Solution for Application and Transport Layer Inconsistency during Adaptive Multimedia Streaming

Edward Casey
Performance Engineering Laboratory
School of Electronic Engineering
Dublin City University
Dublin 9, Ireland
Email: caseye@eeng.dcu.ie

Gabriel-Miro Muntean
Performance Engineering Laboratory
School of Electronic Engineering
Dublin City University
Dublin 9, Ireland
Email: munteang@eeng.dcu.ie

***Abstract*— Universal Multimedia Access has become the driving force behind a significant amount of research. However, current research focuses on application layer adaptation of multimedia content to suit the diverse terminal, network and user requirements. This research has neglected the interaction between the application layer adaptation solutions and transport layer rate / congestion control mechanisms. Due to the lack of communication between these layers during adaptive multimedia streaming, an inconsistency occurs. After the application layer has adapted each of the media streams to suit their specific requirements, the transport layer mechanisms treat all of the media streams equally, causing application-level unfairness. This unfairness mostly affects user perceived quality when streaming high bitrate multimedia content. Therefore, there is a need to allow adaptive streaming applications that support Universal Multimedia Access, to tune the aggressiveness of transport layer mechanisms, in order to create application layer Quality of Experience fairness between competing media streams.**

I. Introduction

Today's multimedia devices are becoming increasingly sophisticated. Devices such as the iPod, XBOX 360 and Apple TV are revolutionizing the way multimedia is experienced. Users now expect larger screens, more storage space, faster response times, higher speed wireless connectivity and most importantly, ease of use. It is difficult to distribute media efficiently to devices with such a diverse range of requirements. Traditional distribution techniques followed a Write Once Read Many (WORM) philosophy. This approach is suited perfectly to the analog broadcast content distribution model. However, the future model of choice for these multimedia services will be over broadband IP based networks, often heterogeneous (wired and wireless), where differentiated treatment is required and the WORM approach will not suffice.

The MPEG-21 framework [1] is expected to be a key component for providing universal access of multimedia resources, to a diverse range of devices over heterogeneous wireless networks. It provides a set of tools that allows the characterization of the environment in which multimedia services are offered. The consequent description takes into account terminal capabilities, network characteristics and user characteristics. Although these descriptions are a step in the right direction to achieving MPEG-21's goal of Universal Multimedia Access (UMA), a number of factors have been overlooked. Streaming processes that adapt media using this framework do so at the application layer based on terminal capabilities, user characteristics and network characteristics. This information is gathered horizontally from devices and vertically from other layers in the OSI stack. However once

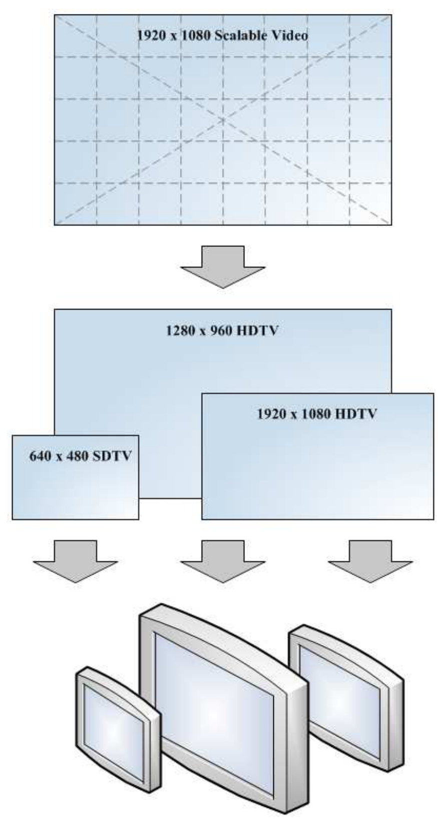

Fig. 1. The MPEG-21 based adaptation process

the adaptation takes place, the streaming process generally does not inform the lower layers of the changes that have taken place. In other words, the application layer has adapted the media to suit the environmental characteristics but lower layers of the stack still treat the media as if no adaptation has taken place. This adaptation must be accounted for by the transport layer rate estimation and/or congestion control mechanism otherwise a fairness anomaly occurs and impacts negatively on the Quality of Experience (QoE) of the end user.

