OPTIMAL FORESIGHTED PACKET SCHEDULING AND RESOURCE ALLOCATION FOR MULTI-USER VIDEO TRANSMISSION IN 4G CELLULAR NETWORKS

Yuanzhang Xiao and Mihaela van der Schaar

Department of Electrical Engineering, UCLA. Email: yxiao@ee.ucla.edu

ABSTRACT

We study joint resource allocation and packet scheduling for multi-user video transmission in a 4G cellular network, where the base station (BS) allocates resources (i.e. bandwidth) among the users and each user schedules its video packets based on the allocated resources. Most existing works either propose *myopic* solutions for multi-user video transmission, in which the resource allocation and packet scheduling is designed to maximize the *short-term* video quality, or propose foresighted packet scheduling solutions for *single*-user video transmission which maximize the *long-term* video quality. In this work, we propose *foresighted* resource allocation and packet scheduling solutions for *multi*-user video transmission. Specifically, we develop a low-complexity algorithm in which the BS updates the prices of resources for each user and the users make individual packet scheduling decisions based on the prices. The algorithm can be implemented by the BS and the users in a decentralized manner, and converges to the optimal prices under which the users' optimal decisions maximize the long-term total video quality subject to per-user minimum video quality guarantees. Simulation results show 7 dB and 3 dB improvements in PSNR (Peak Signal-to-Noise Ratio) over myopic solutions and existing foresighted solutions, respectively.

***Index Terms*—** wireless video transmission, packet scheduling, resource allocation, multi-user communication

1. INTRODUCTION

Video applications, such as multimedia streaming, video chatting, and gaming, have become the major applications deployed over the current cellular networks. Such bandwidth-intensive and delay-sensitive applications require efficient network resource allocation among the users accessing the network, and efficient scheduling of each user's video packets based on its allocated resources.

Most existing works on multi-user video transmission propose *myopic* solutions [4]–[7], in which the resource allocation and packet scheduling is designed to maximize the *short-term* video quality (i.e. the video quality in a given time interval). However, due to time-varying channel conditions and dependency across video packets, current resource allocation and packet scheduling decisions have impact on the future system performance, which is not taken into consideration by the myopic solutions. Hence, the myopic solutions are inferior to *foresighted* solutions that maximize the *long-term* average video quality across different time intervals.

However, most works that propose foresighted solutions [8]–[12] study the packet scheduling of a *single* foresighted video user. In practical networks with multiple users, the solutions developed for a single user cannot be readily applied. A direct extension to the multi-user scenario may be to allocate a fixed amount of resources to each user *a priori*. However, how to optimally allocate resources is not addressed in the above works [8]–[12]. More importantly, such a static allocation of resources may be suboptimal compared to the solutions that dynamically allocate resources among multiple users.

In this paper, we propose a joint *foresighted* resource allocation and packet scheduling solution for *multi*-user video transmission. We study the uplink of a 4G cellular network[1], in which the base station (BS) allocates resources (i.e. bandwidth) to multiple video users who perform packet scheduling given the allocated resources. In our proposed solution, the BS does not directly allocate the resources; instead, it charges each user for resources by a unit "price"[2], based on which each user determines its own optimal packet scheduling and resource acquisition. This approach is desirable, because in this way the users can make optimal decisions in a decentralized fashion. To implement the proposed solution, we propose a low-complexity algorithm in which the BS updates the resource prices and the users make individual decisions based on the prices. We prove that the algorithm can converge to the optimal prices, under which the users' optimal decisions maximize the long-term total video quality in the network (subject to a minimum video quality guarantee for each user).

The rest of the paper is organized as follows. We discuss prior work in Section 2. In Section 3, we describe the system model and formulate the design problem. Then we propose our solution in Section 4. Simulation results in Section 5 demonstrate the performance improvement of the proposed solution. Finally, Section 6 concludes the paper.

