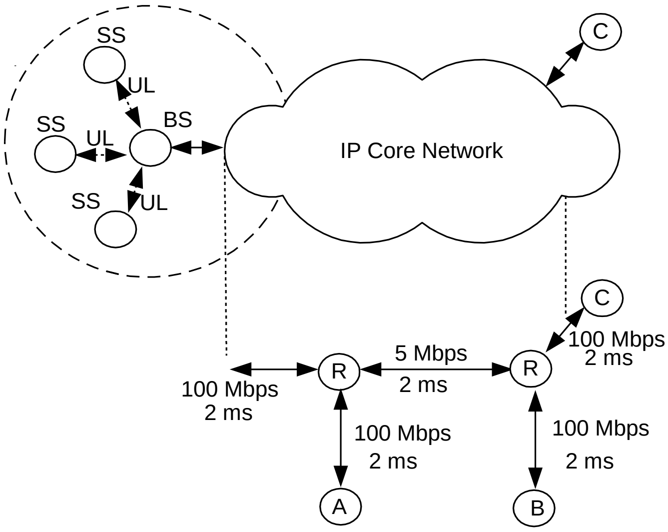

Figure 1. Network scenario with inset showing routing across the core network, A, B and C being sources and sinks, and R = router

Table 1. Simulated WiMAX settings

Parameter	*Value*
PHY	OFDMA
Frequency band	5 GHz
Duplexing mode	TDD
Frame length	5 to 20 ms
Max. packet length	1024 B
Raw data rate	10.67 Mbps
IFFT size	1024
Modulation	16-QAM 1/2
Guard band ratio	1/8
DL/UL ratio	1:3
DL length	1.25 to 5 ms
UL length	3.75 to 15 ms
Channel model	Gilbert-Elliott
MS transmit power	250 mW
BS transmit power	20 W
Approx. range to SS	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

mostly opted for a fixed 5 ms TDD frame size, though this is difficult to verify. The uplink (UL)/downlink (DL) is adaptable at the BS and is set to favor the UL for the purposes of our tests.

2.2 WiMAX traffic characteristics

There were three SSs communicating to the BS, with one of the SS sending a VBR video sequence encoded with the H.264/Advanced Video Codec (AVC) (Wiegand, 2003) and split between the multiple TFRC connections. The other SSs are introduced as sources of contending traffic across the wireless link and do not indicate the likely size of a WiMAX network, which obviously could be larger. A trace file was input to the well-known network simulator ns-2 and packet losses recorded in the output. The output serves to calculate the PSNR. Video quality comparisons were made under the EvalVid (version 2) environment (Klaue et al, 2003). Data points are an average of fifteen runs. The output serves to calculate the Peak-Signal-to-Noise ratio (PSNR). As a test, we used the 'Paris' clip H.264 Variable Bit Rate (VBR)-encoded at 30 Hz (frame/s) at Common Intermediate Format (CIF) (352×288 pixel/frame) with initial quantization parameter set to 26 (from a range 0 to 51). 'Paris' consists of two figures seated around a table in a TV studio setting, with high spatial coding complexity. The Intra-refresh rate was every 15 frames with IPBB…I structure, i.e. the GOP size was 15. 1063 frames were transmitted. Previous Frame Replacement (PFR) was set for error concealment at the decoder for comparison with coding results, which assume PFR. The slice size was fixed at the codec as 900 B. In selecting codec determination of slice size, packet segmentation is avoided, which improves video quality, as slices are not separated from their resynchronization headers.

Table 2 records the simulated traffic characteristics for the three SSs communication with the BS. Network Adaptation Layer units (NALUs) from the H.264 codec were encapsulated with Real Time Protocol (RTP) headers. After the addition of IP headers, these in turn formed a single WiMAX MAC Packet Data Unit (MPDU), which are variable-sized WiMAX packets. For simplicity, a WiMAX MPDU is now referred to as a packet. Coexisting rtPS queue CBR sources were all sent at 1500 kbps, i.e. at a similar rate to the video source. The inter-packet gap was 0.03 s for the CBR traffic. The FTP applications, which continuously supplied data according to available bandwidth, were set up out of convenience as a way of occupying the nrtPS queues; otherwise a Best-Effort (BE) queue might be more appropriate. Likewise, the DL traffic is selected to fully occupy the DL link capacity.

