Distributed Media-Aware Rate Allocation for Wireless Video Streaming

Xiaoqing Zhu and Bernd Girod, *Fellow, IEEE*
Information Systems Laboratory, Stanford University, Stanford, CA 94305, USA

***Abstract*—When multiple video streams share a wireless network, careful rate allocation is needed to prevent congestion, as well as to balance the video qualities among the competing streams. In this paper, we present a distributed media-aware rate allocation protocol, and evaluate its performance for streaming of high-definition (HD) video over 802.11-based wireless home networks.**

Our optimization framework accommodates heterogeneity in wireless link speeds and video rate-distortion (RD) characteristics, as well as traffic contention among neighboring links. The goal of the protocol is to minimize the total video distortion of all participating streams while limiting network utilization. It relies on cross-layer information exchange between video rate controllers at the end hosts and link state monitors at the relaying wireless nodes. Results from various network simulations confirm that the media-aware allocation outperforms TCP-Friendly Rate Control (TFRC) in terms of average video quality and fairness among the streams.

***Index Terms*—distributed rate allocation, video streaming, wireless networking, cross-layer design**

I. Introduction

Video streaming over wireless networks is compelling for many applications, ranging from home entertainment to video surveillance, to audiovisual communication for search-and-rescue operations. Technical challenges abound. The wireless radio channel is subject to interference from nearby transmitters, multipath fading, and shadowing, causing fluctuations in link throughput and sometimes an error-prone communication environment. The traffic patterns of compressed video streams typically change dynamically due to content variations and user behavior, and the received video quality may degrade drastically in the presence of packet losses, due to error propagation in the compressed bitstream. Moreover, video streaming applications typically have high data rates and stringent latency requirements, at odds with the limited bandwidth resources in a wireless network.

When multiple video streaming sessions simultaneously share a wireless network, an effective mechanism for rate allocation among these streams is indispensable to avoid excessive network congestion and to achieve best possible video qualities. Since neighboring links compete for the same wireless radio channel, the rate of a video stream will not only affect the links along its own path, but also contend with traffic over other nearby links. The task of rate allocation is further complicated by heterogeneity in both the rate utilities of video streams and in wireless link qualities. Lack of centralized control in a wireless network also requires that the task be performed in a distributed manner.

In this paper, we present an optimization framework for distribute video rate allocation over wireless, taking into account the above challenges. We focus on streaming pre-encoded video contents over wireless networks consisting of static nodes. In our framework, a wireless network model explicitly captures the effect of traffic contention among neighboring links and heterogeneous link transmission speeds. A parametric video distortion model represents the utility of allocated rate for each stream. Multi-stream rate allocation is formulated as a convex optimization problem, with the goal of minimizing total video distortion while avoiding excessive network utilization. We show that the globally optimal solution is achievable via distributed computation, by iteratively updating video source rates and link congestion prices. We further analyze the dynamics of the proposed distributed solution, and establish system stability under proper parameter choices.

Based on the optimization framework, we then design a distributed media-aware rate allocation protocol to overcome issues found in a practical system. Instead of revamping the entire protocol stack, the proposed scheme leverages cross-layer information exchange between video rate controllers at the end hosts and link state monitors at the relaying wireless nodes. Consequently, each video sender can regulate its rate according to explicit congestion prices accumulated along its path, and can quickly adapt to changes both in wireless network conditions and in video characteristics. Simulation studies of the protocol show that it leads to lower average video distortion and more balanced qualities among the streams than conventional media-oblivious schemes such as TCP-Friendly Rate Control (TFRC) [2].

The rest of this paper is organized as follows. The next section reviews related research in multi-stream rate allocation. Section III presents our optimization framework and stability analysis of the distributed solution. In Section IV, we explain our design of the distributed media-aware rate allocation protocol, based on cross-layer information exchange between video end hosts and wireless nodes. Section V evaluates the protocol performance in various network simulations involving high-definition (HD) video streaming over various 802.11 networks.

II. Related Work

Rate allocation among multiple flows that share a network is an important and well-studied problem. Internet applications typically use the TCP congestion control mechanism for regulating their outgoing rates [3] [4]. Another common approach is to stream video over UDP [5] and RTP [6], while adapting its source rate acccording to the target rate calculated by an equation-based congestion control mechanism, such as TCP-Friendly Rate Control (TFRC) [2] [7]. With this approach, the target rate computed by TFRC is used to adapt the source rate of the video streams [8] [9] [10]. Several modifications have been proposed to improve the media-friendliness of both TCP congestion control and TFRC [11] [12] [13].

Kelly *et al.* have presented a general mathematical framework to incorporate flows with different utilities in distributed pricing-based rate control algorithms [14]. Application of such algorithms has been investigated for elastic traffic over the Internet [15] [16], with extensions for multicast sessions [17]. Differing in details, they all achieve the goal of utility maximization over a network, as summarized by the survey paper [18]. Our work extends Kelly's framework to wireless networks, by considering overlapping interference sets as

Some preliminary results from this work were presented at the Picture Coding Symposium, May 2009, Chicago, IL, USA [1]. The authors can be contacted via email at zhuxq@stanfordalumni.org and bgirod@stanford.edu.

shared resources. Furthermore, our design of the packet-level rate allocation protocol addresses many practical issues not considered in the original mathematical framework.

One of the main challenges in adopting the utility maximization framework for wireless video streaming is to define a good optimization objective to achieve the best tradeoff between individual stream quality and fairness among all participating streams [19] [20] [21]. Subjective test results for multi-stream video distribution over a 802.11 wireless LAN suggest that allocations obtained by minimizing the total MSE distortion of all streams agree sufficiently well with subjective preferences [22]. Results from our own subjective viewing tests confirm this finding [23].

In the context of wireless networking using the 802.11 MAC protocol, many performance issues have been identified [24]. It has been pointed out in [25] that the presence of one stream traversing a slow link significantly reduces the throughput achieved by other streams traversing faster links. Certain network topologies also lead to severe unfairness in throughput among competing TCP flows, in some extreme cases causing complete flow starvation in the middle of the network [26]. The main reason for such anomaly is that TCP relies only on end-to-end observations for network congestion indications and cannot correctly infer the impact of a streams on other streams traversing nearby links [27]. In this paper, we show that the above performance issues can be avoided by explicitly accounting for traffic contention among competing streams when designing the rate allocation protocol.

III. Optimization Framework

A. Wireless Network Model

Consider a network with a set of wireless nodes $\mathcal{N} := \{1, \ldots, N\}$, and a set of video streams denoted by $\mathcal{S} := \{1, \ldots, S\}$. The collection of links between neighboring node pairs is designated by:

$$\mathcal{L} := \{l_{i\to j} \mid \text{Node } j \text{ is within transmission range of Node } i\}. \quad (1)$$

The collection of links traversed by Stream s constitutes its path $\mathcal{P}^s$.

In this work, we assume that transmission errors are effectively remedied by error control techniques such as adaptive modulation, channel coding, and retransmissions [28]. Consequently, the wireless channel exhibits time-varying throughput over each link, with no residual packet losses observed at the application layer [29] [30]. [1] Throughput of Link l is denoted by C_l, and is defined as the payload transmission rate over that link *without* contention from traffic over any other links. In other words, C_l is the maximum achievable data rate over Link l when the rest of the network is silent. In 802.11 networks, the value of C_l depends on the link transmission speed, which, in turn, is dictated by the wireless channel condition over that link. The throughput C_l is also affected by the average payload size and the percentage of control overhead due to MAC-layer header, ACK packets, and, in the case of multi-hop topologies, RTS/CTS handshakes.[2]

Our formulation also allows for some traffic that does not follow the proposed rate allocation scheme. We refer to such traffic as *background traffic*. Let F_l' represent the rate of background traffic

[1] While this assumption is valid for wireless networks with static nodes in our setting, it may not hold in networks with mobile nodes and frequent topological changes. We leave the exploration of mobile networks as future research.

[2] While the analysis and simulations are confined to 802.11 networks throughout this paper, we note that our wireless network model is generic enough to accommodate other types of networks in which neighboring links time-share their transmission opportunities.

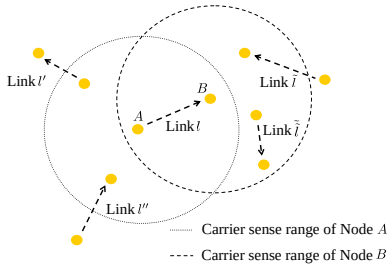


Fig. 1. Illustration of interference set. Given Link l and its source and destination nodes A and B, all links with source or destination node within transmission range of Node A or B constitute interference set $\mathcal{L}_l$.

over Link l and R^s designate the rate of Stream s, the traffic rate over Link l is:

$$F_l = F_l' + \sum_{s:l\in\mathcal{P}^s} R^s. \quad (2)$$

We define *channel time utilization* as the fraction of time that a link is occupied by traffic of rate F_l over its throughput C_l: $u_l = F_l/C_l$. This can be further decomposed as: $u_l = u_l' + \sum_s u_l^s$, with contribution from background traffic $u_l' = F_l'/C_l$, and contribution from each video stream:

$$u_l^s = \begin{cases} \frac{R^s}{C_l}, & l \in \mathcal{P}^s \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

According to (3), the same amount of traffic gives rise to a higher channel time utilization over a slower link than over a faster link.