2. RELATION TO PRIOR WORK

The existing works on wireless video transmission can be classified based on various criteria. In Table 1, we categorize the existing works [1]–[13] based on different criteria. Simply put, most works propose solutions either for multiple *myopic* video users [1]–[7] or for a *single* foresighted video user [8]–[12].

Very few works [13] propose solutions for *multiple foresighted* video users. Since this work [13] is most related to our work, we discuss the differences from [13] in detail. The challenge in foresighted multi-user video transmission is that the users' decisions are dynamic and are coupled through the resource (e.g. bandwidth or time) constraints. Hence, the design problem is much more

This work is supported by NSF grant CCF-1218136.

[1]The work can be easily extended to the downlink, and to wireless LANs (Local Area Networks) in which temporal transmission opportunities are allocated.

[2]Note that the "price" is a control signal, rather than the price for real monetary payment.

Table 1. Comparisons With Related Works.

	Traffic model	Users	Foresighted	Optimal
[1]–[3]	Flow-level	Multiple	No	No
[4]–[7]	Packet-level	Multiple	No	No
[8]–[12]	Packet-level	Single	Yes	No
[13]	Packet-level	Multiple	Yes	No
Proposed	Packet-level	Multiple	Yes	Yes

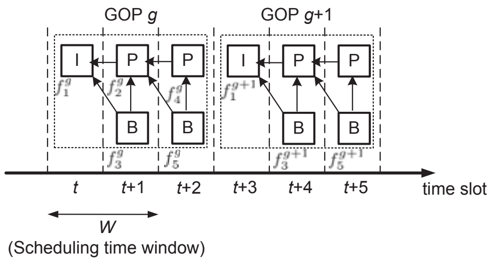

Fig. 1. Illustration of GOP (group of pictures), DU (data unit), and packet scheduling. Since the scheduling time window is $W = 2$, the contexts in different time slots are $C_t = \{f_1^g, f_2^g, f_3^g\}$, $C_{t+1} = \{f_2^g, f_3^g, f_4^g, f_5^g\}$, $C_{t+2} = \{f_4^g, f_5^g, f_1^{g+1}\}$, $C_{t+3} = \{f_1^{g+1}, f_2^{g+1}, f_3^{g+1}\}$, and so on.

complicated than that in myopic multi-user video transmission (where the decision is static) and that in foresighted single-user video transmission (where there is no coupling). The key idea to overcome this challenge is to use Lagrangian multipliers (interpreted as the prices of the resources) and add the product of the Lagrangian multipliers and the constraints onto the users' objective functions [13]. In this way, the users' decision problems are decoupled. However, contrary to standard dual decomposition in myopic video transmission, such a solution is in general suboptimal (i.e. there is a positive duality gap), because they use the *same* Lagrangian multiplier (i.e. a *uniform* price) under *all* the states [13]. In contrast, in our work, we allow the BS to use different prices in different states (i.e. different channel conditions). In this way, we can achieve better performance than MU-MDP with uniform price [13].

3. SYSTEM MODEL AND PROBLEM FORMULATION

A key feature of our model is that we use a packet-level model [13] to the characterize the video traffic (in terms of distortion impacts, delay deadlines, interdependency, etc.), which is distinct from widely-used flow-level models in other papers that characterize only the rate changes of video traffic [1]. Hence, we will first introduce the video traffic model for each user, and then describe the network model.

3.1. The Video Traffic Model For Each User

3.1.1. Characteristics of Video Data

In video transmissions, we usually encode the source data using a GOP (Group of Pictures) structure: the data is encoded into a series of GOPs, indexed by $g = 1, 2, \ldots$, where one GOP consists of N data units (DUs). Each DU $n = 1, \ldots, N$ in GOP g is characterized by its size, distortion impact, delay deadline, and dependency. The DUs in one GOP have different characteristics. However, different GOPs are "the same", in the sense that the nth DU in different GOPs have the same (statistical) characteristics. We denote the nth DU in GOP g by f_n^g, and list its characteristics as below.

