Distributed Link-aware Rate Allocation for R-D Optimal Multiple Video Streaming over Wireless Networks

Zejun Ma[1], Li Song[1], Cheng Zhi[1] and Libo Yang[2]

[1] Institute of Image Communication and Information Processing
Shanghai Jiao Tong University, Shanghai, China

[2] China Mobile Communications Corporation, Beijing, China
zjun.ma@gmail.com, {song_li, zhicheng}@sjtu.edu.cn, yanglibo@chinamobile.com

Abstract—**When multiple video streams are delivered in wireless networks, careful rate allocation is desirable to share the available bandwidth without overflowing the whole network or any individual link, as well as to maximize the total received quality among the competing streams. This paper presents a distributed rate allocation for Rate Distortion (R-D) optimal multiple video streaming in a video- and link-aware fashion. We take into account video R-D characteristics, as well as network and link congestion in the optimization framework, to strive for the minimal total video distortion of all participating streams while limiting network and individual link channel time utilization. *Initial Start* scheme addresses the initialization problem for video streaming effectively, and *Fast Recovery* scheme eliminates network and any individual link congestion as fast as possible. Global optimal rates can be obtained in a distributed manner. Results from numerical experiments confirm excellent performance of this distributed rate allocation.**

Keywords- ***distributed rate allocation; multiple video streaming; link-aware; link congestion; network congestion***

I. INTRODUCTION

As wireless networks have made significant advances in recent years, video streaming over wireless networks is becoming increasingly important. Due to its real-time nature, video streaming typically has bandwidth, delay and loss requirements. In wireless environments, the channel conditions change rapidly due to noise, interference from nearby transmitters, multi-path, and host movement, causing fluctuations in individual link capacity. In such a context, many challenges result from the limited bandwidth of wireless nodes, combined with the demanding rate and stringent delay requirements of video streaming.

There has been some prior work on video streaming. [4][5] propose the TCP congestion control mechanism to control the data rate, relying on packet losses as an implicit congestion signal from overloaded links. One common approach, such as TFRC [6], allocates the source rate at each node, depending on the end-to-end observations. However, all the above mechanisms are unaware of the R-D characteristics of various video sequences. Recently, Zhu et al. [7] present distributed media-aware rate allocation accommodating heterogeneity in wireless link speeds and video R-D characteristics, but this scheme is unaware of any individual link congestion in the network.

Some links may be in heavy congestion while the network is still under the available utilization. When multiple video streams are delivered in wireless networks, careful rate allocation is desirable to share the available bandwidth without overflowing the network or any individual link, as well as to maximize the total received qualities among the competing streams. We accommodate video R-D characteristics as well as network and link congestion in our optimization framework, and present distributed rate allocation in a video- and link-aware manner.

Our contribution is as follows: 1) we propose *Initial Start* scheme to well evaluate the initial link price to address the initialization problem, and *Fast Recovery* scheme to eliminate network and individual link congestion as fast as possible to avoid packet loss; 2) the scheme to accumulate the end-to-end path price for one stream is improved based on the *Cannikin Law.*

The rest of the paper is structured as follows: in the next section, we present optimization framework including wireless network model, video rate distortion model, and optimization objective which takes into account video R-D characteristics as well as network and link congestion; Section III proposes our distributed rate allocation algorithm, which introduces *Initial Start*, *Rate Growth* and *Fast Recovery* schemes in detail; Results from numerical experiment of various topology networks are illustrated in Section IV; Finally, conclusions are drawn in Section V.

II. OPTIMIZATION FRAMEWORK

A. Wireless Network Model

Consider wireless network with a set L of links, and let C_l be the finite capacity of Link $l \in L$. We denote the set of video streams by S, and associate Stream s ($s \in S$) with a user and one routing path P^s, and suppose that if rate R^s is allocated to the Stream s then this has distortion $D^s(R^s)$ to the user.

The collection of links traversed by Stream s constitutes its routing path: $P^s = \{\, l \mid Stream\ s\ traverses\ link\ l \,\}$

For Link l, we define the set of links whose source or destination node is within transmission range of either ingoing or outgoing node of Link l as its interference set: L_l. In a small network where all nodes are within carrier-sense range of each other, all links belong to the same interference set L.