For TFRC, the inter-packet sending time gap was varied according to the TFRC equation (Handley et al, 2003), not the simplified version reported in Chen and Zakhor (2006). As described in Handley et al (2003), TFRC is a receiver-based system in which the packet loss rate is found at the receiver and fed-back to the sender in acknowledgment messages. The sender calculates the round-trip time from the acknowledgment messages and updates the packet sending rate. A throughput equation models TCP New Reno to find the sending rate:

$$TFRC(t_{rtt}, t_{rto}, s, p) = \frac{s}{t_{rtt}\sqrt{\frac{2bp}{3}} + t_{rto}\min\left(1, 3\sqrt{\frac{3bp}{8}}\right)p(1+32p^2)} \tag{1}$$

where t_{rtt} is the round-trip time, t_{rto} is TCP's retransmission timeout, s is the segment size (TCP's unit of output) (herein set to the packet size), p is the normalized packet loss rate, w_m is the maximum window size, and b is the number of packets acknowledged by each ACK. b is normally set to one and $t_{rto} = 4t_{rtt}$. It is important to notice that t_{rto} comes to dominate TFRC's behavior in high packet loss regimes, which is why it is unwise to use a simplified form of (1). General inspection of (1) indicates that if the round-trip time and/or the packet loss rate increase then the throughput reduces as terms containing these parameters exist in the denominator.

In our variant to standard TFRC, the packet size, s, in the TFRC equation (1) was dynamically altered according to EvalVid-created trace file sizes. This variant makes for more responsive control rather than the mean packet length employed in the reference TFRC formulation (Handley et al, 2003). TFRC was originally

Table 2. Simulated WiMAX traffic characteristics

SS-UL	*Service type*	*Traffic type*	*Protocol*	*Packet Size (B)*
1	rtPS	VBR (video)	Multiple TFRC	Variable
		CBR	UDP	1000
	nrtPS	FTP	TCP	
2	rtPS	CBR	UDP	1000
	nrtPS	FTP	TCP	
3	rtPS	CBR	UDP	1000

	nrtPS	FTP	TCP	
SS-DL				
1,2	rtPS	CBR	UDP	1000
3	nrtPS	FTP	TCP	

intended for video-on-demand applications, when it is feasible to calculate the mean packet length from the stored video. Setting a mean packet length is inappropriate for interactive multimedia applications. The underlying TFRC transport protocol was set to UDP, as is normal for real-time applications. Though (1) appears to represent a considerably computational task that could impede real-time performance, it is possible to extract a term parameterized by p, the packet loss rate. Therefore, a look-up table indexed by p represents a practical way to speed up calculations.

2.3 Channel model

A Gilbert-Elliott two-state, discrete-time, ergodic Markov chain (Haßlinger and Hohlfeld, 2008) modeled the wireless-channel error characteristics at the ns-2 physical layer. The result of applying this model is that burst errors typical of known wireless channel conditions appear. 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 modeled 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.

2.4 Core network traffic characteristics

In Figure 1, all links except a bottleneck link within the core network are set to 100 Mbps to easily accommodate the traffic flows entering and leaving the network. The link delays are minimal (2 ms) to avoid confusing propagation delay with re-ordering delay in the results. A bottleneck link with capacity set to 5 Mbps is set up between the two routers. The buffer size in each router was set to 50 packets. This arrangement is not meant to physically correspond to a network layout but to represent the type of bottleneck that commonly lies at the network edge.

Node A sources to node B a CBR stream 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 a CBR stream at 1.5 Mbps with packet size 1 kB (see Table 2 downlink). Other SS sources apart from the video connections do not pass over the core network shown but are assumed to be routed elsewhere after passing the WiMAX BS. Node C in Figure 1 is the sink for the TFRC multiple connections.

2.5 Management of connections

To systematically test the effect of multiple TFRC connections the number of TFRC connections was incrementally stepped up in successive experiments. In MULTTFRC (Chen and Zakhor, 2006), the number of connections is changed over time according to the average round-trip time of all the connections, but this hides the interpretability of results. It is also unclear from Chen and Zakhor (2006) how a single video stream would be apportioned between a varying number of connections. In our experiments, a single queue was segmented into GOPs (one GoP = 15 frames). Each connection was statically allocated its GOPs, which are taken in interleaved manner from the video sequence. This assumes that a re-ordering buffer is available at the receiver, the size of which is discussed in Section 3..

