Cross-Layer Optimized Video Streaming Over Wireless Multihop Mesh Networks

Yiannis Andreopoulos, *Member, IEEE*, Nicholas Mastronarde, and Mihaela van der Schaar, *Senior Member, IEEE*

Abstract—The proliferation of wireless multihop communication infrastructures in office or residential environments depends on their ability to support a variety of emerging applications requiring real-time video transmission between stations located across the network. We propose an integrated cross-layer optimization algorithm aimed at maximizing the decoded video quality of delay-constrained streaming in a multihop wireless mesh network that supports quality-of-service. The key principle of our algorithm lays in the synergistic optimization of different control parameters at each node of the multihop network, *across the protocol layers*—application, network, medium access control, and physical layers, *as well as end-to-end, across the various nodes*. To drive this optimization, we assume an overlay network infrastructure, which is able to convey information on the conditions of each link. Various scenarios that perform the integrated optimization using different levels ("horizons") of information about the network status are examined. The differences between several optimization scenarios in terms of decoded video quality and required streaming complexity are quantified. Our results demonstrate the merits and the need for cross-layer optimization in order to provide an efficient solution for real-time video transmission using existing protocols and infrastructures. In addition, they provide important insights for future protocol and system design targeted at enhanced video streaming support across wireless mesh networks.

Index Terms—Cross-layer strategies, distributed video streaming optimization, quality-of-service (QoS), wireless mesh networks.

I. INTRODUCTION

WIRELESS mesh networks are built based on a mixture of fixed and mobile nodes interconnected via wireless links to form a multihop ad hoc network. The use of existing protocols for the interconnection of the various nodes (hops) is typically desired as it reduces deployment costs and also increases interoperability [1]. However, due to the network and channel dynamics, there are significant challenges in the design and joint optimization of application, routing, medium access control (MAC), and physical (PHY) adaptation strategies for efficient video transmission across such mesh networks.

In this paper, we are addressing some of these challenges by developing an integrated video streaming paradigm enabling cross-layer interaction across the protocol stack and across the multiple hops. The problem of multihop video streaming has recently been studied under a variety of scenarios [2]–[4]. However, the majority of this research does not consider the protection techniques available at the lower layers of the protocol stack and/or optimizes the video transport using purely end-to-end metrics, thereby excluding a significant amount of improvement that can occur by cross-layer design [5]–[7]. Consequently, the inherent network dynamics occurring in a multihop wireless mesh network as well as the interaction among the various layers of the protocol stack are not fully considered in the existing video streaming literature. Indeed, recent results concerning the practical throughput and packet loss analysis of multihop wireless networks [8], [9] have shown that the incorporation of appropriate utility functions that take into account specific parameters of the protocol layers such as the expected retransmissions, the loss rate, and bandwidth of each link [8], as well as expected transmission time [9] or fairness issues [10], can significantly impact the actual end-to-end network throughput. Motivated by this work, we show that, for delay-constrained video streaming over multihop wireless mesh networks, including the lower layer network information and adaptation parameters in the cross-layer design can provide significant improvements in the decoded video quality.

In this paper, we focus on the problem of real-time transmission of an individual video bitstream across a multihop 802.11a/e wireless network and investigate: 1) what is the video quality improvement that can be obtained if an integrated cross-layer strategy involving the various layers of the protocol stack is performed and 2) what is the performance and complexity impact if the optimized streaming solution is performed using only limited, localized information about the network status, as opposed to global, complete information.

We assume that the mesh network topology is fixed over the duration of the video session and that, prior to the transmission, each application (video flow) reserves a predetermined transmission opportunity interval, where contention-free access to the medium is provided.[1] This reservation can be performed following the principles of the HCCA[2] protocol of IEEE 802.11e [12] and can be determined based on the amount of flows sharing the network. Although the design of such a reservation system is an important problem and it affects our results, recent work showed that scheduling of multiple flows in the context of a mesh topology can be done such that the average rate for every flow is satisfied and the interference to neighboring

Y. Andreopoulos and M. van der Schaar are with the Department of Electrical Engineering, University of California at Los Angeles (UCLA), Los Angeles, CA 90095-1594 USA (e-mail: yandreop@ee.ucla.edu; mihaela@ee.ucla.edu).

N. Mastronarde is with the Department of Electrical and Computer Engineering, University of California at Davis (UCDavis), Davis, CA 95616 USA (e-mail: nhmastro@ucdavis.edu).

[1]Existing IEEE standards [12] already support such quality-of-service (QoS) mechanisms, which, barring interference and environment noise, provide guaranteed transmission time for each admitted application (video flow).

[2]HCCA: HCF controlled channel access, where HCF stands for hybrid coordinator function [12].
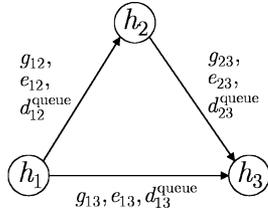
Fig. 1. A simple topology with three hops.

nodes is minimized [13]. Hence, a similar solution can be applied for our case and the available nodes and links within the entire mesh topology can be preestablished by a central coordinator prior to the video streaming session initiation. This minimizes the probability of additional delays and link failures due to routing reconfigurations during the video streaming and also decouples the problem of optimized media streaming and optimized route and link-reservation establishment within the wireless multihop network. Once the available network infra-structure to the video streaming session has been established, we assume that an overlay network topology can convey (in frequent intervals) information about the expected bit-error rate (BER), the queueing delay for each link, as well as the guaran-teed bandwidth under the dynamically changing modulation at the PHY. Several examples of such application-layer overlay networks have been proposed in the literature [19], [20].

Under the above assumptions, this paper makes the following contributions. For video packets of each hop in the mesh net-work, we propose an optimization framework that jointly deter-mines per packet: 1) the optimal modulation at the PHY; 2) the optimal retry limit at the MAC; 3) the optimal path (route) to the receiver in the remaining part of the mesh network; and 4) the application-layer optimized packet scheduling, given a predeter-mined topology and time reservation per link using the concepts of IEEE 802.11e HCCA.

This paper is organized as follows. Section II defines the sce-narios examined in this work and provides the necessary defi-nitions and formulations for the expected bandwidth, transmis-sion error rate as, well as the expected delay for streaming under various network paths. Section III presents the cross-layer op-timization problem. The proposed solutions are presented in Section IV. Section V analyzes the complexity and feedback requirements of the proposed approaches. Section VI presents indicative results, including comparisons with other well-known approaches from the literature. Our conclusions are presented in Section VII.

II. PROPOSED INTEGRATED CROSS-LAYER VIDEO STREAMING

Consider that N nodes (hops) of a wireless multihop mesh network decide to participate in a video streaming session. Ex-ample topologies with $N = 3$ and $N = 7$ are shown in Figs. 1 and 2. Node h_1 represents the original video source, while node h_N is the destination node (video client). Each link ij is associ-ated with the corresponding allocated bandwidth for the video traffic (g_{ij}), the error rate observed on the link (e_{ij}), as well as the corresponding delay due to the video queue (d_{ij}^{queue}). Within the reserved time for the video traffic, each link exhibits a certain throughput given the chosen modulation strategy. Video

packets are lost due to the experienced BER. This error is due to noise and interference in the wireless medium stemming from background noise, node mobility, or simultaneous link transmis-sions. In addition to this error, under delay-constrained video streaming, packets are discarded due to delays incurred in the transmission, e.g., the queueing delay of each link. Notice that Fig. 2 displays different connectivity structures for the network topology, as specified by the indicated links. Obviously, the tightly connected multihop mesh topologies T1 and T2 of Fig. 2 offer more alternative paths for the video traffic that topology T3; however, the overall reserved time across the various nodes of the network is also increased. In general, the decision on the connectivity as well as the number of nodes participating in the video streaming session depends largely on a number of system-related factors that transcend the video streaming problem (e.g., node cooperation strategy/incentives and network coordination and routing policies imposed by the utilized protocols). Hence, in this paper, we investigate cross-layer optimization for video over multihop wireless mesh networks given the network spec-ification (participating nodes and connectivity), as well as the available reservation time on each link for the video traffic.

Under the existence of feedback from an overlay network in-frastructure, the BER and queueing delay per link can be dis-seminated to the remaining network hops at frequent intervals (via a hop-to-hop feedback mechanism[3]), or when the incurred change in network parameters is larger than a preset threshold. Thus, they can be considered to be known (Fig. 1). However, in certain cases, feedback from remote hops may arrive with an intolerable delay, or, alternatively, it can be deemed unreliable due to the rapidly changing network conditions. As a result, a certain "horizon" of information retrieval can be envisaged for each hop (Fig. 2), where network information within the horizon is deemed reliable and can be received in a timely manner, while information beyond the horizon can only be theoretically esti-mated based on average or previous measurements.

