Dynamic Resource Allocation for MGS H.264/AVC Video Transmission over Link-Adaptive Networks

Hassan Mansour*, *Student Member, IEEE,* Yaser Pourmohammadi Fallah, *Member, IEEE,*
Panos Nasiopoulos, *Member, IEEE,* and Vikram Krishnamurthy, *Fellow, IEEE*

***Abstract*—In this paper, we address the problem of efficiently allocating network resources to support multiple scalable video streams over a constrained wireless channel. We present a resource allocation framework that jointly optimizes the operation of the link adaptation scheme in the physical layer (PHY), and that of a traffic control module in the network or medium access control (MAC) layer in multi-rate wireless networks, while satisfying bandwidth/capacity constraints. Multi-rate networks, such as IEEE 802.16 or IEEE 802.11, adjust the PHY coding and modulation schemes to maintain the reliability of transmission under varying channel conditions. Higher reliability is achieved at the cost of reduced PHY bit-rate which in turn necessitates a reduction in video stream bit-rates. The rate reduction for scalable video is implemented using a traffic control module. Conventional solutions operate unaware of the importance and loss tolerance of data and drop the higher layers of scalable video altogether. In this paper, we consider Medium Grain Scalable (MGS) extension of H.264/AVC video and develop new rate and distortion models that characterize the coded bitstream. Performance evaluations show that our proposed framework results in significant gains over existing schemes in terms of average video PSNR that can reach 3dB in some cases for different channel SNRs and different bandwidth budgets.**

EDICS Category: 5-WRLS, 5-STRM

***Index Terms*—Scalable video, multirate wireless networks, rate-distortion modeling, link-adaptation.**

I. Introduction

THE rapid growth of broadband wireless services, along with the advancements in video compression technology, will enable future high quality digital video broadcast and telephony applications. While the promise of high quality video applications seems real, there are still challenges that must be addressed before such applications are efficiently deployed. In particular, the variation in wireless link quality has to be taken into account when designing video streaming systems. The variation in wireless channel quality is usually handled using a link adaptation scheme that maintains a certain level of link reliability under varying channel conditions. The use of link adaptation results in multi-rate operation, where lower physical layer (PHY) transmission rates are used to achieve higher reliability under bad channel conditions. With multi-rate operation, the throughput available to each video stream may vary considerably over time, making it impossible to deliver the video stream without significant over-reservation. To counter this situation, scalable video coding is used, where a single video sequence is encoded into a base layer and several enhancement layers. For continuity of video playback, it is enough to receive the base layer, whereas the enhancement layers are used to increase the quality of the rendered video. When the available throughput decreases, the higher enhancement layers are dropped. This method is the conventional solution for handling capacity variation using scalable video. A traffic control module may be used for this purpose.

To improve the resource allocation, the operation of the traffic control module and the link adaptation scheme can be jointly optimized. In this paper, we offer a new traffic control algorithm as well as an optimization framework that achieve the maximum possible video quality for multiple stations that are using Medium Grain Scalable (MGS) video. MGS is one of the quality scalability features provided by the scalable video coding (SVC) extension of H.264/AVC. The work in this paper differs from the previous works in that it offers a packet drop mechanism tied to an intelligent link adaptation scheme, and it jointly optimizes the operation of these schemes under capacity constraints in a multirate wireless network, for multiple users. We also use a new rate distortion model for MGS video in the optimization framework. Our proposed solution is specifically useful for digital video broadcast services, where a base station (BS) in a wireless domain distributes pre-encoded video content to multiple users in the domain (Fig. 1). The presented algorithms will be deployed in the base station. Intensive packet processing is not required in the proposed method. The algorithm adjusts the parameters of the link adaptation and packet drop schemes dynamically, on a per GOP (Group of Pictures) or per frame basis. When H.264 scalable video coding is used, several sub-streams of encoded units are produced for a single video sequence. Packets of each stream are tagged and treated differently in the PHY (link adaptation) and MAC or Network layers (packet dropping scheme).

For scalable video coding we consider the medium grain scalable extension of the H.264/AVC standard [1], known as SVC [2], [3]. The methods presented in this paper are applicable to all packet based multirate networks such as IEEE 802.11 or IEEE 802.16 [4]–[6]. For demonstration purposes we consider the physical layer of the IEEE 802.11 Wireless

This work was supported under NSERC Strategic Project Grant: STPGP 322075 - 05.

H. Mansour, P. Nasiopoulos, and V. Krishnamurthy are with the Department of Electrical and Computer Engineering at the University of British Columbia, 2356 Main Mall, Vancouver, BC Canada V6T 1Z4. Email: {hassanm, panosn, vikramk}@ece.ubc.ca Phone: 1-604-781-9967. Fax: 1-604-822-9013.

Y. P. Fallah is with the Departments of Electrical Engineering and Computer Sci. and Civil and Environmental Engineering, University of California, Berkeley. Email: yaserpf@berkeley.edu.

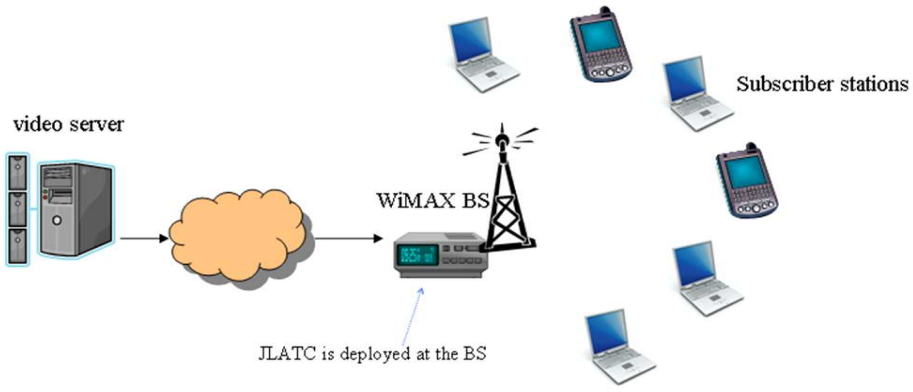

Fig. 1. General diagram of the wireless video streaming system in which a transmitter (base-station or access-point) serves multiple video users/clients. The video streaming service is allocated a fixed share of the bandwidth that the transmitter can support, such that all video users compete for a portion of the allocated bandwidth share.

Local Area Network (WLAN).

TABLE I
GLOSSARY OF ACRONYMS AND VARIABLES

BOF	Batch of Frames
CGS	Coarse Grained Scalability
CLD	Conventional Layer Drop
ILA	Intelligent Link Adaptation
JLATC	Joint Link Adaptation and Traffic Control
MAD	Mean Absolute Difference
MGS	Medium Grained Scalability
MIMO	Multiple Input Multiple Output
QP	Quantization Parameter
PHY	Physical Layer
PSNR	Peak Signal to Noise Ratio
SVC	Scalable Video Coding
l	Scalable video layer index
d_l	Video PSNR after decoding all layers up to layer l
r_l	Video bitrate of layer l
C_l	PHY transmission rate of layer l
τ_l	Temporal share of layer l
η	Video service temporal share

A. Overview of Related Work

Transmission of scalable video over different types of wireless networks has been studied in several recent works [7]–[13]. These works consider different aspects and issues of transmitting SVC over wireless networks. The work in [7] presents an optimal solution for transmission of scalable video traffic over MIMO systems, through optimizing quantization parameters, GOP (Group of Picture) size and channel coding and symbol constellation for a simplified MIMO system. The work presented in [8] attempts to improve the received video quality in a 3G wireless network by optimizing for power and coding rate allocation. Another method relevant to this topic is the work presented in [9], which focuses on employing spatial multiplexing feature of a MIMO system to better deliver scalable video. This method proposes an adaptive channel selection for scalable video transmission over MIMO channels where partial channel information is derived and used to improve the system performance. In another recent work, [10], an adaptive scheme is provided for assigning different FEC codes to packets of each layer and controlling the transmission rate by dropping higher layers of video. The method reported in [11] presents a schemes for transmission of wavelet based scalable video over WLANs. This method is different from the other papers that consider the more general H.264 based scalable video. In another relevant work, [12] proposes a rate distortion optimized method for transmission of scalable video over CDMA networks. Finally, in [13], the authors consider the problem of scheduling and resource allocation of pre-encoded SVC bitstreams for multi-user video downlink over OFDM channels.

Some of the work published after the preparation of this manuscript include [14]–[16]. In [14], a low complexity joint source-channel coding methodology is proposed to minimize the end-to-end distortion of SVC bitstreams transmitted over a lossy channel. Low-density parity check (LDPC) codes are used as channel codes to provide unequal error protection to the scalable bitstream. The topics of distortion estimation and bandwidth allocation for the transmission of SVC bitstreams over MIMO systems are also discussed in [15], [16].

Our proposed solution differs from existing works in that it offers a combined link adaptation and MAC/Network layer packet drop scheme that maximizes the collective quality of the delivered video to multiple users. We present an optimization framework that adjusts the operation of the link adaptation algorithm in the PHY, as well as the network traffic control (packet drop) module in the MAC or network layers, under bandwidth (temporal fairness) constraints. Our method targets multirate wireless networks such as 802.11 and 802.16, and uses H.264 MGS scalable video. Our proposed solution is specifically useful for digital video broadcast services, where a base station in a wireless domain distributes pre-encoded video content to multiple users. The algorithm will be deployed in the base station and does not require intensive packet processing or application layer involvement. We list the main contributions of our work in the following subsection.

