Performance Evaluation of MPEG-4 Video Transmission with the Adaptive Smooth Multicast Protocol (ASMP)

Christos Bouras, Georgios Kioumourtzis
Computer Engineering and Informatics Department
University of Patras
Rion, Greece
bouras@cti.gr , gkioumou@ceid.upatras.gr

Apostolos Gkamas
Research Academic Computer Technology Institute
Rion, Greece
gkamas@cti.gr

Abstract— **We present in this work the performance evaluation of MPEG-4 video transmission with our proposed single rate multicast protocol named Adaptive Smooth Multicast Protocol (ASMP). ASMP key attributes are: a) adaptive scalability to large sets of receivers, b) TCP-friendly behavior, c) high bandwidth utilization, and finally d) smooth transmission rates which are suitable for multimedia applications. We evaluate the performance of ASMP under an integrated simulation environment which extends ns-2 and Evalvid-RA to the multicast domain with the use of the RTP/RTCP protocols. Simulations conducted under this environment combine the measurements of network-centric along with video quality metrics. This "joint" evaluation process provides a better understanding of the benefits and limitations of any proposed protocol for multimedia data transmission.**

Keywords-component; Multicast; congestion control; multimedia transmission; ns-2; simulation

I. Introduction

Video distribution has gained the interest of industry and the research community over the last few years. An efficient way to disseminate a video file to a number of users in terms of bandwidth consumption is through multicasting. However, in the field of multimedia applications there are two main issues that need to be addressed. The first one is the lack of congestion/flow control so that these applications can fairly share the available network resources with TCP-based applications. The second is how to meet the Quality of Services (QoS) guaranties given the bandwidth constraint.

Up to now there is an efficient number of promising approaches in literature based either in analytical models or in network statistics. An early work, TFMCC [1], extends the basic mechanisms of TFRC [2] to support single stream multicast congestion control. The most important attribute of TFMCC is the suppression of feedback receiver reports. PGMCC [3] is a window-based TCP scheme, which is based on positive ACKs between the sender and the group representative (the *acker*). LDA+ [4] is an additive increase and multiplicative decrease (AIMD) algorithm, in which the addition and reduction values are dynamically determined based on current network conditions. Explicit Rate Multicast Congestion Control (ERMCC) [6] implements a congestion control scheme that is based on a new metric named TRAC (Throughput Rate At Congestion). Most of these proposals have been evaluated through simulations conducted mainly with the ns-2 [7] simulator software. However, simulations were not based on any multimedia traffic generation model and in the best case trace files were used instead. The above limitations, as part of the simulation environment, undermine the performance evaluation studies in which quality measurements for multimedia data transmission (e.g. Peak Signal to Noise Ratio (PSNR), Mean Opinion Score (MOS), etc) are missing. Therefore, the only quality indicators were purely based on network-centric metrics (e.g. packet loss ratio, delay jitter etc). At this point we need to mention that the evaluation of multimedia transmission based only on network metrics can lead to debatable results because the perceived multimedia quality at the end user is not measured. One possibility for simulating real video transmission under a rate adaptive control scheme is to use the EvalVid-RA [8] tool-set. However, Evalvid-RA is also restricted to unicast transmission and therefore simulations and performance evaluation studies with multicast protocols are excluded. For the purpose of this paper we have extended EvalVid-RA from the unicast to multicast domain so that it can also support multicast transmission of multimedia data.

We present in this work the performance evaluation of MPEG-4 video transmission with our proposed protocol termed as the Adaptive Smooth Multicast Protocol (ASMP). ASMP is a new single-rate multicast transport protocol for multimedia applications. In ASMP, each receiver calculates a TCP-friendly bandwidth share based on the TCP Friendly Rate Control (TFRC) specification [2]. More details on the functionality of ASMP can be found in [9]. We focus on a detailed evaluation which is based on both network-centric and video quality metrics. This "joint" evaluation process provides a better understanding of the benefits and limitations of any proposed protocol for multimedia data transmission.

The rest of this paper is organized as follows: In the next section we provide an overview of the simulation environment. In section III we present performance evaluation results. We conclude our paper in section IV.

978-1-4244-7755-5/10/$26.00 ©2010 IEEE

II. SIMULATION ENVIRONMENT

In order to set up our test-bed we extend Evalvid-RA to multicast domain and integrate it with the ns-2 simulation software. This simulation environment consists of three parts and is depicted in Figure 1.

