

Informationally Decentralized Video Streaming Over Multihop Wireless Networks

Hsien-Po Shiang and Mihaela van der Schaar, *Senior Member, IEEE*

Abstract—Various packet scheduling, dynamic routing, error-protection, and channel adaptation strategies have been proposed at different layers of the protocol stack to address multi-user video streaming over multihop wireless networks. However, these cross-layer transmission strategies can be efficiently optimized only if they use accurate information about the network conditions and hence, are able to timely adapt to network changes. Due to the informationally decentralized nature of the multihop wireless network, performing centralized optimization for delay-sensitive video streaming application based on global information about the network status is not practical. Distributed solutions that adapt the transmission strategies based on timely information feedback need to be considered. To acquire this information feedback for cross-layer adaptation, we deploy an overlay infrastructure, which is able to relay the necessary information about the network status and incurred delays across different network “horizons” (i.e., across a different number of hops in a predetermined period of time). In this paper, we propose a distributed streaming approach that is optimized based on the local information feedback acquired from the various network horizons. We investigate the distributed cross-layer adaptation at each wireless node by considering the advantages resulting from an accurate and frequent network information feedback from larger horizons as well as the drawbacks resulting from an increased transmission overhead. Based on the information feedback, we can estimate the *risk* that packets from different priority classes will not arrive at their destination before their decoding deadline expires. Subsequently, the various transmission strategies such as packet scheduling, retransmission limit and dynamic routing policies are adapted to jointly consider the estimated risk as well as the impact in terms of distortion of the different priority classes. Our results show that the proposed dynamic routing policy based on timely information feedback outperforms existing state-of-the-art on-demand routing solutions by more than 2 dB in terms of the received video quality.

Index Terms—Decentralized information feedback, dynamic routing, multihop wireless networks, video streaming.

I. INTRODUCTION

EMERGING multihop wireless LAN (WLAN) networks provide a low-cost and flexible infrastructure that can be simultaneously utilized by multiple users for a variety of applications, including delay-sensitive multimedia transmission. However, these wireless networks provide only limited Quality

of Service (QoS) support for real-time multimedia applications. Hence, efficient solutions for multimedia streaming must accommodate time-varying bandwidths and probabilities of error introduced by the shared nature of the wireless medium and quality of the physical connections. In the studied distributed transmission scenario, multimedia users proactively collaborate in sharing the available wireless resources to maximize their video quality. To enable optimal usage of the multihop infrastructure, the various network entities (source nodes, relay nodes, etc.) can timely and accurately exchange information about channel statistics, expected delays, or even packet loss probabilities (due to the deadline expiration of video packets) incurred by previously transmitted multimedia packets from different users and distortion classes across the network. However, this network information feedback usually shares the same resources allocated for the payload (e.g., multimedia) transmission and thus, the resulting overheads need to be explicitly considered for optimized transmission.

Prior research on multi-user multimedia transmission over wireless networks has focused on centralized, flow-based resource allocation strategies based on pre-determined rate requirements and usually neglects the overheads associated with the network information gathering and dissemination [1], [2], [6]. These solutions are not very adaptive to the network size or the number of users and attempt to solve the global end-to-end routing and path selection problem as a Multicommodity Flow [3] problem in a non-scalable fashion. This flow-based optimization does not take into account the fact that video applications are loss tolerant and hence, they can gracefully adjust their quality to accommodate a larger number of users as channel conditions are changing. Importantly, they do not guarantee that the packet delay constraints are met for video applications, since they do not timely adapt the transmission strategies for the various packets based on their delay deadlines and the available network information, which captures the changing network conditions such as congestion/interference.

Alternatively, the majority of the multimedia-centric research optimizes the video streaming using purely end-to-end metrics and does not consider the protection techniques available at the lower layers of the protocol stack. Hence, they do not take advantage of the significant gains provided by cross-layer design [2], [5], [6]. In [7], an integrated cross-layer optimization framework was proposed that considers the video quality impact based on different information horizons. However, the proposed solution in [7] considers only the single user case, where a set of paths and transmission opportunities are statically pre-allocated for each video application. This leads to a suboptimal, non-scalable solution for the multi-user case. Importantly,

Manuscript received December 9, 2006; revised February 11, 2007. This work was supported under NSF Career Grant CCF-0541867 and by UC Micro and HP Research. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anna Hac.

The authors are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: hpshiang@ee.ucla.edu; mihaela@ee.ucla.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2007.902845

the overhead induced by the various information horizons are not investigated in [7], which have essential impact for the delay-sensitive multimedia applications. To enable efficient distributed multi-user video streaming over a wireless multihop infrastructure, nodes need to timely collect and disseminate network information based on which, the various nodes can collaboratively adapt their cross-layer transmission strategies. For instance, based on the available information feedback, a network node can timely choose an alternate (less congested) route for streaming the packets that have a higher contribution to the overall distortion or a more imminent deadline.

Although the information feedback is essential to the cross-layer optimization, the cost of collecting the information is seldom discussed in the literature. Due to the informationally decentralized nature of the multihop wireless network, it is impractical to assume that the global network information and the time-varying application requirements can be relayed to the central (overlay) network manager in a timely manner. Distributed suboptimal solutions that adapt the transmission strategies based on well-designed localized information feedback should be adopted for the delay-sensitive applications.

In summary, no integrated framework has been developed that explicitly considers the impact of accurate and frequent network information feedback from various horizons, when optimizing the resource allocation and the cross-layer transmission strategies for multiple collaborating users streaming real-time multimedia over a wireless multihop network. In this paper, we build on the previous work [24] and investigate the impact of this information feedback on the distributed cross-layer transmission strategies deployed by the multiple video users. We assume a directed acyclic overlay network [8] that can be superimposed over any wireless multihop network to convey the information feedback. Our solution relies on the users' agreement to collaborate by dynamically adapting the quality of their multimedia applications to accommodate the flows/packets of other users with a higher quality impact and/or higher probability to miss their decoding deadlines. Unlike commercial multi-user systems, where the incentive to collaborate is minimal, we investigate the proposed approach in an enterprise network setting where source and relay nodes exchange accurate and trustworthy information about their applications and network statistics.

To increase the number of users that can simultaneously share the same wireless multihop infrastructure as well as to improve their performance given time-varying network conditions, we deploy scalable video coding schemes [21] that enable a fine-granular adaptation to changing network conditions and a higher granularity in assigning the packet priorities. We assume each receiving node performs polling-based contention-free media access control (MAC) [16] that dynamically reserves a transmission opportunity interval in a service interval. The network topology and the corresponding channel condition of each link are assumed to remain unchanged within the service interval.

In this paper, we discuss the required information/parameter exchange among network nodes/layers for implementing a distributed solution for selecting the following cross-layer transmission strategies at each intermediate node – the packet scheduling, and the next-hop relay (node) selection based on routing policies similar to the Bellman-Ford routing algorithm [10], and

the retransmission limit at the MAC layer. In performing the cross-layer adaptation, we explicitly consider the packet deadlines and the relative priorities (based on the quality impact of the packets) encapsulated in the packet headers. Each intermediate node maintains a queue of video packets from various users and determines the cross-layer transmission strategies in a distributed fashion through the information feedback from other intermediate nodes within a certain network horizon and with a certain frequency. While a larger horizon/frequency can provide more accurate network information, this also results in an increased transmission overhead that can have a negative impact on the video performance. Hence, we aim at quantifying the video quality benefit derived by the various users for different network conditions and video application characteristics based on various information feedbacks.

Our paper makes the following contributions.

A. Decentralized Information Feedback Driven Cross-Layer Adaptation

In this paper, we show how the various cross-layer strategies can be adapted based on the information feedback. The solutions of centralized flow-based optimizations [1]–[3] have several limitations. First, the video bitstreams are changing over time in terms of required rates, priorities and delays. Hence, it is difficult to timely allocate the necessary bandwidths across the wireless network infrastructure to match these time-varying application requirements. Second, the delay constraints of the various packets are not explicitly considered in centralized solutions, as this information cannot be relayed to a central resource manager in a timely manner. Third, the complexity of the centralized approach grows exponentially with the size of the network and number of video flows. Finally, the channel characteristics of the entire network (the capacity region of the network) need to be known for this centralized, oracle-based optimization. This is not practical as channel conditions are time-varying, and having accurate information about the status of all the network links is not realistic.

Alternatively, we focus on a fully distributed packet-based solution, where timely information feedback can efficiently drive the cross-layer adaptation for each individual multimedia stream as well as the multi-user collaborations in sharing the wireless infrastructure. To cope with the delay sensitivity of the video traffic, we explicitly consider the delay deadlines of the various packets (packets are dropped whenever their deadlines expire) and estimate the remaining transmission time based on the available information feedback. This approach is better suited for the informationally decentralized nature of the investigated multi-user video transmission problem over multihop infrastructures.

B. Impact of Various Information Horizons/Frequencies

We define the mechanism of information feedback conveyed through a multihop overlay infrastructure and investigate the impact of different information horizons/frequencies on the video quality derived by the various multimedia users. We discuss the tradeoff between the increased transmission overhead and the benefit of larger information horizons, which result in improved predictions of network conditions. More information allows nodes in the network to better estimate the time for each

packet to reach its destination and hence, the chance of missing its deadline.

C. Information Feedback Driven Packet Scheduling and Retransmission Strategies

We introduce the concept of *risk estimation* based on the available information feedback that determines the probability that a packet will miss its delay deadline. Based on the estimated risk and the quality impact of the video packet, we proposed novel information feedback driven scheduling and retransmission strategies for each node in the network.

D. Dynamic Routing Policies versus On-Demand Source Routing

We present a novel dynamic routing policy that adapts based on the information feedback to maximize the overall quality of delay-sensitive multimedia. Besides exploring different information horizons under various network conditions to enable a dynamic routing adaptation, we also compare our solution against on-demand source routing that does the information feedback prior to the video transmission (e.g., the route discovery probing in DSR [25], and AODV [26]). Hence, unlike our dynamic routing solution, where the information feedback is performed during the entire video transmission, in these routing schemes there is no transmission overhead associated to the real-time information feedback.

The paper is organized as follows. Section II defines the video and network specification for multi-user video transmission over multihop wireless networks and provides a cross-layer distributed optimization scheme based on the information feedback. In Section III, we discuss the impact of the information feedback with different information horizons and present an integrated cross-layer adaptation algorithm for the real-time multi-user streaming problem. Section IV introduces a novel information feedback driven scheduling algorithm that takes advantage of the larger information horizons. Section V introduces our information feedback driven retransmission limit calculation, and Section VI presents our dynamic routing algorithms based on the information feedback. In Section VII, we discuss the overheads of the information feedback of various parameters. Simulation results are given in Section VIII. Section IX concludes the paper.

II. MULTI-USER VIDEO TRANSMISSION – PROBLEM FORMULATION AND SYSTEM DESCRIPTION

A. Video Classes

We assume that V video users with distinct source and destination nodes are sharing the same multihop wireless infrastructure. In [21], it has been shown that partitioning a scalable video flow (stream) into several prioritized classes can improve the number of simultaneously admitted stations in a congested 802.11a/e WLAN infrastructure, as well as the overall received quality. Hence, we separate each scalable encoded video stream into a certain number of classes (quality layers). Similarly, in this paper, we adopt an embedded 3-D wavelet codec [23] and construct video classes by truncating the embedded bitstream. We assume that the packets within each video class have the

same delay deadline, similar to [11], [21]. The number of classes for a video sequence v equals C_v , and the total number of classes across all users in the network equals $K = \sum_{v=1}^V C_v$. Each video class f_k is characterized by the following.