B. Main Contributions

The main contributions of this paper can be listed as follows:

1) We show that the joint optimization of traffic control and link adaptation processes considerably improves the received scalable video quality of all video users and enhances the channel bandwidth utilization in a multi-user wireless video streaming system.
2) We formalize the resource allocation problem statement in Section II showing the differences in optimization objectives between existing solutions and our proposed solution. Our proposed technique solves for the PHY modulation $C_l^{(u)}$ and truncation percentage $x_l^{(u)}$, for all video layers l of users u, that maximize the expected sum of video PSNRs D^u while satisfying the service temporal share constraint η. Moreover, the joint optimization determines the optimal time-share among the multiple users that maximize the defined objective.
3) We develop new analytical rate and distortion models in Section III that characterize the real-time behavior of MGS scalable H.264/AVC streams. To our knowledge, these are the first rate and distortion models proposed for MGS coded video streams.
4) We show in Section IV that the joint link-adaptation and traffic-control optimization problem is a non-convex constrained optimization problem with combined discrete and continuous variables. Moreover, we propose an algorithm that first solves a continuous relaxation of the discrete optimization to find the optimal PHY rate points, and next solves for the packet drop ratios using the discrete PHY rate points that lie in the vicinity of the optimal solution of the continuously relaxed problem.

The performance of the proposed joint resource allocation scheme is analyzed in Section V, and finally we present our conclusions in Section VI.

II. General Overview of the Resource Allocation Problem

In this section, we describe the different components of the video streaming system, formalize the problem at hand and discuss the existing solutions and their limitations. We consider a wireless video streaming system in which a transmitter (base-station or access-point) serves multiple video users/clients. The video streaming service is allocated a fixed share of the bandwidth that the transmitter can support, such that all video users compete for a portion of the allocated bandwidth share. Digital video broadcast services are an example of this operation scenario. Fig. 1 illustrates an overview of the transmission system described above.

A. System Overview

Broadband wireless networks are known to be susceptible to variations in wireless link quality. To counter the variation in link quality, a user-defined link adaptation scheme is used in wireless networks such as IEEE 802.11 and IEEE 802.16 networks. This mechanism is also known as rate adaptation in other standards (e.g., 3GPP). The link adaptation scheme adapts transmission parameters to dynamic channel conditions, based on some measured parameters. These parameters are generally measured at the receiver and are delivered to the transmitter through a feedback mechanism. The transmitter uses this information to determine a specific modulation and coding scheme (MCS) for the next packet to be transmitted. Under good channel conditions, more information bits and spectrally efficient modulation schemes can be used; whereas, under bad channel conditions, link adaptation adds resilience to the transmitted signal, reducing the number of delivered information bits per symbol.

Several transmitter parameters can be modified by the link adaptive scheme. The Forward Error Correction (FEC) coding rate and modulation type are the most common. In recent years, several other parameters have also been incorporated into link adaptation, such as transmitter power, cyclic prefix length and switching between space time block coding (STBC), spatial multiplexing (SM) and beam-forming in multiple input multiple output (MIMO) systems.

Link adaptation adds an additional constraint to networks that have to deal with limited resources. In communications scenarios where the available capacity is limited, the nodes connected to a link with limited capacity have to restrict the amount of traffic that passes through that link. We call the module that performs this task "traffic control module". This module works in conjunction with the scheduling and admission control schemes which ensure that resources are properly assigned to flows. With link adaptive wireless networks, the task of traffic control becomes more complicated. In this article we assume the existence of the scheduling and admission control modules and focus on the design of the traffic control module instead. The details of how scheduling is done in multi-rate networks are given in [17]. For the wireless access networks considered in this article, the traffic control module is deployed in the BS of the network. The BS is usually connected to the internet or other networks through higher capacity or wired links; nevertheless, traffic control can be performed on any outgoing link (to stations or to the backbone).

B. Problem Description

Link adaptation and dynamic variation of capacity in wireless networks necessitates dynamic traffic control, which in turn requires the applications to be tolerant of variable throughput. For video applications, scalable video will be used to adapt the video bit-rate to the variable available throughput. With SVC, the simplest solution to adapt to decreased throughput is to drop the higher enhancement layers.

To formulate the problem of dynamic traffic control we consider the scenario described earlier, where a fixed portion of the available bandwidth is dedicated to the video service and is shared by multiple stations. Let $u \in \{1, 2, \cdots, U\}$ be the video user index and let $l \in \{0, 1, ...L\}$ be the scalable video layer index, such that, layer 0 corresponds to the base layer and layer L the highest (lowest priority) enhancement layer. Let r_l^u be the estimated video bit-rate required by layer l of user u, and let C_l^u be the PHY transmission rate allocated

for layer l of user u. We define the temporal share τ_l^u occupied by video layer l of user u as follows:

$$\tau_l^u = \frac{r_l^u}{C_l^u}, \quad \tau_l^u \in [0,1]. \tag{1}$$

In this paper, we assume that the scheduling algorithm is able to assign a fair share of the bandwidth, or a fair time share, to each user (video stream) or an aggregate of users [17]. Moreover, the scheduler handles MAC layer overhead such as retransmissions or Acknowledgements while allocating such service shares. The fair scheduler provides the total time share $\eta \in [0,1]$ for all video users (aggregate of their traffic) that share the streaming service. Within the aggregate service the time shares are distributed according to the method presented in this paper. The aim here is to maximize the total quality of the delivered video under the time share constraint for the entire video service. Here we propose an optimization framework that achieves this goal. The proposed method provides significant improvement over existing mechanisms, which are described in the next section.

C. Overview of Existing Resource Allocation Solutions

The 802.11 or 802.16 standards do not mandate a specific link adaptation scheme. The most common mechanism of selecting between the different transmission rates is to use SNR thresholds for the different transmission rates that meet a maximum error constraint, usually 10% PER for packet length of 1000 bytes [4]. The highest transmission rate that meets the maximum PER requirement is selected. To adapt the scalable video traffic to rate changes, different schemes may be used, as explained below.

1) Conventional Layer Drop: The conventional and simple solution for adapting the bit-rate of the scalable video to the available throughput is to drop the higher layers of video, until the remaining traffic fits in the available throughput for the stream. We call this method "conventional layer drop" (CLD), in this paper. The CLD method can be formulated as follows:

$$\max_{C^u, \hat{L}^u} \quad \hat{L}^u \quad \text{s.t.} \quad \begin{array}{l} \sum_{l=0}^{\hat{L}^u} \tau_l^u(C_l^u) \leq \eta^u, \\ \text{and } p_l(C_l^u) \leq 10\% \end{array} \tag{2}$$

where $\hat{L}^u \in \{0,1,...L\}$ is the highest admissible video layer of user u, L is the maximum number of available video layers, p_l is the packet error probability associated with the selected transmission rate C_l^u for video layer l of user u, $C^u = \{C_1^u, \cdots, C_{\hat{L}^u}^u\}$ is a set containing the PHY transmission rates for all the video layers of user u, and η^u is the temporal share of user u.

With CLD, the transmission rates of all video layers belonging to each user will be the same. The transmission rates are set by the link adaptation scheme regardless of the type of traffic. The objective of CLD is to fit as many layers as possible in the available bandwidth or time share for the flow.

2) Intelligent Link Adaptation: A more intelligent method is to replace the conventional link adaptation, and layer drop schemes, with a combined link adaptation and layer drop scheme that assigns different PHY rates to each layer and drops the enhancement layers that do not fit in the available throughput for the stream. This method is called Intelligent Link Adaptation (ILA) and has been studied in detail in [18]. ILA is based on per-stream traffic control, and aims at improving the video quality of users individually, and on a long term average basis. We formulate the ILA scheme as shown below:

$$\min_{\hat{L}^u, C^u} \bar{D}^u(\hat{L}^u, C^u), \text{ s.t. } \sum_{l=0}^{\hat{L}^u} \tau_l^u(C_l^u) \leq \eta^u, \tag{3}$$

where $\bar{D}^u(\hat{L}, C^u)$ is the time-averaged video distortion model given as a function of the maximum number of admissible video layers $\hat{L}^u$ and the PHY transmission rates $C^u = \{C_1^u, \cdots, C_{\hat{L}^u}^u\}$. Unlike CLD, the ILA scheme can assign different PHY rates C_l^u to different video layers of the same user u.

Note that given the channel SNR, there exists a one-to-one mapping that associates the PHY rate C_l^u with a PER P_l^u. Therefore, the distortion model $\bar{D}^u(\hat{L}^u, C^u)$ is in effect a loss-distortion model estimating the expected value of the video distortion given the PER $P^u = \{P_1^u, \cdots, P_{\hat{L}^u}^u\}$. The formulation in [18] assumes a static time-averaged video distortion model that is derived for pre-encoded content. However, to ensure a fair comparison with the dynamic scheme proposed in this paper, we have used the real-time rate and distortion models described in Section III for both schemes in our performance evaluations.

The solution to the ILA problem consists of a combinatorial search over all achievable PHY rate points (and corresponding PER points) to choose the set that minimizes the video distortion function $\bar{D}^u$. Since ILA considers video streams individually, it can be used at any source that transmits the video, i.e., the base station or the subscriber stations.

D. Brief overview of the proposed solution

In a multi-user system, combining the resources assigned to all users will allow dynamic allocation of these resources on a short term basis, which results in more efficient use of the limited capacity and higher overall video quality. Our proposed technique solves for the PHY modulation $C_l^{(u)}$ and truncation percentage $x_l^{(u)}$, for all video layers l of users u, that maximize the expected sum of video PSNRs $\mathcal{S}^u$ while satisfying the service temporal share constraint η. Moreover, the joint optimization determines the optimal time-share among the multiple users that maximize the defined objective. A detailed description of this technique is presented in section IV.

