Analysis of Management Policies for Multicast Transmission of Scalable Video Content in Next Generation Networks

Iffat Ahmed[*‡], Leonardo Badia[§‡]

[*] IMT Institute for Advanced Studies, Piazza S. Ponziano 6, 55100 Lucca, Italy

[§] University of Padova, Dept. of Information Engineering, via Gradenigo 6B, 35131 Padova, Italy

[‡] Consorzio Ferrara Ricerche, Ferrara, Italy

email: *iffat.ahmed@imtlucca.it, badia@dei.unipd.it*

Abstract—Users of video multicast groups are highly heterogeneous in terms of individual channel conditions and requirements for video transmission. Their experienced quality may vary, making it a challenging task for the network to optimally configure the resource management. In this paper, we consider mathematical model to represent layered video content delivery in a multicast group. We compare various network policies to choose the optimum number of transmit opportunities and we investigate the role of feedback, which, if present, dynamically tunes the resource management. We analyze the actual perceived quality of the users as well as how their satisfaction levels vary in the multicast session. Simulation results show that the presence of feedback generally enhances the overall users quality; however, this improvement is heavily related to the resource allocation policy of the operator.

Index Terms—Video transmission; multicast; feedback; resource allocation.

I. INTRODUCTION

VIDEO contents are very sensitive in nature and need special treatment for being transmitted over heterogeneous communication networks. In the current Internet structure, all the contents are treated equally, hence, there is no support to provide efficient video delivery and enhance the Quality of Experience (QoE), i.e., the actual quality perceived by the users, directly related to the user satisfaction [2]. Due to the proliferation of enhanced wireless technologies and video codecs, the demands are shifting towards user centric approaches, where satisfaction of requirements is of much importance compared to traditional Quality of Service (QoS) improvements through a network centric approach.

Most of the research research on video services usually considers point-to-point scenarios, and less investigation has been done for multicast. Similarly, many papers have studied how to improve the quality of transmitted video by using mechanisms for efficient resource allocation [3], congestion control [4] or bandwidth/rate management [5]. Multicast communication is an important solution for distributed multimedia applications, particularly video, to efficiently use the network resources by exploiting spatial and content redundancy of the user requests.

Among the unique characteristics of video traffic, we consider especially its layered structure, with particular reference to Scalable Video Coding (SVC) [1], [6], which transmits a base layer of video bitstreams and also enhancement layers that further improve the quality. As long as the users receive enough SVC base layer contents, they are able to decode the video, and when they receive the SVC enhancement layers, they actually start improving the quality of received video contents.

In [7], solutions are provided for video multicast in infrastructure based broadband wireless networks, where layered hybrid Forward Error Correction/Automatic Repeat reQuest (FEC/ARQ) [13] is used to combat packet losses and different resources are assigned to multiple video layers using Orthogonal Frequency Division Multiple Access (OFDMA) techniques. An enhanced configuration leveraging Adaptive Modulation and Coding (AMC) schemes is obtained, though feedback mechanisms are not considered.

To incorporate accurate video quality metrics, [8] provides a framework for rate allocation. It also involves cross layer design and tries to provide differentiated QoE for various video sources. The emphasis is on maximizing a weighted QoE value for each video source considering Medium Access Control (MAC) layer scheduling. However, it does not discuss the impact of the number of transmit opportunities on QoE.

Using a cross-layer feedback approach, [9] tries to estimate the user received video quality. The transmitted video signal and received video quality are monitored in three phases, i.e., signal encoding, transmission on lossy channel and the decoded signal. Such an estimation can be used to improve the performance of video transmission from the user perspective.

Feedback of video services in a multicast environment is also investigated by [10] and [11]. The former focuses on congestion control mechanism in Asynchronous Transfer Mode (ATM) networks, whereas, the latter approaches the layered video multicast problem through game theory and deals with error control. Finally, [12] provides a framework to mitigate transmission impairments or avoiding error propagation at the receiver. The focus is experimental and the considered scenario is video conferencing over a Wireless Local Area Network (WLAN). However, this work does not exploit the layered structure of the video content, which induces some packet prioritization, nor it considers different user requirements within the same multicast group.

In this paper, we model the network configuration based on the individual requirements of the users and apply various policies regarding the number of transmit opportunities required for a multicast session. Finally, we investigate the effect of feedback, i.e., how it impacts on user perceived quality, and any improvements brought to the QoE of users. We also define a mathematical model for layered video feedback in a multicast environment, analyzed both theoretically and via simulation.

