Delivering Adaptive Scalable Video over the Wireless Internet

Pavlos Antoniou, Vasos Vassiliou, and Andreas Pitsillides

Networks Research Laboratory, Computer Science Department, University of Cyprus*
{paul.antoniou,vasosv}@cs.ucy.ac.cy,
andreas.pitsillides@ucy.ac.cy

Abstract. A large portion of the emerging and future wireless Internet traffic is foreseen to be consumed by video streams. However, delivering video over wireless networks poses a lot of challenges. Network congestion and wireless channel errors yield tremendous packet loss leading to degraded video quality. One of the most critical issues for video applications is to ensure that the quality of service (QoS) requirement will be maintained at an acceptable level, providing responsiveness to the time-variant network conditions as well as scalability and fairness among concurrent users. In this paper, we study the performance of a novel fuzzy-based adaptive mechanism which takes into account a combination of Network Adaptation Techniques with Content Adaptation Techniques in order to achieve graceful performance degradation when network load increases and network conditions deteriorate. Our performance evaluations indicate that our approach finely adapts the video stream bit rate to the available bandwidth, maintains responsiveness to dynamic changes and achieves scalability and fairness as well as high and stable objective quality of service.

1 Introduction

The overwhelming majority of today's handheld devices like mobile phones, PDAs and laptops are capable of streaming video content. Therefore, video transmission over the Internet is considered to be the prime candidate for being the next killer application.

Needless to say that video communications face a lot of challenges. Compressed video streams (like MPEG) exhibit large variations in their data rates something which makes their management in a packet-based best-effort network like IP extremely difficult. Moreover, the unpredictable nature of various heterogeneous networks within the Internet primarily in terms of bandwidth, latency and loss variation make the transmission of the compressed video streams an even more challenging task. The problem is worsened when we consider mobile users connecting with wireless terminals due to the erroneous and time-variant conditions of the wireless environment.

Under these circumstances, video transmission applications need to be responsive to dynamic changes and different demands. Thus, they need to implement highly scalable and adaptive techniques in terms of content encoding and transmission rates in order to cope with the increased network heterogeneity and complexity. Towards this direction,

* This work was partly funded by the UCY ADAVIDEO project.

the combination of Content Adaptation Techniques (CATs) with Network Adaptation Techniques (NATs) is considered to be an imperative need. CATs deal with adaptation of the video content to the desirable transmission rate using primarily scalable video approaches. Scalable video approaches can solve the variable bandwidth problem only if the streaming architecture is able to track the available bandwidth and react without latency. Thus, we consider NATs which deal with the end-to-end adaptation of real time video application needs to the network parameters using algorithms which take into account the state and/or load of the network and the type of errors.

In this paper, we present a fuzzy-based approach for the adaptive delivery of video streams under variable connection characteristics, which is targeted for video delivery in wireless and mobile environments. Our approach involves a new feedback mechanism that works in conjunction with a fuzzy decision algorithm. We study the performance of our approach with respect to responsiveness to dynamic changes, graceful co-existence with cross traffic, scalability and fairness among concurrent mobile and wireless users.

The remainder of this paper is organized as follows: Section 2 presents and analyzes the architecture of the adaptive mechanism. Section 3 deals with the evaluation setup and scenarios. Section 4 presents some performance results. Section 5 concludes the paper and discusses future work.

2 Adaptive Video Streaming Components

Our approach consists of two basic components, namely a feedback mechanism and a fuzzy-oriented decision algorithm depicted in Fig. 1. The feedback mechanism combines receiver's critical information on the perceived quality as well as measurements obtained by the core network in order to evaluate the available bandwidth of the network path. The estimated available bandwidth is then fed into the decision algorithm which decides in a fuzzy manner the optimal number of layers that should be sent by adding or dropping layers.