This paper focuses on Part 7 [2] of the MPEG-21 Framework which is referred to as Digital Item Adaptation (DIA). It outlines the fairness issues associated with this multimedia adaptation, caused by a lack of interaction between the application layer adaptation process and lower network layers. The paper also describes a solution to these problems in the form of a TCP compatible Greediness Control Algorithm (GCA) that tunes the greediness of the transport layer congestion control mechanism to create application layer fairness in heterogeneous networks. GCA introduces two parameters that complement the MPEG-21 standard and allow the streaming applications to tune the aggressiveness of their rate estimation. As a result, GCA enables quality-driven, cross-protocol fairness to the media streaming process.

The paper is structured as follows. Section II gives an overview of MPEG-21 Digital Item Adaptation and presents material related to wireless multimedia streaming pertinent to the proposed solution. In Section III, a detailed description of the problem including simulation-based testing results is outlined. Section IV presents GCA as the solution to the fairness problem, simulation results and discussion. The paper is concluded in Section V.

II. Related Work

A. Wireless Multimedia Streaming

The majority of streaming solutions employ transport layer rate adaptation schemes based on congestion avoidance mechanisms. The end-to-end approach of these schemes provides low complexity, flexibility and requires no modification to existing network infrastructure. These mechanisms lie at the upper level of the transport layer and provide application layer adaptation engines with the relevant information to adapt the multimedia stream to best suit the available network conditions. TCP-Friendly Rate Control (TFRC) [3] is one such solution that determines an appropriate transmission rate as a function of loss events and round-trip time. However, TFRC is not suitable for streaming media in wireless multimedia networks. Similarly to TCP, TFRC assumes that every packet loss is congestion induced (not always the case in wireless networks) and reacts by lowering the sending rate. A number of derivatives of TFRC (TFRC Wireless [4] and MULTTFRC [5]) have been proposed that address this issue. Recently the Datagram Congestion Control Protocol (DCCP) [6] has been proposed as a complete unreliable transport layer protocol incorporating TFRC based end-to-end congestion control.

Alternatively, the Video Transport Protocol (VTP) proposed in [7] achieves robust wireless performance using a combination of rate estimation and loss discrimination techniques. The Rate Adaptation Protocol (RAP), proposed in [8], is a source based TCP friendly Additive Increase Multiplicative Decrease (AIMD) rate adaptation scheme. Enhanced Loss Delay Adjustment (LDA+) [9] adapts the transmission behaviour of UDP based multimedia streams in accordance with the current network congestion state. However as mentioned above these solutions are based on congestion avoidance mechanisms that try to obtain as much bandwidth as possible while minimizing loss. Usually higher throughput translates into higher quality, however, when media is adapted by application layer UMA engines, greediness of the rate adaptation becomes a serious fairness issue. The authors of [10] have proposed the use of self limiting sources to control the greediness of multimedia traffic. However this also fails to address the greediness issue, as it does not provide protection from greedy background traffic flows.

The original standards from the IEEE 802.11 family [11] [12] [13] (the most popular in WLANs) were designed for best effort services and as a result, lacked support for real-time services. The recently ratified IEEE 802.11e supplement [14] addresses the shortcomings of the original specification. It provides the required service differentiation by associating a priority level with each packet. The higher priority packets then receive preferential access to the wireless medium. However, there is limited service differentiation for traffic types within the same traffic category (e.g. video) so other solutions must be sought for this. The major problems with these lower layer solutions are that they do not provide sufficient service differentiation and/or require modification to the network infrastructure.

B. MPEG-21 DIA Overview

Universal Multimedia Access (UMA) has become the driving concept behind a significant amount of research and standardization [15]. UMA refers to the ability to seamlessly access a rich set of multimedia content on a diverse range of devices over heterogeneous networks. To date, the majority of UMA solutions have concentrated on the constraints imposed by terminals and networks along the multimedia delivery chain; the users who consume the content are rarely considered [16].