A. Wireless Multihop Mesh Topology Specification

For a generic multihop wireless mesh network, we consider the *connectivity structure* $\mathcal{P}$

$$\mathcal{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_M\} \tag{1}$$

where each element $\mathbf{p}_i, 1 \leq i \leq M$ is the *connectivity vector* (end-to-end network path) given by

$$\mathbf{p}_i = \begin{bmatrix} l_{i,1} & l_{i,2} & \cdots & l_{i,\rho_i^{\text{total}}-2} & l_{i,\rho_i^{\text{total}}-1} \end{bmatrix} \tag{2}$$

where each component $l_{i,j}(1 \leq j < \rho_i^{\text{total}})$ indicates a partic-ular wireless link (the jth link of path i), and $\rho_i^{\text{total}} - 1$ is the total number of links participating in the network path $\mathbf{p}_i$. For example, for the topology of Fig. 1 with $N = 3$, we have

$$\mathcal{P} = \{[l_{1,1} \quad l_{1,2}], l_{2,1}\} \tag{3}$$

[3]For example, in order to utilize the medium more efficiently, it is possible to piggyback feedback about the link status information onto the acknowledgment packets.
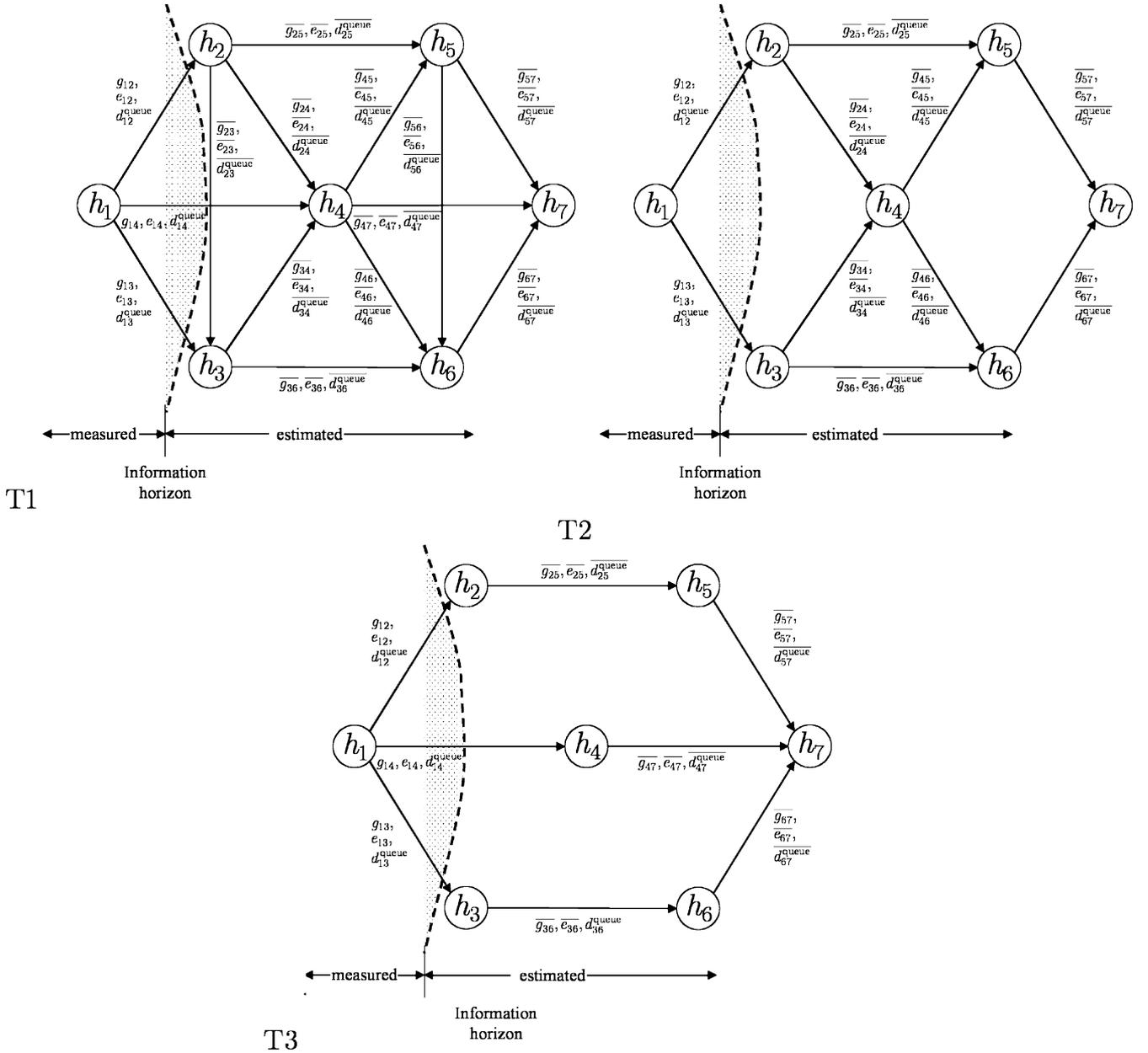
Fig. 2. More complex topologies with seven hops. In these cases, the link information up to hop h_4 is directly conveyed by the overlay network infrastructure, while other link information is inferred based on theoretical estimates using average or past information (where $\overline{m_{ij}}$ indicates the estimated or average value for the metric $\overline{m_{ij}}, m = \{g, e, d\}$).

with $\rho_1^{\text{total}} = 3, \rho_2^{\text{total}} = 2$ and

$$l_{1,1} = (h_1 \rightarrow h_2); l_{1,2} = (h_2 \rightarrow h_3)$$
$$l_{2,1} = (h_1 \rightarrow h_3). \tag{4}$$

Notice that (1) and (2) apply both for the end-to-end topology of interest but also for the topology between any intermediate node and the terminal (client) node in the mesh network utilized for video transmission. For example, if we consider the subnetwork of topology T2 of Fig. 2 consisting of nodes h_4, h_5, h_6, and h_7, there are two paths from h_4 to h_7, and the equivalent definitions apply locally. Hence, the subsequent problem specification and analysis is inherently scalable and can be applied in a similar

fashion to either the entire end-to-end topology or only part of the topology (subnetwork). Finally, it is important to mention that all the proposed algorithms in this paper assume the nonexistence of routing loops, i.e., the mesh network between the current hop and the destination hop can be represented by a tree graph.

B. Link and Path Parameter Specification

For each link $l_{i,j}$, given a certain modulation $m(l_{i,j})$ at the physical layer, we denote the expected BER as $e(l_{i,j})$. Notice that this error is usually estimated based on channel modeling, as well as experimental studies in the network which analyze the effects of interference [11]. As a result, the higher layers of the protocol stack can assume $e(l_{i,j})$ to be independent and

randomly distributed [21]. Under a predetermined negotiation of traffic specification parameters for each link in the mesh network (e.g., following the HCCA protocol [12]), each link can provide a guaranteed bandwidth $g(l_{i,j})$ at the application layer. Following the HCCA specification [12], this bandwidth is linked with the traffic specification parameters by [14]:

$$g(l_{i,j}) = \frac{t_{\text{TXOP}}(l_{i,j}) \cdot \bar{L}}{[\bar{L} \cdot (R_{\text{phy}}(l_{i,j}))^{-1} + T_{\text{overhead}}] \cdot t_{\text{SI}}(l_{i,j})} \quad (5)$$

where $t_{\text{TXOP}}(l_{i,j})$ is the transmission opportunity duration provided by the HCCA admission control for the video flow traffic of link $l_{i,j}$, $\bar{L}$ is the nominal MAC service data unit (MSDU) size,[4] $t_{\text{SI}}(l_{i,j})$ is the specified duration of the service interval [12] for the video flow traffic at link $l_{i,j}$, $R_{\text{phy}}(l_{i,j})$ is the physical-layer rate, and T_{overhead} represents the duration of the required overheads corresponding to polling and acknowledgment policies. As demonstrated by (5), even though the negotiated transmission opportunity duration is constant per link, the guaranteed bandwidth depends on the provided physical-layer rate $R_{\text{phy}}(l_{i,j})$, which in turn makes it dependent on the chosen modulation[5] $m(l_{i,j})$. Finally, depending on the chosen modulation, $R_{\text{phy}}(l_{i,j})$ may change for each MSDU. Hence, the guaranteed bandwidth $g(l_{i,j})$ of (5) can be determined for each MSDU.

Under the aforementioned assumptions for the error model of each link, the probability of error for the transmission of MSDU v of size L_v bits is

$$e_{l_{i,j}}(L_v) = 1 - (1 - e(l_{i,j}))^{L_v}. \quad (6)$$

Consequently, the probability of error for the packet transmission in path $\mathbf{p}_i$ is

$$e_{\mathbf{p}_i}(L_v) = 1 - \prod_{j=1}^{\rho_i^{\text{total}}-1} (1 - e_{l_{i,j}}(L_v))$$

$$= 1 - \prod_{j=1}^{\rho_i^{\text{total}}-1} [(1 - e(l_{i,j}))^{L_v}]. \quad (7)$$

Under a single (successful) MSDU transmission via each link $l_{i,j}$, the transmission delay for path $\mathbf{p}_i$ can be calculated as

$$d_{\mathbf{p}_i}(L_v, 1) = \sum_{j=1}^{\rho_i^{\text{total}}-1} (\frac{L_v}{g(l_{i,j})} + d_{\text{queue}}(l_{i,j}) + T_{\text{overhead}}) \quad (8)$$

where $d_{\text{queue}}(l_{i,j})$ depends on the transmitting-link queue length and will be discussed in the next subsection.