The rate-distortion (R-D) model used in the ILA scheme is based on the model proposed in [19] and extended in [18]. This model has its shortcomings in that it cannot be used to predict the R-D behavior of the next video frame or batch of frames (BOF). Therefore, these models can only be used when long term average quality is concerned. In the next section, we present new real-time rate and distortion models for MGS scalable coded video streams which can be used in online or batch traffic management and resource allocation algorithms on short term (BOF size) resource allocation decision making.

III. Rate-Distortion Modeling of MGS/CGS Scalable Video

In this section, we present new rate and distortion estimation models for the representation of MGS/CGS scalable coded video content. Rate and distortion modeling has become a keystone in model-based video rate and transmission control algorithms. The models are used to predict the average coded video packet size and distortion of the corresponding decoded picture. Therefore, two encoding parameters are commonly used in rate-distortion (R-D) estimation models, namely the quantization parameter (QP) and the mean-absolute-difference (MAD) of the residual signal. The latter parameter is also considered to be a measure of the prediction error, since it quantifies the mismatch between the original un-coded picture and the Intra/Inter prediction.

The rate control algorithm used in H.264/AVC allows for the estimation of the prediction MAD $\tilde{m}$ of a frame using the following first order linear predictor model:

$$\tilde{m}^{(k)} = \alpha m^{(k-1)} + \beta, \tag{4}$$

where k denotes the frame number, $m^{(k-1)}$ is the actual MAD of the previous frame, and α and β are model parameters that are initialized to 1 and 0, respectively. These parameters are updated by a linear regression method similar to that of MPEG-4 Q2 after coding each picture [20].

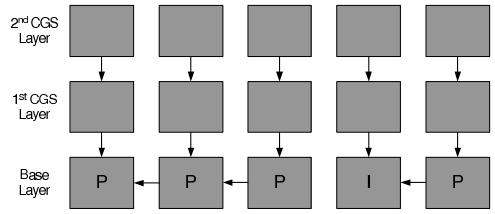

Fig. 2. Example of the scalable video coding structure used in this paper. The loop is closed at the base layer only and no motion refinements are encoded in the quality enhancement layers.

A. Overview of Medium/Coarse-Grain Scalability in SVC

In SVC, quality scalability is achieved using medium-grained or coarse-grained scalability (MGS/CGS) where scalable enhancement packets deliver quality refinements to a preceding layer representation by re-quantizing the residual signal using a smaller quantization step size and encoding only the quantization refinements in the enhancement layer packets [21]. Moreover, MGS/CGS enhancement layer packets are coded using the *key-picture concept* where the coding loop is closed at the highest scalability layer for key-pictures only, while the loop is closed at the base-layer quality for the remaining pictures. This approach achieves a tradeoff between the coding efficiency of enhancement layers and the drift at the decoder [21]. These new features in SVC result in a divergence from existing rate-distortion models, calling for the development of improved models that can accurately capture the runtime rate-distortion behavior. In this work, we consider a coding structure in which the coding loop is closed at the base-layer only. Although this structure results in a loss in coding efficiency at the enhancement layer, it helps to improve the error resilience of the video bitstream. Fig. 2 illustrates the coding structure adopted in this paper.

The model described in (4) is sufficient when both the current and previous frames have the same QP value. However, the QP values of real-time encoded MGS/CGS video frames change in both the base and enhancement layers. Let $q^{(k)}, q^{(k-1)} \in \{0, 1, ...51\}$ be the current and previous QP values, respectively. We have found a simple relationship that estimates $\tilde{m}(q^{(k)})$ given the residual MAD at an initial $q^{(k-1)}$ as shown below:

$$\tilde{m}(q^{(k)}) = \tilde{m}(q^{(k-1)}) 2^{a(q^{(k)} - q^{(k-1)})} \tag{5}$$

where a is a sequence dependent constant that is derived empirically and valued around 0.07 for most video sequences.

B. Proposed scalable distortion model

We present an improved distortion model in this section that estimates the decoded picture luminance peak signal to noise ratio (PSNR). Let q_0 and q_1 be the base and first enhancement layer QP values, respectively, and let $\tilde{m}_0 = \tilde{m}(q_0)$ and $\tilde{m}_1 = \tilde{m}(q_1)$ be the respective prediction MAD estimates. Note that the sub-indices for $\tilde{m}$ and q, from here on, refer to the layer number.

For an H.264/AVC coded video stream and H.264/AVC compliant base-layer SVC stream, we propose the following PSNR expression:

$$\tilde{d}_0(q_0) = d_0(\tilde{m}, q_0) = b_1 \log_{10}\left((\tilde{m}_0)^s + 1\right) \cdot q_0 + b_2, \tag{6}$$

where s is a value that depends on the frame type, b_1 and b_2 are sequence-dependent model parameters. The parameters b_1 and b_2 can be refined for each sequence during encoding. The value of s depends on the frame type, such that, $s = 1$ for Inter-coded frames and $s \approx \frac{5}{6}$ for Intra-coded frames.

For MGS/CGS scalability in SVC, the term $\tilde{m}(q_0)$ in (6) is simply replaced by the estimate of the quality refined prediction error $\tilde{m}(q_1)$, thus resulting in the following MGS/CGS PSNR model:

$$\tilde{d}_1(q_1) = d_1(\tilde{m}, q_1) = b_1 \log_{10}\left((\tilde{m}_1)^s + 1\right) \cdot q_1 + b_2, \tag{7}$$

where, $\tilde{m}_1 = \tilde{m}_0 2^{a(q_1 - q_0)}$. Note that $\tilde{d}_1(q_1)$ corresponds to the video quality (PSNR) achieved after decoding both layers 0 and 1.

To demonstrate the accuracy of our proposed PSNR model we show the performance for the Foreman and Mobile sequences in Figs. 3 (a) and (b).

C. Proposed scalable rate model

In this section, we present the rate model for base and MGS enhancement layer SVC coded video frames. The presented model is based on the work proposed in [22] which we extend to incorporate MGS/CGS scalability. The bit-rate of the H.264/AVC compatible base layer in SVC can be expressed as follows:

$$\begin{aligned} \tilde{r}_0(Q_0) = r_0(\tilde{m}, Q_0) &= c_1 \tilde{m}_0^s / Q_0, \\ &= c_2 \tilde{m}_0^s 2^{-q_0/6} \end{aligned} \tag{8}$$

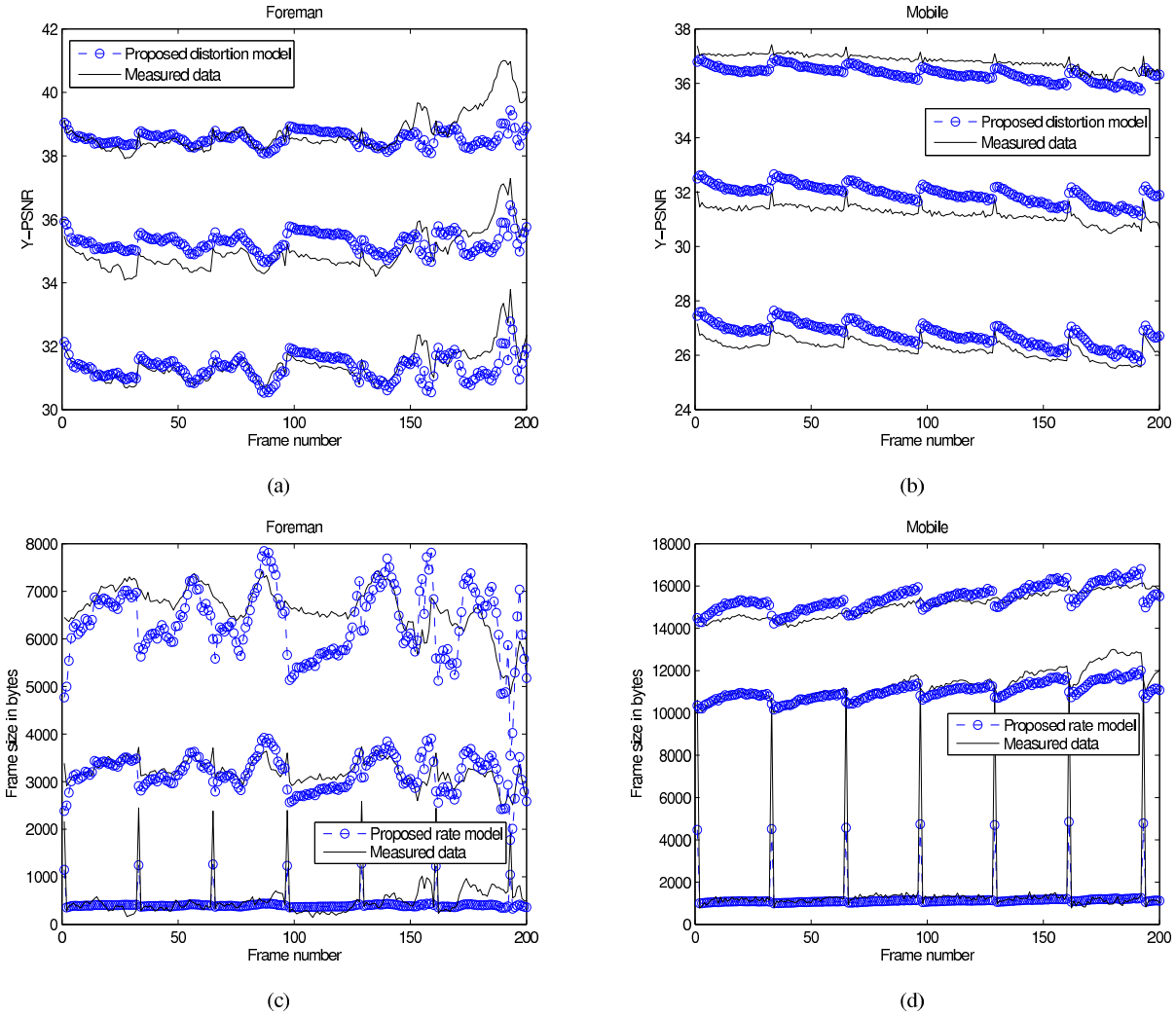

Fig. 3. Illustration of the accuracy of the modeled PSNR estimates (a) and (b), and the modeled rate estimates (c) and (d) for the base layer and two MGS enhancement layers of the sequences Foreman and Mobile.

where c_1 is a model parameter, Q_0 is the quantization step size at the base layer, and s is the same power factor described in (6). Note that instead of the lookup table conversion from quantization parameter to quantization step described in the H.264/AVC standard, we approximate this conversion as follows:

$$Q = 0.625 \cdot 2^{q/6}.$$

We extend the model in (8) to incorporate MGS/CGS packets which only contain refinements on the quantization of residual texture information [21]. Therefore, we express the enhancement layer bit-rate as follows:

$$\tilde{r}_1(q_1) = r_1(\tilde{m}, q_1) = c_3 \tilde{m}_1^s 2^{-q_1/6}, \tag{9}$$

where c_3 is a model parameter. The term $\tilde{r}_1(q_1)$ corresponds to the number of bits required to encode layer 1. Therefore, the total bitrate is given by $\tilde{r}_0(q_0) + \tilde{r}_1(q_1)$.