The rest of the paper is organized as follows: Section II defines our proposed mathematical model, describing unicast and multicast scenarios. Section III develops the analytical model and provides extensive simulation results. Finally, Section IV discusses the findings and concludes the paper.

II. Proposed Framework

Consider a multicast transmission of the SVC source, including a base layer and enhancement layers to provide better video quality. For simplicity, we just assume two layers, i.e., one enhancement layer beyond the base layer. The extension to multiple enhancement layers would be conceptually straightforward. Thereafter, indices "1" and "2" refer to the base and enhancement layers, respectively. This section introduces a taxonomy of user categories according to their satisfaction level and mathematically models the choice of the optimal number of transmit opportunities (TxOps) required for the users to decode the base layer and maximize their perceived quality in a multicast environment. Finally, the framework also describes the network policies evaluated in Section III.

We define the following quantities:

- T is the total number of TxOps available for transmission.
- t_1, t_2 as the number of TxOps for the base and the enhancement layer, respectively; note that $t_1 + t_2 = T$.
- $\hat{n}_1(j), \hat{n}_2(j)$ as the number of packets actually received at BL and EL layers for any user j, respectively.
- threshold θ_j is the required number of base layer packets that each user j needs to correctly decode the stream.
- p_j is the loss probability of user j.
- P_j^d is the decoding probability for user j.

A. User Satisfaction Levels

To check and measure the user satisfaction about the perceived video quality, we define the following labels:

Satisfied — A user j is said to be *satisfied*, if it has *received* enough base layer packets to decode the video within available number of transmit opportunities, i.e., $\hat{n}_1(j) > \theta_j$.

Happy — A user is said to be *happy*, if it *expects* that it will eventually receive the required number of base layer packets to decode the video within the available transmit opportunities. Whenever this expectation is confirmed, this user will become satisfied. Conversely, if the expectation is contradicted, the user will become e-unhappy or hopeless.

UnHappy — A user is said to be *unhappy*, if it will likely *not receive* the required number of base layer packets to decode the video before the end of the available transmit opportunities. Similarly, confirmation of this expectation leads the user to become *Hopeless* or the expectation can change to *Happy* or *Satisfied*.

Hopeless — User j is said to be *Hopeless* if it is *sure* that it will *never be satisfied* within the available number of transmit opportunities. That is, it will not receive the required number of base layer video packets to decode correctly within the available transmit opportunities, i.e., $\hat{n}_{1_j}(T) < \theta_j$.

B. Mathematical Model

The model considers two multicast scenarios. In the former, we consider a single packet transmission per time slot and a single user in the system. This scenario is useful to understand the latter, where multiple packets are transmitted depending on the channel quality, and also multiple users are considered.

1) Scenario A: We focus on a single user j trying to decode the video stream (or, alternatively, we can consider several users identical to each other). We assume that different packets are decoded with independent identically distributed (i.i.d) probabilities for the sake of simplicity. Therefore, the probability to correctly receive a packet is $1-p_j$. The decoding probability of the flow for user j corresponds to the probability of decoding θ_j packets out of t_1 TxOps, i.e.

$$P_j^d = \begin{cases} 0 & \text{if } (1-p_j) \cdot t_1 < \theta_j \\ F & \text{if } (1-p_j) \cdot t_1 \geq \theta_j \end{cases} \tag{1}$$

where F is defined as

$$F = \sum_{k=0}^{t_1-\theta_j} \binom{t_1}{k} (1-p_j)^{t_1-k} (p_j)^k \tag{2}$$

We formalize the expected goodput of user j as:

$$\hat{q}_j = \begin{cases} 0 & \text{if } non\text{-}decoding \\ \alpha \cdot \hat{n}_1(j) + \beta \cdot \hat{n}_2(j) & \text{if } decoding \end{cases} \tag{3}$$

where α and β are properly chosen weights that regulate the relative importance of base layer packets versus enhancement layer packets in terms of supplied quality. Thus, the average perceived goodput is

$$\begin{aligned} \mathbb{E}[\hat{q}_j] &= \left(\alpha t_1 (1-p_j) + \beta t_2 (1-p_j) \right) \\ &\times \sum_{k=0}^{t_1-\theta_j} \binom{t_1}{k} (1-p_j)^{t_1-k} (p_j)^k u[t_1-k] \end{aligned} \tag{4}$$

where $u[\cdot]$ is a unit step, i.e., it is one if its argument is non-negative, zero otherwise.