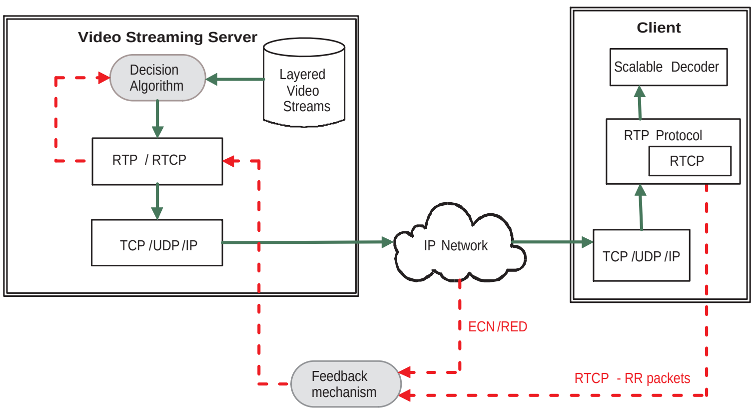


Fig. 1. Adaptive fuzzy-based video streaming architecture.

Fig. 1 illustrates the architecture of a unicast-oriented system using our adaptive fuzzy-based approach. The two outlined components focus on the adaptation of the layered video content to the available network bandwidth. We assume that each video stream is encoded in multiple layers stored at the sender side. The layered video content is transmitted over an RTP/RTCP connection [1]. Dashed arrows track the path of control packets whereas solid arrows track the path of video data packets.

The role of the feedback and adaptation components is to link the quality demand of video-enabled applications to the underlying network. Network adaptation is assisted by a proper content adaptation technique which is carried out by layered video encoding.

2.1 Feedback Mechanism

The feedback mechanism collects QoS information (e.g. loss rate, jitter) from both the core network and the receiver that will be used for the evaluation of the available bandwidth of the path between the sender and a receiver.

Each receiver sends reception statistics using dedicated RTCP packets called Receiver Report (RR) packets which carry reception statistics. Among them, the packet loss fraction within an interval is given by the number of packets expected divided by the number of lost packets during the interval. The loss rate per second (LRPS) can be obtained by dividing the loss fraction by the difference in RRs timestamps. The difference between two successive values of LRPS can be used in order to track the increasing or decreasing trend of packet loss percentage.

Additionally, network elements as, for example, routers within the network path may explicitly notify the sender about the current status of congestion within the core network. These notifications can be efficiently used for the evaluation of the available bandwidth. The Explicit Congestion Notification (ECN) mechanism mentioned in [2] is used for the notification of congestion to the end nodes in order to prevent unnecessary packet drops. ECN option allows active queue management (AQM) mechanisms such as, for example RED [3] or Fuzzy-RED [4] to probabilistically mark packets. The number of marked packets within a given period may provide a meaningful reference about the congestion status. The receiver collects these data and sends them back to the sender using a dedicated field of the RR packets.

2.2 Fuzzy Decision Mechanism

The decision algorithm which is implemented at the sender side, processes the feedback information and decides the optimum number of layers that will be sent using fuzzy logic control. Our fuzzy decision algorithm is based on two linguistic input variables and one output variable. All quantities in our system are considered at the discrete instant kT, with T the decision period.

Our first linguistic input variable involves the LRPS parameter. LRPS(kT) is the loss rate per second at each decision period and LRPS(kT−T) is the loss rate per second with a delay T. The linguistic variable D_{LRPS}(kT) gives the increasing or decreasing trend of the LRPS and can be evaluated by:

$$D_{LRPS}(kT) = LRPS(kT) - LRPS(kT - T) \tag{1}$$

The LRPS parameter is lower and upper bounded by 0 and 1 respectively. Thus, the $D_{LRPS}(kT)$ parameter ranges from -1 to $+1$.