During the pre-processing a raw video file, which is usually stored in YUV format, is encoded with the desired video encoder[1], into 30 different encoded MPEG-4 video clips with quantizer scale values in the range 2 to 31. Quantizer scale 2 provides an encoded video with the highest quality. We use the *ffmpeg* [10] free video encoder for the creation of the video clips. For our simulations we create a raw video that consists of video sequences, which differ in complexity with the following order: News (frame 1-300) has medium complexity, Akiyo (300-600) has very low complexity as it contains more static information, Stefan (600-700) is very complex due to continuous moving pictures, and lastly Paris (700-1753) which has high to medium complexity. All video clips have temporal resolution of 25 frames per second. The frame size of all clips is 352 x 288 pixels, which is known as the Common Intermediate Format (CIF). We use the above files in order to create a large YUV video sequence, which consists of 10000 frames with duration of 400 seconds. By combining various video sequences of different complexity we can better simulate video transmission as video complexity is changed over time. The ns-2 creates the simulated network. The video file is transmitted from the server to the group of multicast receivers. During the simulation time we store the traces for both the server (*sd_be_0*) and the receivers (*rd_be_0*) to enable easy calculation of network metrics and video performance evaluation metrics (PSNR, MOS). The third part of the simulation environment is consisted of the reconstruction of the transmitted video and the measurements for the performance evaluation assessments.

III. PERFORMANCE EVALUATION

As we mentioned in the previous sections the performance evaluation is not only based on "classic" network metrics but also on quality measurements of real video transmission, namely PSNR and MOS. Although, MOS is a subjective evaluation metric we can obtain a MOS grade based on the corresponding PSNR value [8].

The PSNR values of all individual video frames are averaged to produce the mean PSNR of the complete video sequence; which is then mapped to the corresponding MOS value. However, we need to point out, that PSNR mapping to MOS values provides only a rough estimation of the perceived video quality by the end user. For more accurate results real experiments with a sufficient number of viewers should be conducted when possible. We also measure the cumulative inter-frame jitter, and the video frame error rate.

The cumulative inter-frame jitter is defined as the amount of playback delay that must be provided in order to avoid discarding video frames at the client side [11]. The video frame error rate is defined as the amount of lost frames or frames with missing data divided by the total number of transmitted frames by the sender. We conduct several simulations under three different scenarios in order to investigate:

- The TCP-friendliness of ASMP.
- The perceived video quality by the end user.
- The impact of packet losses on the video quality.
- The intra-fairness and the stability of ASMP, and finally
- The responsiveness and the performance of ASMP when sharing network resources with UDP traffic, which does not employ any congestion control mechanism.

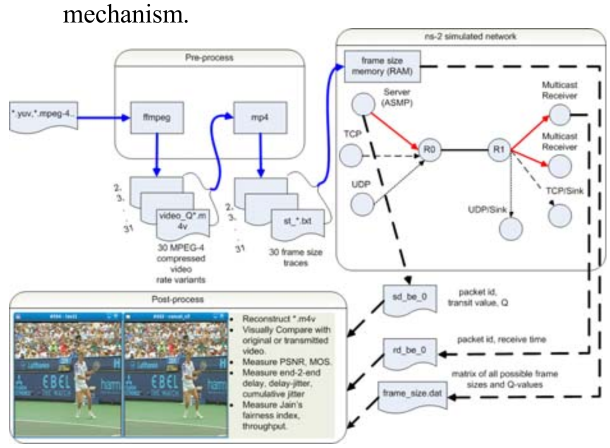

Figure 1. Simulation Environment.

A. TCP-fairness

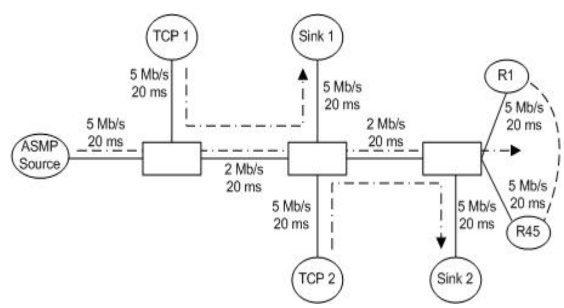

Figure 2. Parking-lot bottleneck scenario.

