Journal of Zhejiang University SCIENCE A
ISSN 1009-3095 (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Joint rate control and scheduling for wireless uplink video streaming*

HUANG Jian-wei[1], LI Zhu[2], CHIANG Mung[1], KATSAGGELOS Aggelos K.[3]

([1]Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544, USA)

([2]Multimedia Research Lab, Motorola Labs, Schaumburg, Illinois 60196, USA)

([3]Department of Electrical Engineering & Computer Science, Northwestern University, Evanston, Illinois 60208, USA)

E-mail: jianweih@princeton.edu; zhu.li@motorola.com; chiangm@princeton.edu; aggk@ece.northwestern.edu

Received Dec. 15, 2005; revision accepted Feb.20, 2006

Abstract: We solve the problem of uplink video streaming in CDMA cellular networks by jointly designing the rate control and scheduling algorithms. In the pricing-based distributed rate control algorithm, the base station announces a price for the per unit average rate it can support, and the mobile devices choose their desired average transmission rates by balancing their video quality and cost of transmission. Each mobile device then determines the specific video frames to transmit by a video summarization process. In the time-division-multiplexing (TDM) scheduling algorithm, the base station collects the information on frames to be transmitted from all devices within the current time window, sorts them in increasing order of deadlines, and schedules the transmissions in a TDM fashion. This joint algorithm takes advantage of the multi-user content diversity, and maximizes the network total utility (i.e., minimize the network total distortion), while satisfying the delivery deadline constraints. Simulations showed that the proposed algorithm significantly outperforms the constant rate provision algorithm.

Key words: Video streaming, Pricing, Uplink communications, CDMA, Cross-layer design, Video summarization
doi:10.1631/jzus.2006.A0801 **Document code:** A **CLC number:** TN919.8

INTRODUCTION

Video streaming is becoming one of the major driving forces of next generation wireless networks. For the currently deployed cellular networks, the practical data rates are not enough to support full rate, high quality video applications. As a result, many research efforts have been devoted to adapting video content to reconcile the conflict between the high demand of video quality and the limited wireless communication resources among users. A large body of literature utilizes the cross-layer approach, which jointly designs the video coding in the application layer and the resource allocation in lower layers (Zhang *et al.*, 2005; van der Schaar *et al.*, 2003; Zheng, 2003; Yoo *et al.*, 2004; Zhao *et al.*, 2002, and the references therein).

To ensure the quality of real-time video streaming, smart video coding and adaptation techniques need to be performed at the application layer to meet the stringent resource constraints at lower layers. For very low bit rate channels, simply coding the video sequences at high quantization distortion levels is unpleasant to viewers. A better solution is to perform content-aware video coding, via summarization, which selects a subset of video frames that best represent the sequence, and encodes them at a higher quality. Various summarization techniques have recently been reported (e.g., Liu and Kuo, 2005; Li *et al.*, 2005a; 2005b). The actual adaptation can be achieved through video summarization/transcoding and/or bit bit stream extraction from scalable video.

This paper develops, analyzes and simulates a new joint rate control and scheduling algorithm for uplink video streaming in CDMA cellular networks. The rate control part of the algorithm relies on adaptive content-aware video summarization, and utilizes a pricing-based approach to distribute the computational burden across the network (i.e., base station)

* Project (Nos. CNS-0427677 and CCF-0448012) supported by the National Science Foundation of USA

and individual mobile users, and fully utilizes the multi-user diversity to minimize users' total distortion. The scheduling part of the algorithm lets users transmit the summary frames in a time-division-multiplexing (TDM) fashion, which avoids excessive mutual interferences, achieve higher rate compared with a simultaneous equal rate transmission scheme.

The pricing approach has been successfully used to efficiently allocate communication resources among elastic data applications in wireless networks (e.g., Saraydar *et al.*, 2002; Liu *et al.*, 2004; Lee *et al.*, 2002; Huang *et al.*, 2004). We previously showed that a pricing-based approach combined with adaptive video summarization techniques can greatly improve the performance of multi-user wireless downlink video transmissions (Li *et al.*, 2006). This paper further extends the pricing framework to the uplink streaming case, which is more complex due to the interference limited nature of the communication channel.

Scheduling of multiple video users in CDMA networks has been considered in previous work (e.g., Xu *et al.*, 2004; Wang *et al.*, 2005; Kam *et al.*, 2001). Our contribution here is to show that a TDM-based transmission among video users leads to better performance compared with the scheme where video users also transmit simultaneously, by avoiding large mutual interference and exploiting multi-user content diversity.

In previous work on multi-user uplink video streaming at very low bit rate (Li *et al.*, 2005c), we try to control the admissible rate profile by iteratively adjusting peak rates and average rates among video users. The drawback is that convergence is not guaranteed in general. The algorithm proposed here, however, has theoretically provable and practically very fast convergence.

The paper is organized as follows. We first introduce the system model in Section 2, and then describe the resource allocation algorithm in Sections 3 and 4. Simulation results are shown in Section 5, and conclusion is given in Section 6.