For the second input linguistic variable we use the number of packets that have the ECN bit set within a period. The receiver calculates periodically this number called $N_{ECN}(kT)$ and send it back using a dedicated field of the RR packet. The sender extracts this value and calculates a scaled parameter, $N_{ECNsc}(kT)$, which ranges from -1 to $+1$, and represents the percentage of packets marked within this period. Eq. 2 is used to obtain the scaled parameter $N_{ECNsc}(kT)$:

$$N_{ECNsc}(kT) = \frac{N_{ECN}(kT)}{N_{ps}(kT)}, \tag{2}$$

where $N_{ps}(kT)$ is the number of packets sent within the same period. Therefore, we calculate the parameter $DN_{ECNsc}(kT)$, which gives the increasing or decreasing trend of the number of marked packets. The $DN_{ECNsc}(kT)$ is upper and lower bounded by $+1$ and -1 respectively, and can be evaluated by:

$$DN_{ECNsc}(kT) = N_{ECNsc}(kT) - N_{ECNsc}(kT - T) \tag{3}$$

Our fuzzy system [5] processes the two linguistic input variables based on the predefined if-then rule statements (rule base) shown in Table 1, and derives the linguistic output variable a(kT), which is defined for every possible combination of inputs. The defuzzified crisp values of a(kT) can be used by the decision algorithm for the evaluation of the available bandwidth using the formula:

$$avail\_bw(kT) = a(kT) * avail\_bw(kT - T) \tag{4}$$

The defuzzified output value is selected to range from 0.5 to 1.5. Thus a 'gradual' increase is allowed when there is available bandwidth and reduced congestion, whereas quick action is taken to reduce the rate to half in case of severe congestion.

Table 1. Linguistic Rules[1].

a(kT)		$DN_{ECNsc}(kT)$						
		NVB	NB	NS	Z	PS	PB	PVB
$D_{LRPS}(kT)$	NVB	H	H	B	B	Z	S	VS
	NB	H	VB	Z	Z	Z	S	VS
	NS	B	Z	B	Z	Z	S	VS
	Z	B	Z	Z	B	Z	S	VS
	PS	Z	Z	Z	Z	S	S	VS
	PB	Z	Z	Z	Z	S	S	VS
	PVB	S	S	S	S	VS	VS	VS

[1] Table Content Notations: Negative/Positive Very Big (NVB, PVB), Negative/Positive Big (NB, PB), Negative/Positive Small (NS, PS), Zero (Z), Very Small/Big (VS, VB), Small/Big (S, B), Medium (M), Huge (H).

Our decision algorithm has to decide which layers should be sent according to the available bandwidth, based on a non aggressive layer selection approach. The server hosts an appropriate number of layers which correspond to different transmission rates. To avoid ping-pong effects there should not be a transition to an upper level layer every time the available bandwidth exceeds the threshold of a specific rate that corresponds to a higher layer. Instead, a time hysteresis is introduced in order to avoid frequent transitions from one layer to another. More detailed description can be found in [5].

3 Evaluation Setup and Scenarios

3.1 Topology

Fig. 2 illustrates the dumbbell topology we used for the performance evaluation of our approach. A bottleneck link was simulated using two routers directly connected with a link having variable characteristics. All the other wired links have constant bandwidth (10Mbps) and propagation delay (1ms). A video streaming server is attached to the first router. Mobile clients are wirelessly connected to an access point which is attached to the second router. In order to make our scenarios more realistic we added FTP and web-like cross traffic initiated by the FTP server and the WEB server which are both connected to the first router. Wired clients were used to initiate cross traffic.

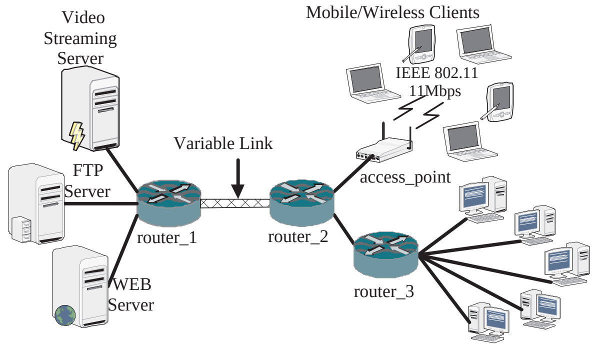


Fig. 2. Evaluation topology.