When two nearby links l and l' cannot transmit at the same time, they are deemed to *interfere with* each other, denoted as $l \bowtie l'$. In 802.11 networks, successful transmission of each packet requires successful transmission of both the data packet and the acknowledgement (ACK) packet. Therefore, for Link $l_{i\to j}$, we define the set of links whose source or destination node is within transmission range of either Node i or Node j as its *interference set* $\mathcal{L}_l$. Note that interference among links is reciprocal, i.e., $l' \bowtie l$ if and only if $l \bowtie l'$. And, by definition, $l' \bowtie l, \forall l' \in \mathcal{L}_l$. Figure 1 illustrates the concept of interference set. In this example network, data packet transmission of Link l conflicts with data packet transmission of Link l', therefore Link l' is within interference set of Link l. Similarly, Link l'' is also in the interference set of Link l, since data packet transmission of Link l'' conflicts with ACK packet transmission of Link l. In a small network where all nodes are within carrier-sense range of each other, all links belong to the same interference set $\mathcal{L}$. In general, a network comprises multiple overlapping interference sets.

For all links within the interference set $\mathcal{L}_l$, total channel time utilization $\tilde{u}_l$ can be computed as: $\tilde{u}_l = \sum_{l'\in\mathcal{L}_l} u_{l'}$. To avoid overloading the network, it is important to constrain $\tilde{u}_l < \gamma$, where $\gamma < 1$ is the target total utilization. The extra headroom is needed to absorb various effects not included in our model, such as random backoff in a CSMA/CA network for resolving link contentions, or inaccurate estimates of instantaneous link throughput.

We use the shorthand:

$$\tilde{u}_l^s = \sum_{l'\in\mathcal{L}_l} u_{l'}^s = \sum_{l'\in\mathcal{L}_l\cap\mathcal{P}^s} \frac{R^s}{C_{l'}} = \frac{R^s}{\tilde{C}_l^s} \quad (4)$$

to denote the total contribution of channel time utilization from Stream s to interference set $\mathcal{L}_l$. Here, the notation $\tilde{C}_l^s$ introduces the concept of *equivalent throughput*:

$$\tilde{C}_l^s = \frac{1}{\sum_{l'\in\mathcal{L}_l\cap\mathcal{P}^s} \frac{1}{C_{l'}}}, \quad (5)$$

Symbol	Explanation
s	index of video stream
l	index of wireless link
i, j	index of wireless node
$\mathcal{S}$	set of video streams
$\mathcal{N}$	set of wireless nodes
$\mathcal{L}$	set of wireless links
$\mathcal{P}^s$	path of Stream s
$\mathcal{L}_l$	interference set of Link l
$\mathcal{L}_{\mathcal{P}^s}$	extended path of Stream s
C_l	throughput of Link l
F'_l	rate of background traffic over Link l
F_l	total traffic rate over Link l
R^s	rate of Stream s
u_l^s	channel time utilization of Stream s over Link l
u'_l	channel time utilization of background traffic over Link l
u_l	total channel time utilization over Link l
$\tilde{u}_l$	total channel time utilization in $\mathcal{L}_l$
$\tilde{u}'_l$	total channel time utilization due to background traffic in $\mathcal{L}_l$
$\tilde{u}_l^s$	total channel time utilization due to Stream s in $\mathcal{L}_l$
$\tilde{C}_l^s$	equivalent throughput for Stream s in $\mathcal{L}_l$

TABLE I
SUMMARY OF NOTATIONS FOR THE WIRELESS NETWORK MODEL.

summarizing the impact of R^s over $\mathcal{L}_l$. Similarly, the total channel time utilization attributed to background traffic can be calculated as:

$$\tilde{u}'_l = \sum_{l' \in \mathcal{L}_i} u'_{l'} = \sum_{l' \in \mathcal{L}_l} \frac{F'_{l'}}{C_{l'}}. \tag{6}$$

Total utilization over interference set $\mathcal{L}_l$ can therefore be decomposed as:

$$\tilde{u}_l = \tilde{u}'_l + \sum_{s \in \mathcal{S}} \tilde{u}_l^s. \tag{7}$$

For ease of notation later on, we designate the set of links *affected* by Stream s as its *extended path*:

$$\mathcal{L}_{\mathcal{P}^s} = \bigcup_{l \in \mathcal{P}^s} \mathcal{L}_l. \tag{8}$$

It consists of links directly traversed by the stream as well as nearby links that *overhear and compete with* Stream s.

Table I summarizes all notations introduced in this section.

B. Video Distortion Model

For each encoded video stream $s \in \mathcal{S}$, its distortion D^s is measured as the Mean Squared Error (MSE) between original and reconstructed pixel values. The MSE distortion typically decreases nonlinearly with increasing rate R^s, and can be fitted to a parametric rate-distortion (RD) model [31]:

$$D^s(R^s) = D_0^s + \frac{\theta^s}{(R^s - R_0^s)}. \tag{9}$$

The parameters D_0^s, θ^s and R_0^s are affected by factors including the coding scheme, the encoder configuration, and the video scene complexity. They need to be updated periodically to track time-varying video contents. A natural choice for the update period is the duration of a Group Of Pictures (GOP), between 0.5 – 2.0 seconds.

For each GOP within a video stream, the RD parameters D_0^s, R_0^s, and θ^s can be estimated from three or more trial encodings using nonlinear regression techniques as in [31]. Alternatively, one can apply least-square methods to estimate them based on a collection of empirical RD pairs: $(R_1^s, D_1^s), \ldots, (R_k^s, D_k^s), \ldots, (R_K^s, D_K^s)$. From (9), we have: $\theta^s = (R^s - R_0^s)(D^s - D_0^s)$. Consequently, the following equation holds for any k and k':

$$(D_k^s - D_{k'}^s)R_0^s + (R_k^s - R_{k'}^s)D_0^s = D_k^s R_k^s - D_{k'}^s R_{k'}^s. \tag{10}$$

The values of D_0^s and R_0^s can be obtained by solving the following set of equations in the least squares sense:

$$(D_k^s - D_{k+1}^s)R_0^s + (R_k^s - R_{k+1}^s)D_0^s = D_k^s R_k^s - D_{k+1}^s R_{k+1}^s, \quad \text{for } k = 1, \ldots, K-1. \tag{11}$$

The value of θ^s can then be estimated as:

$$\theta^s = \frac{1}{K} \sum_{k=1}^{K} (D_k^s - D_0^s)(R_k^s - R_0^s). \tag{12}$$

In this paper, rate adaptation is achieved by means of bitstream switching among multiple pre-encoded versions of each video sequence. As a result, only a finite set of available encoded rates are available for each stream. Note, however, that the optimization framework is general enough to accommodate other types of RD models, as long as they are convex.

C. Optimization Objective

The goal of multi-stream rate allocation is to maximize the overall viewing experience of all video streams without overloading the wireless network. Based on the subjective viewing test results described in [23], we choose to minimize the weighted sum of MSE distortion of all streams:

$$\min_{R^s, s \in \mathcal{S}} \quad \sum_{s \in \mathcal{S}} w^s D^s(R^s) \tag{13}$$

$$\text{s. t.} \quad R^s > R_{min}^s, \qquad \forall s \in \mathcal{S} \tag{14}$$

$$R^s < R_{max}^s, \qquad \forall s \in \mathcal{S} \tag{15}$$

$$\tilde{u}_l = \tilde{u}'_l + \textstyle\sum_{s \in \mathcal{S}} \frac{R^s}{\tilde{C}_l^s} < \gamma, \quad \forall l \in \mathcal{L}. \tag{16}$$

In (13), the scaling factor w^s indicates relative importance of each stream. The constraints (14) and (15) indicate the upper and lower bounds of the allocated video rates, corresponding to the highest and the lowest quality versions of the encoded bitstreams. The last constraint (16) limits the total channel time utilization within each interference set below a prescribed target $\gamma < 1$.

One can easily verify that the optimization in (13) – (16) has a convex objective function with linear constraints. We only consider the case when the problem is feasible, and assume that the case where streams at minimum rates still violate the total channel utilization constraint is handled separately by some access control mechanism. If all link states and all video RD parameters were available at a central entity, the optimal solution could be calculated using standard numerical techniques [32]. However, the overhead in collecting such global information may not scale well with growing network size or stream density. In practice, therefore, a distributed solution is preferable.

D. Distributed Solution

We now extend the pricing-based rate control algorithm proposed in [14] to support video streaming over wireless networks. The original optimization objective of maximizing generic logarithmic utility functions is substituted with minimizing video-specific parametric RD functions. In addition, the resource constraints now correspond to total channel utilization within each interference set, instead of total rate over individual links in a wired network. The algorithm can be decomposed into the following two iterative steps.