The MPEG-21 framework is a key element in this push towards UMA. It provides the resources for standardizing descriptors for multimedia content access and allowing standards-compatible technologies to be used for adapting multimedia content. Its goal is to support users during the exchange, access, consumption, trade, or other manipulation of so-called digital items in an efficient, transparent, and interoperable way. Digital Items (DI) are MPEG-21's fundamental unit for distribution and transaction. Part 7 of the MPEG-21 framework, named Digital Item Adaptation (DIA) (see Figure 2), provides a set of descriptors and tools that aid the adaptation of DI's. DIA includes Usage Environment Descriptors (UED) that specify how to describe the devices in terms of their physical characteristics, codec properties and

I/O capabilities. UED also defines the description formats for the networks through which the DI's are accessed. It also provides means to describe user characteristics and preferences. DIA also specifies Bitstream Syntax Description (BSD) for describing bitstream's high-level structure which endeavors to allow coding-format-independent bitstream adaptations. These descriptors allow media to be efficiently adapted to suit the requirements of the multimedia streaming chain.

It is very important to note that MPEG-21 DIA standard only provides tools for guaranteeing inter-operability for media adaptation purposes. It specifies the interface to adaptation engines and leaves implementation to developers.

III. Problem Statement

A. Overview

Consider a typical residential IEEE 802.11g wireless LAN (WLAN) with a number of video capable devices attached. Access to the wireless network is shared equally among these devices, resulting in them competing and receiving an equal share of the available bandwidth. The streaming solution deployed in the wireless network optimizes video at the application layer using the MPEG-21 DIA specification. Adaptation takes place at the application layer by utilizing the UEDs specified in the MPEG-21 framework. These descriptions contain information that characterizes the current terminal, network and user conditions. This allows the streaming solution to make an informed decision on which adaptation strategy to apply. Although the adaptation has taken place, the lower layers of the streaming solution continue to treat the media as if it was un-adapted. The transport layer rate estimation / congestion control mechanism still assumes that all devices have equal bandwidth requirements. This results in greedy devices unfairly consuming excessive bandwidth and inefficient distribution of available resources at the application layer.

For example, consider the situation where three wireless clients, a 32" HDTV, 20" HDTV and 12" laptop compete for resources on a WLAN. The 32" HDTV and 20" HDTV receive different H.264 video streams, with properties presented in Table I, from a media server adapted to their specific characteristic requirements and the 12" laptop is downloading a

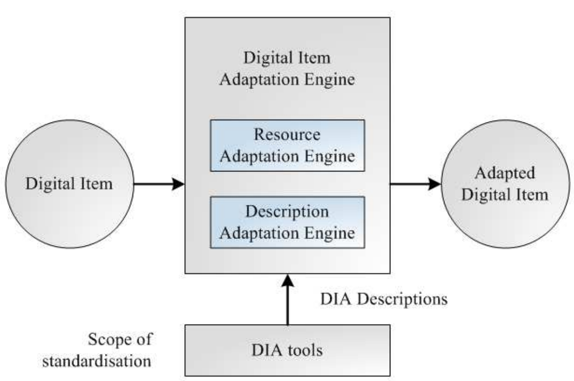

Fig. 2. MPEG-21 Digital Item Adaptation

large file from a Network Attached Storage (NAS) device. If conventional transport layer rate control schemes, such as the ones outlined in section II-A, were deployed in this scenario it would result in all clients receiving an equal share of available bandwidth. Assuming there is only 20 Mbps of available bandwidth this may result in clients 2 and 3 receiving their required bandwidth share while client 1 receives only 70% of what it actually requires. Although adaptation has taken place at the application layer the transport layer rate control mechanism is not taking this adaptation into account. From a transport layer perspective, bandwidth is distributed fairly, however, from an application layers point of view this allocation impacts severely on the user quality, as experienced on the 32" HDTV.