It is straightforward to identify the optimum value of t_1 that maximizes the expected quality as per (4) from an a priori standpoint. Therefore, we define $Opt_{t_1}(j)$ as the value of t_1 between 0 and T that provides maximum expected quality for user j. Even if we have multiple users, but all have identical error probability p_j, it is immediate to derive the best policy for the operator, i.e., the optimal split of T into t_1 and t_2 so that the quality according to (3) is maximized. Thus, formally

$$Opt_{t_1}(j) = \arg\max \mathbb{E}[q_j] \tag{5}$$

subject to

$$\begin{aligned} (1-p_j) \cdot t_1 &\geq \theta_j \\ T &= t_1 + t_2 \end{aligned}$$

We can write the former constraint as $(1 - p_j) \cdot t_1 - \theta_j = g$, and the latter as $t_1 + t_2 - T = h$, resulting in $g \geq 0, h = 0$. Further, the objective function can be named as f. Thus, we will have following objective and constraints:

$$
\begin{aligned}
f(t_1, t_2, p_j) &= \Big(\alpha t_1 (1 - p_j) + \beta t_2 (1 - p_j) \Big) \\
&\quad \times \sum_{k=0}^{t_1 - \theta_j} \binom{t_1}{k} (1 - p_j)^{t_1 - k} (p_j)^k u[t_1 - k] \\
g(t_1, p_j, T) &= (1 - p_j) \cdot t_1 - \theta_j \\
h(t_1, p_j, T) &= t_1 + t_2 - T
\end{aligned}
$$
(6)
(7)
(8)

We can utilize the Lagrange Multiplier optimization method to find out the maxima, i.e., we need to take partial derivative of $f(t_1, t_2, p_j)$, $g(t_1, t_2, p_j)$ and $h(t_1, t_2, p_j)$ with respect to t_1, t_2 and p_j. Finally, using Lagrange Multipliers λ and μ for the constraints, we can find the additional number of TxOps required to correctly decode the video as:

$$
\eta_1 = \frac{\beta p_j T}{(\beta - \alpha)}
$$
(9)

The optimum t_1 would be equal to θ_j plus the additional number of transmit opportunities as in (9):

$$
Opt_{t_1} = \theta_j + \eta_1
$$
(10)

The analytical results of finding the optimum t_1 is presented in Fig. 1. It shows that, as the error probaility is increased, user will need more and more TxOps for BL packets to correctly decode. Also, the higher the threshold θ_j of BL packets, the higher would be requirement for Opt_{t_1}.
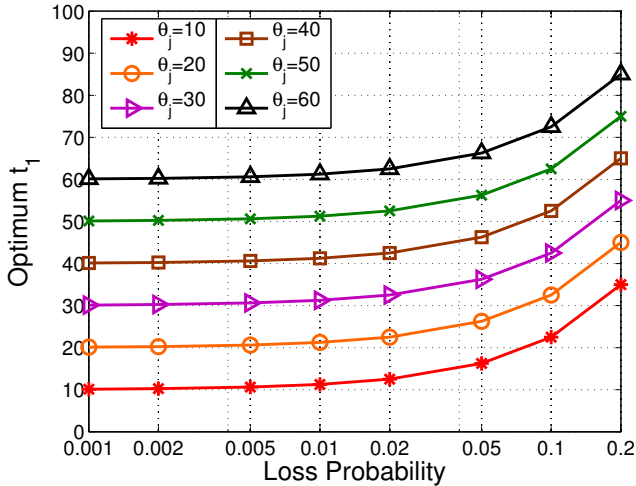

Fig. 1: Optimum t_1, varying θ_j, $T = 100$ and $\beta = 5$, $\alpha = 1$

2) Scenario B: For feedback analysis and simulation purposes, we define another scenario in which we have multiple users and each user can receive from 0 to n packets at each transmit opportunity, according to their loss probability.