Figs. 3 (c) and (d) demonstrate the accuracy of our proposed rate estimation model for the sequences Foreman and Mobile.

The sequence dependent constants $b_1, b_2, c_2,$ and c_3 are estimated only once for each data stream. The only parameter update performed is that of the prediction MAD $\tilde{m}$. We adopt the MAD, rate, and PSNR estimation models and parameter update methods described above to characterize the average MAD, rate, and PSNR of a batch of frames (BOF) instead of individual video frames. This approach allows us to run our resource allocation algorithm for batches of G frames, where G is the number of frames in a BOF. *Note that the defined BOF simply refers to a fixed number of coded video frames and therefore imposes no restrictions on the coding structure (e.g. hierarchical coding, IDR period) of the different video users.* In the remainder of this paper, the time unit will correspond to that of a single BOF.

IV. Proposed Resource Allocation Scheme

In this section, we present a new link-adaptation scheme for real-time resource allocation of multiple MGS scalable video streams. The objective of our optimization framework is to maximize the total quality of multiple video streams that share

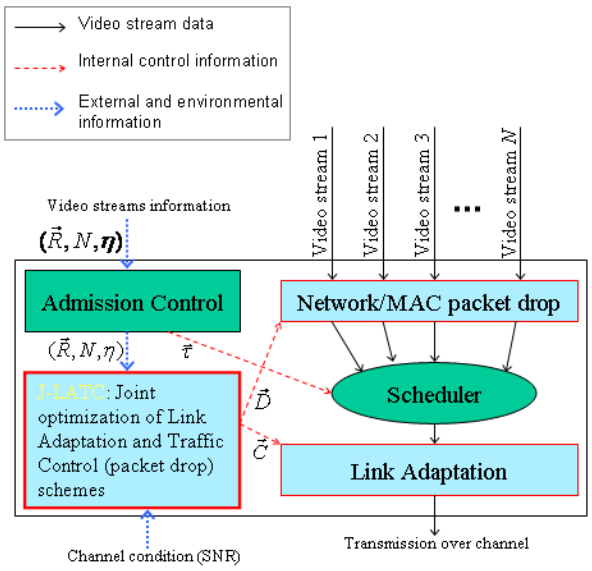

Fig. 4. The Joint operation of the Link-Adaptation and Traffic-Control modules in our proposed framework. The algorithm combines the resources $(\vec{R}, N, \vec{\tau})$ assigned to all users allowing dynamic allocation of these resources on a short term basis. This results in more efficient use of the limited capacity η and higher overall received video quality $\vec{\mathcal{S}}$ of all transmitted streams that share the same wireless channel.

the same wireless channel, while maintaining the total load on the network constant, by adjusting the PHY modulation and video layer truncation percentages.

A. Joint Link Adaptation and MAC/Network Traffic Control

We propose a joint link adaptation and traffic control scheme which strives to optimize the use of the assigned time share (capacity) by achieving maximum video quality given the channel conditions. The optimization framework specifies the PHY transmission rate of each video layer (for the BOF duration), which is then enforced by the link adaptation module. The optimization framework also specifies the percentage of the packets that should be dropped from each layer of each video stream, this decision is then enforced by the Network or MAC layer traffic control module which selectively blocks or drops some packets from each video layer. Fig. 4 illustrates the interaction between the link adaptation and traffic control modules.

We consider a wireless system where U scalable video (MGS) users/clients are served over a single wireless channel with a streaming service constraint η, where $\eta \leq 1$. For simplicity of illustration and without loss of generality, we assume that each video stream u is composed of a base-layer and $L^u = 2$ CGS enhancement layers. The scalable video streams are characterized in terms of their rate-PSNR parameters: $\{\tilde{d}_0^u, \tilde{d}_1^u, \tilde{d}_2^u\}$ and $\{\tilde{r}_0^u, \tilde{r}_1^u, \tilde{r}_2^u\}$ given by equations (6) – (9) above. The rate and PSNR parameters are calculated for fixed values of base and enhancement layer QPs. Note that a PSNR term $\tilde{d}_l^u$ corresponds to the video PSNR achieved by decoding layers 0 to l, whereas a rate term $\tilde{r}_l^u$ corresponds to the number of bits required to encode layer l alone. Moreover, if a lower layer representation of a frame is lost, then the higher layers are discarded.

In a multirate wireless physical layer considered here, different PHY rates are achieved by changing the modulation and coding schemes, resulting in different packet loss ratios under a given SNR γ. We denote such loss ratio as $p_l = f(C_l, \gamma)$, for each video layer assigned a PHY rate C_l. The relationship between PHY rate and PER in a MIMO PHY is not straightforward. Fortunately, we can bypass an exact analysis and construct a one-to-one mapping between the values of C^u and P^u from the bit error rate (BER) performance curves shown in Fig. 5. These curves are obtained through system simulations under typical channel conditions similar to the work done in [18]. BER performance curves for different conditions and networks can be used too, and the formulation of the problem does not depend on these curves, as long as there exists a one to one mapping of C^u and P^u values. The relationship between C_l and p_l can be seen in Fig. 7.

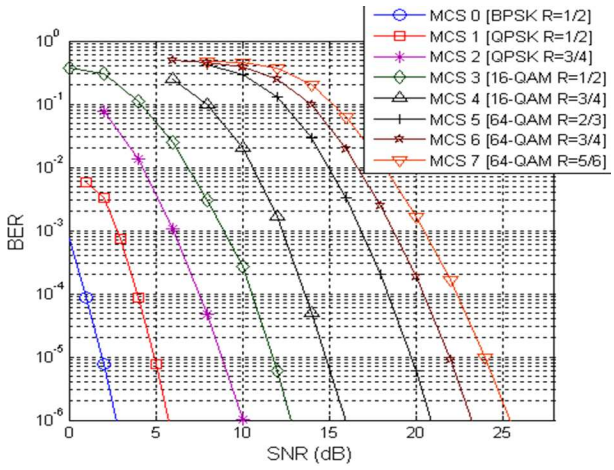

Fig. 5. BER performance of STBC modes of an 802.11n MIMO PHY in quasi-static fading channel, perfect channel estimation assumed [18].

Let t and G be the BOF index and BOF size, respectively. We define $\tilde{\mathcal{S}}_t^u$ to be the predicted average PSNR for BOF t of video stream u, such that

$$\begin{aligned}\tilde{\mathcal{S}}_t(P) = \; & p_0 d_{EC} + (1-p_0)\tilde{d}_0 \\ & +(1-p_1)(1-p_0)(\tilde{d}_1 - \tilde{d}_0) \\ & +(1-p_2)(1-p_1)(1-p_0)(\tilde{d}_2 - \tilde{d}_1)\end{aligned} \tag{10}$$

where d_{EC} is the PSNR due to the error concealment mismatch, and $P = f(C, \gamma) = [p_0 p_1 p_2]^T$ is the packet error rate (PER) vector. Note that we omit the superscript u from the expression above for clarity of illustration.

The predicted temporal time share of a BOF t of stream u, denoted as $\tilde{\tau}_t^u$, can be expressed as follows:

$$\tilde{\tau}_t(C) = \frac{\tilde{r}_0}{C_0} + \frac{\tilde{r}_1}{C_1} + \frac{\tilde{r}_2}{C_2}. \tag{11}$$

We propose to employ a packet dropping/stopping stage in the proposed scheme that blocks the transmission of some packets of each layer, effectively reducing the bit-rate of the layer. The use of MGS coded video bitstreams results in a

linear rate-PSNR relationship within a single CGS layer as illustrated in Fig. 6 . Therefore, a specific loss percentage in video bitrate results in a proportional loss in video PSNR within the same CGS layer.