We use the simulation scenario depicted in Figure 2 in order to simulate a topology with multiple bottleneck links (something very common to the real word) on a path between a sender and receiver(s). In this scenario, there is a long multicast video flow, passing through two bottleneck links. Two additional TCP short flows are passing through only one bottleneck link. R1 through R45 stands for the video receivers. We use in our simulations Drop Tail queue in the routers and set the access link capacity of all agents to 5 Mb/s with an access delay of 20

[1] Currently H.263 and MPEG-4 are supported by Evalvid-RA.

milliseconds. Therefore, the total one way delay on the path from the video source to video receivers is 80 milliseconds giving an RTT of 160 milliseconds. The packet size for both the ASMP and the TCP flows is set to 972 bytes. According to proportional fairness the ASMP video flow should get no more that the one half of the bottleneck bandwidth, which is 1 Mb/s.

Figure 3 depicts the achieved throughput during the simulation time. For easier observation we present only the achieved throughput of one multicast ASMP receiver and the aggregated throughput of TCP receivers (Sink1 and Sink2). All multicast receivers R1 through R45 have similar performance as the sender's transmission rate in ASMP is always driven by the "slowest" receiver. We observe that TCP flows enjoy higher throughput than ASMP receivers. It is one of our design goals, however, to "smooth" the behavior of ASMP in any case of packet losses due to network congestion. As a result, ASMP is "slower" than TCP when recovering from packet losses or from a congested stage. We measure the fairness of ASMP by using the Jain's Fairness Index [12]. By calculating the Jain's fairness index for the three flows (two TCP flows and one ASMP) we obtain a value of 0.982, which means that the simulated network as a system is 98.2% fair when ASMP shares network resources with TCP. These results are very encouraging as they not only verify the TCP-friendliness but also the fairness of ASMP.

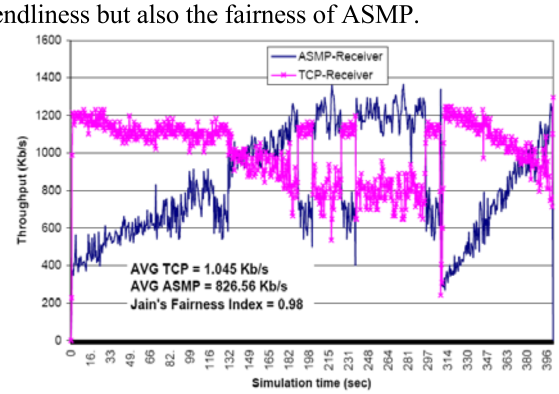

Figure 3. Throughput of all flows.

Figure 4 presents the delay jitter measurements of the video flow, which are bounded between 15 and 50 milliseconds. These results disclose that ASMP has very high performance when competing for network resources with TCP. The low value of delay jitter is an important attribute of any transmission protocol for multimedia data. Someone can claim that delay jitter measurement is not so important and a solution to address high delay jitter is to increase the receiver's buffer size in order to "smooth" the negative effects of late arriving packets. However, this is not true in fact for real time applications (video conferencing and VoIP) as delayed packets are useless and as such are discarded at the end user. What is also important is the effect of the cumulative jitter delay when dealing with video streaming applications. Once the cumulative jitter of a video frame exceeds the playback buffer duration, the video frame is discarded. In our simulations we set a fixed playback buffer of 500 milliseconds. A larger playback buffer allows late arriving frames to be decoded at the expense of longer start-up time experienced by the viewer. In addition, a large playback buffer requires extra storage capacity. Figure 5 depicts the cumulative jitter delay of ASMP receiver R1 and we can observe that values are within the 500 milliseconds buffer limit. The lower the values of cumulative jitter the higher the performance of the rate adaptive transmission scheme. With the thirty different video sequences that require different transmission rates we can address network congestion by assigning the proper transmission rate based on current network conditions. Therefore, we are able to minimize or at least to undermine the effects of a congested link on the perceived video quality at the end user.

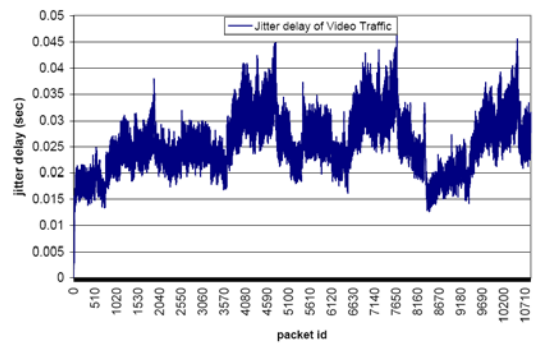

Figure 4. Jitter delay at ASMP receiver R1.hroughput of all flows.