We assume the number of received packets is Bernoulli distributed, related to the individual loss probability. Based on p_j, the loss probability, n probabilities are calculated, $\{\rho_0(j), \rho_1(j), \rho_2(j) \ldots \rho_n(j)\}$. Therefore, at each TxOp, depending on the loss probability of user j, it expects to receive ζ_j packets, where

$$
\zeta_j = \sum_{i=0}^{n} i \rho_i(j)
$$

The decoding probability P_j^d becomes

$$
P_j^d = \begin{cases} 0 & \text{if } \zeta_j \ t_1 < \theta_j \\ q_j & \text{if } \zeta_j \ t_1 \geq \theta_j \end{cases}
$$
(11)

where q_j is the expected quality for user j, defined as

$$
q_j = \alpha \zeta_j t_1 + \beta \zeta_j (T - t_1)
$$

Similar to (5), we can find out the optimum value of t_1 satisfying, $\zeta_j t_1 > \theta_j$ and $t_1 < T$, and which provides the maximum expected quality given P_j^d for user j:

$$
Opt_{t_1}(j) = \arg\max \mathbb{E}[q_j]
$$
(12)

Once each user has requested the individually optimal number of transmit opportunities to the central base station, the base station decides based on a given policy, the number of transmit opportunities for multicast. In this paper, we take into consideration the following two policies.

Average Policy: All users declare their preferred $Opt_{t_1}(j)$ and the central station calculates the average χ_{avg} of all these values, and then for χ_{avg} TXOPs base station transmits packets from the base layer, the rest are enhancement layer packets. Some users will be served with fewer base layer packets than their $Opt_{t_1}(j)$, and some of them can even be unable to decode the stream.

Maximum Policy: The users declare their preferred $Opt_{t_1}(j)$ and the central base station takes a conservative approach, selecting the maximum of the declared values, named χ_{max}. With this policy, all the users may reasonably expect to be satisfied in the end. Therefore, at the beginning of the transmission they are all *happy* as defined in II-A. However, this policy may decrease the QoE because more base layer and fewer enhanced layer packets will be received.

Note that both policies can use feedback, whenever present, to dynamically adapt their behavior. This means that, in the absence of feedback, the choice of t_1 is made by the central base station once at the beginning for the entire transmission, while it can be dynamically adapted if feedback is present. Actually, this reduces to the base station possibly transmitting fewer base layer packets than the initially computed t_1 whenever all the users notify they are either satisfied or hopeless. Therefore, the base station can decide earlier about EL transmission based on the feedback of users, thus some users may start enhancing video quality earlier than what was expected initially.

III. ANALYSIS AND SIMULATION RESULT

After the aforementioned initial computations, we simulated various policies for the multicast allocation and the presence of feedback, analyzing user satisfaction as well as actual perceived quality of the users. We also describe simulation setup and policies for multicast environment regarding the selection of t_1 for the entire multicast users group. The simulations are

performed using C++ and the results are averaged over a large number of runs. We set $\alpha = 1$ and $\beta = 1.5$ (if not specified otherwise). First of all, we analyze *scenario-A* to investigate the effect of varying threshold requirements of each user, the optimum number of transmit opportunities calculated by the system and the respective expected perceived quality of user. Then, we analyze *Scenario-B* with 10 users, and evaluate the performance regarding user satisfaction and perceived quality depending upon the availability of feedback.

A. Scenario A – single user

We take the value of threshold θ_j as a uniformly distributed random variable for each user j. Initially, we considered the case of a user by varying θ_j and β values. As the loss probability increases the quality decreases irrespect of β values, but when β is higher the quality is also higher.
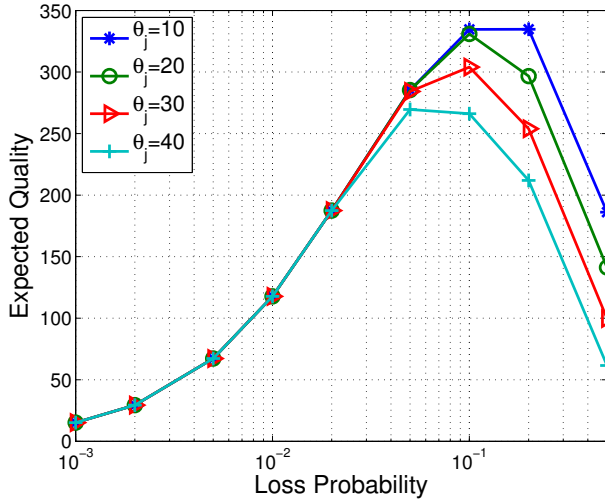
Fig. 2: Optimum quality with variable θ_j, $T = 100$ and $\beta = 5$

In Fig. 2 we show the expected quality corresponding to the optimal choice of t_1, i.e., Opt_{t_1}, versus p_j. The effect of varying values of θ_j for single user and keeping β constant as 5 is presented. Results show that the expected quality is higher when the loss probability is neither too high nor too low. Of course, the higher p_j the lower the achieved quality, but this effect seems to be limited, which is likely due to the fact that t_1 is chosen *optimally*, thereby achieving the most satisfactory allocation between base and enhancement layers. Also note that when $p \ll 1$, then quality of the single packet is not much of an issue; the lower the θ_j, the higher the number of EL packets, therefore, the higher the quality. The impact of the decoding threshold, for which the higher θ_j the lower the quality, seems more relevant.

