System Scheduling for Multi-Description Video Streaming over Wireless Multi-Hop Networks

Liang Zhou, Benoit Geller, Baoyu Zheng, Anne Wei, Jingwu Cui

System Scheduling for Multi-Description Video Streaming over Wireless Multi-Hop Networks

Liang Zhou, Benoît Geller, Baoyu Zheng, Anne Wei, Jingwu Cui

Abstract

Providing real-time multimedia applications over wireless multi-hop networks is a challenging problem because the wireless channels are highly sensitive to delay, interference and topology change. Multiple description coding (MDC), as a new emerging error-resilient technique, has been widely used recently in wireless video transmission. Its fundamental principle is to generate multiple correlated descriptions such that each description approximates the source information with a certain level of fidelity. Inevitably, MDC introduces many description streams which may influence each other and thus, reasonable system scheduling is needed to provide a satisfied video quality. The novelty of this work is to investigate the optimal distributed scheduling for multiple competing MDC streams in a resource-limited wireless multi-hop network. This is achieved by joint optimization of MDC, rate control and multipath routing. Two joint optimal algorithms, namely a distributed rate control and routing (DRCR) and a simplified DRCR algorithm, are proposed to solve this problem with constraints that arise from the multiple description streams among multiple users via multiple paths. Both algorithms are designed in a distributed manner that is amenable to on-line implementation for wireless networks. Theoretical analysis and simulation results are provided which demonstrate the effectiveness of our proposed joint schemes.

Index Terms

multiple description coding; video transmission; rate control; multipath routing

This work is supported by the International Project PRA-SI (financed by France and China government) under Grant No. SI04-03, and the Key Project of Nature Science Foundation of Jiangsu (China) under Grant BK2007729.

L. Zhou, B. Zheng and J. Cui are with the Electronic Engineering Department, Shanghai Jiao Tong University, Shanghai, China; B. Geller is with the Laboratory of LEI, Ecole Nationale Supérieure des Techniques Avancées (ENSTA-PairsTech), Paris, France; A. Wei is with the LATTIS, University of Toulouse II, Toulouse, France.

Main Acronym and Notation

AIMD	Additive Increase Multiplicative Decrease
DRCR	Distributed Rate Control and Routing
MDC	Multiple Description Coding
MPLS	Multi-Protocol Label Switching
QoS	Quality of Service
SNMP	Simple Network Management Protocol
C_l	Capacity of link l
d_s^0	Distortion when both descriptions are received for user s
d_s^i	Distortion when only description i is received for user s, $i = 1, 2$
d_s^3	Distortion when none description is received for user s
f_l	Background traffic in link l
H_s^i	Description i of user s, $i = 1, 2$
K	Average packet length
L,l	L is the set of links, and l is one link in L
p_l	Packet error probability in link l
R_s^i	Rate for the description i of user s
P_s^i	Packet loss rate for the description i of user s
r_l	Traffic in link l
S,s	S is the set of users, and s is one user in S
T	Delay constraints
θ_l	Link utilization for link l

I. Introduction

Rapid growth in wireless networks is fueling the demand that services traditionally available only in wired networks, such as video, be available to mobile users. However, the characteristics of wireless systems provide a major challenge for reliable transport of video since the wireless multimedia transmission is highly sensitive to delay, interference and topology change, which can cause both packet losses and bit-errors. Current and future wireless systems will have to cope with this lack of QoS (Quality of Service) guarantees [1].

The issue of supporting error-resilient video transport over error-prone wireless networks has received considerable attention recently. Along one thread, some works presented some source coding-based error-resilient approaches that divide the original bit-stream into multiple streams, called multiple description coding (MDC), to tradeoff the error-resilience and the coding complexity (see [2], [3] and the references therein). The fundamental principle of MDC is to generate multiple correlated descriptions of the source such that each description approximates the source with a certain level of fidelity [8]. [4] and [5] amplified the benefits of using MDC by combining it with path diversity; in this context, each stream is explicitly transmitted over an independent path to the receiver in order to achieve higher tolerance to packet loss and delay due to network congestion. Along another thread, some researchers studied the network congestion control and optimal routing for wireless video transmission (see [6], [7] and the references therein) so that the network can be stable, robust and the users can have better QoS for the applications. In [6], congestion control and multipath routing were studied and it demonstrated that there are significant advantages when each source randomly selects multiple paths from all its available choices. In [7], it showed that the optimal allocated rate strikes a balance between the selfish motivation of minimizing video distortion and the global goodness of minimizing network congestions, while the routes are chosen over least-congestion links in the network. [11] researched MDC video streaming over wireless network, however, it only considered the multi-routing problem and did not present a systematic scheduling algorithm.

Typically, for real-time video communications over resource-limited wireless networks, the key point is how to allocate the resource to different users to minimize the total video distortion. In this paper, we will employ MDC as our error-resilient target. For such a case, there are many and different description streams in multiple paths which may influence each other and thus the sources should choose optimal rate-distortion points and provide reasonable transmission rate and routing such that the video sources can be both error-resilient and network-adaptive. However, to the best of our knowledge, the current literature considered MDC, congestion control and multipath routing separately and independently. In order to achieve improved video quality supported by wireless networks, and to provide a more robust video delivery system, these factors are jointly considered in this paper.

In this work, we propose a framework to combine multipath routing and rate allocation with the asymmetric MDC in a general wireless multi-hop network in order to minimize the total distortion of all users. Our framework reflects the intrinsic tradeoff between the rates and distortions at source coding, and the tradeoff between the allocated rates and packet loss due to congestion and unreliable links. The main contributions and novelties of this paper are: (1) providing a general framework to combine multipath routing and rate allocation with multiple asymmetric MDC streams in a wireless multi-hop network; (2) proposing a joint routing and rate control algorithm for system scheduling to maximize the end-to-end quality of all the users; (3) extending and simplifying this joint algorithm to improve its convergence rate. Note that both algorithms are designed in a distributed manner that is amenable to on-line implementation for wireless networks.

The rest of the paper is organized as follows. In section II, we formulate the problem for the network with asymmetric MDC. Based on the problem analysis, a distributed joint rate control and routing scheme is proposed in section III. In section IV, an extended algorithm is presented to improve its convergence rate, and related works are given in section V, followed by the concluding remarks in section VI.

II. Problem Formulation

A. Video Distortion

Consider a wireless network with L links, and each link $l \in L$ with a capacity of C_l. In this system, there are S users, and each user $s \in S$ uses asymmetric MDC. In this work, we focus on two-description since it is the most widely used for MDC video. Each description of user s is denoted as H_s^i for $i = 1, 2$, each using a set of links $L(H_s^i)$ from source to destination. Each link l is shared by a set of descriptions $\cup_{i=1}^2 H_s^i(l)$ of each user s.