SYSTEM MODEL

The uplink capacity for the wideband CDMA system is interference limited (Tse and Viswanath, 2005). In the case of mixed voice and streaming video transmissions, the objective is to provide the best possible Quality of Service (QoS) to the video users, without interrupting the transmissions of voice users. This can be translated into a total received power constraint of the video users at the base station.

Consider a single CDMA cell with a set of M voice users and N video users. Assuming perfect power control, each voice user adjusts its transmission power to achieve the common received P_{voice} and the SINR target γ_{voice} at the base station, thus a common transmission rate R_{voice}. A sufficient condition to achieve this is that the total received power at the base station from video users, P_{video}, should satisfy the following condition:

$$\frac{G_{\text{voice}}W}{R_{\text{voice}}}\frac{P_{\text{voice}}}{n_0W+P_{\text{video}}+(M-1)P_{\text{voice}}}\geq\gamma_{\text{voice}}, \tag{1}$$

where W is the channel bandwidth, G_{voice} is a constant that depends on voice users' common modulation scheme (e.g., G_{voice}=1 for BPSK), and n_0 is the background noise density (including both thermal noise and inter-cell interference). Here each voice user treats the received power from other users as Gaussian noise. We can solve for the maximum value of P_{video} that satisfies Eq.(1), denoted as $P_{\max}$:

$$P_{\max}=\left(\frac{G_{\text{voice}}W}{R_{\text{voice}}\gamma_{\text{voice}}}+1-M\right)P_{\text{voice}}-n_0W. \tag{2}$$

This is the maximum total received power constraint from all video users in the cell. For simplicity, we assume that each voice user could achieve the same received power P_{voice}. For more general treatment of different peak power constraints, see (Sampath *et al.*, 1995).

In practice, the value of $P_{\max}$ could change over time to reflect the load change of voice. Here we only focus on resource allocation during a single time segment [0, T], which corresponds to the time window during which users perform one round of video coding and summarization. The typical value of T is around 3 s, which is sufficiently short to ensure that the voice load and $P_{\max}$ do not change much.

Given the constraint of $P_{\max}$, we want to maximize the total network utility defined on the received video qualities (The qualities of the voice transmissions are guaranteed at fixed levels, and thus are not considered in the optimization). The optimization variables are $\{P_j(t)\}_{1\le j\le N}$, where $P_j(t)$ is the received power function of video user j at any time $t\in[0, T]$. The Network Utility Maximization problem is

$$\begin{aligned}&\max_{\{P_j(t)\ge 0\}_{1\le j\le N}} \sum_{j=1}^{N} U_j\left\{\int_0^T R_j[\boldsymbol{P}(t)]\mathrm{d}t\right\},\\ &\text{s.t.}\sum_{j=1}^{N} P_j(t)\le P_{\max}, \forall t\in[0,T],\end{aligned} \tag{3}$$

here $R_j[\boldsymbol{P}(t)]$ is the rate achieved by user j at time $t\in[0, T]$, and is a function of the received power allocation of all users, $\boldsymbol{P}(t)=[P_1(t),\ldots,P_N(t)]$. U_j is the utility function of user j and is increasing and strictly concave in the average transmission rate achieved during time $[0, T]$ (An example of the utility function will be given in Section 3). Here we assume that each video user can achieve peak received power equal to $P_{\max}$ (For a user that cannot achieve $P_{\max}$, it might be better for it to handoff to another base station with a better channel).

Due to the Variable Bit Rate (VBR) nature of the video streaming contents, it is likely that the optimal solution of Eq.(3) consists of time-varying functions. Together with the fact that the utility functions U_j usually do not have analytical forms, finding the optimal solutions of Eq.(3) is quite difficult. To simplify the analysis, we consider a TDM-based transmission scheme, where the video frames of different users are transmitted one at a time without overlapping.

There are two major motivations for using the TDM-based transmission. First, TDM transmission could avoid interferences among video users, who typically need much larger received powers at the base station (compared with voice users) to achieve enough transmission rate. As a result, multiple video users transmitting simultaneously would lead to excess interferences. A more detailed discussion of this point can be found in (Kumaran and Qian, 2003), where the authors show that to maximize the weighted rate in an uplink CDMA transmission, it is optimal to schedule "strong" users to transmit one-at-a-time, and "weak" users to transmit simultaneously in larger groups. Here a "strong" user has a high peak received power, and a "weak" user has a low peak received power. In our context, video users are considered "strong" and voice users are considered "weak".

Second, the VBR nature of the video sources makes it a huge waste of resource to provide a constant rate pipe for each video user without taking the actual video content into consideration. Thus if users transmit simultaneously, the transmission rates need to be adjusted on a frame by frame basis to reflect the VBR nature of traffic. This would require the calculation of the optimal transmission rate for each frame considering its unique size and deadline constraint, and jointly find the optimal power levels that achieve such rates. This is computationally expensive and generate a lot of signaling overhead. On the other hand, the TDM-based transmission simplifies the scheduling problem by letting video frames to be transmitted at the highest rate possible, one by one, in the order of deadlines. This is quite easy to implement since the frame sizes and deadlines for a given time segment are typically available at the beginning of the frame due to buffering. Furthermore, the multi-user content diversity is naturally exploited under TDM-based scheduling.

