VIDEO-AWARE MIMO PRECODING WITH PACKET PRIORITIZATION AND UNEQUAL MODULATION

Amin Abdel Khalek, Constantine Caramanis, and Robert W. Heath Jr.

ABSTRACT

Video streaming over wireless networks can achieve large quality and capacity gains from transmission techniques that are aware of the video content. This paper proposes modifying conventional multiple-input multiple-output (MIMO) systems to realize these gains by introducing video-aware PHY layer decisions. Specifically, a new PHY layer packet prioritization method is introduced that splits video data among spatial streams based on the packet loss visibility and per-stream SNRs. The objective is to maximize throughput weighted by packet loss visibility, a metric coined perceived throughput. The globally-optimal splitting policy among streams is derived. Furthermore, unequal modulation per stream is proposed and visibility-based packet dropping is optimized to satisfy delay constraints. We derive the gains vs. conventional MIMO precoding and prove that the gain is the throughput averaged over streams divided by the throughput of the worst stream.

1. INTRODUCTION

The growth of video traffic demand over wireless networks calls for video-aware transmission techniques that are able to support more video streams with high video quality. While multiple-input multiple-output (MIMO) systems achieve large capacity gains relative to single antenna systems, they can provide higher performance if they are made video-aware.

Video packet loss visibility information serves as a suitable choice of side information for the PHY layer to use in making video-aware decisions. The loss visibility of different videos slices or packets exhibit high variability because state-of-the-art codecs (e.g. [1]) use inter-frame coding and motion compensation. Loss visibility modeling received attention in recent literature [2–4]. In [2], we developed a learning-based approach that assesses the perceptual quality loss due to packet losses from each scalable-coded layer and used that to provide unequal protection for temporally- and quality-scalable video. In [4], a generalized linear model was proposed for loss visibility modeling accounting for temporal and spatial masking effects and was used to provide packet prioritization by applying visibility-based packet dropping.

While no previous work explicitly incorporates loss visibility information into MIMO-based systems, previous work on video-aware MIMO PHY adaptation is found in [5–7]. In [5], the problem of video broadcast in a multiuser MIMO setup is studied subject to delay constraints. In [6], a method for prioritized delivery for layered video is proposed using antenna switching. In [7], the diversity and multiplexing gain of a MIMO system are adjusted to minimize the video distortion.

In this paper, we propose a new MIMO PHY layer architecture that enables video-aware adaptive transmission. A packet prioritization method is proposed that splits the video data among spatial streams based on the post processing SNRs and packet loss visibility. To optimize the splitting policy, we define a video-aware metric, perceived throughput, that generalizes the conventional notion of throughput by weighting each packet by its loss visibility. We derive the optimal splitting policy for video packets among spatial streams as well as the optimal MCS per stream to maximize perceived throughput and satisfy delay constraints. Furthermore, we incorporate visibility-based packet dropping to satisfy real-time delay requirements under stringent channel conditions. Finally, we derive expressions for the gain in comparison to conventional MIMO precoding and show that notable quality and capacity gains are achieved by the proposed approach.

2. SYSTEM OVERVIEW

We consider a MIMO system with N_t transmit antennas and N_r receive antennas. The system uses N_s spatial streams where $N_\text{s} \leq \min(N_\text{t}, N_\text{r})$ along with unitary SVD precoding.

The proposed cross-layer block diagram is shown in Figure 1. The video encoder compresses and packetizes the raw video data. The compressed video data has a source rate R bits/sec and passes through a packet loss visibility estimator that outputs a value $v(p)$ ranging from 0 to 1 and indicating the *perceptual value* or *loss visibility* of the packet. A value $v(p) = 0$ indicates that losing packet p does not have a visible impact on the end video quality. More relevant packets have a higher $v(p)$. The loss visibility $v(p)$ of packet p is communicated to the physical layer as side information.