- λ_k , the expected quality impact of receiving the packets in the class f_k . We prioritize the video classes based on this parameter. In the subsequent part of the paper, we label the K classes (across all users) in descending order of their priorities, i.e., $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K$.
- L_k , the average packet lengths of the class f_k . The expected quality improvement for receiving a video packet in the class f_k is defined as $\lambda_k \cdot L_k$ (see e.g., [11] for more details).
- N_k , the number of packets in the class f_k in one GOP duration of the corresponding video sequence.
- P_k^{succ} , the probabilities of successfully receiving the packets in the class f_k at the destination. Thus, the expected number of the successfully received packets of the class f_k is $N_k \cdot P_k^{succ}$.
- d_k , the delay deadlines of the packets in the class f_k . Due to the hierarchical temporal structure deployed in 3-D wavelet video coders (see [11] and [21]), for a video sequence v , the lower priority packets also have a less stringent delay requirement. This is the reason why we prioritize the video bitstream in terms of the quality impact. However, if the used video coder did not exhibit this property, we need to deploy alternative prioritization techniques $\lambda_k^{video}(\lambda_k, d_k)$ that jointly consider the quality impact and delay constraints (see more sophisticated methods in e.g., [27]).

At the client side, the expected quality improvement for video v in one GOP can be expressed as

$$Q_v^{rec} = \sum_{f_k \in v} (\lambda_k \cdot L_k) \cdot (N_k \cdot P_k^{succ}). \quad (1)$$

Here, we assume that the client implements a simple error concealment scheme, where the lower priority packets are discarded whenever the higher priority packets are lost [11]. This is because the quality improvement (gain) obtained from decoding the lower priority packets is very limited (in such embedded scalable video coders) whenever the higher priority packets are not received. For example, drift errors can be observed when decoding the lower priority packets without the higher priority packets [21]. Hence, we can write

$$P_k^{succ} = \begin{cases} 0, & \text{if } P_{k'}^{succ} \neq 1 \text{ and } f_{k'} \prec f_k \\ (1 - P_k) = E[I(D_k \leq d_k)], & \text{otherwise,} \end{cases} \quad (2)$$

where we use the notation in [27] – $f_{k'} \prec f_k$ to indicate that the class f_k depends on $f_{k'}$. Specifically, if f_k and $f_{k'}$ are classes of the same video stream, $f_{k'} \prec f_k$ means $k' < k$ due to the descending priority ($\lambda_{k'} > \lambda_k$). P_k represents the end-to-end packet loss probability for the packets of class f_k . D_k represents the experienced end-to-end delay for the packets of class f_k . $I(\cdot)$ is an indicator function. Note that the end-to-end probability P_k^{succ} depends on the network resource, competing users'

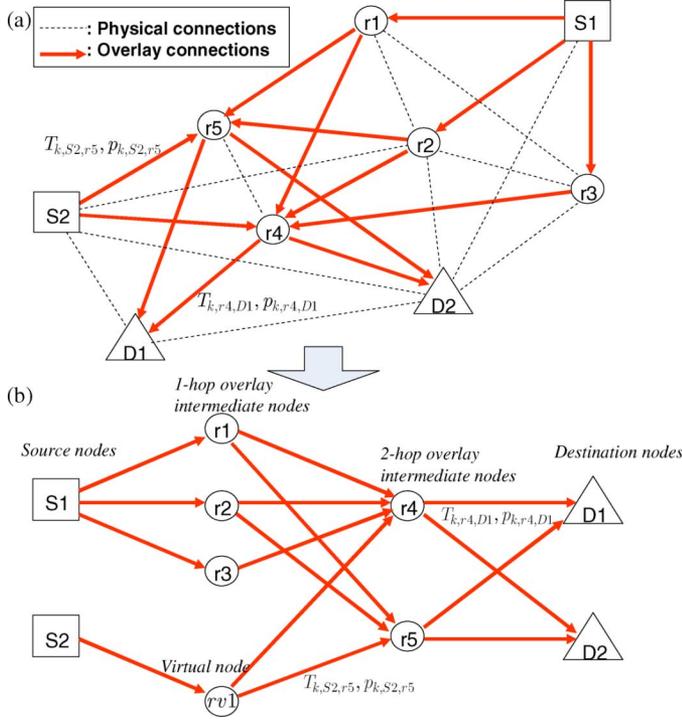


Fig. 1. Directed acyclic multihop overlay network for an exemplary wireless infrastructure. (a) Actual network topology that has two source-destination pairs, five relay nodes. (b) Overlay network topology that has two source-destination pairs, six relay nodes (with one virtual node in the 1-hop intermediate nodes).

priorities as well as the deployed cross-layer transmission strategies. In addition, at the intermediate node m , we assume that the video packets are scheduled in a specific order π_m according to the prioritization associated with the video content characteristics.

B. Overlay Network Specification

We assume that the V source-destination pairs are connected by a directed acyclic multihop overlay network, which is superimposed upon the physical wireless network. This overlay network consists of H hops with M_h intermediate nodes at the h th hop ($1 \leq h \leq H - 1$). The number of source and destination nodes are the same, i.e., $M_0 = M_H = V$, and each node will be tagged with a distinct number m_h ($1 \leq m_h \leq M_h$). Fig. 1 illustrates the overlay network. We define $\beta_{k,m_h,m_{h+1}}$ as the fraction of packets of class f_k at node m_h to select the node m_{h+1} as its relay at the $h + 1$ th hop. We refer to this term as the *relay selecting parameter*. Note that multiple paths could be selected for a class, i.e., $0 \leq \beta_{k,m_h,m_{h+1}} \leq 1$. Whenever an intermediate node m_{h+1} is not reachable for class f_k at node m_h then $\beta_{k,m_h,m_{h+1}} = 0$. Since the total number of intermediate nodes in the $h + 1$ th hop is M_{h+1} , we have $\sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} = 1$.

Importantly, note that the deployed structure is very general and any multihop network that can be modeled as a directed acyclic graph can be modified to fit into this overlay structure by simply adding virtual nodes (virtual hops for different users) [9]. We introduce virtual nodes with zero service time for users

that have a smaller number of hops, and fix the path for particular classes to pass through the virtual node (by enforcing $\beta_{k,m_h,m_{h+1}}$). Fig. 1 gives an example of a 3-hop overlay network with two users ($V = 2$, $H = 3$, $M_0 = M_3 = 2$, $M_1 = 4$, $M_2 = 2$). Methods to construct such overlay structures given a specific multihop network and a set of transmitting-receiving pairs can be found in [19], [20]. Through the multistage overlay infrastructure, the information feedback is performed from the intermediate nodes to all the connected nodes ($\beta_{k,m_h,m_{h+1}} \neq 0$) in the previous hop.

To describe the channel conditions of each transmission link from node m_h to m_{h+1} , i.e., m_h, m_{h+1} , we assume as in [18] that each wireless link is a memoryless packet erasure channel. $p_{k,m_h,m_{h+1}}$ represents the packet error rate over the link m_h, m_{h+1} for the class f_k and $T_{k,m_h,m_{h+1}}^{goodput}$ represent the effective transmission rate (goodput). They can be approximated using the sigmoid function [18]

$$p_{k,m_h,m_{h+1}} = \frac{1}{1 + e^{\zeta_k(SINR - \delta_k)}} \quad (3)$$

$$T_{k,m_h,m_{h+1}}^{goodput} = (1 - p_{k,m_h,m_{h+1}}) T_{k,m_h,m_{h+1}} = \frac{T_{k,m_h,m_{h+1}}}{1 + e^{-\zeta_k(SINR - \delta_k)}} \quad (4)$$

where $SINR$ is the measured Signal-to-Interference-Noise-Ratio (SINR), and ζ_k and δ_k are empirical constants corresponding to the modulation and coding schemes for a given packet length L_k of class f_k . $T_{k,m_h,m_{h+1}}$ represents the maximum transmission rate supported by the optimal modulation and coding scheme. To cope with the packet error, at the MAC layer, we assume the network deploys a protocol similar to that of IEEE 802.11a/e, which enables packet-based retransmission. Let $\gamma_{k,m_h,m_{h+1}}^{MAX}$ represent the maximum number of retransmissions for the packets of class f_k over the link (m_h, m_{h+1}) .

C. Centralized Cross-Layer Optimization for Multi-User Wireless Video Transmission

We define STR_{m_h} as the cross-layer transmission strategy vector for packets at the node m_h consisting of the packet scheduling policy π_{m_h} , the relay selecting parameters $\beta_{k,m_h,m_{h+1}}$ for routing, the MAC retransmission limit $\gamma_{k,m_h,m_{h+1}}^{MAX}$ per link, i.e., $STR_{m_h} = [\pi_{m_h}, \beta_{k,m_h,m_{h+1}}, \gamma_{k,m_h,m_{h+1}}^{MAX}] \in \mathcal{A}_{tot}$. And $\mathcal{A}_{tot} = \mathcal{A}_{APP} \times \mathcal{A}_{NET} \times \mathcal{A}_{MAC}$ represents the set including all the feasible cross-layer transmission strategy vector, where \mathcal{A}_{APP} is the set of all feasible packet scheduling strategies, \mathcal{A}_{NET} is the set of all possible selections of relays, and \mathcal{A}_{MAC} is an integer set from 0 to the maximum retransmission limit supported by the MAC protocol. Then, assuming the global information \mathcal{I}_{global} is available, the investigated multi-user wireless video transmission problem can be formulated as a centralized delay-driven cross-layer optimization

$$\begin{aligned} \mathbf{STR}^{opt} &= \arg \max_{\mathbf{STR} \in \mathcal{A}_{tot}^{|M|}} \sum_{v=1}^V Q_v^{rec}(\mathbf{STR}, \mathcal{I}_{global}) \\ &= \arg \max_{\mathbf{STR} \in \mathcal{A}_{tot}^{|M|}} \sum_{v=1}^V \sum_{f_k \in v} \lambda_k L_k N_k P_k^{succ}(\mathbf{STR}, \mathcal{I}_{global}) \end{aligned} \quad (5)$$

where $\mathbf{STR} = [STR_{m_h} \mid m_h \in \mathbf{M}]$, and \mathbf{M} represents the set of nodes at which the transmission strategies decisions can be made for the video packets. $|\mathbf{M}| = \sum_{h=0}^{H-1} M_h$ is the number of the nodes in \mathbf{M} . Since the successfully received packets of each class f_k must have their end-to-end delay D_k smaller than their corresponding delay deadline d_k , the constraint of the optimization is $D_k(\mathbf{STR}) < d_k, k = 1, \dots, K$. Due to the priority queuing and the error concealment scheme in (2), the optimal solution of (5) serves the more important packets instead of transmitting as many packets as possible. Although the centralized optimization provides optimal solution for the multi-user video streaming problem, it suffers from the unrealistic assumption of collecting timely global information across the multihop network for the delay-sensitive applications. Due to the informationally decentralized nature of the multihop wireless networks, the centralized solution is not practical for the multi-user video streaming problem. For instance, the optimal solution depends on the delay incurred by the various packets across the hops, which cannot be timely relayed to a central controller. Moreover, the complexity of the centralized optimization grows exponentially with number of classes and nodes in the network. Hence, the optimization might require a large amount of time to process and the collected information might no longer be accurate by the time transmission decisions need to be made.