B. Scenario B – multiple users

We simulate the system for different values of T. Each user has a random loss probability with uniform distribution between 0.01% and 1%. There are four possible combinations to evaluate, resulting from the choice between Average and Maximum policies, and including the presence or the absence

of feedback in multicast environment, the unicast case would be straight forward. The test cases are described as follows.

Average Policy vs Maximum Policy: In the Average Policy, the base station averages all $Opt_{t_1}(j)$ values and selects χ_{avg} for all the users. When feedback from each user is sent after each TXOP about the number of packets received and the perceived quality, it updates its value. For the Maximum policy instead, the maximum value amongst $Opt_{t_1}(j)$ for all users is selected as χ_{max} for multicast session. In fact, thanks to the presence of feedback, the $Opt_{t_1}(j)$ are updated at each transmission and the maximum value is selected again.

Feedback vs Non-feedback cases: In the absence of feedback, the base station cannot update the $Opt_{t_1}(j)$ at each TXOP; therefore, it will assume the same $Opt_{t_1}(j)$ values and the averaged value χ_{avg} which is calculated initially, if using average policy. If Maximum policy is used, then the base station selects the maximum $Opt_{t_1}(j)$ amongst all users, but does *not* update its value at each transmission. Therefore, even though the policy is meant to be conservative, the chosen value of χ_{max} may still dissatisfy some users because it is determined based on a priori expectations.

End-user Satisfaction Results— First, we present the results for user satisfaction, considering the effect presence or absence of feedback. Fig. 3 which represents user satisfaction for both cases with feedback (Fig. 3a) and without feedback (Fig. 3b) with T set to 10. Here, we can see that the expected happy users (labeled as "HAP") at the beginning of the transmission are closely related to actual number of satisfied user at the end of total transmit opportunities. Similarly, unhappy users (labeled as "UnHap") are approximately equal to number of Hopeless (labeled as HOP) users at the end. However, note that the curves for the feedback case are spiky, which is likely due to the dynamic change in policy when receiving feedback regarding packet losses.

Another choice of the total number of TXOPs is shown in Figs. 4(a) and 4(b). Here we can see that increasing the total number of transmit opportunities will eventually affect the total number of satisfied users and hopeless users at the end. User satisfaction result for maximum policy is presented in Fig. 4(b). This time, all of the users are expected to be happy initially before transmission, because the base station selected maximum number of transmit opportunities as multicast value. Note that the user satisfaction level is the same regardless of whether we have feedback or not.

QoE Results— Now, we present the results for actual quality perceived by users, which is averaged over all users for each TXOP and for all simulation runs. Fig. 5(a) presents the perceived quality, where we compare both maximum and average policies, as well as checking the difference about presence and absence of feedback. From Fig. 5(a), we can see that the presence of feedback increases the overall quality compared to the case without feedback. Note that the Maximum policy refers to satisfying all the users, therefore, the quality is eventually lower than the average policy, because the system will transmit BL packets for most of the TXOPs, and less EL packets as compared to the average policy. Similarly, Figs. 5(b) and 5(c) show the results for $T = 10$ and 7 TXOPs, respectively. From Figs. 5(a), 5(b) and 5(c) we can
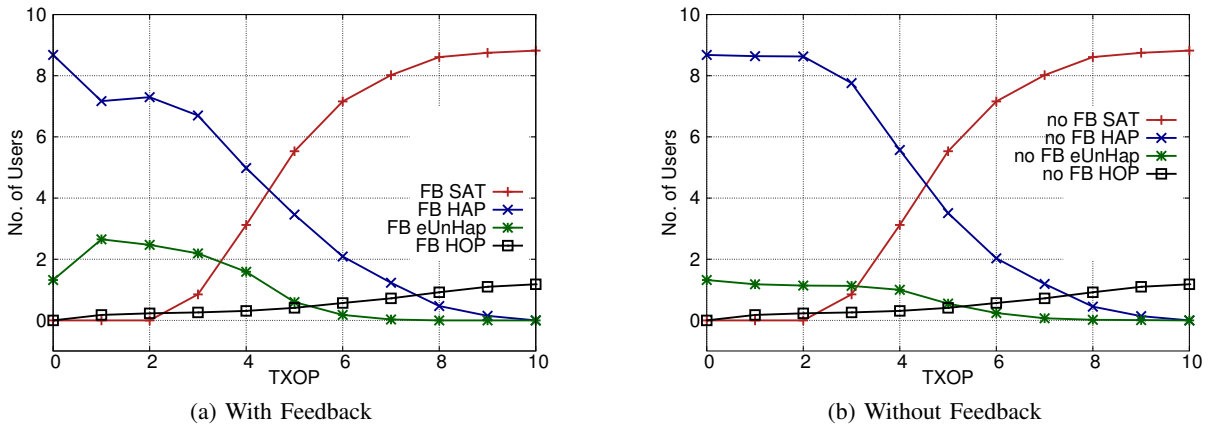
(a) With Feedback