Considering an end-to-end scenario, if we denote the retransmission limit for each MSDU v (transmitted via path $\mathbf{p}_i$) as $\text{T}_{\mathbf{p}_i}^{\max}$, the average number of transmissions over path $\mathbf{p}_i$ until

---

[4]In this paper, we assume that one video packet is encapsulated in one MSDU and the two terms are used interchangeably.

[5]For notational simplicity, we do not particularly indicate the dependence of $e(l_{i,j})$ and $g(l_{i,j})$ on the modulation $m(l_{i,j})$.

the packet is successfully transmitted, or the retransmission limit is reached, can be calculated as

$$t_{\mathbf{p}_i}^{\text{mean}}(\text{T}_{\mathbf{p}_i}^{\max}) = \sum_{j=1}^{\text{T}_{\mathbf{p}_i}^{\max}+1} \left(j \cdot [e_{\mathbf{p}_i}(L_v)]^{j-1} [1 - e_{\mathbf{p}_i}(L_v)] \right)$$

$$+ [\text{T}_{\mathbf{p}_i}^{\max} + 1] \cdot [e_{\mathbf{p}_i}(L_v)]^{\text{T}_{\mathbf{p}_i}^{\max}+1}$$

$$= \frac{1 - [e_{\mathbf{p}_i}(L_v)]^{\text{T}_{\mathbf{p}_i}^{\max}+1}}{1 - e_{\mathbf{p}_i}(L_v)}. \quad (9)$$

Consequently, the (end-to-end) expected delay for the transmission of an MSDU of size L_v through $\mathbf{p}_i$ under $t_{\mathbf{p}_i}^{\text{mean}}(\text{T}_{\mathbf{p}_i}^{\max})$ transmissions can be approximated by

$$d_{\mathbf{p}_i}\left(L_v, t_{\mathbf{p}_i}^{\text{mean}}(\text{T}_{\mathbf{p}_i}^{\max})\right) = t_{\mathbf{p}_i}^{\text{mean}}(\text{T}_{\mathbf{p}_i}^{\max})$$

$$\cdot \sum_{j=1}^{\rho_i^{\text{total}}-1} \left(\frac{L_v}{g(l_{i,j})} + T_{\text{overhead}} \right) + \sum_{j=1}^{\rho_i^{\text{total}}-1} d_{\text{queue}}(l_{i,j}). \quad (10)$$

The last equation derives the end-to-end delay estimate by joining all links of path $\mathbf{p}_i$ via the summation terms, thereby forming a "virtual" link from the sender to the receiver node in the multihop network. We follow this approach since the maximum number of retransmissions $\text{T}_{\mathbf{p}_i}^{\max}$ required on each path $\mathbf{p}_i$ can only be defined end-to-end, based on the maximum permissible delay from the sender node to the receiver. We remark that in our experiments, the retransmission limit for any part of a path or even for one link $l_{i,j}$ is set equal to $\text{T}_{\mathbf{p}_i}^{\max}$, since, in principle, all possible retransmissions (until the MSDU expires due to delay violation) could occur at an individual link. Following the analysis of (7)–(10), it is straightforward to define the average MSDU transmissions and the expected delay for subpaths that include only a subset of links, or even for an individual link. This will be proven to be very useful for some of the derivations of this paper.

C. Application and Network-Layer Parameter Specification

Since we are considering real-time video streaming through the multihop wireless network, each MSDU v is associated with a corresponding delay limit d_v^{deadline}, before which the video data encapsulated in the MSDU should arrive at the destination node h_N. In addition, decoding the video data at the video receiver incurs a reduction in the perceived distortion, which is represented by Δ_v. Several models exist for the definition of Δ_v (e.g., see [15]). Recent results [16], [17] demonstrated that acknowledgment-based transmission of scalable video under a strict distortion-reduction prioritization of the video packets leads to an additive distortion-reduction model at the receiver side under packet losses in the multihop network. This additive model is codec-specific and typically expresses the expected mean square error (MSE) reduction at the video decoder instead of the visual distortion reduction, since the latter is harder to quantify. See [15] for an example and [18] for further details on linking the distortion-reduction estimates (Δ_v) with the packetization process at the application layer. In this paper, we assume that an optimized packet scheduling is performed at the application layer, where all packets v with the same

delay deadline d_v^{deadline} are ordered at the encoder (sender) side according to their expected distortion reduction Δ_v [15], [3]. The delay deadline d_v^{deadline} will also be considered as a parameter in the proposed optimization strategy and it will be defined based on the application requirements.

At each link in the mesh wireless network, each video flow is subjected to a queueing delay,[6] which depends on: 1) the MSDUs from a particular flow (user) that are scheduled for transmission via the link of interest at the moment when MSDU v arrives and 2) the queue output rate. The queue output rate depends on the quality of the link [error probability given in (6)] and the average number of retransmissions for each MSDU in the particular link [given by (9) with the replacement of $\mathbf{p}_i$ by $l_{i,j}$]. If we assume that link $l_{i,j}$ is shared among multiple paths, then at the arrival of MSDU v at the queue of link $l_{i,j}$, another u MSDUs (where, typically, $0 \leq u \leq v-1$) will be in the same queue. For each $l_{i,j}$, by indicating the group of u MSDUs by vector $\mathbf{v}_{\text{queue}}(l_{i,j})$, the queueing delay can be estimated as

$$d_{\text{queue}}(l_{i,j}) = \sum_{\forall u \in \mathbf{v}_{\text{queue}}(l_{i,j})} d_{l_{i,j}}\left(L_u, t_{l_{i,j}}^{\text{mean}}\left(\mathbf{T}_{\mathbf{p}_i}^{\max}\right)\right). \tag{11}$$

For the optimization of the routing strategy of each MSDU v (presented in the next section), the determination of (11) can be performed dynamically during the path estimation, under the knowledge of the previous decisions for the MSDUs that were transmitted by the current node. Alternatively, each node can independently calculate (11) based on the queue contents of the particular link and disseminate the result at frequent intervals in the mesh network via the overlay infrastructure.

III. PROBLEM FORMULATION

Assume a set of N wireless hops (nodes), with h_1 being the video encoder (server) and h_N the video decoder (client), and a connectivity structure $\mathcal{P}$ with M paths, where each path $i, 1 \leq i \leq M$ consists of ρ_i^{total} hops. In addition, assume a predefined HCCA transmission opportunity duration $t_{\text{TXOP}}(l_{i,j})$ for each link $l_{i,j}$, with $1 \leq j \leq \rho_i^{\text{total}}$, and a link adaptation mechanism at the physical layer that can operate at an MSDU granularity. The end-to-end cross-layer optimization which determines the chosen path (routing), the maximum MAC retry limit, and the chosen modulation (at the PHY layer) for the transmission of each MSDU is

$$\forall v : \left\{ \mathbf{p}_i^*, \mathbf{T}_{\mathbf{p}_i}^{\max *}, m(l_{i,j}) \right\}$$
$$= \arg \max_{\forall \mathbf{p}_i \in \mathcal{P}, \forall l_{i,j} \in \mathbf{p}_i} \left[\min_{1 \leq j < \rho_i^{\text{total}}} \{c(l_{i,j})\} \cdot \Delta_{v,\text{expected}} \right] \tag{12}$$

where

$$\Delta_{v,\text{expected}} = \left[1 - [e_{\mathbf{p}_i}(L_v)]^{\mathbf{T}_{\mathbf{p}_i}^{\max}} \right] \cdot \Delta_v \tag{13}$$

[6] In this paper, we assume that the MSDUs of each flow are accommodated with an independent queue at each link.

with $e_{\mathbf{p}_i}(L_v)$ given by (7), $\mathbf{T}_{\mathbf{p}_i}^{\max}$ the maximum number of retransmissions for MSDU v if scheduled via path $\mathbf{p}_i$, and $c(l_{i,j})$ corresponding to the remaining time interval for which link $l_{i,j}$ can support the video-flow traffic under HCCA. For the transmission opportunity intervals belonging to the current service interval, $t_{\text{SI}}, c(l_{i,j})$ can be calculated as

$$c(l_{i,j}) = \max \left\{ \sum_{t_{\text{SI}}} t_{\text{TXOP}}(l_{i,j}) - d_{\text{queue}}(l_{i,j}), 0 \right\}. \tag{14}$$