We refer to some traffic flow over Link l which does not take part in the distributed rate allocation as background traffic

---

This work was supported in part by NSFC (60702044, 60632040, 60625103), MIIT of China(2010ZX03004-003) and 973 (2010CB731401,2010CB731406)

F_l'. Then the total flow traffic F_l over Link l is denoted as: $F_l = F_l' + \sum_{s:l\in P^s} R^s$

We denote the fraction of time that Link l is occupied by traffic F_l over the link capacity C_l as *channel time utilization:* $u_l = F_l / C_l$. Then the total channel time utilization over interference set L_l is as follows: $\bar{u}_l = \sum_{l'\in L_l} u_{l'}$, and the contribution from background traffics is denoted as: $\bar{u}_l' = \sum_{l'\in L_l} F_{l'}' / C_{l'}$, while the total contribution from Stream s is : $\bar{u}_l^s = \sum_{l'\in P^s \cap L_l} R^s / C_{l'}$.

The notation $\bar{C}^s$ introduces the concept of *equivalent capacity* for Stream s : $\bar{C}^s = \sqrt[N]{C_1 \bullet C_2 \bullet \ldots \bullet C_N}$, where C_n ($n \in \{1,2,\ldots,N\}$) is the capacity of links which Stream s traverses over.

All notations in this section are summarized in Table I.

TABLE I. SUMMARY OF NOTATIONS IN WIRELESS MODLE

Symbol	Description
l	index of wireless link
s	index of video stream
L	set of wireless links
S	set of video streams
P^s	set of links Stream s traverses over
L_l	interference set of Link l
C_l	capacity of Link l
$\bar{C}^s$	equivalent capacity of Stream s
R^s	allocated rate of Stream s
F_l'	background traffic over Link l
F_l	total flow traffic over Link l
u_l	channel time utilization of Link l
$\bar{u}_l$	total channel time utilization over L_l

B. Video Rate Distortion Model

Stuhlmüller et al. [1] derive a rate distortion (R-D) model, which is well suited for the average R-D behavior of the state-of-the-art video coders. For each Stream $s \in S$, we associate a mean square error (MSE) distortion D^s when the video stream is encoded at the rate R^s. The R-D characteristics of each stream can be fitted to a parametric model [1]:

$$D^s(R^s) = D_0^s + \frac{\theta^s}{R^s - R_0^s} \tag{1}$$

Where D_0^s, R_0^s and θ^s depend on the video content, coding scheme, as well as the encoder configuration, and can be estimated from empirical R-D curves using least-square methods.

C. Optimization Objective

Results in [8] show that MIN-MSE objective leads to allocations that most closely match subjective preference. To reduce the call blocking probability and call dropping probability of video streaming in wireless networks, it is frequent to employ network resource reservation to achieve the abovementioned goal as well as to avoid long latency of path rebuilding. Particularly, it is of great significance to set aside adequate bandwidth for a bottleneck link which several streams traverse over to provide a high quality guarantee for these streaming videos. The goal of our multi-stream rate allocation is to minimize the overall distortion of all video streams without overloading the wireless network and any individual link in the network.

$$\min_{R^s, s\in S} \quad \sum_{s\in S} \omega^s \bullet D^s(R^s) \tag{2}$$

$$s.t. \quad u_l = \frac{F_l' + \sum_{s:l\in P^s} R^s}{C_l} \le \alpha_l, \forall l \in L \tag{3}$$

$$\bar{u}_l = \bar{u}_l' + \sum_{s\in S} \bar{u}_l^s \le \beta_l, \forall l \in L \tag{4}$$

$$R_{\min}^s \le R^s \le R_{\max}^s, s \in S \tag{5}$$

In (2), the scaling factor ω^s indicates relative importance of each stream. In (3) and (4), the constraints limit the channel time utilizations of Link l and Link l 's interference set below the prescribed target $\alpha_l < 1 (l \in L)$ and $\beta_l < 1 (l \in L)$ to reserve certain bandwidth, respectively. The constraint (5) limits the allocated rate of each stream to the lower and upper rate bounds which the encoder can offer.