D. Proposed Distributed Cross-Layer Adaptation Based on Information Feedback

Instead of gathering the global information \mathcal{I}_{global} , we propose a distributed suboptimal solution that collects the local information feedback \mathcal{I}_{local} at the node m_h to maximize the expected quality of the various users sharing the same multihop wireless infrastructure

$$STR_{m_h}^{opt} = \arg \max_{STR \in \mathcal{A}_{tot} \forall f_k \text{ at } m_h} \sum \lambda_k L_k N_{k,m_h} E[P_k^{succ}(STR_{m_h}, \mathcal{I}_{local})] \quad (6)$$

where N_{k,m_h} represents the number of packets of class f_k present in the queue at the node m_h .

In this paper, we define \mathcal{I}_{local} with the following information feedback parameters.

- $SINR$, the SINR to calculate the channel conditions over each link of the overlay network [see (3) and (4)].
- P_{k,m_h} , the packet loss probability of the class f_k through the intermediate node m_h . The parameter illustrates the bottleneck identification for various video classes. This information can be used by the application layer to decide how many quality layers are transmitted or to adapt its encoding parameters (in the case of real-time encoding) to improve its video quality performance given the current number of users, priorities of the competing streams and network conditions, but also, importantly, to alleviate the network congestion.
- $E[Delay_{k,m_h}]$, the expected delay from the intermediate node m_h to the destination node of the class f_k to convey the congestion information of the network, which is essential for the delay-sensitive applications.

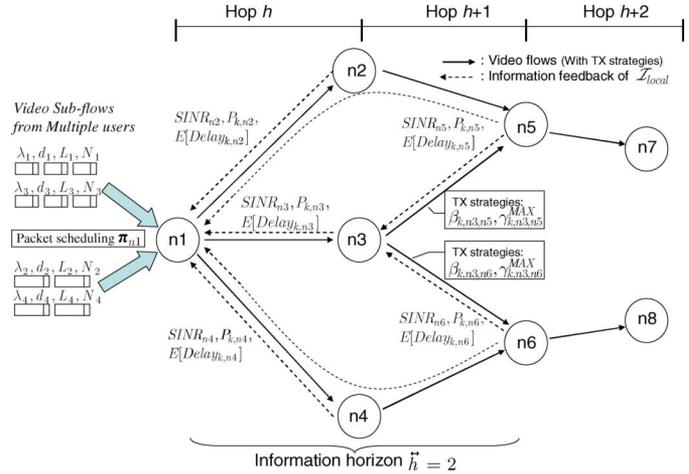


Fig. 2. Illustrative example of an application layer overlay network with information horizon $\bar{h} = 2$.

Let us consider the simple example in Fig. 2 that illustrates how information feedback is deployed. The term information horizon will be defined in Section III. In this example, node n1 is an intermediate node that needs to relay multiple video classes from various users. In order for the relay n1 to determine the optimized cross-layer transmission strategies, at least 1-hop information feedback is required. The network status information can be disseminated at frequent intervals over the overlay infrastructure, and it is considered to be known at the decision relay n1. However, in certain cases, feedback information from some hops (beyond the information horizon) may arrive with an intolerable delay, and may be unreliable due to the rapidly-changing network conditions.

In this paper, we make the following assumptions for performing the information feedback and the delay estimation $E[Delay_{k,m_h}]$. First, we assume a polling-based contention-free media access (which is similar to the deployed IEEE 802.11e [16] and 802.11s [17] standards) that dynamically reserves transmission opportunities within a service interval t_{SI} [16], and the network status (such as the topology, the transmission rate $T_{k,m_h,m_{h+1}}$ and the packet error rate $p_{k,m_h,m_{h+1}}$ for each link) remains unchanged in t_{SI} . Second, because of the retransmission in the MAC layer protection, the effective packet transmission time can be formulated as a geometric distribution [30] with $T_{k,m_h,m_{h+1}}$, $p_{k,m_h,m_{h+1}}$, and packet length L_k (as discussed in Section III-B). Third, for simplification, the arrival of the packets at each intermediate node is regarded as a Poisson arrival process, which is reasonable if the number of intermediate nodes is large enough and the selection of paths is relatively balanced. Fourth, we assume that the queue waiting time dominates the overall delay. Under these assumptions, we can estimate the risk that packets from different priority classes will not arrive at their destination before their decoding deadline expires (see Section IV for more detail). The adaptation of π_{m_h} , $\gamma_{k,m_h,m_{h+1}}^{MAX}$, and the dynamic routing policies for $\beta_{k,m_h,m_{h+1}}$ can be deployed in a distributed manner based on the information feedback. Next, we discuss the mechanism of performing the information feedback through the directed acyclic overlay network.

III. IMPACT OF ACCURATE NETWORK STATUS

Since the network conditions can rapidly vary in multihop network infrastructures, the performance of any video streaming solution will significantly depend on the availability of accurate network information. Three key aspects for multi-user video streaming are influenced by the availability, accuracy, and timeliness of this information feedback.

- *Decentralized decision making* – network nodes can be implemented to improve their adopted cross-layer strategies based on information feedback about the channel conditions and regional network congestion to avoid unnecessary queuing delay and hence, packet drops.
- *Timely adaptation* – information feedback enables timely adaptation to network changes (e.g., nodes leaving or sources of interference appearing or disappearing), which is essential for delay-sensitive multimedia transmission.
- *Inter-user collaboration* – based on information feedback, network resources can be effectively managed and users are able to effectively collaborate to achieve the desired global optimal utility. For instance, in the absence of such information, an intermediate node may waste precious resources by allocating time to packets from classes that will miss their deadlines, thereby preventing other classes which can meet their delay constraint from being transmitted.

A. Information Feedback Frequencies and Information Horizon

The information feedback should be performed in a distributed (per hop) fashion that explicitly considers the dissemination delay. We assume that the information feedback is periodically transmitted to the previous hop every t_{info}^1 seconds during each t_{SI} ($0 < t_{\text{info}} \leq t_{SI}$). We define $f_{\text{info}}(1)$ as the frequency of the information feedback within one hop

$$f_{\text{info}}(1) = \frac{1}{t_{\text{info}}}. \quad (7)$$

We also define the vector $\mathbf{b} = (b_1, b_2, \dots, b_H)$ of the dissemination factors over the network. Let $t_{\text{info}}(h)$ represent the time it takes for the information to be disseminated over h hops

$$t_{\text{info}}(h) = b_h \times t_{\text{info}}, \quad \text{where } b_h \geq 1, \quad \text{for } h \geq 1. \quad (8)$$

Since the network information requires time to pass through the various hops, we have $b_h > b_{h-1}$. We set $b_1 = 1$. Because the information is conveyed hop by hop, $t_{\text{info}}(h)$ also depends on the per-hop information feedback frequency $f_{\text{info}}(1)$. We define $f_{\text{info}}(h)$ as the information feedback frequency when the information is conveyed over h hops in the following way:

$$f_{\text{info}}(h) = \frac{1}{t_{\text{info}}(h)} = \frac{c}{t_{SI} \times b_h} \quad (9)$$

where c is defined as t_{SI}/t_{info} . Since the network conditions are assumed to be unchanged within the service interval t_{SI} , we

¹The time interval is not the time fraction for transmitting the information feedback in a service interval, but rather the time between two subsequent information feedbacks (which includes time for transmitting the video packet, the information feedback and also the protocol overheads).

define the information horizon $\overset{\leftrightarrow}{h}$ as the number of hops from which the information feedback can be accurately disseminated during t_{SI}

$$\overset{\leftrightarrow}{h}(\mathbf{b}, f_{\text{info}}(1)) = \text{maximize } h \\ \text{subject to } t_{\text{info}}(h) \leq t_{SI}, \quad h = 1, \dots, H. \quad (10)$$

In [7] and [18], the dissemination time for the information feedback is proportional to the number of hops across which the information feedback is traversed, i.e., $\mathbf{b}_p = (1, 2, 3, \dots, H - 1, H)$, and if we assume that c is an integer, the relationship between $\overset{\leftrightarrow}{h}$ and $f_{\text{info}}(1)$ becomes a linear function

$$\overset{\leftrightarrow}{h}(\mathbf{b}_p, f_{\text{info}}(1)) = f_{\text{info}}(1) \times t_{SI} = c. \quad (11)$$

We focus on the impact of different information horizons directly on the video qualities of multiple users sharing the same multihop wireless network. Note that $f_{\text{info}}(1)$ can be converted into an information horizon based on (10), as long as the information dissemination factors (i.e., \mathbf{b} vector) are given. Thus, for simplicity, in the remainder part of the paper, we denote the information horizon $\overset{\leftrightarrow}{h}(\mathbf{b}, f_{\text{info}}(1))$ by simply $\overset{\leftrightarrow}{h}$. An example of $\overset{\leftrightarrow}{h} = 2$ is shown in Fig. 2. The local information feedback in (6) for a larger information horizon becomes a vector

$$\mathcal{I}_{\text{local}} = [SINR_{m_h}, P_{k,m_h}, E[\text{Delay}_{k,m_h}]]|_{m_h} \\ \in \mathbf{M}_{k,h+1} \cup \dots \cup \mathbf{M}_{k,h+\overset{\leftrightarrow}{h}}, \forall k]$$

where $\mathbf{M}_{k,h+1}$ represents a set of nodes in the $h + 1$ th hop that feedback the information for the class f_k traffic to the decision nodes (e.g., node “n1” in the example in Fig. 2).

B. Impact of Various Information Horizons

With a larger information horizon, more accurate network status can be obtained, which can be used to adapt the cross-layer transmission strategies at various layers. A larger information horizon ensures that the information can be obtained in a timely manner and network status can be estimated more accurately. For example, a better routing decision can be determined to avoid the congested regions in the network. This decreases the packet loss probability P_k for each class, thus increasing the P_k^{succ} for the important classes and improving the received video qualities. However, the penalty of the overhead is seldom jointly considered in the prior works. Let $E[\text{Time}_{k,m_h}^{\text{packet}}(\overset{\leftrightarrow}{h})]$ represent the expected transmission time for a video packet in class f_k at node m_h to the next hop with the information feedback of horizon $\overset{\leftrightarrow}{h}$. Based on the geometric assumption, we can write (12), as shown at the bottom of the next page, which is calculated as an average transmission time over all the possible relays m_{h+1} in the next hop. $\text{Time}^{\text{over}}(\overset{\leftrightarrow}{h})$ denotes the time overhead introduced by the various protocols [16] including the time of waiting for the MAC acknowledgements etc., and also the information feedback. Consequently, a larger information horizon can induce larger overheads for the packet transmission time, and hence increases the end-to-end delay D_k , which can lead to higher packet losses P_k as they can miss their deadline [see (2)]. In this paper, we assume that the time overhead is a

known function of the information horizon and we will discuss this in more detail in Section VII.