- *Size*: The size of a DU is the number of packets (assumed to be of equal length as in [12]) in the DU. We denote the size of DU f_n^g by l_n^g (packets). The size l_n^g of DU f_n^g is an random variable following the probability mass function PMF_n. As in [13], we assume that the sizes of different DUs are independent random variables. Note that the distributions of the sizes of the nth DUs in different GOPs are the same.
- *Distortion impact*: Each DU f_n^g has a distortion impact of q_n^g per packet. The distortion impact measures how much distortion is added to the video if a packet is not received or cannot be successfully decoded at the receiver. The nth DUs in different GOPs have the same distortion impact per packet, namely $q_n^g = q_n^{g'}, \forall g, g'$.
- *Delay deadline*: The delay deadline d_n^g of DU f_n^g is the time before which the DU should be decoded. The relative differences between delay deadlines of DUs are fixed across different GOPs, namely $d_n^g - d_{n+1}^g = d_n^{g'} - d_{n+1}^{g'}, \forall g, g'$. Moreover, the relative differences between delay deadlines of the same DUs in adjacent GOPs are fixed as the length of the GOP, namely $d_n^g - d_n^{g+1} = d_m^g - d_m^{g+1}, \forall m, n$.
- *Dependency*: Since the DUs in one GOP are encoded using techniques such as motion estimation/compensation, the DUs may have complex dependency relationships. We say that DU f_n^g depends on DU f_m^g, if DU f_n^g is encoded based on the prediction from DU f_m^g (in other words, DU f_n^g cannot be decoded without decoding DU f_m^g first). We represent the dependency among DUs in one GOP by a directed acyclic graph (DAG) [8], where the vertices are DUs and an edge from DU f_m^g to DU f_n^g indicates that DU f_n^g depends on DU f_m^g. As in [13], we assume that if DU f_n^g depends on DU f_m^g, we have $d_n^g \geq d_m^g$ and $q_n^g \leq q_m^g$, namely DU f_m^g should be decoded before DU f_n^g and has a higher distortion impact than DU f_n^g. Note that there is no dependency between DUs in different GOPs.

3.1.2. Traffic State

We introduce the traffic state, which completely characterizes the state of the video traffic at each time slot $t = 1, 2, \ldots$. Note that in this paper, we take the perspective from the application layer, and denote t as the time slot in the application layer (which could be divided into smaller time slots in the physical layer). At time slot t, as in [8][12], we assume that the wireless user will only consider for transmission the DUs in the range of $[t, t + W - 1]$, where W is referred to as the scheduling time window (STW) and assumed to be determined a priori[3]. We further assume that the STW is chosen to satisfy the following condition: if DU f_j^g directly depends on DU $f_{j'}^g$, then $d_j^g - d_{j'}^g < W$. This assumption ensures that we can choose to transmit DUs f_j^g and $f_{j'}^g$ in the same time slot. Following the model in [12], at time slot t, we introduce the *context* to represent the set of DUs that are considered for transmission, i.e., whose delay deadlines are within the range of $[t, t + W - 1]$. We denote the context by $C_t = \left\{f_j^g | d_j^g \in [t, t + W - 1]\right\}$. Since the GOP structure is fixed, the context C_t is periodic with the period

[3]The STW can be determined based on the channel conditions experienced by the user in each time slot. For example, the STW can be set small when the channel conditions are poor, and large whenever the channel conditions are good.

of T (i.e. the length of a GOP), namely C_t and C_{t+T} have the same types of DUs and the same DAG between these DUs. Since the context represents the set of DUs to be transmitted, it implicitly represents the dependency among the DUs. The transition from context C_t to C_{t+1} is deterministic.

Given the current context C_t, we let $x_{f,t}$ denote the number of packets in the buffer associated with DU $f \in C_t$. We denote the buffer state of the DUs in C_t by $x_t = \{x_{f,t} | f \in C_t\}$. The traffic state at time slot t is then defined as (C_t, x_t), where the context C_t represents the types of DUs, the dependency among them, and the buffer state x_t represents the amount of packets remaining for transmission. Hence, the traffic state is able to capture heterogeneous multimedia traffic and is a super-set of existing well-known priority-buffer models.