We propose a thresholding-based policy whereby the highest priority packets are transmitted through the best spatial stream and vice versa. For that purpose, we define a vector of thresholds $\hat{\mathbf{v}} = \{\hat{v}_i\}_{i=2}^{i=N_\text{s}}$ where $0 \leq \hat{v}_i \leq \hat{v}_{i+1} \leq 1$. We fix $\hat{v}_{N_\text{s}+1} = 1$. At the PHY layer, the packet prioritization demultiplexer routes packet p through stream i if $\hat{v}_i \leq v(p) \leq \hat{v}_{i+1}$. Furthermore, if $v(p) \leq \hat{v}_1$, packet p is

The authors are with the Wireless Networking & Communications Group in the Department of Electrical and Computer Engineering at The University of Texas at Austin, 1 University Station C0803, Austin, TX 78712-0240. Email: akhalek@utexas.edu, caramanis@mail.utexas.edu, rheath@ece.utexas.edu. This work was supported by the Intel-Cisco Video Aware Wireless Networks (VAWN) Program.

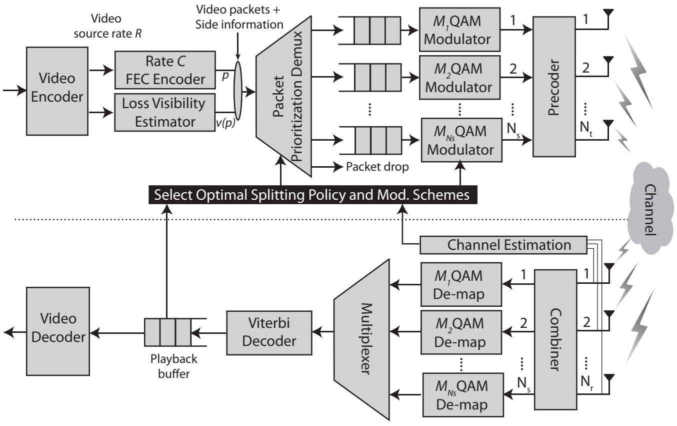

Fig. 1. System block diagram.

dropped. Packets are queued to absorb any mismatch between the source rate and the transmission rate.

An $N_\mathrm{t} \times N_\mathrm{s}$ linear precoding matrix $\mathbf{F}$ maps a N_s-dimensional symbol $\mathbf{s}$ to a N_t-dimensional spatial signal. The signal encounters a channel matrix $\mathbf{H}$ and an AWGN noise vector $\mathbf{n}$. Thus, the input-output relationship is $\mathbf{y} = \sqrt{E_\mathrm{s}/N_\mathrm{t}}\mathbf{HFs} + \mathbf{n}$ where the entries of $\mathbf{H}$ are distributed according to $\mathcal{CN}(0,1)$ and the entries of $\mathbf{v}$ are distributed according to $\mathcal{CN}(0,N_0)$. Furthermore, we assume a block-fading model whereby the channel realization is fixed over a set of P packets and then independently takes a new realization. The transmitter is assumed to have perfect channel knowledge. We apply unitary precoding whereby $\mathbf{F}$ is a normalized version of the right singular vectors of $\mathbf{H}$, that is, $\mathbf{F} = [\mathbf{V}]_{:,1:N_\mathrm{s}}/N_\mathrm{s}$ where $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^*$ is the singular value decomposition of $\mathbf{H}$. Combining at the receiver reduces to matched filtering by $[\mathbf{U}]^*_{:,1:N_\mathrm{s}}$ and performing N_s independent detections per stream. Thus, the post processing SNR for the i^{th} stream simplifies to $\gamma_i(\mathbf{H}) = (E_\mathrm{s}/N_0) \times (\sigma_i^2/N_s)$ where σ_i is the i^{th} singular value of $\mathbf{H}$. Given $\mathbf{H}$, the streams are re-indexed decreasingly according to their post processing SNRs, i.e., $\gamma_i(\mathbf{H}) \geq \gamma_{i+1}(\mathbf{H})$.

We apply unequal modulation per stream. The data through stream i is modulated with a QAM constellation of size $M_i \in \mathcal{M}$. Each constellation is normalized such that the average symbol energy is unity. The vector of modulation schemes is denoted $\mathbf{M} = \{M_i\}_{i=1}^{i=N_\mathrm{s}}$. The data through all streams are coded with a convolutional encoder with coding rate $C \in \mathcal{C}$. The probability that a packet through stream i is in error is denoted by $\rho_i = \rho(M_i, C, \gamma_i(\mathbf{H}))$ where

$$\rho(M_i, C, \gamma_i(\mathbf{H})) = 1-(1-\mathrm{P}(s_i \neq \hat{s}_i|M_i, C, \gamma_i(\mathbf{H})))^{\frac{b}{C\cdot\log_2 M_i}}$$

where $\mathrm{P}(s_i \neq \hat{s}_i|M_i, C, \gamma_i(\mathbf{H}))$ is the symbol error probability on the i^{th} stream, and b is the packet size in bits. We compute $\mathrm{P}(s_i \neq \hat{s}_i|M_i, C, \gamma_i(\mathbf{H}))$ by modifying the uncoded M-QAM expressions to account for the coding gain of the convolutional codes in use.