(b) Without Feedback

Fig. 3: User Satisfaction Comparison for feedback/non feedback cases for *Average Policy* $T = 10$
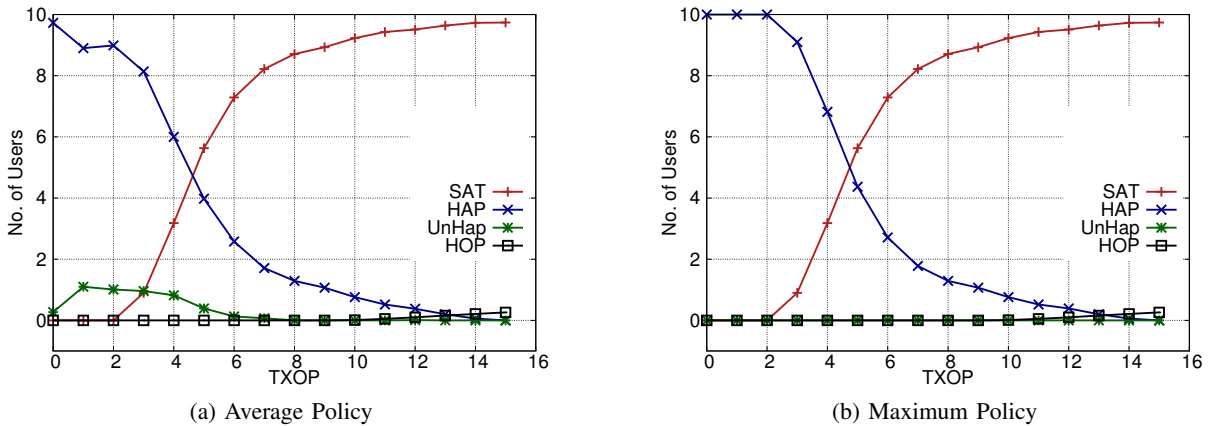

(a) Average Policy

(b) Maximum Policy

Fig. 4: User Satisfaction Comparison for Average and Maximum Policy using feedback case with *T=15*

see that the lower the total number of TXOPs, the smaller the difference between perceived qualities for policies; however, there is an increase in standard deviation of all runs if we decrease the number of allowed TXOPs. If we increase β factor to 10, the trend is preserved but the difference between the curves increases as illustrated in Figs. 6(a), 6(b) and 6(c) for $T = 15$, 10 and 7 respectively.

IV. CONCLUSIONS

Exploiting the layered video contents for enhancing the video QoE is a challenging task. We considered various network management policies in multicast environment with layered video contents. Each user has its own channel state, packet loss probability, and thus its own requirement about the number of base layer packets needed to decode the video. The strong point of the proposed solution is to give management options for policy selection in the network depending on operator requirements. We also provide a framework for evaluating the importance of feedback in multicast session for layered video contents delivery. We plan to apply this study to specific network scenarios, such as the Long Term Evolution of third generation cellular systems, by using specific network simulators. The idea is to further explore the channel state of users, and utilize efficient resource management techniques in addition to the network policies and feedback mechanisms in

such multicast layered video scenario. Another evolution will involve investigating the effect of the number of users in a multicast group and how their joining and leaving can effect the quality of all users. Further, it can also be investigated to compute the PSNR [14] of the video receiver and investigate the effect of feedback on PSNR and delays.

V. ACKNOWLEDGEMENT

REFERENCES

[1] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard, " *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103 -1120, 2007.