3.1.3. Packet Scheduling

At each time slot t, the wireless user experiences a channel condition $h_t \in \mathcal{H}$, where $\mathcal{H}$ is the set of finite possible channel conditions and h_t is referred to as the channel state. Note that the channel condition is the quality of the channel perceived by the application layer, rather than the channel gain from transmitter to receiver measured in the physical layer. In this paper, we assume that the wireless channel is slow-fading (i.e. remains the same in one time slot) and that the channel condition h_t can be modeled as a finite-state Markov chain with transition probability $p_h(h'|h) \in [0, 1]$. We further define the state which the wireless user experiences at each time slot t as $s_t = (C_t, x_t, h_t)$, which includes the current context, buffer state and channel state. At time slot t, the wireless user decides how many packets should be transmitted from each DU in the current context. The decision is represented by $a_t(C_t, x_t, h_t) = \{a_{f,t} | f \in C_t, a_{f,t} \in [0, x_{f,t}]\}$, where $a_{f,t}$ represents the amount of packets transmitted from DU f. We consider the following payoff at each time slot t:

$$u(s_t, a_t) = \sum_{f\in C_t} q_f \cdot a_{f,t} - \rho \left(h_t, \sum_{f\in C_t} a_{f,t} \right) . \quad (1)$$

In the above payoff function, the first term represents the distortion reduction obtained by transmitting the data from the DUs in the current context. The second term represents the disutility of the energy consumption by transmitting the data. The energy consumption function $\rho(h, a)$ is assumed to be a convex function of a given the channel condition h.

3.2. The Network Model

In the previous subsection, we discuss the model for a single user. In this work, we consider a 4G cellular network, where there are I wireless video users transmitting to the BS indexed by 0. The users access the channels in a FDMA (frequency-division multiple access) manner. We normalize the total bandwidth to be 1, and will be divided and shared by the users. The BS knows the channel states of all the users (this information can be obtained by channel estimation from pilot signals sent from the users). We write the BS's state as $s_0 = (h_t^1, \ldots, h_t^I)$, where h_t^i is the channel state of user i. We will hereafter use superscript i to denote user i.

We assume that each user i uses adaptive modulation and coding (AMC) based on its channel condition. In other words, each user i chooses a data rate r_t^i under the channel state h_t^i. Note that the rate selection is done by the physical layer and is not a decision variable in our framework. Then as in [2][3], we have the following resource constraint:

$$\sum_{i=1}^{I} \frac{\sum_{f\in C_t} a_{f,t}}{r_t^i(h_t^i)} \le 1 \ . \quad (2)$$

Finally, note that although we consider cellular networks in this paper, the model can be readily applied to video transmission in wireless LANs operating under the IEEE 802.11e protocol, where the users access the channel in a TDMA (time-division multiple access) manner.

3.3. The Design Problem

Each user performs packet scheduling based on its state s_t. Hence, each user i's strategy can be defined as a mapping $\pi_i(s_t^i) \in A^i(s_t^i)$, where $A^i(s_t^i)$ is the set of actions available under state s_t^i. We allow the set of available actions to depend on the state, in order to capture the minimum video quality guarantee. For example, we may have a minimum distortion impact reduction requirement for each user at any time, which imposes constraints on the users' actions. The joint strategy profile is $\boldsymbol{\pi} = (\pi^1, \ldots, \pi^I)$.

The users aim to maximize their expected long-term payoff. The initial state $(s_0^0, \ldots, s_0^I)$ induce a probability distribution over the sequences of states, and hence a probability distribution over the sequences of total payoffs $u_i^0, u_i^1, \ldots$. Taking expectation with respect to the sequences of stage-game payoffs, we have user i's expected long-term payoff given the initial state as

$$U_i(\boldsymbol{\pi} | (s_0^0, \ldots, s_0^I)) = \mathbb{E} \left\{ (1-\delta) \sum_{t=0}^{\infty} \left(\delta^t \cdot u_t^i \right) \right\}, \quad (3)$$

where $\delta \in [0, 1)$ is the discount factor.

