The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

ADIVIS: A NOVEL ADAPTIVE ALGORITHM FOR VIDEO STREAMING OVER THE INTERNET

Pavlos Antoniou, Vasos Vassiliou and Andreas Pitsillides
Networks Research Laboratory, Computer Science Department, University of Cyprus*
75 Kallipoleos Street, P.O. Box 20537, 1678 Nicosia, Cyprus

ABSTRACT

The transfer of information over the wireless networks is emerging as a promising business model. In addition, new applications require the transfer of video content in real time. However, the unpredictable nature of wireless and mobile networks in terms of bandwidth, end-to-end delay and packet loss has tremendous impact on the transmission of video streams. In this paper, we adopt Network Adaptation Techniques applied together with Content Adaptation Techniques to achieve graceful performance degradation when network load increases and network conditions deteriorate. We present a new feedback mechanism that provides video adaptation to network parameters, working together with a fuzzy-based decision algorithm. Our preliminary performance evaluations indicate that our algorithm can finely adapt the video stream bit rate to the available bandwidth while providing fairness as well as high and stable objective quality of service.

I. INTRODUCTION

The unprecedented spread of heterogeneous, video-enabled devices such as computers, mobile phones, and PDAs has multiplied the need for video streaming. Increasingly, there is a voiced need for efficient and effective techniques for adapting compressed video streams to suit better the different constraints, capabilities, and requirements of wireless and mobile networks, applications, services and end users.

The unpredictable nature of wireless and mobile networks in terms of bandwidth, end-to-end delay and loss variation, remains one of the most significant problems in video communications. In this context, video streaming applications need to implement highly scalable and adaptive techniques in terms of content encoding and transmission rates in order to cope with the erroneous and time variant conditions of the network.

Our approach aims at combining Network Adaptation Techniques (NATs) with Content Adaptation Techniques (CATs) in order to finely adapt the video stream bit rate to the changing network parameters. NATs deal with the end-to-end adaptation of real time multimedia 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. CATs deal with adaptation of content to the desirable transmission rate using primarily scalable video approaches.

We propose a new feedback mechanism that works in conjunction with a fuzzy decision algorithm. Their performance is discussed taking into account the influence of a critical control knob, namely, the decision period T, on the objective quality.

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

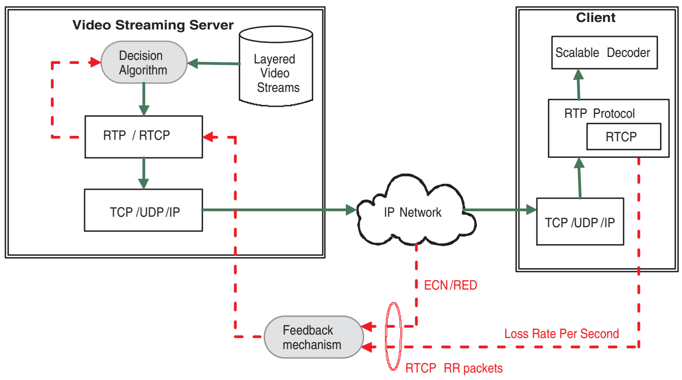


Figure 1: ADIVIS-based system.

The rest of the paper is organized as follows: Section II. analyzes the ADIVIS algorithm. Section III. deals with the evaluation setup and scenarios. Section IV. presents some preliminary results. Section V. concludes the paper.

II. ADAPTIVE FEEDBACK ALGORITHM FOR INTERNET VIDEO STREAMING

ADIVIS involves an adaptive feedback mechanism for Internet video streaming and a fuzzy decision algorithm. 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 connection.

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 illustrates a unicast-oriented ADIVIS-based system. The two outlined components, namely, feedback mechanism and decision algorithm, focus on the adaptation of the layered video content to the available network bandwidth. Dashed arrows track the path of control packets whereas solid arrows track the path of video data packets.

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

The remainder of this section is organized as follows: Section II-A. deals with layered encoding. Section II-B. presents the feedback mechanism in detail. Section II-C. analyzes the main aspects of the fuzzy decision algorithm. Finally Section II-D. deals with the evaluation of the fuzzy rate controller.

A. Layered Encoding

Layered encoding is suitable for adapting the quantity of data transmitted by a video server to the capacity of a given network path. Video streams are encoded in a layered manner in a way that every additional layer increases the perceived quality of the stream. Usually a layered video stream consists of a base layer and several additional enhancement layers. Base layers should be encoded in a very low rate so as to accommodate for a large variety of mobile handheld devices as well as terminals connected to the Internet through low bandwidth modem connections. Additional enhancement layers are added, or dropped, in order to adapt the content rate to the desirable transmission rate.