III. Distributed Rate Allocation Algorithm

While the optimization objective is mathematically fairly tractable in a centralized manner, it is desirable to adopt distributed rate allocation to share computation burden among all the participating nodes in a practical wireless network. This objective involves the R-D characteristics of various videos, which are unlikely to be known by the network. To solve this problem, we present a pricing policy, in which users pay a price to obtain a particular rate for streaming, and the network updates the link price in a given period.

We extend the dual rate control algorithm [3] to share available information between individual user and the network, and define our new link price update rules, using Zhu's congestion price update rules [7] for reference, to better reflect users' rate demands and network's resource supplies. Then the rate allocation problem is addressed by adjusting source video rate according to the path price for each stream.

A. Link Price Update

We define Link l price at time t as: $\lambda_l(t)$, and $\lambda_l(t-T)$ is the last link price, where T is the updating interval between two discrete time. The link price update rules lie in that various link price changing magnitudes and tendencies are adopted, depending on different network situations and congestion levels.

Here we calculate the end-to-end accumulated path price for Stream s :

$$\Lambda^s = \sum_{l\in P^s} \lambda_l \bullet \frac{\bar{C}^s}{C_l} \tag{6}$$

Where the contribution of the link price to the path price is weighted by the ratio of $\overline{C}^s / C_l$. The intuition behind this is that the link with the smallest capacity plays a critical role in determining the allocated rate for Stream s , in the same way that how much water a bucket can contain is determined by the shortest board, rather than the tallest board in the *Cannikin Law*. When all the links have the same capacity, the path price will be $\Lambda^s = \sum_{l \in P^s} \lambda_l$, only summing to all the link prices of all the links for Stream s , where all link prices have the same weight to the path price.

In general, the network states fall into three cases below.

Initial Start: user makes video streaming demand for the first time and is unknown of the network situation, then it is desirable to evaluate the link price for all the links which the stream traverses through.

Rate Growth: when neither the network nor any individual link is in congestion, a decreasing link price is needed to encourage higher rates from all contributing streams.

Fast Recovery: when network detects network or individual link congestion, *Fast Recovery* is of great importance to deal with any congestion as fast as possible. Large link price is needed to slow down the end-to-end allocated rate immediately.

1) *Rate Growth*

When neither the network nor Link l is in congestion, a decreasing link price is preferable to encourage higher streaming rate, and we update link price:

$$\lambda_l(t) = \max[\lambda_l(t-T) + K_1 \bullet (\alpha_l - u_l) \bullet (\bar{u}_l - \beta_l) \bullet C_l, 0], \quad s.t. \quad u_l - \alpha_l \le 0, \bar{u}_l - \beta_l \le 0, l \in L \tag{7}$$

In (7), K_1 is updating coefficient. The intuition behind (7) is that as the utilizations of network and Link l approach their prescribed targets, the decrement of $\lambda_l(t)$ will become smaller in every iteration, which helps to converge algorithm due to fewer rate fluctuations around the prescribed channel time utilization .

2) *Fast Recovery*

Video streaming is extremely sensitive to packet loss. Some important frames such as I frame, will make neighboring frames undecodable if lost, and user's viewing quality will suffer severely from this kind of loss. It is desirable to get rid of any congestion as fast as possible to avoid packet loss. Congestion occurs due to overwhelming the network or any individual link.

a) When both network and Link l are in congestion, we have:

$$\lambda_l(t) = \max[\lambda_l(t-T) + p_1 \bullet (u_l - \alpha_l) \bullet (\bar{u}_l - \beta_l) \bullet C_l, 0], \quad s.t. \quad u_l - \alpha_l > 0, \bar{u}_l - \beta_l > 0, l \in L \tag{8}$$

b) When only the network is in congestion, the link price is updated as:

$$\lambda_l(t) = \max[\lambda_l(t-T) + p_2 \bullet (\bar{u}_l - \beta_l) \bullet C_l, 0], \quad s.t. \quad u_l - \alpha_l \le 0, \bar{u}_l - \beta_l > 0, l \in L \tag{9}$$

c) When Link l is in congestion, we obtain:

$$\lambda_l(t) = \max[\lambda_l(t-T) + p_3 \bullet (u_l - \alpha_l) \bullet C_l, 0], \quad s.t. \quad u_l - \alpha_l > 0, \bar{u}_l - \beta_l \le 0, l \in L \tag{10}$$