For each user s, denoting d_s^0 as the central distortion when both descriptions are received, d_s^i for $i = 1, 2$ as the side distortion if description H_s^i is received, and d_s^3 as the distortion if none of the description is received. Besides, denoting R_s^i and P_s^i for $i = 1, 2$ as the rate and the packet loss probability of the path for description H_s^i, respectively. Our system scheduling aims at finding the optimal rate-distortion operating points to minimize the total distortion

$$\min \qquad \sum_{s\in S} D_s, \tag{1}$$

where D_s is the distortion of each user s, the expected average video distortion at the receiver can be approximated as:

$$D_s = d_s^0\big(1 - P_s^1\big)\big(1 - P_s^2\big) + d_s^1\big(1 - P_s^1\big)P_s^2 + d_s^2\big(1 - P_s^2\big)P_s^1 + d_s^3 P_s^1 P_s^2. \tag{2}$$

In this paper, assuming Gaussian sources with zero and variance σ^2, we employ the following distortion-rate regions [11]:

$$
\begin{cases}
d_s^0 \geq \frac{2^{-2(R_s^1+R_s^2)}}{2^{-2R_s^1}+2^{-2R_s^2}-2^{-2(R_s^1+R_s^2)}} \cdot \sigma^2 \\
d_s^1 \geq 2^{-2R_s^1} \cdot \sigma^2 \\
d_s^2 \geq 2^{-2R_s^2} \cdot \sigma^2 \\
d_s^3 = \sigma^2.
\end{cases}
\tag{3}
$$

B. Packet Loss Approximation

In wireless video communications, the reconstructed video quality is affected by quality degradation due to packet losses either caused by late arrival or transmission error. Late arrival is induced by the packet end-to-end delay exceeding delay constraints T and transmission error is caused by unreliable wireless links. In this paper, we don't employ any retransmission policy when transmission errors occur, so we can view the two previous kinds of distortion as independent and additive. Thus, we can calculate P_s^i as:

$$
P_s^i = P_s^{i,late} + P_s^{i,error}, \tag{4}
$$

where $P_s^{i,late}$ and $P_s^{i,error}$ refer to the packet losses due to late arrival and transmission error, respectively. Assume that on each link l, in addition to the MDC traffic r_l we are interested in, there is some background traffic f_l. In a bandwidth-limited network, this combined loss rate can be further modeled based on the M/G/1 queuing model. In this case, the delay distribution of packets over a single link is exponential [12]. Note that, since the end-to-end delay of packet delivery in wireless network is dominated by the queuing delay at the bottleneck link, the empirical delay distribution for realistic traffic patterns can still be modeled by an exponential formulation:

$$
P_s^{i,late} \approx \max_{l \in L(H_s^i)} e^{-\mu_l(1-\theta_l)T} \tag{5}
$$

where T reflects the maximum tolerate delay, μ_l is the service rate C_l/K (K is the average packet length) and θ_l is the link utilization $(r_l + f_l)/C_l$ for link l. If each link is independent with each other, the end-to-end packet loss probability induced by transmission error for description H_s^i is

$$
P_s^{i,error} = 1 - \prod_{l \in L(H_s^i)} (1 - p_l), \tag{6}
$$

where p_l is the packet error probability of link l. According to [9], p_l is an increasing function of r_l, and a bound on this function is

$$
p_l \geq \frac{1}{2} 2^{-KC_l(1-\theta_l)}. \tag{7}
$$

Assuming that the error probability of each link l is small, we can approximate $P_s^{i,error}$ as

$$P_s^{i,error} \approx \sum_{l \in L(H_s^i)} p_l \approx \sum_{l \in L(H_s^i)} \frac{1}{2} 2^{-KC_l(1-\theta_l)}. \tag{8}$$

Therefore, the overall end-to-end packet loss P_s^i can be expressed as

$$P_s^i \approx \max_{l \in L(H_s^i)} e^{-\mu_l(1-\theta_l)T} + \sum_{l \in L(H_s^i)} \frac{1}{2} 2^{-KC_l(1-\theta_l)} = F(r_l), \tag{9}$$

where F is just a function chosen for notation convenience.

C. Optimization Problem

We investigate the optimal system scheduling by jointly optimizing the rate-distortion adaptation and congestion control. Based on the previous discussion, the optimal problem is:

$$\begin{aligned} \min \quad & \sum_{s \in S} D_s \\ \textit{subject to} \quad & P_s^i \geq F(r_l), l \in L(H_s^i), \forall s, i = 1,2 \\ & r_l = \textstyle\sum_{i=1}^2 \sum_{H_s^i \in H_s^i(l)} R_s^i \leq \theta_l C_l - f_l, \forall l \\ & d_s^0, d_s^i, R_s^i, P_s^i, r_l, \theta_l, \forall s, i = 1,2, \forall l \end{aligned} \tag{10}$$

where D_s is defined by (2). The first constraint is for the packet loss and the second is the flow constraint. This problem is a non-convex optimization problem. But if the ordering of R_s^i ($i = 1,2$) is known for every user s[1], the optimization problem is a convex optimization if we apply a logarithmic change of variable [13]. With the additional log change of variable to P_s^i, $P_s^i = exp(\widetilde{P}_s^i)$, $\widetilde{P}_s^i \leq 0$, the optimization problem (10) becomes:

$$\begin{aligned} \min \quad & \sum_{s \in S} D_s \\ \textit{subject to} \quad & \widetilde{P}_s^i \geq \log F(r_l) = \widetilde{F}(r_l), l \in L(H_s^i), \forall s, i = 1,2 \\ & \textstyle\sum_{i=1}^2 \sum_{H_s^i \in H_s^i(l)} R_s^i \leq \theta_l C_l - f_l - \xi r_l^2, \forall l \\ & d_s^0, d_s^i, R_s^i, \widetilde{P}_s^i, r_l, \theta_l, \forall s, i = 1,2, \forall l \end{aligned} \tag{11}$$

[1]In practical, R_s^i is known by setting its original value for all users. It should be noted that when one stream updates in current time slot, the other steams which share links with this stream unchange in this time slot. One iteration contains all of the streams completing one update.

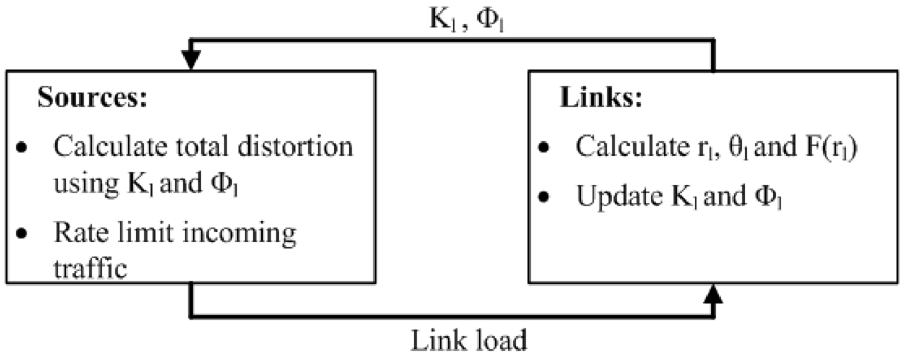


Fig. 1. The diagram for DRCR algorithm.

To make (10) strictly convex in R_s^i, we add $-\xi r_l^2$ to the right side of the constraints in (11), where ξ is a small number such that ξr_l^2 is small compared with r_l. Since there is a fixed closed interval for R_s^i, the optimization problem (11) is a convex optimization given the ordering of possible R_s^i for every user s [24]. In the next section, we propose a distributed algorithm where each source and each link solve their own problem with only local information through the standard dual decomposition.

III. Distributed Rate Control and Routing

In this section, we first describe the Distributed Rate Control and Routing (DRCR) algorithm. Then, we evaluate DRCR 's convergence and robustness.

A. DRCR Algorithm

DRCR is a joint multipath routing and rate allocation scheme where the sources split traffic for each source-destination pair over multiple paths. The key challenges in designing DRCR are how to select optimal paths $L(H_s^i)$ as well as allocated rates R_s^i to ensure that the resulting system is both stable and optimal. We illustrate the interplay between the sources (that determine $L(H_s^i)$ and compute the rates R_s^i) and the network links (that feedback link price ϕ_l and congestion price κ_l) in Fig. 1.