B. Feedback Mechanism

Each receiver sends reception statistics using RTCP packets. According to [1], special RTCP packets called Receiver Report (RR) packets are sent from participants that are not active senders carrying reception statistics. 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 LRPS values can be used to track the increasing/decreasing trend of packet loss percentage.

Additionally, network elements (i.e. routers) 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 [2], [3] is used for the notification of congestion to the end nodes in order to prevent unnecessary packet drops. ECN option allows active queue management mechanisms as, for example, RED [4] or Fuzzy-RED [5] to probabilistically mark packets. The number of marked packets within a given period may provide a meaningful reference about the congestion status.

C. Fuzzy Decision Algorithm

Fuzzy control may be viewed as a way of designing feedback controllers [6], [5] in situations where rigorous control theoretic approaches can not be applied due to difficulties in obtaining formal analytical models.

Linguistic variables, a key concept of fuzzy logic control, take on linguistic values which are words (linguistic terms) used to describe characteristics of the variables. Our fuzzy control system 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:

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

The LRPS parameter is lower and upper bounded by 0 and 1 respectively. Thus, D_{LRPS}(kT) 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, as a strong indication for congestion. The receiver calculates periodically this number called N_{ECN}(kT). The sender extracts this value from an RR packet 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):

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

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:

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

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. The output of the fuzzy system could have been a discrete value indicating directly the number of layers that should be sent. Instead, we chose to obtain a crisp value because we wanted our algorithm to be applicable not only in cases where the video streams are not encoded in a coarse grained manner but also when fine grained scalability encoding techniques are applied. The defuzzified crisp values of a(kT) can be used by the decision algorithm for the evaluation of the available bandwidth using the following formula:

$$avail\_bw(kT) = a(kT) * avail\_bw(kT - T) \quad (4)$$

Table 1 involves if-then rule statements which are used to formulate the conditional statements that comprise fuzzy logic.

Our algorithm has to decide which layers should be sent according to the available network bandwidth, based on a non aggressive layer selection approach. The server will host an appropriate number of layers. Each layer corresponds to a different transmission rate. 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 transmission rate that corresponds to a higher layer. Instead, a time hysteresis is introduced in order to avoid frequent transitions from one layer to another. In a lower layer transition the effect is immedi-

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

ate, as we seek quick relief from possible congestion. The time hysteresis is equal to the time interval between the reception of two successive RR packets. If the available bandwidth exceeds the threshold of a specific transmission rate that corresponds to an upper level layer, then the hysteresis variable is set. When a new RR packet arrives, if the available bandwidth is still at the same levels, a transition occurs. More detailed description of the algorithm can be found in [6].

D. *Fuzzy Rate Controller Evaluation*

The ability of the fuzzy rate controller to sense the available bandwidth of a bottleneck link in the presence of multiple CBR connections which are superimposed progressively, and adapt the transmission rate of a 1Mbps CBR non trace-based layered video stream is shown in Fig.2. The bottleneck link bandwidth is 1Mbps and the CBR cross traffic rate ranges from 200Kbps to 800Kbps. Fig. 2 depicts the instantaneous transmission rate of the layered CBR video stream as the CBR cross traffic rate changes over the time. As shown, the video transmission rate driven by the fuzzy rate controller, evolves at a slow and smooth pace in order to respond to network conditions, but also prevent unnecessarily fluctuations.

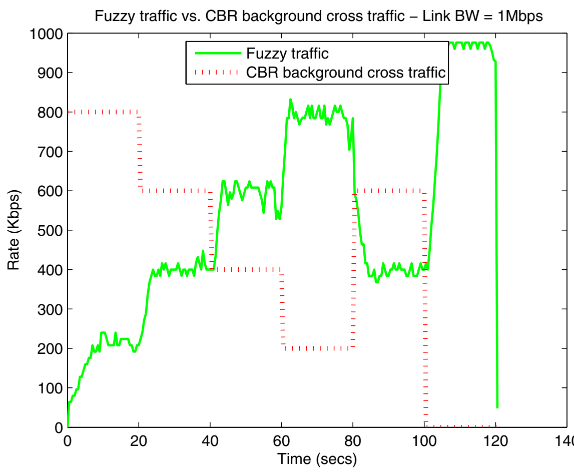


Figure 2: Instantaneous fuzzy traffic rate for CBR cross traffic.