3.2 Variable Test Parameters

The different parameter values used to characterize the variable link between the routers are shown in Table 2. The bandwidth of the variable link ranges from 64Kbps to 4Mbps, while the propagation delay varies from 10ms to 800ms. The choice of the parameters used in the video quality evaluations is based on the representative characteristics of wired and wireless networks. For example, the link bandwidth can be considered as either the last hop access link bandwidth or the available bandwidth to the user. The values chosen can represent typical wired home access rates (e.g. modem, ISDN, xDSL).

The maximum buffer capacity was set to 50 packets, and RED parameters as shown: $(min_{th}, max_{th}, p_{max}) = (10, 30, 0.1)$. Moreover the interval T between transmissions of RR packets was set to 0.5 seconds. The selection of 0.5 seconds is dictated by the desire to maintain responsiveness to changes in the network state.

3.3 Test Sequences

The video sequence used in this study was the well known real test video named *Foreman*, which is a stream with a fair amount of movement and change of background. The sequence has temporal resolution 30 frames per second, GoP (Group of Pictures) pattern IBBPBBPBBPBB, and spatial resolution 176x144. We encoded this sequence using a publicly available MPEG4 encoder [6] in 8 different bit rates as shown in Table 2. Each encoded video stream corresponds to a separate layer. Since the encoding of the sample video sequences is based on MPEG4, individual frames have variable sizes.

Table 2. Variable Link and Video Parameters.

Video Stream Bit Rate		*Link Bandwidth*		*Propagation Delay*
64 Kbps	384 Kbps	64 Kbps	768Kbps	10 ms
96 Kbps	512 Kbps	128 Kbps	1 Mbps	100 ms
128 Kbps	768 Kbps	256 Kbps	2 Mbps	200 ms
192 Kbps		384 Kbps	4 Mbps	400 ms
256 Kbps		512 Kbps		800 ms

3.4 Data Collection

All the aforementioned experiments were conducted with an open source network simulator tool ns2 [7]. Due to the inadequacy of the existing ns2 modules, we implemented some new software modules [5]. Based on the open source framework called EvalVid [8] we were able to collect all the necessary information needed for the objective video quality evaluation like PSNR values. Video quality is measured by taking the average Peak Signal-to-Noise Ratio (PSNR) over all the decoded frames. Some new functionalities were implemented in ns2 from [9] in order to support EvalVid.

4 Results

In this section we present and investigate the performance of our approach based on the results obtained from the above scenario evaluations. The time varying behavior of the network environment is carried out through cross traffic patterns. Section 4.1 studies the responsiveness of the proposed approach to dynamic changes of the network environment. Section 4.2 investigates the effect of link bandwidth and propagation delay on the received video stream quality in terms of PSNR. Section 4.3 deals with scalability and fairness issues. Finally Section 4.4 focuses on the system capacity with respect to the number of users that can be supported by a video streaming server.

Objective quality metrics like, PSNR, cannot characterize fully the response and the end satisfaction of the viewer. Subjective quality assessment is more a reliable method, as the perceived measure of the quality of a video is done through the human ”grading” of streams which helps collect and utilize the general user view (Mean Opinion Score, MOS). To this end, the relationship between the MOS and the PSNR, based on the same *Foreman* video sequence, in a similar network environment is demonstrated in [10].

4.1 Responsiveness to Dynamic Network Changes

We investigate the ability of the fuzzy rate controller to sense the available bandwidth of a bottleneck link in the presence of various cross traffic patterns, and adapt the transmission rate of a 1Mbps scalable CBR non trace-based video stream. The video stream is transmitted over a bottleneck link having constant bandwidth of 1Mbps to a mobile user. We consider three kinds of traffic patterns, namely, (a) multiple CBR connections which are superimposed progressively, (b) FTP traffic, and (c) Web-like traffic.