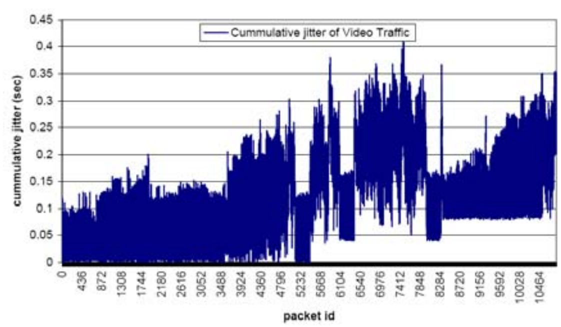

Figure 5. Cumulative jitter delay at ASMP receiver R1.

Following our evaluation we present in Figure 6 the obtained PSNR values by directly comparing the encoded video at the sender with the one "as seen" by receivers. MOS values are matched to PSNR values (Figure 7) to get an indication on the perceived video quality at the end user. Although the average MOS cannot directly lead us to assess the quality of the whole video sequence at the receiving end, it is mentioned here to indicate that on average the video experience could be regarded by the end user as "good". We observe from MOS values that the biggest part of the video sequence is graded as "excellent" with the exception of some frames that are assessed as "poor".

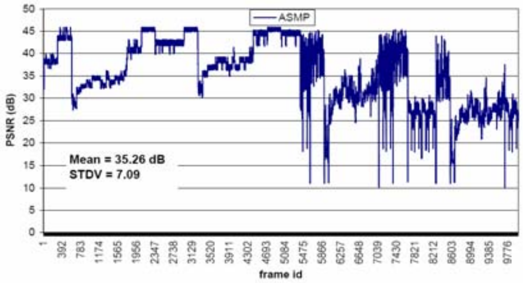

Figure 6. PSNR values of the received MPEG-4 video.

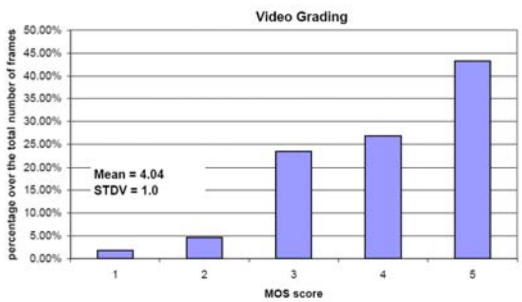

Figure 7. MOS grading of the received MPEG-4.

B. *Intra-fairness*

In this simulation scenario we connect two ASMP sources that transmit the same video file to two different multicast groups as shown in Figure 8.

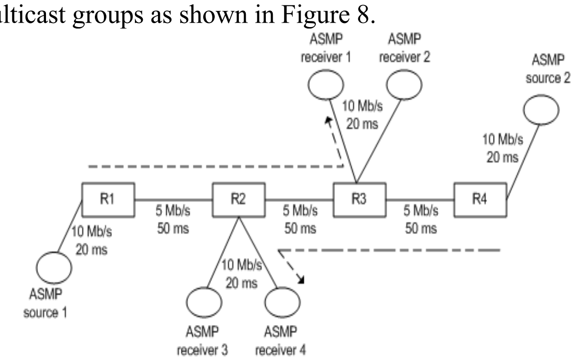

Figure 8. Network topology for ASMP-intra fairness evaluation.

The two sources transmit to opposite directions via links with the same bandwidth and propagation delay. We use Drop Tail queues in the routers R1 to R4 and set the same packet size for the two video sources. Simulation results from two representative receivers of each group (Figure 9) disclose that ASMP sources fairly share the available bandwidth in the bottleneck links with a measured Jain's fairness index of 0.99. Jitter delay is also consider low and is bounded between 5 and 20 milliseconds (Figure 10). PSNR values (Figure 11) show that at the beginning of the transmission the two ASMP sources attempt to transmit at high bit rates which provides PSNR values above 40 dBs. However, shortly after the transmission of the first two thousand frames both sources adjust their transmission rates as they observe higher RTT values due to congestion. MOS grading (Figure 12) also shows a "good" user experience. The frame loss ratio is very low resulting at a value of 0.0007 (7 lost frames out of total 10000 transmitted). Furthermore, we measure the stability of ASMP by using the coefficients of variation (CoV)[2] of the throughput values and plot the results in Figure 13. We observe that ASSP shows good stability.