PRICING-BASED RATE CONTROL

In this section, we aim at allocating averaged transmission rates among users to maximum total utility. The delivery deadlines of video frames will be considered in the scheduling algorithm described in the next section. Based on the discussion above, we can rewrite Eq.(3) in the following form

$$\max_{\{t_j\ge 0\}_{1\le j\le N}} \sum_{j=1}^{N}\tilde{U}_j(t_j),\ \text{s.t.}\sum_{j=1}^{N} t_j\le T, \tag{4}$$

where

$$\tilde{U}_j(t_j)=U_j(R_{\mathrm{TDM}}t_j). \tag{5}$$

Here R_{TDM} is the transmission achieved by letting only one user transmit with a received power $P_{\max}$. A user j's average transmission rate during time $[0, T]$ is determined by the product of R_{TDM} and the active transmission time allocated to it, denoted by t_j.

Eq.(4) could be solved by the standard dual decomposition technique. First relax the total transmission time constraint by associating it with a dual price, λ. Then solving Eq.(4) is equivalent to maximizing the following Lagrangian, i.e.,

$$\max_{t\geq 0} J(\boldsymbol{t},\lambda)=\sum_{j=1}^{N}\tilde{U}_j(t_j)-\lambda\left(\sum_{j=1}^{N}t_j-T\right) \quad (6)$$

for some optimal non-negative value λ. Here $\boldsymbol{t}=[t_1,\ldots,t_N]$.

Let us first consider how to solve Eq.(6) for a fixed λ. We observe that Eq.(6) can be decomposed into several subproblems, one for each user. In other words, Eq.(6) can be solved by letting each user j find an optimal value of $t_j(\lambda)$ such that

$$t_j(\lambda)=\arg\max_{t_j}[\tilde{U}_j(t_j)-\lambda t_j]. \quad (7)$$

Finding $t_j(\lambda)$ in Eq.(7) requires more information on how the utility function is defined. Although any formulation with increasing and concave utility functions would work, here we focus on the case where the utility function is defined on the video quality in terms of summarization distortion level. Let $V=\{f_0, f_1,\ldots, f_{n-1}\}$ be a sequence of video frames encoded at pre-determined PSNR quality levels from some video clip. The distortions introduced by the encoding process are given but not considered in the optimization in this paper. Further consider a summary sequence of m frames $S=\left\{f_{l_0}, f_{l_1},\ldots, f_{l_{m-1}}\right\}$ of the sequence V, where $m\leq n$. The sequence S is then transmitted through the wireless channel. After receiving S (assuming error free), the receiver reconstructs the original sequence V as $V_S'=\{f_0', f_1',\ldots, f_{n-1}'\}$ by substituting the missing frames with the most recent frame that is in the summary S. The video summary quality, which is defined as the average distortion caused by the missing frames, is given as:

$$D(S)=\frac{1}{n}\sum_{k=0}^{n-1}d(f_k, f_k'), \quad (8)$$

where $d(f_k, f_k')$ is the distance between the kth original frame in V and the corresponding constructed frame in V'. Therefore, the optimization problem in Eq.(7) can be transformed into the problem of summarization with a penalty in the transmission time,

$$S_j^*(\lambda)=\arg\min_{S_j} D(S_j)+\lambda t_j(S_j), \quad (9)$$

where the total transmission time t_j is determined by the resulting video summary sequence and the transmission rate. Eq.(9) can be solved with a Dynamic Programming approach at the video sources. More details can be found in (Li *et al.*, 2005b).

It is known from information theory (Cover and Thomas, 1991) that a variety of practical signal sources have strictly convex rate-distortion function. This is also the case empirically for the rate-summarization distortion function, $D(S_j)$, as shown in (Li *et al.*, 2005a). As a result, the utility U_j is an increasing and strictly concave function of the rate, so is $\tilde{U}_j(t_j)$ in transmission time t_j. An example of the rate-distortion trade-off curve is plotted in Fig.1, where the video sequence corresponds to frames 150~239 from the "Foreman" sequence. The distortion per frame is the total distortion divided by 90 (i.e., the total number of frames in the time segment).

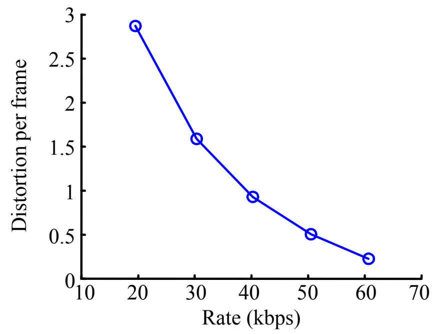

Fig.1 Rate-distortion trade-off curve of frames 150~239 from the "Foreman" sequence in a 3 s time segment

Once $S_j^*(\lambda)$ is found, the corresponding transmission time $t_j[S_j(\lambda)]$ can be computed assuming a rate of R_{TDM}. Each user j sends the value of $t_j[S_j(\lambda)]$ to the base station, who wants to solve the following dual problem

$$\max_{\lambda\geq 0} J[\boldsymbol{t}(\lambda),\lambda]. \quad (10)$$

This can be solved by a subgradient search on λ, where the subgradient is determined by the degree of violation of the time constraint. To be specific, the price λ could be updated according to Eq.(11):

$$\lambda^{i+1}=\max\left\{0,\lambda^i+\alpha^i\left\{\sum_{j=1}^{N}t_j[S_j(\lambda^i)]-T\right\}\right\},\qquad(11)$$

where α^i is the step-size at price iteration i. In Eq.(11), if the requested total transmission time is larger than T, the price increases in the next iteration, and the price decreases if the requested total time is below T.