The design problem can be formulated as

$$\begin{aligned} \min\nolimits_{\boldsymbol{\pi}} \quad & \sum_{s_0^0,\ldots,s_0^I} \sum_{i=1}^{I} U_i(\boldsymbol{\pi} | (s_0^0, \ldots, s_0^I)) \\ s.t. \quad & \sum_{i=1}^{I} \frac{\|\pi^i(s^i)\|_1}{r_t^i(h_t^i)} \le 1, \ \forall s^0, \end{aligned} \quad (4)$$

where the constraint in the above design problem is an abstraction of the bandwidth constraint (2). We write the solution to the design problem as $\boldsymbol{\pi}^\star$ and the optimal value of the design problem as $U^\star$.

4. OPTIMAL FORESIGHTED VIDEO TRANSMISSION

In this section, we derive the optimal foresighted video transmission.

4.1. Decomposition of The Users' Decision Problems

Contrary to the designer, each user aims to minimize its own long-term total payoff $U_i(\boldsymbol{\pi} | (s_0^0, \ldots, s_0^I))$. In other words, each user i solves the following problem:

$$\pi^i = \arg\max_{\tilde{\pi}^i} U_i(\tilde{\pi}^i, \boldsymbol{\pi}^{-i} | (s_0^0, \ldots, s_0^I)).$$

Assuming that the user knows all the information, the optimal solution to the above problem should satisfy the following:

$$\begin{aligned} V(s^i) = \quad & \max_{a^i \in A^i(s^i)} \quad (1-\delta) u^i(s^i, a^i) + \delta \cdot \sum_{s^{i\prime}} \rho^i(s^{i\prime} | s^i, a^i) V(s^{i\prime}) \\ & s.t. \quad \sum_{i=1}^{I} \frac{\|a^i\|_1}{r_t^i(h_t^i)} \le 1. \end{aligned} \quad (5)$$

Note that the above equations would be the Bellman equations, if the user knew all the information such as the other users' actions $\boldsymbol{a}_{-i}$,

Table 2. Distributed algorithm to compute the optimal decentralized video transmission strategy.

Input: Each user's performance loss tolerance ϵ^i

Initialization: Set $k=0$, $\bar{a}_i(0)=0, \forall i$, $\lambda_i(0)=0, \forall i$.

repeat

Each user i solves

$$\tilde{V}^{i,\lambda^{i,(k)}(s^0)}(s^i) = \max_{a^i \in A^i(s^i)} (1-\delta)\left[u^i(s^i,a^i) - \lambda^{i,(k)}(s^0)\cdot a^i\right] +\delta \cdot \sum_{s^{i\prime}} \left[\rho^i(s^{i\prime}|s^i,a^i)\tilde{V}^{i,\lambda^{i,(k)}(s^0)}(s^{i\prime})\right]$$

Each user i submits its bandwidth request $\pi^{i,\lambda^{i,(k)}(s^0)}(s^i)$

The BS updates $\bar{a}^i(k+1) = \bar{a}^i(k) + \pi_i^{\lambda^{i,(k)}(s^0)}(s^i)$ for all $i=1,\ldots,I$

The BS updates the prices:

$$\lambda^{i,(k+1)}(s^0) = \lambda^{i,(k)}(s^0) + \frac{1}{k+1}\cdot\frac{1}{r_t^i(h_t^i)}$$

until $\|\tilde{V}^{i,\lambda^{i,(k+1)}(s_0)} - \tilde{V}^{i,\lambda^{i,(k)}(s_0)}\| \le \epsilon^i$

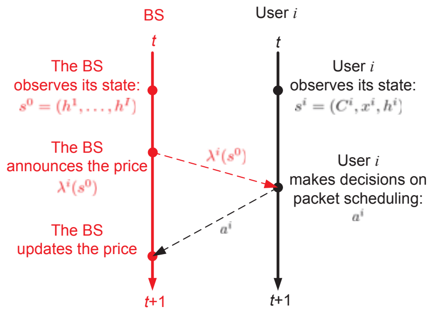

Fig. 2. Illustration of the interaction between the BS and user i (i.e. their decision making and information exchange) in one period.

and the BS's state s_0. However, such information is never known to the user. Hence, we need to separate the influence of the other entities from each user's decision problem.