TABLE I

Device characteristic video requirements

	Client 1	Client 2	Client 3
Device type	32' HDTV	20' HDTV	12" Laptop
Format	H.264	H.264	H.264
Resolution (pixels)	1920x1080	1280x720	640x480
Average Bit Rate (Mbps)	9	6	3
Max Bit Rate (Mbps)	20	14	8

The problem occurs because application layer MPEG-21 adaptation engine adapts video content based on the UEDs obtained from various layers. However, once the adaptation has taken place, it does not inform those layers of the action it has taken. This causes the transport layer to treat all media streams equally even though media has unequal requirements.

B. Simulation

In order to illustrate the problem statement outlined above, a number of simulations were carried out. These simulations were designed to illustrate the effect that transport layer fairness provided by conventional rate control mechanisms has on adapted application layer video quality fairness. Simulations were carried out using Network Simulator 2 (NS-2) [17]. The simulations considered a varying number of clients connected to a centralized media server via the a IEEE 802.11g WLAN (see Figure 3). Clients are assumed to be devices with the characteristics and requirements detailed in Table I. A background traffic source (BG) is also included to suitably load the network by simulating the transfer of a large file in parallel with multimedia streaming. Simulation results demonstrate the problem using a conventional TFRC based streaming solution with media adapted at the application layer using an MPEG-21 like framework to suit the relevant terminal characteristics.

Clients 1, 2, 3 and the background traffic source join the WLAN in 100.0s increments beginning at $t = 1s$ and subsequently leave the network in 100.0s increments beginning at $t = 700s$. The simulation period is $1000s$. The throughput analysis of this simulation is illustrated in Figure 4 and a comprehensive summary of results can be found in Table II. The simulation begins with the addition of the background traffic source at $t = 1s$. This client has requested to download a large file from a wired server via the WLAN. During the period $1s < t < 100s$ this client has exclusive access to the wireless medium and acquires 10.1 Mbps of available bandwidth. At $t = 100s$ the 12" Laptop computer joins the WLAN. As outlined in Table I, this device requires an application layer adapted video stream 3 Mbps average bit rate to achieve optimal quality and efficient use of wireless resources. During the period $100s < t < 200s$ this client receives its required throughput to provide this optimal quality.

At $t = 200s$ the 20" HDTV with an average throughput requirement of 6 Mbps. The 20" HDTV receives approximately 5.7Mbps throughput from the wireless medium at the expense of the background traffic source while the 12" laptop remains largely unaffected. From a transport layer perspective this bandwidth allocation is fair. However the application layer metrics suggest unfair distribution of estimated user perceived quality. The 12" laptop is achieving high PSNR scores of 95dB while the 20" HDTV is only achieving 85 dB. As a result the bandwidth distribution at the transport layer is fair but the quality distribution at the application layer is unfair.

The effect of this anomaly becomes more prominent when the 32" HDTV is added at $t = 300s$. During the period $300s < t < 700s$ all four clients compete for access to the WLAN. From a transport layer perspective all clients receive approximately equal access to the WLAN resulting in fair throughput distribution. The two HDTVs receive approximately 4.5 Mbps video stream each while the 12" Laptop still maintains its required 3 Mbps. The background traffic source also receives less throughput, but due to its best effort non qualitative nature, this reduction is acceptable. The application layer quality metrics for this period clearly illustrate inequality experienced by the users. The 12" Laptop continues to receive multimedia at near perfect quality of 89dB, while the 20" HDTV and 32" HDTV receive media at a quality of only 36dB and 66dB respectively.

Although the transport layer network related metrics show equality between clients, the application layer quality metrics show huge differences which directly impact the overall user QoE. These results imply that in order to observe the best quality using this streaming solution, it is better to have small screen devices.

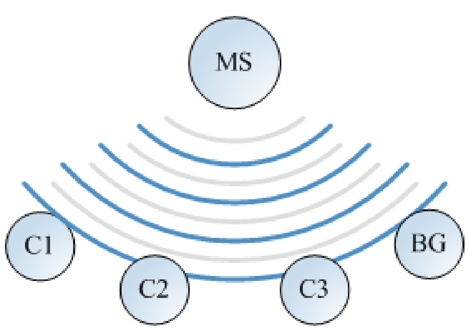

Fig. 3. Simulation topology

IV. Proposed Solution

A. Overview

In order to eradicate the fairness anomaly described in section III it is necessary to introduce a mechanism that allows application layer to inform the transport layer rate estimation algorithm that adaptation has taken place. The proposed solution utilizes parameterized media information communicated from the application layer to tune the aggressiveness of the transport layer rate estimation equation. This adapts the transport layer to suit the previously adapted application layer data and thus solve the fairness anomaly.