Proposition 1 If the step-sizes satisfy $\lim_{i\to\infty}\alpha^i=0$ and $\sum_i \alpha^i\to\infty$, then updating Eq.(7) to Eq.(11) converge to the optimal solution of Eq.(4).

The proposition can be shown by techniques similar to those of (Srikant, 2004), thus is omitted here.

In the rate control algorithm, we vertically decompose the NUM problem in Eq.(3) into video summarization Eq.(6) that can be solved in the application layer, and rate control Eq.(10) that can be solved in the transport layer. Furthermore, horizontal decomposition is used such that Eq.(6) can be solved in a distributed fashion by letting each video user solve a subproblem Eq.(7). More on the framework of "Layering as Optimization Decomposition" can be found in (Chiang *et al*., 2006).

Let us denote the price that corresponds to the optimal solution of Eq.(4) as λ^*, then the resulting $[S_j(\lambda^*)]$ or $[t_j(\lambda^*)]$ are just indications of the resource consumption levels for delivering certain level of utility for each user. The actual transmission schedule of individual frames is computed by the scheduling algorithm in the next section.

TDM-BASED GREEDY SCHEDULING

To ensure satisfactory reception of the video streaming application, each video summary frame has to be delivered to the receiver before a certain deadline. The pricing-based rate control algorithm leads to "optimal" averaged rate allocation without considering the deadlines of video summary frames. The GREEDY scheduling algorithm in this section aimes at meeting all the deadline requirements.

Since here we only consider the uplink transmission, and would like to upper-bound the delay of the frames received at the base station. The deadline constrained downlink video streaming considered in (Li *et al*., 2006) can be jointly used with the approach in this paper to provide bounded end-to-end delay guarantees.

The delivery deadline of each summary frame is determined by three components: the initial delay of all frames F_{ini}, its position delay $F_{j,\text{position}}$, and the total length of the current time segment T.

The initial delay F_{ini} is typically determined by the sizes of users' first frames. Since each user's first frame in a time segment is intra-coded with a typical large size, it has to be included in its summary frames. For these first frames, their deadlines equal the values of initial delay, since the values of $F_{j,\text{position}}$ for these frames are zero. For other summary frames, $F_{j,\text{position}}$ is determined as follows. If frame j is the 45th frame in its original source sequence V, and the sampling rate of the frames is 30 Hz, then $F_{j,\text{position}}=(45-1)/30$ s. The -1 here is due to the fact that we consider the position of the first frame as time 0. Finally, the entire summary frames need to be delivered within the current time segment, i.e., not interfering with the transmission of frames in the next time segment. Thus the deadline of packet j is

$$T^j=\min(F_{\text{ini}}+F_{j,\text{position}},T).\qquad(12)$$

Assume all video summary frames from different users which are to be transmitted in time segment $[0, T]$ are available at time 0 by proper buffering, the users communicate the frame sizes and deadlines to the base station. The GREEDY algorithm works as follows. The base station first sorts the frames in increasing order of the delivery deadline and then lets the frames to be transmitted one-at-a-time, at constant rate R_{TDM}, such that the received SINR of each frame meets the target value

$$\frac{G_{\text{video}}W}{R_{\text{TDM}}}\frac{P_{\max}}{n_0W+MP_{\text{voice}}}\geq\gamma_{\text{video}},\qquad(13)$$

here G_{video} is a constant that depends on the modulation scheme used for the video traffic (e.g., $G_{\text{video}}=2$ for QPSK). Assume the jth frame in the sorted sequence (containing summary frames from all users) has a frame size B^j, it then takes time equal to B^j/R_{TDM} to transmit this frame.

GREEDY algorithm is simple and is optimal among all TDM-based algorithms:

Proposition 2 If any TDM-based scheduling algorithm can meet the deadlines of all video frames, the GREEDY scheduling algorithm also can.

Proposition 2 can be proved as follows: pick any TDM-based scheduling algorithm where all deadlines are met and one or more packets are transmitted out of the deadline order. Then by rearranging the corresponding out of order packets by the deadline as in the GREEDY algorithm, all the deadline constraints are still satisfied.