[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).

III. Evaluation Setup and Scenarios

Fig. 3 illustrates the topology we used in the ADIVIS performance evaluation. The topology consists of two routers directly connected with a link having variable characteristics. A video streaming server is attached to the first router. Mobile wireless clients are connected to the second router over wireless links. In order to make our scenarios more realistic we added background traffic initiated by the FTP server.

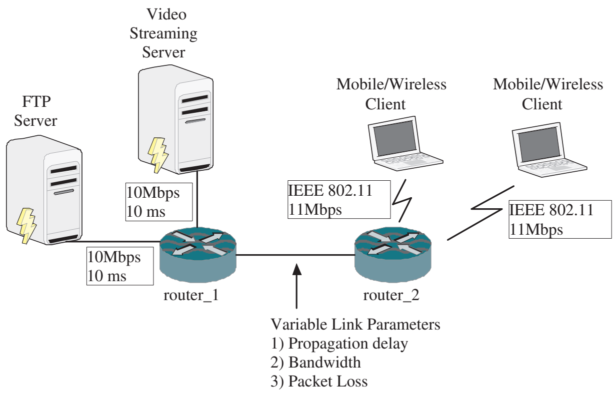


Figure 3: Evaluation Topology for ADIVIS.

In order to simulate the video traffic patterns, we used ns2 [7] and a well known real test video sequence named Foreman. It is a stream with a fair amount of movement and change of background. The sequence has temporal resolution 30 fps and spatial resolution 176x144. We encoded this sequence using a publicly available MPEG4 encoder [8] in 8 different bit rates: 64, 96, 128, 192, 256, 384, 512 and 768Kbps. Each encoded video stream corresponds to a separate layer.

We set the maximum buffer capacity 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. Further analysis of T compared with the quality of the received video stream is given in Section IV-C.

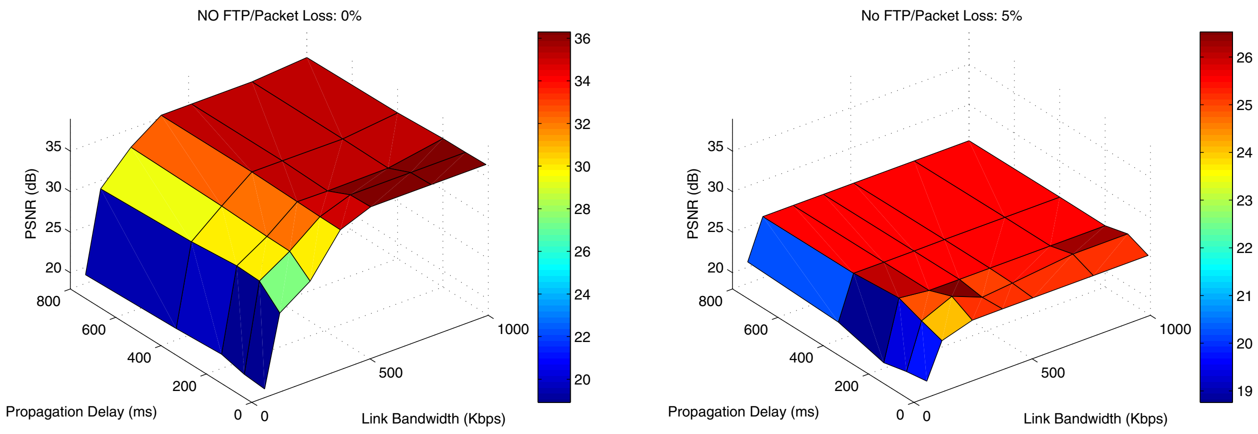
Moreover, the link bandwidth is selected to range from 64Kbps to 1Mbps and the propagation delay from 10ms to 800ms. The choice of these parameters was based on the representative characteristics of wired and wireless networks.

IV. Results

In this section we analyze the results obtained from the above scenario evaluations. In Sections IV-A. and IV-B. we present scenarios involving one and two wireless users respectively. Section IV-C. presents the inter-RR time influence on objective quality. Video quality is measured by taking the average Peak Signal-to-Noise Ratio (PSNR) over all the decoded frames.

A. *One mobile user*

The effect of propagation delay and link bandwidth on the PSNR in the absence of cross traffic is presented in Fig. 4. The results obtained by scenarios where the packet loss is 0% (Fig.