DRCR proceeds first by determining the available paths between the source and destination, and then by deciding paths according to the minimum distortion. It proceeds in two phases, the path discovery and path reservation phases, respectively. To this aim, control messages are exchanged between the source and the destination via forwarding by the intermediate nodes. In order to derive exact bounds on the performance of DRCR, we assume that the control channel is reliable, and the nodes are synchronized[2] (i.e., there is a bounded time interval in which all nodes receive all dedicated control packets).

TABLE I

THE DRCR ALGORITHM

Source s: determine the optimal path for each source-destination pair

$$\min \quad \sum_{s\in S} D_s - \sum_i \sum_{l\in L(H_s^i)} \phi_l(t)\widetilde{P}_s^i - \sum_i R_s^i \kappa_s^i(t)$$

where $\phi_l(t)$ denotes the price for link l and $\kappa_s^i(t)$ refers to the end-to-end congestion price for H_s^i at iteration t. Assuming the congestion price for link l is $\kappa_l(t)$, $\kappa_s^i(t)$ can be expressed as: $\kappa_s^i(t) = \sum_{l\in L(H_s^i)} \kappa_l(t)$

• Congestion Price Update:
$\kappa_l(t+1) = [\kappa_l(t) + \lambda_\kappa(t)\big(r_l'(t) - \theta_l C_l + f_l + \xi r_l^2\big)]^+$
where $[x]^+ = \max(x,0)$, $\lambda_\kappa(t)$ is the step size and
$r_l'(t) = \sum_{i=1}^2 \sum_{H_s^i \in H_s^i(l)} R_s^i(t)$ is the aggregate rate of all the sources on link l at iteration t.

Link l: determine the optimal traffic in each link

$$\min \quad \phi_l'(t)\widetilde{F}(r_l) + \kappa_s^i(t)\big(\theta_l C_l - f_l - \xi r_l^2\big)$$

where $\phi_l'(t) = \sum_{i=1}^2 \sum_{H_s^l \in H_s^i(l)} \phi_l(t)$ is the aggregate traffic load reduction price paid by sources using link l.

• Link Price Update:
$\phi_l(t+1) = [\phi_l(t) + \lambda_\phi(t)\widetilde{F}(r_l)]^+$, $l \in L(H_s^i)$
where $\lambda_\phi(t)$ is the step size.

The source sends all outgoing links with path discovery messages, which are forwarded by the intermediate nodes on the control channel. At each intermediate node, the path discovery

[2]We use the method proposed in [14] to guarantee the node synchronization in one practical experiment.

messages contain the information of congestion price and link price related to every possible flow between the source and intermediate nodes. Each intermediate node then extends the path as the source does. Upon reception of path discovery messages from the destination, the source determines the possible paths $L(H_s^i)$ between the source and destination based on explicit feedback from the links, in the form of link price ϕ_l and congestion price κ_l. In particular, the source minimizes the total distortion while balancing the price of using path $L(H_s^i)$. The path price is the product of the source rate with the price per load for path $L(H_s^i)$ (computed by summing ϕ_l over all the links in the path). In fact, it is similar to the standard TCP dual algorithm except that the maximization problem is conducted over a vector not a scalar, to reflect the multi-path nature of DRCR.

The assignment of R_s^i at the source is determined by the total traffic that traverses each link in $L(H_s^i)$. The resulting aggregate traffic load price on link $l \in L(H_s^i)$ is $\sum_i \sum_{H_s^i \in H_s^i(l)} \phi_l \widetilde{F}(r_l)$, which serves as an implicit feedback that the link uses to compute the congestion price κ_l. By the standard dual decomposition approach [9], DRCR is realized by each source and each link solving their own problem with only local information to get an optimal solution for Eq. (11). In this case, each source adjusts its offered congestion price per unit traffic load for each description and each link in its path determines its total traffic that maximize the "net income" of the network based on its link price. The whole DRCR algorithm is shown in Table I.

Since the fixed descriptions for each source, the computations at the sources and the links are linear with the number of the sources. In addition to computation overhead, there are three new functionalities required by DRCR. First, DRCR requires MPLS (Multi-Protocol Label Switching) for splitting traffic over multiple paths. Second, DRCR requires frequent link-load measurements which is possible using the SNMP (Simple Network Management Protocol). Finally, DRCR requires an explicit rate limit for the incoming traffic, and this can be done by dropping packets sent above the allowed rate.

B. Optimality and Stability Characteristics

In this subsection and Appendix I, we provide the analytical derivation and theoretical foundation of the DRCR algorithm.

Theorem 1: The algorithm DRCR converges to the joint global optimum (D_s, R_s^i) of Eq. (11) for sufficiently small step sizes λ_κ and λ_ϕ.

Outline of the Proof: The idea of DRCR algorithm is to decouple the coupled objective function in Eq. (11) by introducing auxiliary variables and additional constraints, and then use Lagrange dual decomposition to decouple all of the constraints. There are two exact steps: (1) Introducing new variables to enable decoupling; (2) Employing dual decomposition and gradient descent method to derive the DRCR algorithm. See Appendix I for the detailed proof. ■

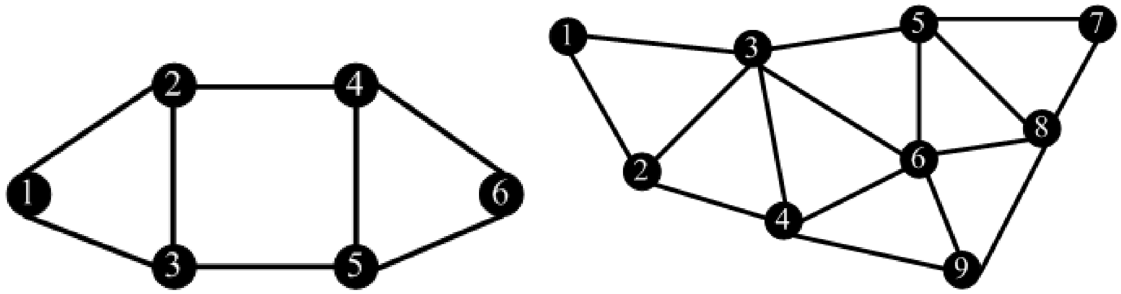
So far, we have taken a deterministic model with a static population of sources, and stability here means global asymptotic convergence. Given the dynamic nature of DRCR, it is natural to wonder whether it would also behave well with stochastic variations in traffic. Consider sessions arriving according to a Poisson process with exponentially-distributed file sizes. A session leaves the network after it finishes transmitting a file. The service rates are determined by the solution of the DRCR algorithm. Note that sessions may arrive and depart even before the DRCR algorithm converges, i.e., we do not assume time-scale separation between the algorithm convergence and the stochastic stability of DRCR (whether the number of active sessions and the sizes of the queues in the network remain finite for DRCR in such dynamic environment). The answer is positive, as summarized in the following theorem.

Theorem 2: The DRCR algorithm is stochastically stable if the average arrival load in each link is smaller than its capacity, i.e., the stochastic region of DRCR is the largest possible one: the interior of the feasible region of problem (11).