Under the constraints set by the video codec and the mesh wireless network infrastructure, the optimization of (12) attempts to find the cross-layer parameters that maximize a capacity-distortion utility function. This function is formulated as the product of the minimum path capacity (expressed by the remaining reserved time within the current service interval at the most congested link) and the expected source distortion-reduction of (13). In this way, we minimize congestion across the various links [since the path whose worst link is having the highest capacity is selected under $\Delta_{v,\text{expected}}$ given by (13)], and concurrently maximize the expected distortion reduction (under the current path's minimum link capacity $\min_{1 \leq j < \rho^{\text{total}}} \{c(l_{i,j})\}$). The granularity of this optimization is one MSDU. However, coarser granularities could also be considered, in order to reduce complexity. The problem constraints can be expressed for each MSDU v as

$$\forall v, \text{and } \forall \mathbf{p}_i \in \mathcal{P} : d_{\mathbf{p}_i}\left(L_v, \mathbf{T}_{\mathbf{p}_i}^{\max}\right) \leq d_v^{\text{deadline}} \tag{15}$$

i.e., the maximum transmission delay through each possible path must be below or equal to the MSDU deadline (d_v^{deadline}) in order for the video data to be useful to the decoder. Moreover, the timing constraint set from HCCA scheduling is

$$\forall v, \text{and } \forall \mathbf{p}_i \in \mathcal{P} : \frac{\left[\mathbf{T}_{\mathbf{p}_i}^{\max} + 1\right] \cdot L_v}{g(l_{i,j})} \leq \min_{\forall l_{i,j} \in \mathbf{p}_i} \{c(l_{i,j})\}. \tag{16}$$

The two constraints of (15) and (16) can provide two bounds for the maximum number of retries for each MSDU v for each link $l_{i,j}$. Since both the MAC-layer scheduling and the application-layer deadline constraints are concurrently imposed, if $\min_{\forall l_{i,j} \in \mathbf{p}_i} \{c(l_{i,j})\}(g(l_{i,j})/L_v) - 1 > 0$, we set the tightest bound for the maximum retry limit

$$\mathbf{T}_{\mathbf{p}_i}^{\max} = \min \left\{ \frac{d_v^{\text{deadline}} - \sum_{j=1}^{\rho_i^{\text{total}}-1} d_{\text{queue}}(l_{i,j})}{\sum_{j=1}^{\rho_i^{\text{total}}-1} \left(\frac{L_v}{g(l_{i,j})} + T_{\text{overhead}} \right)}, \right.$$
$$\left. \frac{g(l_{i,j}) \cdot \min_{\forall l_{i,j} \in \mathbf{p}_i} \{c(l_{i,j})\}}{L_v} - 1 \right\}. \tag{17}$$

Obviously, if there is no path $\mathbf{p}_i$ for which $d_v^{\text{deadline}} > \sum_{j=1}^{\rho_i^{\text{total}}-1} d_{\text{queue}}(l_{i,j})$, then the current MSDU may be dropped.

```
1.    For each node that has non-expired MSDUs in its queue
2.        Extract the network connectivity structure 𝒫 (eq. (1), (2))
3.        For the MSDU v existing at output of the queue of the sender node
4.            For each path p_i (network topology emanating from the sender node)
5.                For each link l_{i,j} of path p_i
6.                    For each modulation strategy m(l_{i,j})
7.                        Calculate e_{l_{i,j}}(L_v), e_{p_i}(L_v), d_queue(l_{i,j}) from eq. (6),(7),(11)
8.                        Calculate T^max_{p_i} from (15)-(17)
9.                        Under the calculated T^max_{p_i}, evaluate eq. (12)
10.                       Compare with previous best choice, retaining the maximum (eq. (12))
11.          Schedule the MSDU according to the established {p_i^*, T^{max*}_{p_i}, m(l_{i,j})}
```

Fig. 3. Exhaustive algorithm for the determination of the cross-layer optimized mesh-network path selection, MAC retry limit, and physical-layer modulation. The algorithm is applied for each MSDU existing in the queue of each node in the multihop wireless network.

IV. VIDEO STREAMING OPTIMIZATION IN THE MULTIHOP MESH NETWORK

In this section, we derive an algorithm that determines the optimal parameters for (12) under a predetermined deadline for each MSDU v (given by d_v^{deadline}) and a predetermined transmission opportunity duration per link, given by $t_{\text{TXOP}}(l_{i,j})$, which is set by the HCCA admission control once the video flow is scheduled for transmission. Moreover, although the conditions of the various links vary over time, we assume the network topology to be fixed for the duration of the video transmission.

A. End-to-End Optimization

The optimization of (12) can be performed for each node of the mesh wireless network under the assumption that, for every link $l_{i,j}$, the parameters $g(l_{i,j}), e(l_{i,j})$ are determined based on the chosen modulation $m(l_{i,j})$, and the experienced signal-to-interference-plus-noise ratio (SINR) [21]. In addition, we assume that $d_{\text{queue}}(l_{i,j})$ is communicated to the sender node via frequent feedback using an overlay network infrastructure [19], [20] that uses real-time protocols for conveying information from different layers.

The proposed optimization algorithm is given in Fig. 3. Notice that, although an entire path is selected at the sender node, the algorithm is executed for each node in the network independently by assuming each node is the sender and considering only the network (and MSDU) subset corresponding to the node of interest. This ensures that the algorithm can scale well under a variety of topologies. In addition, in this way, potential network variations that invalidate the error, bandwidth, or queueing-delay assumptions used when scheduling at the sender node, can be incorporated/corrected during the scheduling of a subsequent node. Finally, the independent algorithm execution at each node ensures that expired MSDUs will not propagate through the entire network unnecessarily. This facilitates the conservation of network resources in the mesh topology and reduces link congestion.

The algorithm of Fig. 3 searches through all the possible routing configurations (line 4) that emerge under varying modulation strategies (line 6) and determines the retransmission limit for each case (line 8). The utility function is evaluated (line 9)

and the overall maximum is retained. Although this is a greedy approach, it is guaranteed to obtain the maximum under dynamic feedback from the overlay network (parameters calculated in line 7).

B. Optimization Under a Certain Horizon of Network Information

In this case, we are only considering the part of the mesh network topology that immediately connects to the node of interest. This may be advantageous in comparison to the previous case, since a limited set of network parameters needs to be communicated to the sender node.

For analytical purposes, this can be considered as the previous case with $2 \leq \rho_i^{\text{sub}} < \rho_i^{\text{total}}$, where ρ_i^{total} is the total length of the path that was used in the end-to-end optimization of the previous section. In this case, every path i originating from the current node consists of one or more links, but we do not consider the entire path to the destination. The advantage offered by this scenario is that the required information for the MSDU scheduling is localized (limited).

For each path $\mathbf{p}_i$, we assume that the information for the optimization process is known only for links $l_{i,1}, \ldots, l_{i,\rho_i^{\text{sub}}-1}$. For the remaining links of each path $(l_{i,\rho_i^{\text{sub}}}, \ldots, l_{i,\rho_i^{\text{total}}-1})$, we assume that the allocated transmission opportunity duration available for the MSDUs of each link is known, as well as the limits for the SINR experienced by each link. Our goal is to establish d_v^{deadline} for the video transmission up to links $l_{i,\rho_i^{\text{sub}}-1}$, i.e., the known network "horizon," in order to perform the optimization of Fig. 3 locally. With respect to IEEE 802.11a networks, it can be shown [21], [22] that the physical-layer throughput of each link $l_{i,j}$ can be approximated by

$$R_{\text{phy}}(l_{i,j}) = \frac{R_{\text{phy}}^{\text{max}}(l_{i,j})}{1 + e^{-\mu \cdot (s(l_{i,j}) - \delta)}} \qquad (18)$$

where $R_{\text{phy}}^{\text{max}}(l_{i,j})$ is the maximum achievable data rate for each modulation $m(l_{i,j})$, $s(l_{i,j})$ is the observed SINR, and μ, δ are constants whose values for each modulation $m(l_{i,j})$ can be extracted based on the observation for s and predetermined experimental points [21]. Assuming that, for every link $l_{i,j}$, the SINR

s is a random variable following a certain probability distribution $\mathcal{F}(s)$, we have

$$\mathbf{E}\{R_{\mathrm{phy}}(l_{i,j})\} = R_{\mathrm{phy}}^{\max}(l_{i,j}) \int_{s_{\min}}^{s_{\max}} \frac{1}{1 + e^{-\mu\cdot(s-\delta)}} \cdot \mathcal{F}(s)ds \tag{19}$$

where $s_{\max}, s_{\min}$ are the bounds of the observable values for the SINR for each link $l_{i,j}$. In addition, under a given or estimated probability distribution $\mathcal{F}(s)$, (5) can be used in order to derive the expectation for the guaranteed bandwidth of each link, which, after a few straightforward manipulations is shown in (20) at the bottom of the page, since the remaining parameters of (5) are constants (in our analysis, we consider a nominal MSDU size $\bar{L}$). In a similar manner, the expected error of the subpath $l_{i,\rho_i^{\mathrm{sub}}}, \ldots, l_{i,\rho_i^{\mathrm{total}}-1}$ within each path $\mathbf{p}_i$ is

$$\mathbf{E}\left\{e_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,\rho_i^{\mathrm{total}}-1}\right]}(\bar{L})\right\}$$
$$= \int_{s_{\min}}^{s_{\max}} \left(1 - \prod_{j=\rho_i^{\mathrm{sub}}}^{\rho_i^{\mathrm{total}}-1} \left[\left(1 - \frac{\kappa}{1+e^{\mu\cdot[s-\delta]}}\right)^{\bar{L}}\right]\right)$$
$$\cdot \mathcal{F}(s)\,ds. \tag{21}$$

The last equation was derived based on (7) under the assumption that the bit-error probability can be approximated by [21], [22]

$$e(l_{i,j}) = \frac{\kappa}{1 + e^{\mu\cdot[s(l_{i,j})-\delta]}} \tag{22}$$

where κ, μ, δ are derived experimentally depending on the observation for $s(l_{i,j})$ and the chosen modulation $m(l_{i,j})$ [21], with $\rho_i^{\mathrm{sub}} \le j < \rho_i^{\mathrm{total}}$.