It can be shown that if no TDM-based scheduling algorithm can meet all deadline constraints, then the GREEDY algorithm incurs the least deadline violation. To formally state the result, let us define

$$\Delta^{\Pi}=\max_{j}(T_{\Pi}^{j}-T^{j}). \tag{14}$$

as the maximum delay violation under TDM-based scheduling policy Π, where T_{Π}^{j} denotes the actual delivery time of the jth packet under TDM-based algorithm Π. If $\Delta^{\Pi}\leq 0$, then all deadline constraints are met. We have

Proposition 3 Among all TDM-based scheduling algorithms, the GREEDY algorithm yields the smallest value of Δ^{Π}.

In fact, Proposition 2 is just a special case of Proposition 3, and the same proof technique can be generalized to prove the latter.

In the case of $\Delta^{\mathrm{GREEDY}}>0$, the base station must increase price so that video users request less rate. One way of adjusting price is by the following

$$\lambda^{i+1}=\max\{0,\lambda^{i}+\beta\max[\Delta^{\mathrm{GREEDY}}(\lambda^{i}),0]\}, \tag{15}$$

where β is a small step-size. In other words, the price is increased until the resulting frame sequences are schedulable (i.e., all deadline constraints can be met under GREEDY algorithm). There is a trade-off between the value of β and convergence speed. If β is large, then the schedulability can be achieved by a few adjustments; however, a significant portion of the time segment [0, T] might be wasted. If β is small, then it takes longer to achieve schedulability, and the resource utilization will be higher. In either case, since users' average transmission rates decrease with λ, the price adjustment process in Eq.(15) always converges.

SIMULATION RESULTS

We choose four video clips with different content activity levels, similar as in (Li *et al.*, 2006). Clips 1 and 2 are frames 150~239 and frames 240~329 from the "Foreman" sequence, and clips 3 and 4 are frames 50~139 and 140~229 from the "Mother-daughter" sequence, respectively. There are 90 frames within each video clip at a sampling frequency of 30 Hz, which corresponds to a time segment of T=3 s. Besides the GREEDY scheduling algorithm, we also simulate a simultaneous transmission scheme with equal constant rate (SIMCONST), where all four video users are allowed to transmit simultaneously with equal rates. In other words, the received power from each of the video user is the same as that at the base station in the SIMTRANS scheme, and no scheduling across users is needed due to simultaneous transmission.

Table 1 lists the simulation parameters that are kept constant throughout this section. These values are just chosen for illustration purpose instead of being from any particular standard, and our proposed algorithm is applicable to any version of CDMA networks.

Table 1 Simulation parameters

Entity	Symbol	Value
Bandwidth	W	1.228 MHz
Noise density	n_0	8.3×10^{-7} mW/Hz
Voice target SINR	γ_{voice}	6 dB
Voice modulation		BPSK
Voice received power	P_{voice}	1 mW
Voice spreading gain	G_{voice}	128
Voice transmission rate	R_{voice}	9.6 kbps
Video target SINR	γ_{video}	6 dB
Video modulation		QPSK

We first compare the video users' total achievable rate under GREEDY and SIMCONST algorithms for different voice user load. Under GREEDY, we plot the maximum rate achieved by allowing only one user to transmit. Under SIMCONST, we plot the total rate achieved by all four users. Fig.2 shows that video users' total achievable rate decreases with the number of voice users, and becomes zero when there are more than 31 voice users in the system. In other words, the system's ability to support video users depends heavily on the current voice load in the cell. It is also clear that GREEDY algorithm always outperforms SIMCONST, at a rate gap of more than 200% when the voice load is low. In that case, the mutual interference among video users becomes the bottleneck in achieving high rate under SIMCONST.

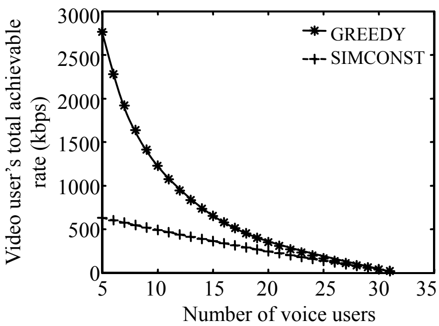

Fig.2 Total achievable rate comparison between GREEDY and SIMCONST

As we argued in Section 2, a TDM-based transmission scheme among "strong" uplink users not only achieves higher rate but also efficiently exploits multi-user content diversity. To illustrate this point, let us consider a cell with 28 voice users, where the GREEDY algorithm offers almost the same total rate as the SIMCONST. For the rest of the simulation, we will further assume the following Table 2 parameters.

Table 2 Further simulation parameters

Entity	Symbol	Value
Number of voice users	M	28
Maximum video received power	$P_{\max}$	4 mW
Video rate under GREEDY	R_{TDM}	84.7 kbps
Video rate per user under SIMCONST	R_{CR}	19.2 kbps

First consider the pricing-based rate control algorithm. Based on the assumption of TDM scheduling, pricing on transmission time is equivalent to pricing on the achievable rate. We start from an initial price λ=0.1, and use diminishing step-sizes α^i=0.05/i that satisfy the conditions in Proposition 1. The iteration stops when the total transmission time of four video users achieves more than 99% of the time segment length (i.e., 3 s). Fig.3 shows the convergence of price in 6 iterations, with a final optimal price λ^*= 8.674×10^{-3}.