In (8)(9)(10), p_1 , p_2 , p_3 are punishing coefficients which always are large enough to dispel congestion in as fewer iterations as possible. The fewer iterations to take to recover normal network status, the fewer packet losses there will be. However, it is crucial to choose proper coefficients, since too small coefficients will not achieve *fast recovery* while too lager ones will make it difficult for algorithm to converge.

3) *Initial Start*

For user streaming for the first time, he may allocate any available rate because of being unknown of network states. If the allocated rate is so low that neither network congestion nor link congestion occurs, we obtain $\lambda_l = 0$ $(l \in P^s)$ using (7), which makes Λ^s zero according to (6). Then we can't calculate the next optimal allocated rate for Stream s by (13). Therefore it is necessary to deal with this kind of initialization problem by defining a particular link price initialization.

Here is the solution: Firstly, for the first allocated rate R^s , using (13) by substituting R^s_{opt} with R^s , we calculate the path price $\Lambda^s = \dfrac{\omega^s \bullet \theta^s}{(R^s - R^s_0)^2}$; Secondly, we obtain $\lambda_l(t)$ $(l \in P^s)$ by the following defined formula:

$$\lambda_l(t) = \frac{\Lambda^s}{\overline{C}^s} \bullet \frac{1}{\sum_{l' \in P^s} 1/C_{l'}} \bullet \frac{C_l}{\overline{C}^s}, \; l \in P^s \tag{11}$$

Given the initial rate, it is preferable to allocate next rate no smaller than the current allocated rate for Stream s . In view of the fact that the allocated rate for Stream s is calculated in inversely proportion to the path price of Stream s , the next path price $\Lambda^s(t+T)$ for Stream s should be no bigger than the current path price $\Lambda^s(t)$. One key advantage of our defined formula (11) is that it naturally ensures our expected relation, which is proved as follows.

$$\Lambda^s(t+T) = \sum_{l \in P^s} \lambda_l(t+T) \bullet \frac{\overline{C}^s}{C_l} = \sum_{l \in P^s} \left(\frac{\Lambda^s(t)}{\overline{C}^s} \bullet \frac{1}{\sum_{l' \in P^s} 1/C_{l'}} \bullet \frac{C_l}{\overline{C}^s}\right) \bullet \frac{\overline{C}^s}{C_l}$$

$$= \left(\frac{\Lambda^s(t)}{\overline{C}^s} \bullet \frac{1}{\sum_{l' \in P^s} 1/C_{l'}}\right) \bullet \sum_{l \in P^s} \left(\frac{C_l}{\overline{C}^s} \bullet \frac{\overline{C}^s}{C_l}\right) = \left(\frac{\Lambda^s(t)}{\sqrt[N]{C_1 \bullet C_2 \bullet \ldots \bullet C_N}} \bullet \frac{1}{1/C_1 + 1/C_2 + \ldots + 1/C_N}\right) \bullet N$$

$$\le \left(\frac{\Lambda^s(t)}{\sqrt[N]{C_1 \bullet C_2 \bullet \ldots \bullet C_N}} \bullet \frac{1}{N \bullet \sqrt[N]{1/(C_1 \bullet C_2 \bullet \ldots \bullet C_N)}}\right) \bullet N = \Lambda^s(t)$$