From a transport layer perspective, the proposed Greediness Control Algorithm (GCA) is a TCP compatible rate estimation / congestion control mechanism for unicast flows. GCA extends the TFRC protocol. It inherits many of TFRC's characteristics which make it suitable for multimedia streaming applications. GCA determines its transmission rate based on a simplified version of the TCP Reno throughput equation. It also resembles the TFRC protocol mechanism as it involves a sender transmitting data packets to the receiver, which periodically returns feedback to the sender. It determines the sending rate using Equation 1, where sending rate is a function of Round Trip Time (RTT), loss event rate (p) and packet size (s). α and β ($\delta = 1/\beta$) are specially added parameters that tune the aggressiveness of the rate estimation. The aggressiveness parameters are derived from the stochastic TCP model presented in [18] and the methodology used in [19].

$$X = \frac{s}{RTT(\sqrt{\frac{2p(\delta-1)}{\alpha(\delta+1)}} + 12 \times p\sqrt{\frac{p(\delta-1)(\delta+1)}{2\alpha\delta^2}}\,(1+32p^2))} \tag{1}$$

Using this equation and by varying α and β, it is possible to configure GCA flows so that they are either more or less aggressive, thus adapting the transport layer rate estimation to suit the adapted application layer multimedia process.

B. Simulation

The proposed GCA based adaptive multimedia streaming solution was evaluated using the same scenario and conditions that were used to evaluate the TFRC based streaming solution in section III-B. Results are illustrated in Figure 5 and a detailed summary is presented in Table III. Clients were assigned priorities to account for the device characteristics outlined in Table I.

The most noticeable difference between the results is the level of service differentiation achieved between clients. Each of the clients receive transport layer equality in terms of delay and jitter, each experiencing an average of approximately 22 ms delay and 0.6 ms jitter. However, the service differentiation introduced by the α and β parameters, as expected, has resulted in a throughput inequality.

At $t = 1s$ background TCP traffic source begins transmission. During the interval $1s < t < 100s$ this traffic obtains maximum throughput of $\approx$ 10Mbps. At $t = 100s$, the 12"