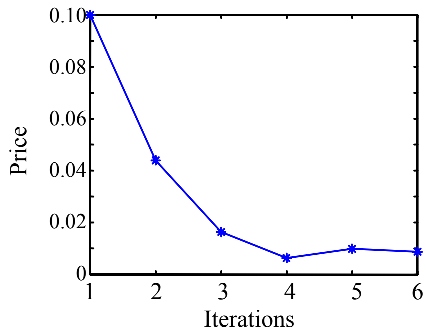

Fig.3 Pricing iteration

Fig.4 shows how the summarized distortion per frame of each individual user decreases (or increases) as the price decreases (or increases). Depending on the video contents that determine the specific rate-distortion functions, users experience different levels of distortions under the same price. Among the four users, user 2 experiences the largest distortion due to the large temporal variations of its contents. Users 3 and 4 achieve similar distortions that are much smaller than those of users 1 and 2, due to the small time variations in the contents.

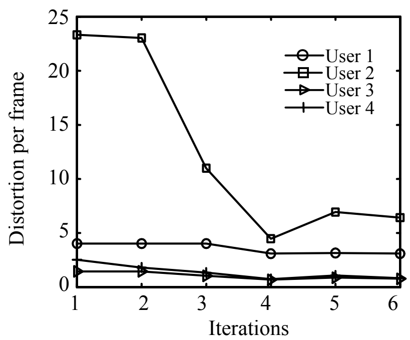

Fig.4 Distortion iteration

Fig.5 shows how the total transmission time of the summarized packets changes during the iteration. If we relax the convergence criterion from 99% to 80% (i.e., the price converges when it first enters the region bounded by the two dashed lines in Fig.5), then the convergence time can be shortened by half. This reflects a trade-off between the computational complexity and resource utilization efficiency.

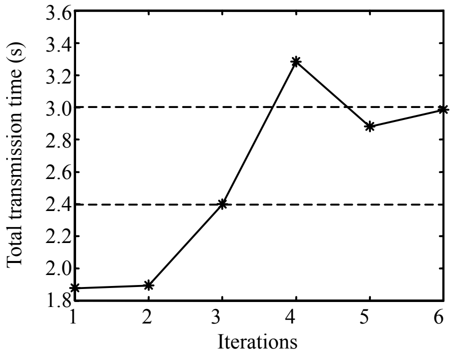

Fig.5 Transmission time iteration

The pricing algorithm will introduce additional delay into the system due to its iterative nature, and this delay can be leveraged by having sufficient buffer at the transmitter of each mobile device. Assume that each video transmitter has enough buffers to hold all the packets over two time segments. During the transmissions of the frames for the first time segment, the price iteration and summarizations for the frames to be transmitted in the second time window can take place in a distributed fashion among the base station and the mobile users. The iterative pricing process will not lead to any additional reception delay at the base station if the convergence can be achieved within one time segment length.

If there are not enough buffers available at the transmitters' side, then a system designer needs to carefully choose the iteration parameters to trade-off the convergence speed and performance. In general, the convergence speed of the pricing algorithm depends on the video contents, the initial price, the choice of step-sizes and the stopping criterion. Except the video contents that cannot be adjusted by the system, all other factors can be continuously tuned based on experiences to offer the best trade-off between convergence and performance. Typically the requirement of faster convergence inevitably leads to degraded performance since the resource (transmission time) may not be fully unutilized (e.g., reducing the stopping criterion from 99% to 80% as explained before). This trade-off becomes more important as the number of video users increases. We want to emphasize that the convergence time does not increase linearly with the number of users, since the most time consuming operation is the summarization process, which is performed by users in parallel.

The resulting video summary distortions based on the optimal price λ^* are plotted in Fig.6. The vertical arrows indicate video summary frame locations in the sequence. Notice that the distortion is zero at summarized frame locations, since the received frames are exactly the same as the original frames before summarization. The optimal price gives a good trade-off between total transmitting time and total video summary distortions. Clips 1 and 2 are coded at an average *PSNR* of 27.8 dB, and clips 3 and 4 at 31.0 dB. The resulting average bit rates for 4 clips are 18.26, 47.79, 8.04 and 10.22 kbps, respectively.

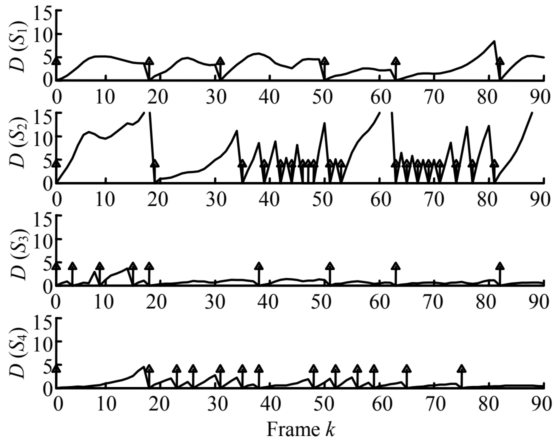

Fig.6 Resulting video summary distortion at the optimal pricing

Given the summarization results, the GREEDY algorithm performs scheduling based on sorted packet deadlines. The corresponding received power functions of users are plotted in Fig.7, and the corresponding delivery deadlines are plotted in Fig.8. Under an initial delay of 30 frames (1 s), the GREEDY algorithm successfully transmits all packets within 3 s and meets all deadline requirements.