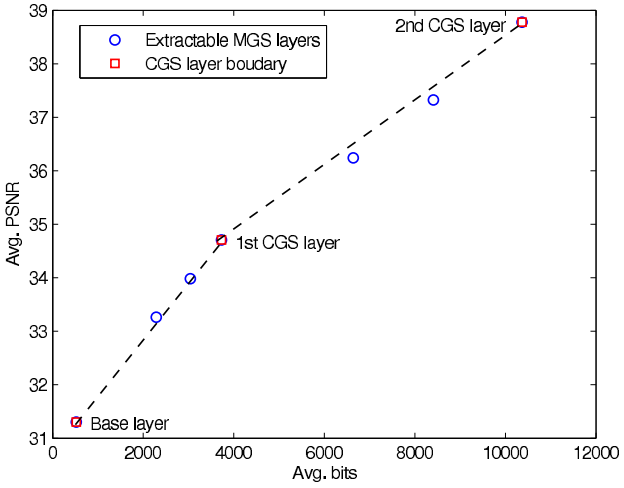

Fig. 6. Example of the operational R-D points achievable with MGS scalability. Foreman sequence is encoded at base-layer QP = 38, and two CGS enhancement layers at QP = 32 and 26. Each CGS layer is divided into three MGS layers. The rate-PSNR behavior is linear within a single CGS layer.

Let x_l be the percentage of blocked packets of layer l, resulting in an effective video layer bit-rate $\tilde{R}_l = \tilde{r}_l(1-x_l)$. Therefore, we can rewrite the video stream's predicted PSNR and corresponding time share as shown below:

$$\begin{aligned}\tilde{\mathcal{S}}_t(X,C) &= p_0 d_{EC} + (1-p_0)\tilde{d}_0 \\ &\quad +(1-x_1)(1-p_1)(1-p_0)(\tilde{d}_1-\tilde{d}_0) \\ &\quad +(1-x_2)\textstyle\prod_{l=0}^{2}(1-p_l)(\tilde{d}_2-\tilde{d}_1) \\ \tilde{\tau}_t(X,C) &= \tfrac{\tilde{r}_0}{C_0} + (1-x_1)\tfrac{\tilde{r}_1}{C_1} + (1-x_2)\tfrac{\tilde{r}_2}{C_2} \\ & \\ & 0 \le x_1 \le x_2 \le 1,\end{aligned} \tag{12}$$

where l is the video layer index, and C_l is the PHY transmission rate of layer l. The video layer dependency is accounted for in the above formulation by setting $x_1 \le x_2$, such that, if x_1 of the packets are dropped from layer 1 then their dependent packets are also dropped from layer 2. Notice that we do not assign a drop rate for the base layer to avoid the costly quality degradation associated with error concealment mismatch and to satisfy quality of service (QoS) guarantees.

The new multi-user resource allocation problem can now be formulated as the following constrained optimization problem with variables $\mathbf{X}$ and $\mathbf{C}$:

$$\begin{aligned}\max_{\mathbf{X},\mathbf{C}} \quad & \textstyle\sum_{u=1}^{U} \tilde{\mathcal{S}}_t^u(X^u,C^u) \\ \text{subject to} \quad & \textstyle\sum_{u=1}^{U} \tilde{\tau}_t^u(X^u,C^u) \le \eta_t, \\ & C_0^u \le C_1^u \le C_2^u, \\ & 0 \le x_1^u \le x_2^u \le 1 \text{ for all } u,\end{aligned} \tag{13}$$

where $\mathbf{X} = [X^1 X^2 ... X^U]$, $X^u = [x_1^u x_2^u]^T$, $\mathbf{C} = [C^1 C^2 ... C^U]$, and $C^u = [C_0^u C_1^u C_2^u]^T$ defined in (12), $u \in \{1,2,...U\}$ is the user index.

The constrained optimization problem defined in (13) is a mixed integer non-linear programming problem whose solution is non-trivial. The variables $X^u \in [0,1]$ are continuous, while $C^u \in \mathcal{C}$ are discrete and $\mathcal{C} = \{\mathcal{C}_0, \mathcal{C}_1, ...\mathcal{C}_M\}$ is the set of feasible PHY transmission rates associated with the current channel SNR, and M is the cardinality of $\mathcal{C}$. In the next subsection, we will discuss our proposed solution to the problem in Eq. (13).

B. Solution of the Global Optimization Problem

Mixed integer programming problems are typically NP-hard. Therefore, we develop an algorithm in this section to solve the above mentioned problem (defined in (13)) by first solving the continuously-relaxed problem illustrated below, then projecting the solution of the continuous relaxation on the set of achievable PHY rate points and solving for the layer drop rates X^u that satisfy the temporal share constraint η.

1) Continuous relaxation of the optimization problem: The continuous relaxation of the discrete constrained optimization problem of equation (13) can be performed by replacing the discrete variable C^u with a continuous valued variable $W^u = [w_0^u w_1^u w_2^u]^T \in \mathcal{W}^3$, where $\mathcal{W} = [\mathcal{C}_0, ...\mathcal{C}_M]$, and finding the continuous approximation of the function $P(C^u, \gamma^u)$, where γ^u is the channel SNR of user u.

We re-write the continuously relaxed constraint and objective functions as follows:

$$\begin{aligned}\tilde{\mathcal{S}}_t(W) &= \rho(w_0)d_{EC} + (1-\rho(w_0))\tilde{d}_0 \\ &\quad +(1-\rho(w_1))(1-\rho(w_0))(\tilde{d}_1-\tilde{d}_0) \\ &\quad +\textstyle\prod_{l=0}^{2}(1-\rho(w_l))(\tilde{d}_2-\tilde{d}_1) \\ \text{and} & \\ \tilde{\tau}_t(W) &= \tfrac{\tilde{r}_0}{w_0} + \tfrac{\tilde{r}_1}{w_1} + \tfrac{\tilde{r}_2}{w_2},\end{aligned} \tag{14}$$

where $\rho(W^u, \gamma^u)$ is the continuous approximation of the PER function $P(C^u, \gamma^u)$.

2) Normal approximation of the PER curves: The PER vs PHY transmission rate relationship is given by a discrete set of achievable points that are a direct consequence of the channel SNR and the employed modulation and coding scheme. We have found that this relationship can be modeled using the continuous complementary CDF of the normal distribution $N(np_c, np_c(1-p_c))$ shown below:

$$\rho(n, C, p_c) = \frac{1}{2}\left[1 - \text{erf}\left(\frac{n - C - \mu}{\sqrt{2\sigma^2}}\right)\right], \tag{15}$$

where C is the PHY transmission rate, $\mu = np_c$ and $\sigma^2 = np_c(1-p_c)$, erf(.) is the Gauss error function, and n and p_c are model parameters calculated offline.

Table II shows an example of the valid MCS indices for a 2x2 802.11n WLAN channel used in the construction of the mapping between C^u and P^u.

Fig. 7 shows the achievable PER vs PHY rate curves and their continuous approximation using the gaussian complimentary CDF for different channel SNRs. The packet error rate values are calculated for an average packet length of 1000 bytes. It can be seen from the figure that the function ρ of (15)

TABLE II
VALID MCS INDICES FOR 2X2 802.11N WLAN

MCS	Bitrate (Mbps)	Code Rate	SM streams	Modulation
0	6.5	1/2	1	BPSK
1	13	1/2	1	QPSK
2	19.5	3/4	1	QPSK
3	26	1/2	1	16-QAM
4	39	3/4	1	16-QAM
5	52	2/3	1	64-QAM
6	58.5	3/4	1	64-QAM
7	65	5/6	1	64-QAM

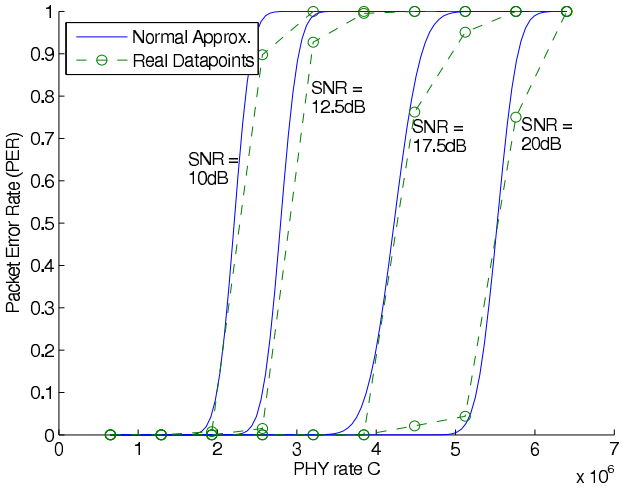

Fig. 7. Example of the achievable packet error rate (PER) curves and their continuous approximation as a function of the PHY transmission rates for different SNRs.

is non-convex over the full range of W. Fig. 8 illustrates the non-convex shape of ρ over the full range of W, which renders the constrained optimization problem non-convex. Since ρ is either 0 or 1 for most of the range of W, and a PER greater than 50% is counter productive, we have narrowed down the feasible range of W to the interval where $0 < \rho \leq 0.5$. This interval corresponds to the following bounds for W:

$$n - \mu - \text{erf}^{-1}(0.99)\sqrt{2\sigma^2} \leq W \leq n - \mu, \qquad (16)$$

where μ and σ^2 are the mean and variance of the approximate normal distribution function shown in equation (15). The new feasible region is now convex and corresponds to the highlighted area in Fig. 8.

3) Joint Link-Adaptation and Traffic-Control (JLATC) Algorithm: For simplicity of presentation, we describe the algorithm for a three layer scenario. The continuously-relaxed constrained optimization problem can now be written as shown below:

$$\begin{aligned} \max_{\mathbf{W}} \quad & \textstyle\sum_{u=1}^{U} \tilde{\mathcal{S}}_t^u(W^u) \\ \text{subject to} \quad & \textstyle\sum_{u=1}^{U} \tilde{\tau}_t^u(W^u) \leq \eta_t, \\ & n - \mu - 2\sqrt{2\sigma^2} \leq w_l^u \leq n - \mu, \\ & \text{and } w_0^u \leq w_1^u \leq w_2^u, \\ & \text{for all } u \in \{1, 2...U\}, l \in \{0, 1, 2\}, \end{aligned} \qquad (17)$$

where $\mathbf{W} = [W^1 W^2 ... W^U], W^u = [w_0^u w_1^u w_2^u]^T$ is the continuous valued PHY rate defined in (14).