3. A VIDEO-AWARE METRIC: PERCEIVED THROUGHPUT

We define a video-aware metric, coined *perceived throughput*, as the total perceptual value of packets transmitted per unit time. The perceived throughput expression is

$$PT = \frac{\sum_{i=1}^{N_\mathrm{s}} V_i}{\max_i t_i} \tag{1}$$

where V_i is the cumulative value of packets transmitted through stream i and t_i is the time to transmit through stream i. While V_i could be computed based on the values of the instantaneous set of queued packets $\mathcal{P}$, it is undesirable because it dictates a slow adaptation timescale so that the set of buffered packets is representative of the loss visibility variation. Furthermore, it would result in (1) being non-differentiable, making it difficult to derive optimal policies that exhibit structure. Instead, we propose to estimate the *loss visibility distribution*, update it over time using the values of incoming packets, and use it to adapt the splitting policy. This allows the PHY layer adaptation to operate at the channel coherence timescale. It also enables deriving the optimal splitting policy and the packet prioritization gain expressions for any continuous loss visibility distribution. We note that the loss visibility distribution can be inexpensively computed and updated using kernel density estimation (KDE) [8].

Let $f_v(.)$ and $f_b(.)$ denote the distribution of the loss visibility values $v(p)$ and packet sizes $b(p)$ respectively. We assume packetization ensures packet sizes and values are uncorrelated. This assumption is reasonable because PHY layer packets are typically of comparable sizes irrespective of the content. Furthermore, we assume the number of retransmissions is a geometric random variable with mean $1/(1-\rho_i)$.

Given a representative set a packets $\mathcal{P}$, the fraction of packets transmitted through stream i is $|\mathcal{P}|\int_{\hat{v}_i}^{\hat{v}_{i+1}} f_v(v)dv = |\mathcal{P}|(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))$. Furthermore, for geometric retransmissions, the mean time to transmit a packet of size $b(p)$ through stream i is $b(p)/(CB\log_2 M_i(1-\rho_i))$. Thus,

$$t_i = |\mathcal{P}|\frac{\mathbb{E}[b]}{CB\log_2 M_i(1-\rho_i)}(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i)) \tag{2}$$

where $\mathbb{E}[b] = \int bf_b(b)db$ is the mean packet size and B is the bandwidth. Furthermore, $V_i = |\mathcal{P}|\int_{\hat{v}_i}^{\hat{v}_{i+1}} vf_v(v)dv$. Therefore, PT reduces to

$$\begin{aligned} PT &= \frac{\int_{\hat{v}_1}^{1} vf_v(v)dv}{\mathbb{E}[b]} \cdot \min_i \left\{\frac{CB\log_2 M_i(1-\rho_i)}{F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i)}\right\} \\ &= \frac{\mathbb{E}[v] - \int_0^{\hat{v}_1} vf_v(v)dv}{\mathbb{E}[b]} \cdot \frac{CB\log_2 M_{\tilde{i}}(1-\rho_{\tilde{i}})}{(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))} \end{aligned} \tag{3}$$

with $\tilde{i} = \mathrm{argmax}_i\{(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))/C\log_2 M_i(1-\rho_i)\}$.

4. PACKET PRIORITIZATION POLICY

4.1. Problem Formulation

The objective is to select the set of thresholds $\mathbf{v}$, the set of modulation schemes $\mathbf{M}$, and the coding rate C to maximize the perceived throughput as expressed in (3) subject to a delay constraint. The delay constraint should ensure that the queuing delay d_i through each stream does not exceed the deadline D. Thus, the problem is formulated as follows

$$\max_{\hat{\mathbf{v}},\mathbf{M},C} \quad PT(\hat{\mathbf{v}},\mathbf{M},C) \tag{4}$$
$$\text{s.t.} \quad d_i \leq D \; \forall i = 1,\cdots,N_{\text{s}} \tag{5}$$
$$0 \leq \hat{v}_i \leq \hat{v}_{i+1} \leq 1 \; \forall i = 1,\cdots,N_{\text{s}} \tag{6}$$
$$M_i \in \mathcal{M};\; C \in \mathcal{C}. \tag{7}$$