In general, the information horizon might be different for various users or classes and also can vary per node, depending on its location, congestion level, etc. Thus, a scalable information feedback can be implemented (e.g., the information horizon can depend on class f_k and node m_h). For instance, to reduce the overhead associated with the information feedback, some less important classes can have smaller horizons. However, for simplicity, the information horizon is assumed to be the same for all classes (users) in the rest of the paper. The topic of implementing the scalable information feedback and the analysis of its impact form a topic of our future research.

C. Distributed Cross-Layer Adaptation Based on the Information Feedback With Larger Information Horizon

Instead of performing the exhaustive search for the distributed optimization in (6), we present the following iterative cross-layer adaptation to solve the multi-user video streaming problem. Based on the information feedback, the goal of the distributed cross-layer adaptation is to determine an optimal packet present in the queue (from π_{m_h}) to be transmitted through the optimal relay m_{h+1} (from $\beta_{k,m_h,m_{h+1}}$) in the next hop with the optimal retransmission limit (from $\gamma_{k,m_h,m_{h+1}}^{\text{MAX}}$).

1) To determine a packet of a specific class f_k for transmission, the packet scheduling policy π_{m_h} in the queue of the intermediate node m_h is optimized to first transmit the video packets with larger λ_k , since they have a higher impact on the overall video quality. With a larger information horizon, such packet scheduling can be improved as we will discuss it in **Section IV**.

2) To solve the routing problem, we deploy a priority queuing approach based on the information feedback and apply dynamic routing policies similar to the Bellman-Ford routing algorithm [10]. We exploit the $E[\text{Delay}_{k,m_h}]$ in the local information feedback \mathcal{I}_{local} . The selection of $\beta_{k,m_h,m_{h+1}}$ is based on the $E[\text{Delay}_{k,m_{h+1}}]$ value that minimizes the end-to-end packet loss probability P_k for the transmitted packet. We will discuss the routing problem in detail in **Section VI**.

3) At the MAC layer, we choose the appropriate retransmission limit $\gamma_{k,m_h,m_{h+1}}^{\text{MAX}}$ per packet based on the $T_{k,m_h,m_{h+1}}^{\text{goodput}}$ such that its delay constraint is satisfied. Based on our prior results [11] in one-hop network, the optimal retransmission strategy is to send the highest priority packet until it is successfully received by the next relay or until its delay deadline expires. Specifically, let d^{curr} represent the current delay incurred by a particular packet at the current node m_h . The maximum retransmission limit for the packet of class f_k over the link

from m_h to m_{h+1} is determined based on the delay deadline d_k (where $\lfloor \cdot \rfloor$ is the floor operation)

$$\gamma_{k,m_h,m_{h+1}}^{\text{MAX}} = \left\lfloor \frac{T_{k,m_h,m_{h+1}}^{\text{goodput}} (d_k - d^{\text{curr}})}{L_k} \right\rfloor - 1. \quad (13)$$

With a larger information horizon, the retransmission limit can be improved as we will discuss it in **Section V**.

4) Then, we measure the SINR and estimate the corresponding $E[\text{Delay}_{k,m_h}]$ and P_{k,m_h} for each class f_k at the node m_h and feed back this information to the nodes in the previous hops within the information horizon \overleftrightarrow{h} .

IV. INFORMATION FEEDBACK DRIVEN SCHEDULING OF PRIORITIZED VIDEO PACKETS

At each intermediate node m_h , in order to optimize the scheduling of the various video packets, we determine the risk $Risk_{k,m_h}$ ($0 \leq Risk_{k,m_h} \leq 1$) that the packets of class f_k will miss their delay deadline, based on the probability that the estimated received time at the destination is *after* their delay deadlines. Higher probabilities of packet loss over the network (due to interference, congestion, nodes leaving etc.) will lead to higher risks of packets missing their delay deadlines. Based on this risk, the scheduling of the various packets of the different classes can be determined to ensure a maximized system quality.

To compute the risk estimation for a packet, we need to consider both the delay deadlines d_k as well as the expected delay $E[\text{Delay}_{k,m_h}]$ in the information feedback \mathcal{I}_{local} conveyed from the intermediate nodes within the information horizon \overleftrightarrow{h} . The video packets at an intermediate node can be divided into three categories:

- packets that will certainly be dropped (“dropped” packets);
- packets that have very high probability to be dropped (“almost-dropped” packets);
- packets that have low probability to be dropped (“seldom-dropped” packets).

“Dropped” packets are video packets with a current cumulative delay d^{curr} exceeding their delay deadline ($d^{\text{curr}} > d_k$). These packets will be dropped at the current node and hence, there is no need to compute their risk. The “almost-dropped” packets have not yet exceeded their delay deadline ($d^{\text{curr}} < d_k$), but their current cumulative delay plus the expected delay to reach the destination does exceed their delay deadline, i.e., $d^{\text{curr}} + E[\text{Delay}_{k,m_h}] > d_k$. We set the risks for these “almost-dropped” packets to be 0, as they have a very high probability of being dropped and hence, they will unnecessarily waste resources that could be used for the successful transmission of “seldom-dropped” packets. The remaining video packets are

$$E[\text{Time}_{k,m_h}^{\text{packet}}(\overleftrightarrow{h})] = \sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} \left(\frac{1 - p_{k,m_h,m_{h+1}}^{\gamma_{k,m_h,m_{h+1}}^{\text{MAX}} + 1}}{1 - p_{k,m_h,m_{h+1}}} \right) \left(\frac{L_k + L_{\text{header}}}{T_{k,m_h,m_{h+1}}} + \text{Time}^{\text{over}}(\overleftrightarrow{h}) \right) \quad (12)$$

“seldom-dropped” packets. Their current cumulative delay plus the expected delay from the current node to the destination is lower than the delay deadline, i.e., $d^{curr} + E[Delay_{k,m_h}] < d_k$. Hence, these packets have a high probability of arriving at the destination on time and their scheduling needs to be optimized to maximize the video quality across the various users. Next, we discuss how to estimate the risk for these seldom-dropped packets.

A. Risk Estimation Based on Priority Queuing Analysis

The risk estimation for the seldom dropped packets is determined based on the priority queuing analysis, by using the approximation of the waiting time tail distribution. Let W_{k,m_h} represent the queue waiting time for class f_k at intermediate node m_h . The waiting time tail distribution can be approximated as [13], [14]

$$\text{Prob}(W_{k,m_h} > t) \approx \left(\sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}] \right) \cdot \exp \left(- \frac{\sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}]}{E[W_{k,m_h}]} \times t \right) \quad (14)$$

where η_{k,m_h} is the measured average input rate and $E[X_{i,m_h}]$ is the average service time of class f_k at the intermediate node m_h . The expected average queue waiting time of the priority queue is

$$\begin{aligned} \int_{t=-\infty}^{\infty} \text{Prob}(W_{k,m_h} > t) dt &= E[W_{k,m_h}] \\ &= \frac{\sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}^2]}{2 \left(1 - \sum_{i=1}^{k-1} \eta_{i,m_h} E[X_{i,m_h}] \right) \left(1 - \sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}] \right)}. \end{aligned} \quad (15)$$

Equation (15) is determined based on the Mean Value Analysis (MVA) of a preemptive-priority M/G/1 queue [10]. Until now, we do not consider the interference incurred in wireless multihop networks (orthogonal transmission channels are available for adjacent wireless links), the average service time $E[X_{k,m_h}]$ is the average packet transmission time $E[Time_{k,m_h}^{packet}(\vec{h})]$ in (12). If the influence of interference is

considered, the average service time $E[X_{k,m_h}]$ can be approximated using a virtual queue analysis similar to the “service on vacation” concept in queuing theory [10], [15]. Using (14), the proposed risk estimation² for the packets in class f_k can be computed as (16), shown at the bottom of the page, where $E[d_k^{left}] = d_k - d^{curr} - E[Delay_{k,m_h}]$ represents the expected time remaining after a packet reaches its destination. We can determine the probability that the waiting time W_{k,m_h} plus a pre-determined time duration $Time^I$, which is a general variable for risk estimation, exceeds the expected time left $E[d_k^{left}]$, and thus, that the packet will be lost. The time duration $Time^I$ can be viewed as an extension of the waiting time for the packet. Larger $Time^I$ values lead to higher risks. An example of the risk estimation is given in Section IV-B. Note that the accuracy of computing the expected time left $E[d_k^{left}]$ increases with a larger information horizon. Thus, the $Risk_{k,m_h}(Time^I, \vec{h})$ also depends on the information horizon \vec{h} and can be better estimated given a larger \vec{h} .

B. Information Feedback Driven Scheduling

In a priority queue, the packet scheduler at an intermediate node transmits first the most important packets (i.e., the packets with the largest λ_k). Each packet is transmitted until the packet is successfully received by the next hop node or until its deadline expires. Assume that there are L total video packets at the intermediate node m_h . Let the application layer packet scheduling $\boldsymbol{\pi}_{m_h} = (\pi_1, \dots, \pi_l, \dots, \pi_L)$, where π_l represents the scheduling order for the video packet $l \in \{1, \dots, L\}$. The basic priority scheduling can be written as

$$\begin{aligned} \boldsymbol{\pi}_{m_h}^{PRI} &= \arg \max_{\boldsymbol{\pi}_{m_h}} \left\{ \sum_{k=1}^K \lambda_k \times N_{k,m_h}(\boldsymbol{\pi}_{m_h}, Time_{\pi}^I) L_k \right\} \\ \text{subject to } \boldsymbol{\pi}_{m_h} &= (\pi_1, \dots, \pi_l, \dots, \pi_L), \\ \pi_l &= drop, \text{ if } l \in f_k, \text{ and } d_l^{curr} \geq d_k \end{aligned} \quad (17)$$

where $N_{k,m_h}(\boldsymbol{\pi}_{m_h}, Time_{\pi}^I)$ is the number³ of packets of the class f_k that are transmitted during a period of time $Time_{\pi}^I$ using a specific packet scheduling $\boldsymbol{\pi}_{m_h}$. The notation $\pi_l = drop$

²The higher risk packets should be sent earlier, since they are with high probability to exceed their deadlines. However, we do not want to waste our resources on those almost-drop packets, hence the risk estimation for these packets are set to zero.