One way to decouple the interaction among the users is to penalize the constraint onto the objective function. Denote the Lagrangian multiplier (i.e. the "price") associated with the constraint under state s^0 as $\lambda^i(s^0)$, we can rewrite user i's decision problem as

$$\tilde{V}^{\lambda^i(s^0)}(s^i) = \max_{a^i\in A^i(s^i)}(1-\delta)\left[u_i(s^i,a^i) - \lambda^i(s^0)\cdot a^i\right] +\delta\cdot\sum_{s_i'}\left[\rho^i(s^{i\prime}|s^i,a^i)\tilde{V}^{\lambda^i(s^0)}(s^{i\prime})\right] . \quad (6)$$

Clearly, we can see from the above equations that given the price λ^i, each user can make decisions based only on its local information.

The remaining question is how to determine the optimal prices, such that when each user reacts based on its price, the resulting strategy profile maximizes the social welfare.

4.2. The Optimal Decentralized Video Transmission Strategy

The optimal prices depend on the BS's state, which is known to the BS only. Hence, we propose a distributed algorithm used by the BS to iteratively update the prices and by the users to update their optimal strategies. The algorithm will converge to the optimal prices and the optimal strategy profile that achieves the minimum total system payoff $U^\star$. The algorithm is described in Table 2.

Theorem 1 *The algorithm in Table 2 converges to the optimal strategy profile, namely*

$$\lim_{k\to 0}\left|\sum_{s_0^0,\ldots,s_0^I}\sum_{i=1}^I U_i(\boldsymbol{\pi}^{\lambda^{(k)}}|(s_0^0,\ldots,s_0^I)) - U^\star\right| = 0 .$$

Table 3. The proposed framework reduces to existing frameworks.

Framework	Simplification of the proposed framework
Myopic [4]–[7]	$\delta=0$
MU-MDP [13]	The price $\boldsymbol{\lambda}(s_0)=\boldsymbol{\lambda}$ under all the state s_0

Table 4. Comparisons of PSNR under different energy consumptions.

Energy consumption (Joule)	0.08	0.10	0.15
Myopic [4]–[7]	(30, 24, 24) dB	(40, 35, 25) dB	(45, 45, 34) dB
MU-MDP [13]	(34, 28, 27) dB	(44, 40, 28) dB	(50, 48, 38) dB
Proposed	(37, 32, 30) dB	(46, 43, 32) dB	(54, 51, 42) dB

Proof 1 *See [14, Appendix A].*

We illustrate the algorithm (i.e. the BS's and users' updates and their information exchange) in one period in Fig. 2. We can see that the amount of information exchange at each iteration is small ($O(I)$), compared to the amount of information unavailable to each entity ($\prod_{j\neq i}|S_i|$ states plus the strategies $\boldsymbol{\pi}_{-i}$). In other words, the algorithm enables the entities to exchange a small amount ($O(I)$) of information and reach the optimal video transmission strategy that achieves the same performance as when each entity knows the complete information about the system.

5. SIMULATION RESULTS

We consider a wireless network with three users streaming video sequences "Foreman" (CIF resolution, 30 Hz), "COstguard" (CIF resolution, 30 Hz), "Mobile" (CIF resolution, 30 Hz), respectively. The energy consumption function is set as $\rho(h,a) = \sigma^2(2^a-1)/|h|^2$, where $|h|^2/\sigma^2 = 1.4$ (≈ 1.5 dB) [12]. We set the discount factor as $\delta = 0.95$. Note that different video sequences have heterogeneous distortion impact, delay deadlines, and dependency among packets. Hence, the simulation will demonstrate that our proposed solution can accommodate heterogeneous video streams.