For the duration of a block of P packets, the queue ahead of each stream can be modeled as an M/M/1 queue for exponential packet sizes since the packet prioritization demultiplexer effectively performs random splitting and the channel state and service rate are fixed for the duration of one block. Thus, the average queuing delay for a packet to be served is $d_i = 1/(\mu_i - \lambda_i)$ where λ_i and μ_i are the arrival rate and service rate on the i^{th} stream respectively. Given $\mathbf{M}$ and C, we have $\mu_i = CB\log_2 M_i(1-\rho(M_i, C, \gamma_i(\mathbf{H})))/\mathbb{E}[b]$ packets/sec and $\lambda_i = RC(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))/\mathbb{E}[b]$ packets/sec. Thus, the queuing delay is

$$d_i = \frac{\mathbb{E}[b]}{C[B\log_2 M_i(1-\rho_i) - R(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))]}. \tag{8}$$

The gradient of PT can be written as $\partial PT/\partial\hat{v}_i = (h\partial g/\partial\hat{v}_i - g\partial h/\partial\hat{v}_i)/h^2$ where $g = \mathbb{E}[v] - \int_0^{\hat{v}_1} v f_v(v)dv$ and $h = \mathbb{E}[b](F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))/(CB\log_2 M_{\tilde{i}})$ are the numerator and denominator of (3). We have,

$$\frac{\partial g}{\partial \hat{v}_i} = \begin{cases} 0 & \text{if } i > 1 \\ -\hat{v}_1 f_v(\hat{v}_1) & \text{if } i = 1 \end{cases}$$

where we used the fact that $\partial(\int_0^{\hat{v}_1} v f_v(v)dv)/\partial\hat{v}_1 = \lim_{\epsilon\to 0}(\int_{\hat{v}_1}^{\hat{v}_1+\epsilon} v f_v(v)dv/\epsilon) = \hat{v}_1 f_v(\hat{v}_1)$. Furthermore,

$$\frac{\partial h}{\partial \hat{v}_i} = \begin{cases} \mathbb{E}[b] f_v(\hat{v}_i)/(CB\log_2 M_{i-1}(1-\rho_{i-1})) & \text{if } i = \tilde{i}+1 \\ -\mathbb{E}[b] f_v(\hat{v}_i)/(CB\log_2 M_i(1-\rho_i)) & \text{if } i = \tilde{i} \\ 0 & \text{otherwise.} \end{cases}$$

The following Lemma will be used to derive the necessary conditions for optimality in Theorem 1.

Lemma 1. *The gradient $\partial PT/\partial\hat{v}_i \leq 0 \; \forall i \neq \tilde{i}$.*

Proof. Investigating the terms of the gradient $\partial PT/\partial\hat{v}_i = (h\partial g/\partial\hat{v}_i - g\partial h/\partial\hat{v}_i)/h^2$, we have $h \geq 0$ and $\partial g/\partial\hat{v}_i \leq 0$ unconditionally. Furthermore, $\partial h/\partial\hat{v}_i \geq 0 \; \forall i \neq \tilde{i}$ and $g \geq 0$ if $\hat{v}_1$ satisfies $\int_0^{\hat{v}_1} v f_v(v)dv \leq \mathbb{E}[v]$. □

4.2. The Optimal Splitting Policy

Theorem 1 provides the optimal solution to the stream splitting problem and applies for any continuous loss visibility distribution. Intuitively, the solution balances the packet load among streams in proportion to the throughput of each stream.