1) Congestion Price Update: A non-negative congestion price λ_l is associated with each interference set $\mathcal{L}_l$. Its value is updated periodically, according to instantaneous residual rate over Link l:

$$\lambda_l(t) = \max[\lambda_l(t - \tau) + \kappa(\tilde{u}_l - \gamma)C_l, 0]. \tag{17}$$

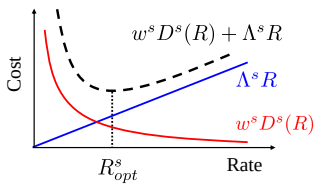


Fig. 2. The allocated video rate R^s_{opt} is determined by the end-to-end accumulated congestion price Λ^s, by the weight of importance w^s, and by the video RD function $D^s(R)$.

In (17), τ indicates the price update interval and κ is a scaling factor controlling the update step sizes.[3] For $\lambda_l > 0$, the price update is proportional to instantaneous *excess* total utilization over $\mathcal{L}_l$ and link throughput C_l. The intuition behind this is that λ_l increases if total channel time utilization $\tilde{u}_l$ temporarily exceeds the specified limit γ in order to induce rate reduction by all streams affecting $\mathcal{L}_l$. Conversely, as long as $\tilde{u}_l$ is below the target γ, the corresponding congestion price should keep decreasing to encourage higher rates from all contributing streams.

2) Video Rate Update: The rate of Stream s is updated as:

$$R^s_{opt} = \underset{R}{\operatorname{argmin}}[w^s D^s(R) + \Lambda^s R] = R^s_0 + \sqrt{\frac{w^s \theta^s}{\Lambda^s}}, \quad (18)$$

where $D^s(R)$ follows the parametric video RD model (9) and

$$\Lambda^s = \sum_{l \in \mathcal{L}_{\mathcal{P}^s}} \frac{\lambda_l C_l}{\tilde{C}^s_l} \quad (19)$$

is the end-to-end accumulated price. Note that the contribution of congestion price from each link is weighted by the ratio $C_l/\tilde{C}^s_l$. When Link l is the only link traversed by Stream s within its own interference set, $\tilde{C}^s_l = C_l$ and these two terms cancel out. Otherwise, $\tilde{C}^s_l < C_l$ when Stream s affects multiple links within $\mathcal{L}_l$; in this case the contribution of congestion price λ_l for $\mathcal{L}_l$ is inflated by a factor of $C_l/\tilde{C}^s_l = 1 + \sum_{l' \neq l, l' \in \mathcal{L}_l} C_l/C_{l'}$.

As shown in Fig. 2, the end-to-end congestion price Λ^s determines the slope of the linear term counterbalancing the RD tradeoff. Intuitively, a higher video rate is encouraged when the end-to-end accumulated congestion price Λ^s is lower. When the network becomes congested, increase in congestion prices at bottleneck interference sets leads to a higher accumulated price Λ^s and subsequently lower allocated video rates.

Note that the iteration between (17) and (18) naturally constitutes a distributed algorithm: the congestion price update only requires local observation of $\tilde{u}_l$ while the video rate update depends only on the end-to-end accumulated price Λ^s and the video RD parameters of individual streams. At equilibrium, the congestion prices satisfy the following:

$$\begin{cases} \lambda_l > 0, & \tilde{u}_l = \gamma, \text{ or} \\ \lambda_l = 0, & \tilde{u}_l < \gamma. \end{cases} \quad (20)$$

In other words, only congestion prices of fully utilized interference sets have strictly positive values. This corresponds to the Karush-Kuhn-Tucker (KKT) conditions for the constraint (16) in the original optimization problem, thereby guaranteeing optimality of the solution [33].

[3]It will become clear later in (18)-(19) that the unit of the congestion price λ should be MSE distortion divided by rate. Consequently, the unit of κ should be MSE distortion divided by rate squared. From now on, we will denote the unit of λ with MSE/Mbps, and that of κ with MSE/Mbps^2.

E. Stability Analysis

We now investigate stability of the proposed distributed solution based on a continuous flow model, for the simplest case of a single video stream traversing a single bottleneck link. Our analysis extends to the more general case of multiple video streams over multiple bottleneck interference sets, with details given in Appendix A. Moreover, it can be shown that in networks with low round trip times the system is globally stable. In other words, it is guaranteed that *any* trajectory following the system dynamics will converge to the equilibrium point. The proof of global stability is presented in [34].

Consider a single video stream traversing a single wireless link with throughput C.[4] If $\gamma C < R_{min}$, no feasible solution exists; the congestion price keeps increasing while the allocated rate remains at R_{min}. On the other hand, if $\gamma C > R_{max}$, then regardless of the choice of κ the congestion price will keep decreasing until it reaches zero, whereby the allocated rate reaches R_{max} and the link remains under-utilized. Between these two extreme cases, the system (17) – (18) can be expressed in continuous form for a small enough price update interval τ:

$$\dot{\lambda}(t) = \frac{\kappa}{\tau}(R(t) - \gamma C) \quad (21)$$

$$R(t) = R_0 + \sqrt{\frac{\theta}{\lambda(t-\tilde{\tau})}}, \quad (22)$$

where $\tilde{\tau}$ designates the round trip time experienced by the stream.

Linearizing the system around its equilibrium point ($R_{opt} = \gamma C$, $\lambda_{opt} = \frac{\theta}{(\gamma C - R_0)^2}$), we obtain:

$$\delta\dot{\lambda}(t) = \frac{\kappa}{\tau}\delta R(t) \quad (23)$$

$$\delta R(t) = -\frac{1}{2}\frac{\theta^{\frac{1}{2}}}{\lambda_{opt}{}^{\frac{3}{2}}}\delta\lambda(t-\tilde{\tau}), \quad (24)$$

for $\delta\lambda(t) = \lambda(t) - \lambda_{opt}$ and $\delta R(t) = R(t) - R_{opt}$. Substituting (24) into (23), the system becomes:

$$\delta\dot{\lambda}(t) = -\frac{\kappa}{2\tau}\frac{\theta^{\frac{1}{2}}}{\lambda_{opt}{}^{\frac{3}{2}}}\delta\lambda(t-\tilde{\tau}) = -\frac{\kappa}{2\tau}\frac{(\gamma C - R_0)^3}{\theta}\delta\lambda(t-\tilde{\tau}). \quad (25)$$

The open-loop frequency response of the system is (see Fig. 3):

$$\mathcal{G}(j\omega) = -\frac{\kappa\tilde{\tau}}{2\tau}\frac{(\gamma C - R_0)^3}{\theta}\frac{e^{-j\omega\tilde{\tau}}}{j\omega\tilde{\tau}}. \quad (26)$$

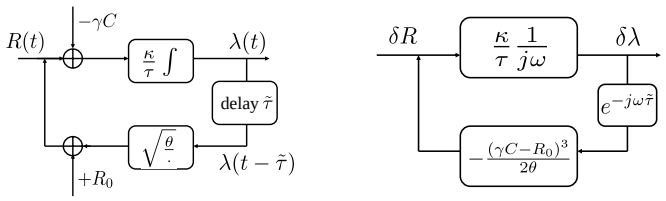
Denoting $\psi = \omega\tilde{\tau}$, we can rewrite (26) as $\mathcal{G}(j\omega) = \alpha e^{j\psi}/j\psi$, where

$$\alpha = \kappa\tilde{\tau}(\gamma C - R_0)^3/(2\tau\theta). \quad (27)$$

Given a fixed price update interval τ, the only remaining free parameter affecting the open-loop gain α is the price update scaling factor κ. All other parameters are determined either by the video stream (RD parameters θ and R_0) or by the network condition (link throughput C and round-trip-time $\tilde{\tau}$).

We apply the Nyquist plot analysis of the open-loop system (26) to guide our choice of κ to ensure stability of the closed-loop system. Figure 4 shows the Nyquist plot for the function $e^{j\psi}/j\psi$. It can be noted that the curve crosses the real axis at (-$\frac{2}{\pi}$,0), on the right side of the critical point (-1, 0). The Nyquist stability criterion states that the closed-loop system is stable, if and only if the Nyquist plot of the open-loop system does not encircle the critical point (-1, 0) [35]. Therefore, to ensure that the closed-loop system in Fig. 3 (b) is stable, one would need the scaling factor $\alpha < \pi/2$. In other words,

[4]For notational simplicity, we omit in this subsection the stream superscript s and the link subscript l. Without loss of generality, the weight of the stream is chosen as $w = 1$.



(a) nonlinear system (b) linearized system

Fig. 3. Block diagram of rate control for a single stream traversing a single bottleneck link.

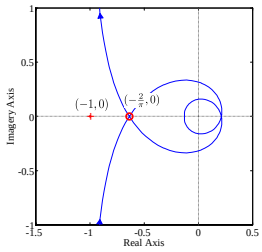


Fig. 4. Nyquist plot of the function $e^{-j\psi}/j\psi$.

a sufficient and necessary condition for stability for the single-link, single-stream system is:

$$\frac{\kappa\tilde{\tau}}{2\tau}\frac{(\gamma C - R_0)^3}{\theta} < \frac{\pi}{2}. \tag{28}$$

The value of κ should therefore be chosen as:

$$\kappa < \frac{\pi\theta\tau}{(\gamma C - R_0)^3\tilde{\tau}}. \tag{29}$$

In general, a longer round trip time $\tilde{\tau}$, a higher allocated rate at equilibrium $R_{opt} = \gamma C$, or a lower video scene complexity θ would require a lower value of κ to ensure stability.