When $C_1 = C_2 = \ldots = C_N$, $\Lambda^s(t+T) = \Lambda^s(t)$. □

B. *Video Rate Allocation*

In this section, we extend the decomposition scheme proposed in [2], and the user-end's optimization problem is translated into the following problem, similar with Zhu's [7]:

$$R^s = \begin{cases} R^s_{opt} = \arg\min_R[\omega^s \bullet D^s(R) + \Lambda^s \bullet R], R^s_{\min} \le R^s_{opt} \le R^s_{\max} \\ R^s_{\min}, & , R^s_{opt} < R^s_{\min} \\ R^s_{\max}, & , R^s_{\max} < R^s_{opt} \end{cases} \quad (12)$$

Combined with (1), we have the optimal rate R^s_{opt} from (12) as:

$$R^s_{opt} = R^s_0 + \sqrt{\frac{\omega^s \bullet \theta^s}{\Lambda^s}} \quad (13)$$

It is obvious that *Section A* and *B* naturally solve the optimization problem in a distributed manner. Information exchange among local neighboring nodes helps to update link price, and the optimal allocated rate R^s_{opt} is calculated based on only the path price Λ^s and R-D parameters of each Stream s. Iterations between these two steps end with matching user rate demand and network resource supply.

IV. NUMERICAL EXPERIMENT

In this section, numerical experiment is carried out for various topology networks to illustrate our distributed rate allocation algorithm. Fig. 1 shows the wireless network topologies. We consider single-hop parallel, multi-hop symmetrical and multi-hop topologies, and assume that all the links are within the transmitting range of each other. The Harbour and Crew HD sequences are encoded using $x264$, with GOP length of 30 at 60 fps, and are delivered in the three topology networks in Fig. 1.

The configuration is as follows: The prescribed utilization β[1] of interference set is 95%, while the ones of links are set to 80% except that the limited utilization α_1 of Link 1 varies from 20 % to 80%; The coefficients are set: $K_1 = 0.015$, $p_1 = 10$, $p_2 = 2$, $p_3 = 12$.

We illustrate our distributed rate allocation has link-aware performance and video quality improvement in the following two parts: multi-stream link utility and video quality.

TABLE II. COMPARISON OF LINK 1 UTILIZATION BETWEEN OUR AND ZHU'S DISTRIBUTED SOLUTION

target α_1 (%)		20	30	40	50	60	70	80
experiment α_1 (%)	Zhu's	81.5	81.5	81.5	81.5	81.5	81.5	81.5
	Proposed	20.0	30.0	34.8	50.0	50.1	50.2	50.4

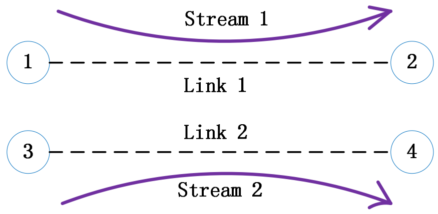
a. Single-hop Parallel Topology

target α_1 (%)		20	30	40	50	60	70	80
experiment α_1 (%)	Zhu's	70.6	70.6	70.6	70.6	70.6	70.6	70.6
	Proposed	20.0	30.0	40.0	46.2	47.1	48.0	49.0

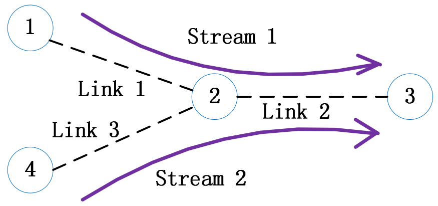
b. Multi-hop Symmetrical Topology

target α_1 (%)		20	30	40	50	60	70	80
experiment α_1 (%)	Zhu's	62.2	62.2	62.2	62.2	62.2	62.2	62.2
	Proposed	20.0	29.8	39.9	38.4	23.5	28.0	34.7

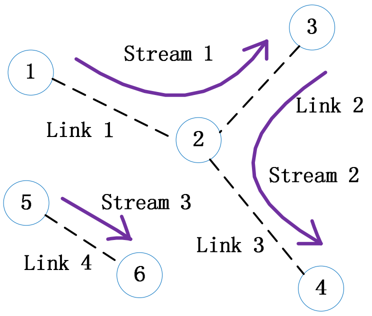
c. Multi-hop Topology



(a) Single-hop Parallel Topology



(b) Multi-hop Symmetrical Topology



(c) Multi-hop Topology

Figure 1: Wireless Network topologies. In example wireless networks, Link 1 has capacity of 5 Mbps, while Link 2, Link 3, and Link 4 have capacity of 50 Mbps. The HD sequence Harbour traverses over Stream 1, while the HD sequence Crew traverses over Stream 2 and Stream 3.