[2] S. Jelassi, G. Rubino, H. Melvin, H. Youssef, G. Pujolle, "Quality of experience of VoIP service: a survey of assessment approaches and open issues, " *IEEE Communications Surveys & Tutorials*, vol. 14, no. 2, pp. 491-513, 2012.

[3] P. Li, H. Zhang, B. Zhao, and S. Rangarajan, "Scalable video multicast in multi-carrier wireless data systems, *Proc. IEEE ICNP*, 2009.

[4] I. Yutaka, T. Yoshiro and T. Shuji, "Responsiveness of layered multicast and feedback control for video traffic in the Internet", *Proc. IEEE ICC*, vol. 2, pp. 846 - 852, 2000.

[5] L. Dounis, I. Politis, T. Dagiuklas, "On the comparison of real-time rate control schemes for H.264/AVC video streams over IP- based networks using network feedbacks", *Proc. IEEE ICC*, 2011.
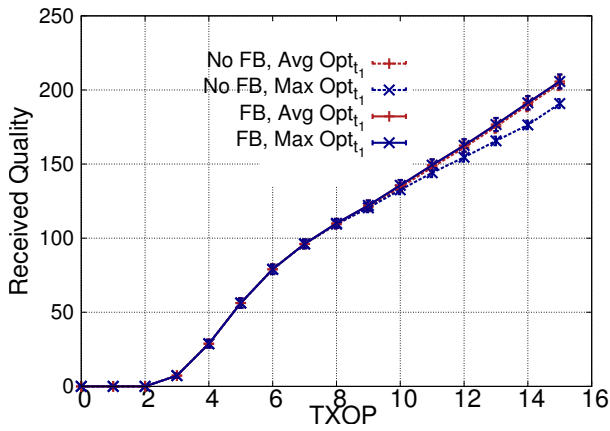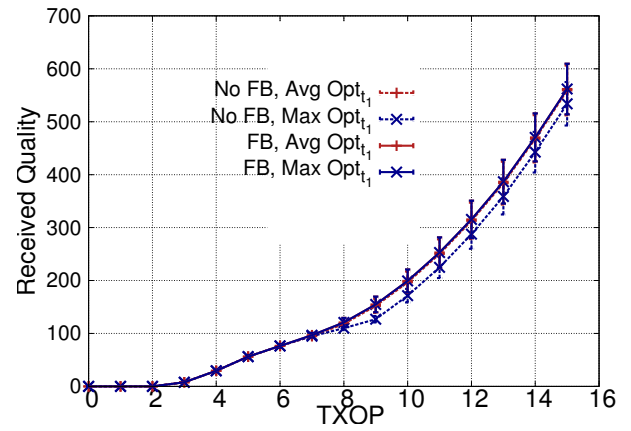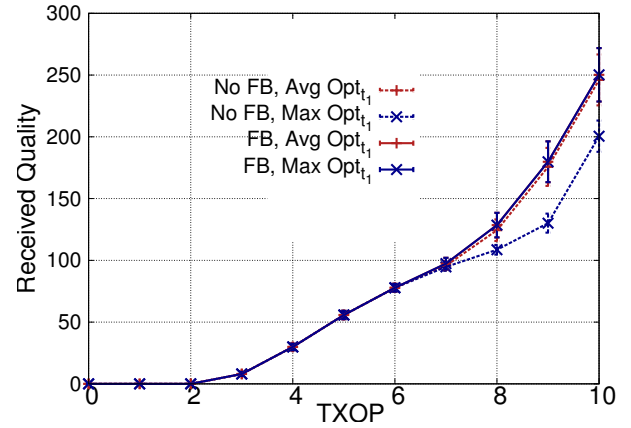
(a) $T = 15$, $\beta = 1.5$
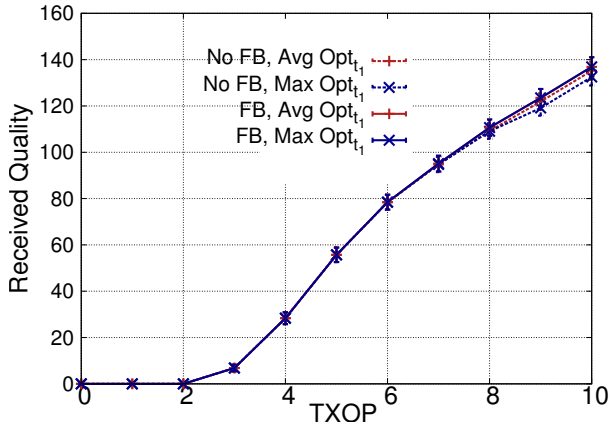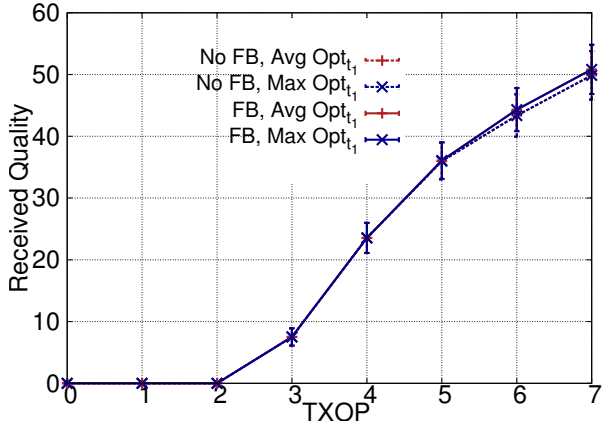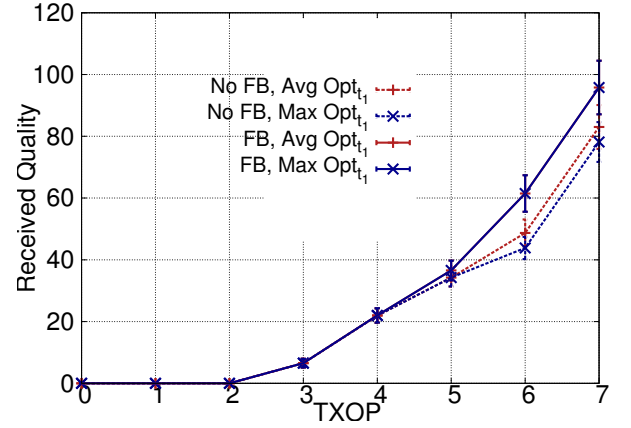