IV. Distributed Rate Allocation Protocol

Additional challenges arise when deploying the proposed distributed rate allocation scheme in a practical system. Wireless link throughput experiences abrupt changes due to fluctuations in channel conditions. Video RD characteristics vary over time in the presence of scene cuts. When the video rate adaptation is achieved by bitstream switching, the approach introduces additional constraints: the available rates of each stream constitute a discrete set whereas rate adaptation can only occur at boundaries of encoded groups of pictures. In the following, we explain how to address the above issues in the design of a distributed media-aware rate allocation protocol for video streaming over wireless networks.

A. System Overview

Figure 5 provides an overview of a system containing multiple simultaneous video streaming sessions over a wireless network. For the purpose of distributed rate allocation, each node consists of a Link State Monitor (LSM) at the MAC layer and a Video Rate Controller (VRC) at the application layer. Cross-layer information exchange is achieved by granting intermediate wireless nodes access to a set of special video packet header fields listed in Table II.[5]

[5]This paper focuses on the conceptual design of the rate allocation protocol, without getting into details of the packet header field assignment. In a real implementation, it is possible to map these special fields as extensions of existing transport protocol headers, for instance, as supported by RTP [6] header extensions.

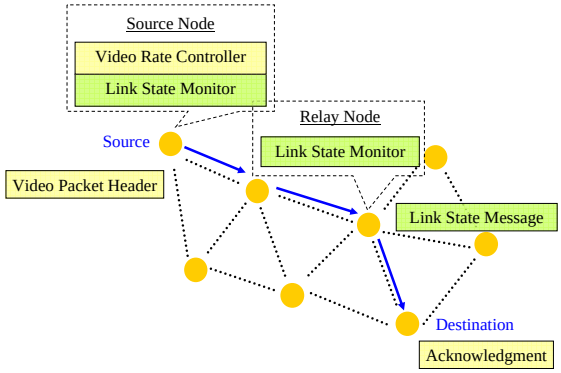


Fig. 5. Cross-layer information exchange among the MAC-layer Link State Monitors (LSMs) and the application-layer Video Rate Controllers (VRCs). The wireless nodes periodically exchange link state messages with their neighbors to collect utilization and congestion prices information of other links within the same interference set. The video packet header contains fields of advertised rate and accumulated congestion price to facilitate optimal rate allocation at the sender.

Symbol	Content	Size	Range
s	Stream ID	1 byte	0 – 255
R^s	Advertised Rate	2 bytes	0 – 65535×10^{-3} Mbps
Λ^s	Congestion Price	2 bytes	0 – 65535×10^{-2} MSE/Mbps

TABLE II
Fields for Cross-Layer Information Exchange in the Video Packet Header.

In addition, neighboring nodes exchange knowledge of their local utilization and congestion prices by periodically broadcasting link state messages containing entries shown in Table III.

This protocol is fully distributed: calculation of the video rates is confined to the VRCs at the senders while update of congestion prices in LSMs depends on utilization over local interference sets only. Although each stream does not share its video RD characteristics with others, the rate allocation result is media-aware, in that the solution minimizes the total video distortion of all participating streams.

B. Link State Monitor

At the MAC layer, the goal of the link state monitor on a wireless node is multi-fold: to estimate link state information such as throughput and background traffic rate, to track total utilization within interference sets centered around its outgoing links, to update congestion prices according to local utilization, to report congestion information in traversing video packet headers, and to collect utilization and congestion price information from neighboring nodes. We describe these tasks in greater detail below.

1) Link State Estimation: The link state monitor at each wireless node estimates the throughput of each outgoing link as: $C_l = B_l/(T_o + \frac{B_l}{C_l^o})$, where B_l is average packet payload size, T^o is the average MAC-layer overhead in terms of occupied channel time, and C_l^o is the *nominal link speed* of Link l, i.e., data rate for transmitting payload information. In wireless networks following the IEEE 802.11 standard, T^o can be further decomposed into: time for transmitting MAC-layer control packets (RTS/CTS/ACK) and headers, typically at a lower *basic* rate, guard time slots (SIFS/DIFS/EIFS) between subsequent transmission of control or data packets, and expected random backoff period before each transmission.

Symbol	Content	Size	Range
l	Link ID	1 byte	0 – 255
C_l	Link Throughput	2 bytes	0 – 65535$\times 10^{-3}$ Mbps
u_l	Link Utilization	1 byte	0 – 255%
λ_l	Congestion Price	2 bytes	0 – 65535$\times 10^{-3}$ Mbps

TABLE III
ENTRIES IN THE LINK STATE MESSAGE.

Estimation of background traffic rate on Link l is straightforward: $F_l' = B_l'/T_l$, where B_l' is average packet size at the application layer, and T_l is average packet inter-arrival time. Both statistics are obtained from logging consecutive packet arrival times over each link.

2) Congestion Price Update: The link state monitor maintains a congestion price λ_l for each interference set $\mathcal{L}_l$ centered around one of its outgoing links. At intervals of τ triggered by a price update timer, the LSM updates the congestion price λ_l based on total utilization within the interference set $\tilde{u}_l$ according to (17). Calculation of $\tilde{u}_l$, in turn, requires collection of the total link utilization reported in link state message from neighboring nodes.

Upon relaying a video packet from Stream s over Link l, the LSM records the advertised rate R^s in the video packet header and refreshes its local cache of $u_l^s = R^s/C_l$ and $u_l = u_l' + \sum_s u_l^s$ accordingly. It also updates the field of accumulated congestion price Λ_s in the video packet header as:

$$\Lambda^s := \Lambda^s + \sum_{l' \in \mathcal{L}_l} \frac{\lambda_{l'} C_{l'}}{C_l}. \tag{30}$$

Here, the values of $\lambda_{l'} C_{l'}$ are collected from link state messages from neighboring nodes. As the video packet traverses all links along its path, the final accumulated congestion price is the same as in (19):

$$\Lambda^s = \sum_{l \in \mathcal{P}^s} \sum_{l' \in \mathcal{L}_l} \frac{\lambda_{l'} C_{l'}}{C_l} = \sum_{l' \in \mathcal{L}_{\mathcal{P}^s}} \lambda_{l'} C_{l'} \sum_{l \in \mathcal{P}^s \cap \mathcal{L}_l} \frac{1}{C_l} = \sum_{l' \in \mathcal{L}_{\mathcal{P}^s}} \frac{\lambda_{l'} C_{l'}}{\tilde{C}_{l'}^s}. \tag{31}$$

3) Link State Message Exchange: Each node informs its neighbors of its existence via periodic broadcast of a link state message. The message also contains updated values of the estimated link throughput C_l, the channel time utilization u_l, and the congestion price λ_l of its outgoing and incoming links, as shown in Table III. Such information allows the wireless node to construct a local view of the interference sets of its outgoing links. Note that a node within the transmission range of both the source and the destination of a link may asynchronously receive multiple versions of its link state information. When such cases arise, it adheres to the information originated from the source of the link.

C. Video Rate Controller

At the application layer, the video rate controller is in charge of tracking the video RD characteristics over time, of re-calculating the advertised video rate upon receipt of every ACK packet, and of adapting the video streaming rate once the allocation has converged.

We assume that the video sequence is pre-encoded at different quality levels, resulting in a discrete set of RD tradeoff points $(R_1, D_1), \ldots, (R_K, D_K)$. The VRC records one set of RD data points and the fitted model parameters (θ, R_0, D_0) according to (9) for each group of pictures (GOP) in the encoded bitstream to track video content changes. The GOP duration is typically 0.5 to 2.0 seconds, therefore the rate allocation should converge within the same time frame.

Upon receipt of an ACK packet, the VRC records the accumulated congestion price Λ^s reported in the packet header. It then re-calculates the allocated rate R^s according to (18), using RD parameters for the *next* GOP in the sequence.[6] The updated value of R^s is then advertised in subsequent video packet headers until the arrival of the next ACK packet.

The convergence criterion for the rate allocation process is that the fluctuation between consecutive allocations is smaller than the difference between adjacent available rate points. Meanwhile, the received end-to-end congestion price should not exhibit an increasing or decreasing trend. When the allocation has converged, the VRC chooses the quality level k for the next GOP such that $R_k^s \le R_{opt}^s < R_{k+1}^s$. The actual bitstream switch occurs later, when the first frame in the next GOP is transmitted.

V. PERFORMANCE EVALUATION

A. Simulation Setup

We evaluate the proposed rate allocation protocol in various simulations of high-definition (HD) video streaming over wireless networks. All simulations are carried out in `ns-2`, an event-based packet-level network simulator [36]. The transmission power and receiving threshold of the nodes are adjusted to achieve a transmission range of 55 m. MAC parameters such as SIFS/DIFS/EIFS slot time, random backoff window size, and retry limits are chosen according to specifications of the IEEE 802.11a standard [37]. The basic rate for header and control packet transmissions is set to 6 Mbps; the nominal link speed for payload transmissions varies between 6 and 54 Mbps. RTS/CTS handshakes are enabled in multi-hop topologies. Routing is specified manually and is kept static throughout each simulation.