(a) Packet Loss = 0%

(b) Packet Loss = 5%

Figure 4: Mean PSNR vs. Link BW and Prop. Delay, No FTP.

4(a)) reveal that the PSNR values are increasing at a steady pace (up to 37dB) as the link bandwidth increases. PSNR values are decreased (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. Fig. 4(b) presents the results obtained by scenarios involving packet loss of 5%. Obviously, PSNR values have been significantly decreased compared to those of Fig. 4(a). This is because the decision algorithm recognizes the high packet drop rates and strives to maintain an acceptable level of video quality, whilst satisfying the worsening network state, by sending fewer layers, resulting in lower PSNR values. As shown in Fig. 4, when the link bandwidth is high enough to sustain the video transmission rate, PSNR values are slightly increased for low delay values because the adaptation evolves at a faster pace. In the case of low bandwidth links, delayed decisions caused by longer propagation delays 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 shows PSNR values for scenarios involving background FTP traffic while the packet loss is 0%. 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. We perceive lower 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 drop rates.

B. Two mobile users

Here we compare some of the results obtained previously with those concerning scenarios involving two mobile users in the absence of FTP traffic and packet loss. Fig. 6(a) depicts

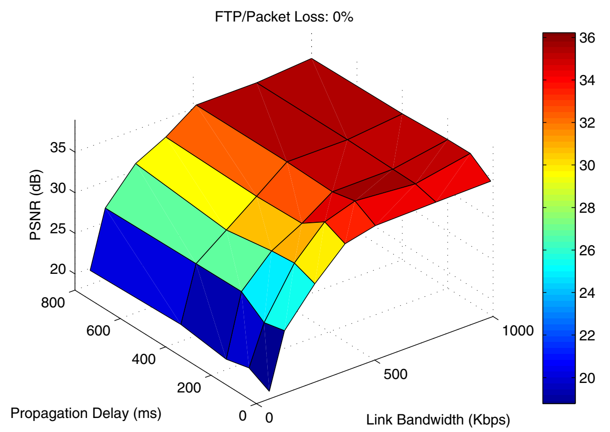


Figure 5: Mean PSNR vs. Link BW and Prop. Delay, FTP, Packet Loss = 0%.

the objective quality evaluations for scenarios involving two users compared with results involving one mobile user, when the propagation delay is 10ms. In the case of two users (Fig. 6(a)) our algorithm provides fairness, because no one of the two users takes advantage over the other, as both users receive almost the same quality. Expectedly, both users receive lower video quality than a single user because the two video streams compete with each other. Fig. 6(b) shows that in two mobile users scenarios, User 1 receives almost the same quality for both values of propagation delay. Specifically, the PSNR values for 10ms are slightly higher than for 400ms whenever the link bandwidth is high enough to sustain the video transmission rate. The same behavior is observed in scenarios involving one mobile user. In general, the quality of video received by each one of the two mobile users is lower than the quality received by one mobile user, regardless the propagation delay, because the two mobile users share fairly the available bandwidth as implied by the PSNR. Further analysis of the quality received by multiple users, is planned for future work.

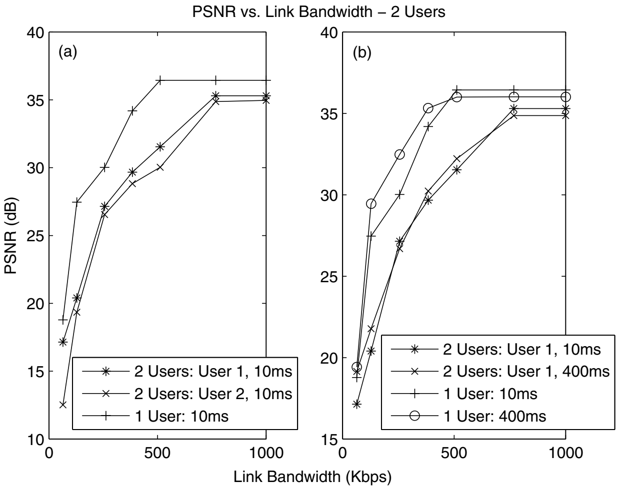


Figure 6: Mean PSNR vs. Link BW.