3. EVALUATION

Initial investigations considered the WiMAX link alone in Figure 1. Table 3 shows the average data-rate when transmitting the Paris clip over one or more connections, for two different WiMAX frame lengths: 5 ms and 20 ms. Allowable frame lengths are specified in the Standard (IEEE, 2005), ranging from 2.5 to 20 ms. Clearly, TFRC is able to multiplex more data onto a link as the number of connections increases, though observation of a time-wise plot of throughput shows that during transmission TFRC sharply reduces its overall sending rate in response to packet loss. Because the sending period for one connection with the shorter frame duration is more than the display period of the 'Paris' clip, the longer frame length is preferable if only one connection were to be used. However, with more than one connection, throughput and, hence, wireless channel utilization by the congestion-controlled video streams increases significantly. There is a marked difference if the larger frame length is used whether one or four connections. As smaller frame lengths than 20 ms are generally used for WiMAX, this is an important observation. In fact, the UL proportion of the frame length, that is 15 ms, is more than the total 5 ms frame length that appears to be usually implemented.

Table 3. Sending periods and throughputs when streaming from a mobile SS to the WiMAX BS

No. of connections	*SS to BS (s) (frame length 5 ms)*	*Throughput (kbps)*	*SS to BS (s) (frame length 20 ms)*	*Throughput (kbps)*
1-conn	71.4	217	33.5	467
2-conn	35.8	437	20.5	754
3-conn.	23.3	663	17.7	874
4-conn.	17.4	889	14.6	1059

Table 4. Streaming periods, throughputs and mean packet end-to-end delays from mobile SS to node C in Figure 1 (frame length 20 ms)

No. of connections	*Sending Period (s)*	*Throughput (kbps)*	*Mean end-to-end delay (s)*
1-conn	35.2	444	0.035
2-conn	22.4	690	0.036
3-conn.	21.6	716	0.039
4-conn.	15.6	991	0.062

In respect to the longer frame length of 20 ms, an interesting comparison is with the throughput when the core network is included. In Table 4, there is a similar pattern to the throughputs in Table 3 but the rates are reduced to when streaming only over the WiMAX link. We interpret this effect as not being due to TFRC's response to packet loss but being due to its response to the increased round trip time caused by queuing delay in the buffer prior to the bottleneck link in Figure 1. Notice that TFRC uses reliable TCP to return ACKs, which will tend to add to the round-trip time. Recall also that from equation (1) that round-trip-time is one of the parameters determining TFRC's sending rate. This interpretation is confirmed by the increase in per slice/packet end-to-end delay as more connections are added. In effect, the packets from other connections intervene in the router buffers causing an increase in latency. However, even though the delay is larger for four connections the mean is still less than 100 ms for this scenario.

More significant than end-to-end delay for reconstruction of the video stream is GOP arrival ordering, as this ordering has the potential to introduce interruptions to the display. GOP arrival ordering for four connections is shown in Figure 2. Firstly, a few points about this Figure are explained. Notice that the first H.264/AVC GOP contains parameters that are fixed throughout the sequence (Wiegand et al, 2003). Therefore, this GOP is transported more quickly. Secondly, to avoid a sudden injection of traffic into the network, connection starting times were offset by 0.5 s. In respect to the general findings, a noticeable feature of this Figure is the lengthier start-up periods in sending initial GOPs on each of the connections. We attribute this to the loss of packets at an early stage, which causes TFRC to sharply reduce its rate in a similar manner to TCP's slow-start mechanism. This does mean that about 6 s of frames (amounting to 90 frames) should be stored in the reordering buffer, to avoid the possibility of subsequent underflow in the decoder's playout buffer. As the destination is on the fixed network the reorder buffer is not expected to be a drain on energy resources, as it might be on an SS. Of course, data is not physically reordered in the buffer but accessed through pointers. 6 s is longer than a typical start-up time of around 2 s but is not too large to be objectionable to the user.