³Packet loss is considered in this number due to the delay constraint that drops packets.

$$\begin{aligned} Risk_{k,m_h}(Time^I) &= \begin{cases} \text{Prob}(W_{k,m_h} + Time^I > E[d_k^{left}]), & \text{if } E[d_k^{left}] > 0 \text{ (seldom-dropped packets)} \\ 0, & \text{if } E[d_k^{left}] \leq 0 \text{ (almost-dropped packets)} \end{cases} \\ &= \begin{cases} \left(\sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}] \right) \exp \left(- \frac{\sum_{i=1}^k \eta_{i,m_h} E[X_{i,m_h}]}{E[W_{k,m_h}]} \times (Time^I - E[d_k^{left}]) \right), & \text{if } E[d_k^{left}] > 0 \\ 0, & \text{if } E[d_k^{left}] \leq 0 \end{cases} \end{aligned} \quad (16)$$

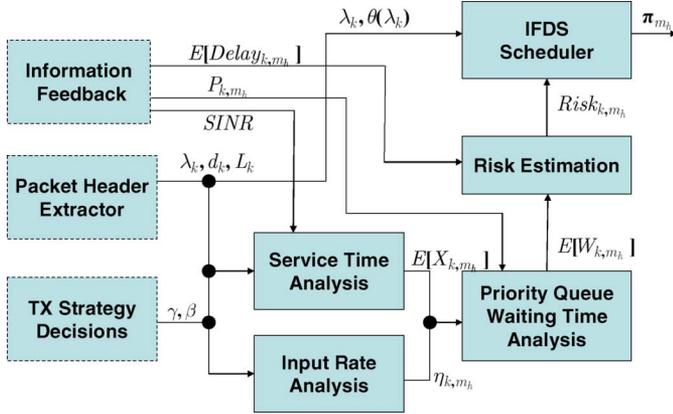


Fig. 3. System map for the IFDS packet scheduling.

indicates that the packet l is not scheduled due to its deadline expiration.

A packet could be dropped in the future hops, as its deadline is exceeded at these hops, and the transmission time of this packet is wasted. This may result in the loss of other packets that would have arrived on time at their destination. Thus, enabled by the information feedback, an intermediate node gathers the network status and makes a scheduling decision. Instead of always transmitting the most important packet in the queue, some other video packets of the different users that are less important but have a higher packet loss probability (risk) can be sent first. Based on this, we propose a novel Information Feedback Driven packet Scheduling (IFDS). The system map of the IFDS scheduling at an intermediate node is illustrated in Fig. 3. The risk is estimated using the information feedback $E[Delay_{k,m_h} | l]$ and the waiting time distribution [see (16)].

For the IFDS scheduling, the video packets ordered in $\pi_{m_h}^{IFDS}$ are transmitted for a pre-determined period of time $Time_{\pi}^{Interval}$. The IFDS scheduling is determined as (18), shown at the bottom of the page. As opposed to the priority queuing scheduling ((17)), the risk of losing a certain class $Risk_{k,m_h}(Time_{\pi}^I, h)$ is considered jointly with the packet quality impact. The scheduler sends the packets in the order that maximizes the output video quality weighted by $\lambda_k Risk_{k,m_h}$ within the time interval $Time_{\pi}^I$. Since different traffic classes have different packet transmission times $E[Time_{k,m_h}^{packet}]$ [see (12)], the number of packets being transmitted per class $N_{k,m_h}(\pi_{m_h}, Time_{\pi}^I)$ depends on which packets are sent (scheduling decision). However, the $Risk_{k,m_h}(Time_{\pi}^I, h)$ remains constant and is independent of the scheduling decision within $Time_{\pi}^I$. Recall that with a larger information horizon \overleftrightarrow{h} , the risk is estimated more accurately because the node is

able to obtain more accurate information from nodes which are closer to the destination. Hence, the packet scheduling policy $\pi_{m_h}^{IFDS}$ is more accurate and adaptive to the network changes than the priority scheduling strategy of (17). Finally, the IFDS scheduling has the following constraint:

$$\pi_{m_h}^{Final} = \left\{ \begin{array}{l} \pi_{m_h}^{IFDS} = (\pi_1, \dots, \pi_l, \dots, \pi_{l'}, \dots, \pi_L) | \pi_l \ll \pi_{l'}, \\ \text{only if } l \in f_k, l' \in f_{k'} \text{ and } \lambda_k > \theta(\lambda_{k'}) \end{array} \right\} \quad (19)$$

where the notation $\pi_l \ll \pi_{l'}$ represents that packet l is scheduled before packet l' . If λ_k belongs to user v , the $\theta(\lambda_k)$ is a class dependent threshold, which can be defined as

$$\theta(\lambda_k \in \text{user } v) = \max_{u \in \{k+1 \dots K\}} \{ \lambda_u | f_u \in \text{the same user } v \}. \quad (20)$$

Equation (20) provides a threshold for a particular class, which is the quality impact value of the next important class of the same user. The reason for the constraint in (19) is to avoid sending an unimportant class with high risk (i.e., for the classes of the same user, packets with higher λ_k must be sent first). This is important since the less important classes depend on the more important classes of the same user and hence, their distortion will be significantly impacted if the higher priority packets are lost [11] [see (2)].

An example of the risk estimation at an intermediate node m_h with fixed \overleftrightarrow{h} is given in Fig. 4 for a case of two users and four classes with the quality impact parameters $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$. User 2 (with classes f_2 and f_3) has a smaller expected time left $E[d_k^{left}]$ than user 1 (having classes f_1 and f_4). Note that when $Time_{\pi}^I \geq E[d_k^{left}]$, $Risk_{k,m_h}(Time_{\pi}^I) = 1$ for all the classes, because they miss their deadlines after waiting for $Time_{\pi}^I$. Let us now adopt the IFDS packet scheduling algorithm, and set the $Time_{\pi}^{Interval}$ between $E[d_1^{left}]$ and $E[d_2^{left}]$. From Fig. 4, we can observe that $Risk_{2,m_h}(Time_{\pi}^I) = 1$ and $Risk_{1,m_h}(Time_{\pi}^I) = 0$. Hence, the packets of class f_1 can wait for $Time_{\pi}^I$ without significantly increasing the packet loss, while the packets of class f_2 that are less important ($\lambda_2 < \lambda_1$) are transmitted.

From the example, we see that the setting of $Time_{\pi}^I$ affects the risk estimation and hence the scheduling decision. Note that if we set $Time_{\pi}^I$ larger than the maximum delay deadline of all the users, the risk will be 1 for all the seldom-dropped packets, and thus the information feedback driven scheduling will only depend on λ_k . If $Time_{\pi}^I$ is set too small, the risk estimations

$$\pi_{m_h}^{IFDS}(\overleftrightarrow{h}) = \arg \max_{\pi_{m_h}} \left\{ \sum_{k=1}^K \lambda_k \times N_{k,m_h}(\pi_{m_h}, Time_{\pi}^I) L_k \times Risk_{k,m_h}(Time_{\pi}^I, \overleftrightarrow{h}) \right\}$$

subject to $\pi_{m_h} = (\pi_1, \dots, \pi_l, \dots, \pi_L)$, $\pi_l = drop$, if $l \in f_k$, and $d_l^{curr} \geq d_k$ (18)

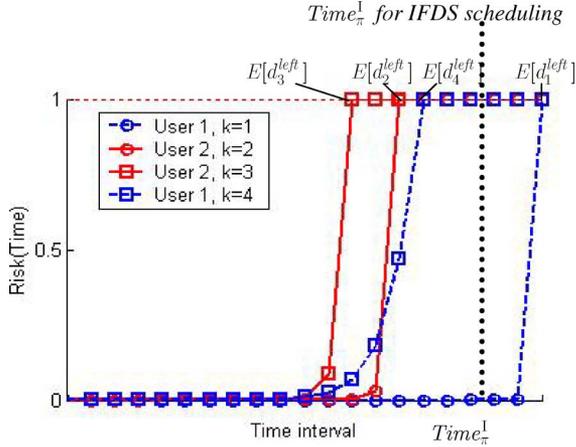


Fig. 4. Risk estimation versus time interval for two users.

will not affect the original priority decision. Thus, we define a lower and an upper bound of the $Time_{\pi}^I$

$$\min_k E[d_k^{left}] \leq Time_{\pi}^I \leq \max_k E[d_k^{left}],$$

$$\text{where } E[d_k^{left}] > 0 \text{ (for seldom-dropped packets)} \quad (21)$$

since the risk estimations are large enough to take effect within this interval. For the example in Fig. 4 $E[d_3^{left}] \leq Time_{\pi}^I \leq E[d_1^{left}]$.

V. INFORMATION FEEDBACK DRIVEN MAC LAYER RETRANSMISSION STRATEGY

For protection over an error-prone wireless link, a retransmission scheme at the MAC layer is adopted. In [11], it was shown that for the scalable video coders such as [23], the video packets should be retransmitted by the MAC until they are received without error or their deadline expires in order to maximize the received video quality. However, if a packet approaches its delay deadline, the risk that it will not reach its destination increases. Hence, similarly to the application layer scheduling strategies discussed in the previous section, we propose a MAC layer information feedback driven retransmission strategy $\gamma_{k,m_h,m_{h+1}}^{IFDS}(\vec{h})$ that explicitly considers the risk of losing a packet based on the available information feedback \mathcal{I}_{local} .

Let *retry* be an integer variable that represents the number of retransmissions for a packet. If the transmission of the packet repeatedly fails, the retransmission should last only until another class of video packets starts to have a higher impact in terms of overall video quality. In both scheduling policies in the previous section, the scheduler will send packets of class f_k having a larger $\lambda_k Risk_{k,m_h}$ value [see (18)]. Therefore, the information feedback driven retransmission limit becomes

$$\begin{aligned} \gamma_{k,m_h,m_{h+1}}^{IFDS}(\vec{h}) &= \text{maximize } \gamma \\ &\text{subject to } \lambda_k Risk_{k,m_h} \left(Time_{\gamma}^I, \vec{h} \right) \\ &\geq \lambda_j Risk_{j,m_h} \left(Time_{\gamma}^I, \vec{h} \right), \\ &\text{for all } j \text{ that } \lambda_j > \theta(\lambda_k) \\ Time_{\gamma}^I &= (\gamma+1) \times Time_{k,m_h,m_{h+1}}^{packet}, \gamma \in \mathbb{N} \end{aligned} \quad (22)$$

which states that the retransmission limit is the maximum number of retries such that the transmitting packet (of class f_k) has a greater $\lambda_k Risk_{k,m_h}$ than other packets in the queue. Due to the scheduling constraint in (19), we only need to check the classes that have a quality impact value larger than the threshold $\theta(\lambda_k)$ in (22). Note that the information feedback driven retransmission limit is always smaller than the retransmission limit in (13) ($\gamma_{k,m_h,m_{h+1}}^{IFDS} \leq \gamma_{k,m_h,m_{h+1}}^{MAX}$), since when a packet approaches the deadline, it will first belong to the ‘‘almost-dropped’’ packets class ($d^{curr} + E[Delay_{k,m_h}] > d_k$), for which $Risk_{k,m_h} = 0$. Thus, another class of packets will be transmitted, thereby terminating the retransmission of the current packet. Consequently, a packet retransmission will first reach the information feedback driven retransmission limit $\gamma_{k,m_h,m_{h+1}}^{IFDS}$ before the delay deadline. Thus, other packets that have a better chance to reach the destinations could be sent earlier.