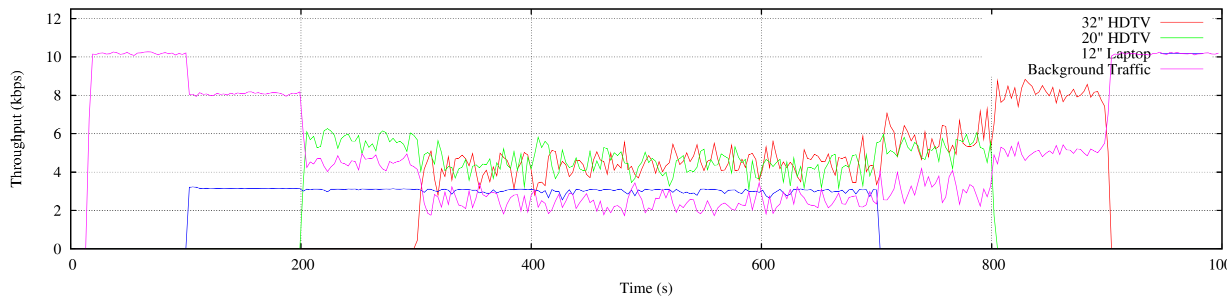

Fig. 4. TFRC based streaming solution throughput

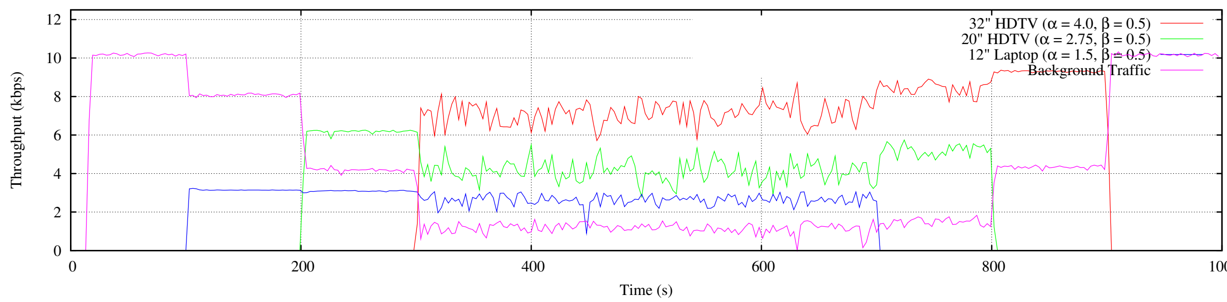

Fig. 5. GCA based streaming solution throughput

TABLE II
TFRC BASED STREAMING SOLUTION SIMULATION RESULTS SUMMARY FOR $300.0s < t < 700.0s$

	Client 1	Client 2	Client 3	BG Traffic
Throughput (Mbps)	4.56	4.23	2.91	2.53
Delay (ms)	23.46	22.78	23.56	22.95
Loss (%)	1.65	1.71	1.69	1.56
PSNR (dB)	36.64	66.24	89.25	-

TABLE III
GCA BASED STREAMING SOLUTION RESULTS SUMMARY FOR $300.0s < t < 700.0s$

	Client 1	Client 2	Client 3	BG Traffic
Throughput (Mbps)	7.09	4.19	2.56	1.23
Delay (ms)	23.32	22.96	22.45	23.15
Loss (%)	1.68	1.65	1.75	1.62
PSNR (dB)	74.23	61.52	77.32	-

laptop joins the wireless network and requests an adapted multimedia video stream, that suits its characteristic requirements, from the media server. This new source competes with the background traffic for the wireless medium. As a result the background traffic source receives 20% less throughput while the new multimedia client receives 100% in terms of its average bit rate requirement. These results compare favorably with results obtained in section III-B for the same period. The positive effect of the GCA can now be clearly seen from $t > 200s$. When the 20” HDTV joins the network, it receives higher throughput than the 12” Laptop and the background traffic, where as before it received approximately the same average throughput as the background and greater throughput than the 12” Laptop PC. This interaction illustrates the positive effect that GCA has on prioritized traffic. The higher priority HDTV video stream acquires more bandwidth share at the expense of the background traffic source, which results in a clear increase in estimated user QoE for this client.

For the period $300s < t < 700s$, a fourth client (32” HDTV) joins the network. As per Table I, this client has 9 Mbps average throughput requirement for achieving maximum QoE. This addition has a significant impact on the throughput distribution of existing clients and acquires approximately 7 Mbps bandwidth share from the network. As a result of this addition the 20” HDTV has 33% lower throughput, the 12” laptop multimedia stream has about 30% lower throughput while the background traffic source experiences a 75% drop in throughput. Greater throughput fluctuation is experienced during this period due to the increased competition for resources.

From an application layer perspective, each of the multimedia streams is receiving approximately 80% of their characteristics requirements, which inevitability leads to higher overall user QoE. PSNR analysis of the period $300s < t < 700s$ shows very high levels of quality for all devices of approximately 70 dB. Although no application layer statistics are obtained for the background traffic source it is expected that due to the best effort nature of the service there is very little impact on its quality. These results show important increases in the overall user QoE obtained by tailoring the transport layer rate control mechanism to suit the application layer optimization of the media being carried. It can be concluded that GCA-based adaptation brings bandwidth efficiency, maintains TCP compatibility and determines an important overall increase in end-user perceived quality.

V. Conclusion

This paper presents the Greediness Control Algorithm (GCA) for alleviating the fairness inconsistency introduced between application layer adaptation engines and transport layer rate control mechanisms by the MPEG-21 Framework Part 7: Digital Item Adaptation. The paper highlights the need for communication between these entities in order to allow application layer adaptation mechanisms to manage the greediness of transport layer rate control processes. GCA solves this problem by introducing parameters that allow the application layer to control the greediness and generosity of the transport layer rate control processes. Simulation results show that the proposed solution achieves its desired goal while maintaining transport layer stability.