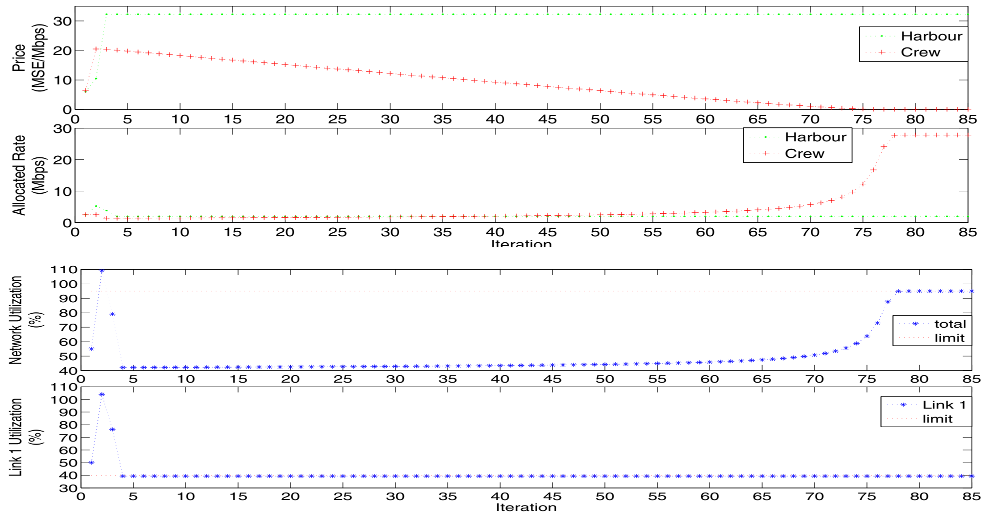
A. *Multi-stream Link Utility*

Fig. 2 shows traces of path prices, allocated rates, network utilization, and Link 1 utilization for three wireless network topologies in Fig. 1. The allocated rates for each stream are diverse relying on their respective streaming path prices and R-D parameters. It can be noted from Fig. 2(a) that network utilization recovers from network congestion in only one iteration, and from Fig. 2(a) and 2(c) that Link 1 utilization falls below the limited target α_1 in three and one iteration respectively, due to *Fast Recovery* scheme. Fig. 2(b) shows that network utilization gradually approaches $\beta = 95\%$ while Link 1 utilization changes slightly, which reflects that channel time utilization is shifted from bottleneck Link 1 to other idle links.

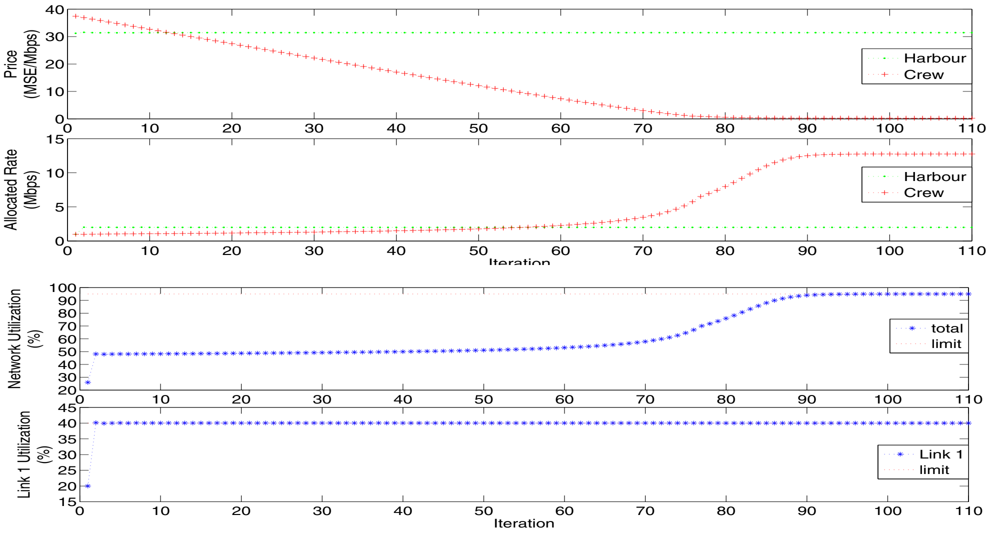
Link 1 has the smallest capacity, and take bottleneck Link 1 for example. Table II compares Link 1 utilization in numerical experiment between our distributed allocation algorithm and Zhu's distributed solution [7] for various network topologies in Fig. 1. It shows that Zhu's solution achieves the same Link 1 utilization α_1, irrespective of the various given target α_1, which incurs packet loss when Link 1 is in congestion, whereas our solution is aware of link congestion, and always keeps Link 1 utilization α_1 under its limited target.