As we mentioned in Section 4, if the current summary frames cannot be scheduled (i.e., deadline violation occurs), then the base station needs to increase the price and let the users re-compute the summarizations. However, in all the simulations that we perform, the summarization result from the pricing-based rate control is always schedulable. This is due to the fact that by taking advantage of the multi-user content diversity, the deadline requirements of the summary frames are typically spread out through the time segment, thus is relatively easy to satisfy. This implies that as long as there are enough content differences among the video users, the two stages of the algorithm can be operated separately in practice. This avoids unnecessary iterations among the two stages and ensures fast convergence of the algorithm.

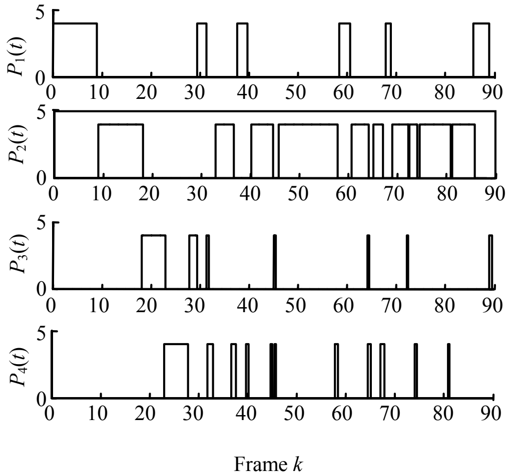

Fig.7 Video users' received powers at the base station under GREEDY scheduling algorithm (power is measured in mW)

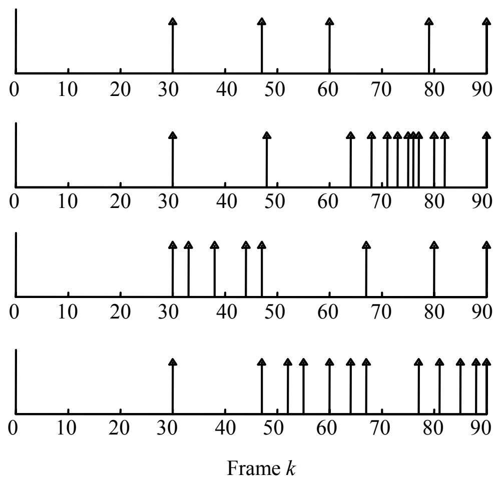

Fig.8 Frame delivery deadlines under GREEDY scheduling algorithm

For comparison purpose, we also simulate the SIMCONST scheme, where all four video users are allowed to transmit simultaneously. The base station can only guarantee a constant rate of 19.2 kbps for each user. The users perform summarizations based on the guaranteed rates, so that all the summary frames can be transmitted within their individual deadline constraints. The resulting summary distortions are shown in Fig.9. The averaged distortions per frame for all users are 2.85, 31.43, 0.059 and 0.068, respectively, with a total distortion per frame of 34.4. As comparison, the averaged distortions per frame for all users achieved under pricing-based rate control are 3.09, 6.42, 0.76 and 0.81, respectively (Fig.4), with a total distortion per frame of 11.09. Under SIMCONST, user 2 suffers a much larger distortion due to its busy contents. As a result, the total distortion per frame increases more than 200% in the pricing-based approach to SIMCONST.

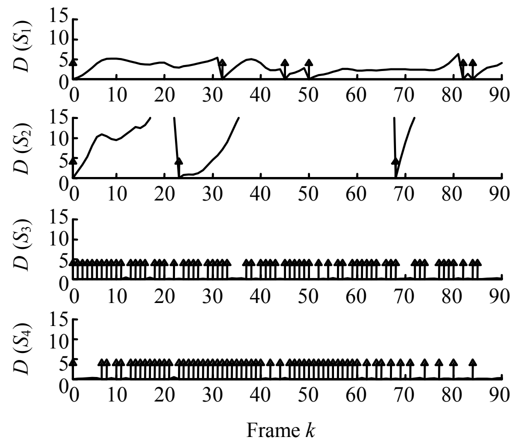

Fig.9 Resulting summary distortion under SIMCONST scheme

CONCLUSION AND REMARKS

In this paper, we consider a cross-layer design approach for uplink video streaming in a single CDMA cell. The application layer (video coding), transport layer (rate control) and data link layer (scheduling) are jointly optimized. We propose a two-stage resource allocation scheme, which includes a pricing-based rate control algorithm and TDM-based GREEDY scheduling algorithm. In the rate control algorithm, the base station announces a price for the rate, and the mobile video devices independently choose their average rate by performing optimal content-aware video summarization based on both the price and their utility functions. In other words, the operations in the application layer (video coding and summarization) and transport layer (rate control) are coupled only through a single price signal. In the data link layer, the base station performs TDM-based GREEDY scheduling based on the deadlines of the summary frames, and adjusts the price if it is not schedulable. Simulation results showed that it significantly improves the network utility compared with the constant rate transmission scheme.