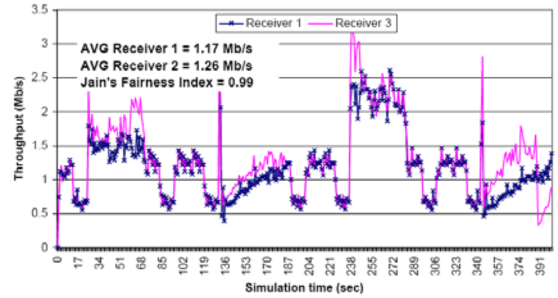

Figure 9. Throughput ASMP receivers.

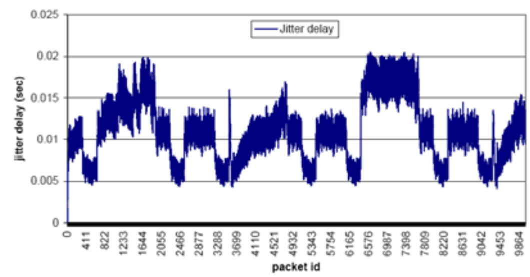

Figure 10. Jitter delay measurements of ASMP intra-fairness evaluation.

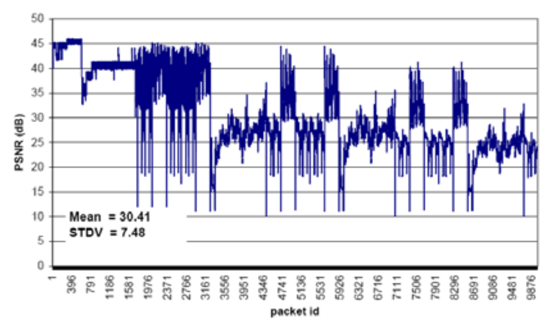

Figure 11. PSNR values of the received MPEG-4 video.

To obtain a better understanding on the dependencies between the video quality, the achieved throughput and the propagation delay we run several simulations with different RTT values. Figure 14 shows the throughput variations over different RTTs. We observe that ASMP achieves higher throughput with lower RTT values. On the other hand it is interesting to notice that the low throughput values in the case of RTT of 560 milliseconds provide better video quality as the frame loss ratio is low. This is the case of all simulation results when directly

[2] Coefficient of Variation (CoV) is the standard deviation divided by the mean.

comparing the achieved video quality to frame loss ratio values. Therefore, our assessment is that the aggressiveness of any transport protocol for multimedia applications should be carefully defined as it may lead to frequent packet losses that undermine the video quality. The aggressiveness is defined in RFC 5166 [13] as the maximum increase in the sending rate in one RTT, in packets per second, in the absence of congestion.

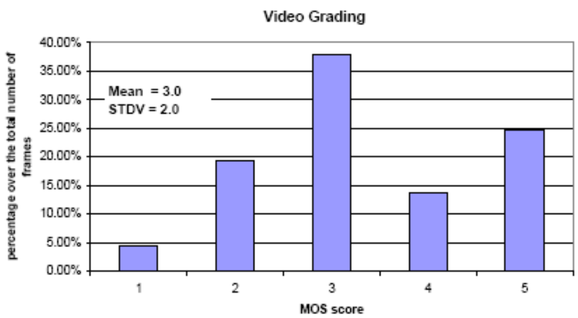

Figure 12. MOS grading of the received MPEG-4 video.

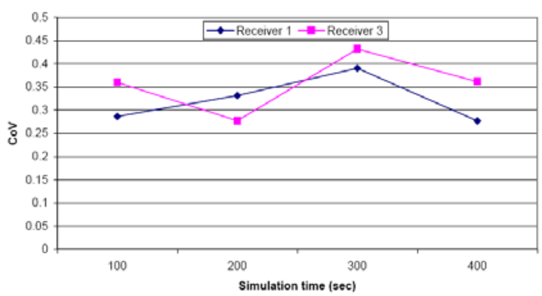

Figure 13. CoV of the two ASMP flows.

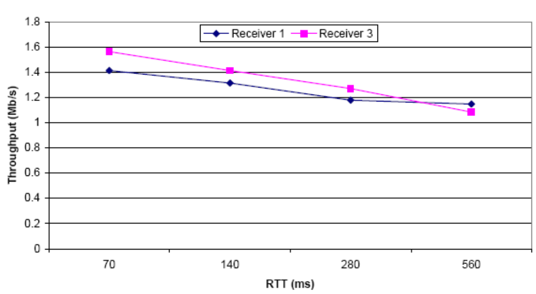

Figure 14. Achieved throughput of ASMP flows (different RTTs).