VI. ADAPTIVE DYNAMIC ROUTING BASED ON INFORMATION FEEDBACK

A. Self-Learning Dynamic Routing With 1-Hop Information Feedback

In this section, we propose a dynamic routing policy for the relay selecting parameters $\beta_{k,m_h,m_{h+1}}$. The decision is based on the information feedback of the expected delay from the current node m_h to the destination, $E[Delay_{k,m_h}]$, for each class of traffic (class f_k ; see Fig. 2). The decisions are made by the following policy:

$$\beta_{k,m_h,m_{h+1}} = \frac{Coeff_k}{1 + \kappa E[Delay_{k,m_{h+1}}]^{\varphi}}. \quad (23)$$

$Coeff_k$ are normalized coefficients to ensure that the summation equals to one

$$Coeff_k = \left[\sum_{m_{h+1} \in \mathbf{M}_{k,h+1}} \frac{1}{1 + \kappa E[Delay_{k,m_{h+1}}]^{\varphi}} \right]^{-1} \quad (24)$$

with κ and φ being constants, which are determined similarly to the balking arrival probability in queuing theory [15]. The value of κ is set depending on the arrival rate according to [15]. The term φ weights the average delay $E[Delay_{k,m_{h+1}}]$ such that the routing policy favors paths leading to a significant lower delay to the destination. Recall that $\mathbf{M}_{k,h+1}$ represents a set of nodes m_{h+1} in the $h+1$ th hop that feedback the information $E[Delay_{k,m_{h+1}}]$. We set $\beta_{k,m_h,m_{h+1}} = 0$ for the nodes whose information feedback is not received, indicating that node m_{h+1} is not connected to node m_h using the overlay infrastructure [8]. We refer to this relay selecting policy as *self-learning policy*, since the decision of $\beta_{k,m_h,m_{h+1}}$ will influence the future information feedback. Recall that $0 \leq \beta_{k,m_h,m_{h+1}} \leq 1$, hence the relay selecting parameters provide a routing description across the network with multipath capability.

The expected delay to the destination of each class is periodically updated at each node using the information feedback from the next hop. If the current node is node m_h at the h th hop, the

1. Assume that changes happen at the hop h' .
The intermediate nodes from $\max(h' - \vec{h}, 0)$ to $h' - 1$ receive the information feedback $E[Delay_{k,m_h}^{new}]$
2. For every class f_k , $k = 1, \dots, K$
3. Let $h = h'$
4. If $h > \max(h' - \vec{h}, 0)$, goto Step 5, else goto Step 8.
5. Recalculate $\beta_{k,m_h,m_{h+1}}^{new} \in \mathbf{B}_k^{new}(\vec{h}, h')$ using Eq.(26)
6. Estimate the $E[W_{k,m_h}^{new}]$ and $E[Delay_{k,m_h}^{new}]$ using Eq.(15), and Eq.(25)
7. Set $h \leftarrow h - 1$ goto Step 4.
8. Using the new relay selecting parameters $\beta_{k,m_h,m_{h+1}}^{new}$ for routing.

Fig. 5. Self-learning dynamic routing algorithm for large information horizon.

expectation of delay to the destination of each class is as follows:

$$E[Delay_{k,m_h}] = E[W_{k,m_h}] + \sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} E[Delay_{k,m_{h+1}}] \quad (25)$$

where $E[Delay_{k,m_{h+1}}]$ is provided from the information feedback of the nodes of the next hop, and the relay selecting parameter $\beta_{k,m_h,m_{h+1}}$ is calculated based on (23). $E[W_{k,m_h}]$ is the average queuing delay at the current queue, which can be either calculated from (15), or measured from the average queue size. Fig. 5 gives the algorithm of the proposed self-learning policy. The self-learning policy will dynamically adapt the relay selection to minimize the delay through the network.

This method is inspired by the Bellman-Ford shortest path (delay) routing algorithm [10] that minimizes the end-to-end delay across the network. Our routing algorithm reduces to the Bellman-Ford algorithm when $\beta_{k,m_h,m_{h+1}} = 1$ to the node m_{h+1} that feedbacks the smallest $E[Delay_{k,m_{h+1}}]$. The differences are that our distributed routing policy enables multi-path capability, and the delay of class f_k is influenced by only the same or higher priority traffic. Since the packet losses P_k result from the violation of the delay constraint, the minimum end-to-end delay D_k provided by our routing policies leads to the minimum end-to-end packet loss [see (2)]. Combined with the optimal packet scheduling and the retransmission limit (to satisfy the delay constraint), the self-learning policy can consequently maximize the P_k^{succ} of the traffic class f_k with higher quality impact [see (6)].

B. Dynamic Routing With Larger Information Horizon

The self-learning policy is the shortest delay routing policy with one hop information horizon ($\vec{h} = 1$). However, if the network conditions or topologies change, this information will be fed back with a certain delay. A larger information horizon ensures that any network changes in the horizon can be timely learned and the cross-layer transmission strategies can quickly adapt to the latest network conditions. With the

information feedback parameters $E[Delay_{k,m_{h+1}}]$, if some significant changes happen in the network at hop h' (such as node failure, link status change etc.), the corresponding relay selecting parameters $\beta_{k,m_{h'},m_{h'+1}}^{new}$ are recalculated based on the new information feedback $E[Delay_{k,m_{h'}}^{new}]$. Note that the expected waiting time and the expected delay toward destination ($E[W_{k,m_{h'-1}}^{new}]$ from (15), and $E[Delay_{k,m_{h'-1}}^{new}]$ from (25)) depend on $\beta_{k,m_{h'},m_{h'+1}}^{new}$ and can then be recalculated recursively. Define $\mathbf{B}_k^{new}(\vec{h}, h')$ as a vector of relay selecting parameters $\beta_{k,m_h,m_{h+1}}^{new}$ of the nodes between the changing hop h' and the hop $h' - \vec{h}$ (if $h' - \vec{h} > 0$, otherwise the recalculation ended at the source nodes)

$$\mathbf{B}_k^{new}(\vec{h}, h') = [\beta_{k,m_h,m_{h+1}}^{new} | h = h', h' - 1, \dots, h' - \vec{h}],$$

where $\beta_{k,m_h,m_{h+1}}^{new} = \frac{Coeff_k}{1 + \kappa E[Delay_{k,m_{h+1}}^{new}]^\varphi}$. (26)

In Fig. 5, we provide the corresponding self-learning dynamic routing algorithms for the increased information horizon.

C. Channel-Aware Self-Learning Routing Policy

In this section, we provide another dynamic routing policy called *channel-aware self-learning* policy that not only determines the relay selecting parameters based on the information feedback parameters $E[Delay_{k,m_{h+1}}]$ (to avoid congested area), but also depends on the channel conditions of transmitting packets to different relays. Denote the operation $\min_N(X)$ as the smallest N choices from a set X. Assume $\mathbf{N}_{k,h+1}^{relay}$ is the set of the intermediate nodes that are selected as the relays for class f_k at the $h + 1$ hop

$$\mathbf{N}_{k,h+1}^{relay} = \arg \min_{m_{h+1}} \{E[Delay_{k,m_{h+1}}]\}, m_{h+1} = 1 \dots M_{h+1}. \quad (27)$$

Note that if $N = 1$, then the algorithm is again equivalent to the Bellman-Ford shortest delay algorithm, since the only relay that will be selected is the one with the smallest $E[Delay_{k,m_{h+1}}]$. If $N > 1$, we determine the relay selecting parameters to minimize the second moment of the service time of the class f_k at node m_h

$$\beta_{k,m_h,m_{h+1}}^{ch} = \begin{cases} \arg \min_{\beta \in \mathcal{A}_{NET}} E[X_{k,m_h}^2], & \text{if } m_{h+1} \in \mathbf{N}_{k,h+1}^{relay} \\ 0, & \text{if } m_{h+1} \notin \mathbf{N}_{k,h+1}^{relay} \end{cases}. \quad (28)$$

Note that the second moment of the service time is

$$E[X_{k,m_h}^2] \approx \sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} \times E[(Time_{k,m_h,m_{h+1}}^{packet})^2]. \quad (29)$$

Using the same soft minimum (probabilistic) approach in the previous two sections, we rewrite (28) as

$$\beta_{k,m_h,m_{h+1}}^{ch} = \begin{cases} \frac{Coeff_{m_h,k}}{1 + \kappa E[(Time_{k,m_h,m_{h+1}}^{packet})^2]}, & \text{if } m_{h+1} \in \mathbf{N}_{k,h+1}^{relay} \\ 0, & \text{if } m_{h+1} \notin \mathbf{N}_{k,h+1}^{relay} \end{cases}. \quad (30)$$

$Coeff_{m_h,k}$ are normalized coefficients to make sure that the summation equals to one

$$Coeff_{m_h,k} = \left[\sum_{m_{h+1} \in \mathcal{N}_k^{relay}} \frac{1}{1 + \kappa E[(Time_{k,m_h,m_{h+1}}^{packet})^2]^\varphi} \right]^{-1}. \quad (31)$$

This routing policy is different from the self-learning routing policy that minimizes the end-to-end cumulative queuing delay. The channel-aware self-learning routing policy distributes the traffic through the links with better channel conditions that directly affect the $E[(Time_{k,m_h,m_{h+1}}^{packet})^2]$, and the influence of channel conditions can be emphasized by using a larger $\mathcal{N} = |\mathcal{N}_{k,h+1}^{relay}|$.

The channel-aware self-learning routing policy has a property to reduce the uncertainty of the packet loss probability. Note that the packet loss probability has the following upper bound of (32), shown at the bottom of the page. The new routing policy selects the relay selecting parameters $\beta_{k,m_h,m_{h+1}}^{ch}$ to decrease the upper bound of the packet loss probability. First, we reduce the term $\sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} E[Delay_{k,m_{h+1}}]$ by enforcing $\beta_{k,m_h,m_{h+1}}^{ch} > 0$ only for nodes m_{h+1} with a smaller $E[Delay_{k,m_{h+1}}]$ [see (27)]. Second, we try to decrease the second moment of the queue waiting time W_{k,m_h} , which is related to the higher moments of the service time distribution, i.e., $\beta_{k,m_h,m_{h+1}}^{ch}$ are selected to have a service time (transmission time of a packet) with small variation. In order to decrease those higher moments of the service time distribution, besides limiting the number of nonzero relay selection parameters for the next hop nodes, we also determine $\beta_{k,m_h,m_{h+1}}^{ch}$ by minimizing $E[X_{k,m_h}^2]$ to decrease the variation of the service time.

The channel-aware self-learning routing policy for a larger information horizon can also apply the same algorithm stated in Fig. 5 by substituting (26) with (30).

D. Proposed Dynamic Priority Hybrid Routing Algorithm

The self-learning dynamic routing policy introduced in Section VI-A attempts to minimize the average end-to-end delay toward the destination, while the channel-aware self-learning routing policy introduced in Section VI-C attempts to decrease the upper bound of the packet loss probability (resulting from deadline expiration). Usually, the important video packets (with larger quality impact λ_k) will be sent first. To reduce the uncertainty of the packet loss rate is more important to these packets, since they experience significantly small average queuing delay (as well as the packet loss rate). The channel-aware self-learning routing policy is more suitable for these packets. On the other hand, some other packets of the less important classes experience larger queuing delay in

the network, and the self-learning routing policy that ensures shortest delay is essential for these packets. Consequently, with a pre-determined threshold λ^{TH} , we choose the channel-aware self-learning routing for the important classes with $\lambda_k > \lambda^{TH}$, and the self-learning dynamic routing for the rest of the classes. We refer to this hybrid routing policy as Dynamic Priority Hybrid (DPH) routing algorithm. Our DPH routing can also apply the algorithm in Fig. 5 for larger information horizons.