Outline of the Proof: The key idea is to show that dual variables are scaled versions of queue lengths, and then to find that the DRCR algorithm is a special case of dual-based algorithms for generalized minimal distortion whose stochastic stability has already been established [14]. ■

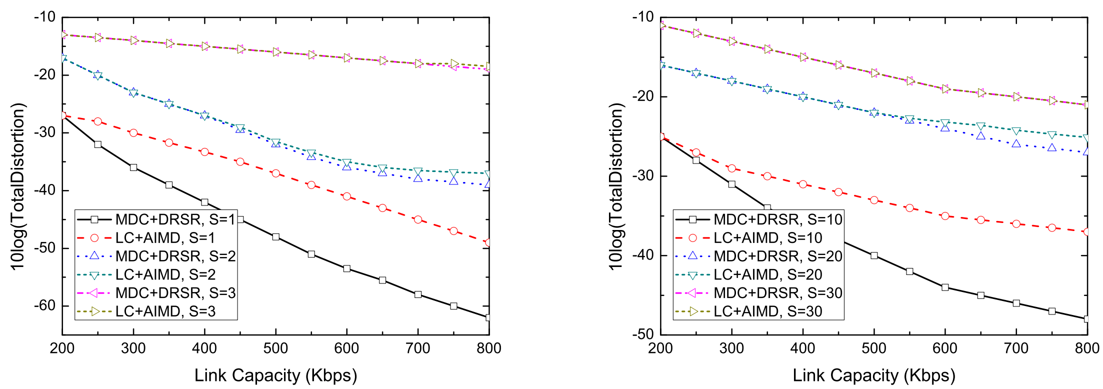
C. Simulations

To study the proposed DRCR algorithm, we display some experiment with two wireless multi-hop networks shown in Fig. 2. Fig. 2(a) is a geometrical-mesh topology, which is representative of a specific network structure. Of the many possible source-destination pairs, we choose 1-6. If the number of user $S = 2$ and $S = 3$, the source-destination pair is replaced by 2 and 3 uses, respectively; Fig. 2(b) is a general network, and we choose 10 source-destination pairs: 1-4, 2-4, 1-3, 2-6, 3-8, 4-5, 7-5, 9-5, 4-6, and 6-9. Similarly, if the number of user $S = 20$ and $S = 30$, the source-destination pair is replaced by 2 and 3 uses, respectively. It should be noted that, in order to avoid negative values, the link capacity of each network is assumed to



(a) 8-link network (b) 16-link network

Fig. 2. Two network topologies.



(a) 8-link network (b) 16-link network

Fig. 3. Distortion comparison of different methods for different sources.

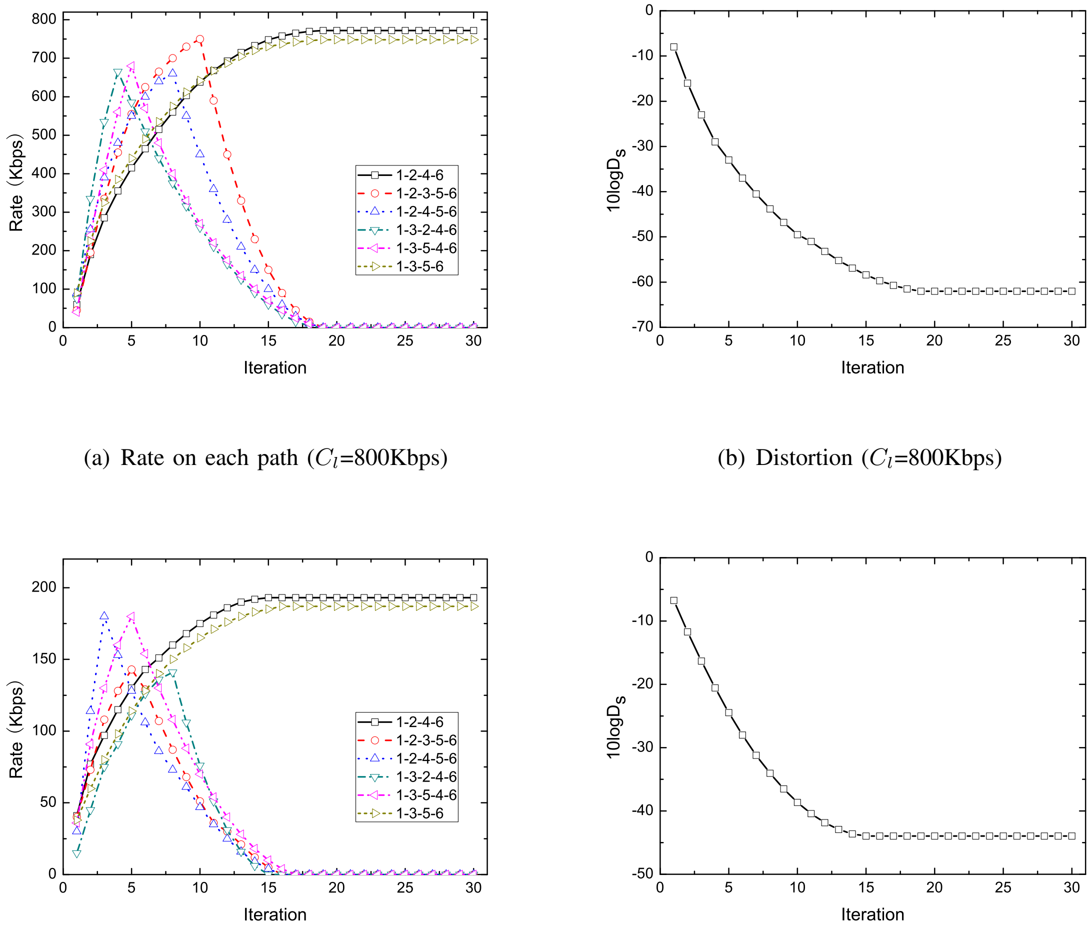
follow a truncated Gaussian distribution, with an average varying from 200Kbps to 800Kbps. In all experiments, we start with an initial routing configuration (i.e., the earliest path known by the source) that splits the traffic evenly among the paths for each source-destination pair. For background flow f_l, it is generated according to an on/off source model with exponential distribution of staying time, and average rates spanning between 0 and 100Kbps. To simulate the video application, one HD (High-Definition) sequence (*City*) is used. In terms of HD video, the sequence has spatial resolution of 1280×720 pixels, and the frame rate of 60 frames per second. In addition, we employ the 3-D SPIHT coder, an asymmetric MDC coder introduced in [8]. In the following, we set $K = 40$Kbits, $T = 300$ms unless otherwise specified.

TABLE II

PERFORMANCE COMPARISON FOR DIFFERENT METHODS IN DIFFERENT NETWORKS WITH VARIABLE LINK CAPACITIES

User Number	Network Type	Average Distortion of Different Methods				Average Improvement
		MDC+DRSR	SDC+DRSR	LC+AIMD	Ref. [11]	
1	8-link	-48	-44	-38	-38	4∼10
	16-link	-77	-70	-64	-60	7∼17
2	8-link	-47	-41	-33	-33	6∼14
	16-link	-72	-65	-55	-53	7∼19
3	8-link	-45	-36	-30	-27	9∼18
	16-link	-66	-58	-50	-53	8∼16
10	8-link	-33	-15	-24	-20	9∼18
	16-link	-47	-27	-37	-33	10∼20
20	8-link	-31	-11	-19	-21	10∼20
	16-link	-44	-22	-30	-32	12∼22
30	8-link	-27	-7	-15	-5	12∼22
	16-link	-41	-16	-26	-23	15∼25