Therefore, moderate and stable transmission rates with minimum losses provide a better service to the end user in terms of video quality. Otherwise, error resilience mechanisms should be applied to avoid at least the losses of I-frames that have the highest importance in the video sequence. ASMP performs better in topologies with realistic RTT of several hundreds of milliseconds as the protocol prevents the high oscillations of the transmission rate which definitely lead to packet losses.

C. Competing with UDP Traffic

In this simulation we investigate the ability of ASMP to react to congestion and adjust the sender's transmission rate when competing for network recourses with UDP traffic. We use the network topology in Figure 2 and replace the TCP background traffic with one UDP flow that passes through the two bottleneck links. To better test the responsiveness of ASMP we vary the available bandwidth as a square wave by injecting UDP traffic throughout the simulation lifetime. UDP traffic is transmitted by a Constant Bit Rate (CBR) source at 1 Mb/s. We observe in Figure 15 that ASMP demonstrates its ability to control the sender's transmission rate in such way that packet losses due to congestion are minimized. This becomes more obvious when we measure the frame loss ratio to be 0.0041 (41 lost frames out of 10000 total transmitted).

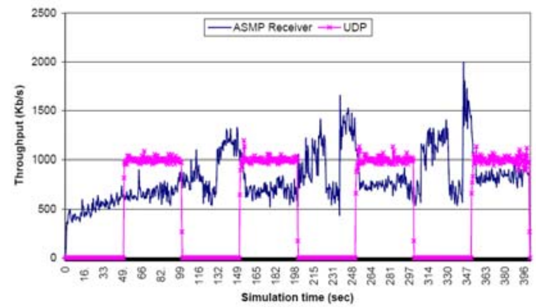

Figure 15. Throughput of ASMP and UDP flows.

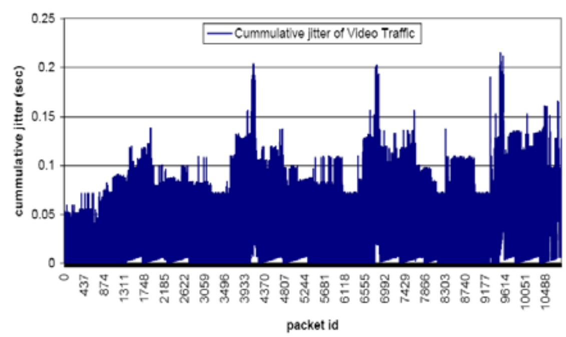

Figure 16. Cumulative jitter delay at ASMP receiver R1.

Figure 16 depicts the cumulative jitter results. We observe that values are lower than 0.2 seconds which indicate that we can succeed a lower video start-up time at the end user and at the same time lower buffer requirements. PSNR values indicate on average a "fair" video experience by the end user.

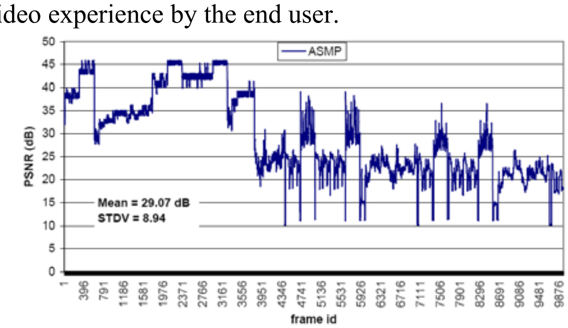

Figure 17. PSNR values of the received MPEG-4 video.

Values of PSNR (Figure 17) would have been higher if we would have set up bottleneck links with higher capacity in our simulation scenario. This is due to the fact that for the specific tested video sequence the lowest quantizer value requires at least available bandwidth of 1642 Kb/s in order to provide the best quality video frame. Our intention was, however, to stress ASMP and investigate its performance in scenarios that involve links with low bandwidth. MOS values in Figure 18 indicate a rather "fair" experience by the end user. The grading in this simulation is lower than in the previous simulation with competing TCP traffic. However, we expected these results as UDP does not employ any flow control mechanism.