Theorem 1. *The optimal splitting policy $\hat{\mathbf{v}}^* = \{\hat{v}_i^*\}$ satisfies $F_v(\hat{v}_{i+1}^*) - F_v(\hat{v}_i^*) = (1 - F_v(\hat{v}_1^*))\log_2 M_i(1-\rho_i)/\sum_j \log_2 M_j(1-\rho_j) \; \forall i = 1,\cdots,N_s$.*

Proof. We present a convergent method that takes as input any feasible solution and obtains a solution with an improved objective satisfying the Theorem statement. We denote the set of streams with the longest average transmission time by $\mathcal{I}$. For any initial policy $\hat{\mathbf{v}}$, we have $\mathcal{I} = \{\tilde{i}\}$. For all $i \notin \mathcal{I} \cup \{1\}$, reduce $\hat{v}_i$ until $t_i = \max_j\{t_j\}$. Reducing $\hat{v}_i$ improves the objective by Lemma 1. Repopulate $\mathcal{I}$ and repeat until $\mathcal{I} = \{2,\cdots,N_{\text{s}}\}$. Next, reduce $\hat{v}_1$ until either $t_1 = \max_j\{t_j\}$ or $\hat{v}_1 = 0$. If the former occurs first, the Lemma is shown since now $\mathcal{I} = \{1,\cdots,N_{\text{s}}\}$. Otherwise, if $\hat{v}_1 = 0$, the objective can be further improved by jointly scaling up all $\hat{v}_2,\cdots,\hat{v}_{N_{\text{s}}}$ while preserving $\mathcal{I} \supseteq \{2,\cdots,N_{\text{s}}\}$ until $\mathcal{I} = \{1,\cdots,N_{\text{s}}\}$. Thus, $(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))/(\log_2 M_i(1-\rho_i)) = (F_v(\hat{v}_2) - F_v(\hat{v}_1))/(\log_2 M_1(1-\rho_1)) \; \forall i$. By taking $1 - F_v(\hat{v}_1) = \sum_i F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i)$, the Theorem follows. We note that the delay constraint is not jeopardized since the worst case delay is only improved in the process. □

Theorem 1 enables finding the optimal $\hat{\mathbf{v}}^* = \{\hat{v}_i^*\}_{i=2}^{i=N_s}$. Next, Theorem 2 provides the optimal packet drop threshold $\hat{v}_1^*$ to satisfy the delay constraints.

Theorem 2. *The optimal dropping threshold $\hat{v}_1^*$ is given by*

$$\hat{v}_1^* = \max\left\{F_v^{-1}\left(1 - \frac{\kappa}{R}\left[1 - \frac{\mathbb{E}[b]}{BDC\min_i\{\log_2 M_i(1-\rho_i)\}}\right]\right), 0\right\} \tag{9}$$

where $\kappa = B\sum_i \log_2 M_i(1-\rho_i)$.

Proof. For a given choice of $\mathbf{M}$ and C, the problem of finding the optimal packet dropping threshold reduces to

$$\min \quad \hat{v}_1 \tag{10}$$
$$\text{s.t.} \quad \frac{\mathbb{E}[b]}{C\log_2 M_i(1-\rho_i)} \times \frac{1}{1 - R(1 - F_v(\hat{v}_1))/\kappa} \leq D \forall i \tag{11}$$
$$0 \leq \hat{v}_1 \leq 1 \tag{12}$$

where $\kappa = B\sum_i \log_2 M_i(1-\rho_i)$. The Theorem follows as the unique solution of the optimization problem. □

Applying Theorem 1 and Theorem 2, the optimal PT is

$$PT(\hat{\mathbf{v}}^*, \mathbf{M}, C) = K(\hat{v}_1^*)CB\sum_i \log_2 M_i(1-\rho_i) \tag{13}$$

where $K(\hat{v}_1^*) = (\mathbb{E}[v] - \int_0^{\hat{v}_1^*} v f_v(v) dv)/(\mathbb{E}[b](1 - F_v(\hat{v}_1^*)))$. Furthermore, the optimal MCS $(\{M_i^*\}, C^*)$ is as follows: $\tilde{M}_i(C) = \text{argmax}_{M_i \in \mathcal{M}}\{\log_2 M_i(1 - \rho(M_i, C, \gamma_i(\mathbf{H})))\}$, $C^* = \text{argmax}_{C \in \mathcal{C}}\{C \sum_i \log_2 \tilde{M}_i(C)(1 - \rho(M_i, C, \gamma_i(\mathbf{H})))\}$, and $M_i^* = \tilde{M}_i(C^*)$. Thus,

$$PT^* = K(\hat{v}_1^*) B \max_{C \in \mathcal{C}} \left\{ C \sum_i \max_{M_i \in \mathcal{M}} \{\log_2 M_i (1 - \rho_i)\} \right\}. \quad (14)$$