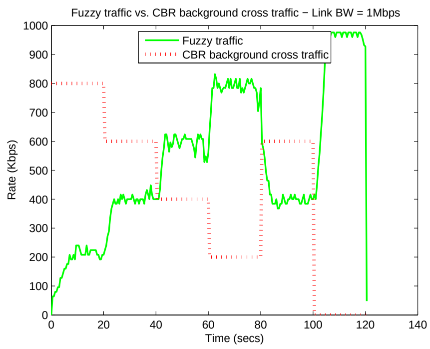
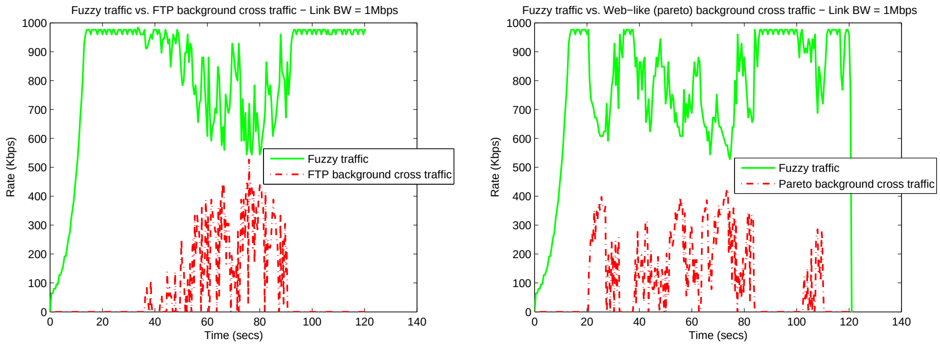


Fig. 3. Instantaneous rate for 1Mbps bottleneck link with CBR cross traffic.

Fig. 3 depicts the instantaneous transmission rate of the layered CBR video stream as the CBR traffic rate changes over the time. The CBR cross traffic rate ranges from 200Kbps to 800Kbps. As can be seen, the video transmission rate driven by the fuzzy rate controller, evolves at a slow and smooth pace in order to respond to the network and quality conditions, but also prevent unnecessarily many fluctuations.

Fig. 4(a) shows the transmission rate of the layered CBR video stream in the presence of FTP traffic. Although the FTP cross traffic is more bursty than CBR shown in Fig. 3, the fuzzy controller senses the available capacity of the bottleneck link and finely adapts the video rate to it. The fuzzy-controlled flow appears to be TCP-friendly against an FTP flow, as it does not aggressively consume the available bandwidth.

Fig. 4(b) depicts the instantaneous transmission rate of the layered CBR video stream in the presence of web-like cross traffic. We simulated web-like traffic using



(a) FTP cross traffic.

(b) Web-like (pareto) cross traffic.

Fig. 4. Instantaneous rate for 1Mbps bottleneck link with FTP/Web-like cross traffic.

a single ON/OFF UDP source, with ON and OFF times drawn from a Pareto distribution. The mean ON time is 350ms, the mean OFF time is 650ms, and during ON time, the UDP sending rate is 400Kbps. The shape parameter of the Pareto distribution is set to 1.11. Even though the web-like traffic is extremely bursty and unpredictable, the fuzzy controlled flow maintains responsiveness during heavy and time-variant loads.

4.2 Effect of Link Bandwidth and Propagation Delay on the QoS

In order to study the effect of link bandwidth and propagation delay on the received QoS, we conducted scenarios involving one wireless mobile user that receives streaming video over the topology shown in Fig. 2 in the absence and presence of cross traffic.