Returning to the effect of frame length, video quality (PSNR) and mean packet end-to-end delay were found for a range of WiMAX frame lengths. However, the standard deviation (stdv) over the runs is relatively large (but similar to those reported in Chen and Zachor (2006)). This is explained by the strong effect resulting from the position of error bursts. From Table 5, video quality is generally 'good', as there is an approximate equivalence of PSNR's over 31 dB and above to the ITU's subjective scale. Again the larger TDD frame size results in better and surprisingly in this instance improves in the mean with an increasing number of connections. However, we take this to signify that using four connections produces equivalent video quality at the destination to using one connection, provided the larger frame size is employed. A 5 ms frame size consistently reduces the quality by one or two dB, which on a logarithmic scale is significant.

Examination of the total packet losses (congestion and channel loss), Figure 3, shows that losses also are generally higher for a 5 ms TDD frame length than a 20 ms frame length. However, between the connections, it is *not* the case that mean PSNR is a direct reflection of mean packet loss. As might be expected, employing four connections leads to an increase in congestion loss and also channel loss (because error bursts affect

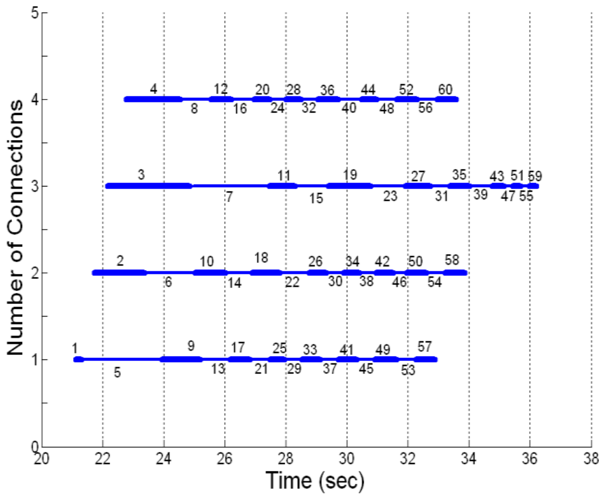

Figure 2. Example GoP arrival sequence at the receiver (node C in Figure 1) showing the start and end times of each GoP

Table 5. Mean PSNR for a range of frame lengths when streaming from a mobile SS to node C in Figure 1

Frame length:	*5ms*		*8ms*		*10ms*		*12.5ms*		*20ms*	
	PSNR (dB)									
Connections	*Mean*	*stdv*	*Mean*	*stdv*	*Mean*	*stdv*	*Mean*	*stdv*	*Mean*	*stdv*
1-conn	29.95	2.90	32.85	3.32	32.44	3.45	33.22	3.45	31.84	3.78
2-conn	28.88	3.07	31.28	3.81	32.17	3.38	33.07	3.47	32.34	3.49
3-conn	29.54	3.25	31.07	3.04	31.83	3.14	30.79	3.25	33.15	3.68
4-conn	28.12	3.11	31.92	3.51	31.20	3.45	33.31	3.80	33.34	3.74

more than one connection). Examining the relative breakdown between frame types, shows that anchor frames (I-frames) and reference frames (I- and P-frames) are evenly affected whatever the number of connections. Therefore, we conclude that the differences in the mean PSNRs are explained by the relatively low number of packet losses when using congestion control and possibly the volatility in the pattern of packet losses when burst errors occur.

From Figure 3, packet loss is particularly high for three connections. The reasons for this anomaly in this scenario are unclear. More generally, the advantages of using four connections in terms of improved wireless utilization and video quality equivalent to one connection are offset by the increased mean end-to-end packet delay, Table 6. However, as remarked earlier, the mean is still below 100 ms in this scenario.

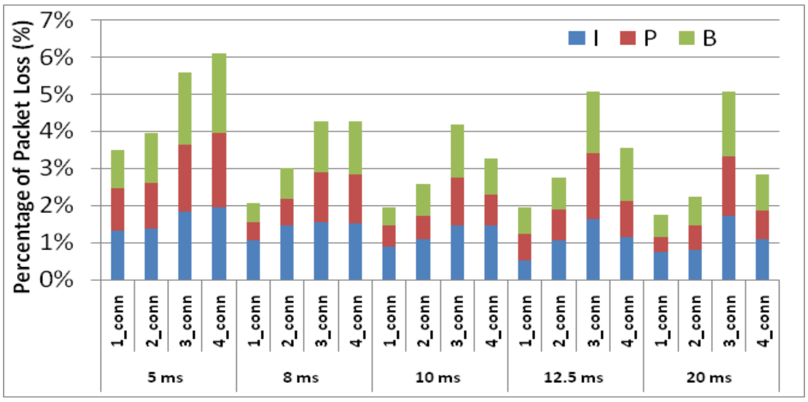

Figure 3. Mean percentage packet loss broken down according to frame type, by connection and TDD frame length, I = Intra-coded frame, P = Predictively-coded frame, B = Bi-predictively coded frame.