---

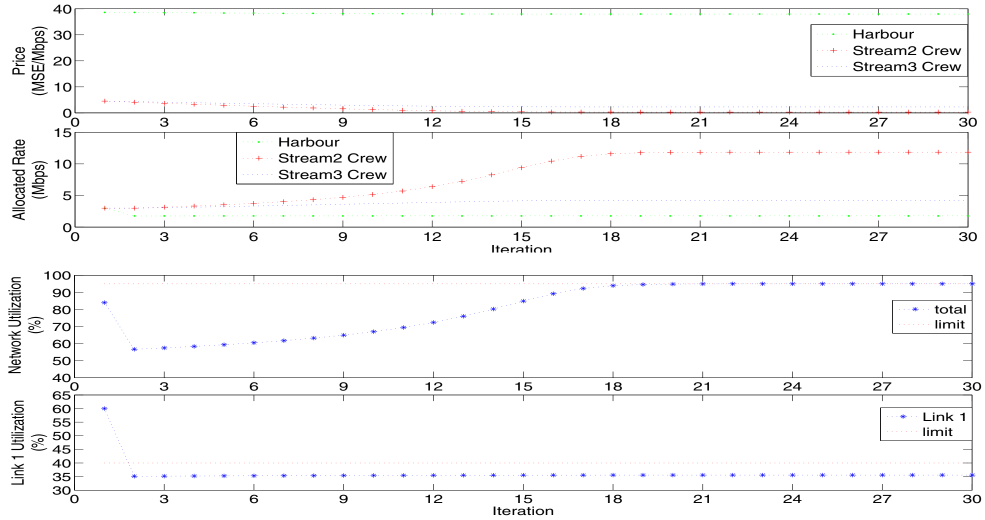
[1] Subscript is deleted, since there is only one interference set where all the links are within the transmitting range of each other.



(a) Single-hop Parallel Topology

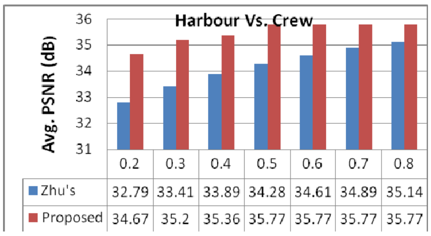


(b) Multi-hop Symmetrical Topology

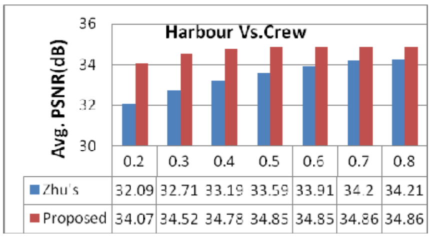


(c) Multi-hop topology

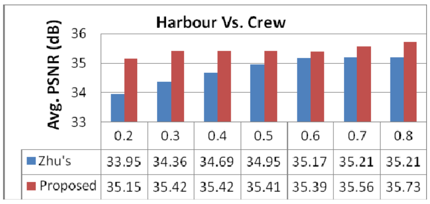
Figure 2: Traces of path prices, allocated rates, network utilization, and Link 1 utilization over these three topologies. All the link limited channel time utilizations are 80% expect that Link 1 limited utilization $\alpha_1 = 40\%$.