Fig. 5 reveals that in the absence of cross traffic, the PSNR values are increasing at a steady pace (up to 36.5dB) as the link bandwidth increases. PSNR values are significantly lower (less than 20dB) in scenarios where the link bandwidth is equal to the bit rate of the lowest layer (64Kbps), since there is a strong possibility of packet loss. In high bandwidth links (above 512Kbps), the PSNR values are slightly higher for low delay values. On the other hand, in medium bandwidth links (between 256Kbps and 512Kbps), the PSNR values are slightly lower for low delay values. This observation is attributed to the fact that the longer the propagation delay the longer the interval between reception of two successive RR packets. Under these circumstances, the system will experience delayed decision-making that will influence the quality of the video stream. If the link bandwidth is high enough to sustain the video transmission rate, a delayed decision will result to smaller PSNR values because the content adaptation evolves at a slow pace. In the contrary, low delay values will result to higher PSNR since the content adaptation to network parameters evolves at a faster pace. In the case of a low bandwidth link, delayed decisions will benefit the system since the sending rate will be kept in lower levels. This results to higher PSNR values due to the small number of packets lost, since rapid changes in the number of layers sent are avoided.

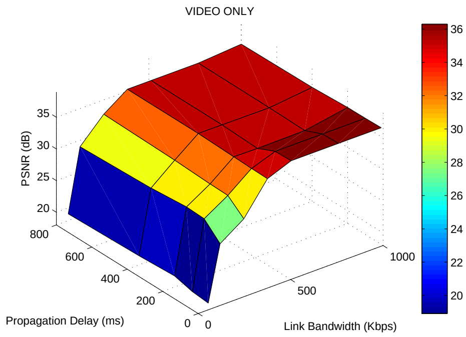
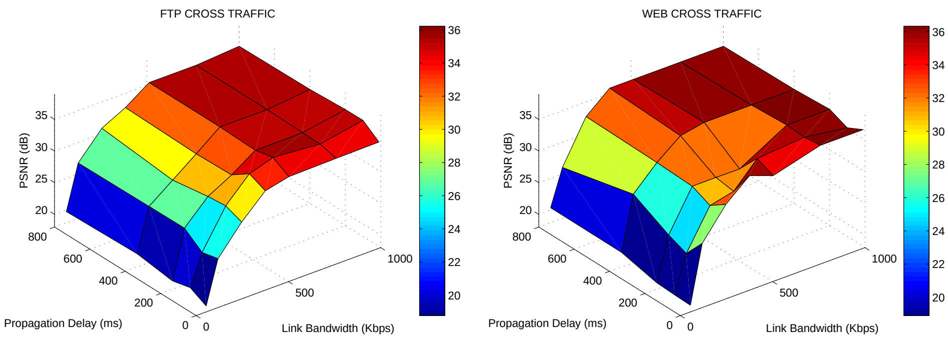


Fig. 5. Mean PSNR vs. Link BW and Prop. Delay.

Fig. 6(a) shows PSNR for scenarios involving FTP cross traffic. We observe a slight decrease in PSNR for scenarios having link bandwidth less or equal to 256Kbps due to the excessive FTP traffic load. As the link bandwidth increases (more than 256Kbps), the quality of a video stream is not severely affected by the FTP traffic since the decision algorithm adjusts the number of layers sent, according to the variable network conditions. Moreover, we perceive a lower objective quality for low propagation delay values, because the FTP rate evolves at a faster and more aggressive pace than in scenarios with longer delay due to the inherent characteristics of the underlying TCP protocol, resulting in high packet drop rates.



(a) FTP cross traffic.

(b) Web cross traffic.

Fig. 6. Mean PSNR vs. Link BW and Prop. Delay, with FTP/Web cross traffic.

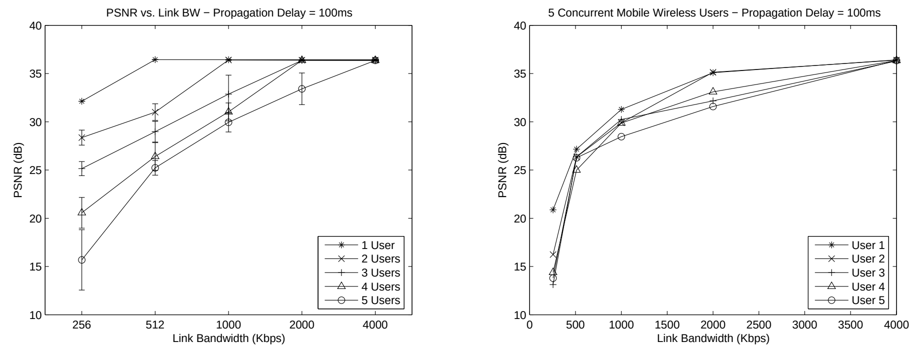
The effect of propagation delay and link bandwidth on the PSNR in the presence of web traffic is presented in Fig. 6(b). Five fixed users (see Fig. 2) are used to simulate web traffic. Each user initiates two sessions and each session consists of 10 web pages.