VII. OVERHEAD ANALYSIS FOR INFORMATION FEEDBACK

The information feedback can enable the cross-layer adaptation of video streaming over a multihop network. As the information horizon increases, the network status can be estimated more timely and accurately, and the cross-layer strategies can be improved for the delay-sensitive applications. However, a larger information horizon also consumes more network resources for video transmission and results in an increased time overhead per packet transmission, $Time_k^{over}(\overleftrightarrow{h})$ [see (12)]. Various information feedback parameters have different transmission overheads. In this paper, we take the three information feedback parameters illustrated in Fig. 2 as examples.

Assuming a certain topology, let us perform a worst-case analysis to quantify the maximum information feedback. We assume that the information feedback overheads are I_{SINR} , I_{Ploss} , $I_{E[Delay]}$ for the three information feedback parameters, respectively. We assume that the average number of nodes in one hop is M , the number of total classes is K , and we set the information horizon as $\overleftrightarrow{h}_{max}$ for all users (classes). The SINR information is fed back from potential receivers to the transmitters to enable the link adaptation as well as to facilitate the polling control signaling. Thus, an information horizon of only 1 hop is sufficient for the adopted overlay infrastructure, and the overhead in terms of the information feedback unit is $M^2 I_{SINR}$. As for the other two information feedback parameters, the parameters are required across the whole information horizon and different for all the classes. An aggregation scheme \mathcal{G} can be applied to reduce the repeated information (as in e.g., [8], [31]). The worst-case overheads in terms of the information feedback unit are $\overleftrightarrow{h}_{max} KM \cdot I_{Ploss}(\overleftrightarrow{h}_{max}, \mathcal{G})$ and $\overleftrightarrow{h}_{max} KM \cdot I_{E[Delay]}(\overleftrightarrow{h}_{max}, \mathcal{G})$, respectively. $I_{Ploss}(\overleftrightarrow{h}_{max}, \mathcal{G})$ and $I_{E[Delay]}(\overleftrightarrow{h}_{max}, \mathcal{G})$ represents the functions of aggregated information feedback over $\overleftrightarrow{h}_{max}$ hops for these two information feedback parameters. In conclusion, the information feedback overhead increases with the information horizon.

VIII. SIMULATION RESULTS

To assess the importance of information feedback, we consider several multi-user video transmission scenarios. Two

$$P_{k,m_h} \leq P_{k,m_h}^{bound}(\beta) = \text{Prob} \left(W_{k,m_h} > d_k - d^{curr} - \sum_{m_{h+1}=1}^{M_{h+1}} \beta_{k,m_h,m_{h+1}} E[Delay_{k,m_{h+1}}] \right) \quad (32)$$

TABLE I
 DESCRIPTION FOR THE FOUR CASES OF THE SIMULATION RESULTS ($t_{SI} = 100$ ms)

Scenario	Information horizon \vec{h}	Information feedback interval t_{info}	Overhead per packet $Time^{\text{over}}(\vec{h})$	Equivalent overhead/packet ratio of "Coastguard"	Equivalent overhead/packet ratio ($T_m=300$ Kbps)
$c=1$	1 hop	100ms	0.1 ms	$2.1 \times 10^{-8} T_m$	6.3/1000
$c=2$	2 hops	50ms	0.2 ms	$4.2 \times 10^{-8} T_m$	12.6/1000
$c=3$	3 hops	33ms	0.3 ms	$6.3 \times 10^{-8} T_m$	18.9/1000
$c=4$	4 hops	25ms	0.4 ms	$8.4 \times 10^{-8} T_m$	25.2/1000

 TABLE II
 SIMULATION RESULTS FOR IFDS SCHEDULING WITH VARIOUS INFORMATION HORIZONS AND DIFFERENT NETWORK EFFICIENCIES

Mobile (1668 Kbps)			Packet loss probability P_k (delay deadline $d_k = 500$ ms) for different T_m (Kbps)												
f_k	λ_k (dB/Kb)	N_k	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Optimal Value
			300	400	500	300	400	500	300	400	500	300	400	500	
f_1	0.0340	278	0%	0.3%	0%	5.5%	2.0%	0%	7.8%	1.2%	0.2%	1.3%	0.8%	0%	0%
f_3	0.0128	166	21%	8.1%	3.3%	62%	18%	3.0%	68%	16%	4.1%	51%	18%	4.2%	0%
f_6	0.0084	166	79%	30%	12%	100%	69%	15%	100%	52%	19%	100%	55%	19%	0%
f_8	0.0062	222	100%	95%	83%	100%	100%	83%	100%	100%	82%	100%	100%	90%	0%
Y-PSNR (dB)			29.46	30.98	31.66	28.66	30.02	31.24	28.30	30.16	31.22	28.23	29.97	31.16	33.12
Coastguard (1500 Kbps)			Packet loss probability P_k (delay deadline $d_k = 300$ ms) for different T_m (Kbps)												
f_k	λ_k (dB/Kb)	N_k	Scenario 1			Scenario 2			Scenario 3			Scenario 4			Optimal Value
			300	400	500	300	400	500	300	400	500	300	400	500	
f_2	0.0210	250	33%	10%	11%	7.9%	5.7%	1.7%	8.1%	4.8%	3.9%	8.0%	6.4%	1.2%	0%
f_4	0.0128	150	100%	96%	41%	95%	65%	51%	97%	67%	27%	99%	43%	38%	0%
f_5	0.0096	150	100%	100%	96%	100%	100%	100%	100%	100%	98%	100%	99%	96%	0%
f_7	0.0084	200	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	0%
Y-PSNR (dB)			28.51	29.35	30.34	29.92	30.49	31.13	29.81	30.76	31.84	29.73	30.68	31.67	35.61

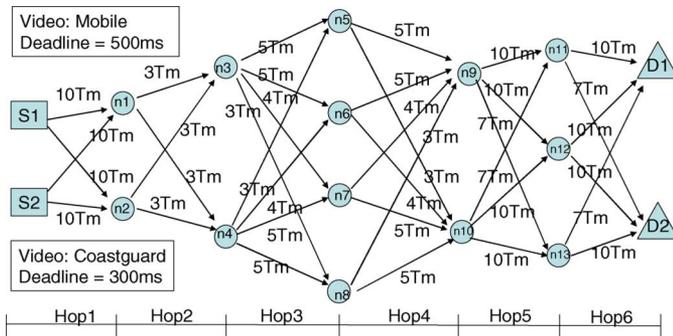


Fig. 6. Simulation settings of a 6-hop overlay network with two video sequences.

video sequences, "Mobile" and "Coastguard" (16 frames per GOP, frame rate of 30 Hz, CIF format) compressed using a scalable video codec [23] are sent from distinct sources to their corresponding destinations through the multihop wireless network shown in Fig. 6. We consider four different scenarios with various information horizons and information feedback overheads as stated in Table I. Each video sequence is divided into four classes ($C_v = 4$, $K = 8$). The quality impact parameters λ_k and the number of packets N_k in one group of picture for each class are shown in the left side of Table II.

In our simulation, we captured the packet-loss pattern under different channel conditions (described in the paper by the link SINR) using our wireless streaming test-bed [8]. In this way, we can assess the efficiency of our system under real wireless channel conditions and link adaptation mechanisms currently deployed in state-of-the-art 802.11a/g wireless cards with 802.11e extension. Link adaptation selects one appropriate physical-layer mode (modulation and channel coding) depending on the link condition, in order to continuously maximize the experienced goodput [8]. Hence, each link in our network settings shown in Fig. 6 is assigned with an effective transmission rate measured from the test-bed. The parameter T_m represents the streaming efficiency of the network. The various efficiency levels are represented by varying the available time fraction for the contention-free period in the polling-based MAC protocol, which induces the various available transmission rates for the video packets over the links. In our event-driven simulation, these network efficiency levels range from 300 Kbps to 500 Kbps. A larger T_m gives higher network efficiency. We set $t_{SI} = 100$ ms, $\mathbf{b} = \mathbf{b}_p$ (see Section III) and \vec{h} varies from 1 to 4 for the four scenarios. The information feedback overheads are set as $Time^{\text{over}}(\vec{h}) = t_{SI}/1000 \times \vec{h}$ for all the classes. Note that the time overhead is limited, i.e., 2.5% of the average packet transmit time when $T_m = 300$ Kbps, and $\vec{h} = 4$.

A. IFDS Scheduling and Retransmission Limit

Note that the effect of the IFDS scheduling depends on many factors, such as the network topology, application characteristics, network transmission efficiency, and congestion/interference conditions, etc. Here, we would like to assess the importance of the risk consideration in resource-constrained networks. We set the application playback delay deadlines are set to 500 ms and 300 ms for the classes of the two video sequences respectively. The transmission rates of the links in the first hop are, relatively higher than the subsequent links. Consequently, most of the packets of the various classes will be queued at the specific intermediate nodes $n1$ and $n2$ (some of them will still be left in the source queues), and the effect of risk can be highlighted for two streams with different delay deadlines.

We adopt the IFDS scheduling and the retransmission limit algorithm in Sections IV and V for cases with larger information horizons ($\vec{h} \geq 2$). In scenario 1, we make the packet scheduling first transmit the packets with the highest quality impact parameter λ_k until the transmission success or delay deadline expiration [i.e., (17)]. In scenario 2, the risk estimation is considered jointly with the quality impact parameters using (18). In scenarios 3 and 4, larger information horizons are used in (18) for risk estimation. However, with larger information horizon, the performance degrades due to larger information feedback overheads. The simulation results of the packet loss rate of each class at their destinations are shown in Table II under various network transmission efficiencies. Since the delay deadline of the ‘‘Coastguard’’ sequence is smaller, it has higher packet loss rate, especially in networks with low transmission efficiency. However, it is shown that as the information horizon increases, the IFDS scheduling sends more ‘‘Coastguard’’ packets to improve its video quality without degrading significantly the video quality of the ‘‘Mobile’’ sequence.

To observe the impact of the various information horizons on the overall video quality, the average Y-PSNR decoded at the destinations of the two sequences are shown in Fig. 7. It shows that the optimal choice of information horizon varies with the network transmission efficiency. For networks with high transmission efficiency, a larger information horizon ($\vec{h} \geq 3$) makes the IFDS scheduling more efficient, and improves the video qualities. However, for a network with low transmission efficiency that is more congested, a shorter information horizon ($\vec{h} \leq 2$) results in better performance since the limited network resource can be focused on the video transmission (payload).