(a) $T = 15$, $\beta = 10$

(b) $T = 10$, $\beta = 1.5$

(b) $T = 10$, $\beta = 10$

(c) $T = 7$, $\beta = 1.5$

(c) $T = 7$, $\beta = 10$

Fig. 5: QoE comparison for all policies, with and without feedback with beta = 1.5

Fig. 6: QoE comparison for all policies, with and without feedback with $\beta = 10$

[6] H. Sohn, H. Yoo, W. D. Neve, C. S. Kim, and Y. M. Ro, "Full-reference video quality metric for fully scalable and mobile SVC content, " *IEEE Trans. on Broadcasting*, vol. 56, no. 3, 2010.

[7] J. Jin, Y. Ji, B. Zhao, Z. Hao and Z. Liu, "Error-resilient video multicast with layered hybrid FEC/ARQ over broadband wireless networks" *Proc. IEEE Globecom*, 2011.

[8] P. Goudarzi, M. R. N. Ranjbar and F. Ayatollahi "Differentiated QoE provisioning for video delivery over wireless networks" *Proc. International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, 2011.

[9] M. Loiacono, "Estimating an end-users received video signal at the transmitter using cross-layer feedback" *Proc. IEEE ICC*, 2011.

[10] V. J. Brett, L. Meejeong, and S. Tatsuya, "Feedback control mechanisms

for real-time multipoint video services", *IEEE JSAC*, vol. 15, no. 3, pp. 512-530, 1997.

[11] L. Zhengye, W. Zhenyu, L. Pei, L. Hang, W. Yao, "Layer bargaining: multicast layered video over wireless networks" *IEEE JSAC*, vol. 28 , no. 3, pp. 445 - 455, 2010.

[12] M. Freytes, C.A. Marques; "Error recovery for multicast conversational video over error-prone networks (with application to ad hoc WLANs)", *Proc. IEEE WCNC*, vol. 4, pp. 2245 - 2251, 2005.

[13] L. Badia, "On the impact of correlated arrivals and errors on ARQ delay terms," *IEEE Trans. on Commun.*, vol. 57, no. 2, pp. 334-338, 2009.

[14] S. Khan, S. Duhovnikov, E. Steinbach, W. Kellerer, "MOS-based multiuser multiapplication cross-layer optimization for mobile multimedia communication," *Adv. MultiMedia*, no. 1, pp. 1–11, 2007.