Future development will focus on further refinements of the proposed GCA together with a solution for mapping the two greediness control parameters to actual device requirements. Further testing will also be carried out with a more diverse range of background traffic types. Subjective assessment of end-user perceived quality is also envisaged.

Acknowledgment

The support of the Irish Research Council for Science Engineering and Technology (IRCSET) Embark Initiative and Enterprise Ireland Commercialization Fund is gratefully acknowledged.

References

[1] *Information Technology Multimedia framework (MPEG-21) Part 1: Vision, technologies and strategy*, International Organization for Standardization ISO/IEC TR 21000-1:2001, 2004.

[2] *Information Technology Multimedia framework (MPEG-21) Part 7: Digital Item Adaptation*, International Organization for Standardization ISO/IEC 21000-7:2004, 2004.

[3] M. Handley, S. Floyd, J. Padhye, and J. Midmer, *TCP Friendly Rate Control (TFRC): Protocol Specification*, IETF RFC 3448, Std., 2003.

[4] S. Cen, P. Cosman, and G. Volelker, “End-to-end differentiation of congestion and wireless losses,” *IEEE/ACM Transactions on Networking*, vol. 11, no. 5, pp. 703–717, October 2003.

[5] M. Chen and A. Zakhor, “Rate control for streaming video over wireless,” in *Proceedings of the IEEE Infocom*, Hong Kong, China, March 2004.

[6] E. Kohler, M. Handley, and S. Floyd, *Datagram Congestion Control Protocol (DCCP)*, IETF RFC 4340, Std., 2006.

[7] G. Yang, M. Gerla, and M. Y. Sanadidi, “Adaptive video streaming in the presence of wireless errors,” in *Proceedings of the IPIF/IEEE MMNS*, San Dieigo, USA, Octomber 2004.

[8] R. Rejaie, M. Handley, and D. Estrin, “RAP: an end-to-end rate-based congestion control mechanism for realtime streams in the Internet,” *Proceedings of IEEE INFOCOM*, vol. 3, no. 3, pp. 1337–1345, 1999.

[9] D. Sisalem and A. Wolisz, “LDA+: A TCP friendly adaptation scheme for multimedia communication,” in *Proceedings of the International Conference on Multimedia and Expo*, New York, USA, July 2000, pp. 1619–1622.

[10] T. Phelan, “TFRC with self-limiting sources,” 2004. [Online]. Available: http://www.phelan-4.com/dccp/

[11] *IEEE 802.11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification*, IEEE Standard, 2000.

[12] *IEEE 802.11b: Higher-Speed Physical Layer (PHY) Extension in the 2.4GHz Band*, IEEE Standard, 2001.

[13] *IEEE 802.11g: Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band*, IEEE Standard, 2003.

[14] *IEEE 802.11e: Medium Access Control (MAC) Quality of Service Enhancements*, IEEE Standard, 2005.

[15] A. Vetro, C. Chritopolos, and T. Ebrahimi, “Special issue on universal multimedia access,” *IEEE Signal Processing Magazine*, vol. 20, no. 2, March 2003.

[16] C. Timmerer and H. Hellwagner, “Interoperable adaptive multimedia communication,” *IEEE Multimedia*, vol. 12, no. 1, pp. 74–79, March 2005.

[17] Network simulator 2. [Online]. Available: http://www.isi.edu/nsnam/ns/

[18] J. Padhye, V. Firoiu, D. Towsley, and J. Kurose, “Modeling TCP Reno performance: a simple model and its empirical validation,” *IEEE ACM Transactions on Networking*, vol. 8, no. 2, pp. 133–145, 2000.

[19] A. Kuzmanovic and E. W. Knightly, “A performance vs. trust perspective in the design of end-point congestion control protocols,” in *Proceedings of the 12th IEEE International Conference on Network Protocols*, Berlin, Germany, October 2004, pp. 96–107.