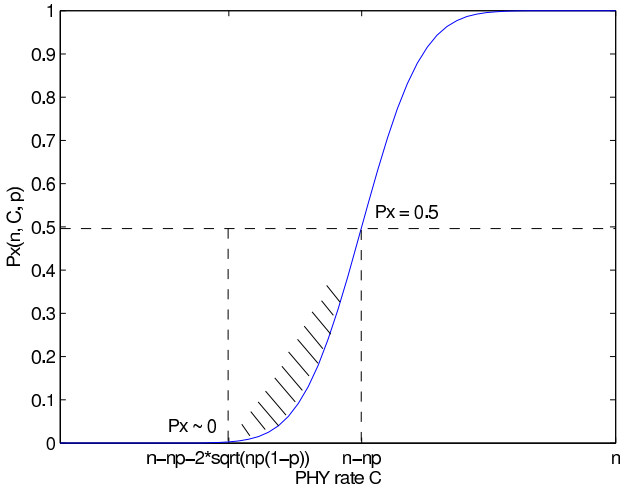

Fig. 8. Plot of the function $\rho(.)$ showing the non-convex shape of the function over the full range of W and the convex region over the reduced feasible interval.

The solution to Eq. (17) is very straightforward, since the problem is convex, and results in the continuous valued PHY rate allocation matrix $\mathbf{W}^*$. The optimal rates $\mathbf{W}^*$, however, are not realizable since the transmitter can only use the predefined set of discrete PHY rates $\mathcal{C}$. Moreover, each optimal rate point w_l^{u*} falls between two achievable PHY rate points in $\mathcal{C}$. Therefore, for each optimal rate vector $W^{u*} = [w_0^{u*} w_1^{u*} w_2^{u*}]^T$ of user u, we propose the approach shown in Algorithm 1.

Algorithm 1 Joint Link-Adaptation and Traffic-Control

1: Solve Eq. (17) to find the continuous valued PHY rates $\mathbf{W}$.
2: Find the rate points C_l' and C_l'', for all $0 \leq l \leq 2$, such that $\{C_l' \leq w_l^{u*} \leq C_l''$, and $C_l', C_l'' \in \mathcal{C}\}$ are the discrete achievable PHY rate points adjacent to w_l^{u*}.
3: Form the set of vectors

$$\hat{C} = \begin{bmatrix} \hat{C}_0 \\ \hat{C}_1 \\ \hat{C}_2 \end{bmatrix},$$

where $\hat{C}_0 = C_0'$, $\hat{C}_1 \in \{C_1', C_1''\}$, and $\hat{C}_2 \in \{C_2', C_2''\}$.
4: For every vector $\hat{C}$, find the layer drop rates $\hat{X}^u$ that solve the following equation:

$$\begin{aligned} \max_{\mathbf{X}} \quad & \textstyle\sum_{u=1}^{U} \tilde{\mathcal{S}}_t^u(\hat{C}, \hat{X}^u) \\ \text{subject to} \quad & \textstyle\sum_{u=1}^{U} \tilde{\tau}_t^u(\hat{C}, \hat{X}^u) \leq \eta_t, \\ \text{and} \quad & 0 \leq x_1^u \leq x_2^u \leq 1 \text{ for } u \in \{1, 2, \ldots, U\}. \end{aligned} \qquad (18)$$

5: Choose the set of mapped PHY rate points C^u and layer drop rates X^u that results in the maximum value for the objective function in (18) as shown below:

$$\{C^u, X^u\} = \arg \max_{\{\hat{C}, \hat{X}\}} \tilde{\mathcal{S}}_t^u(\hat{C}, \hat{X}).$$

Fig. 9 illustrates the mapping process from the optimal rate vector W^* to the achievable PHY rate points. Note that w_0^*

is mapped to the lower rate point C_0' to avoid a possible increase in PER for the base layer, while w_1^* and w_2^* can each be mapped to two possible rate points C_1', C_1'' and C_2', C_2'', respectively.

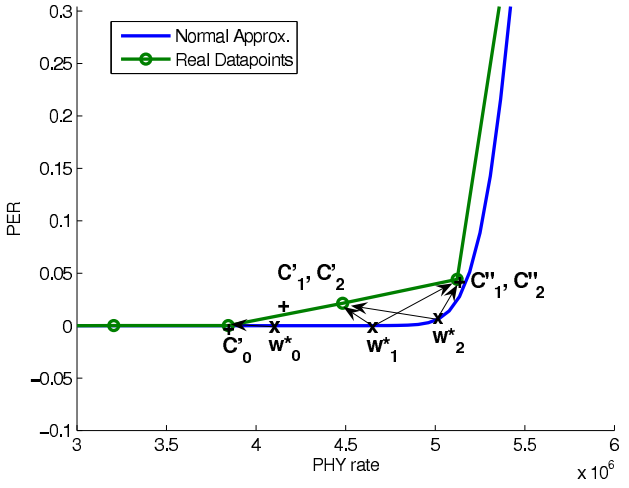

Fig. 9. Example of the mapping process involved in discretizing the continuous valued optimal rate points w_l^* to the achievable discrete PHY rates C_l' or C_l'', where $l \in \{0, 1, 2\}$.

This algorithm, in effect, narrows down the combinatorial search range limiting it to the achievable PHY rate points that are adjacent to the optimal solution given by $\mathbf{W}^*$.

C. Analysis of the JLATC scheme

In this section, we analyze the JLATC algorithm compared to the ILA algorithm in terms of complexity and optimality. In both algorithms, the same video rate and distortion model parameters should be available.

1) Computational cost: In this subsection, we analyze the computational complexity of the JLATC algorithm and compare the cost with that of the ILA algorithm and an exhaustive combinatorial search approach that guarantees optimality.

The JLATC algorithm can be separated into two main computational stages. The first stage constitutes the continuously relaxed $L \times U$ dimensional convex optimization problem which can be solved in polynomial time using interior point methods such as the simplex algorithm. This stage requires a computational cost of $O\left(p(LU)\right)$, where $p(.)$ is a polynomial function. The second stage requires solving a set of L linear programming problems each of which can also be solved in polynomial time with an $O\left((L-1)^2U^2\right)$ operations. Therefore, the total cost of the JLATC algorithm can be approximated by $O\left(p(LU)\right) + L \times O\left((L-1)^2U^2\right)$ operations.

Compared to the JLATC algorithm, an exhaustive combinatorial search algorithm requires solving $|\mathcal{C}|^{LU}$ linear programming problems to reach an optimal solution, where $\mathcal{C}$ is the set of achievable PHY rate points and $|.|$ denotes the cardinality of a set. The computational cost associated with such an approach can be approximated by $|\mathcal{C}|^{LU} \times O\left((L-1)^2U^2\right)$ which far exceeds the cost of the JLATC algorithm.

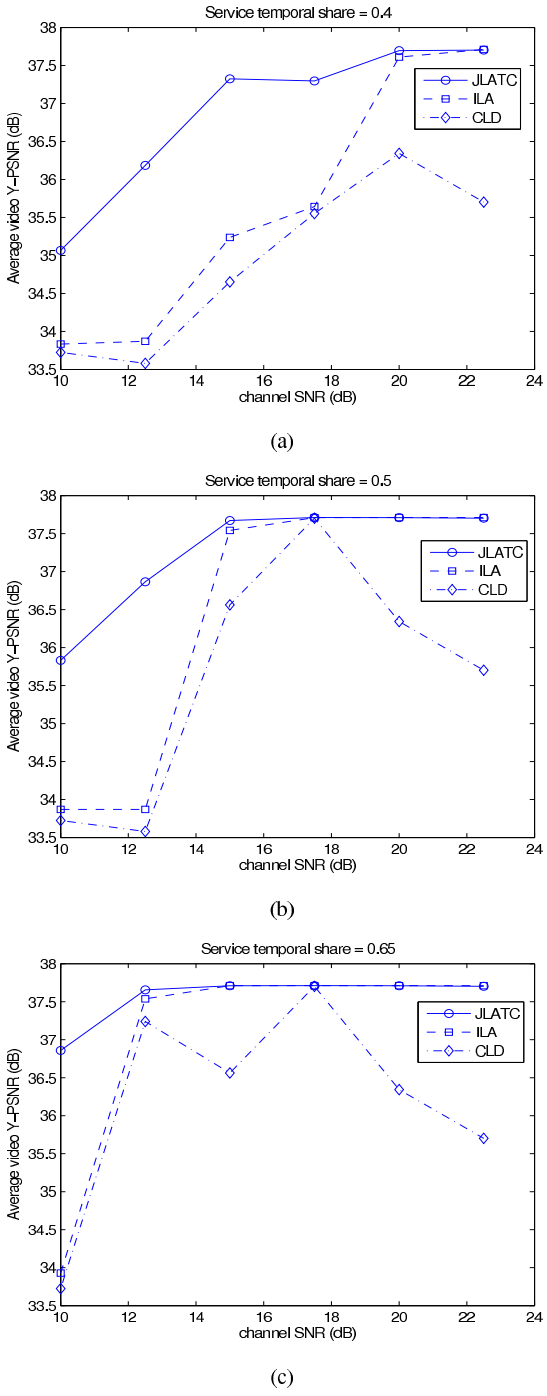

Fig. 10. Comparison in average Y-PSNR performance between the three resource allocation schemes: JLATC (proposed), ILA, and CLD. The performance of each scheme is plotted versus different channel SNRs and the following service temporal share values: (a) $\eta = 0.4$, (b) $\eta = 0.5$, (c) $\eta = 0.65$.