References

Chiang, M., Low, S.H., Calderbank, A.R., Doyle, J.C., 2006. Layering as Optimization Decomposition. Proceedings of IEEE.

Cover, T., Thomas, J., 1991. Elements of Information Theory. Wiley and Sons, New York.

Huang, J., Berry, R., Honig, M.L., 2004. Auction-based Spectrum Sharing. Proceedings of Allerton Conference of Communication, Computing and Control.

Kam, A.C., Minn, T., Siu, K.Y., 2001. Supporting rate guarantee and fair access for bursty data traffic in W-CDMA. *IEEE Journal on Selected Areas in Communications*, **19**(11):2121-2130. [doi:10.1109/49.963799]

Kumaran, K., Qian, L., 2003. Uplink Scheduling in CDMA Packet-data Systems. Proceedings of IEEE INFOCOM.

Lee, J.W., Mazumdar, R.R., Shroff, N.B., 2002. Downlink Power Allocation for Multi-class CDMA Wireless Networks. Proceedings of IEEE INFOCOM.

Li, Z., Katsaggelos, A.K., Schuster, G., Gandhi, B., 2005a. Rate-distortion optimal video summary generation. *IEEE Trans. on Image Processing*, **14**(10):1550-1560. [doi:10.1109/TIP.2005.854477]

Li, Z., Zhai, F., Katsaggelos, A.K., 2005b. Video Summarization for Energy Efficient Wireless Streaming. Proceedings of SPIE Visual Communication and Image Processing (VCIP).

Li, Z., Cheng, Q., Katsaggelos, A.K., Ishtiaq, F., 2005c. Video Summarization and Transmission Adaptation for Very Low Bit Rate Multiuser Wireless Uplink Video Communication. Proceedings of IEEE International Workshop on Multimedia Signal Processing (MMSP).

Li, Z., Huang, J., Katsaggelos, A.K., 2006. Pricing Based Collaborative Multi-user Video Streaming over Power Constrained Wireless Downlink. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP).

Liu, S., Kuo, C.J., 2005. Joint temporal-spatial bit allocation for video coding with dependency. *IEEE Trans. on Circuits & System for Video Tech.*, **15**(1):15-26. [doi:10.1109/TCSVT.2004.839996]

Liu, P., Zhang, P., Jordan, S., Honig, M.L., 2004. Single-cell forward link power allocation using pricing in wireless networks. *IEEE Transactions on Wireless Communications*, **3**(2):533-543. [doi:10.1109/TWC.2004.825363]

Sampath, A., Kumar, P.S., Holtzman, J.M., 1995. Power Control and Resource Management for a Multimedia CDMA Wireless System. Proceedings of IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC).

Saraydar, C., Mandayam, N.B., Goodman, D., 2002. Efficient power control via pricing in wireless data networks. *IEEE Trans. on Wireless Communications*, **50**(2):291-303. [doi:10.1109/26.983324]

Srikant, R., 2004. The Mathematics of Internet Congestion Control. Birkhauser, Boston.

Tse, D., Viswanath, P., 2005. Fundamentals of Wireless Communication. Cambridge University Press.

van der Schaar, M., Krishnamachari, S., Choi, S., Xu, X., 2003. Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs. *IEEE Journal on Selected Areas in Communications*, **21**(10): 1752-1763. [doi:10.1109/JSAC.2003.815231]

Wang, L., Aghvami, A., Chambers, W., 2005. Design issues of uplink media access control (MAC) protocols for multimedia traffic over DS-CDMA systems. *IEEE Trans. on Multimedia*, **7**(3):551-562. [doi:10.1109/TMM.2005.846786]

Xu, L., Shen, X., Mark, J., 2004. Dynamic fair scheduling with QoS constraints in multimedia wideband CDMA cellular networks. *IEEE Trans. on Wireless Communications*, **3**(1):60-73. [doi:10.1109/TWC.2003.819028]

Yoo, T., Setton, E., Zhu, X., Goldsmith, A., Girod, B., 2004. Cross-Layer Design for Video Streaming over Wireless Ad hoc Networks. Proceedings of IEEE Multimedia Signal Processing Workshop (MMSP).

Zhang, Q., Zhu, W., Zhang, Y.Q., 2005. End-to-end QoS for video delivery over wireless Internet. *Proceedings of the IEEE*, **93**(1):123-134. [doi:10.1109/JPROC.2004.839603]

Zhao, S., Xiong, Z., Wang, X., 2002. Joint error control and power allocation for video transmission over CDMA networks with multiuser detection. *IEEE Trans. on Circuits & System for Video Tech.*, **12**(6):425-437. [doi:10.1109/TCSVT.2002.800336]

Zheng, H., 2003. Optimizing Wireless Multimedia Transmissions through Cross Layer Design. Proceedings of IEEE International Conference on Multimedia & Expo (ICME).