Six video sequences are considered for streaming. They are encoded using `x264` [38], a fast implementation of the H.264/AVC standard [39], at various quantization step sizes. The GOP length is 30 frames; the GOP structure is IBBPBBP..., similar to that used in MPEG-2 streams. Different RD tradeoff points are obtained by varying the quantization parameter (QP). Figure 6 shows their relative rate-distortion and rate-PSNR performances. Encoded video frames are segmented into packets with maximum size of 1500 bytes for streaming. Packet transmission intervals are evenly spread out within each GOP to avoid unnecessary queuing delays due to large intra-coded frames.

All simulations presented in this paper adhere to the same set of protocol parameter choices: the target utilization is γ = 85%; the price update scaling factor is κ = 0.01 MSE/Mbps^2; the price update interval is τ = 10 ms; the video ACK interval is τ_{ACK} = 16 ms, corresponding to one ACK per frame; and the LSM message exchange interval is τ_{LSM} = 20 ms. These values are chosen to achieve the best tradeoff between convergence speed, protocol overhead, and network congestion level, based on simulation studies over a wide range of video contents and network topologies.

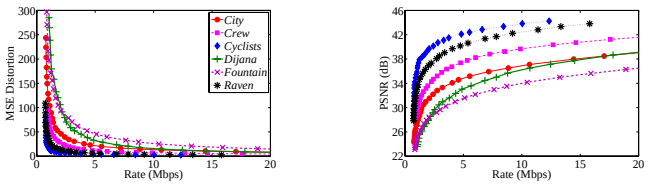
B. Comparison Schemes

Performance of the proposed media-aware rate allocation protocol is compared against the conventional approach of regulating video streaming rate according to the TCP-Friendly Rate Control (TFRC) equation [2]:

$$R = k \frac{B}{RTT\sqrt{p}}. \tag{32}$$

In (32), the rate is determined by estimated round trip time RTT, average packet size B, packet loss ratio p, and a scaling factor k.

[6]Since rate adaptation can only be performed at GOP boundaries, the video rate controller needs to calculate the rate allocation *before* the actual transmission of each GOP starts.



(a) (b)

Fig. 6. Rate-distortion (a) and rate-PSNR (b) tradeoff curves of the six test video sequences in high-definition (HD) resoution. They are encoded by `x264` [38], a fast implementation of the H.264/AVC standard [39].

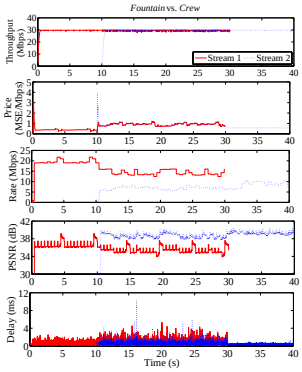
We enhance the TFRC-based scheme by introducing a *virtual* Random-Early-Detection (RED) mechanism at the relay nodes. Such mechanism helps to avoid quality degradation caused by actual packet losses due to queue overflow. The relay node monitors its queue size, and calculates the probability for randomly marking a packet as *virtual loss* according to the same principles in RED queues [40]. Such virtual loss is marked in the the 1-bit ECN field in the IP packet header [41]. The receiver calculates the average percentage of marked packets as the *virtual* packet loss ratio p and reports this information back to the sender via ACK packets.

We also compare results from the proposed distributed protocol against the performance upper bound achieved by a hypothetical omniscient centralized rate controller, which has global information of wireless link states throughout the network and video RD parameters of all streams.

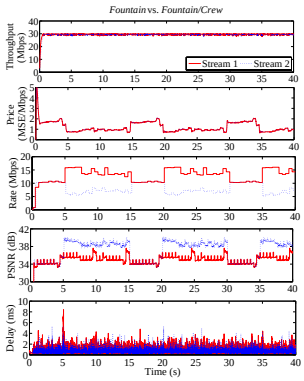
C. Adaptation to Changes

We first investigate transient behavior of the proposed media-aware rate allocation protocol in the simple scenario of two HD streams over two parallel single-hop wireless links. In the first example, the *Fountain* HD sequence streams over the first link and the *Crew* HD sequence streams over the second link 10 seconds later. Figure 7 (a) presents traces of estimated link throughput, accumulated congestion price, allocated video rates, corresponding video qualities in PSNR, and packet delivery delays. The accumulated congestion prices observed by both streams react to the arrival of the second stream by quickly rising to a new equilibrium value, inducing lower allocated rate for the existing *Fountain* stream. In a similar fashion, after the *Fountain* stream has finished transmission at time $t = 30$s, rate and quality of *Crew* increase in accordance with decreased accumulated congestion price. In both cases, the updated congestion prices and the allocated video rates of these streams reach convergence within 0.5 – 1 seconds, comparable to their GOP durations. When both streams are active in the network, difference in their allocated rates reflects difference in their respective video RD characteristics: the more complex *Fountain* streams at a higher rate with a lower video quality while the relatively static *Crew* streams at a lower rate with a higher video quality.

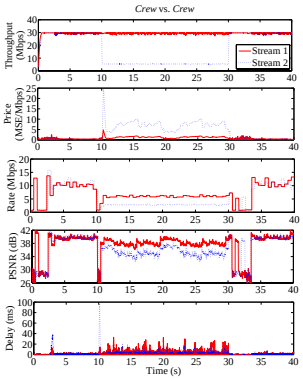
Next, we examine the case involving periodic scene cuts in one of the video streams, with alternating content from *Fountain* and *Crew*. Figure 8 shows the time-varying rate-PSNR curves of this composite sequence. As can be observed from the traces in Fig. 7 (b), the allocated rates and accumulated congestion prices react quickly to such periodic scene cuts. During the periods when the composite sequence shows contents from *Fountain*, the allocated rates and qualities are the same for both streams. During the periods when the composite sequence displays scenes from the less complex *Crew* sequence, the allocation is re-adjusted to higher rate for *Fountain* and lower rate for *Fountain/Crew*.



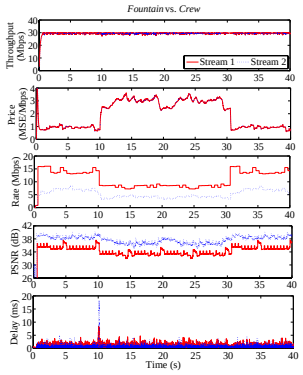
(a) Stream arrival and departure



(b) Video scene cut



(c) Abrupt change of link speed



(d) Presence of background traffic

Fig. 7. Traces of estimated link throughput, accumulated congestion prices, allocated video rates, resulting video qualities, and packet delivery delays. In this scenario, two HD sequences stream over two parallel single-hop wireless connections. Both wireless links operate at a nominal link speed of 54 Mbps.

The impact of abrupt changes in the wireless link speed is demonstrated in Fig. 7 (c). In this experiment, the *Crew* HD sequence streams over both links. Initially both links operate at the same nominal speed of 54 Mbps and both streams receive the same rates. At time $t = 10$s, nominal speed of the second link drops abruptly to 6 Mbps, reducing its throughput from around 30 Mbps to 5 Mbps. The media-aware rate allocation protocol responds to such changes by increasing the congestion prices and reducing the rates of both streams, meanwhile favoring the sequence over the faster link. Shortly after the link speed recovers to 54 Mbps at time $t = 30$s, both streams resume their initial video rates and qualities.

The traces in Fig. 7 (d) illustrate how the media-aware rate allocation protocol reacts to the presence of non-video background traffic. In this experiment, the *Fountain* and *Crew* sequences stream over two parallel wireless links, both operating at a nominal speed of 54 Mbps. A UDP flow of rate 10 Mbps enters on the second link at time $t = 10$s and finishes transmission at time $t = 30$s. During this period, the congestion prices observed by both streams increase to a higher level, leading to reduced video rates and qualities in both streams. Convergence time for adapting to the arrival and departure of background traffic is around 1 second, comparable to the GOP durations of these streams. Irrespective of the presence of background traffic, allocation for the more complex *Fountain* sequence is consistently higher than for *Crew*, reflecting difference in their respective video RD characteristics.

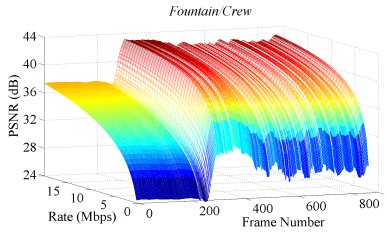


Fig. 8. Rate-PSNR curves in the composite sequence *Fountain/Crew*. The initial 300 frames from *Fountain* contains highly complex scenes; the remaining 600 frames from *Crew* sequence contains moderate motion.