5. VIDEO-AWARE MIMO PRECODING GAINS

We derive the gains in comparison to conventional MIMO signaling whereby the bits are interleaved among spatial streams in a round robin fashion. The operating modes are denoted by PP for packet prioritization, SM for non-prioritized precoded spatial multiplexing, UM for unequal modulation, and EM for equal modulation. To assess the gains due to packet prioritization separately from the gains due to visibility-aware packet dropping, we consider the non-delay-limited regime whereby $\hat{v}_1^* = 0$. In this regime, we have $K(0) = \mathbb{E}[v]/\mathbb{E}[b]$. Thus, (14) reduces to

$$PT_{\text{UM}}^{\text{PP}*} = \frac{\mathbb{E}[v]}{\mathbb{E}[b]} B \max_{C \in \mathcal{C}} \left\{ C \sum_i \max_{M_i \in \mathcal{M}} \{\log_2 M_i (1 - \rho_i)\} \right\}.$$

As a baseline, we consider the case when no packet prioritization is applied at the PHY layer, i.e., with conventional MIMO signaling and precoding. In this case, for a set of packets $\mathcal{P}$, the mean value of transmitted packets is $\mathbb{E}[v]|\mathcal{P}|$. Since $|\mathcal{P}|/N_{\text{s}}$ packets are transmitted through stream i, the worst-case time to transmit among all streams is $\max_i\{\mathbb{E}[b]/CB \log_2 M_i(1 - \rho(M_i, C, \gamma_i(\mathbf{H})))\}|\mathcal{P}|/N_{\text{s}}$. Thus, the expression is

$$\begin{aligned} PT_{\text{UM}}^{\text{SM}} &= \frac{\mathbb{E}[v]|\mathcal{P}|}{\max_i\{\mathbb{E}[b]/CB \log_2 M_i(1 - \rho_i)\}|\mathcal{P}|/N_{\text{s}}} \\ &= \frac{\mathbb{E}[v]}{\mathbb{E}[b]} B N_{\text{s}} C \min_i \{\log_2 M_i (1 - \rho_i)\} \end{aligned} \quad (15)$$

Thus, the corresponding optimal perceived throughput is

$$PT_{\text{UM}}^{\text{SM}*} = \frac{\mathbb{E}[v]}{\mathbb{E}[b]} B N_{\text{s}} \max_{C \in \mathcal{C}} \{C \min_i \{\max_{M_i \in \mathcal{M}} \log_2 M_i (1 - \rho_i)\}\}.$$

Now, we write the gain due to packet prioritization $G_{\text{PP}} = \mathbb{E}_{\mathbf{H}}\left[PT_{\text{UM}}^{\text{PP}*}\right] / \mathbb{E}_{\mathbf{H}}\left[PT_{\text{UM}}^{\text{SM}*}\right]$ as follows

$$G_{\text{PP}} = \frac{\mathbb{E}_{\mathbf{H}}[\max_C \{C \sum_i \max_{M_i}\{\log_2 M_i (1 - \rho_i)\}\}]}{N_{\text{s}} \mathbb{E}_{\mathbf{H}}[\max_C \{C \min_i \{\max_{M_i}\{\log_2 M_i (1 - \rho_i)\}\}\}]}. \quad (16)$$

Interestingly, this expression is interpreted as the achievable throughput averaged over all streams divided by the achievable throughput of the worst stream. Intuitively, the more variability among streams, the more we gain from packet prioritization. Furthermore, it does not depend on the loss visibility distribution and is achieved for any such distribution whereby the splitting solution in Theorem 1 is applied. We further note that an extra gain is achieved due to visibility-aware dropping in the delay limited regime. It is greater than 1 and depends on the distribution. For instance, with a uniform distribution, the gain is $\mathbb{E}_{\mathbf{H}}[(1 + \hat{v}_1^*)]/\mathbb{E}_{\mathbf{H}}[(1 - \hat{v}_1^*)]$.