Having the expected values for the full path's guaranteed bandwidth, the maximum transmission delay for an MSDU L_v transmitted through links $l_{i,\rho_i^{\mathrm{sub}}}, \ldots, l_{i,\rho_i^{\mathrm{total}}-1}$ can be derived based on (10) as

$$d_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,\rho_i^{\mathrm{total}}-1}\right]}\left(L_v, \mathrm{T}_{\mathbf{p}_i}^{\max}\right)$$
$$= \mathrm{T}_{\mathbf{p}_i}^{\max} \sum_{j=\rho_i^{\mathrm{sub}}}^{\rho_i^{\mathrm{total}}-1} \left(\frac{L_v}{\mathbf{E}\{g(l_{i,j})\}} + T_{\mathrm{overhead}}\right)$$
$$+ \sum_{j=\rho_i^{\mathrm{sub}}}^{\rho_i^{\mathrm{total}}-1} d_{\mathrm{queue}}(l_{i,j}). \tag{23}$$

Notice that (23) involves also the knowledge of the queueing delay $d_{\mathrm{queue}}(l_{i,j})$ of each link $l_{i,j}, \rho_i^{\mathrm{sub}} \le j < \rho_i^{\mathrm{total}}$, i.e., for the subsequent links after the "horizon." For each link $l_{i,j}$,

the expected $d_{\mathrm{queue}}(l_{i,j})$ can be updated within intervals of $d_{[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}]}(L_v, t_{[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}]}^{\mathrm{mean}}(\mathrm{T}_{\mathbf{p}_i}^{\max}))$ (that correspond to the expected time required for MSDU v to reach link $l_{i,j}$, after it passes the link $l_{i,\rho_i^{\mathrm{sub}}-1}$ which is at the "horizon") as

$$d_{\mathrm{queue}}(l_{i,j}) \leftarrow d_{\mathrm{queue}}(l_{i,j}) + t_{l_{i,j}}^{\mathrm{mean}}\left(\mathrm{T}_{l_{i,j}}^{\max}\right)$$
$$\cdot \left(\frac{L_v}{\mathbf{E}\{g(l_{i,j})\}} + T_{\mathrm{overhead}}\right)$$
$$\cdot \varphi_{\mathrm{enter}} - d_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}\right]}$$
$$\times \left(L_v, t_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}\right]}^{\mathrm{mean}}\left(\mathrm{T}_{\mathbf{p}_i}^{\max}\right)\right) \cdot \varphi_{\mathrm{exit}} \tag{24}$$

where operator $q_0 \leftarrow q_0 + u_0$ indicates the update of quantity q_0 by u_0 and

$$\varphi_{\mathrm{enter}}$$
$$= 1 - \mathbf{E}\left\{e_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}\right]}(L_v)\right\}^{t_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,j-1}\right]}^{\mathrm{mean}}\left(\mathrm{T}_{\mathbf{p}_i}^{\max}\right)} \tag{25}$$

$$\varphi_{\mathrm{exit}}$$
$$= 1 - \mathbf{E}\{e_{l_{i,j}}(L_v)\}^{t_{l_{i,j}}^{\mathrm{mean}}\left(\mathrm{T}_{\mathbf{p}_i}^{\max}\right)}. \tag{26}$$

The derivation of (24) is performed as follows. The queueing delay of the previous iteration is incremented by the product of the factor which indicates the expected delay due to retries for the new MSDU in link $l_{i,j}$ with the probability that the MSDU will reach link $l_{i,j}$ successfully (after maximally $\mathrm{T}_{\mathbf{p}_i}^{\max}$ retries are performed at links $l_{i,\rho_i^{\mathrm{sub}}}, \ldots, l_{i,j-1}$), which is given by φ_{enter}, defined in (25). At the same time, the queueing delay is decremented by the product of the factor indicating the time duration for the possible successful MSDU transmissions with the probability of a successful MSDU transmission, which is given by φ_{exit}, defined in (26).

Assuming that the value for $\sum_{j=\rho_i^{\mathrm{sub}}}^{\rho_i^{\mathrm{total}}-1} d_{\mathrm{queue}}(l_{i,j})$ is provided based on (24), (23) can be used in the constraint of (15) by updating the delay deadline

$$d_v^{\mathrm{deadline}} \leftarrow d_v^{\mathrm{deadline}} - d_{\left[l_{i,\rho_i^{\mathrm{sub}}} \cdots l_{i,\rho_i^{\mathrm{total}}-1}\right]}\left(L_v, \mathrm{T}_{\mathbf{p}_i}^{\max}\right) \tag{27}$$

and the optimization process follows, as explained in the beginning of this section.

The analytical formulation of this section is also useful in defining low-complexity scheduling algorithms at each node without the need for real-time network feedback. For example,

$$\mathbf{E}\{g(l_{i,j})\} = \frac{t_{\mathrm{TXOP}}(l_{i,j})}{t_{\mathrm{SI}}(l_{i,j})} \cdot \int_{s_{\min}}^{s_{\max}} \frac{\mathcal{F}(s)}{\left[R_{\mathrm{phy}}^{\max}(l_{i,j})\right]^{-1} \cdot \left(1 + e^{-\mu\cdot(s-\delta)}\right) + (\bar{L})^{-1} \cdot T_{\mathrm{overhead}}} ds \tag{20}$$

1. The optimal modulation parameters $\kappa^*(l_{i,j})$, $\mu^*(l_{i,j})$, $\delta^*(l_{i,j})$ are estimated only once per link during the optimization of the first MSDU and they remain constant throughout the remaining MSDUs until an update is received for the interval $[s_{\min}, s_{\max}]$ or for the values of $R_{\text{phy}}^{\max}(l_{i,j})$, $t_{\text{TXOP}}(l_{i,j})$, $t_{\text{SI}}(l_{i,j})$.

2. Per MSDU, the expected physical layer rate and guaranteed bandwidth per link are estimated from (19), (20), and the error for each path is estimated from (21).

3. Having the calculated queuing delay for each link from (23), as well as the available time interval from (27), and the estimated MSDU deadline from (27), the maximum retry limit is established from (17) and the average number of retries, $t_{\mathbf{p}_i}^{\text{mean}}(\text{T}_{\mathbf{p}_i}^{\max})$ from (9).

4. For each node, the link that maximizes $\min_{1 \leq j < \rho_i^{\text{total}}} \{c(l_{i,j})\} \cdot \left[\mathbf{1} - \mathbf{E}\{e_{\mathbf{p}_i}(L_v)\}^{\text{T}_{\mathbf{p}_i}^{\max}}\right]$ is selected.

Fig. 4. Algorithm for cross-layer optimization under an estimation-based framework.

if we assume that, due to the random interference caused by the simultaneous operation of the wireless nodes in the mesh network, the probability distribution $\mathcal{F}(s)$ for each link $l_{i,j}$ is uniform within an interval of $[s_{\min}, s_{\max}]$, we have

$$\mathcal{F}(s) = \frac{1}{s_{\max} - s_{\min}}. \tag{28}$$

With the explicit expression of $\mathcal{F}(s)$ from (28), we can derive the expected values for the physical-layer rate, the guaranteed bandwidth and the path error rate by solving (19) and (20), and approximating the solution of (21) with numerical methods. Based on these values, we derive a less complex solution for the scheduling of each group of MSDUs corresponding to a video GOP. The algorithm is given in Fig. 4. Based on this algorithm, for every new MSDU, all the cross-layer parameters are established analytically for each path (lines 1–3 of Fig. 4) and only the search through all the possible paths (i.e., line 4 of Fig. 4) is required in order to derive the optimal solution. Consequently, this optimization has minimal complexity. In the following section, we formulate the complexity requirements of the three different optimization solutions, while Section VI presents comparative experimental results.

V. COMPLEXITY AND INFORMATION REQUIREMENTS OF THE DIFFERENT ALTERNATIVES

Each proposed cross-layer optimization approach explores a different search space in order to determine the optimal parameters and also requires a varying amount of feedback on the conditions of the various links in the multihop mesh network. This results in varying computational and communication requirements for the presented algorithms.