Session 1 has the following characteristics: exponential inter-page interval (mean = 200ms), Pareto II web page size (mean = 4 objects, shape = 1.5), exponential inter-object interval (mean = 10ms), and Pareto II object size (mean = 4 packets, shape = 1.2). In addition, Session 2 has the following characteristics: exponential inter-page interval (mean = 300ms), constant web page size (1 object), exponential inter-object interval (mean = 10ms), and Pareto II object size (mean = 10 packets, shape = 1.2). As can be seen, the shape of the quality surface obtained from these scenarios is somehow similar to this concerning FTP traffic (Fig. 6(a)). The quality of the received video stream seems to deteriorate more than in FTP for low propagation delay values when link bandwidth ranges from 64Kbps to 768Kbps. As mentioned in the case of FTP traffic, this is justified by the aggressive behavior of the TCP protocol on which the web traffic is based, as well as by the aggressive characteristics (small intervals between web pages and embedded objects) of the web traffic.

The aforementioned scenarios reveal that our approach can finely adapt the video stream bit rate to the available bandwidth. Based on subjective evaluations presented in [10], the *Good* and *Excellent* categories of MOS define the lowest limit for acceptable objective quality, which is 27dB. Thus, our results demonstrate that our system provides high objective quality (above 27dB) both in the absence and in the presence of cross traffic.

4.3 Scalability and Fairness in Multiple Concurrent Mobile/Wireless Users Scenarios

We investigate the ability of our unicast-oriented system to provide scalability and fairness, taking into account that the decision algorithm operates individually for each user. Our scenarios involve multiple concurrent wireless and mobile users, having the same characteristics and requirements. Fig. 7(a) depicts the mean PSNR between all users in each scenario, for scenarios involving one, two, three, four, and five users, when the propagation delay over the bottleneck link is 100ms.



(a) Scalability in scenarios with multiple users. (b) Fairness in scenario with 5 concurrent users.

Fig. 7. Scalability and fairness concerning users with the same characteristics/requirements.

As can be seen, our system achieves scalability by sharing the available bandwidth to all active users, even in the cases where the link bandwidth is not high enough to sustain the aggregated video transmission rate. As the number of concurrent users scales up, more users can be supported by diminishing the received quality per user thus offering graceful degradation. Similarly, fairness is achieved when link bandwidth is inadequate of handling aggregated traffic. Fig. 7(b) shows that in the case of 5 concurrent users, the available bandwidth is fairly shared among them as they receive almost the same quality in terms of PSNR.

4.4 System Capacity

Fig. 8 provides an intuition for the capacity of the system with respect to the number of wireless and mobile users that can be supported by a video streaming server, taking into account the bottleneck link bandwidth. The diagram depicts the mean quality of service in terms of PSNR that is experienced by multiple identical users having the same connection characteristics, with respect to the bandwidth of the bottleneck link.

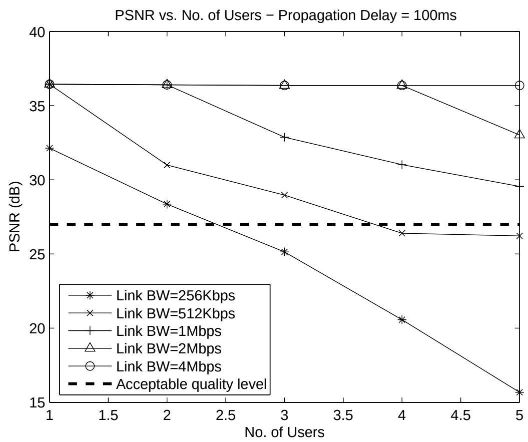


Fig. 8. Mean PSNR vs. Number of active users.