Next, we assess the unequal modulation gain. With equal modulation, the perceived throughput expression is

$$PT_{\text{EM}}^{\text{PP}*} = \frac{\mathbb{E}[v]}{\mathbb{E}[b]} B \max_{C \in \mathcal{C}} \left\{ C \max_{M \in \mathcal{M}} \{\log_2 M \sum_i (1 - \rho_i)\} \right\}.$$

Thus, the gain due to unequal modulation, defined as $G_{\text{UM}} = \mathbb{E}_{\mathbf{H}}\left[PT_{\text{UM}}^{\text{PP}*}\right] / \mathbb{E}_{\mathbf{H}}\left[PT_{\text{EM}}^{\text{PP}*}\right]$, can be expressed as follows

$$G_{\text{UM}} = \frac{\mathbb{E}_{\mathbf{H}}[\max_{C \in \mathcal{C}} \{C \sum_i \max_{M_i \in \mathcal{M}}\{\log_2 M_i (1 - \rho_i)\}\}]}{\mathbb{E}_{\mathbf{H}}[\max_{C \in \mathcal{C}} \{C \max_{M \in \mathcal{M}}\{\log_2 M \sum_i (1 - \rho_i)\}\}]}. \quad (17)$$

6. RESULTS AND ANALYSIS

In this section, we quantify the packet prioritization and unequal modulation gains in comparison to a non-video-aware MIMO precoding system for the same N_{t}, N_{r}, and N_{s}. The set of possible M-QAM constellations is $\mathcal{M} = \{2, 4, 16, 64\}$ and set of possible coding rates is $\mathcal{C} = \{1/2, 2/3, 3/4, 5/6\}$.

In Figure 2a, the packet prioritization gain G_{PP}, defined in (16), is plotted vs. E_{s}/N_0 for $N_{\text{s}} = 2$ spatial streams with different antenna configurations. Recall from the gain expression (16) that the largest gain is achieved when the post processing SNR statistics exhibit the highest variability among streams. Thus, for $N_{\text{s}} = 2$, a 2×2 system gains more than a 4×4 system. In a 4×4 system with $N_{\text{s}} = 2$, the diversity and channel hardening reduce the gains from the proposed policy.

In Figure 2b, we plot the gain from packet prioritization for a 4×4 system for different values of N_{s}. In the medium to high SNR regime, for the same $N_{\text{t}} \times N_{\text{r}}$ configuration, more streams provide higher gains versus non-video aware approaches since the condition number of the effective channel $\mathbf{HF}$ is likely to be higher making it possible for the Algorithm to utilize the diverse channel statistics among streams.

To explain the oscillatory behavior in Figure 2a and 2b, we show the fractional use of each modulation scheme at the peak operating points in Figure 2b. For $N_s = 2$ at $E_{\text{s}}/N_0 = -1$ dB, the best stream can support 4-QAM for most realizations while the worst stream can only support BPSK. A similar observation follows at 8 dB and 15 dB for 16-QAM and 64-QAM. Conversely, when both streams use the same modulation for most channel realizations, the gain drops to 1.

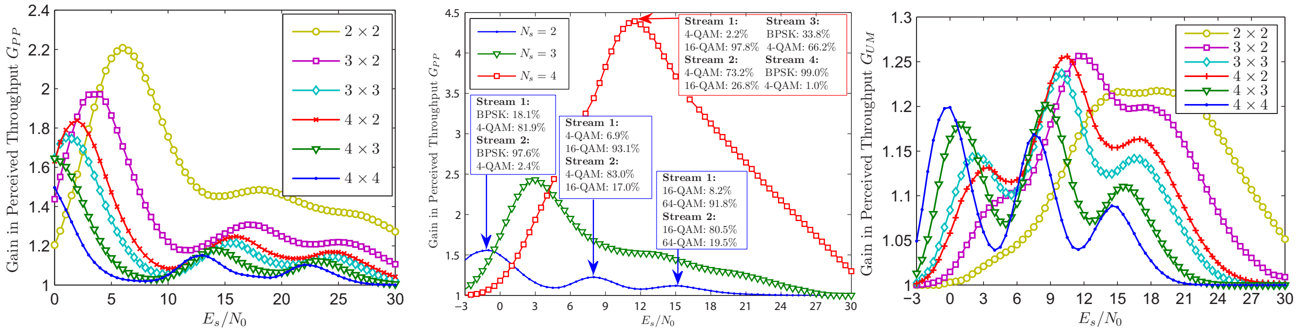

Fig. 2. (a) Packet Prioritization Gain G_{PP} vs. E_s/N_0 for $N_s = 2$ and different antenna configurations, (b) Packet prioritization gain for a 4×4 system with different values of N_s. For $N_s = 2$ and $N_s = 4$, the fractional use of each modulation scheme is shown at the peak operating points, (c) Unequal Modulation Gain G_{UM} vs. E_s/N_0 for $N_s = 2$ with different antenna configurations.