In order to evaluate the performance of the joint scheme, the proposed MDC and DRCR scheme (noted as MDC+DRCR) is benchmarked against joint Layered Coding (LC) with Additive-Increase-Multiplicative-Decrease (AIMD) rate allocation scheme (noted as LC+AIMD). LC, which is also called hierarchical coding, can be considered as an extreme case of MDC, where streams are useful only after successful reception of higher priority layers; AIMD-based rate allocation method is used by TCP congestion control [10]. Fig. 3 shows the distortion comparison for different users under different networks. The proposed MDC+DRCR scheme can be seen to achieve a higher performance in terms of end-to-end distortion ($10\log\sum_s D_s$) compared to the competing scheme. For the 8-link network, when $S = 1$ and $C_l = 800$Kbps, the average distortion using the proposed scheme is -62 while it is -49 for the case of LC+AIMD, thus, 13 performance gain can be achieved on average using the proposed scheme. Similarly, for the 16-link network, when $S = 1$ and $C_l = 800$Kbps, around 11 performance gain can be achieved on the average. It should be noted that as the user number increases or link capacity decreases, the gap between the proposed MDC+DRCR and LC+AIMD reduces. That is because when users are numerous or link capacities are limited, the total congestion price of each possible path is large and relatively approaches to each other. In this case, the performance difference between



(a) Rate on each path (C_l=800Kbps) (b) Distortion (C_l=800Kbps)

(c) Rate on each path (C_l=200Kbps) (d) Distortion (C_l=200Kbps)

Fig. 4. Plots of rate and distortion versus time for source pair 1-6 in 8-link network for $C_l = 800$Kbps and $C_l = 200$Kbps, respectively. Step sizes: $\lambda_\kappa = 4 \times 10^{-5}$, $\lambda_\phi = 2 \times 10^{-5}$.

the DRCR and AIMD is small.

Then, we test the efficiency of the proposed scheduling scheme when the link capacity is time varying from 200Kbps to 800Kbps. To demonstrate the advantages of using MDC, we compare MDC with single description coding (SDC) and the method in [11]. Similarly, our MDC+DRSR is benchmarked against SDC+DRSR, LC+AIMD, and Ref. [11]. The simulation results are presented in Table II. From Table II, we can see that the proposed MDC+DRSR scheme has some performance improvement compared to other competing schemes. It should be noted that although SDC has high coding efficiency, however, it is very sensitive to network conditions, especially to the case of limited network resources. For example, for the 8-link network, when $S = 1$, the performance gap between MDC+DRSR and SDC+DRSR is 4, while it increases to 20 when $S = 30$. This in turn illustrates the benefits of using MDC in wireless networks. Moreover, we can also observe that MDC+DRSR also outperforms Ref. [11]. That is because Ref. [11] only considers the routing problem, while our paper joint considers the video distortion, rate allocation and multi-path routing.

Then, to present a clear picture of how DRCR works, the proposed DRCR algorithm is operated over the 8-link network when $f_l = 0$, $S = 1$ and all of the links are with the same capacity. The graphs in Fig. 4 illustrate both the rates and the total distortion of each iteration for 800Kbps and 200Kbps link capacity, respectively. It can be observed that the possible paths for 1-6 pair have changed over the iterations, and re-dispensing its own traffic over the network to avoid already congested links. Changes in the paths also affect the congestion-increment information, which in turn leads to changes in the rate allocation decisions. Note that, DRCR needs numerous iterations to get the optimal value even when $S = 1$. Therefore, to make DRCR suitable for a more general condition, it is necessary to improve its convergence rate. In addition, a number of interesting observations can be made from these graphs. For example, the final selected paths are both disjoint and the optimal rates in these selected paths are very close to the link capacity. In the next section, we will study some possible rules to simplify DRCR.

IV. Extension: Simplified DRCR for Improving Convergence Rate

So far, we have discussed the equilibrium behavior of the DRCR algorithm. Since the convergence rate for any distributed algorithm on wireless networks is particularly important because the network resource and system traffic are dynamic and source traffic may exhibit low degree of stability, this section simplifies the DRCR algorithm to improve its convergence rate.

A. Design Guide

In this subsection, we derive several theorems to guide the design of a simplified DRCR strategy. This subsection shows that, in the optimal rate allocation, a flow is either used at full bandwidth or not used at all. Furthermore, the optimal rate allocation always chooses the lowest congestion price paths, i.e., a path is selected because there are no other paths with a lower congestion price. We start from an ideal streaming scenario with adequate disjoint paths, and then add the bottleneck links and streaming constraints.

First, assuming that there are adequate disjoint paths for all of the descriptions H_s^i, $\forall s \in S$, $i = 1, 2$. From the minimum distortion point of view, any of the two different descriptions H_s^i streaming over different paths with the same packet loss rate P_s^i, then, they can be viewed as a single flow with an aggregated rate with packet loss rate P_s^i. In this case, we first claim that the optimal rate either uses a full bandwidth path, or does not use it at all.

Lemma 1: Given any description H_s^i having rate $R_s^i \in [0, R_{max}]$ (R_{max}: maximum potential rate for R_s^i) and a distortion metric $\sum_{s \in S} D_s$ (D_s is in Eq. (2)), the optimal solution of the rate allocation problem when all the paths are disjoint and the optimal value of R_s^i is either R_{max} or 0, $\forall s$, $i = 1, 2$.

Outline of the Proof: The proof consists of two main steps. First, we show that the optimal rate allocation problem is achieved when paths are disjoint. In this case, we need to prove that $R_s^i = r_l$, $l \in L(H_s^i)$, $\forall s \in S$, $i = 1, 2$. We derive the distortion function with respect to rate R_s^i and r_l, respectively. Then, we observe and analyze the conditions of an extremum $\partial \sum_{s \in S} D_s / \partial R_s^i = 0$ for any R_s^i and r_l. See Appendix II for the detailed proof. ■

Next, we consider the case when there exists joint paths (bottleneck links) in the network. In order to get the minimum distortion, we choose the paths with minimum total network congestion.

Proposition 1: If there are some joint paths in the network, the optimal path selection strategy is to choose the "least" joint paths which leads to the minimum total network congestion, and the rate allocation in this case is the same as *Lemma 1*.

Outline of the Proof: From the point view of distortion, the less congestion, the less distortion. Therefore, the basic proof of *Proposition 1* is identical to *Lemma 1*, and the only difference is that the flow constraints have changed. Then we just show that the optimal solution is achieved when R_s^i gets to the largest possible value approaching to r_l. ■

Note that the previous theorem deals with how to choose optimal paths when "joint paths" occurs. The following we address how to allocate the rate for the bottleneck link. Assuming l_b is a bottleneck link, and $\mathbf{L}_{l_b} = \{L(H_s^i)\}$, $\forall s, i : l_b \in H_s^i$, be set of paths (at least two distinct descriptions) sharing the bottleneck link l_b. Note that l_b may, or may not be a bottleneck link for any of the paths $L(H_s^i)$, treated independently. The following theorem regulates the sharing the bandwidth of l_b among these joint paths:

Theorem 3: Let l_b be a bottleneck link for the set of paths $\mathbf{L}_{l_b} = \{L(H_s^i)\}, \forall s, i : l_b \in H_s^i$, the bottleneck link bandwidth shall be shared among paths $L(H_s^i)$ in a greedy way, starting with the path with the lowest congestion price.