D. Comparison with TFRC

We now compare performance of the proposed media-aware rate allocation protocol against the media-oblivious TFRC scheme, as described in Section V-B. In the simple scenario of two HD sequences streaming over two parallel links, the first link operates at a nominal speed of 54 Mbps whereas the second link varies its speed from 6 Mbps to 54 Mbps. Figure 9 shows the allocated rates, video quality of each stream, and the average video quality of both streams for three sequence pairs. Since the TFRC scheme relies only on end-to-end observations, it allocates similar rates for both streams irrespective of their video RD characteristics and underlying link speeds. In contrast, the proposed media-aware allocation yields a higher rate for the more complex stream traversing a faster link. This leads to a higher average video quality of both streams and more balanced qualities among the streams. For instance, in the case of *Fountain* over a fast link competing against *Cyclists* over a slow link, improvement in average video quality ranges between 0.7 – 4.3 dB in PSNR.

Similar observations hold for Fig. 10, when three HD sequences streaming over three parallel links. Two of the streams bear identical video contents and traverse links operating at 54 Mbps. The nominal link speed of the third link supporting the third stream varies between 6 Mbps and 54 Mbps. In comparison with TFRC, the media-aware allocation achieves higher rates for streams with more demanding video RD characteristics or traversing faster links. As a result, it improves the average video quality over TFRC by 0.2 dB to 3 dB in PSNR, for various video sequence pairs and wireless link speeds.

We now investigate the more general scenario of multiple video streams sharing a multi-hop network. Figure 11 shows three representative topologies, together with the average video quality achieved by both the media-aware and TFRC schemes. Media-aware allocation consistently outperforms the TFRC in all three networks. The gain in average video quality ranges between 0.8 – 1.5 dB in the PARALLEL network, between 2.6 – 5.5 dB in the JOINT network, and between 2.1 – 4.1 dB in the MESH network.

E. Comparison with Centralized Allocation

We verify effectiveness of the distributed rate allocation protocol by comparing the resulting video qualities against results obtained from an oracle-aided centralized rate controller as described in Section V-B. Figure 12 (a) shows the comparison for three HD streams competing over three parallel single-hop connections. Figure 12 (b) shows the comparison for two HD streams sharing the multi-hop PARALLEL network. In both cases, the per-stream video qualities achieved by the distributed media-aware scheme match closely with those from the centralized solution. Note that the video qualities from the distributed scheme are slightly lower than their centralized counterparts, due to the presence of protocol overhead introduced by video acknowledgment streams and link state message exchanges.

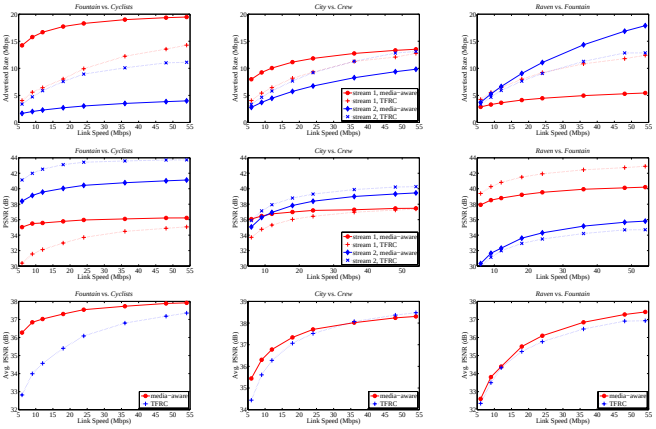


Fig. 9. Allocated video rate of each stream, video quality of each stream, and average video quality of both streams, as achieved by the media-aware and TFRC schemes. The first HD stream traverses a single-hop wireless link with a nominal speed of 54 Mbps, whereas the second HD stream traverses a wireless link with a nominal speed varying from 6 Mbps to 54 Mbps. In all simulations, the results are averaged over 600 seconds after convergence.

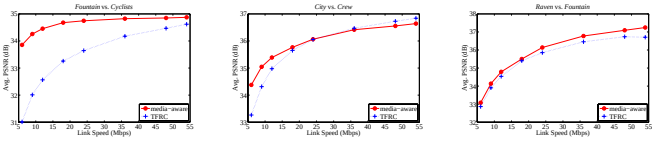


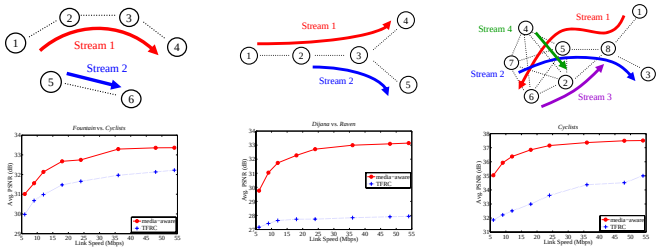
Fig. 10. Average Video quality in PSNR of all streams, as achieved by the media-aware and TFRC schemes, when three HD streams compete over three parallel wireless links. The first two streams traverse links with a nominal speed of 54 Mbps; the third stream traverses a link with a nominal speed varying from 6 Mbps to 54 Mbps. In all simulations, the results are averaged over 600 seconds after convergence.

F. Fairness Among Streams

In the following, we show that the proposed media-aware rate allocation protocol can also effectively resolve fairness issues encountered by TCP flows over 802.11 wireless networks. We compare rate allocation results from the media-aware scheme against TFRC and TCP in several network topologies with known fairness issues reported in literature [24]. In all simulations, all video streams contain the same *Fountain* HD sequence and all links operate at the same nominal speed of 54 Mbps.

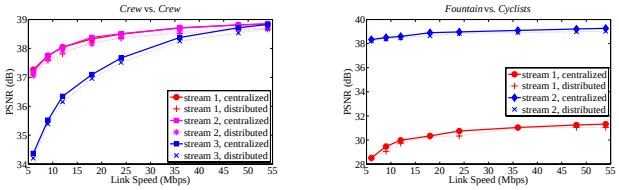
Figure 13 (a) shows the ASYMMETRY topology where only the receiver of the first stream and the sender of the second stream can overhear each other's RTS/CTS messages. When TCP is used, since both streams always have packets waiting to be transmitted and transmissions are initiated by the sender, the second stream has a higher success probability of reserving the wireless channel for transmission. It therefore achieves a significantly higher throughput than the first stream. In the case of streaming over TFRC, the first stream achieves a higher allocated rate due to lower observed round trip times and virtual packet loss ratios. In comparison, the proposed media-aware rate allocation protocol explicitly accounts for the presence of a neighboring link by means of link state message exchanges. It therefore allocates equal rates for these two streams traversing links with equal nominal speeds.

Figure 13 (b) illustrates the STACK topology leading to the well-known *flow-in-the-middle* problem [26]. Again we compare the allocated stream rates achieved by TCP, TFRC and the media-aware scheme. When TCP is used, the two outside streams are unaware of each other's presence and their transmissions completely block the



(a) PARALLEL (b) JOINT (c) MESH

Fig. 11. Topologies of wireless multi-hop networks and average video qualities resulting from the media-aware and TFRC schemes. (a) In the PARALLEL network, two HD video streams compete over a 3-hop path and an 1-hop path respectively. The nominal link speed from Node 1 to Node 2 varies from 6 Mbps to 54 Mbps; all other links operate at 54 Mbps. (b) In the JOINT network, two HD video streams compete over a 3-hop path and an 2-hop path respectively. The nominal link speed from Node 1 to Node 2 varies from 6 Mbps to 54 Mbps; all other links operate at 54 Mbps. (c) In the MESH network, four HD sequences with the same content stream over paths of different hop counts. All links operate at 54 Mbps except the one from Node 1 to Node 8, the nominal speed of which varies between 6 Mbps and 54 Mbps. In all simulations, the results are averaged over 600 seconds after convergence.



(a) three single-hop streams (b) PARALLEL network

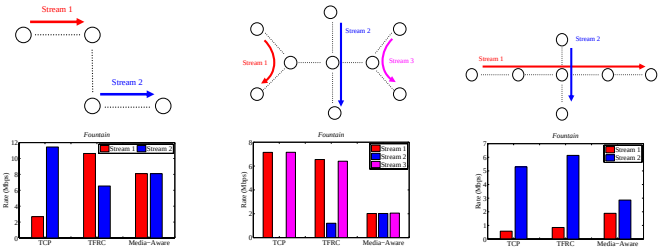
Fig. 12. Comparison of distributed and centralized schemes in terms of average video quality of each stream. (a) three HD streams compete over three parallel single-hop wireless connections; (b) two HD streams compete over the multi-hop PARALLEL network as shown in Fig. 11 (a). In all simulations, the results are averaged over 600 seconds after convergence.

stream in the middle, which needs to avoid collision with packets from either stream. Such unfairness is alleviated by TFRC. Since the source rate of each stream is regulated according to the TFRC equation in (32), the two outside streams no longer fully occupy all transmission opportunities, leaving some room for the middle stream. Nevertheless, rate of the middle stream is significantly lower, due to higher observed round trip times and virtual packet loss ratios. In contrast, the media-aware scheme allocates equal rates among all three streams by inferring from link state message exchanges that their underlying links belong to the same interference set.