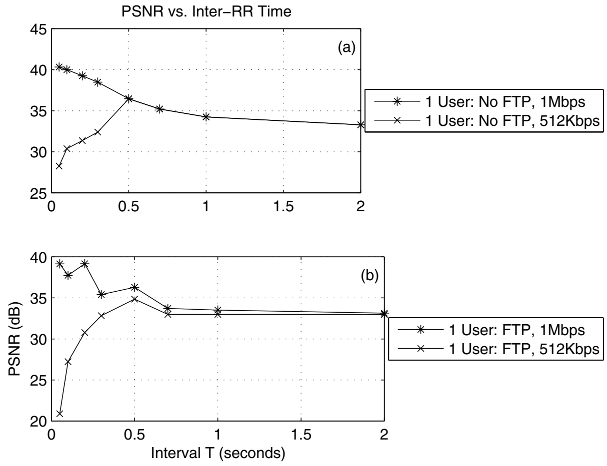


Figure 7: Mean PSNR vs. Inter-RR Time.

C. *Inter-RR Time Influence on Objective Quality*

In all the aforementioned scenarios, the interval T between successive RR reports was 0.5 seconds. The reason for this choice is made clear below. Here we study the impact of the decision period T on the objective quality of the received video streams. Fig. 7(a) reveals that in the absence of FTP cross traffic when bandwidth is 1Mbps and T is close to 0, the PSNR reaches 40dB followed by a slight decrease approaching 33dB as the T increases, because layers are superimposed at a slower pace. In the case of 512Kbps, the PSNR reaches its peak value when T is 0.5 seconds. There is a gradual decrease of PSNR down to 28dB as the T diminishes to 0 as a result of the huge number of reports which cause instant changes of the transmission rate which cannot be accommodated by the link bandwidth. For intervals larger than 0.5 seconds, the PSNR curve follows the same behavior as in the case of 1Mbps because the reduction of T implies that congestive phenomena (due to limited bandwidth) will not occur frequently. In the presence of FTP cross traffic (Fig. 7(b)), the PSNR values are relatively smaller compared with previous graphs. For a link of 1Mbps, some fluctuations occur due to the dynamically changing available bandwidth. As for 512Kbps, a steep decrease of PSNR is observed when T is reduced as a consequence of congestion caused by FTP cross traffic and fast transitions in layer selection.

V. Conclusions and Future Work

In this paper we present ADIVIS, an adaptive video transmission algorithm specifically designed for video streaming over the Internet. Our main objective is to provide a framework that incorporates both Content Adaptation and Network Adaptation Techniques. Towards this direction, we introduce two new components; a feedback mechanism and a decision algorithm, that deal with layered video streams.

We evaluated ADIVIS under error-free and error-prone environments and our results indicate that the algorithm can finely adapt the video stream bit rate to the available bandwidth, while providing high and stable objective quality of service. Moreover, simulations showed that ADIVIS performs best in the absence of cross traffic while the objective quality remains acceptable in the presence of cross traffic. It seems that our algorithm provides fairness, however, this is an issue which will be further investigated in the presence of more than two users. We studied the impact of the inter-RR time T on the objective quality and we found that small values of T drive the system to instability resulting in lower PSNR, whereas large values make the system unresponsive to transient changes. Thus, the selection of 0.5 seconds is dictated by the desire to maintain responsiveness to changes in the network state providing stability and avoiding fast transitions in layer selection.

For future work we are planning to evaluate our algorithm using more extensive scenarios, taking into account other types of cross traffic e.g., web traffic. Moreover, our proposed fuzzy-based approach should be compared with other existing relevant approaches in order to assess its advantages, by looking at the interaction between our adaptive flow and other flows sharing the same routers.

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, Oct. 1994, pp. 8-23.

[3] K. Ramakrishnan and S. Floyd, "A proposal to add explicit congestion notification (ECN) to IP", *RFC 2481*, Jan. 1999.

[4] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance", *IEEE/ACM Trans. on Networking*, Vol. 1, Aug. 1993.

[5] C. Chrysostomou, A. Pitsillides, G. Hadjipollas, A. Sekercioglou and M. Polykarpou, "Fuzzy Explicit Marking for Congestion Control in Differentiated Services Networks", *8th IEEE ISCC'03*, 2003, pp. 312-319.

[6] P. Antoniou, A. Pitsillides, and V. Vassiliou, "Adaptive Feedback Algorithm for Internet Video Streaming based on Fuzzy Rate Control", *12th IEEE ISCC'07*, July, 2007.

[7] NS2 site, http://www.isi.edu/nsnam/ns/.

[8] FFmpeg System site, http://ffmpeg.mplayerhq.hu/.