The ILA algorithm defined in [18] is an integer programming problem that solves for $L \times U$ discrete variables $\mathbf{C}$. The JLATC algorithm, on the other hand, solves a mixed integer programming problem with $L \times U$ discrete variables $\mathbf{C}$ and $(L-1) \times U$ continuous variables $\mathbf{X}$. In order to handle the increase in the dimensionality of the problem, the JLATC algorithm solves a continuous and convex relaxation

of the mixed integer programming problem resulting in the continuous valued solution $\mathbf{W}$. Then it finds the discrete variables $\mathbf{C}$ that fall in the vicinity of $\mathbf{W}$.

Up to this point, these steps can be considered as an efficient way of solving the ILA problem. However, the JLATC algorithm also solves L continuous linear programming problems to find the variables X^u for each user u independently. Therefore, the improved performance of the JLATC algorithm over ILA comes at a cost of solving the L linear programming problems.

2) Comments on optimality: The JLATC algorithm employs a series of relaxations that allow us to effectively find a solution to the mixed integer programming problem. Therefore, the possibility of reaching a non-optimal solution can only arise if the solution to the continuous and convex relaxed problem falls beyond two discrete locations away from the global optimal solution. For example, if we examine the locations of discrete PHY rate points shown in Fig. 7 and the continuous solutions w_l^* relative to the discrete points in Fig. 9, it can be seen that the possibility of significant divergence from the global optimal solution is unlikely to occur. However, we emphasize that the JLATC algorithm is still sub-optimal and optimality can only be guaranteed after the very costly combinatorial search approach described in the previous subsection.

V. Performance Evaluation

In this section, we compare the performance of our proposed streaming approach to existing streaming schemes. For our comparisons, we used the JSVM-9-8 reference software [23] to encode 7 reference video sequences in CIF resolution: Foreman, Football, Bus, Mobile, Soccer, Crew, and City. Each sequence is composed of 150 frames encoded in SVC, with an H.264/AVC compatible base layer, and two MGS enhancement layers. The coded video streams are composed of I and P frames only, with an IDR period of 32 frames, a frame rate of 30 fps, and minimum QoS guarantee corresponding to the base-layer quantization parameter value of $q_0 = 38$. The first and second MGS enhancement layers have QP values of 32 and 26, respectively.

The channel BER performance curves were obtained using an elaborate MATLAB model developed at UBC for the 802.11a/n PHY, based on the work in [24] and [25]. For the IEEE 802.11n model Channel Model D of [24] was used. A more accurate doubly exponentially decaying multi-cluster PDP is also used. In [25], a MATLAB implementation of MIMO channels based on [24] is provided. This program is used to generate channel responses which are fed into our model to obtain BER performance curves. Using bit error rate (BER) curves of all valid combinations of modulation and coding, a look up table is constructed. This look up table is used to generate one-to-one mapping between the PHY rate and PER which are approximated using the normal cumulative CDF curve as described earlier in section IV-B2. For simplicity of simulations and experiments we have only included the performance curves for the Space Time Block Coding (STBC) modes of a 2X2 802.11n MIMO system. This does not have any effect on the generality of the optimization problem, and the results are easily extendable. BER curves for all STBC MCS indices for a 2x2 MIMO system can be seen in Fig. 5 and Table II. The values of MCS are 0 to 7, with corresponding bitrates of: $\{6.5, 13, 19.5, 26, 39, 52, 58.5, 64\}$ Mbps.

Furthermore, we assume that video frames are already packetized and no intrusive packet processing is done. This is of particular importance, since most intermediate network devices, such as WiMAX base stations, should avoid deep packet processing. Instead, we use all available options in the MAC and PHY layers (which are closely monitored and controlled by base and subscriber stations) to provide the best possible service.

To demonstrate the advantages of our proposed Joint Link Adaptation and Traffic Control (JLATC) algorithm, we compare its performance with that of the ILA and CLD schemes. Resource allocation using each of the three schemes is performed for every group of eight pictures. Therefore, the ILA model parameters are calculated for every BOF of size $G = 8$ and stored at the transmitter. A major drawback for the ILA scheme is that the model parameters of future BOFs cannot be predicted from the parameters of the current BOF. As a result, the ILA scheme suffers in real-time applications and can only use the initial model estimates for the entire streaming session. Our JLATC, on the other hand, requires only the model parameters of the first BOF which it uses consequently to predict and update the model parameter of future BOFs using the process described in Section III.

For our performance evaluations we use the video luminance peak signal-to-noise ratio (Y-PSNR) as the picture quality metric, and compare the received video PSNR from each resource allocation scheme for different channel SNR values and different bandwidth budgets η for the streaming service. Fig. 10 illustrates the results of the comparison in terms of the average Y-PSNR for all seven streams. The figure shows that for a temporal share $\eta = 0.4$ which corresponds to tight bandwidth constraints, the proposed JLATC scheme significantly outperforms its closest contender, the ILA scheme, by an average of 1.2dB for different channel SNR values. The performance gain peaks at 2.3dB in average Y-PSNR in the case where the channel SNR = 12.5dB. In the case where the channel is not congested the plots show that both the JLATC and ILA schemes approach the maximum achievable video PSNR that is given by the video stream encoding. The effect of the three resource allocation schemes on the performance of each individual stream can also be seen in Fig. 11.

The JLATC performs better than the ILA scheme because it can utilize the available channel bandwidth, given by the temporal share η, more efficiently than the ILA scheme. Fig. 12 demonstrates the improved temporal share utilization of our proposed JLATC scheme over the ILA and CLD schemes. The figure shows the temporal share utilization and corresponding average video PSNR for every BOF of streamed video frames.

If we consider the temporal share utilization and corresponding PSNR performance of the three schemes shown in Fig. 12 (a) and (b), it can be seen that the JLATC scheme results in considerable improvement in PSNR over the ILA scheme for a small improvement in temporal share utilization.

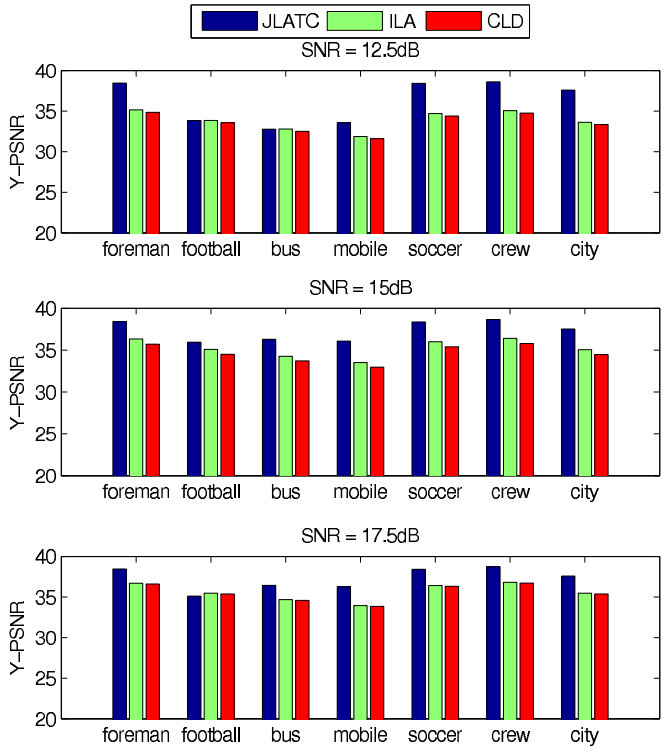

Fig. 11. Individual stream performance in terms of Y-PSNR using the JLATC, ILA, and CLD resource allocation schemes for different channel SNR values and a service temporal share $\eta = 0.4$.

Such behavior arises from the different PHY rate allocation schemes involved in each scheme. The ILA scheme solves the combinatorial optimization problem defined in (3) which provides the admitted video layers with high error protection sacrificing the higher video layers. The JLATC scheme performs the PHY rate allocation by solving the continuously-relaxed problem and then mapping the optimal solution to the achievable PHY rate points. This procedure results in a different PHY rate allocation than the ILA scheme which has clearly shown to produce better video PSNR performance. The reason for such performance is that when the ILA scheme cannot admit a video layer or is forced to allocate a PHY rate with high PER to the video layer, a large portion of the available temporal share is used to provide more error protection for the lower video layers. This problem is avoided by the JLATC scheme by enforcing the constraint given by Eq. (16) which limits the feasible interval of the PHY rate to the convex region of the PER curves.

VI. Conclusion

In this paper, we showed that jointly optimizing the link-adaptation and traffic control of a multi-user wireless video streaming system considerably improves the received video quality for all video users and enhances the channel bandwidth utilization. Multi-rate networks such as IEEE 802.11 or 802.16 use link adaptation mechanisms in the physical layer to adjust the modulation and coding schemes that control the reliability of transmission for varying channel conditions. Conventional schemes to link-adaptation and quality control drop the enhancement layers of scalable coded video streams in order to match the reduced PHY bit-rate required to achieve higher reliability. We propose a new resource allocation scheme that takes advantage of the robust and efficient delivery provided by scalable coded video streams to jointly optimize the link-adaptation scheme in the physical layer (PHY) with a new traffic control module in the network or MAC layer. The traffic control module operates by dropping the necessary percentage of packets from enhancement layers that cannot be fully transmitted, in order to fully utilize the available transmission bandwidth. We show that the resulting joint optimization problem is a non-convex constrained optimization problem with combined discrete and continuous variables. Since, the solution to this problem is non-trivial, we propose an algorithm that first solves a continuous relaxation of the discrete optimization to find the optimal PHY rate points, and next solves for the packet drop ratios using the discrete PHY rate points that lie in the vicinity of the optimal solution of the continuously relaxed problem. Moreover, the algorithm defines operational intervals that maintain the convexity of the continuously relaxed problem. Performance evaluations comparing our proposed scheme with existing solutions show that the proposed approach results in significant gains in terms of average video PSNR that can reach 3dB in some cases for different channel SNRs and different bandwidth budgets.