In Fig. 13 (c), the two video streams traverse different number of links along their paths in the CROSS topology. Allocation from both TCP and TFRC tend to favor the stream over the shorter path as it encounters shorter round trip time. In the media-aware scheme, the allocated rates of these streams are also inversely weighted by their hop counts, since a stream traversing a longer path incurs higher level of network congestion. Nevertheless, the media-aware allocation is more balanced than TCP and TFRC.

G. Protocol Scalability

Finally, we investigate how the distributed media-aware rate allocation protocol scales with growing video stream density and network size. We vary both the number of competing streams and the hop count traversed by each stream, over the GRID wireless network



(a) ASYMMETRY (b) STACK (c) CROSS

Fig. 13. Network topology and comparison of the allocated stream rates from TCP, TFRC and the media-aware schemes. In all simulations, all video streams contain the same *Fountain* HD sequence and all links operate at the same nominal speed of 54 Mbps.

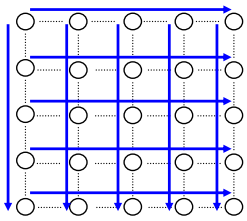


Fig. 14. A grid network, carrying competing video streams along the horizontal and vertical paths. All streams have the same content of *City* and all links have the same nominal speed of 54 Mbps. For the simulation investigating protocol parameter choices, the grid size varies from 3×3 to 7×7.

in Fig. 14. All streams carry the same *City* sequence in standard-definition (SD) resolution, and all links operate at the same nominal speed of 54 Mbps. In Fig. 15, the allocated stream rates, the average video quality of all streams in PSNR, and the protocol overhead in percentage of channel time utilization are plotted as function of network grid size. When the grid size is greater than 3×3, the video streams in the center of the network contribute more to the overall network congestion, therefore they receive lower rates than the streams on the peripheral. In general, increasing the network grid size lowers the allocated video rates and qualities, while increasing the protocol overhead introduced by video ACK streams. The protocol

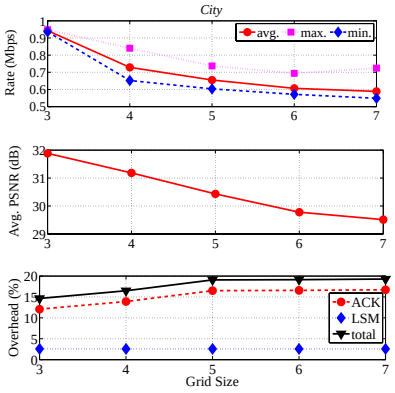


Fig. 15. The allocated video rates, the average video quality of all streams, and the protocol overhead in percentage of channel time utilization, as the size of the GRID network varies from 3×3 to 7×7. The same *City* SD sequence is streamed across parallel horizontal and vertical paths in the grid. All links operate at a nominal speed of 54 Mbps. In all simulations, the results are averaged over 600 seconds after convergence.

overhead due to link state message exchanges remains constant, since the number of neighbors for each wireless node is constant in a grid. Furthermore, both the video qualities and the protocol overhead tend to saturate for grid sizes greater than 5 × 5. Beyond this size, the network constitutes of multiple overlapping interference sets repeating the same topological structure.

VI. Conclusions

We present in this paper a distributed rate allocation protocol for wireless video streaming, with the goal of minimizing total mean-squared-distortion (MSE) distortion of all participating video streams. The proposed scheme is media-aware, in that it tends to allocate higher rates for streams with more demanding RD characteristics at the expense of lower rates and reduced qualities for less complex streams. The proposed scheme is also network-aware, in that it explicitly accounts for the impact of heterogeneous wireless link speeds into the calculation of congestion prices, thereby favoring streams over faster links. The adoption of cross-layer information exchange further facilitates fast convergence of the protocol in various transient events, such as abrupt changes in wireless link speeds, video scene cuts, arrival/departure of other video streams, and presense of non-video background traffic.

Performance of the proposed media-aware rate allocation protocol is compared against the conventional media-oblivious TCP-Friendly Rate Control (TFRC) scheme in various network simulations. The media-aware allocation consistently outperforms TFRC in terms of average video quality of all participating streams, and achieves more balanced video qualities among the streams. Comparison against a hypothetical oracle-aided centralized rate controller confirms that the distributed allocation results closely approximate their centralized counterparts. The proposed protocol can also effectively resolve the fairness issues experienced by TCP and TFRC flows in several representative wireless network topologies.

While we focus on streaming of pre-encoded high-definition (HD) video over 802.11-based wireless home networks throughout this paper, the insights gained may apply more generally. It is straightforward to extend the media-aware rate allocation scheme for applications involving live encoding. We have also extended the optimization framework for video multicast over wireless networks, and plan to report the results in a separate paper. For future research, we intend to investigate joint optimization of video rate allocation and error control for streaming over more error-prone networks.

Appendix A
Stability Proof for the Equilibrium Analysis of the General Case

In this section, we establish stability of the media-aware rate allocation scheme for the general case of multiple video streams sharing multiple bottleneck interference sets. Using the same linearization as in Section III-E, we obtain:

$$\delta\dot{\lambda}_l(t) = \frac{\kappa}{\tau}\sum_{s:l\in\mathcal{L}_{\mathcal{P}^s}}\frac{\delta R^s(t)C_l}{\tilde{C}_l^s}, \qquad l\in\mathcal{L} \quad (33)$$

$$\delta R^s(t) = -\frac{1}{2}\frac{R^s_{opt}-R^s_0}{\Lambda^s_{opt}}\sum_{l\in\mathcal{L}_{\mathcal{P}^s}}\frac{\delta\lambda_l(t-\tilde{\tau}^s)C_l}{\tilde{C}_l^s}, \quad s\in\mathcal{S}, \quad (34)$$

with $\delta\lambda_l(t)=\lambda_l(t)-\lambda_l^{opt}$, $\delta R^s(t)=R^s(t)-R^s_{opt}$ and $\delta\Lambda^s(t)=\Lambda^s(t)-\Lambda^s_{opt}$. The open-loop frequency response of the multiple-input multiple-output system (33) – (34) can be expressed in matrix form:

$$\mathcal{G}(s) = -\frac{\kappa}{2\tau}\mathcal{A}\mathcal{M}\Psi(j\omega)\mathcal{A}^{\mathcal{T}}, \quad (35)$$

where

$$\mathcal{M} = \text{diag}\left(\frac{(R^s_{opt}-R^s_0)\tilde{\tau}^s}{\Lambda^s_{opt}}\right), \quad (36)$$

$$\Psi(s) = \text{diag}\left(\frac{e^{-j\omega\tilde{\tau}^s}}{j\omega\tilde{\tau}^s}\right), \quad (37)$$

and matrix $\mathcal{A}$ has elements:

$$A^s_l = \begin{cases}\frac{C_l}{\tilde{C}^s_l}, & l\in\mathcal{L}_{\mathcal{P}^S}\\ 0, & \text{otherwise.}\end{cases} \quad (38)$$

For any eigenvalue of $\mathcal{G}$ with corresponding eigenvector x such that $\nu x=\mathcal{G}x$, we have:

$$\nu\mathcal{A}^T x = -\frac{\kappa}{2\tau}(\mathcal{A}^T\mathcal{A}\mathcal{M}\Psi)\mathcal{A}^T x. \quad (39)$$

Therefore, ν is also an eigenvalue for the matrix $\mathcal{H}=-\frac{\kappa}{2\tau}\mathcal{A}^T\mathcal{A}\mathcal{M}\Psi$, with corresponding eigenvector $y=\mathcal{A}^{\mathcal{T}}x$. It follows that $\mathcal{H}$ has the same eigenvalues as $\mathcal{G}$. We further decompose $\mathcal{H}$ as $\mathcal{H}=\mathcal{W}\Psi$, where $\mathcal{W}=-\frac{\kappa}{2\tau}\mathcal{A}^T\mathcal{A}\mathcal{M}$. The spectral radius of $\mathcal{W}$ is upper bounded by:

$$\rho(\mathcal{W}) \le ||\mathcal{W}|| \le \frac{\kappa}{2\tau}||\mathcal{A}^T\mathcal{A}||\cdot||\mathcal{M}|| = \frac{\kappa}{2\tau}||\mathcal{A}^T\mathcal{A}||\max_s\frac{(R^s_{opt}-R^s_0)\tilde{\tau}^s}{\Lambda^s_{opt}}. \quad (40)$$

We now invoke the following lemma from [42] to show that a sufficient condition for system stability is to ensure $\rho(\mathcal{W})<\frac{\pi}{2}$.

Lemma A.1: Given matrices $\mathcal{W}\succeq 0$ and Ψ=diag(ψ^s) of size $S\times S$, the eigenvalues of matrix $\mathcal{H}=\mathcal{W}\Psi$ belong to the convex hull of $\{0, \psi^1\dots\psi^S\}$, scaled by the spectral radius $\rho(\mathcal{W})$.

