Video Quality Based Adaptive Wireless Video Streaming to Multiple Users

Maria G. Martini

WiMM, MINT Centre
Faculty of Computing, Information Systems and Mathematics
Kingston University London
KT12EE, UK
Email: m.martini @ kingston.ac.uk

Velio Tralli

CNIT, ENDIF, University of Ferrara
Via Saragat, 1
Ferrara, IT
Email: trv @ unife.it

***Abstract*—The paper proposes a cross-layer approach for joint closed-loop control of wireless video transmission to multiple users. The approach allows joint adaptation of the multiple users source rates by maximizing the worst served user video quality, taking into account the characteristics of the video sources, channel conditions and user requirements. The controlling task is split in two parts: resource allocation between users and cross-layer optimization for each user. A novel simple algorithm for dynamic resource allocation between users is proposed in conjunction with single user rate control.**

***Index Terms*—Rate-control, video coding, multiple users, congestion/capacity management**

I. Introduction

A critical problem in wireless video transmission is how to efficiently serve video streaming sessions over a multiple access wireless channel with shared communications resources. The problem can be afforded at different layers: dynamic rate control strategies optimized across the users can be considered at application layer in order to allocate resources according to users requirements and transmission conditions. Packet scheduling schemes across multiple users can be considered below in the protocol stack in order to adapt each stream to the available resources [1]. Content-aware scheduling can also be considered, as in [2].

Cross-layer design solutions should be investigated in order to optimize the global system. As an example, in [3] a cross layer design approach is considered with multiuser diversity which explores source and channel heterogeneity for different users; the focus is on downlink OFDMA scenario. Similarly, in [4] a distortion-based cross-layer optimization scheme is proposed.

In this paper, we specifically focus on controlling application layer parameters according to multi-layer information for downlink video transmission to multiple users.

A closed loop approach, based on the received video quality value, is considered in order to perform joint adaptation across users. Our approach realistically considers all the system layers also with their contribution to bandwidth occupancy.

Work performed when M. G. Martini was with CNIT and ENDIF, University of Ferrara, IT

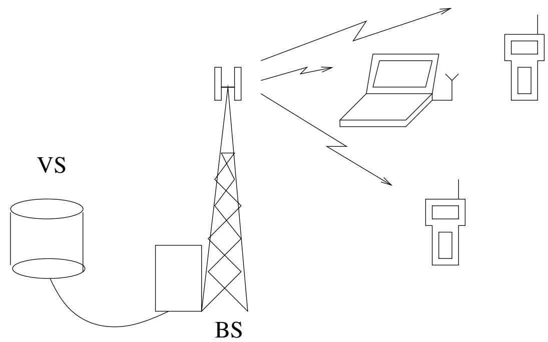

Fig. 1. Reference scenario

This paper is organized as follows. Section II describes the problem afforded and presents the proposed system controller and the relevant key algorithm. Section III reports simulation results. Finally, conclusions are drawn in Section IV.

II. Problem Statement and Proposed Adaptive Multi-User Controller

We consider a scenario as in Figure 1, where a video server (VS) is directly linked to the base station (BS), which transmits different video streams to multiple users. Different users are characterized by specific user requirements, different channel conditions and by the reception of video sequences with different characteristics in terms of video activity.

Such characteristics are taken into account in this paper through the parameters reported in Table I.

TABLE I
User characterization - parameters used to take into account the user characteristics

USER CHARACTERISTICS	symbol
User requirements (QoS)	w_i
Channel conditions	ρ_i
Video sequences activity	μ_i

The goal of our system controller is to perform dynamic rate control for each user's source and to adapt channel protection with the constraint of the total bandwidth B available for users.

The objective is to maximize the minimum weighted quality (user requirements are considered) among the users (or equivalently minimizing the maximum distortion), *i.e.*, the minmax criterium is adopted.

We consider control is performed dynamically, in subsequent time steps (*e.g.*, one second each) and we assume the controller behaves according to a Markov decision process, by selecting at each step one of the possible states, each characterized by a different rate for each user. The adaptive controller monitors the system and accordingly decides on actions for the subsequent time step. Since system conditions varies with time, the process is repeated at any time step

We split the cross-users / cross-layers controlling task in two steps: a multi-user controller takes the decision on resource allocation among users and a multi-layer single-user controller performs cross-layer adaptation.

Figure 2 reports a schematization of the controllers roles. The multi-user (MU) and cross-layer (CL) single-user (SU) controllers are shown together with the relevant main input and output parameters.