Outline of the Proof: Let the path $L(H_s^i) \in \mathbf{L}_{l_b}$ be arranged in an increasing order of their congestion price $\kappa_s^i = \sum_i \sum_{l \in L(H_s^i)} \kappa_l$. Let $\mathbf{R}_k = \{R_k\}_{k \in \mathbf{L}_{l_b}}$ denote a valid rate allocation among these joint paths. Recall that a valid rate allocation has to satisfy the multiple flow constraints. Let $L(H_s^i)$ be the path with lowest congestion price in $\mathbf{L}_{l_b}$. If the rate in $L(H_s^i)$ is not the largest one, one can always find a better solution by transferring rate from other flows sharing the same bottleneck link. Since the total rate stays constant at that moment, the rate transfer does not violate the multiple flow constraints. It however changes the total source distortion, resulting in a decreased overall performance. By induction, the proof can be extended to all the joint paths. This shows that, for any valid rate allocation $\mathbf{R}_k = \{R_k\}_{k \in \mathbf{L}_{l_b}}$, there exists a best solution that fills up in priority the lowest congestion price path. ■

B. Simplified DRCR

The previous theorems represent the keys for designing a simplified DRCR algorithm to improve its convergence rate. There are two main changes in the simplified DRCR: (1) cancel the link price. r_l value just depends on whether it serves for source-destination pairs or not and on each link's capacity; (2) decouple the joint paths. For the joint paths, each R_s^i is re-allocated in a greedy way starting with the path with the lowest congestion price, and recalculate the corresponding congestion price of each path again. Therefore, the convergence rate of the DRCR can be improved dramatically by simplifying the rate control procedure. Table III proposes a sketch of the simplified DRCR.

In the following, we show some properties of the simplified DRCR: (1) algorithm converges in one round if paths are disjoint; (2) algorithm terminates in a finite number of rounds in any condition; (3) algorithm converges to the joint global optimum value of Eq. (11).

Property 1: If the paths requested by the sources do not share any bottleneck joint link l_b, the simplified DRCR converges in one round.

Proof: Let $\mathbf{L}^*(H_s^i) = \{L(H_s^i)\}$ be the optimal set of paths chosen by the destination for transmission. According to *Lemma 1* and *Proposition 1*, for any available path rate R_s^i is less than the available bandwidth of each link $l \in L(H_s^i)$. Since, by hypothesis, the chosen paths

$\mathbf{L}^*(H_s^i)$ do not contain any joint bottleneck link, this means that $r_l \geq \sum_s \sum_i R_s^i$, $l \in L(H_s^i)$. This means that any node, upon the reception of reservation packets, can allocate the request rate on the outgoing links for all requested flows. Therefore, the source can compute the optimal allocation after one round of the protocol. ■

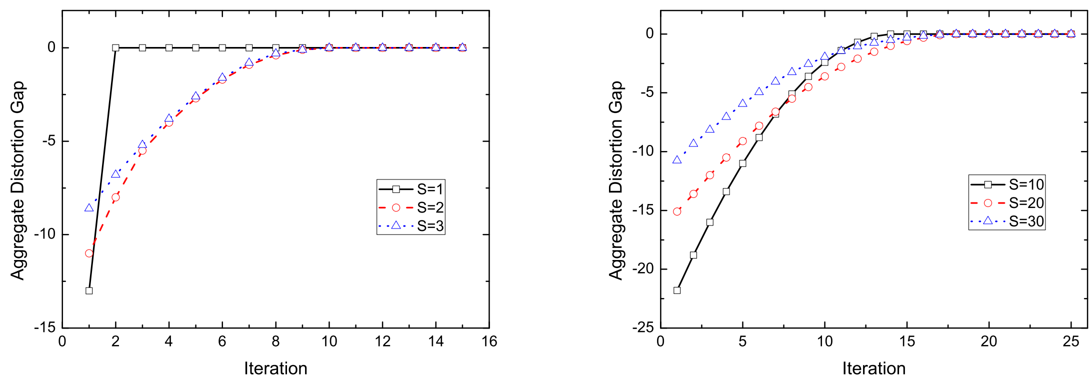
Property 2: If the paths requested by the sources have shared bottleneck joint links and their number are N_{l_b} ($N_{l_b} \geq 1$), simplified DRCR terminates in $N_{l_b}+1$ rounds.

Proof: This result can be seen as an extension of *Property 1*. For the simplified DRCR, since using the greedy way based on the known congestion price of each possible path, the algorithm can deal with at least one joint link, and the available rate of the links and the corresponding

TABLE III

SIMPLIFIED DRCR ALGORITHM

Input:
source, destination, available network topology;
Output:
optimal source-destination paths and corresponding allocated rates;
Initialization:
$R_s^i = 0$, $\kappa_s^i = 0$, f_l=0, θ_l=1, $\forall s, i, l$;
Procedure JointRoutingRate
while (true)
for s=1 to S **do**
    for i=1 to 2 **do**
        if L_s^i is a disjoint path **then**
            $r_l = \theta_l C_l - f_l$, $l \in L(H_s^i)$;
            $R_s^i = \min_{l \in L(H_s^i)} r_l$;
            $\kappa_l(t+1) = [\kappa_l(t) + \lambda_\kappa(t)\big(\sum_{i=1}^2 \sum_{H_s^i \in H_s^i(l)} R_s^i(t) - \theta_l C_l + f_l + \xi r_l^2\big)]^+$;
            $\kappa_s^i(t) = \sum_{H_s^i \in H_s^i(l)} \kappa_l(t)$;
            select path with minimum κ_s^i;
        else
            Decouple joint paths and allocate each R_s^i according to *Theorem 3*;
            Recalculate κ_s^i and reselect the path with minimum κ_s^i;
    end for
end for



(a) 8-link network

(b) 16-link network

Fig. 5. Plots of aggregate distortion gap using simplified DRCR ($C_l = 200$Kbps, $f_l = 0$).

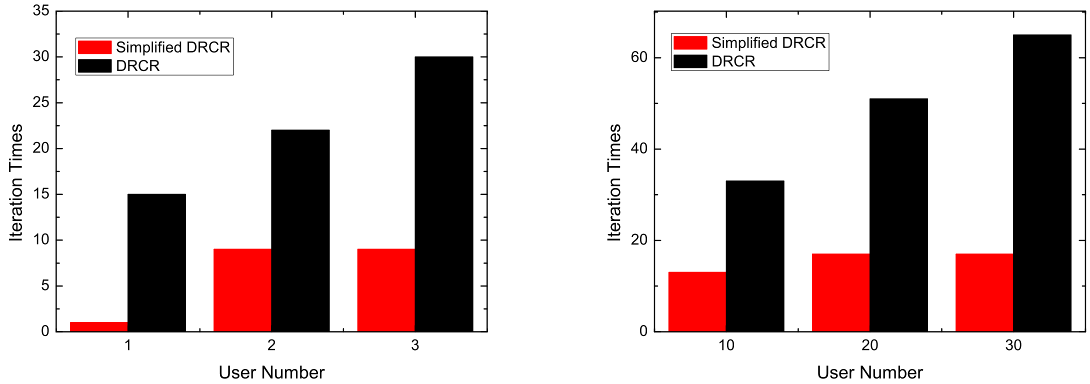
rate allocation for each H_s^i can be adjusted at each round. Hence, on subsequent rounds of the algorithm, the sources will be able to deal with a finite number of flows and the algorithm terminates in $N_{l_b} + 1$ rounds. ■

Property 3: The simplified DRCR converges to the joint global optimum value of Eq. (11).

Proof: In fact, as for the path selection, the simplified DRCR scheme is identical to the DRCR. The only difference between them is how to allocate the rate to the given possible paths. From *Lemma 1*, *Proposition 1*, and *Theorem 3*, we can know that the final allocated rates by simplified DRCR is the same with the DRCR. From *Theorem 1*, we can find that simplified DRCR also converges to the joint global optimum value of Eq. (11). ■

C. Simplified DRCR Versus DRCR

Following the experimental set-up in subsection.III-C, we compare the simplified DRCR with DRCR. At first, we target DRCR as an optimal solution, and test the simplified DRCR in both networks. In Fig. 5, we plot the distortion gap using simplified DRCR with $C_l = 200$Kbps. We can observe that the aggregate distortion gap follows an increasing concave trajectory, converging close to the optimum in a finite iterations. While the graphs in Fig. 5 are for one particular initial conditions, we have done simulations for a variety of initial conditions to verify that the convergence and distortion gap are independent of the initial conditions. Fig. 6 illustrates the



(a) 8-link network

(b) 16-link network

Fig. 6. Iteration times comparison used by Simplified DRCR and DRCR ($C_l = 200$Kbps, $f_l = 0$).

number of iteration times corresponding to Fig. 5. Clearly, the simulation results are consistent with the analytical properties in subsection.IV-B. In particular, the simplified method can reduce the number iteration times dramatically compared to DRCR.