Proof: See [42]. ∎

Note that the points $\psi^s=\frac{e^{-j\omega\tilde{\tau}^s}}{j\omega\tilde{\tau}^s}$ belong to the curve in the Nyquist plot on Fig. 4, and that its convex hull intersects the negative real axis in the segment $[-\frac{2}{\pi}, 0]$. Therefore, so long as $\rho(\mathcal{W})<\frac{\pi}{2}$, the eigenvalues of $\mathcal{H}=\mathcal{W}\Psi$ will not cause the Nyquist plot to encircle the critical point at $(0,-1)$ and the system is stable [35]. From (40), a sufficient condition for $\rho(\mathcal{W})<\frac{\pi}{2}$ is:

$$\frac{\kappa}{2\tau}||\mathcal{A}^T\mathcal{A}||\frac{(R^s_{opt}-R^s_0)\tilde{\tau}^s}{\Lambda^s_{opt}} < \frac{\pi}{2}, \ \forall s\in\mathcal{S}. \quad (41)$$

The price update scaling factor κ should therefore be chosen as:

$$\kappa < \frac{\pi\tau}{||\mathcal{A}^T\mathcal{A}||}\min_{s\in\mathcal{S}}\frac{w^s\theta^s}{(R^s_{opt}-R^s_0)^3\tilde{\tau}^s}. \quad (42)$$

References

[1] X. Zhu and B. Girod, "Distributed media-aware rate allocation for wireless video streaming," in *Picture Coding Symposium (PCS'09)*, Chicago, IL, USA, May 2009.

[2] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the Internet," *IEEE/ACM Trans. on Networking*, vol. 7, pp. 458–472, Aug. 1999.

[3] V. Jacobson, "Congestion avoidence and control," in *Proc. ACM Conference on Communications Architectures and Protocols (SIGCOMM'88)*, vol. 18, no. 4, Stanford, CA, USA, Aug. 1988, pp. 157–173.

[4] M. Allman, V. Paxson, and W. R. Stevens, *TCP Congestion Control*, RFC 2581 (Proposed Standard), Updated by RFC 3390, Apr. 1999.

[5] J. Postel, *User Datagram Protocol*, RFC 768 (Standard), Aug. 1980.

[6] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson., "RTP: A transport protocol for real-time applications," RFC 3550 (Standard), Jul. 2003.

[7] M. Handley, S. Floyd, J. Pahdye, and J. Widmer, "TCP friendly rate control (TFRC): Protocol specification," RFC 3448 (Proposed Standard), Jan. 2003.

[8] H. Kanakia, P. P. Mishra, and A. Reibman, "An adaptive congestion control scheme for real-time packet video transport," in *Proc. ACM Conference on Communications Architectures and Protocols (SIGCOMM'93)*, Ithaca, NY, USA, Sep. 1993, pp. 20–31.

[9] N. Wakamiya, M. Miyabayashi, M. Murata, and H. Miyahara, "MPEG-4 video transfer with TCP-friendly rate control," *Lecture Notes In Computer Science; Proc. 4th IFIP/IEEE International Conference on Management of Multimedia Networks and Services (MMNS'01)*, vol. 2216, pp. 29–42, 2001.

[10] A. Szwabe, A. Schorr, F. J. Hauck, and A. J. Kassler, "Dynamic multimedia stream adaptation and rate control for heterogeneous networks," *Journal of Zhejiang University, Science A*, vol. 7, no. 5, pp. 63–69, May 2006.

[11] Z. Wang, S. Banerjee, and S. Jamin, "Media-friendliness of a slowly-responsive congestion control protocol," in *Proc. 14th International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'04)*, Kinsale, County Cork, Ireland, Jun. 2004, pp. 82–87.

[12] J. Yan, K. Katrinis, M. May, and B. Plattner, "Media- and TCP-friendly congestion control for scalable video streams," *IEEE Trans. on Multimedia*, vol. 8, no. 2, pp. 196–206, Apr. 2006.

[13] E. Casey and G.-M. Muntean, "TCP compatible greediness control algorithm for wireless multimedia streaming," in *Proc. IEEE 65th Vehicular Technology Conference (VTC'07)*, Dublin, Ireland, Apr. 2007, pp. 160–164.

[14] F. P. 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, Mar. 1998.

[15] R. J. La and V. Anantharam, "Utility-based rate control in the Internet for elastic traffic," *IEEE Trans. on Networking*, vol. 10, no. 2, pp. 272–285, Oct. 2002.

[16] T. Alpcan and T. Basar, "A utility-based congestion control scheme for Internet-style networks with delay," in *Proc. 22nd Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03)*, Jul. 2003, pp. 2039–2048.

[17] K. Kar, S. Sarkar, and L. Tassiulas, "Optimization based rate control for multirate multicast sessions," in *Proc. 20th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'01)*, vol. 1, Anchorage, AK, USA, Apr. 2001, pp. 123–132.

[18] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, Jan. 2007.

[19] M. van der Schaar and N. S. Shankar, "New fairness paradigms for wireless multimedia communication," in *Proc. IEEE International Conference on Image Processing (ICIP'05)*, vol. 3, Genova, Italy, Sep. 2005, pp. 704–707.

[20] T. Ozcelebi, M. O. Sunay, M. R. Civanlar, and A. Tekalp, "Application QoS fairness in wireless video scheduling," in *Proc. IEEE 14th Signal Processing and Communications Applications (SIU'06)*, Antalya, Turkey, Apr. 2006, pp. 1–4.

[21] S. Khan, S. Duhovnikov, E. Steinbach, and W. Kellerer, "MOS-based multiuser multiapplication cross-layer optimization for mobile multimedia communication," *Advances in Multimedia*, vol. 2007, no. 1, pp. 1–11, Jan. 2007.

[22] M. Kalman and 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 (MMSP'05)*, Shanghai, China, Oct. 2005, pp. 1–4.

[23] X. Zhu and B. Girod, "Subjective evaluation of multi-user rate allocation for streaming heterogeneous video contents over wireless networks," in *Proc. IEEE International Conference on Image Processing (ICIP'08)*, San Diego, CA, USA, Oct. 2008, pp. 3092–3095.

[24] C. Chaudet, D. Dhoutaut, and I. G. Lassous, "Performance issues with IEEE 802.11 in ad hoc networking," *IEEE Communications Magazine*, vol. 43, no. 7, pp. 110–116, Jul. 2005.

[25] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda, "Performance anomaly of 802.11b," in *Proc. INFOCOM'03, Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, San Francisco, CA, U.S.A., Mar. 2003, pp. 836–843.

[26] A. Jindal and K. Psounis, "Achievable rate region and optimality of multi-hop wireless 802.11-scheduled networks," in *Proc. Information Theory and Applications Workshop (ITA'08)*, San Diego, CA, USA, Jan. 2008.

[27] K. Xu, M. Gerla, L. Qi, and Y. Shu, "Enhancing TCP fairness in ad hoc wireless networks using neighborhood RED," in *Proc. 9th annual international conference on Mobile computing and networking (MOBICOM'03)*, San Diego, CA, USA, Sep. 2003, pp. 16–28.

[28] D. Wu, T. Hou, and Y.-Q. Zhang, "Scalable video coding and transport over broadband wireless networks," *Proceedings of the IEEE, Special Issue on Multi-Dimensional Broadband Wireless Technologies and Applications*, vol. 89, no. 1, pp. 6–20, January 2001.

[29] T. Stockhammer, H. Jenkac, and G. Kuhn, "Streaming video over variable bit-rate wireless channels," *IEEE Trans. on Multimedia*, pp. 268–277, April 2004.

[30] Y.Li, A.Markopoulou, N.Bambos, and J.Apostolopoulos, "Joint power-playout control for media streaming over wireless links," *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 830–843, August 2006.

[31] 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, Jun. 2000.

[32] N. Z. Shor, *Minimization Methods for Non-differentiable Functions*. New York, USA: Springer Series in Computational Mathematics. Springer, 1985.

[33] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, United Kindom: Cambridge University Press, 2004.

[34] X. Zhu, "Distributed rate allocation for video streaming over wireless networks," Ph.D. dissertation, Stanford University, USA, 2009. [Online]. Available: http://msw3.stanford.edu/~zhuxq/thesis/

[35] A. E.-N. G. Franklin, J. D. Powell, *Feedback Control of Dynamic Systems*. NJ, USA: Prentice Hall, 2006.

[36] "The Network Simulator NS-2." [Online]. Available: http://www.isi.edu/nsnam/ns/

[37] *IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, P802.11*, Nov. 1997.

[38] "x264." [Online]. Available: http://developers.videolan.org/x264.html

[39] *Advanced Video Coding for Generic Audiovisual services, ITU-T Recommendation H.264 - ISO/IEC 14496-10 (AVC)*, ITU-T and ISO/IEC JTC 1, 2003.

[40] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, Aug. 1993.

[41] K. K. Ramakrishnan, S. Floyd, and D. Black, *The Addition of Explicit Congestion Notification (ECN) to IP*, RFC 3168 (Proposed Standard), Sep. 2001.

[42] G. Vinnicombe, "On the stability of end-to-end congestion control for the Internet," Technical Report, Cambridge University, CUED/F-INFENG/TR.398, Dec. 2000. [Online]. Available: http://www-control.eng.cam.ac.uk/gv/internet/