B. Adaptive Dynamic Routing

We compare our DPH routing algorithm with an on-demand multipath routing algorithm AOMDV [29] that selects several loop-free link-disjoint paths based on a well-known on-demand routing algorithm AODV [26]. For the on-demand routing, each video source requires a route discovering period (depends on the round-trip delay to the destination; here, we assume 100 ms) [29] before streaming the video. Although this on-demand routing saves additional time overheads for packet transmissions, it sacrifices the tolerable delay when transmitting the delay-sensitive video packets. In Table III, we also compare the

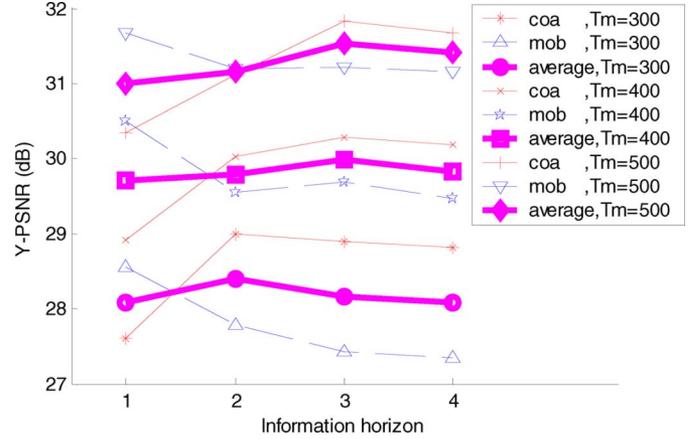


Fig. 7. Y-PSNR versus various information horizon cases under different network transmission efficiencies.

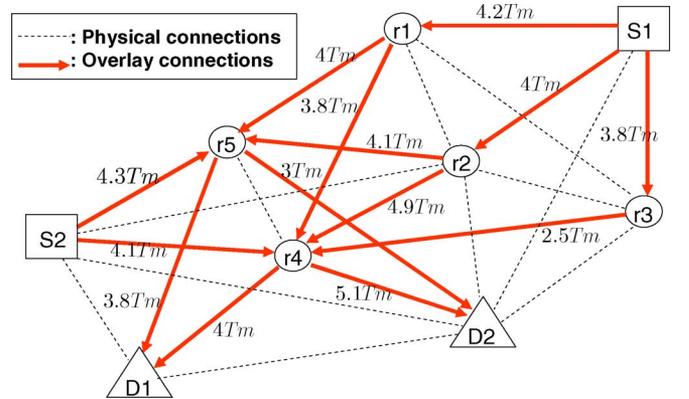


Fig. 8. Overlay network simulation with two video streams over the wireless mesh network.

TABLE III
COMPARISON OF THE DPH ROUTING POLICY WITH THE ON-DEMAND MULTIPATH-ALGORITHMS

Method	$T_m = 0.3$ (Mbps) Low Network Efficiency		$T_m = 0.5$ (Mbps) Medium Network Efficiency	
	‘‘Mobile’’ Y-PSNR (dB)	‘‘Coastguard’’ Y-PSNR (dB)	‘‘Mobile’’ Y-PSNR (dB)	‘‘Coastguard’’ Y-PSNR (dB)
Proposed DPH	29.54	30.71	32.00	33.10
AOMDV [29]	28.22	28.50	31.31	31.52

proposed DPH routing with the on demand routing upon a more general overlay structure over wireless mesh network shown in Fig. 8. The same two video sequences are streamed from node S1 to D1 and from node S2 to D2 with delay deadline 500 ms, respectively. In summary, the simulation results show that our dynamic routing policy outperforms existing on-demand routing solutions by more than 2 dB in terms of the received video quality in low efficiency wireless networks.

IX. CONCLUSION

In this paper, we investigate the impact of information feedback with different network horizons on the video quality of multiple users sharing the same multihop wireless network. We illustrate how the various cross-layer strategies can be adapted to take advantage of the available information feedback from a larger network horizon through the proposed information feedback driven scheduling, retransmission limit and the dynamic

priority hybrid routing algorithm. Unlike the end-to-end feedback that exists in today's networking protocols (such as the rate control in TCP), the information feedback is performed in a distributed (per hop) fashion that explicitly considers the instantaneous delays, which is essential for supporting delay-sensitive multimedia applications. We investigate the tradeoff between the increased transmission overhead and the benefit of larger information horizons leading to an improved prediction of network conditions. The results show that in a network with higher transmission efficiency, a larger information horizon can lead to an improved performance in terms of video quality, which leads to more than 2 dB improvement in video quality as a result of balancing the effect of different delay deadline among users. However, with lower transmission efficiency, smaller information horizon performs better by ensuring limited overhead of the information feedback.

REFERENCES

- [1] W. Wei and A. Zakhor, "Multipath unicast and multicast video communication over wireless ad hoc networks," in *Proc. Int. Conf. Broadband Networks*, 2002, pp. 496–505.
- [2] E. Setton, T. Yoo, X. Zhu, A. Goldsmith, and B. Girod, "Cross-layer design of Ad hoc networks for real-time video streaming," *IEEE Wireless Commun. Mag.*, pp. 59–65, Aug. 2005.
- [3] B. Awerbuch and T. Leighton, "Improved approximation algorithms for the multi-commodity flow problem and local competitive routing in dynamic networks," in *Proc. 26th ACM Symp. Theory of Computing*, May 1994.
- [4] S. Toumpis and A. J. Goldsmith, "Capacity regions for wireless Ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 736–748, Jul. 2003.
- [5] A. Butala and L. Tong, "Cross-layer design for medium access control in CDMA Ad hoc networks," *EURASIP J. on Applied Signal Processing*, to be published.
- [6] Y. Wu, P. A. Chou, Q. Zhang, K. Jain, W. Zhu, and S. Y. Kung, "Network planning in wireless Ad Hoc networks: A cross-layer approach," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 136–150, Jan. 2005.
- [7] Y. Andreopoulos, N. Mastrorade, and M. van der Schaar, "Cross-layer optimized video streaming over wireless multi-hop mesh networks," *IEEE J. Select. Areas Commun.*, to be published.
- [8] D. Krishnaswamy and J. Vicente, "Scalable adaptive wireless networks for multimedia in the proactive enterprise," *Intel Technol. J.*
- [9] J. R. Evans and E. Minieka, *Optimization Algorithms for Networks and Graphs*. New York: Marcel Dekker, 1993.
- [10] D. Bertsekas and R. Gallager, *Data Networks*. Upper Saddle River, NJ: Prentice-Hall, 1987.
- [11] M. van der Schaar and D. S. Turaga, "Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission," *IEEE Trans. Multimedia*, vol. 9, pp. 185–197, Jan. 2007.
- [12] D. Qiao, S. Choi, and K. G. Shin, "Goodput analysis and link adaptation for IEEE 802.11a wireless LAN," *IEEE Trans. Mobile Comput.*, vol. 1, no. 4, pp. 278–292, Oct.-Dec. 2002.
- [13] T. Jiang, C. K. Tham, and C. C. Ko, "An approximation for waiting time tail probabilities in multiclass systems," *IEEE Commun. Lett.*, vol. 5, pp. 175–177, Apr. 2001.
- [14] J. Abate, G. L. Choudhury, and W. Whitt, "Exponential approximations for tail probabilities in queues I: Waiting times," *Oper. Res.*, vol. 43, no. 5, pp. 885–901, 1995.
- [15] L. Kleinrock, *Queueing Systems Volume I: Theory*. New York: Wiley, 1975.
- [16] Draft Supplement to Part 11: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Medium Access Control (MAC) Enhancements for Quality of Service (QoS) 2003, IEEE 802.11e/D5.0.
- [17] S. M. Faccin, "Mesh WLAN networks: Concept and system design," *IEEE Wireless Commun. Mag.*, pp. 10–17, Apr. 2006.
- [18] D. Krishnaswamy, "Network-assisted link adaptation with power control and channel reassignment in wireless networks," in *Proc. 3G Wireless Conf.*, 2002, pp. 165–170.
- [19] M. Waldvogel and R. Rinaldi, "Efficient topology-aware overlay network," *ACM SIGCOMM Comput. Comm. Rev.*, vol. 33, no. 1, pp. 101–106, Jan. 2003.
- [20] J. Jannotti, "Network-layer support for overlay networks," in *Proc. IEEE Conf. Open Architectures and Network Programming*, Jun. 2002.
- [21] M. van der Schaar, Y. Andreopoulos, and Z. Hu, "Optimized scalable video streaming over IEEE 802.11a/e HCCA wireless networks under delay constraints," *IEEE Trans. Mobile Comput.*, vol. 5, no. 6, pp. 755–768, Jun. 2006.
- [22] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 23–50, Nov. 1998.
- [23] Y. Andreopoulos, A. Munteanu, J. Barbarien, M. van der Schaar, J. Cornelis, and P. Schelkens, "In-band motion compensated temporal filtering," *Signal Processing: Image Communication (Special Issue on "Subband/Wavelet Interframe Video Coding")*, vol. 19, no. 7, pp. 653–673, Aug. 2004.
- [24] H.-P. Shiang and M. van der Schaar, "Multi-user video streaming over multi-hop wireless networks: A cross-layer priority queuing approach," in *Proc. IEEE Conf. Intelligent Information Hiding and Multimedia Signal Processing (IHH-MSP 2006)*, Dec. 2006, pp. 255–258.
- [25] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*. Norwell, MA: Kluwer, 1996.
- [26] C. E. Perkins and E. M. Royer, "Ad hoc on-demand distance vector routing," in *Proc. 2nd IEEE Workshop on Mobile Computing Systems and Applications*, Feb. 1999, pp. 90–100.
- [27] P. A. Chou and M. Zhouong, "Rate-distortion optimized streaming of packetized media," *IEEE Trans. Multimedia*, vol. 8, pp. 390–404, Apr. 2006.
- [28] D. Jurca and P. Frossard, "Packet selection and scheduling for multi-path video streaming," *IEEE Trans. Multimedia*, vol. 9, Apr. 2007.
- [29] M. K. Marina and S. R. Das, "Ad hoc on-demand multi-path distance vector routing (AOMDV)," in *Proc. Int. Conf. Network Protocols (ICNP)*, 2001, pp. 14–23.
- [30] A. G. Konheim, "A queuing analysis of two ARQ protocols," *IEEE Trans. Commun.*, vol. COM-28, Jul. 1980.
- [31] B. Krishnamachari, D. Estrin, and S. Wicker, "The impact of data aggregation in wireless sensor networks," in *Proc. IEEE Int. Conf. Distributed Computing Systems Workshop*, 2002, pp. 575–578.

Hsien-Po Shiang received the B.S. and M.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., in 2000 and 2002, respectively. He is currently pursuing the Ph.D. degree in the Electrical Engineering Department, University of California at Los Angeles.

During his graduate study, he was with Intel Corp., Folsom, CA, in 2006, and researched the overlay network infrastructure over wireless mesh networks. His research interests are the cross-layer optimizations/adaptations for multimedia transmission over wireless mesh networks, and the dynamic routing based on collaborative information exchange for delay-sensitive applications.

Mihaela van der Schaar (SM'04) is currently an Assistant Professor in the Electrical Engineering Department at the University of California, Los Angeles (UCLA). Since 1999, she was an active participant to the ISO MPEG standard to which she made more than 50 contributions and for which she received three ISO recognition awards.

Dr. van der Schaar received the NSF CAREER Award in 2004, IBM Faculty Award in 2005, Okawa Foundation Award in 2006, Best IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY Paper Award in 2005, and Most Cited Paper Award from EURASIP Journal Signal Processing: Image Communication between the years 2004–2006. She holds 28 granted U.S. patents. She is currently an associate editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SIGNAL PROCESSING LETTERS and *IEEE Signal Processing e-Newsletter*. She is also the editor (with Phil Chou) of the book *Multimedia over IP and Wireless Networks: Compression, Networking, and Systems*.