The dashed line illustrates the limit for acceptable video quality (27dB) as mentioned in Section 4.2. As the link bandwidth is high enough to sustain the aggregated video transmission rate, all users are supported by the video streaming server at equal quality levels. In particular, results show that at most two, three and four users can be supported at high quality (above 27dB) when the bottleneck link bandwidth is 256Kbps, 512Kbps, and 1Mbps respectively. If there is additional traffic, the number of the supported users will be intuitively smaller.

5 Conclusions and Future Work

In this paper we present a fuzzy-based adaptive video transmission approach specifically designed for scalable video streaming over the Internet. Our main objective is to combine NATs with CATs in order to achieve acceptable QoS levels in unpredictable mobile and wireless network environments. Thus, we introduce two new components: a feedback mechanism and a decision algorithm, that deal with layered video streams.

We evaluated our approach under various cross traffic patterns and our results indicate that the algorithm can finely adapt the video stream bit rate to the available bandwidth. Simulations showed that the proposed algorithm maintains responsiveness to various traffic patterns like CBR, FTP, and web-like cross traffic. In addition, we studied the effect of the link bandwidth and propagation delay on the QoS, and we discovered that the objective quality remains acceptable even in the presence of FTP and Web cross traffic. We demonstrated that our system is able to scale up offering graceful performance degradation and the same time the available bandwidth is fairly shared between active users who receive almost the same quality in terms of PSNR. We investigated the capacity of the system with respect to the number of users that can be supported by a video server. We showed that 2, 3 and 4 users can be supported at high quality when the bottleneck link bandwidth is 256Kbps, 512Kbps, and 1Mbps respectively.

For future work we are planning to provide a comparative study between our approach and other existing approaches in order to assess its advantages, by looking at the interaction between our adaptive flow and other flows sharing the same routers. In addition, we will investigate the capability of our approach to cope with handoff issues.

References

1. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2003.
2. S. Floyd, "TCP and explicit congestion notification," ACM Computer Communication Review, Vol. 24, No. 5, October 1994, pp. 8-23.
3. S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," IEEE/ACM Trans. on Networking, Vol. 1, August 1993.
4. C. Chrysostomou, A. Pitsillides, G. Hadjipollas, A. Sekercioglou and M. Polykarpou, "Fuzzy Explicit Marking for Congestion Control in Differentiated Services Networks," 8th IEEE Symposium on Computers and Communications (ISCC'03), 2003, pp. 312-319.
5. P. Antoniou, A. Pitsillides, and V. Vassiliou, Adaptive Feedback Algorithm for Internet Video Streaming based on Fuzzy Rate Control, 12th IEEE Symposium on Computers and Communications (ISCC'07), Aveiro, Portugal, July 1-4, 2007.
6. FFmpeg Multimedia System site, http://ffmpeg.mplayerhq.hu/.
7. Network Simulator (NS2), http://www.isi.edu/nsnam/ns/.
8. J. Klaue, B. Rathke and A. Wolish, "EvalVid - A Framework for Video Transmission and Quality Evaluation," Proc. of the 13th International Conference on Modelling Techniques and Tools for Computer Performance Evaluation, Illinois, USA, September 2003, pp. 255-272.
9. C.-H. Ke, C.-K. Shieh, W.-S. Hwang, and A. Ziviani, "An Evaluation Framework for More Realistic Simulations of MPEG Video Transmission," Journal of Inf. Sc. and Eng. (accepted).
10. V. Vassiliou, P. Antoniou, I. Giannakou, and A. Pitsillides "Requirements for the Transmission of Streaming Video in Mobile Wireless Networks," International Conference on Artificial Neural Networks (ICANN), Athens, Greece, September 10-14, 2006.