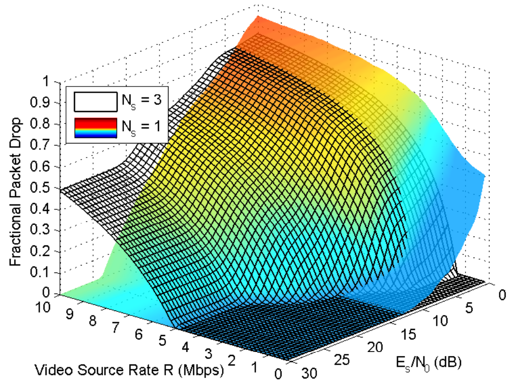

Fig. 3. Fractional packet drop $\mathbb{E}[v_1^*]$ vs. source rate R and E_s/N_0 for a 4×4 system with $B = 1$ MHz for $N_s = 1$ (transmit beamforming) and $N_s = 3$.

Figure 2c shows the unequal modulation gain G_{UM}, defined in (17), vs. E_s/N_0 for $N_s = 2$. In the low-SNR regime, the gains are most pronounced for a 4×4 system since the frequently good SNR on the better stream can be leveraged with a high order modulation scheme. In the high-SNR regime, channel hardening causes the modulation schemes per stream to be equivalent, thus limiting the unequal modulation gain.

Finally, we analyze delay-constrained video transmission with different MIMO modes and source rates. We allow a 150 ms startup delay. For subsequent decisions, the delay constraint D is updated based the source and channel rate to meet the playback deadline. Figure 3 shows the fractional packet drop $\mathbb{E}[v_1^*]$ to satisfy the delay constraint with a uniform loss visibility distribution for a 4×4 MIMO system with bandwidth $B = 1$ MHz vs. the video source rate R and the average SNR E_s/N_0 for $N_s = 1$ (transmit beamforming) and $N_s = 3$. To satisfy a 10% drop rate at $E_s/N_0 = 15$ dB, we can support at most a 4 Mbps video source with beamforming and a 6 Mbps video source with $N_s = 3$. Conversely, at $E_s/N_0 = 6.5$ dB, we can support a 2 Mbps video source with beamforming and only a 1.2 Mbps video source with $N_s = 3$. Thus, beamforming is preferred at low SNRs and source rates.

7. CONCLUSION

We proposed a PHY layer architecture that supports prioritized packet delivery based on packet loss visibility. The proposed architecture requires minimal additional cross-layer overhead, is easy to implement in current streaming systems, and achieves notable quality and capacity gains.

8. REFERENCES

[1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H. 264/AVC video coding standard," *IEEE Transactions on circuits and systems for video technology*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[2] A. Abdel Khalek, C. Caramanis, and R. W. Heath Jr., "A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection," *Accepted to IEEE Journal on Selected Areas in Communications*, vol. 30, no. 7, pp. 456–466, Aug. 2012.

[3] A. Abdel Khalek, C. Caramanis, and R. W. Heath Jr., "Joint source-channel adaptation for perceptually optimized scalable video transmission," *In Proceedings of IEEE Globecom*, Dec. 2011.

[4] T. Lin, S. Kanumuri, Y. Zhi, D. Poole, P. Cosman, and A. Reibman, "A versatile model for packet loss visibility and its application to packet prioritization," *IEEE Transactions on Image Processing*, vol. 19, no. 3, pp. 722–735, Mar. 2010.

[5] J. Xu, R. Hormis, and X. Wang, "MIMO video broadcast via transmit-precoding and SNR-scalable video coding," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 3, pp. 456–466, Apr. 2010.

[6] 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, pp. 1218 –1226, Sep. 2007.

[7] R. Hormis, E. Linzer, and X. Wang, "Adaptive mode and diversity-control for video transmission on MIMO wireless channels," *IEEE Transactions on Signal Processing*, vol. 57, no. 9, pp. 3624–3637, Sep. 2009.

[8] D. Scott, *Multivariate density estimation*, vol. 139. Wiley Online Library, 1992.