Consider the case of a mesh network consisting of N nodes. Each node $h_n, 1 \leq n < N$, is the origin of M_n paths. Each path $\mathbf{p}_i$ stemming from node h_n consists of $\rho_{M_n}^{\text{total}}$ nodes, with $1 \leq \rho_{M_n}^{\text{total}} \leq N$. For each link $l_{i,j}$ of these paths, with $1 \leq i \leq M_n$ and $1 \leq j \leq \rho_{M_n,i}^{\text{total}} - 1$, there are S_{mod} possible modulations at the physical layer, which result in a different error rate and different guaranteed bandwidth at the MAC layer. For the end-to-end cross-layer optimization with network feedback from each node (Section IV-A), the overall complexity for the scheduling of an MSDU at node h_n is

$$C_{\text{exhaustive}}^{\text{full}}(h_n) = \sum_{i=1}^{M_n} \sum_{j=1}^{\rho_{M_n,i}^{\text{total}}-1} \sum_{m=1}^{S_{\text{mod}}} C_{\text{exhaustive}}(n) \tag{29}$$

where $C_{\text{exhaustive}}(n)$ represents the complexity for the dissemination of the necessary network information from node h_n, as well as the execution of the algorithm of Fig. 3.

Similarly, considering a scenario with partial network information, i.e., when the overlay network provides feedback only until node ρ_i (with $1 \leq \rho_i < \rho_{M_n,i}^{\text{total}}$), we have

$$C_{\text{exhaustive}}^{\text{partial}}(h_n)$$
$$= \sum_{i=1}^{M_n'} \sum_{j=1}^{\rho_i} \sum_{m=1}^{S_{\text{mod}}} (C_{\text{exhaustive}}(n) + C_{\text{estimation}}(n)) \tag{30}$$

where $C_{\text{estimation}}(n)$ represents the complexity for the estimation of the various parameters based on the analysis of the previous section and M_n' is the number of different paths (within the partial network topology under examination) originating from node h_n, with $1 \leq M_n' \leq M_n$. Finally, for the optimization of Fig. 4 where the best modulation strategy is *a priori* determined

$$C_{\text{estimation}}(h_n) = \sum_{i=1}^{M_n''} C_{\text{estimation}}(n) \tag{31}$$

where M_n'' is the number of links that are directly connected to node h_n. As an indication of the different complexity requirements, as well as the different information requirements of each case, Table I presents numerical results for the three mesh network topologies of Fig. 2 based on (29)–(31), and we set $S_{\text{mod}} = 8$ [21]. The normalized information requirements are expressed in terms of the number of links in all possible paths (whose error, guaranteed bandwidth, and queueing delay is conveyed by the overlay network) multiplied by the total number of times this information is updated by the overlay network per MSDU (I_{refresh} with $0 < I_{\text{refresh}} \leq 1$). First, we considered the case of the first node [$h_1, n = 1$ in (29)–(31)] since this includes all the possible paths and all the links in the mesh topology (top of the table). Hence, the results of the top part of Table I show the worst case complexity and information requirements from the viewpoint of an individual node.

Notice that the information cost depends on the frequency of updates received by the overlay network per MSDU, denoted by I_{refresh}. Given I_{refresh} and the required bytes for conveying the status of each link via hop-to-hop feedback, it is straightforward to convert the provided information cost for each case into actual bandwidth overhead for the overlay infrastructure in the multihop wireless mesh network. Since $I_{\text{refresh}} \simeq 0$ for the

TABLE I
COMPLEXITY COMPARISON AND THE ASSOCIATED INFORMATION REQUIREMENTS OF DIFFERENT ALTERNATIVES FOR CROSS-LAYER
OPTIMIZATION FOR: (TOP) THE FIRST NODE OF EACH OF THE THREE TOPOLOGIES OF FIG. 2; (BOTTOM) ALL NODES
IN THE TOPOLOGY. THE BASIC COMPLEXITY UNIT FOR (29)–(31) IS SET TO $C_{exhaustive}(1) = 1$ AND WE
ADDITIONALLY SET $C_{estimation}(1) = 0.2 \cdot C_{exhaustive}(1)$ BASED ON EXPERIMENTAL OBSERVATIONS

Method/Topology Node h_1	Normalized Complexity			Normalized Information requirements		
	T1	T2	T3	T1	T2	T3
End-to-end	624	176.0	64	$14I_{refresh}$	$10I_{refresh}$	$8I_{refresh}$
Localized $\forall i : \rho_i = 1$	28.8	19.2	28.8	$3I_{refresh}$	$2I_{refresh}$	$3I_{refresh}$
Estimation based	0.6	0.4	0.6	0	0	0
Method/Topology All nodes in each topology	Normalized Complexity			Normalized Information requirements		
	T1	T2	T3	T1	T2	T3
End-to-end	1120	352.0	120	$40I_{refresh}$	$28I_{refresh}$	$15I_{refresh}$
Localized $\forall i : \rho_i = 1$	134.4	96	76.8	$14I_{refresh}$	$10I_{refresh}$	$8I_{refresh}$
Estimation based	2.8	2	1.6	0	0	0

estimation-based case, the information cost of this case is practically negligible.

As a second step, we considered the cumulative complexity and information requirements for all the nodes in the multihop mesh network in order to estimate the streaming complexity and information overhead at the system level; the results are presented in the bottom of Table I. We remark that, depending on the topology specification (i.e., average node connectivity) and the chosen method, the estimated complexity scales up to three orders of magnitude. Similarly, there is a large gap between the lowest and highest information requirements for the various approaches among the different topologies. As expected, the more complex the mesh topology, the higher the rate of increase of complexity and information requirements.

VI. EXPERIMENTAL RESULTS

Although we experimented with a variety of video content, we present results with one typical video sequence ("Foreman," 300 frames, CIF format, with 30 Hz replay rate) since this experiment captures the average behavior of our system very well. We used a fully scalable codec [23] and the produced bitstream was extracted at an average bitrate of 2 Mb/s and packetized into MSDUs of data payload not larger than 1000 bytes. The end-to-end delay for the MSDUs of each GOP was set to 0.54 s, which corresponds to the replay duration of one GOP. We remark that although the utilized video coder is not a member of the MPEG family of coders, the assumptions made in Section II-C for the distortion-reduction estimation and the application-layer packet scheduling are also valid for the scalable coder currently standardized by the JSVM group of MPEG/VCEG [24] since it is based on open-loop motion-compensated prediction and update steps followed by embedded quantization and context-based entropy coding. Hence, our methods and experiments are relevant to future systems that will utilize such scalable video coding technology in the context of mesh networks.

We simulated the cases of the multihop mesh network topologies of Fig. 2, labeled T1–T3, under predetermined transmission intervals for each link. Our simulation took into account the different parameters for the various layers, such as varying SINR, transmission overheads at the MAC layer due to MSDU acknowledgments, and polling overheads, as well as queueing and propagation delays in the various links of the mesh network. In order to incorporate the effect of noise and interference, we performed a number of simulations using random values for the SINR of each link, chosen between 15 and 25 dB. Network feedback via the overlay network was conveyed to each hop whenever a significant change in the experienced channel condition occurred. For the end-to-end optimization with network feedback (termed "end-to-end" in our results) this includes the information conveyed from all hops. However, we also considered a localized case where the information horizon was set to the direct neighborhood of each hop (termed "localized" in our results—this information horizon is shown pictorially in Fig. 2) and the remaining network parameters were estimated, as explained in Section IV-B. In addition, a purely estimation-based case was also considered with no "horizon," where the only available information is the channel SINR range [s_{min}, s_{max} of (28)] for each link, communicated by the overlay network infrastructure whenever the channel variation exceeded 2 dB (termed "estimation based") from the estimated value given by (28). This ensured that the information cost for the dissemination of the network information is minimal compared with the other alternatives, as indicated in Table I. Notice that, both for the "localized" case, as well as for the "estimation-based" case, the theoretical framework of (18)–(28) was used.