Acknowledgment

The authors would like to thank the reviewers for their invaluable comments and suggestions. Their input has been extremely appreciated and has helped in improving the conciseness and preciseness of this work.

References

[1] *ITU Rec. H.264/ISO IEC 14996-10 AVC, Advanced Video Coding for Generic Audiovisual Services, , organization = Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, number = version 3, year = 2005*.

[2] J. Reichel, H. Schwarz, and M. Wien, "Joint Scalable Video Model JSVM-9," ISO/IEC JTC 1/SC 29/WG 11, Marrakech, Morocco, Tech. Rep. N 8751, January 2007.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable h.264/mpeg4-avc extension," in *IEEE International Conference on Image Processing,*, Oct. 2006, p. 161164.

[4] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*, ANSI/IEEE Std 802.11: 1999 (E) Part 11, ISO/IEC 8802-11, 1999.

[5] *Amendment 8, Medium Access Control (MAC) Quality of Service (QoS) Enhancements*, ANSI/IEEE Std 802.11e, July 2005.

[6] *IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems.*, ANSI/IEEE Std 802.16-2004, 2004.

[7] M. Jubran, M. Bansal, R. Grover, and L. Kondi, "Optimal Bandwidth Allocation for Scalable H.264 Video Transmission Over MIMO Systems," in *MILCOM*, Oct. 2006, pp. 1–7.

[8] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel-adaptive resource allocation for scalable video transmission over 3G wireless network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 8, p. 1049 1063, Aug. 2004.

[9] D. Song and C. W. Chen, "Scalable H.264/AVC Video Transmission Over MIMO Wireless Systems With Adaptive Channel Selection Based on Partial Channel Information," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, p. 1218 1226, Sept. 2007.

[10] S. Y. Shi, W. C. Wu, and D. J. Du, "A novel unequal loss protection approach for scalable video streaming over wireless networks," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 363 – 368, March 2007.

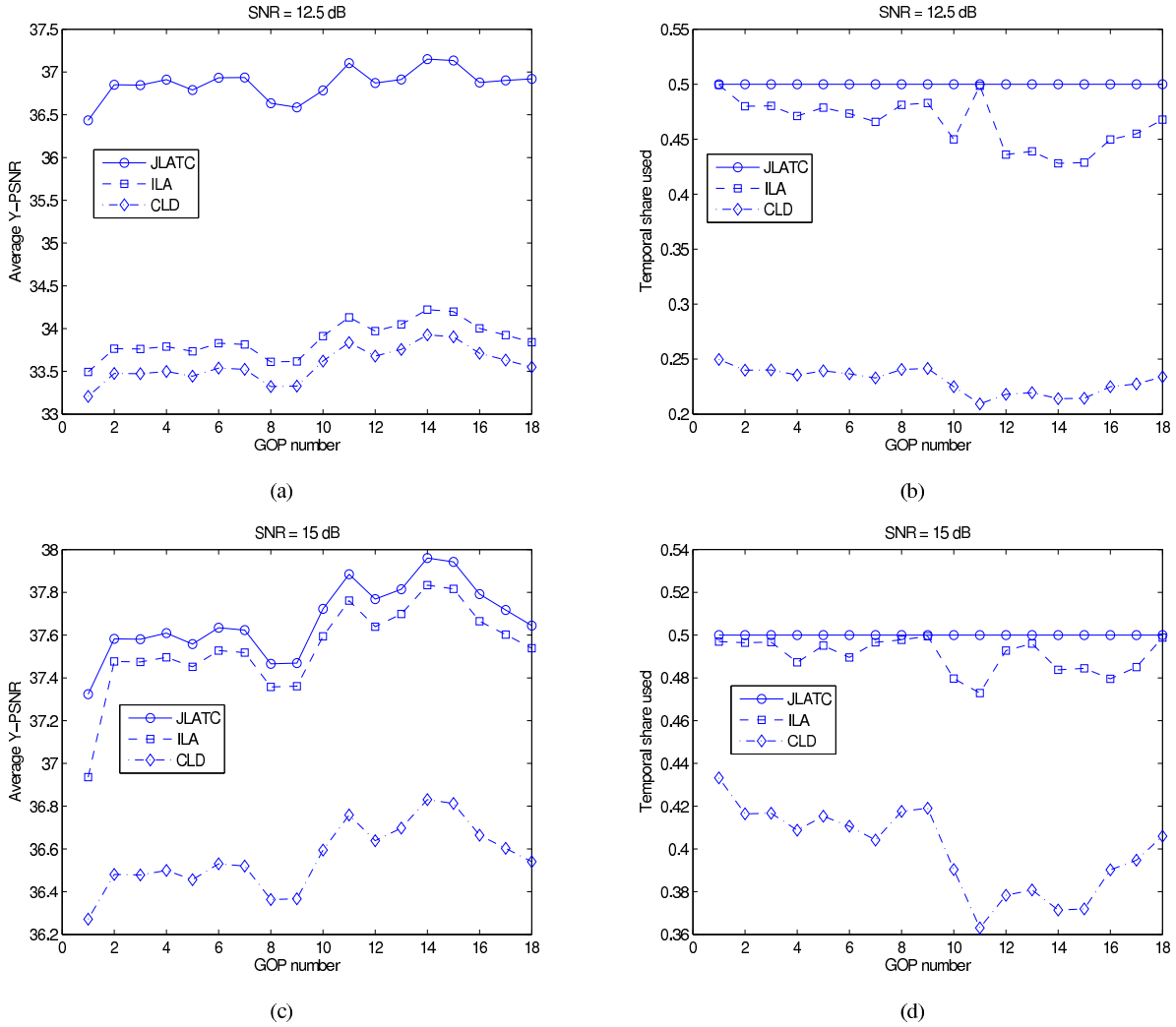

Fig. 12. Comparison of bandwidth utilization and corresponding video Y-PSNR between the three resource allocation schemes: JLATC (proposed), ILA, and CLD. The plots correspond to a streaming temporal share budget $\eta = 0.5$ and channel SNR = 12.5dB for (a) and (b), and SNR = 15dB for (c) and (d).

[11] C. H. Foh, Y. Zhang, Z. Ni, J. Cai, and K. N. Ngan, "Optimized cross-layer design for scalable video transmission over the ieee 802.11e networks," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 17, pp. 1665 – 1678, Dec. 2007.

[12] D. Srinivasan and L. Kondi, "Rate-distortion optimized video transmission over ds-cdma channels with auxiliary vector filter single-user multirate detection," *Wireless Communications, IEEE Transactions on*.

[13] X. Ji, J. Huang, M. Chiang, and F. Catthoor, "Downlink ofdm scheduling and resource allocation for delay constraint svc streaming," May 2008, pp. 2512–2518.

[14] M. Stoufs, A. Munteanu, J. Cornelis, and P. Schelkens, "Scalable joint source-channel coding for the scalable extension of h.264/avc," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 12, pp. 1657–1670, Dec. 2008.

[15] M. Jubran, M. Bansal, L. Kondi, and R. Grover, "Accurate distortion estimation and optimal bandwidth allocation for scalable h.264 video transmission over mimo systems," *Image Processing, IEEE Transactions on*, vol. 18, no. 1, pp. 106–116, Jan. 2009.

[16] M. Jubran, M. Bansal, and L. Kondi, "Low-delay low-complexity bandwidth-constrained wireless video transmission using svc over mimo systems," *Multimedia, IEEE Transactions on*, vol. 10, no. 8, pp. 1698–1707, Dec. 2008.

[17] Y. Pourmohammadi-Fallah and H. Alnuweiri, "Analysis of Temporal and Throughput Fair Scheduling in Multi-Rate IEEE 802.11e WLANs," *Journal of Computer Networks, Elsevier*, vol. 52, no. 16, pp. 3169–3183, November 2008.

[18] Y. Pourmohammadi-Fallah, H. Mansour, S. Khan, P. Nasiopoulos, and H. Alnuweiri, "A Link Adaptation Technique for Efficient Transmission of H.264 Scalable Video Over Multirate WLANs," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 7, pp. 875–887, July 2008.

[19] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of Video Transmission over Lossy Channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, June 2000.

[20] G. Sullivan, T. Wiegand, and K.-P. Lim, "Joint Model Reference Encoding Methods and Decoding Concealment Methods," ISO/IEC JTC1/SC29/WG11 and ITU-T Q6/SG16 JVT-I049, September 2003.

[21] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 17, no. 9, pp. 1103–1120, 2007.

[22] D.-K. Kwon, M.-Y. Shen, and C. C. J. Kuo, "Rate Control for H.264 Video With Enhanced Rate and Distortion Models," *IEEE Trans. Circuits Syst. Video Techn.*, vol. 17, no. 5, pp. 517–529, 2007.

[23] ISO/IEC JTC 1/SC 29/WG 11 N8964, "JSVM-10 software," 2007. [Online]. Available: http://wg11.sc29.org/mpeg/docs/_listwg11_80.htm

[24] V. Erceg, L. Schumacher, P. Kyritsi, A. Molisch, D. S. Baum, and et al., "Tgn channel models," IEEE 802.11-03/940r2, Jan. 2004.

[25] L. Schumacher, "Wlan mimo channel matlab program." [Online]. Available: http://www.info.fundp.ac.be/~lsc/Research/IEEE_80211_HTSG_CMSC/distribution_terms.html