We compare against the myopic solution [4]–[7] and the foresighted solution with uniform price [13] in terms of the peak signal-to-noise ratio (PSNR) and energy consumption. The results are listed in Table 4. We can see that the proposed solution can achieve on average a 7 dB PSNR improvement for all the users, compared to the myopic solution, and a 3 dB PSNR improvement for all the users, compared to the MU-MDP solution with uniform price.

6. CONCLUSION

We propose the optimal foresighted resource allocation and packet scheduling solution for multi-user video transmission. The proposed solution achieves the optimal long-term total video quality subject to each user's minimum video quality guarantee, by dynamically allocating bandwidth among the users and dynamically scheduling the users' packets while taking into account the dependency among the packets and the time-varying channel conditions. We develop a low-complexity algorithm that can be implemented by the BS and the users in a decentralized manner and can converge to the proposed optimal solution. Simulations show that our proposed solution can achieve significant improvements in PSNR of up to 7 dB compared to myopic solutions and of up to 3 dB compared to state-of-the-art foresighted solutions.

7. REFERENCES

[1] A. Reibman, and A. Berger, “Traffic descriptors for VBR video teleconferencing over ATM networks,” *IEEE/ACM Trans. Netw.*, vol. 3, no. 3, pp. 329–339, Jun. 1995.

[2] P. Li, H. Zhang, B. Zhao, and S. Rangarajan, “Scalable video multicast with joint layer resource allocation in broadband wireless networks,” in *Proc. IEEE International Conference on Network Protocols (ICNP)*, pp. 295–304, Oct. 2010.

[3] Y.-J. Yu, P.-C. Hsiu, and A.-C. Pang, “Energy-efficient video multicast in 4G wireless systems,” *IEEE Trans. Mobile Comput.*, vol. 11, no. 10, pp. 1508–1522, Oct. 2012.

[4] M. van der Schaar and D. Turaga, “Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission,” *IEEE Trans. Multimedia*, vol. 9, no. 1, pp. 185C197, Jan. 2007.

[5] G-M. Su, Z. Han, M. Wu, and K.J.R. Liu, “Joint uplink and downlink optimization for real-time multiuser video streaming over WLANs,” *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 2, pp. 280–294, Aug. 2007.

[6] X. Zhang and Q. Du, “Cross-layer modeling for QoS-driven multimedia multicast/broadcast over fading channels in mobile wireless networks,” *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 62–70, Aug. 2007.

[7] J. Huang, Z. Li, M. Chiang, and A.K. Katsaggelos, “Joint source adaptation and resource allocation for multi-user wireless video streaming,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 5, 582–595, May 2008.

[8] P. Chou and Z. Miao, “Rate-distortion optimized streaming of packetized media,” *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 390–404, Apr. 2006.

[9] S. Khan, M. Sgroi, Y. Peng, E. Steinbach, W. Kellerer, “Application-driven cross layer optimization for video streaming over wireless networks,” IEEE Commun. Mag. , pp. 122–130, Jan. 2006.

[10] C. De Vleeschouwer and P. Frossard, “Dependent packet transmission policies in rate-distortion optimized media scheduling,” *IEEE Trans. Multimedia*, vol. 9, no. 6, pp. 1241–1258, Oct. 2007.

[11] H. Wang and A. Ortega, “Rate-distortion optimized scheduling for redundant video representations,” *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 225–240, Feb. 2009.

[12] F. Fu and M. van der Schaar, “Structural solutions for dynamic scheduling in wireless multimedia transmission,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 5, pp. 727–739, May 2012.

[13] F. Fu and M. van der Schaar, “A systematic framework for dynamically optimizing multi-user video transmission,” *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 308–320, Apr. 2010.

[14] Y. Xiao and M. van der Schaar, “Optimal Foresighted Multi-User Wireless Video,” Available at: http://arxiv.org/pdf/1311.4227.pdf