Apart from the various alternatives of the proposed optimization, we also derive results with streaming under two other optimization algorithms. The first case is optimization based on the expected transmission count (ETX) [8], where the utility function is chosen such that the retransmission limit of each MSDU is set based on the effective network bandwidth and the expected error rate. This case considers the MSDU delay deadline from a purely network-centric approach [8], i.e., it does not use the constraints set in (15) and (16), but rather restricts the MSDU delay deadline based on link loss ratios and the available throughput [8]. It was termed "ETX optimized" in our results. Second, the case of selecting the link with the highest effective bandwidth

TABLE II
AVERAGE PSNR RESULTS (Y-CHANNEL −25 RUNS WITH 300 VIDEO FRAMES PER RUN) FOR VIDEO STREAMING IN THE MULTIHOP NETWORKS OF FIG. 2

Method/Topology	Medium bandwidth case PSNR (dB)			Low bandwidth case PSNR (dB)		
	T1	T2	T3	T1	T2	T3
End-to-end	37.08	36.04	35.59	32.93	31.86	31.70
ETX optimization	36.09	34.20	34.33	31.39	30.63	30.79
Highest Bandwidth	33.94	31.30	31.41	29.25	27.52	28.47
Localized	34.31	32.35	30.53	30.44	29.37	29.29
Estimation based	33.81	29.88	28.69	28.98	27.06	28.35

TABLE III
AVERAGE LINK UTILIZATION FACTOR AND PLR FOR EACH LINK OF THE THREE TOPOLOGIES OF FIG. 2 FOR THE CASE OF
MEDIUM- AND LOW-BANDWIDTH SIMULATION. THE NOMINAL MSDU SIZE $\bar{L} = 1000$ BYTES WAS ASSUMED

Link/Topology	Medium bandwidth case						Low bandwidth case					
	Average Utilization (%)			Average PLR (%)			Average Utilization (%)			Average PLR (%)		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
$h_1 \rightarrow h_2$	9.49	10.40	10.18	11.30	14.71	13.13	6.86	6.81	7.42	8.82	10.75	13.14
$h_1 \rightarrow h_3$	10.16	11.07	11.46	10.38	16.04	13.49	6.77	7.17	6.72	9.22	14.15	16.68
$h_1 \rightarrow h_4$	10.19	N/A	11.63	13.15	N/A	13.60	6.37	N/A	7.74	10.74	N/A	13.40
$h_2 \rightarrow h_3$	N/A	N/A	6.40	N/A	N/A	5.22	N/A	N/A	5.85	N/A	N/A	13.21
$h_2 \rightarrow h_4$	N/A	9.17	7.56	N/A	13.11	9.62	N/A	5.95	6.20	N/A	14.00	9.66
$h_2 \rightarrow h_5$	10.38	10.71	9.13	13.57	16.57	11.07	6.87	6.66	6.86	12.26	12.80	13.90
$h_3 \rightarrow h_4$	N/A	9.69	11.32	N/A	15.51	10.09	N/A	6.76	7.12	N/A	12.92	15.33
$h_3 \rightarrow h_6$	10.61	10.95	10.79	10.90	11.65	10.46	6.84	6.97	7.08	12.25	19.83	15.35
$h_4 \rightarrow h_5$	N/A	12.20	8.86	N/A	13.33	11.86	N/A	6.86	6.40	N/A	13.90	11.95
$h_4 \rightarrow h_6$	N/A	8.93	8.09	N/A	13.27	8.15	N/A	7.17	6.39	N/A	12.05	5.97
$h_4 \rightarrow h_7$	11.94	N/A	12.54	14.23	N/A	13.32	6.98	N/A	7.70	15.06	N/A	14.95
$h_5 \rightarrow h_6$	N/A	N/A	7.00	N/A	N/A	5.42	N/A	N/A	3.75	N/A	N/A	5.55
$h_5 \rightarrow h_7$	11.49	13.81	12.76	15.52	10.43	11.89	7.18	7.46	7.79	16.30	17.05	15.80
$h_6 \rightarrow h_7$	11.46	12.26	13.83	14.33	12.25	9.25	7.49	8.38	7.82	15.31	13.16	13.30
Average:	**10.72**	**10.92**	**10.11**	**12.92**	**13.69**	**10.47**	**6.92**	**7.02**	**6.77**	**12.50**	**14.06**	**12.73**

was realized for the routing of each MSDU, since it corresponds to the popular solution for optimized routing [25] (termed as the "highest bandwidth" solution). Notice that, in both cases, the best modulation was established as in the "end-to-end" case, and each link's status information was also used for these cases, as conveyed by the overlay network infrastructure. As a result, the differences in performance stem purely from the different performance utilities that were chosen during the MSDU routing and path selection. Effectively, this separates the fully network-aware methods (proposed "end-to-end," "highest bandwidth" [25], and "ETX optimized" [8]) from the partial network-aware approaches (proposed "localized" and "estimation based"). In addition, within the fully network-aware methods, the difference in the performance utilities means that only the "end-to-end" approach fully utilizes application-layer, MAC, and PHY parameters via the optimization framework of (12)–(16).

Indicative results for the obtained average peak-signal-to-noise ratio (PSNR) of each method are given in Table II (25 runs per method/case). Two representative cases of medium and low average transmission bandwidth were chosen. For each case, the average percentage of bandwidth utilization and the packet loss rate (PLR) per link is shown in Table III. These results were generated with the "end-to-end" case but similar results were obtained for the remaining cases as well. The results of Table III demonstrate that for each case (medium or low bandwidth), the predetermined time reservation for the video flow packets per link leads to similar average link utilization in all three topologies, i.e., within a 4%–8% margin. Nevertheless, the obtained average PLRs differ for each topology, as well as the obtained average PSNR for each method, as indicated in Table II. In general, since the average link utilization is similar within each of the medium- and low-bandwidth utilization cases, the topology with the most active links for the video traffic is expected to provide the highest video quality. This is indeed true for the corresponding results of each method as seen from the rows of Table II. Nevertheless, the relationship between PSNR and PLR for each topology (as reported in Table III) is not obvious.

In order to understand better the relationship between the obtained PLR for each case and the derived PSNR, the percentage of losses for the video packets when clustered into eight distinct distortion categories is presented in Fig. 5. The second topology of Fig. 2 was used for these results; similar results have been obtained for the remaining topologies. Notice that our choice of eight distinct categories is only performed for illustration pur-
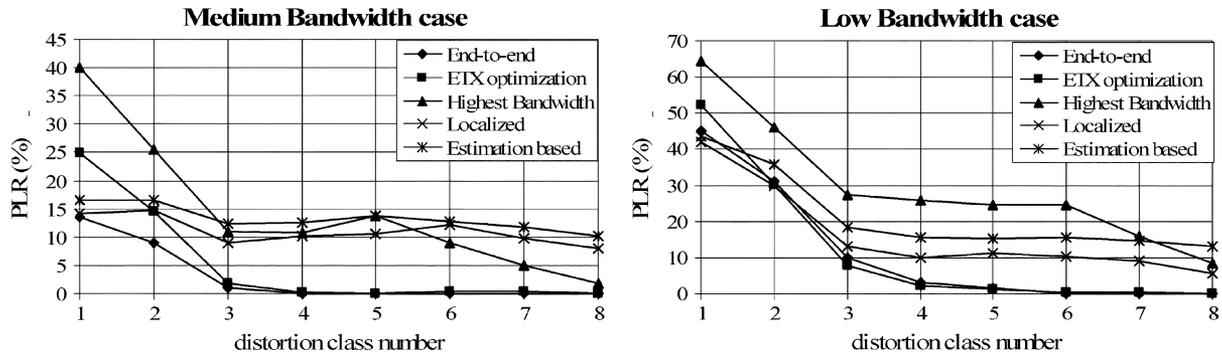
Fig. 5. Percentage of losses for each packet distortion-reduction class ($\mathrm{Cat.1}$ = least significant packets; $\mathrm{Cat.8}$ = most significant). (a) Medium-bandwidth case. (b) Low-bandwidth case. The results correspond to topology T2 of Fig. 2.

poses, since each packet is associated with its own distortion-reduction. In our simulations, the packet losses were mainly due to deadline violation, since each hop drops the packets which have already expired. The results of Fig. 5 indicate, for all the scenarios under consideration, that scheduling at the application layer by expected distortion-reduction leads to reduced losses for the most significant classes of packets. This justifies our use of a scalable video coder that permits such a scheduling. However, each method achieves different PSNR performance and PLRs depending on its chosen utility and the presence of network feedback.

As shown in the results of Table II, the "end-to-end" case outperforms all other methods by a significant margin. The "ETX optimization" appears to perform relatively well, even though it is outperformed by approximately 1.3 dB by the "end-to-end" case. The "localized" case appears to outperform the popular "highest bandwidth" case in the vast majority of cases, even though the latter uses full feedback for the status of all the links in each multihop topology. Although the "estimation-based" case performs worst, this case requires almost no network feedback and, as shown in Table I, has the lowest complexity. Moreover, in the case of low average bandwidth, this case is comparable to the "highest bandwidth." This is expected since the "highest bandwidth" approach provides less intelligent decisions when most of the links have low effective throughput. Finally, a comparison of the results for the different topologies reveals that, as expected, the higher the connectivity, the better the average performance of all methods. Nevertheless, this comes at a higher allocation of resources in the multihop mesh topology, and it additionally has higher complexity and requires more feedback for the condition of all the links, as demonstrated in Table I.

Our results highlight several important issues in network design and infrastructure. First, it was shown that having frequent feedback via an overlay network about the link conditions and performing end-to-end optimization with the appropriate utility function offers significant improvements in the achievable video quality. Indeed, the "end-to-end" and "ETX optimized" cases outperform the remaining algorithms by 3–5 dB, in all cases (Table II). Second, the importance of choosing a cross-layer distortion-capacity utility function is highlighted by the fact that both methods outperform the conventional "highest bandwidth" scenario. Moreover, the proposed utility of (12) and the derivation of the MSDU retransmission limit based on the delay limit for the video transmission [(15)–(17)] appear to be

the best choice for video streaming applications. Third, higher connectivity in the multihop mesh topology leads to better video streaming performance, at the expense of complexity and network feedback requirements.