V. Related Works

In this section, we present some related works on the multi-path routing and multi-streaming rate control that have not been discussed earlier in the paper. In addition, we also indicate the difference between our proposed method with the previous works.

A. Multi-Path Routing

Multi-path routing has been an active research topic over the years. For example, various polynomial time algorithms have been proposed to compute multiple shortest paths. Other important works include, node- or link-disjoint path routing, and braided multiple path routing. However, most of these algorithms do not explicitly consider optimizing performance for the video streams. The problem of path selection for multiple description video streams has recently been explored in [15]. [15] studies the problem of path selection for double-description video in the context of overlay networks, where path selection is formulated as an optimization problem that minimizes video distortion. The problem is solved by an exhaustive search over the exponential solution space. In a recent work, [16] presents a distributed heuristic for finding two maximally disjoint source trees for double description video streaming in ad hoc networks.

Unlike the aforementioned works that just consider multi-path routing for data traffic over wireless networks, we take into account the specific video characteristics in the routing and rate control scheme. Network congestion is considered in the route selection metric, to meet the stringent delay requirement for video transmission. In addition, each source's rate-distortion characteristic is also incorporated in the joint routing and rate control procedure to provide multiple streams with various contents and complexities.

B. Multi-Streaming Rate Control

The issue of multi-streaming rate control is still an open problem and has received considerable attention recently. A mathematical framework of multi-user rate allocation is presented in [17], where the authors also analyzed two classes of distributed solutions, corresponding to the primal and dual decomposition of the optimization objective. In wireless networks, adaptive transmission techniques are typically used to protect the video stream against the time-varying channel [18]. When multiple streams are involved, centralized channel time allocation among multiple wireless stations has been investigated in [19]. Distributed algorithms have also been proposed, using rate-distortion optimized packet scheduling in [20] for rate allocation among streams sharing a bottleneck link. What's more, rate allocation algorithm combined with a packet partitioning algorithm has been proposed to support video streaming from multiple sources to a receiver over the Internet [21]. The rates are chosen to adapt the available network bandwidth for each stream, and the packet partitioning is designed to minimize the start up delay. For video streaming over a wireless multihop networks, a rate control scheme has been proved to efficiently utilize the available wireless link capacity [22].

Our proposed scheme jointly considers the rate control and the routing, and the optimization function contains both network congestion and video distortion. This differs from previous works where routing and rate allocation are considered separately.

VI. Concluding Remarks

In this paper, we have studied a distributed scheduling for multiple competing MDC streams interacting in a resource-limited wireless multi-hop network. The framework is based on the availability of asymmetric MDC, multipath routing and rate control to jointly optimize the end-to-end video distortion of all the users. As detailed in the paper, our proposed distributed rate control and routing scheme as well as its simplified version can be adapted to dynamic wireless networks by adjusting the routing and the allocated rate for each video stream. The theoretical analysis and simulation results demonstrate the effectiveness of our proposed joint scheme for multi-description video streaming transmission over wireless multi-hop networks.

APPENDIX I
PROOF OF THEOREM 1

Since (11) is a convex optimization problem satisfying Slater's condition, the duality gap is zero. Therefore, a distributed algorithm for (11) can be derived through the Lagrange dual problem. First we form the following Lagrangian:

$$\begin{aligned} L(D_s, R_s^i, \phi_l, \kappa_l) &= \sum_{s\in S} D_s - \sum_i \sum_{l\in L(H_s^i)} \phi_l(t)\big(\widetilde{P}_s^i - \widetilde{F}(r_l)\big) \\ &+ \sum_i \sum_{H_s^i \in H_s^i(l)} \kappa_l(t)\big(R_s^i - \theta_l C_l + f_l + \xi r_l^2\big) \end{aligned} \tag{12}$$

- Each Source s:

$$\min \quad \sum_{s\in S} D_s - \sum_i \sum_{l\in L(H_s^i)} \phi_l(t)\widetilde{P}_s^i + \sum_i \sum_{H_s^i\in H_s^i(l)} \kappa_l(t) R_s^i \tag{13}$$

- Each Link l:

$$\min \quad \phi_l'(t)\widetilde{F}(r_l) - \kappa_s^i(t)\big(\theta_l C_l - f_l - \xi r_l^2\big) \tag{14}$$

Recall that $\kappa_s^i(t) = \sum_{l\in L(H_s^i)} \kappa_l(t)$ and $\phi_l'(t) = \sum_{i=1}^2 \sum_{H_s^i\in H_s^i(l)} \phi_l(t)$ refer to the end-to-end congestion price for H_s^i and the aggregate traffic load reduction price paid by sources using link l at iteration t, respectively. The Lagrangian dual function $L_d(\phi, \kappa)$ is defined as the maximized $L(D_s, R_s^i, \phi, \kappa)$ over D_s and R_s^i for given ϕ and κ. Each source can compute an optimizer D_s^* and each link l can compute an optimizer $r_l^*(\phi, \kappa)$. The Lagrange dual problem of (11) is:

$$\min \quad L_d(\phi_l, \kappa_l) = L(D_s^*, r_l^*(\phi_l, \kappa_l), \phi_l, \kappa_l), \tag{15}$$

where (ϕ_l, κ_l) are the dual variables. Note that (15) is a convex minimization. Since $L_d(\phi_l, \kappa_l)$ may be non-differentiable, an iterative subgradient method can be used to update the dual variables to solve (15):

• Link Price Update:

$$\phi_l(t+1) = [\phi_l(t) + \lambda_\phi(t)\widetilde{F}(r_l)]^+, \tag{16}$$

where $\lambda_\phi(t)$ represents the link price step size.

• Congestion Price Update:

$$\kappa_l(t+1) = [\kappa_l(t) + \lambda_\kappa(t)\big(\sum_{i=1}^{2}\sum_{H_s^i \in H_s^i(l)} R_s^i(t) - \theta_l C_l + f_l + \xi r_l^2\big)]^+, \tag{17}$$

where $\lambda_\kappa(t)$ represents the congestion price step size.