Table 6. Mean packet end-to-end delay according to frame length

Frame length:	*5 ms*	*8ms*	*10ms*	*12.5ms*	*20ms*
Connections	*Mean end-to-end packet delay (s)*				
1-conn	0.0195	0.017	0.017	0.020	0.035
2-conn	0.0446	0.029	0.029	0.027	0.036
3-conn	0.0725	0.024	0.028	0.023	0.039
4-conn	0.0982	0.069	0.067	0.073	0.062

4. CONCLUSION

Multi-connection congestion control adapts existing congestion controllers to all-IP networks that include a broadband wireless access link. In effect, they allow the congestion controller to accommodate wireless channel losses but still respond to congestion with the network edge and possibly the core. This in turn leads to improved network utilization, whereas previous observers have noticed a marked drop in utilization if congestion controllers are employed. However, for any wireless technology there still remain issues about how many connections should be used if the disadvantages of multi-connections are to be avoided. This study has found that though there is a small percentage increase in packet loss with four connections over just one, video quality remains equivalent because of the differential effect of packet loss patterns when burst errors are present. There was also a small (in practical terms) increase in packet end-to-end delay. An important observation is that a longer WiMAX TDD frame size is favorable to video transport, though this may not be apparent unless tests are conducted across the whole of a network path and not just the wireless link. An advantage of the multi-connection method of congestion control is the reduction in state when it comes to handoff in a cellular WiMAX, which is important for a delay-intolerant application. Further investigation will examine this issue. Another advantage of the multi-connection method is that a portion of the additional throughput that results is available for error protection, either application layer FEC, or more promising, in terms of compatibility with existing physical layer FEC, the use of source-coded error resilience.

REFERENCES

Chen, M. and A. Zakhor, 2006, Multiple TFRC Connection Based Rate Control for Wireless Networks. *In IEEE Trans. Multimedia*, Vol. 8, No. 5, pp. 1045-1062.

Ekstrom, K. et al, 2006. Technical Solutions for the 3G Long-Term Evolution. *IEEE Commun. Mag*., Vol. 44, No. 3, pp. 38-45.

Haßlinger, G. and Hohlfeld, O., 2008. The Gilbert-Elliott Model for Packet Loss in Real Time Services on the Internet. *Proc. of 14th GI/ITG Conf. on Measurement, Modelling, and Eval. of Comp. and Commun. Systs.*, pp. 269-283.

Handley, J. et al, 2003. TCP-Friendly Rate Control (TFRC): Protocol Specification. *IETF*, RFC 3448.

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

Hsiao, P.-H., Kung, H. T., and Tan, K.-S., 2003. Streaming Video over TCP with Receiver-based Delay Control. *IEICE Trans. on Communications*, Vol. 86, No. 2, pp. 572-584.

Klaue, J. et al, 2003. EvalVid - A Framework for Video Transmission and Quality Evaluation. *Proc. Int. Conf. on Modeling Techniques and Tools for Computer Performance*, pp. 255-272.

Lin, Y.-B. and Pang, A.-C., 2005. *Wireless and Mobile all-IP networks*, Wiley and Sons, Indianapolis, IN.

Shen et al, 2009. Receiver Playout Buffer Requirement for TCP Video Streaming in the presence of Burst Packet Drops, *Proc. of London Communications Symposium*.

Tappayuthpijam, K. et al, 2009. Adaptive Video Streaming over a Mobile Network with TCP-Friendly Rate Control. *Proc. of Int. Conf. on Wireless Communs. and Mobile Computing*, pp. 1325-1329.

Wiegand, T., et al, 2003. Overview of the H.264/AVC Video Coding Standard. In IEEE Trans. Circuits Syst. Video Technol., Vol. 13, No. 7, pp. 560-576.