Fourth, the study of the PLRs reported in Fig. 5 in conjunction with the PSNR results of Table II reveals that prioritization of video packets with respect to distortion-reduction incurred in the decoded video is extremely important. For example, even though the "estimation-based" case has lower average PLR from the "highest bandwidth" case, it performs worse in terms of PSNR since the latter achieves lower PLR for the most significant classes of packets. This result emphasizes the fact that, in the case of analysis of multimedia transmission over wireless, average PLRs that do not consider the significance of the various packets for the application are not always relevant metrics for the system performance.

Finally, it appears that even a limited horizon of information in the network infrastructure can be extremely beneficial. We believe that the determination of an appropriate "horizon" of information that provides the optimal tradeoff between the overhead at the overlay network versus the improvement offered by utilizing dynamic network feedback is an interesting research direction. Moreover, the dynamic adaptation of such a "horizon" in function of the network variations or the mesh topology specification (i.e., simple versus complex mesh networks and static versus dynamic scenarios) could be examined.

VII. CONCLUSION

Delay-constrained video streaming over multihop wireless mesh networks is an application that deserves considerable attention due to the research challenges imposed by such a service, as well as due to the important role that robust and efficient multimedia services have when it comes to commercial deployment of such networks in office and residential areas. We investigated a framework where QoS guarantees are provided for video transmission over a variety of links in a multihop network using IEEE 802.11a/e. The integrated cross-layer solution that maximizes the product of the expected video quality with the link utilization appears to provide significant improvement over other optimized solutions. Moreover, the utilization of network information (for the dynamically changing conditions of the various hops) gathered via overlay-network feedback, appears to be of paramount importance for the overall video quality at the receiver hop.

Although the proposed algorithm operates per video packet and can potentially incur significant complexity and communication overhead for the overlay network infrastructure, there is a significant potential for improved video streaming performance. This motivates us to investigate the problem further and attempt to explore the best granularity for the optimization, as well as the network feedback that provides optimal quality/complexity/robustness in a distributed video streaming scenario over the hops of the mesh network. Finally, under the proposed paradigm, the issues of collaborative streaming of multiple flows and fairness deserve significant attention in future research.

REFERENCES

[1] K. Holt, "Wireless LAN: Past, present, and future," *Proc. IEEE Design, Autom., Test in Europe*, vol. 3, pp. 92–93, 2005.

[2] W. Wei and A. Zakhor, "Multipath unicast and multicast video communication over wireless ad hoc networks," in *Proc. Int. Conf. Broadband Netw.*, 2002, pp. 496–505.

[3] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 390–404, Apr. 2006.

[4] 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.*, vol. 12, no. 4, pp. 59–65, Aug. 2005.

[5] M. van der Schaar, S. Krishnamachari, S. Choi, and X. Xu, "Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1752–1763, Dec. 2003.

[6] A. Butala and L. Tong, "Cross-layer design for medium access control in CDMA ad-hoc networks," *EURASIP J. Appl. Signal Process.*, to be published.

[7] 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. Sel. Areas Commun.*, vol. 23, no. 1, pp. 136–150, Jan. 2005.

[8] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high throughput path metric for multi-hop wireless routing," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, 2003, pp. 134–146.

[9] R. Draves, J. Padhye, and B. Zill, "Routing in multi-radio, multi-hop wireless mesh networks," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, 2004, pp. 114–128.

[10] V. Gambiroza, B. Sadeghi, and E. W. Knightly, "End-to-end performance and fairness in multihop wireless backhaul networks," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, 2004, pp. 287–301.

[11] R. L. Cruz and A. V. Santhanam, "Optimal routing, link scheduling and power control in multihop wireless networks," in *Proc. Joint Conf. IEEE Comput. Commun. Society*, 2003, vol. 1, pp. 702–711.

[12] *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)*, IEEE 802.11e/D5.0, Jun. 2003.

[13] M. Kodialam and T. Nandagopal, "Characterizing achievable rates in multi-hop wireless networks: The joint routing and scheduling problem," in *Proc. ACM Int. Conf. Mobile Comput. Netw.*, 2003, pp. 42–54.

[14] P. Ansel, Q. Ni, and T. Turletti, "An efficient scheduling scheme for IEEE 802.11e," in *Proc. IEEE Workshop on Model. and Opt. in Mobile, Ad-Hoc and Wireless Netw.*, Cambridge, U.K., Mar. 2004.

[15] M. Wang and M. van der Schaar, "Operational rate-distortion modeling for wavelet video coders," *IEEE Trans. Signal Process.*, to be published.

[16] D. Taubman and J. Thie, "Optimal erasure protection for scalably compressed video streams with limited retransmission," *IEEE Trans. Image Process.*, vol. 14, no. 8, pp. 1006–1019, Aug. 2005.

[17] M. van der Schaar and D. Turaga, "Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission," *IEEE Trans. Multimedia*, to be published.

[18] Y. Andreopoulos, R. Keralapura, M. van der Schaar, and C.-N. Chuah, "Failure-aware, open-loop, adaptive video streaming with packet-level optimized redundancy," *IEEE Trans. Multimedia*, to be published.

[19] D. Krishnaswamy and J. Vicente, "Scalable adaptive wireless networks for multimedia in the proactive enterprise," *Intel Technol. J.*, vol. 8, no. 4, Nov. 2004. [Online]. Available: www: http://developer.intel.com/technology/itj/2004/volume08issue04/art04_scalingwireless/p01_abstract.htm

[20] S. Roy, M. Covell, J. Ankcorn, S. Wee, and T. Yoshimura, "A system architecture for managing mobile streaming media services," in *Proc. IEEE Distrib. Comput. Syst. Workshop*, 2003, pp. 408–413.

[21] D. Krishnaswamy, "Network-assisted link adaptation with power control and channel reassignment in wireless networks," in *Proc. 3G Wireless Conf.*, 2002, pp. 165–170.

[22] K.-B. Song and S. A. Mujtaba, "On the code-diversity performance of bit-interleaved coded OFDM in frequency-selective fading channels," in *Proc. IEEE Veh. Technol. Conf.*, 2003, vol. 1, pp. 572–576.

[23] Y. Andreopoulos, A. Munteanu, J. Barbarien, M. van der Schaar, J. Cornelis, and P. Schelkens, "In-band motion compensated temporal filtering," *Signal Process.: Image Commun., Special Issue on "Subband/Wavelet Interframe Video Coding"*, vol. 19, no. 7, pp. 653–673, Aug. 2004, .

[24] J. R. Ohm, "Advances in scalable video coding," in *Proc. IEEE*, Jan. 2005, vol. 93, no. 1, pp. 42–56.

[25] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Chapter in Mobile Computing*, T. Imielinski and H. Korth, Eds. Norwell, MA: Kluwer, 1996.

Yiannis Andreopoulos (M'00) was born in Aeghion, Greece, in 1977. He received the Electrical Engineering Diploma and the M.Sc. degree in signal and image processing systems from the University of Patras, Patras, Greece, and the Ph.D. degree from the Vrije Universiteit Brussel, Brussels, Belgium.

Currently, he is working as a Postdoctoral Researcher at the University of California at Los Angeles (UCLA). During 2002–2003, he contributed to the ISO/IEC JTC1/SC29/WG11 (MPEG) committee (Scalable Video Coding Group). His research interests are in the fields of transforms, complexity modeling for media systems, video coding, and video transmission through unreliable media, e.g., wireless networks and the Internet.

Nicholas Mastronarde received the B.Sc. and M.Sc. degrees in electrical engineering from the University of California at Davis (UCD), in 2005 and 2006, respectively. He is currently working towards the Ph.D. degree at the University of California at Los Angeles (UCLA).

During his graduate study at UCD, he conducted research in video streaming over multihop wireless mesh networks and published several international conference papers and a journal paper. His research interests include wireless multimedia transmission, video coding, signal processing, and information theory.

Mihaela van der Schaar (SM'04) received the Ph.D. degree from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2001.

She is now an Assistant Professor in the Department of Electrical Engineering, University of California at Los Angeles (UCLA). Prior to this, she was a Senior Researcher at Philips Research, The Netherlands, and the U.S., where she led a team of researchers working on multimedia compression, networking, communications, and architectures. In 2003, she was also an Adjunct Assistant Professor at Columbia University. From 2003 to 2005, she was an Assistant Professor in the Department of Electrical and Computer Engineering, University of California at Davis. Since 1999, she was an active participant to the ISO Motion Picture Expert Group (MPEG) standard to which she made more than 50 contributions and for which she received three ISO recognition awards. She was also chairing for three years the ad hoc group on MPEG-21 scalable video coding, and also co-chairing the MPEG Ad Hoc Group on multimedia testbed. She holds 22 granted U.S. patents and several more pending. She has published extensively on multimedia compression, processing, communications, networking, and architectures.

Prof. van der Schaar received the National Science Foundation (NSF) Career Award (2004) and the IBM Faculty Award (2005). She was an Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA and *SPIE Electronic Imaging Journal* from 2002 to 2005. Currently, she is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEM FOR VIDEO TECHNOLOGY and an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS.