

Cross-layer Video Streaming Over 802.11e-Enabled Wireless Mesh Networks

N. Mastronarde¹, Y. Andreopoulos², M. van der Schaar², D. Krishnaswamy³ and J. Vicente³

¹University of California Davis
Dept. of Electrical and Computer Engineering
2346 Ac. Surge III, One Shields Ave. 3129
Davis, CA, 95616, USA

²University of California Los Angeles
Dept. of Electrical Engineering
54-147 Eng. IV Building, 420 Westwood Plaza,
Los Angeles, CA 90095-1594, USA

³Intel Corporation
1900 Prairie City Road
Folsom CA, 95630, USA

ABSTRACT

We propose an integrated cross-layer optimization algorithm for maximizing the decoded video quality of delay-constrained streaming in a quality-of-service (QoS) enabled multi-hop wireless mesh network. The key to our algorithm is the synergistic optimization of control parameters at each node of the multi-hop network, across the protocol layers - application, network, medium access control (MAC) and physical (PHY) layers, as well as end-to-end, i.e. across the various network nodes. To drive this optimization, we assume an overlay network infrastructure, which conveys information on the conditions of each link. Quantitative results are presented that demonstrate the merits and the need for cross-layer optimization in an efficient solution for real-time video transmission using existing protocols and infrastructures.

1. INTRODUCTION

The problem of multi-hop video streaming has recently been studied under a variety of scenarios [1] [2]. 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 [3] [4] [5]. Consequently, the inherent network dynamics occurring in a multi-hop 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 multi-hop wireless networks [6] 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 [6], as well as expected transmission time or fairness issues [7], can significantly impact the actual end-to-end network throughput. Motivated by this work, we show that network aware, cross-layer, transmission strategies for delay-constrained video streaming over multi-hop wireless mesh networks, can provide significant improvements in the decoded video quality.

In this paper we are concerned with developing an integrated video streaming paradigm enabling cross-layer interaction across the protocol stack and across the multiple hops of immobile, fixed-infrastructure, wireless mesh networks. This problem is considered under the constraints of existing protocols for the interconnection of the mesh's nodes thereby reducing potential deployment costs and also increasing interoperability.

Our focus is on real-time transmission of an individual video bitstream across a multi-hop 802.11a/e wireless network. Our optimization algorithm assumes that each application (video flow) reserves a predetermined transmission opportunity (TXOP) interval at each node, during which contention-free access to the medium is provided. This reservation can be performed using the

HCCA protocol of IEEE 802.11e [9] and can be determined based on the amount of flows sharing the network. We further assume that an overlay network topology can convey (in real-time) per-link information about the expected bit error rate (BER), the queuing delay, and the guaranteed bandwidth under the dynamically-changing modulation at the PHY.

Under the above assumptions, the paper makes the following contributions. For video packets of each node in the mesh network, we propose an optimization framework that jointly determines per packet: (a) the optimal modulation at the PHY, (b) the optimal retry limit at the MAC, (c) the optimal path (route) to the receiver in the remaining part of the mesh network and (d) application-layer optimized packet scheduling, given a predetermined topology and time reservation per link using IEEE 802.11e.

The paper is organized as follows. Section II presents the necessary definitions and formulations for the network parameters. Section III poses the cross-layer optimization problem. The proposed solutions are presented in Section IV while Section V demonstrates indicative results, including comparisons with another well known approach from the literature. Conclusions are presented in Section VI.

2. WIRELESS MULTI-HOP MESH TOPOLOGY SPECIFICATION

Consider a wireless multi-hop mesh network consisting of N nodes. One example topology with $N = 7$ is shown in Figure 1. Node h_1 represents the original video source, while node h_N is the destination node (video client).

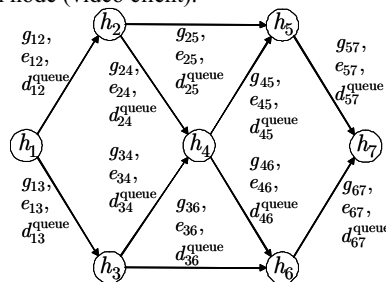


Figure 1: A network topology with seven nodes. Each link ij is associated with the corresponding allocated bandwidth (g_{ij}), the error rate observed on the link (e_{ij}), and the corresponding delay due to the video queue (d_{ij}^{queue}).

For a generic multi-hop wireless mesh network, we consider the connectivity structure \mathcal{P} :

$$\mathcal{P} = \{\mathbf{p}_1, \dots, \mathbf{p}_M\} \quad (1)$$

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

$$\mathbf{p}_i = [l_{i,1} \quad l_{i,2} \quad \dots \quad l_{i,\rho_i^{total}-2} \quad l_{i,\rho_i^{total}-1}] \quad (2)$$

where each component $l_{i,j} = (p_{i,j}, p_{i,j+1})$, $1 \leq j < \rho_i^{\text{total}}$, indicates the j -th link of path i between in-path nodes $p_{i,j}$ and $p_{i,j+1}$, and $\rho_i^{\text{total}} - 1$ is the total number of links participating in the network path \mathbf{p}_i . Notice that (1) and (2) apply both to the end-to-end topology of interest and to the topology between any intermediate node and the terminal (client) node in the mesh network utilized for video transmission. 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 (sub-network). Finally, it is important to mention that all the proposed algorithms assume the non-existence of routing loops, i.e. the mesh network between the current node and the destination node can be represented by a tree graph.

2.1 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 bit error rate as $e(l_{i,j})$ and the physical layer rate as $R_{\text{phy}}(l_{i,j})$. Both the bit error rate and the physical layer rate are estimated based on SINR models of the wireless links using the following approximations:

$$R_{\text{phy}}(l_{i,j}) = R_{\text{phy}}^{\text{max}}(l_{i,j}) \cdot \left[1 + e^{-\mu(s(l_{i,j})-\delta)}\right]^{-1}. \quad (3)$$

$$e(l_{i,j}) = \left[1 + e^{\mu(s(l_{i,j})-\delta)}\right]^{-1}. \quad (4)$$

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.

Under a predetermined negotiation of traffic specification parameters for each link in the mesh network (using the HCCA protocol [9]), each link can provide a guaranteed bandwidth $g(l_{i,j})$ at the application layer. Following the HCCA specification [9], this bandwidth is linked with the traffic specification parameters by¹:

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

where $t_{\text{TXOP}}(l_{i,j})$ is the TXOP duration over link $l_{i,j}$ obtained by dividing the TXOP provided by the HCCA admission control for the video flow traffic of node $p_{i,j}$ over each of $p_{i,j}$'s forwarding links, \bar{L} is the nominal MAC service data unit (MSDU) size², $t_{\text{SI}}(l_{i,j})$ is the specified duration of the service interval [9] 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.

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:

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

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

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

We also define the parameter $c(l_{i,j})$ corresponding to the remaining time interval for which link $l_{i,j}$ can support the video-flow traffic under HCCA. For each TXOP interval, $c(l_{i,j})$ can be calculated as:

$$c(l_{i,j}) = \max\{t_{\text{TXOP}}(l_{i,j}) - d_{\text{queue}}(l_{i,j}), 0\}. \quad (8)$$

where $d_{\text{queue}}(l_{i,j})$ is the expected time required to transmit the MSDUs ahead of MSDU v in link $l_{i,j}$'s queue.

3. PROBLEM FORMULATION

Assume a set of N wireless nodes, with h_1 being the video encoder (server) and h_N the video decoder (client), and a connectivity structure \mathbf{P} with M paths, where each path i , $1 \leq i \leq M$, consists of ρ_i^{total} nodes. In addition, although in practice each node has an allocated HCCA TXOP, we assume a predefined TXOP duration $t_{\text{TXOP}}(l_{i,j})$ for each link $l_{i,j}$, with $1 \leq j < \rho_i^{\text{total}}$, and a link adaptation mechanism at the physical layer that can operate at an MSDU granularity. The end-to-end cross-layer optimization that 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:

$$\forall v : \left\{ \mathbf{p}_i^*, T_{\mathbf{p}_i}^{\text{max}*}, m(l_{i,j}) \right\} = \arg \max_{\mathbf{p}_i \in \mathcal{P}, \forall l_{i,j} \in \mathbf{p}_i} \left[\min_{1 \leq j < \rho_i^{\text{total}}} \{c(l_{i,j})\} \Delta_{v,\text{expected}} \right] \quad (9)$$

where:

$$\Delta_{v,\text{expected}} = \left[1 - [e_{\mathbf{p}_i}(L_v)]^{t_{\mathbf{p}_i}^{\text{mean}}(T_{\mathbf{p}_i}^{\text{max}}(L_v))} \right] \cdot \Delta_v \quad (10)$$

and $e_{\mathbf{p}_i}(L_v)$ is from (7), $t_{\mathbf{p}_i}^{\text{mean}}(T_{\mathbf{p}_i}^{\text{max}}(L_v))$ is the expected number of retransmissions for an MSDU of size L_v on path \mathbf{p}_i , $T_{\mathbf{p}_i}^{\text{max}}(L_v)$ will be explained below, and Δ_v is a coder dependent value proportional to an expected gain in video quality if the decoder receives MSDU v before d_v^{deadline} , the time after which the video data encapsulated in the v is useless to the decoder.

If MSDU v incurs an error during transmission it is retransmitted by the sender node a maximum of $T_{\mathbf{p}_i}^{\text{max}}(L_v)$ times where $T_{\mathbf{p}_i}^{\text{max}}(L_v)$ is the minimum of the following two quantities:

- The number of times a packet of size L_v can be transmitted from the current node to the destination node before exceeding the delay limit d_v^{deadline} .
- The number of times a packet of size L_v can be transmitted from the current node to the destination node before the transmitting node's TXOP for the specific video flow ends.

Note that all $T_{\mathbf{p}_i}^{\text{max}}(L_v)$ retransmissions may be exhausted over one link or over the whole path. Additionally, MSDU v will be dropped if it exhausts all $T_{\mathbf{p}_i}^{\text{max}}(L_v)$ retransmissions, if:

$$d_v^{\text{deadline}} < \sum_{j=1}^{\rho_i^{\text{total}}-1} d_{\text{queue}}(l_{i,j}), \text{ or if } \min_{\forall l_{i,j} \in \mathbf{p}_i} \{c(l_{i,j})\} = 0.$$

While respecting network constraints and considering video codec properties, the optimization of (9) 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 time within the reserved transmission period at the most congested link) and the expected source distortion-reduction of (10). In this way, we minimize congestion across the various links (since the minimum link capacity is maximized), and concurrently maximize the expected distortion reduction. The granularity of this optimization is one MSDU. However, coarser granularities could also be considered, in order to reduce complexity.

4. VIDEO STREAMING OPTIMIZATION IN THE MESH NETWORK

In this section we derive an algorithm that determines the optimal parameters for (9) under a predetermined deadline for each MSDU v (given by d_v^{deadline}) and a predetermined TXOP duration per node, which is set by the HCCA admission control once the video flow is scheduled for transmission and then divided among

¹ For notational simplicity we do not particularly indicate the dependence of $e(l_{i,j})$, $R_{\text{phy}}(l_{i,j})$, $R_{\text{phy}}^{\text{max}}(l_{i,j})$, and $g(l_{i,j})$ on the modulation $m(l_{i,j})$.

² In this paper, we assume that one video packet is encapsulated in one MSDU and the two terms are used interchangeably.

each forwarding link into $t_{\text{TXOP}}(l_{i,j})$. 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.

4.1 End-To-End Optimization

The optimization of (9) 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 noise-interference ratio (SINR). In addition, we assume that $d_{\text{queue}}(l_{i,j})$ is communicated to the sender node via continuous feedback using an overlay network infrastructure [10] that uses real time protocols for conveying information from different layers.

The proposed optimization algorithm is given in Figure 2. 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 queuing-delay assumptions used when scheduling at the sender node, can be incorporated at 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 Figure 2 performs an exhaustive search to find the optimal path for transmission of the MSDU.

-
- 1 For each node that has non-expired MSDUs in its queue
 - 2 Extract the network connectivity structure \mathcal{P} (eq. (1), (2))
 - 3 For the MSDU v at the output of the queue of the sender node
 - 4 For each path \mathbf{p}_i (topology emanating from the sender node)
 - 5 For each link $l_{i,j}$ of path \mathbf{p}_i
 - 6 For each modulation strategy $m(l_{i,j})$
 - 7 Calculate $e_{l_{i,j}}(L_v)$, $e_{\mathbf{p}_i}(L_v)$, $t_{\mathbf{p}_i}^{\text{mean}}(T_{\mathbf{p}_i}^{\text{max}}(L_v))$, $d_{\text{queue}}(l_{i,j})$.
 - 8 Calculate $T_{\mathbf{p}_i}^{\text{max}}(L_v)$.
 - 9 Under the calculated $T_{\mathbf{p}_i}^{\text{max}}(L_v)$, evaluate (9)
 - 10 Compare with previous best choice, retaining the max.
 - 11 Schedule the MSDU according to the established $\{\mathbf{p}_i^*, T_{\mathbf{p}_i^*}^{\text{max}}, m(l_{i,j})\}$
-

Figure 2. 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 multi-hop wireless network.

4.2 Localized Optimization

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 end-to-end 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 and limited.

5. EXPERIMENTAL RESULTS

A typical video sequence was selected for our streaming purposes (“Foreman”, 300 frames, CIF format, with 30 Hz replay rate). We used a fully-scalable codec [11] and the produced bitstream was extracted at an average bitrate of 2 Mbps 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 sec, which corresponds to the replay duration of one GOP.

We simulated the case of the multi-hop mesh network shown in Figure 1 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 acknowledgements and polling overheads, as well as queuing 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 node 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 nodes. However, we also considered a localized case where information was available only in the direct neighborhood of each node (termed “Localized” in our results) and the remaining parameters were estimated based on the expected transmissions via each node.

We also derive results for the popular solution for optimized routing that selects the link with the highest effective bandwidth for routing each MSDU (termed as the “Highest Bandwidth” solution). Notice that, in all cases, the best modulation was established as in the proposed algorithm. As a result, the differences in performance stem from the fact that different performance utilities were chosen during the MSDU routing and path selection.

Method	Medium bandwidth case PSNR (dB)	Low bandwidth case PSNR (dB)
End-to-end	36.04	31.86
Highest Bandwidth	31.30	27.52
Localized	32.35	29.37

Table 1. Average PSNR results (Y-channel – 25 runs with 300 video frames per run) for video streaming in the multi-hop network of Figure 1.

Indicative results for the obtained average PSNR of each method are given in Table 1 (25 runs per method/case) for two representative cases of medium and low expected bandwidth are shown. The average bandwidth and packet loss rate (PLR) per link is shown in Table 2. Finally, the percentage of losses for each case for the video packets belonging to various distortion categories is presented in Figure 3, for the example of the medium-bandwidth case. In our simulations, the packet losses were mainly due to deadline violation, since each node drops the packets which have already expired. The results of Figure 3 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.

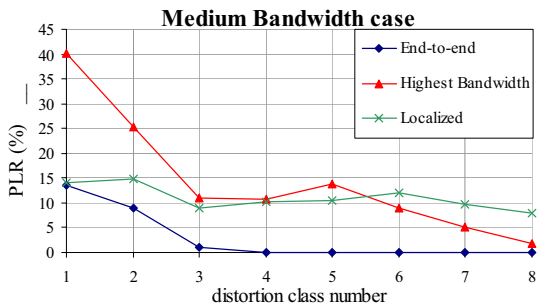


Figure 3. Percentage of losses for each packet distortion-reduction class (Cat.1=least significant packets; Cat.8=most significant).

Link	Medium bandwidth case		Low bandwidth case	
	Average Utilization (%)	Average PLR (%)	Average Utilization (%)	Average PLR (%)
$h_1 \rightarrow h_2$	10.40	14.71	6.81	10.75
$h_1 \rightarrow h_3$	11.07	16.04	7.17	14.15
$h_2 \rightarrow h_4$	9.17	13.11	5.95	14.00
$h_2 \rightarrow h_5$	10.71	16.57	6.66	12.80
$h_3 \rightarrow h_4$	9.69	15.51	6.76	12.92
$h_3 \rightarrow h_6$	10.95	11.65	6.97	19.83
$h_4 \rightarrow h_5$	12.20	13.33	6.86	13.90
$h_4 \rightarrow h_6$	8.93	13.27	7.17	12.05
$h_5 \rightarrow h_7$	13.81	10.43	7.46	17.05
$h_6 \rightarrow h_7$	12.26	12.25	8.38	13.16

Table 2. Average link utilization factor and PLR for each link of Figure 1 or the case of medium and low bandwidth simulation. The nominal MSDU size $\bar{L} = 1000$ bytes was assumed.

Our results highlight several important issues in network design and infrastructure. Firstly, it was shown that network aware end-to-end optimization with the appropriate utility function offers significant improvements in the achievable video quality. Secondly, the importance of choosing a distortion-capacity utility function is highlighted by the fact that both methods outperform the conventional “Highest Bandwidth” scenario. Moreover, the proposed utility of (9) and the derivation of the MSDU retransmission limit based on delay constraints for the video transmission appear to be the best choice for video streaming applications. Finally, it appears that even limited information in the network infrastructure can be extremely beneficial.

6. CONCLUSIONS

Delay-constrained video streaming over multi-hop 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 multi-hop 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 previous solutions. Moreover, the utilization of network

information (for the dynamically changing conditions of the various nodes) gathered via overlay-network feedback, appears to be of paramount importance for the overall video quality at the receiver node.

Although the proposed algorithm operates per video packet and can potentially incur significant complexity and communication overhead for the overlay network infrastructure, through further research there is a significant potential for improved video streaming performance.

ACKNOWLEDGEMENTS

This work was supported in part by NSF CAREER, Intel IT Research and UC-MICRO.

REFERENCES

- [1] W. Wei, and A. Zakhor, “Multipath unicast and multicast video communication over wireless ad hoc networks,” *Proc. Internat. Conf. on Broadband Networks*, pp. 496-505, 2002.
- [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 Communications Mag.*, vol. 12, no. 4, pp. 59-65, Aug. 2005.
- [3] M. van der Schaar, et al., “Adaptive cross-layer protection strategies for robust scalable video transmission over 802.11 WLANs,” *IEEE J. on Select. Areas in Commun.*, vol. 21, no. 10, pp. 1752-1763, Dec. 2003.
- [4] A. Butala, and L. Tong, “Cross-layer design for medium access control in CDMA ad-hoc networks,” *EURASIP J. on Applied Signal Processing*, to appear.
- [5] 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. on Select. Areas in Commun.*, vol. 23, no. 1, pp. 136-150, Jan. 2005.
- [6] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, “A high throughput path metric for multi-hop wireless routing,” *Proc. ACM Internat. Conf. on Mob. Computing and Networking*, MOBICOM, pp. 134-146, 2003.
- [7] V. Gambiroza, B. Sadeghi, and E. W. Knightly, “End-to-end performance and fairness in multihop wireless backhaul networks,” *Proc. ACM Internat. Conf. on Mob. Computing and Networking*, MOBICOM, pp. 287-301, 2004.
- [8] R. L. Cruz, and A. V. Santhanam, “Optimal routing, link scheduling and power control in multihop wireless networks,” *Proc. Joint Conf. of IEEE Comput. and Commun. Societ.*, INFOCOM, vol. 1, pp. 702-711, 2003.
- [9] *IEEE 802.11e/D5.0, 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)*, June 2003.
- [10] D. Krishnaswamy, and J. Vicente, “Scalable adaptive wireless networks for multimedia in the proactive enterprise.” *Intel Technology Journal*.
- [11] 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.