This is exactly the DRCR algorithm described in Table I. Certain choices of step sizes , such as $\lambda_\kappa(t) = \lambda_1/t$, $\lambda_\phi(t) = \lambda_2/t$ where $\lambda_1 > 0$, $\lambda_2 > 0$, guarantee that this algorithm will converge to the joint optimum. As to the relationship between the step size and iteration bounds, please refer to [23]. In this case, the convergent point is a globally optimal (D_s, R_s^i) to the problem (11) since we have shown that the problem can be written as convex optimization. ■

Appendix II
Proof of Lemma 1

First, we view R_s^i and r_l, $l \in L(H_s^i)$ as two different variables and $r_l \geq R_s^i$. Using the Lagrange dual function Eq. (12), we derive the $L(D_s, R_s^i, \phi_l, \kappa_l)$ with respect to R_s^i and r_l respectively, we obtain:

$$\begin{aligned}\frac{\partial L(D_s, R_s^i, \phi_l, \kappa_l)}{\partial R_s^i} &= (1-P_s^1)(1-P_s^2)\frac{\partial d_s^0}{\partial R_s^i} + (1-P_s^1)P_s^2\frac{\partial d_s^1}{\partial R_s^i} \\ &+ (1-P_s^2)P_s^1\frac{\partial d_s^2}{\partial R_s^i} + \sum_{H_s^i \in H_s^i(l)} \kappa_l(t)\end{aligned} \tag{18}$$

and

$$\begin{aligned}\frac{\partial L(D_s, R_s^i, \phi_l, \kappa_l)}{\partial R_l} &= \sum_{l \in L(H_s^i)} \phi_l(t)\frac{\partial \widetilde{F}(r_l)}{\partial r_l} + 2\sum_{H_s^i \in H_s^i(l)} \kappa_l(t)\xi r_l \\ &- \sum_{H_s^i \in H_s^i(l)} \kappa_l(t) C_l \frac{\partial \theta_l}{\partial R_l}\end{aligned} \tag{19}$$

Then, in order to get the optimal value of the $L(D_s, R_s^i, \phi_l, \kappa_l)$, it is necessary to get the following equations at the same time:

$$\frac{\partial L(D_s, R_s^i, \phi_l, \kappa_l)}{\partial R_s^i} = 0, \frac{\partial L(D_s, R_s^i, \phi, \kappa)}{\partial R_l} = 0. \tag{20}$$

However, it is easy to find that Eq. (20) can't be realized when $r_l \neq R_s^i$. Likewise, if $r_l = R_s^i$, taking into account the previous results in (16) and (17), when $\frac{\partial L(D_s, R_s^i, \phi_l, \kappa_l)}{\partial R_l} = 0$ we can find that the optimal solution for R_l is:

$$r_l^* = \theta_l C_l - f_l - 2 \sum_{H_s^i \in H_s^i(l)} \xi r_l^*. \tag{21}$$

Since ξr_l are small values for r_l, so r_l^* is:

$$r_l^* = \theta_l C_l - f_l. \tag{22}$$

Likewise, the optimal solution for R_s^i is:

$$R_s^{*i} = \min_{l \in L(H_s^i)} (\theta_l C_l - f_l). \tag{23}$$

Therefore, the optimal solution of rate the (11) when $R_s^i = r_l = \min(\theta_l C_l - f_l)$, for $l \in L(H_s^i)$, $\forall s$, $i = 1, 2$. If l does not belong to any of the description paths, obviously, $r_l = 0$. ■

References

[1] L. Zhou, B. Zheng, A. Wei, B. Geller and J. Cui, "Joint QoS Control for Video Streaming over Wireless Multihop Networks: A Cross-Layer Approach", *International Journal of Electronics and Communications*, article in press, doi: 10.1016/j.aeue.2008.05.004.

[2] E. Akyol, A. M. Tekalp, and M. R. Civanlar , "A Flexible Multiple Description Coding Framework for Adaptive Peer-to-Peer Video Streaming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 2, pp. 231-245, 2007.

[3] I. Radulovic and P. Frossard, "Multiple Description Coding with Redundant Expansions and Application to Image Communication," *EURASIP Journal on Image and Video Processing*, vol. 2007, Article ID 24863, 15 pages, 2007.

[4] J. Kim, R. M. Mersereau, Y. Altunbasak, "Distributed Video Streaming Using Multiple Description Coding and Unequal Error Protection," *IEEE Transactions on Image Processing*, vol. 14, no. 7, pp. 849-861, 2005.

[5] R. Tian, Q. Zhang, Z. Xiang, Y. Xiong, X. Li, and W. Zhu, "Robust and efficient path diversity in application-layer multicast for video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 8, pp. 961-972, 2005.

[6] P. Key, L. Massoulie, and D. Towsley, "Combining multipath routing and congestion control for roubustness," *proc. of CISS*, 2006.

[7] X. Zhu, J. P. Singh, and B. Girod, "Joint Routing and Rate Allocation for Multiple Video Streams in Ad Hoc Wireless Networks," *Journal of Zhejiang University*, Science A, vol. 7, no. 5, pp. 727-736, May 2006.

[8] L. Zhou, B. Zheng, A. Wei, B. Geller, and J. Cui, "A roubust resolution-enhancement scheme for video transmission over mobile ad-hoc networks," *IEEE Trans. Broadcasting*, vol. 54, no. 2, pp. 312-321, June 2008.

[9] J. W. Lee, M. Chiang, and R. A. Calderbank, "Price-based distributed algorithm for optimal rate-reliability tradeoff in network utility maximization," *IEEE J. Sel. Area Comm.*, vol. 24, no. 5, pp. 962-976, May 2006.

[10] E. Altman, K. Avrachenkov, C. Barakata, P. Dube, "Performance analysis of AIMD mechanisms over a multi-state Markovian path," *Computer Networks*, vol. 47, no. 3, pp. 307-326, 2005.

[11] S. Kompella, S. Mao, Y. T. Hou, and H.D. Sherali, "Cross-layer optimized multipath routing for video communications in wireless networks," *IEEE J. Sel. Area Comm.*, vol. 25, no. 4, pp. 831-840, May 2007.

[12] X. Zhu, E. Setton and B. Girod, "Congestion-distortion optimized video transmission over ad hoc networks," *Journal of Signal Processing: Image Communication*, vol. 20, pp. 773-783, 2005.

[13] M. Chiang, "Balancing Transport and Physical Layers in Wireless Multihop Networks: Jointly Optimal Congestion Control and Power Control," *IEEE J. Sel. Area Comm.*, vol. 23, no. 1, pp. 104-116, 2005.

[14] J. Liu, A. Stolyar, "Distributed Queue Length Based Algorithms for Optimal End-To-End Throughput Allocation and Stability in Multi-Hop Random Access Networks," *Proc. 45th Allerton Conference*, Urbana-Champaign, IL, Sep 2007.

[15] A. C. Begen, Y. Altunbasak, O. Ergun, and M.H. Ammar, "Multi-path selection for multiple description encoded video streaming," *EURASIP Signal Processing: Image Commun.*, vol. 20, no. 1, pp. 39-60, Jan. 2005.

[16] W. Wei, A. Zakhor, "Multiple Tree Video Multicast Over Wireless Ad Hoc Networks," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 17, no. 1, pp. 2-15, Jan. 2007.

[17] F. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," *Journal of Operations Research Society*, vol. 49, no. 3, pp. 237-252, 1998.

[18] Y. Shen, P. C. Cosman, and L. B. Milstein, "Error resilient video communications over CDMA networks with a bandwidth constraint," *IEEE Trans. on Image Processing*, vol. 15, no. 11, pp. 3241-3252, Nov. 2006.

[19] M. van der Schaar and N. Sai Shankar, "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms," *IEEE Wireless Communications*, vol. 12, no. 4, pp. 50-58, Aug. 2005.

[20] J. Chakareski and P. Frossard, "Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources," *IEEE Trans. on Multimedia*, vol. 8, no. 2, pp. 207-218, Apr. 2006.

[21] T. Nguyen, A. Zakhor, "Multiple sender distributed video streaming," *IEEE Trans. on Multimedia*, vol.6, no.2, pp.315-326, 2004.

[22] M. Chen, A. Zakhor, "Rate Contcrol for Streaming Video over Wireless," *Proc. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM04)*, Hong Kong, China, pp.1181-1190, 2004.

[23] G. B. Dantzig, "Linear Programming and Extensions," Princeton University Press, Princeton, NJ, 1963.

[24] S. Boyd and L. Vandenberghe, "Convex Optimization," Cambridge University Press, United Kindom, 2004.