We assume we have N users receiving $L \leq N$ different video streams. We consider that the system controller has availability of:

- reduced channel state information (R-CSI), *i.e.*, non-instantaneous CSI, but CSI averaged over longer time steps, since at the application layer instantaneous information on the channel are not available;
- user requirements in terms of acceptable video quality;
- feed back information on the video quality in the previous controlling steps.

According to such information, a decision on the rates (quantization parameters, frame rate, group of pictures (GOP) rate) for the N users video streams is taken, by assigning in first instance the bandwidth to each user and subsequently performing single user adaptation with such bandwidth constraint. The time step between two subsequent decisions can be selected according to channel conditions and includes in any case the transmission of a whole GOP, i.e. at least an I frame. A model as the one in [5] is considered for the L video sources: the rate/distortion amd error-rate/distortion characteristics of the video source are modeled together with the sensitivity to errors of the different portions of the source bitstream.

Such controller can be seen as a generalization of the controller described in [6]. A cross-layer approach for a single link is considered there, where video quality driven adaptation is performed through a controlling unit collecting information from the different system blocks at different layers and providing controls to the source encoder, channel encoder and modulator.

A further external controlling unit allocating resources across users is then considered here.

A. Problem statement

We assume we have a constraint on the total bandwidth B. The goal of the controller is to determine at each time interval the bandwidth B_i allocated for each user such that the weighted quality of the worst served user is maximized. The single-user controller is then applied for single-user cross-layer adaptation.

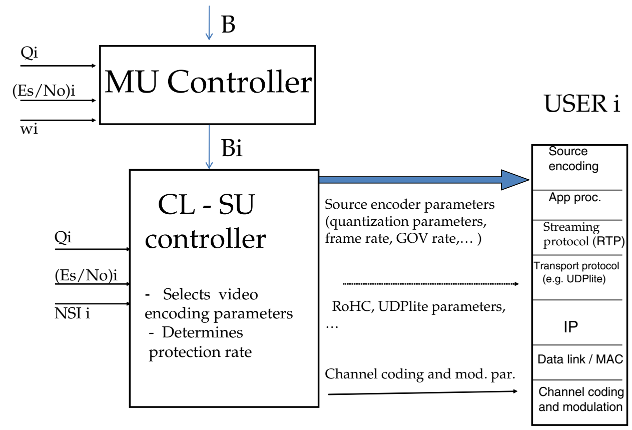

Fig. 2. System controllers

We assume each user i receives the video sequence with an average quality $Q_i(n)$ in the time interval n. In order to take into account user requirements we weight the quality value with a weight w_i. The parameter to optimize for each user is thus the weighted quality $WQ_i(n) = w_iQ_i(n)$ where w_i depends on user requirements.

In order to make the problem easily tractable, we do not consider any possible value for bandwidth B_i, but we consider a finite set of K possible B_i values, i.e.

$$B_i \epsilon \{\tilde{B}_1, \tilde{B}_2, ..., \tilde{B}_K\} \quad (1)$$

where

$$\tilde{B}_1 < \tilde{B}_2 < ... < \tilde{B}_K \quad (2)$$

For simplicity we will assume in the following that $B_{i,j}$ values are equally spaced, i.e. :

$$\tilde{B}_{j+1} - \tilde{B}_j = \tilde{B}_j - \tilde{B}_{j-1} = \Delta \quad (3)$$

for $j = 2, ..., K - 1$.

B. Proposed multi-user controller algorithm

At the first time step, feedback on received video quality is not present, thus bandwidth is allocated according to channel conditions of each user, the a-priori information about the video sequences, and user requirements. Alternative solutions consist in having the bandwidth equally shared between users or assigned according to capacity as initial condition.

At any subsequent time step n, the feedback on the received video quality is also considered and the bandwidth is allocated in order to maximize

$$\min_i \{w_1Q_1(n), w_2Q_2(n), ..., w_iQ_i(n), ..., w_NQ_N(n)\}, \quad (4)$$

where w_1,..., w_N are weights linked to user requirements and $Q_1(n), ..., Q_N(n)$ are the expected video quality values of the N users.

We provide in the following a more detailed description of the algorithm as considered in the subsequent simulation results.

1) FIRST CONTROLLER STEP (n=1): Only channel conditions ρ_i, weights w_i and the total bandwidth B are known. A criterion could be equally sharing the bandwidth and leaving to the dynamic adaptive controller the task to adapt to system conditions in subsequent steps.

An alternative criterion could be assigning the bandwidth according to different channel conditions with the goal of having equal channel capacities for different users. In this case, the goal is to have

$$C_1 = C_2 = ... = C_N \tag{5}$$

with the constraint

$$\sum_i B_i = B \tag{6}$$

Using the formula

$$C_i = B_i \log_2(1 + \rho_i * C_i/B_i) \tag{7}$$

where ρ_i is E_b/N_o for the i-th user, we have:

$$B_i = f^{-1}(\rho_i)/(\sum_k f^{-1}(\rho_k)) \tag{8}$$

where $f(x) = [2^{(1/x)} - 1]x$ and f^{-1} is the inverse function. This would require knowledge of channel conditions.

2) SECOND CONTROLLER STEP (n=2): Decision are taken according to quality values achieved in the previous time interval $Q_i(1)$ and the signal-to-noise ratio E_s/N_0 for the different users $\rho_i(1)$.

The algorithm performs the following steps:

- Select the user with the minimum value of $WQ_i(n)$
- If
$$\lambda_0 < WQ_{imax}(n) - WQ_{imin}(n) < \lambda_1 \tag{9}$$
then
$$B_{i_{min}}(n) = B_{i_{min}}(n-1) + \Delta \tag{10}$$
$$B_{i_{max}}(n) = B_{i_{max}}(n-1) - \Delta \tag{11}$$
i.e., increase its bandwidth one step and decrease one step the bandwidth of the user with the maximum value of $WQ_i(n)$.
We indicated as $B_{i_{min}}(n)$ and $B_{i_{max}}(n)$ the bandwidth of the users with the lowest and the highest weighted video quality, respectively. λ_0 and λ_1 are threshold values selected according to the dynamic range of video quality values.
- If
$$WQ_{imax}(n) - WQ_{imin}(n) > \lambda_1 \tag{12}$$
then
$$B_{i_{min}}(n) = B_{i_{min}}(n-1) + 2\Delta \tag{13}$$
$$B_{i_{max}}(n) = B_{i_{max}}(n-1) - 2\Delta \tag{14}$$
i.e., increase the bandwidth of the user with the minimum value of $WQ_i(n)$ two steps and decrease two steps the bandwidth of the user with the maximum value of $WQ_i(n)$.

The threshold can depend on the statistics of the $WQ_i(n)$ values and can be dynamically adapted. In order to avoid continuous changes of bandwidth, if the algorithm selects an increment / decrement of bandwidth for the same user it was decreased/increased in the previous step, the bandwidth for the user is kept constant and the relevant threshold is increased for the subsequent steps. Similarly, if the bandwidth assigned to each user is kept constant for a long time, thresholds are decreased.

3) SUBSEQUENT STEPS n:

- select the user with the minimum value of $WQ_i(n)$
- the algorithm can operate as in the second step and consider also quality values in the previous time interval(s).

After bandwidth has been allocated between users, each single-user cross-layer controller optimizes source parameters and channel parameters with the constraint of the assigned bandwidth, similarly as in [6].

III. Simulation Results

Example results in the case of wireless streaming to multiple users of MPEG-4 video are reported in the following. We assume a total available bit-rate on the channel of 1.6Mbps, AWGN channel, BPSK modulation and two users in the system. The two users experience different channel conditions, expressed in terms of E_s/N_0. We consider two different portions of the Foreman test video sequence are received by the two users. CIF format is considered for both the sequences.

Table II reports comparative simulation results, showing the video quality received by the two users both in the jointly adapted case and in the reference set-up. In the reference case, the bandwidth is equally shared among users and only single user joint source/channel adaptation is performed. Video quality is expressed in terms of average PSNR of the luminance component (Y-PSNR):

$$YPSNR = 20 \log(\frac{255}{RMSE_Y}). \tag{15}$$

where $RMSE_Y$ is the root mean square error calculated on the luminance component.

In this case we consider five controlling time intervals of one second each (5s in total) and we considered the average E_s/N_0 constant during the interval. In this case the algorithm adapts to channel conditions and to the varying characteristics of the video source.

Note that although the algorithm is exemplified here with the quality metric PSNR, any different objective quality metric can be considered. For example a metric better representing the perceived video quality (e.g. the structure similarity metric (SSIM) [7]) can be substituted to PSNR. Since in that scheme the metric is evaluated "on-the-fly" at the receiver side and then fed back towards the server at the transmission side, it is important that it is possible to evaluate the quality metric without reference to the original transmitted video frames. Although PSNR is considered here for simplicity and since it is more widely known in the scientific community, allowing an easily evaluation of results, a no-reference metric (see e.g. [8]) should be used here.