(a) Single-hop Parallel Topology



(b) Multi-hop Symmetrical Topology



(c) Multi-hop Topology

Figure 3: Comparison of video average PSNR performances between the proposed distributed allocation scheme and Zhu's distributed solution. Axis of abscissa is the change of Link 1 limited channel time utilization α_1 .

B. *Video Quality*

We compare video quality of the proposed distributed rate allocation algorithm against Zhu's distributed solution [7] in terms of average PSNR of all participating streams from the perspective of the whole network. When the allocated rate R^s exceeds the prescribed bandwidth (say, $\alpha_l \bullet C_l$) of Link l from the Stream s path, we calculate the end-to-end video distortion by substituting R^s with $\alpha_l \bullet C_l$ according to (1), under the assumption that the decoder can deal with any packet loss. It is easy to figure that this calculated PSNR performance is the highest achieved video quality when Stream s is delivered at the allocated rate R^s . We obtain Zhu's *equivalent PSNR* for the Stream s which traverses through congested link in this way.

Fig. 3 compares average PSNR between the proposed distributed algorithm and Zhu's distributed solution for three wireless network topologies in Fig. 1. As shown in Fig. 3, improvement of the proposed distributed algorithm against Zhu's distributed solution ranges between 0.63–1.88 in PSNR for single-hop parallel topology, between 0.65–1.98 in PSNR for multi-hop symmetrical topology, and between 0.22–1.2 in PSNR for multi-hop topology:

1) An average PSNR gain is achieved for video streaming over the network without congested link (when α_1 is 80% in experiment with multi-hop symmetrical topology, and 80% or 70% in experiment with multi-hop topology, seen from Table II b and c, respectively), due to our link price update scheme and accumulating path price reasonably;

2) The average PSNR gain is more pronounced for video streaming over the network with much heavier congested links. Zhu's solution is irrespective of link congestion which incurs packet loss, reducing the video quality for loss-sensitive video streaming, whereas the proposed distributed rate allocation algorithm shifts the network utilization from congested links to idle links, improving the total network video quality.

V. CONCLUSION

We present a distributed rate allocation in a video- and link-aware fashion, which aims to share the available bandwidth without overloading the whole network or any individual link, as well as to maximize the total received quality among the competing streams. *Initial Start*, *Rate Growth*, *Fast Recovery* schemes are proposed to deal with various network situations. The scheme to accumulate the path price for one stream is improved based on the *Cannikin Law*. Results from numerical experiments confirm excellent performance of our distributed rate allocation.

REFERENCES

[1] K. Stuhlmüller, N. Färber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," IEEE Journal on Selected Areas in Communications, vol. 18, no. 6, pp. 1012–32, June 2000.

[2] F. P. Kelly, "Charging and Rate Control for Elastic Traffic," European Transactions On Telecommunications, vol. 8, no. 1, pp. 33–37, 1997.

[3] 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.

[4] V. Jacobson, "Congestion avoidance and control," in Proc. ACM Conference on Communications Architectures and Protocols, vol. 18, no.4 , pp. 157–173, Standford, CA, USA, Aug. 1988.

[5] M. Allman, V. Paxson, and W. R. Stevens, "TCP Congestion Control," RFC 2581 (Proposed Standard), Updated by RFC3390, Apr. 1999.

[6] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "TCP Friendly Rate Control (TFRC): Protocol Specification," RFC 3448, Jan. 2003.

[7] X. Zhu, B. Girod, "Distributed Media-Aware Rate Allocation for Wireless Video Streaming," IEEE Transactions on Circuits and Systems for Video Technology. Submitted August 2009.

[8] M. Kalman, B. Girod, "Optimal channel-time allocation for the transmission of multiple video streams over a shared channel," in Proc. IEEE 7th Workshop on Multimedia Signal Processing , Shanghai, China, pp.1-4, Oct.2005.