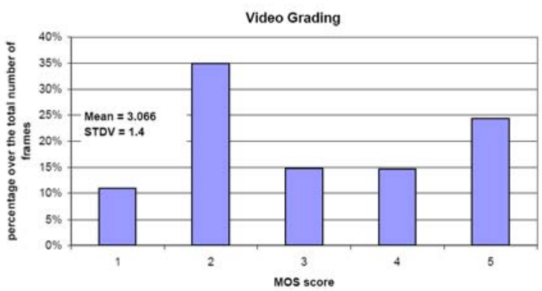

Figure 18. MOS grading of the received MPEG-4 video.

IV. CONCLUSIONS-FUTURE WORK

We presented in this work the performance evaluation of adaptive MPEG-4 video transmission with our proposed multicast protocol named ASMP. Simulation results under that new simulation environment came closer to a real video experimental evaluation process. PSNR values were measured throughout the simulations to better assess the perceived video quality by the end user. MOS grading based on mapping to PSNR values provided only a rough estimation on the video quality. When possible, experiments with real human interactions would provide more precise results. ASMP demonstrated its TCP-fair behavior while being able to deliver MPEG-4 video files with high quality. The small cumulative jitter values reduced large play-out delays and increased QoS. The aggressiveness of a transport protocol should be balanced between achieved throughput and packet losses as high aggressiveness leads to congestion. This situation creates also high oscillations in the transmission rates which are an undesired effect for video transmission applications. Simulation results with competing UDP traffic disclosed that ASMP was able to adjust and reduce the transmission rate to avoid packet losses. Uncontrolled video transmission without any flow/congestion control mechanisms should be avoided. Otherwise, there is always the possibility that TCP-based applications will starve. A second negative effect is the poor quality that is finally offered to the end user. In our future work we will extend the performance evaluation experiments in the field of multi-rate multicast transmission to increase bandwidth utilization. ASMP could also become the building block of multi-rate multicast schemes. It would be also interesting to employ ASMP in wireless communications in which the packet losses are not highly coupled to congestion. Finally, sources simulation scripts and results are publicly available in [14].

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and Mr. Andy Bolt for his assistance in proof reading this paper.

REFERENCES

[1] RFC 4654, "On TCP-friendly Multicast Congestion Control (TFMCC)", J. Widmer, M. Handley, August 2006.

[2] RFC 5348, M. Handley, S. Floyd, J. Padhye, J. Widmer, "TCP Friendly Rate Control (TFRC)", Network Working Group, September 2008.

[3] L. Rizzo, "pgmcc: A TCP-friendly single-rate multicast congestion control scheme," in Proc. of ACM SIGCOMM '00, Stockholm, Sweden, 2000.

[4] Sisalem D., Wolisz A., "LDA + TCP - Friendly Adaptation: A Measurement and Comparison Study," in Proceedings the 10th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'2000), June 25 - 28, 2000, Chapel Hill,NC, USA.

[5] RFC 3550, "RTP: A Transport Protocol for Real-Time Applications", H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, July 2003.

[6] L. Jiang, M.Yuksel, S. Kalyanaraman, "Explicit rate multicast congestion control", Computer Networks: The International Journal of Computer and Telecommunications Networking, Volume 50, Issue 15, (October 2006), Pages: 2614 – 2640, 2006.

[7] The Network simulator ns-2. Available from: http://www.isi.edu/nsnam/ns/.

[8] Lie A., Klaue J., "Evalvid-RA: Trace Driven Simulation of Rate Adaptive MPEG-4 VBR Video" ACM/Springer Multimedia Systems Journal autumn 2007.

[9] C. Bouras, A. Gkamas, G. Kioumourtzis "Adaptive Smooth Multicast Protocol for Multimedia Transmission: Implementation Details and Performance Evaluation" International Journal of Communication Systems, Wiley InterScience, Published on line Oct 2009, DOI: 10.1002/dac.1059.

[10] FFmpeg source code. Available from http://sourceforge.net/projects/ffmpeg/

[11] L. Tionardi and F. Hartanto, "The Use of Cumulative Inter-Frame for Adapting Video Transmission Rate", Proceedings of TENCON, Bangalore, India, 2003.

[12] R. Jain, D. Chiu, and W. Hawe, "A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems," DEC Research Report TR-301, Tech. Rep., Sept 1984.

[13] RFC 5166, S. Floyd, "Metrics for the Evaluation of Congestion Control Mechanisms", Network Working Group, March 2008.

[14] ASMP source code. Available from: http://ru6.cti.gr/ru6/ns_rtp_home.php