TABLE II
COMPARATIVE RESULTS: JOINT ADAPTATION VERSUS SINGLE USER ADAPTATION.

	E_s/N_0	Joint adaptation	Video quality (Y-PSNR)
User 1	0dB	yes	28.5
User 1	0dB	no	26.7
User 2	2dB	yes	30.9
User 2	2dB	no	31.5

From Table II we observe that the video quality of the "worst served" user is sensibly higher in the case where the proposed joint strategy is applied, and the quality of the "best served" user is still compliant with user requirements in such jointly adapted case.

We also compared two possible policies for the initialization of the algorithm: equal bandwidth for all users and bandwidth selected according to channel conditions with capacity considerations. Bandwidth may be also selected according to channel conditions and error-rate/distortion considerations using suitable models of the source, but since we decided to split the controller tasks in bandwidth allocation ("master" or multi-user (MU) controller) and cross-layer (CL) single-user adaptation (CL controller), the BER is not known by the MU controller allocating bandwidth, since the choice on channel protection is performed by the CL controller.

We consider for simplicity the case of two users sharing resources. We assume a total 2.2Mbps available for the users receiving the same video sequences as above. Results are reported in Table III.

TABLE III
COMPARATIVE RESULTS: INITIALIZATION POLICIES

	E_s/N_0	Policy	Video quality (Y-PSNR)
User 1	0dB	equal bandwidth	27.6
User 1	0dB	equal capacity	28.1
User 2	2dB	equal bandwidth	32.5
User 2	2dB	equal capacity	32

In the case shown, the equal capacity approach provides better results in terms of fairness between users. Since information about channel conditions are not always available in the initialization phase and since the system tends to improve the performance in subsequent steps, in many cases the equal bandwidth approach can be anyway chosen without a significant loss in performance.

IV. CONCLUSIONS

A cross-layer approach for joint closed-loop control of wireless video transmission to multiple users is presented in the paper. The approach allows to jointly adapt the multiple users source rates with the goal of maximizing the worst served user video quality, by taking into account the characteristics of the video sources, channel conditions and user requirements. The problem is separated in two parts: "fair" bandwidth allocation among users and cross-layer optimization for each user. A simple algorithm adaptive to channel conditions and to variable video characteristics has been proposed for resource allocation between users and such algorithm is used in link with the controller proposed in [6]. It has been shown that the proposed approach allows a remarkable improvement in the quality of the worst served user, by keeping the quality of other users to acceptable values. Such approach results thus in a better quality experience for the system users.

V. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the European Community for support of the 7th Framework project OPTIMIX.

REFERENCES

[1] D. Jurca and P. Frossard, “Video packet selection and scheduling for multipath streaming,” *IEEE Trans. on Multimedia*, vol. 9, no. 3, pp. 629–641, 2007.

[2] P. V. Pahalawatta, R. Berry, T. N. Pappas, and A. K. Katsaggelos, “Content-aware resource allocation and packet scheduling for video transmission over wireless networks,” *JSAC*, vol. 25, no. 4, pp. 749–759, 2007.

[3] G.-M. Su, Z. Han, and K. Liu, “Multiuser cross-layer resource allocation for video transmission over wireless networks,” *IEEE Network*, March/April 2006.

[4] L. Choi, W. Kellerer, and E. Steinbach, “On cross-layer design for streaming video delivery in multiuser wireless environments,” *Eurasip Journal on Wireless Communication and Networking*, vol. 2006, pp. 1–10, 2006.

[5] M. G. Martini and M. Chiani, “Rate-distortion models for unequal error protection for wireless video transmission,” in *Proc. IEEE Vehic. Tech. Conf. (VTC) 2004*, Milan, Italy, May 2004.

[6] M. G. Martini, M. Mazzotti, C. Lamy-Bergot, J. Huusko, and P. Amon, “Content adaptive network aware joint optimization of wireless video transmission,” *IEEE Communications Magazine*, vol. 45, no. 3, pp. 1–10, 2007.

[7] Z. Wang, L. Lu, and A. C. Bovik, “Video quality assessment based on structural distortion measurement,” *Signal processing: Image communication*, vol. 29, no. 1, Jan. 2004.

[8] Z. Wang, H. R. Sheikh, and A. C. Bovik, “No-reference perceptual quality assessment of JPEG compressed images,” in *Proceedings of IEEE International Conference on Image Processing*, Rochester, New York, September 2002.